Geometry & Topology

Volume 29 (2025)

Issue 3 (pages 1115–1691)

msp

GEOMETRY & TOPOLOGY

msp.org/gt

See inside back cover or msp.org/gt for submission instructions.

The subscription price for 2025 is US $865/year for the electronic version, and $1210/year (+$75, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP. Geometry & Topology is indexed by Mathematical Reviews, Zentralblatt MATH, Current Mathematical Publications and the Science Citation Index.

Geometry & Topology (ISSN 1465-3060 printed, 1364-0380 electronic) is published 9 times per year and continuously online, by Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840. Periodical rate postage paid at Oakland, CA 94615-9651, and additional mailing offices. POSTMASTER: send address changes to Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840.

GT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Geometry & Topology **29**:3 (2025) 1115–1170
DOI: 10.2140/gt.2025.29.1115
Published: 31 May 2025

A cubical model for (∞, n)-categories

TIM CAMPION
KRZYSZTOF KAPULKIN
YUKI MAEHARA

We propose a new model for the theory of (∞, n)-categories (including the case $n = \infty$) in the category of marked cubical sets with connections, similar in flavor to complicial sets of Verity. The model structure characterizing our model is shown to be monoidal with respect to suitably defined (lax and pseudo) Gray tensor products; in particular, these tensor products are both associative and biclosed. Furthermore, we show that the triangulation functor to precomplicial sets is a left Quillen functor and is strong monoidal with respect to both Gray tensor products.

18N65; 18N40, 55U35, 55U40

Introduction

The theory of (∞, n)-categories is becoming an important tool in a number of areas of mathematics, including manifold topology, where it is used in the definition and classification of extended topological quantum field theories [Lurie 2009], and in derived algebraic geometry, where it is used to capture certain properties of the "category" of correspondences [Gaitsgory and Rozenblyum 2017a, 2017b]. There are several equivalent models for this theory, including: n-trivial saturated complicial sets [Verity 2008b, 2007; Riehl and Verity 2020; Loubaton 2022], n-quasicategories [Ara 2014], Θ_n-spaces [Rezk 2010] and n-fold complete Segal spaces [Barwick 2005].

We propose a new model for the theory of (∞, n)-categories, using *comical* (*com*position + cub*ical*) *sets*. Comical sets are certain marked cubical sets (having marked n-cubes for all values $n \geq 1$), just like complicial sets are certain marked simplicial sets. Our model allows for a particularly elegant and simple treatment of the (lax and pseudo) Gray tensor products since they are inherently cubical in nature. One can find drawings of cubes in Gray's book [1974], and the simplest ways of defining the lax Gray tensor product of strict ω-categories are via cubical sets [Crans 1995; Al-Agl et al. 2002].

Because of the obvious similarities with complicial sets, there is a natural comparison functor to marked simplicial sets. To obtain it, we extend the usual triangulation functor $T\colon \mathsf{cSet} \to \mathsf{sSet}$ from cubical sets to simplicial sets to a marked version $T\colon \mathsf{cSet}^+ \to \mathsf{PreComp}$. Here T is valued not in the whole category sSet^+ of marked simplicial sets but in the reflective subcategory $\mathsf{PreComp}$ of precomplicial sets, so that our results hold up to isomorphism rather than homotopy. $\mathsf{PreComp}$ supports a model structure that is Quillen equivalent to the complicial model structure on sSet^+ and the lax Gray tensor product on sSet^+ is more well-behaved when restricted to $\mathsf{PreComp}$.

With that, our main results (see Theorems 3.3, 2.16, 6.5, 6.10 and 7.1) can be summarized as follows:

Theorem *The category cSet^+ of marked cubical sets carries a model structure whose cofibrations are the monomorphisms and whose fibrant objects are the comical sets. This model structure is monoidal with respect to both the lax and pseudo Gray tensor products, which are simultaneously associative and biclosed.*

Furthermore, the triangulation functor $T\colon \mathsf{cSet}^+ \to \mathsf{PreComp}$ is left Quillen and strong monoidal with respect to both Gray tensor products.

Since this paper was first made available in 2020, a slight adaptation of our model was proven to be Quillen equivalent via the triangulation functor to n-trivial saturated complicial sets in [Doherty et al. 2023]. The "special cases" of this result had previously been known for $(\infty, 0)$-categories (ie ∞-groupoids) [Cisinski 2006] and $(\infty, 1)$-categories [Doherty et al. 2024], although these papers consider slightly different versions of the cubical site from us.

In particular, our model validates the assertions [Gaitsgory and Rozenblyum 2017a, Propositions 10.3.2.6 and 10.3.2.9], given there without a proof or a reference. They are essentially the desiderata of a convenient model of $(\infty, 2)$-categories used throughout [Gaitsgory and Rozenblyum 2017a, 2017b], and in that sense our model in cSet^+ is a convenient such model. We should note however that these assertions were previously proven in [Verity 2008b] and [Maehara 2021] in the contexts of complicial sets and 2-quasicategories, respectively.

Finally, our work owes a great deal to [Steiner 2006], where the (semi)cubical nerves of strict ω-categories are analyzed. In particular, our definition of comical open boxes in Section 3 follows [Steiner 2006, Example 2.9].

Organization of the paper We begin in Section 1 by reviewing the necessary background on model categories, cubical sets and complicial sets. In Section 2, we introduce marked cubical sets, study their basic properties, and construct both the lax and the pseudo Gray tensor products. In Section 3, we define comical sets and construct the model structure for them. As a proof of concept, we define in Section 4 the homotopy 1-category of a comical set and show that it has expected properties. We then turn our attention to the comparison between the cubical and the simplicial approaches. We extend the triangulation functor to marked cubical sets in Section 5, show that it is strong monoidal with respect to both the lax and the pseudo Gray tensor products in Section 6, and that it is a left Quillen functor in Section 7.

Acknowledgements The authors benefited greatly from conversations about related matters with Alexander Campbell, Emily Riehl and Dominic Verity. The paper was greatly improved by the comments of the referee. This material is based upon work supported by the National Science Foundation under grant 1440140, while the authors were in residence at the Mathematical Sciences Research Institute in Berkeley, California, in the program *Higher categories and categorification* in Spring 2020. We would like to thank the MSRI for its hospitality, and the program organizers for giving us the opportunity to participate. Above all, we thank (again) Emily Riehl for continued support and encouragement.

1 Background

In this section we introduce the notation and collect preliminary results to be used later in the paper.

1.1 Model categories

In this subsection, we review (a special case of) the theory of Olschok [2009] for constructing combinatorial model structures with all objects cofibrant, which generalizes the theory of Cisinski [2006] for constructing combinatorial model structures on presheaf categories with cofibrations the monomorphisms. This theory will be used to construct the model structures for comical sets.

Definition 1.1 [Simpson 2012] We say a set Λ of trivial cofibrations in a model category $\mathcal{M}$ is *pseudogenerating* if and only if any map f that has a fibrant codomain and the right lifting property with respect to Λ is a fibration.

Now fix a locally presentable category $\mathcal{K}$.

Given a bifunctor $\odot\colon \mathcal{K} \times \mathcal{K} \to \mathcal{K}$ and maps $f\colon A \to A'$ and $g\colon B \to B'$ in $\mathcal{K}$, we denote by

$$f \mathbin{\widehat{\odot}} g\colon (A' \odot B) \amalg_{A \odot B} (A \odot B') \to A' \odot B'$$

the *Leibniz $\odot$-product* of f and g.

Similarly, for any natural transformation $\phi\colon F \Rightarrow G$ between endofunctors $F, G\colon \mathcal{K} \to \mathcal{K}$ and for any $f\colon A \to A'$ in $\mathcal{K}$, we denote by

$$\widehat{\phi}_f\colon G(A) \amalg_{F(A)} F(A') \to G(A')$$

the *Leibniz product* of ϕ and f.

By the *cellular closure* of a set S of maps in $\mathcal{K}$, we mean the closure of S under pushouts along arbitrary maps and transfinite composition. In the rest of this subsection, assume that we are given a small set I of maps in $\mathcal{K}$ whose cellular closure is precisely the monomorphisms.

Definition 1.2 A *functorial cylinder* on $\mathcal{K}$ is a functor $C\colon \mathcal{K} \to \mathcal{K}$ equipped with natural transformations $\partial^0, \partial^1\colon \mathrm{Id} \rightrightarrows C$ and $\sigma\colon C \to \mathrm{Id}$ such that $\sigma\partial^0 = \sigma\partial^1 = \mathrm{id}$. We also write $\partial_X = [\partial^0_X, \partial^1_X]\colon X + X \to CX$. We say that C is a *cartesian* cylinder if the functor C preserves colimits and moreover ∂_X is a monomorphism for all X.

Definition 1.3 Suppose that $\mathcal{K}$ admits a cartesian functorial cylinder $C = (C, \partial^0, \partial^1, \sigma)$. Let S be a set of morphisms in $\mathcal{K}$. We define $\Lambda(\mathcal{K}, I, C, S) \subseteq \operatorname{Mor}\mathcal{K}$ to be the smallest class of morphisms containing

$$S \cup \{\widehat{\partial}^0_i \mid i \in I\} \cup \{\widehat{\partial}^1_i \mid i \in I\},$$

closed under the operation $f \mapsto \widehat{\partial}_f$.

Theorem 1.4 [Olschok 2009, Theorem 2.2.5, Lemma 2.2.20] *Let $\mathcal{K}$ and I be as above. Suppose we are given a cartesian functorial cylinder C on $\mathcal{K}$ and a set S of monomorphisms in $\mathcal{K}$. Then there exists a model structure on $\mathcal{K}$ uniquely characterized by the following properties*:

- *The cofibrations are the monomorphisms.*
- *The set $\Lambda(\mathcal{K}, I, C, S)$ is a pseudogenerating set of trivial cofibrations.*

This model structure is combinatorial and left proper.

Proposition 1.5 *Let $\mathcal{K}$ and I be as above. Suppose $\mathcal{K}$ admits a model structure whose cofibrations are the monomorphisms*, *and a pseudogenerating set Λ of trivial cofibrations. Suppose further that $\mathcal{K}$ is equipped with a tensor product $\odot\colon \mathcal{K} \times \mathcal{K} \to \mathcal{K}$ that forms part of a biclosed monoidal structure. Then these data form a monoidal model structure if and only if*

- *$f \mathbin{\widehat{\odot}} g$ is a cofibration whenever $f, g \in I$,*
- *$f \mathbin{\widehat{\odot}} g$ is a trivial cofibration whenever $f \in \Lambda$ and $g \in I$, and*
- *$f \mathbin{\widehat{\odot}} g$ is a trivial cofibration whenever $f \in I$ and $g \in \Lambda$.*

Proof This is an instance of [Maehara 2021, Proposition A.4]. See also [Henry 2020, Appendix B]. □

1.2 Cubical sets

We will define cubical sets as presheaves on the *box category*, denoted by $\square$. The category $\square$ is the (nonfull) subcategory of the category of posets whose objects are the posets of the form $[1]^n := \{0 \le 1\}^n$, and whose maps are generated by the cubical operators

- *faces* $\partial^n_{i,\varepsilon}\colon [1]^{n-1} \to [1]^n$ for $i = 1, \dots, n$ and $\varepsilon = 0, 1$ given by
$$\partial^n_{i,\varepsilon}(x_1, x_2, \dots, x_{n-1}) = (x_1, x_2, \dots, x_{i-1}, \varepsilon, x_i, \dots, x_{n-1}),$$
- *degeneracies* $\sigma^n_i\colon [1]^n \to [1]^{n-1}$ for $i = 1, 2, \dots, n$ given by
$$\sigma^n_i(x_1, x_2, \dots, x_n) = (x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n),$$
- *max-connections* $\gamma^n_{i,0}\colon [1]^n \to [1]^{n-1}$ for $i = 1, 2, \dots, n-1$ given by
$$\gamma^n_{i,0}(x_1, x_2, \dots, x_n) = (x_1, x_2, \dots, x_{i-1}, \max\{x_i, x_{i+1}\}, x_{i+2}, \dots, x_n),$$
- *min-connections* $\gamma^n_{i,1}\colon [1]^n \to [1]^{n-1}$ for $i = 1, 2, \dots, n-1$ given by
$$\gamma^n_{i,1}(x_1, x_2, \dots, x_n) = (x_1, x_2, \dots, x_{i-1}, \min\{x_i, x_{i+1}\}, x_{i+2}, \dots, x_n).$$

We will omit the superscript n when no confusion is possible.

A straightforward computation shows that cubical operators satisfy the following *cubical identities*. These maps obey the following *cubical identities*:

$$\partial_{j,\varepsilon'}\partial_{i,\varepsilon} = \partial_{i+1,\varepsilon}\partial_{j,\varepsilon'} \quad \text{for } j \le i,$$
$$\sigma_i\sigma_j = \sigma_j\sigma_{i+1} \quad \text{for } j \le i,$$
$$\sigma_j\partial_{i,\varepsilon} = \begin{cases} \partial_{i-1,\varepsilon}\sigma_j & \text{for } j < i, \\ \mathrm{id} & \text{for } j = i, \\ \partial_{i,\varepsilon}\sigma_{j-1} & \text{for } j > i, \end{cases}$$
$$\gamma_{j,\varepsilon'}\gamma_{i,\varepsilon} = \begin{cases} \gamma_{i,\varepsilon}\gamma_{j+1,\varepsilon'} & \text{for } j > i, \\ \gamma_{i,\varepsilon}\gamma_{i+1,\varepsilon} & \text{for } j = i, \varepsilon' = \varepsilon, \end{cases}$$
$$\gamma_{j,\varepsilon'}\partial_{i,\varepsilon} = \begin{cases} \partial_{i-1,\varepsilon}\gamma_{j,\varepsilon'} & \text{for } j < i-1, \\ \mathrm{id} & \text{for } j \in \{i-1, i\}, \varepsilon = \varepsilon', \\ \partial_{i,\varepsilon}\sigma_i & \text{for } j \in \{i-1, i\}, \varepsilon = 1-\varepsilon', \\ \partial_{i,\varepsilon}\gamma_{j-1,\varepsilon'} & \text{for } j > i, \end{cases}$$
$$\sigma_j\gamma_{i,\varepsilon} = \begin{cases} \gamma_{i-1,\varepsilon}\sigma_j & \text{for } j < i, \\ \sigma_i\sigma_i & \text{for } j = i, \\ \gamma_{i,\varepsilon}\sigma_{j+1} & \text{for } j > i. \end{cases}$$

Let us point out that this is only one of the many choices of the box category that appear in the literature. References such as [Maltsiniotis 2009; Kapulkin et al. 2019] consider a box category that is spanned by faces, degeneracies and one of the connections, specifically the max-connection (although dual arguments can be used to work with min-connections as well). In [Cisinski 2006; Jardine 2006], a subcategory of our $\square$ is considered that is generated by the face and degeneracy maps, but no connections; and in [Steiner 2006], an even smaller subcategory is considered, as it is spanned by the face maps alone. At the other extreme, [Kapulkin and Voevodsky 2020] works with $\square$ as the full subcategory of Cat.

Our choice is intentional. Since our (marked) cubical sets will be used to model (∞, n)-categories, all of our cubes need to have an orientation, and hence the symmetry and diagonal maps appearing in the choices strictly larger than ours are undesirable. On the other hand, the box category with at least one connection is known to have better categorical properties than the ones without connections; see [Tonks 1992; Maltsiniotis 2009]. Finally, allowing for at least one connection, we choose to work with both to allow for a convenient description of the opposite (∞, n)-category.

Given our choice of the box category, we have the following *normal form* of morphisms in $\square$.

Theorem 1.6 [Grandis and Mauri 2003, Theorem 5.1] *Every map in the category $\square$ can be factored uniquely as a composite*

$$(\partial_{c_1,\varepsilon'_1} \cdots \partial_{c_r,\varepsilon'_r})(\gamma_{b_1,\varepsilon_1} \cdots \gamma_{b_q,\varepsilon_q})(\sigma_{a_1} \cdots \sigma_{a_p}),$$

where $1 \le a_1 < \cdots < a_p$, $1 \le b_1 \le \cdots \le b_q$, $b_i < b_{i+1}$ *if* $\varepsilon_i = \varepsilon_{i+1}$ *and* $c_1 > \cdots > c_r \ge 1$. □

With this, one can describe $\square$ as the category generated by the cubical operators, subject to the cubical identities.

Remark 1.7 In particular, any composite face map can be written uniquely as $\alpha = \partial_{k_1,\varepsilon_1} \ldots \partial_{k_t,\varepsilon_t}$ with $k_1 > \cdots > k_t$. Geometrically, such an α is the intersection of all the $\partial_{k_s,\varepsilon_s}$.

This theorem allows us to prove the following key property of $\square$:

Theorem 1.8 *The category $\square$ is an EZ Reedy category with the structure defined as*

- $\deg[1]^n = n$,
- $\square_-$ *is generated under composition by degeneracies and (both kinds of) connections*,
- $\square_+$ *is generated under composition by face maps.*

The key difficulty in proving this theorem lies in showing that each map in $\square_-$ is determined by its sections. This is done by induction on the length of the decomposition of such a map given in Theorem 1.6. Before proceeding with the proof, we state two lemmas. The first of these contains the base case of induction, whereas the second contains the technical heart of the proof — a case analysis allowing us to complete the inductive step.

Lemma 1.9 (1) *The sections of σ_i are $\partial_{i,0}$ and $\partial_{i,1}$.*

(2) *The sections of $\gamma_{i,0}$ are $\partial_{i,0}$ and $\partial_{i+1,0}$.*

(3) *The sections of $\gamma_{i,1}$ are $\partial_{i,1}$ and $\partial_{i+1,1}$.* □

Lemma 1.10 *Given two distinct maps $p, p' \colon [1]^n \to [1]^{n-k}$ in $\square_-$, there is a face map*

$$\partial_{i,\varepsilon} \colon [1]^{n-k} \to [1]^{n-(k-1)}$$

such that $p\partial_{i,\varepsilon} \ne p'\partial_{i,\varepsilon}$ and at least one of $p\partial_{i,\varepsilon}$ and $p'\partial_{i,\varepsilon}$ is in $\square_-$.

Proof We proceed by induction with respect to k with the base case of $k = 1$ handled by Lemma 1.9.

For the inductive step, we may use Theorem 1.6 to write

$$p = \gamma_{i_1,\varepsilon_1} \cdots \gamma_{i_l,\varepsilon_l} \sigma_{j_1} \cdots \sigma_{j_m} \quad \text{and} \quad p' = \gamma_{i'_1,\varepsilon'_1} \cdots \gamma_{i'_{l'},\varepsilon'_{l'}} \sigma_{j'_1} \cdots \sigma_{j'_{m'}},$$

and without loss of generality we may assume that $m \le m'$.

We first suppose that there is an index j_i that does not appear in the set $j'_1, \ldots, j'_{m'}$, ie there is a degeneracy in the decomposition of p that is not present in the decomposition of p'. Then we may note that the normal form of $p\partial_{j_i,0}$ is obtained by removing σ_{j_i} from the normal form of p, and hence the resulting map is in $\square_-$. On the other hand, the normal form $p'\partial_{j_i,0}$ will contain more degeneracy maps than that of $p\partial_{j_i,0}$, since $\partial_{j_i,0}$ will not cancel with any of the degeneracy maps present in the normal form of p' and we assumed $m \le m'$.

If there is no such j_i, then the string $\sigma_{j_1} \cdots \sigma_{j_m}$ is a substring of $\sigma_{j'_1} \cdots \sigma_{j'_{m'}}$. By precomposing with different face maps, we may assume that $m = 0$. We proceed by case analysis, addressing $m' \ge 2$, $m' = 1$ and $m' = 0$ in order.

For $m' \ge 2$, we can write

$$p = \gamma_{i_1,\varepsilon_1} \cdots \gamma_{i_\ell,\varepsilon_\ell} \quad \text{and} \quad p' = \gamma_{i'_1,\varepsilon'_1} \cdots \gamma_{i'_{\ell'},\varepsilon'_{\ell'}} \sigma_{j'_1} \cdots \sigma_{j'_{m'}}.$$

Now observe that $p\partial_{i_\ell,\varepsilon_\ell} = \gamma_{i_1,\varepsilon_1} \cdots \gamma_{i_{\ell-1},\varepsilon_{\ell-1}}$ is in the normal form (and belongs to $\square_-$), but the normal form of $p\partial_{i_l,\varepsilon_l}$ must end with at least one degeneracy.

For $m' = 1$, we can write

$$p = \gamma_{i_1,\varepsilon_1} \cdots \gamma_{i_\ell,\varepsilon_\ell} \quad \text{and} \quad p' = \gamma_{i'_1,\varepsilon'_1} \cdots \gamma_{i'_{\ell-1},\varepsilon'_{\ell-1}} \sigma_{j'}.$$

To treat this case, we will precompose both p and p' with $\partial_{j',\varepsilon}$ to cancel the degeneracy appearing in the normal form of p', yielding a normal form of $p'\partial_{j',\varepsilon}$, which then clearly belongs to $\square_-$. However, some care is needed to choose the correct ε in order to ensure that the normal form of $p\partial_{j',\varepsilon}$ is different from that of $p'\partial_{j',\varepsilon}$. If j' appears in the sequence: $i_1, \ldots, i_\ell$, then we pick $\varepsilon = 1 - \varepsilon_{j'}$. With this choice, the normal form of $p\partial_{j',\varepsilon}$ will end with a degeneracy, making it distinct from $p\partial_{j',\varepsilon}$. If on the other hand j' does not appear on the list of indices: $i_1, \ldots, i_l$, we first need to determine whether when using cubical identities to write $p\partial_{j',\varepsilon}$ in normal form, we will encounter a connection with which our face map will cancel: if not, then we can pick either ε; otherwise, we pick ε in such a way as to ensure that as a result of commuting the face and the connection, we obtain a degeneracy map.

At this point, it remains to treat the case when $m' = 0$, ie the normal forms of p and p' consist solely in connections:

$$p = \gamma_{i_1,\varepsilon_1} \cdots \gamma_{i_k,\varepsilon_k} \quad \text{and} \quad p' = \gamma_{i'_1,\varepsilon'_1} \cdots \gamma_{i'_k,\varepsilon'_k}.$$

Note that the two decompositions have the same length, since both p and p' are maps $[1]^n \to [1]^{n-k}$. Without loss of generality, we may assume that $i_k \le i'_k$, and we proceed by case analysis based on $i'_k - i_k$, considering three cases: $i'_k = i_k$, $i'_k = i_k + 1$ and $i'_k \ge i_k + 2$.

If $i_k = i'_k$, we precompose p and p' with $\partial_{i_k,\varepsilon_k}$. In the case of $\varepsilon_k = \varepsilon'_k$, this reduces us to the inductive hypothesis. If however $\varepsilon_k \ne \varepsilon'_k$, then the normal form of $p\partial_{i_k,\varepsilon_k}$ will be $\gamma_{i_1,\varepsilon_1} \cdots \gamma_{i_{k-1},\varepsilon_{k-1}}$, making it an element of $\square_-$, whereas the normal form of $p'\partial_{i_k,\varepsilon_k}$ will end in a degeneracy.

Next, suppose that $i'_k = i_k + 1$. Then the cases of $\varepsilon_k = \varepsilon'_k$ and $\varepsilon_k \ne \varepsilon'_k$ need to be treated separately. In the former, we precompose with $\partial_{i_k,\varepsilon_k}$. Then $p\partial_{i_k,\varepsilon_k} = \gamma_{i_1,\varepsilon_1} \cdots \gamma_{i_{k-1},\varepsilon_{k-1}}$ is the normal form, making it an element of $\square_-$, but the normal form of $p'\partial_{i_k,\varepsilon_k}$ ends with one of: a degeneracy, a connection of first index greater than i_k, or a connection $\gamma_{i_k,\varepsilon'_k}$, making it distinct from $p\partial_{i_k,\varepsilon_k}$. In the latter case, we see that the normal form of $p\partial_{i_k+1,\varepsilon'_k}$ ends with a degeneracy, but the normal form of $p'\partial_{i_k+1,\varepsilon'_k}$ ends with $\gamma_{i'_{k-1},\varepsilon'_{k-1}}$ and this element belongs to $\square_-$.

Finally, if $i'_k \ge i_k + 2$, then we precompose both p and p' with $\partial_{i_k,\varepsilon_k}$. This gives the normal form of $p\partial_{i_k,\varepsilon_k}$ as $\gamma_{i_1,\varepsilon_1} \cdots \gamma_{i_{k-1},\varepsilon_{k-1}}$, making it an element of $\square_-$. But the normal form of $p'\partial_{i_k,\varepsilon_k}$ ends with either a degeneracy or the connection $\gamma_{i'_k,\varepsilon'_k}$, making it distinct. □

Proof of Theorem 1.8 The Reedy part follows immediately from Theorem 1.6. Every morphism in $\square_-$ is a split epimorphism by Lemma 1.9.

It remains to show that maps in $\square_-$ are determined by their sections. To do this, we pick two such maps $p, p' \colon [1]^n \to [1]^{n-k}$ and proceed by induction with respect to k. The base case of $k = 1$ is handled by Lemma 1.9, whereas the inductive step is handled by Lemma 1.10. □

The category $\square$ carries a canonical strict monoidal product $\otimes$ given by $[1]^m \otimes [1]^n := [1]^{m+n}$ with unit given by $[1]^0$. Note that this product is not cartesian since, for instance, there is no "diagonal" map $[1]^1 \to [1]^2$ in $\square$. This monoidal structure leads to another characterization of our box category, due to Grandis and Mauri [Grandis and Mauri 2003, Section 5], as a certain kind of a free monoidal category.

A *cubical monoid* in a monoidal category $(\mathcal{C}, \otimes, I)$ is an object X equipped with maps

$$\partial_0, \partial_1 \colon I \to X, \quad \sigma \colon X \to I, \quad \gamma_0, \gamma_1 \colon X \otimes X \to X,$$

subject to the axioms

$$\begin{aligned}
\sigma\partial_\varepsilon &= \mathrm{id} \quad \text{for } \varepsilon = 0, 1, \\
\sigma\gamma_\varepsilon &= \sigma(\sigma \otimes \mathrm{id}_X) = \sigma(\mathrm{id}_X \otimes \sigma) \quad \text{for } \varepsilon = 0, 1, \\
\gamma_\varepsilon(\gamma_\varepsilon \otimes \mathrm{id}_X) &= \gamma_\varepsilon(\mathrm{id}_X \otimes \gamma_\varepsilon) \quad \text{for } \varepsilon = 0, 1, \\
\gamma_\varepsilon(\partial_\varepsilon \otimes \mathrm{id}_X) &= \mathrm{id}_X = \gamma_\varepsilon(\mathrm{id}_X \otimes \partial_\varepsilon) \quad \text{for } \varepsilon = 0, 1, \\
\gamma_\varepsilon(\partial_\delta \otimes \mathrm{id}_X) &= \partial_\delta\sigma = \gamma_\varepsilon(\mathrm{id}_X \otimes \partial_\delta) \quad \text{for } \delta \ne \varepsilon.
\end{aligned}$$

Theorem 1.11 [Grandis and Mauri 2003, Theorem 5.2(d)] *The box category $\square$ is the free strict monoidal category equipped with a cubical monoid.* $\square$

Having established basic properties of the box category, we can now define cubical sets and fundamental constructions on them.

Definition 1.12 A *cubical set* is a presheaf $X\colon \square^{\mathrm{op}} \to \mathsf{Set}$. A *cubical map* is a natural transformation of such presheaves. The category of cubical sets and cubical maps will be denoted by cSet.

We write $\square^n$ for the cubical set represented by $[1]^n$ and call it the (*generic*) *n-cube*. The *boundary* of the n-cube, denoted by $\partial\square^n \to \square^n$, is the maximal proper subobject of the representable $\square^n$, ie the union of all of its faces. The subobject of $\square^n$ given by the union of all faces except the $(i, \varepsilon)^{\text{th}}$ one is called the (i, ε)-*open box* and denoted by $\sqcap^n_{i,\varepsilon} \to \square^n$.

Proposition 1.13 *The monomorphisms of* cSet *are the cellular closure of the set*

$$\{\partial\square^n \hookrightarrow \square^n \mid n \geq 0\}.$$

Proof This follows from Theorem 1.6. $\square$

The monoidal product $\otimes$ can be extended via Day convolution from $\square$ to cSet, making $(\mathsf{cSet}, \otimes, \square^0)$ a biclosed monoidal category. We refer to this monoidal product as the *geometric product* of cubical sets.

We adopt the convention of writing the action of cubical operators on the right, eg the $(1,0)$-face of an n-cube $x\colon \square^n \to X$ will be denoted by $x\partial_{1,0}$.

Proposition 1.14 *The geometric product $X \otimes Y$ of cubical sets X and Y admits the following description.*

- *For $n \geq 0$, the n-cubes in $X \otimes Y$ are the formal products $x \otimes y$ of pairs $x \in X_k$ and $y \in Y_\ell$ such that $k + \ell = n$, subject to the identification $(x\sigma_{k+1}) \otimes y = x \otimes (y\sigma_1)$.*
- *For $x \in X_k$ and $y \in Y_\ell$, the faces, degeneracies and connections of the $(k+\ell)$-cube $x \otimes y$ are computed as follows*:

$$(x \otimes y)\partial_{i,\varepsilon} = \begin{cases} (x\partial_{i,\varepsilon}) \otimes y & \text{if } 1 \leq i \leq k, \\ x \otimes (y\partial_{i-k,\varepsilon}) & \text{if } k+1 \leq i \leq k+\ell; \end{cases}$$
$$(x \otimes y)\sigma_i = \begin{cases} (x\sigma_i) \otimes y & \text{if } 1 \leq i \leq k+1, \\ x \otimes (y\sigma_{i-k}) & \text{if } k+1 \leq i \leq k+\ell+1; \end{cases}$$
$$(x \otimes y)\gamma_{i,\varepsilon} = \begin{cases} (x\gamma_{i,\varepsilon}) \otimes y & \text{if } 1 \leq i \leq k, \\ x \otimes (y\gamma_{i-k,\varepsilon}) & \text{if } k+1 \leq i \leq k+\ell. \end{cases}$$

Proof This is proven in [Doherty et al. 2024, Proposition 1.20] in the case of cubical sets with one kind of connection. The proof given there works almost verbatim in our case. $\square$

Given cubes $x \in X_k$ and $y \in Y_\ell$, we may regard them as cubical maps $x\colon \Box^k \to X$ and $y\colon \Box^\ell \to Y$. Then applying the geometric product to these maps yields a map $x \otimes y\colon \Box^{k+\ell} \to X \otimes Y$, which corresponds precisely to the $(k+\ell)$-cube with the same name. Moreover, every n-cube of $X \otimes Y$ arises via this construction for some, perhaps nonunique, pair of cubes $(x\colon \Box^k \to X, y\colon \Box^\ell \to Y)$ for $k+\ell = n$.

Since the identification in Proposition 1.14 only concerns degenerate cubes, we obtain the following corollary.

Corollary 1.15 *A pair of nondegenerate cubes $x \in X_k$, $y \in Y_\ell$ yields a nondegenerate $(k+\ell)$-cube $x \otimes y$ in $X \otimes Y$. Conversely, every nondegenerate cube in $X \otimes Y$ arises this way from a unique pair of nondegenerate cubes.* □

Remark 1.16 In particular, when $X = \Box^m$ and $Y = \Box^n$ are representable, this pairing is given by the formula

$$(\partial_{k_1,\varepsilon_1} \cdots \partial_{k_t,\varepsilon_t}) \otimes (\partial_{\ell_1,\eta_1} \cdots \partial_{\ell_s,\eta_s}) = \partial_{m+\ell_1,\eta_1} \cdots \partial_{m+\ell_s,\eta_s} \partial_{k_1,\varepsilon_1} \cdots \partial_{k_t,\varepsilon_t},$$

where all strings of ∂'s are in the normal form specified by Theorem 1.6. The factors are permuted because the geometric product lists cubes in order (in the sense that x in $x \otimes y$ corresponds to smaller values of i) whereas the normal form lists faces in reverse order.

Proposition 1.17 *For natural numbers k, m and n, and $\varepsilon = 0, 1$, we have natural isomorphisms*

$$(\partial\Box^m \hookrightarrow \Box^m) \hat{\otimes} (\partial\Box^n \hookrightarrow \Box^n) \cong (\partial\Box^{m+n} \hookrightarrow \Box^{m+n}),$$
$$(\sqcap^m_{k,\varepsilon} \hookrightarrow \Box^m) \hat{\otimes} (\partial\Box^n \hookrightarrow \Box^n) \cong (\sqcap^{m+n}_{k,\varepsilon} \hookrightarrow \Box^{m+n}),$$
$$(\partial\Box^m \hookrightarrow \Box^m) \hat{\otimes} (\sqcap^n_{k,\varepsilon} \hookrightarrow \Box^n) \cong (\sqcap^{m+n}_{m+k,\varepsilon} \hookrightarrow \Box^{m+n}).$$

Proof This follows from [Doherty et al. 2024, Lemma 1.26] and the associativity of $\hat{\otimes}$. □

Using the above proposition and the fact that $\Box$ is an elegant Reedy category, we obtain:

Corollary 1.18 *If f and g are monomorphisms in* cSet, *then $f \hat{\otimes} g$ is again a monomorphism.* □

The category $\Box$ admits two canonical identity-on-objects automorphisms $(-)^{\mathrm{co}}, (-)^{\mathrm{co\text{-}op}}\colon \Box \to \Box$. The first one takes $\partial^n_{i,\varepsilon}$ to $\partial^n_{n+1-i,\varepsilon}$, σ^n_i to σ^n_{n+1-i}, and $\gamma^n_{i,\varepsilon}$ to $\gamma^n_{(n-1)+1-i,\varepsilon}$. The second one takes $\partial^n_{i,\varepsilon}$ to $\partial^n_{i,1-\varepsilon}$, σ^n_i to σ^n_i and $\gamma^n_{i,\varepsilon}$ to $\gamma^n_{i,1-\varepsilon}$. (Their names are motivated by the fact that, according to the source/target distinction described in Section 3 below, $(-)^{\mathrm{co}}$ reverses the direction of even-dimensional cubes and $(-)^{\mathrm{co\text{-}op}}$ reverses the direction of all cubes.) Precomposition with these automorphisms induces functors also denoted by $(-)^{\mathrm{co}}, (-)^{\mathrm{co\text{-}op}}\colon \mathsf{cSet} \to \mathsf{cSet}$. Moreover, $(-)^{\mathrm{co}} \circ (-)^{\mathrm{co\text{-}op}} = (-)^{\mathrm{co\text{-}op}} \circ (-)^{\mathrm{co}}$, yielding a third automorphism $(-)^{\mathrm{op}} := (-)^{\mathrm{co}} \circ (-)^{\mathrm{co\text{-}op}}$.

The "contravariant" behavior of these automorphisms with respect to the cubical structure can be seen via their interaction with the geometric product.

Proposition 1.19 (1) *The functor $(-)^{\mathrm{co}} \colon \mathsf{cSet} \to \mathsf{cSet}$ is strong antimonoidal, ie $(X \otimes Y)^{\mathrm{co}} \cong Y^{\mathrm{co}} \otimes X^{\mathrm{co}}$, naturally in X and Y.*

(2) *The functor $(-)^{\mathrm{co\text{-}op}} \colon \mathsf{cSet} \to \mathsf{cSet}$ is strong monoidal, ie $(X \otimes Y)^{\mathrm{co\text{-}op}} \cong X^{\mathrm{co\text{-}op}} \otimes Y^{\mathrm{co\text{-}op}}$, naturally in X and Y.*

(3) *The functor $(-)^{\mathrm{op}} \colon \mathsf{cSet} \to \mathsf{cSet}$ is strong antimonoidal, ie $(X \otimes Y)^{\mathrm{op}} \cong Y^{\mathrm{op}} \otimes X^{\mathrm{op}}$, naturally in X and Y.* □

Finally, the composite $\Box \to \mathsf{Cat} \to \mathsf{sSet}$ given by $\Box^n \mapsto (\Delta^1)^n$ defines a cocubical object in the category of simplicial sets. Taking the Yoneda extension, we obtain an adjoint pair

$$T \colon \mathsf{cSet} \rightleftarrows \mathsf{sSet} : U.$$

We will call $T \colon \mathsf{cSet} \to \mathsf{sSet}$ the *triangulation* functor.

The triangulation functor can also be seen through the lenses of Theorem 1.11. The simplicial faces $\partial_1, \partial_0 \colon [0] \to [1]$ and degeneracy $\sigma_0 \colon [1] \to [0]$ maps, along with $\max, \min \colon [1]^2 \to [1]$, equip $[1]$ with the structure of a cubical monoid. Since the nerve functor preserves products, this gives a structure of a cubical monoid on Δ^1 in sSet. The triangulation functor $T \colon \mathsf{cSet} \to \mathsf{sSet}$ arises from this cubical monoid via Theorem 1.11.

We conclude this section by recording some properties of the triangulation functor.

Proposition 1.20 (1) *T is strong monoidal.*

(2) *T preserves monomorphisms.*

Proof The first statement follows by the fact that T preserves colimits and sSet is cartesian closed.

The second statement follows from first, since T takes boundary inclusions, ie the elements of the cellular model, to monomorphisms. □

1.3 Complicial sets

In this section, we recall marked simplicial sets and model structures for (n-trivial) complicial sets from [Verity 2008a, 2008b]. The reader is referred to those papers for more detail on the subject. The theory developed in Section 3 draws great insight from this simplicial precursor.

Just as in the case of cubical sets, when working with simplicial sets, we will write the action of simplicial operators on the right.

Definition 1.21 A *marked simplicial set* is a simplicial set X equipped with a subset eX of its simplices called the *marked* simplices such that

- no 0-simplex is marked, and
- every degenerate simplex is marked.

A *map* of marked simplicial sets $f \colon X \to Y$ is a map of simplicial sets which carries marked simplices to marked simplices. We denote sSet^+ for the category of marked simplicial sets with maps for morphisms.

Marked simplicial sets used to be called *stratified simplicial sets* (see eg [Verity 2008a]), but the name "marked" is more descriptive and has since become more popular.

There is a natural forgetful functor $\mathsf{sSet}^+ \to \mathsf{sSet}$, which has both left and right adjoints. The left adjoint $X \mapsto X^\flat$ endows a simplicial set X with the *minimal marking*, marking only the degenerate simplices. The right adjoint $X \mapsto X^\sharp$ endows a simplicial set X with the *maximal marking*, marking all simplices.

If X is a simplicial set, we will by default consider it as a marked simplicial set *with its minimal marking* $X^\flat$.

Definition 1.22 We say that $X \in \mathsf{sSet}^+$ is *n-trivial* if every simplex of dimension $\geq n+1$ is marked.

Given a marked simplicial set X, we will write $\mathrm{core}_n X$ for its maximal n-trivial subset. In other words, the k-simplices of $\mathrm{core}_n X$ are precisely those k-simplices x in X such that $x\alpha$ is marked in X for any $\alpha\colon [m] \to [k]$ with $m > n$. This assignment extends to a functor $\mathrm{core}_n\colon \mathsf{sSet}^+ \to \mathsf{sSet}^+$, which admits a left adjoint $\tau_n\colon \mathsf{sSet}^+ \to \mathsf{sSet}^+$. Explicitly, τ_n acts as the identity on the underlying simplicial set and a k-simplex is marked in $\tau_n X$ if either $k \leq n$ and x is marked in X or $k \geq n+1$.

Definition 1.23 A map $X \to Y$ of marked simplicial sets is

- *regular* if it creates markings, ie for an n-simplex x of X we have that $x \in eX_n$ if and only if $f(x) \in eY_n$, and
- *entire* if the induced map between the underlying simplicial sets is invertible.

We now define several distinguished objects and maps in sSet^+. These will be essential to the description of various model structures we will be considering.

We denote by $\widetilde{\Delta}^n = \tau_{n-1}(\Delta^n)$ the n-simplex with the nondegenerate n-simplex marked and no other nondegenerate simplices marked. We call the canonical map $\Delta^n \to \widetilde{\Delta}^n$ the *n-marker*.

For $n \geq 1$ and $0 \leq k \leq n$, we denote by Δ^n_k the n-simplex with the following marking: a nondegenerate simplex is marked if and only if it contains all of the points $\{k-1, k, k+1\} \cap [n]$ among its vertices. We call Δ^n_k the *k-complicial n-simplex*. We denote by $\Lambda^n_k \subset \Delta^n_k$ the k-horn of dimension n (ie the simplicial subset missing the nondegenerate n-simplex and the k^{th} $(n-1)$-face) endowed with the marking making it a regular subset of Δ^n_k. We call Λ^n_k the *complicial k-horn of dimension n*. We call the inclusion $\Lambda^n_k \to \Delta^n_k$ the *k-complicial horn inclusion of dimension n*. We write $\Delta^{n\prime\prime}_k = \tau_{n-2}\Delta^n_k$, and we write $\Delta^{n\prime}_k = \Delta^n_k \amalg_{\Lambda^n_k} \tau_{n-2}\Lambda^n_k$. The canonical inclusion $\Delta^{n\prime}_k \to \Delta^{n\prime\prime}_k$ is called the *elementary k-complicial marking extension of dimension n*.

Let Δ^3_{eq} denote the marked simplicial set obtained from Δ^3 by marking the 1-simplices $\{0,2\}$, $\{1,3\}$ and all 2- and 3-simplices. By a *saturation map*, we mean a map of the form $\Delta^m \star \Delta^3_{\mathrm{eq}} \to \Delta^m \star (\Delta^3)^\sharp$ for $m \geq -1$ (where Δ^{-1} is interpreted as $\varnothing$).

There are two standard model structures on marked simplicial sets:

Theorem 1.24 *The category* sSet^+ *carries two model structures*:

(1) *The* ***complicial model structure*** *characterized by the following properties*:
- *The cofibrations are the monomorphisms.*
- *The set of*
  - *complicial horn inclusions*, *and*
  - *elementary complicial marking extensions*

  forms a pseudogenerating set of trivial cofibrations.

(2) *The* ***saturated complicial model structure*** *characterized by the following properties*:
- *The cofibrations are the monomorphisms.*
- *The set of*
  - *complicial horn inclusions*,
  - *elementary complicial marking extensions*, *and*
  - *saturation maps*

  forms a pseudogenerating set of trivial cofibrations.

Both of these model structures are cartesian.

Proof This is a combination of [Verity 2008b, Lemma 72, Theorem 100 and Lemma 105] and [Ozornova and Rovelli 2020, Appendix B]. □

Note that since the terminal object is always fibrant, this includes a characterization of the fibrant objects of the model structures, which are called (*saturated*) *complicial sets*.

Definition 1.25 A map of marked simplicial sets $X \to Y$ is

- a *complicial marking extension* if it is in the cellular closure of the elementary complicial marking extensions, and
- *complicial* if it is in the cellular closure of the complicial horn inclusions and the elementary complicial marking extensions.

There is also the n-trivial version of the (saturated) complicial model structure.

Theorem 1.26 *The category* sSet^+ *carries two model structures*:

(1) *The* ***n-trivial complicial model structure*** *characterized by the following properties*:
- *The cofibrations are the monomorphisms.*
- *The set of*
  - *complicial horn inclusions*,
  - *elementary complicial marking extensions of dimension* $\leq n+1$, *and*
  - *markers of dimension* $> n$

  forms a pseudogenerating set of trivial cofibrations.

(2) *The* ***saturated n-trivial complicial model structure*** *characterized by the following properties*:
- *The cofibrations are the monomorphisms.*
- *The set of*
  - *complicial horn inclusions*,
  - *elementary complicial marking extensions of dimension* $\leq n+1$,
  - *markers of dimension* $> n$, *and*
  - *saturation maps*

  forms a pseudogenerating set of trivial cofibrations.

Proof Essentially the proof of Theorem 1.24, but combined with [Verity 2008b, Example 104]. □

In sSet^+, the pseudo Gray tensor product is modeled by the cartesian product. We will adopt the following notation from [Verity 2008b], which emphasizes this view.

Notation 1.27 The cartesian product on sSet^+ (and its reflective subcategory $\mathsf{PreComp}$ described below) is denoted by $\circledast$.

Thus Theorem 1.24 in particular says that Verity's model structure is monoidal with respect to the pseudo Gray tensor product.

The following proposition will be useful later.

Proposition 1.28 *Let $f\colon A \to X$ and $g\colon B \to Y$ be entire maps in* sSet^+. *Then their Leibniz pseudo Gray tensor $f \mathbin{\widehat{\circledast}} g$ is a complicial marking extension.*

Proof Since the forgetful functor $\mathsf{sSet}^+ \to \mathsf{sSet}$ preserves colimits and products, $f \mathbin{\widehat{\circledast}} g$ is entire. We assume for the sake of simplicity that $f \mathbin{\widehat{\circledast}} g$ is an inclusion. Let (x, y) be an n-simplex that is marked in $X \circledast Y$ but not in $\operatorname{dom}(f \mathbin{\widehat{\circledast}} g) = (A \circledast Y) \cup (X \circledast B)$. Equivalently, x is marked in X but not in A, and y is marked in Y but not in B. Then we must have $n \geq 1$, so the $(n+1)$-simplex $z = (x\sigma_0, y\sigma_1)$ is well-defined. We claim that this simplex z extends as indicated below:

$$\begin{tikzcd} \Delta^{n+1} \arrow[r, "z"] \arrow[d] & (A \circledast Y) \cup (X \circledast B) \\ \Delta_1^{n+1\prime} \arrow[ur, dashed, "\exists"'] & \end{tikzcd}$$

To see that z at least extends to Δ_1^{n+1}, let $\alpha\colon [m] \to [n+1]$ be a simplicial operator with $0, 1, 2 \in \operatorname{im}\alpha$. Then both $x(\sigma_0 \circ \alpha)$ and $y(\sigma_1 \circ \alpha)$ are degenerate, so $z\alpha$ is marked in $\operatorname{dom}(f \mathbin{\widehat{\circledast}} g)$. Since the face $z\partial_0 = (x, y(\partial_0 \circ \sigma_0))$ is marked in $X \circledast B$ and the face $z\partial_2 = (x(\partial_1 \circ \sigma_0), y)$ is marked in $A \circledast Y$, we indeed have an extension as indicated. Therefore we have a pushout square

$$\begin{tikzcd} \coprod \Delta_1^{n+1'} \arrow[r] \arrow[d] & (A \circledast Y) \cup (X \circledast B) \arrow[d] \\ \coprod \Delta_1^{n+1''} \arrow[r] & X \circledast Y \end{tikzcd}$$

where the coproducts are taken over all n-simplices (x, y) that are marked in $X \circledast Y$ but not in $\mathrm{dom}(f \hat{\circledast} g)$, and both horizontal maps are induced by the simplices of the form $(x\sigma_0, y\sigma_1)$. $\square$

Definition 1.29 Let $[n] \in \Delta$ and let $0 \le p, q \le n$ be such that $p + q = n$. Then we write $\amalg_1^{p,q} : [p] \to [n]$ for the simplicial operator $i \mapsto i$, and $\amalg_2^{p,q} : [q] \to [n]$ for the operator $i \mapsto p + i$.

In the following definition, we use slightly different terminology from Verity's original one [2008a, Definitions 127 and 128].

Definition 1.30 Let $X, Y \in \mathsf{sSet}^+$, let $(x, y) \in X_n \times Y_n$ be a simplex of $X \times Y$, and let $0 \le i \le n$. We say that (x, y) is *i-cloven* if *either* $x \amalg_1^{i,n-i}$ is marked in X or $y \amalg_2^{i,n-i}$ is marked in Y. We say that (x, y) is *fully cloven* if it is i-cloven for all $0 \le i \le n$.

The *Gray tensor product* of X and Y, denoted by $X \otimes Y$, is defined to be the marked simplicial set with underlying simplicial set $X \times Y$, where a simplex $(x, y) \in X_n \times Y_n$ is marked if and only if it is fully cloven.

Theorem 1.31 [Verity 2008a, Lemma 131] *The Gray tensor product endows the category sSet^+ with a (nonsymmetric) monoidal structure such that the forgetful functor $(\mathsf{sSet}^+, \otimes) \to (\mathsf{sSet}, \times)$ is strict monoidal.*

Definition 1.32 A *precomplicial set* is a marked simplicial set X with the right lifting property with respect to the complicial marking extensions. These form a reflective subcategory of sSet^+, which we will denote by $\mathsf{PreComp}$. We will denote the localization functor by $X \mapsto X^{\mathrm{pre}}$.

Proposition 1.33 *The unit of the reflection $X \to X^{\mathrm{pre}}$ is a complicial marking extension for any $X \in \mathsf{sSet}^+$.*

Proof Obtain a complicial marking extension $f : X \to Y$ into a precomplicial set Y by applying the small object argument to the unique map $X \to 1$ with respect to the elementary complicial marking extensions. Then any map $X \to Z$ into a precomplicial set Z factors through f. Moreover, since f is an epimorphism, such a factorization is necessarily unique. In other words, f has the universal property of the unit $X \to X^{\mathrm{pre}}$. $\square$

Theorem 1.34 *For each of the four model structures in Theorems 1.24 and 1.26, the category* PreComp *carries an analogous model structure characterized by the following conditions*:

- *The cofibrations are the monomorphisms.*
- *A pseudogenerating set of trivial cofibrations can be obtained by taking that for the corresponding model structure on* sSet^+ (*described in Theorem 3.3 or Theorem 1.26*), *removing the complicial marking extensions*, *and then applying the precomplicial reflection.*

The localization $(-)^{\mathrm{pre}}$ *is a left Quillen equivalence between the complicial model structures* (*resp. the* n*-trivial complicial model structures*).

Proof Fix one of the four model structures on sSet^+. Observe that if we factor a map between precomplicial sets into a cofibration followed by a fibration (one of which is trivial) with respect to that model structure, then the middle object must also be precomplicial. Thus we obtain a restricted model structure on PreComp.

We obtain the characterization of its cofibrations by observing that the reflective inclusion $\mathsf{PreComp} \hookrightarrow \mathsf{sSet}^+$ preserves and reflects monomorphisms. It is straightforward to check that the precomplicial reflection preserves pseudogenerating sets, and moreover it inverts all (elementary) complicial marking extensions.

It follows that the adjunction $\mathsf{sSet}^+ \rightleftarrows \mathsf{PreComp}$ is a Quillen adjunction. Finally, since the unit is a natural weak equivalence this is in fact a Quillen equivalence. $\square$

Remark 1.35 We believe that the precomplicial reflection does not actually affect the remaining members of the pseudogenerating set. However we do not provide a proof as it is not essential.

Now we analyze the precomplicial reflection of the Gray tensor product on PreComp.

Definition 1.36 We write $\otimes^{\mathrm{pre}}\colon \mathsf{PreComp} \times \mathsf{PreComp} \to \mathsf{PreComp}$ for the precomplicial Gray tensor product functor $(X, Y) \mapsto (X \otimes Y)^{\mathrm{pre}}$.

Theorem 1.37 *The bifunctor* $\otimes^{\mathrm{pre}}$ *is part of a biclosed monoidal structure on* PreComp. *Moreover each model structure on* PreComp *described in Theorem 1.34 is monoidal with respect to* $\otimes^{\mathrm{pre}}$.

Proof The first assertion is [Verity 2008a, Theorem 148]. It is straightforward to check that the Leibniz Gray tensor product preserves monomorphisms. Since the complicial model structure on sSet^+ is monoidal with respect to the Gray tensor product [Verity 2008b, Theorem 109] (although it is not biclosed on sSet^+) and the unit of the precomplicial reflection is a levelwise complicial marking extension, it follows that the complicial model structure on PreComp is monoidal. The n-trivial and saturated versions follow from [Ozornova and Rovelli 2020, Section 2]. $\square$

2 Marked cubical sets and Gray tensor products

In this section, we introduce marked cubical sets and define their Gray tensor product.

2.1 Marked cubical sets

In order to define marked cubical sets, we need to introduce certain enlargement $\Box^+$ of the box category. The objects of $\Box^+$ consist of: $[1]^n$ for every $n \geq 0$ and $[1]^n_e$ for every $n \geq 1$. The morphisms of $\Box^+$ are generated by the maps

$$
\begin{aligned}
\partial^n_{i,\varepsilon} \colon [1]^{n-1} &\to [1]^n && \text{for every } n \geq 1, i = 1, \ldots, n \text{ and } \varepsilon = 0, 1,\\
\sigma^n_i \colon [1]^n &\to [1]^{n-1} && \text{for } n \geq 1 \text{ and } i = 1, \ldots, n,\\
\gamma^n_{i,\varepsilon} \colon [1]^n &\to [1]^{n-1} && \text{for } n \geq 2, i = 1, \ldots, n-1 \text{ and } \varepsilon = 0, 1,\\
\varphi^n \colon [1]^n &\to [1]^n_e && \text{for } n \geq 1,\\
\zeta^n_i \colon [1]^n_e &\to [1]^{n-1} && \text{for } n \geq 1 \text{ and } i = 1, \ldots, n,\\
\xi^n_{i,\varepsilon} \colon [1]^n_e &\to [1]^{n-1} && \text{for } n \geq 1, i = 1, \ldots, n \text{ and } \varepsilon = 0, 1,
\end{aligned}
$$

subject to the usual cubical identities and the following additional relations:

$$
\zeta_i \varphi = \sigma_i, \qquad \xi_{i,\varepsilon}\varphi = \gamma_{i,\varepsilon}, \qquad \sigma_i \zeta_j = \sigma_j \zeta_{i+1} \quad \text{for } j \leq i,
$$

$$
\gamma_{j,\varepsilon}\xi_{i,\delta} = \begin{cases} \gamma_{i,\delta}\xi_{j+1,\varepsilon} & \text{for } j > i,\\ \gamma_{i,\delta}\xi_{i+1,\delta} & \text{for } j = i \text{ and } \delta = \varepsilon, \end{cases} \qquad \sigma_j \xi_{i,\delta} = \begin{cases} \gamma_{i-1,\delta}\zeta_j & \text{for } j < i,\\ \sigma_i \zeta_i & \text{for } j = i,\\ \gamma_{i,\delta}\zeta_{j+1} & \text{for } j > i. \end{cases}
$$

Proposition 2.1 *The category $\Box^+$ is an EZ Reedy category with the following Reedy structure*:

- $\deg[1]^0 = 0$, $\deg[1]^n = 2n-1$ *for* $n \geq 1$, *and* $\deg[1]^n_e = 2n$ *for* $n \geq 1$.
- $\Box^+_-$ *is generated by the maps* σ^n_i, $\gamma^n_{i,\varepsilon}$, ζ^n_i *and* $\xi^n_{i,\varepsilon}$.
- $\Box^+_+$ *is generated by the maps* $\partial^n_{i,\varepsilon}$ *and* φ^n.

The proof of this fact follows the one in [Ozornova and Rovelli 2020, Appendix C]. We begin by noting the following simple lemma:

Lemma 2.2 (1) *The are no nonidentity maps in $\Box^+_-$ whose target is in $\Box^+ \setminus \Box$.*

(2) *The are no nonidentity maps in $\Box^+_+$ whose source is in $\Box^+ \setminus \Box$.* □

Proof of Proposition 2.1 We first note that the sections of ζ_i are $\varphi\partial_{i,0}$ and $\varphi\partial_{i,1}$, the sections of $\xi_{i,1}$ are $\varphi\partial_{i,0}$ and $\varphi\partial_{i+1,0}$, and the sections of $\xi_{i,0}$ are $\varphi\partial_{i,1}$ and $\varphi\partial_{i+1,1}$. Thus all maps in $\Box^+_-$ have sections.

Using the techniques of [Grandis and Mauri 2003, Theorem 5.1], we can then extend Theorem 1.6 to write normal forms for maps in $\Box^+$. These are established separately for the four cases:

(1) The normal form of a map of the form $[1]^m \to [1]^n$ is given by its normal form in $\square$. If such a form were to be nonunique, we would need to have $[1]^m \to [1]^k_e \to [1]^n$ with $[1]^k_e \to [1]^n \in \square^+_+$, which is impossible by Lemma 2.2(2).

(2) The normal form of a map $[1]^m \to [1]^n_e$ is obtained by observing that it is necessarily a composite $[1]^m \to [1]^n \xrightarrow{\varphi} [1]^n_e$ and taking the normal form of the first map in $\square$. Again, Lemma 2.2(2) implies uniqueness.

(3) The normal form of a map of the form $[1]^m_e \to [1]^n$ is obtained by factoring it as $[1]^m_e \to [1]^{m-1} \to [1]^n$, where the first map is either ζ_i or $\xi_{i,\varepsilon}$, and taking the normal form of Theorem 1.6 of $[1]^{m-1} \to [1]^n$ in $\square$. Note that the choice of ζ_i or $\xi_{i,\varepsilon}$ as the first map may not be unique, but it can be made so by imposing the additional compatibility requirement with the factorization of Theorem 1.6 — this is because of the additional relations relating the ζ_i to the σ_i, and the $\xi_{i,\varepsilon}$ to the $\gamma_{i,\varepsilon}$. Put differently, we may precompose $[1]^m_e \to [1]^n$ with φ, use the normal form in $\square$, and replace the last element by ζ_i or $\xi_{i,\varepsilon}$ as appropriate.

(4) $[1]^m_e \to [1]^n_e$. In this case, we obtain the normal form by combining the techniques from the previous two cases, namely factoring

$$[1]^m_e \to [1]^{m-1} \to [1]^n \xrightarrow{\varphi} [1]^n_e,$$

where again the first map is one of ζ_i or $\xi_{i,\varepsilon}$, and the composite $[1]^{m-1} \to [1]^k \to [1]^n$ is obtained in $\square$.

Having established the normal forms, we proceed in a manner analogous to the proof of Theorem 1.8. $\square$

Definition 2.3 A *structurally marked cubical set* is a presheaf $X\colon (\square^+)^{\mathrm{op}} \to \mathsf{Set}$. A *map of structurally marked cubical sets* is a natural transformation of such presheaves.

Given a structurally marked cubical set X, we will write X_n for $X([1]^n)$ and eX_n for $X([1]^n_e)$. Just as in the case of cubical sets, we adopt the convention of writing cubical operators on the right, eg for $x \in eX_1$, we write $x\varphi$ for the resulting element of X_1.

Definition 2.4 A *marked cubical set* is a structurally marked cubical set $X\colon (\square^+)^{\mathrm{op}} \to \mathsf{Set}$ for which the map $X\varphi\colon eX_n \to X_n$ is a monomorphism for all $n \geq 1$. We write cSet^+ for the full subcategory of $\mathsf{Set}^{(\square^+)^{\mathrm{op}}}$ spanned by the marked cubical sets.

We think of a marked cubical set X as a cubical set in which certain n-cubes have been designated as *equivalences*, ie those in $eX_n \subseteq X_n$. The maps ζ_i and $\xi_{i,\varepsilon}$ ensure that every degenerate cube is marked.

We may apply the same intuition to structurally marked cubical sets. However, failure of the maps $X\varphi$ to be monomorphisms (in an arbitrary structurally marked cubical set X) means that being an equivalence is not a property of an n-cube of X, but a structure on it, as there can be multiple markings on a single cube.

Every (structurally) marked cubical set has an underlying cubical set, defining a functor $\upsilon\colon \mathsf{cSet}^+ \to \mathsf{cSet}$. Given a cubical set X, we can form a marked cubical set in two ways:

- The *minimal marking* functor takes a cubical set X to a marked cubical set $X^\flat$, where only degenerate n-cubes are marked.
- The *maximal marking functor* assigns to X the marked cubical set $X^\sharp$ in which all cubes marked (ie all maps $X\varphi^n$ are identities).

This gives two functors $(-)^\flat, (-)^\sharp \colon \mathsf{cSet} \to \mathsf{cSet}^+$. A straightforward verification shows:

Proposition 2.5 *We have a string of adjoint functors* $(-)^\flat \dashv \upsilon \dashv (-)^\sharp$. □

Remark 2.6 (limits and colimits of marked cubical sets) The proposition above gives a recipe for computing limits and colimits of diagrams $F \colon \mathscr{I} \to \mathsf{cSet}^+$. In both cases, we first compute the underlying cubical set by taking the (co)limit of υF in cSet, and then equipping it with the minimal marking making the colimit inclusions maps of marked cubical sets, or the maximal marking making the limit projections maps of marked cubical sets. It follows, for example, that a cube in a colimit is marked if and only if it is in the image of a marked cube under one of the colimit inclusions.

Furthermore, the canonical embedding $\mathsf{cSet}^+ \hookrightarrow \mathsf{Set}^{(\Box^+)^{\mathrm{op}}}$ of marked cubical sets into structurally marked cubical sets admits a left adjoint, denoted by $\mathrm{Im} \colon \mathsf{Set}^{(\Box^+)^{\mathrm{op}}} \to \mathsf{cSet}^+$. Explicitly, $\mathrm{Im}\, X$ is obtained by factoring all the φ_n via their image $eX_n \to (eX_n)\varphi_n \to X_n$ and taking the resulting object as a new set of marked n-cubes. We may summarize it with the following statement:

Proposition 2.7 *Marked cubical sets form a reflective subcategory of the structurally marked cubical sets with the reflector given by* $\mathrm{Im} \colon \mathsf{Set}^{(\Box^+)^{\mathrm{op}}} \to \mathsf{cSet}^+$. □

Corollary 2.8 *The category* cSet^+ *of marked cubical sets is locally presentable.* □

Definition 2.9 A map $f \colon X \to Y$ of marked cubical sets is

- *regular* if it creates markings, ie for an n-cube x of X we have: $x \in eX_n$ if and only if $f(x) \in eY_n$,
- *entire* if the induced map between the underlying cubical sets is invertible.

Definition 2.10 We say that $X \in \mathsf{cSet}^+$ is *n-trivial* if every cube of dimension $\geq n+1$ is marked.

Given a marked cubical set X, we will write $\mathrm{core}_n X$ for its maximal n-trivial subset. In other words, the k-cubes of $\mathrm{core}_n X$ are precisely those k-cubes x such that $x\alpha$ is marked for all $\alpha \colon [1]^m \to [1]^k$ with $m > n$. This assignment extends to a functor $\mathrm{core}_n \colon \mathsf{cSet}^+ \to \mathsf{cSet}^+$, which admits a left adjoint $\tau_n \colon \mathsf{cSet}^+ \to \mathsf{cSet}^+$. Explicitly, τ_n acts as the identity on the underlying cubical set and a k-cube is marked in $\tau_n X$ if either $k \leq n$ and x is marked in X or $k \geq n+1$.

When a cubical set is considered as a marked cubical set, it will *almost* always be considered with its minimal marking. The only exception is the open boxes; see Section 3. We denote by $\Box^n = (\Box^n)^\flat$ the

n-cube regarded as a marked cubical set and likewise $\partial\Box^n = (\partial\Box^n)^\flat$. Just as in the case of cubical sets, we call the inclusion map $\partial\Box^n \to \Box^n$ the *boundary inclusion*. We denote by $\widetilde{\Box}^n = \tau_{n-1}(\Box^n)$ the n-cube with the nondegenerate n-cube marked and no other nondegenerate cubes marked. We call the canonical map $\Box^n \to \widetilde{\Box}^n$ the *n-marker*.

Proposition 2.11 *The monomorphisms of* cSet^+ (*and* $\mathsf{Set}^{(\Box^+)^{\mathrm{op}}}$) *are the cellular closure of the set*

$$\{\partial\Box^n \hookrightarrow \Box^n \mid n \geq 0\} \cup \{\Box^n \hookrightarrow \widetilde{\Box}^n \mid n \geq 1\}.$$

The functors $(-)^{\mathrm{co}}, (-)^{\mathrm{co\text{-}op}}, (-)^{\mathrm{op}} \colon \mathsf{cSet} \to \mathsf{cSet}$ generalize to the marked setting in the straightforward manner. For $(-)^{\mathrm{co}}$ we send φ^n to itself, ζ_i^n to ζ_{n+1-i}^n, and $\xi_{i,\varepsilon}^n$ to $\xi_{n+1-i,\varepsilon}^n$. For $(-)^{\mathrm{co\text{-}op}}$, we send φ^n and ζ_i^n to themselves and $\xi_{i,\varepsilon}^n$ to $\xi_{i,1-\varepsilon}^n$. These then induce functors by precomposition $(-)^{\mathrm{co}}, (-)^{\mathrm{co\text{-}op}}, (-)^{\mathrm{op}} \colon \mathsf{cSet}^+ \to \mathsf{cSet}^+$.

2.2 Gray tensor products

The following definition makes use of Corollary 1.15.

Definition 2.12 The (*lax*) *Gray tensor product* $X \otimes Y$ of two marked cubical sets $X, Y \in \mathsf{cSet}^+$ is the geometric product $\upsilon X \otimes \upsilon Y$, wherein a nondegenerate cube $x \otimes y$ is marked if and only if either x is marked in X or y is marked in Y. This extends to a functor $\otimes \colon \mathsf{cSet}^+ \times \mathsf{cSet}^+ \to \mathsf{cSet}^+$ in the obvious way.

Definition 2.13 The *pseudo Gray tensor product* $X \circledast Y$ is the geometric product $\upsilon X \otimes \upsilon Y$, wherein a nondegenerate cube $x \otimes y$ is unmarked if and only if

- x is a 0-cube and y is unmarked in Y, or
- x is unmarked in X and y is a 0-cube.

This extends to a functor $\circledast \colon \mathsf{cSet}^+ \times \mathsf{cSet}^+ \to \mathsf{cSet}^+$ in the obvious way.

Remark 2.14 Since no 0-cubes are marked, one can easily check that $X \circledast Y$ may be obtained from $X \otimes Y$ by marking those nondegenerate $x \otimes y$ such that $x \in X_m$, $y \in Y_n$ with $m, n \geq 1$. Thus the identity at $\upsilon X \otimes \upsilon Y$ lifts to an entire map $\mu_{X,Y} \colon X \otimes Y \to X \circledast Y$. This map is clearly natural in X and Y, and moreover $\mu_{X,Y}$ is invertible if either X or Y is 0-trivial.

Remark 2.15 The Gray tensor products $\otimes$ and $\circledast$ share many properties, and often a statement or a proof applies equally well to both tensor products. In such situations, we write $\odot$ to mean either. Of course the interpretation of $\odot$ is to be kept consistent within each statement and its proof.

Theorem 2.16 (1) *The Gray tensor product $\odot$ forms part of a biclosed monoidal structure on* cSet^+ *such that the forgetful functor $\upsilon\colon (\mathsf{cSet}^+, \odot) \to (\mathsf{cSet}, \otimes)$ is strict monoidal.*

(2) *The entire inclusions $\mu_{X,Y}\colon X \otimes Y \to X \circledast Y$ together with $\mu_0 = \mathrm{id}_{\square^0}$ equip the identity functor with a monoidal structure $(\mathrm{id}_{\mathsf{cSet}^+}, \mu)\colon (\mathsf{cSet}^+, \circledast) \to (\mathsf{cSet}^+, \otimes)$.*

(3) *The minimal marking functor $(-)^\flat\colon (\mathsf{cSet}, \otimes) \to (\mathsf{cSet}^+, \otimes)$ is strict monoidal.*

(4) *The maximal marking functor $(-)^\sharp\colon (\mathsf{cSet}, \otimes) \to (\mathsf{cSet}^+, \odot)$ is strict monoidal.*

Proof We first check the associativity of the tensor product. Suppose we are given nondegenerate cubes $x \in X_m$, $y \in Y_n$, $z \in Z_k$ in $X, Y, Z \in \mathsf{cSet}^+$. Then the $(m+n+k)$-cube $(x \otimes y) \otimes z$ in $(X \odot Y) \odot Z$ is unmarked if and only if

($\otimes$) none of x, y, z is marked, or

($\circledast$) (at least) two of x, y, z are 0-cubes and the last is unmarked.

One can give a similar characterization of when $x \otimes (y \otimes z)$ is unmarked, and it follows that the associativity isomorphism $(\upsilon X \otimes \upsilon Y) \otimes \upsilon Z \cong \upsilon X \otimes (\upsilon Y \otimes \upsilon Z)$ in cSet lifts to an isomorphism $(X \odot Y) \odot Z \cong X \odot (Y \odot Z)$ in cSet^+. The unit isomorphisms can be lifted similarly, and moreover these lifted isomorphisms are suitably natural and coherent. Thus we indeed obtain a monoidal structure on cSet^+ such that υ is strict monoidal. The clauses (2)–(4) are then obvious from the definitions of the tensor products.

It remains to show that this monoidal structure is biclosed. Equivalently, we must show that $\odot$ preserves colimits in each variable separately. So let $F\colon \mathscr{I} \to \mathsf{cSet}^+$ and $X \in \mathsf{cSet}^+$. Since the geometric product is cocontinuous in each variable and υ is cocontinuous and strict monoidal, the canonical comparison map

$$\operatorname{colim}(X \odot F) \to X \odot \operatorname{colim} F$$

is υ-invertible. Moreover one can check using Remark 2.6 that a nondegenerate cube in either side is marked if and only if it is the image of some marked cube under the canonical map from $X \odot Fi$ for some $i \in \mathscr{I}$. It follows that this comparison map is invertible. Dually, $(-) \odot X$ preserves colimits. Since cSet^+ is locally finitely presentable, the existence of the desired biclosed structure now follows. $\square$

Lemma 2.17 *Let $f\colon A \to X$ and $g\colon B \to Y$ be monomorphisms in* cSet^+. *Then $f \mathbin{\hat{\odot}} g$ is again a monomorphism. Moreover*:

(1) *If both f and g are regular, then so is $f \mathbin{\hat{\odot}} g$.*

(2) *If either f or g is entire, then so is $f \mathbin{\hat{\odot}} g$.*

(3) *If both f and g are entire, then $f \mathbin{\hat{\odot}} g$ is invertible.*

(4) *If either f or g is entire, then the square*

$$\begin{tikzcd}
(X \otimes B) \cup (A \otimes Y) \arrow[r] \arrow[d, "f \hat{\otimes} g"'] & (X \circledast B) \cup (A \circledast Y) \arrow[d, "f \hat{\circledast} g"] \\
X \otimes Y \arrow[r, "\mu_{X,Y}"] & X \circledast Y
\end{tikzcd}$$

is a pushout in cSet$^+$, *where the upper horizontal map is induced by μ.*

Proof Since a map in cSet$^+$ is a monomorphism if and only if its underlying map in cSet is a monomorphism, the first (unnumbered) assertion follows from Corollary 1.18, Theorem 2.16(1) and the cocontinuity of υ. We will assume for the sake of simplicity that $f \hat{\odot} g$ is an inclusion.

(1), case ($\otimes$) Let $x \otimes y$ be a nondegenerate cube in $\mathrm{dom}(f \hat{\otimes} g)$. By duality, we may assume that x is in A. If $x \otimes y$ is marked in $X \otimes Y$, then either x is marked in X or y is marked in Y. It follows (by the regularity of f in the former case) that $x \otimes y$ is marked in $A \otimes Y$. This shows that $f \hat{\otimes} g$ is regular.

(1), case ($\circledast$) Let $x \otimes y$ be a marked nondegenerate cube in $X \circledast Y$. Suppose that $x \otimes y$ is in the image of $f \hat{\circledast} g$. The case ($1\otimes$) combined with the commutativity of the square in (4) imply that if $x \otimes y$ is marked in $X \otimes Y$ then it is also marked in $\mathrm{dom}(f \hat{\circledast} g)$. Thus by Remark 2.14, it suffices to consider the case where $x \in X_m$ and $y \in Y_n$ for some $m, n \geq 1$. But in this case $x \otimes y$ is marked in $\mathrm{dom}(f \hat{\circledast} g)$ by the definition of $\circledast$.

(2) Since υ preserves colimits, we have $\upsilon(f \hat{\odot} g) \cong \upsilon f \hat{\odot} \upsilon g$. Thus this assertion follows from the fact that the pushout of an isomorphism along any map is itself an isomorphism.

(3), case ($\otimes$) We know from (2) that $f \hat{\otimes} g$ is entire, so it suffices to show that this map is also regular. Let $x \otimes y$ be a marked nondegenerate cube in $X \otimes Y$. Then either x is marked in X or y is marked in Y. The cube $x \otimes y$ is then marked in $X \otimes B$ in the first subcase and it is marked in $A \otimes Y$ in the second subcase. Thus $f \hat{\otimes} g$ is indeed regular.

(4) By (2), each map in this square is entire. Thus its image under υ is trivially a pushout in cSet. Moreover, for each of the horizontal maps, Remark 2.14 implies that the codomain is obtained from the domain by marking those cubes $x \otimes y$ such that $x \in X_m$ and $y \in Y_n$ with $m, n \geq 1$. Now the assertion follows by Remark 2.6.

(3), case ($\circledast$) This case follows from (3), case ($\otimes$), and (4). □

Proposition 2.18 *For any $m, n \geq 0$, the Leibniz Gray tensor product $(\partial\square^m \hookrightarrow \square^m) \hat{\otimes} (\partial\square^n \hookrightarrow \square^n)$ of boundary inclusions in* cSet$^+$ *is isomorphic to $\partial\square^{m+n} \hookrightarrow \square^{m+n}$.*

Proof This is a straightforward consequence of Proposition 1.17, Theorem 2.16(3) and the cocontinuity of $(-)^\flat$. □

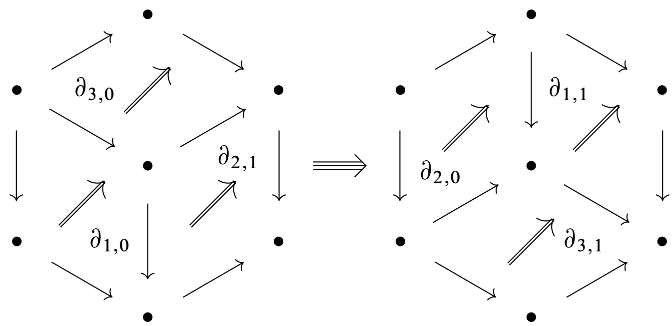
3 Model structure for comical sets

In this section, we construct two families of model structures on the category cSet^+ of marked cubical sets. The former of those has as its fibrant objects (*saturated*) *comical sets*, which we will define, and it is our tentative model for the theory of weak ω-categories. The fibrant objects of the latter are the n-trivial comical sets, and it is our tentative model for the theory of (∞, n)-categories.

A comical set is to be thought of as a kind of weak ω-category, and an n-cube therein represents an n-dimensional morphism. The $(n-1)$-source of such an n-cube is the "composite" of the faces $\partial_{k,\varepsilon}$ with $k+\varepsilon$ odd, and similarly the $(n-1)$-target is given by the even faces. (This idea of parity-based decomposition into source and target goes back to [Street 1987], where Street considers the free ω-categories on simplices. In the case of cubes, see eg [Aitchison 1986; Street 1991; Steiner 1993; Al-Agl et al. 2002]). For instance, a 2-cube can be seen as a morphism of the form

$$\begin{tikzcd} \bullet \arrow[r, "\partial_{2,0}"] \arrow[d, "\partial_{1,0}"'] \arrow[dr, phantom, "\Nearrow"] & \bullet \arrow[d, "\partial_{1,1}"] \\ \bullet \arrow[r, "\partial_{2,1}"'] & \bullet \end{tikzcd}$$

and a 3-cube represents a morphism between the composites



Marked n-cubes are to be thought of as being (weakly) invertible, although not every invertible cube is marked unless the comical set is saturated.

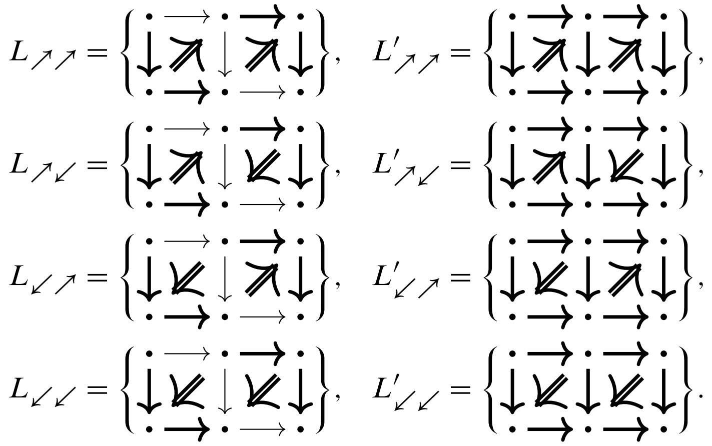
Before defining comical sets, we will need a few auxiliary definitions.

For $n \geq 1$, $1 \leq k \leq n$ and $\varepsilon \in \{0, 1\}$, we denote by $\Box^n_{k,\varepsilon}$ the n-cube with the following marking: a nondegenerate cube $\partial_{k_1,\varepsilon_1} \cdots \partial_{k_t,\varepsilon_t}$, written in the form specified by Theorem 1.6, is marked whenever this string does not contain $\partial_{k-1,\varepsilon}$, $\partial_{k,\varepsilon}$, $\partial_{k,1-\varepsilon}$ or $\partial_{k+1,\varepsilon}$. (This is exactly the marking described in [Steiner 2006, Example 2.9].) We call this the (k, ε)-*comical n-cube*. We denote by $\sqcap^n_{k,\varepsilon} \subset \Box^n_{k,\varepsilon}$ the (k, ε)-open box of dimension n (ie the cubical subset missing the nondegenerate n-cube and the $(k, \varepsilon)^{\text{th}}$ $(n-1)$-face) endowed with the marking making it a regular subset of $\Box^n_{k,\varepsilon}$. We call $\sqcap^n_{k,\varepsilon}$ the *comical (k, ε)-open box of dimension n*. We call the inclusion $\sqcap^n_{k,\varepsilon} \to \Box^n_{k,\varepsilon}$ the (k, ε)-*comical open box inclusion of dimension n*.

The *elementary (k,ε)-comical marking extension of dimension n*, denoted by ${\square^n_{k,\varepsilon}}' \to {\square^n_{k,\varepsilon}}''$, is the Leibniz product of the unit $\mathrm{id} \to \tau_{n-2}$ and the comical box inclusion $\sqcap^n_{k,\varepsilon} \to \square^n_{k,\varepsilon}$, ie the dashed map in

$$\begin{tikzcd}
\sqcap^n_{k,\varepsilon} \arrow[r] \arrow[d] \arrow[dr, phantom, "\text{p.o.}"] & \tau_{n-2}\sqcap^n_{k,\varepsilon} \arrow[d] \arrow[ddr, bend left] & \\
\square^n_{k,\varepsilon} \arrow[r] \arrow[drr, bend right] & \cdot \arrow[dr, dashed] & \\
& & \tau_{n-2}\square^n_{k,\varepsilon}
\end{tikzcd}$$

For each $x, y \in \{\nearrow, \swarrow\}$, we define the *basic Rezk map* $L_{xy} \hookrightarrow L'_{xy}$ as the entire inclusion depicted below:



Here thick arrows indicate marked cubes. More precisely, $L_{\nearrow\nearrow}$ is the pushout of the span

$$X \xleftarrow{\partial_{1,1}} \square^1 \xrightarrow{\partial_{1,0}} Y,$$

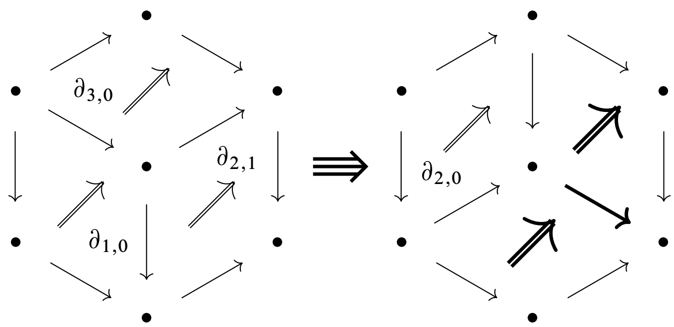
where X is obtained from $\widetilde{\square}^2$ by marking $\partial_{1,0}$ and $\partial_{2,1}$, and Y is obtained from $\widetilde{\square}^2$ by marking $\partial_{1,1}$ and $\partial_{2,0}$. The codomain $L'_{\nearrow\nearrow}$ is the 0-trivialization $\tau_0(L_{\nearrow\nearrow})$. The marked cubical sets L_{xy}, L'_{xy} are defined similarly for other choices of $x, y \in \{\nearrow, \swarrow\}$. By a *Rezk map* we mean any map of the form

$$(\partial\square^m \hookrightarrow \square^m)\hat{\otimes}(L_{xy} \hookrightarrow L'_{xy})\hat{\otimes}(\partial\square^n \hookrightarrow \square^n).$$

Definition 3.1 (1) A *comical set* is a marked cubical set with the right lifting property with respect to the comical open box inclusions and the elementary comical marking extensions.

(2) A *saturated comical set* is a marked cubical set with the right lifting property with respect to the comical open box inclusions, the elementary comical marking extensions and the Rezk maps.

Remark 3.2 We briefly explain how the definition of (saturated) comical set should be interpreted. In the comical n-cube $\square^n_{k,\varepsilon}$, any subcube not contained in $\partial_{k-1,\varepsilon}$, $\partial_{k,\varepsilon}$, $\partial_{k,1-\varepsilon}$ or $\partial_{k+1,\varepsilon}$ is marked. In particular, the unique nondegenerate n-cube is marked, so it can be thought of as an equivalence between the composite of its odd faces and the composite of even faces. In other words, the comical n-cube $\square^n_{k,\varepsilon}$ exhibits $\partial_{k,\varepsilon}$ as a composite of $\partial_{k-1,\varepsilon}$, $\partial_{k,1-\varepsilon}$ and $\partial_{k+1,\varepsilon}$. For example, $\square^3_{2,0}$ looks like



One can thus interpret the right lifting property with respect to the comical box inclusions and the comical marking extensions respectively as the existence of composites and the closure of marked cubes under composition. In Section 4, we show how these conditions additionally encode such expected properties of composition as the unit and associative laws, at least for 1-cubes.

There are two standard model structures on marked cubical sets:

Theorem 3.3 (model structure for comical sets) *The category cSet^+ carries two model structures*:

1. *The **comical model structure**, characterized by the following properties*:
   - *The cofibrations are the monomorphisms.*
   - *The set of*
     - *comical open box inclusions*, *and*
     - *elementary comical marking extensions*

     forms a pseudogenerating set of trivial cofibrations.
2. *The **saturated comical model structure**, characterized by the following properties*:
   - *The cofibrations are the monomorphisms.*
   - *The set of*
     - *comical open box inclusions*,
     - *elementary comical marking extensions*, *and*
     - *Rezk maps*

     forms a pseudogenerating set of trivial cofibrations.

   Both of these model structures are combinatorial, *left proper*, *monoidal with respect to either of the Gray tensor products*, *and have all objects cofibrant.*

The proof of this theorem is an application of the Cisinski–Olschok theory and verification of the closure of anodyne maps under pushout-product. The latter part is contained in Lemma 3.5 below.

Definition 3.4 We say that a map of marked cubical sets $X \to Y$ is

(1) a *comical marking extension* if it is in the cellular closure of the elementary comical marking extensions, and

(2) *comical* if it is in the cellular closure of the comical open box inclusions and the elementary comical marking extensions.

Lemma 3.5 *For any $1 \le k \le m$, $\varepsilon \in \{0,1\}$ and $n \ge 0$ (or $n \ge 1$ for g), the Leibniz Gray tensor products*

$$\begin{aligned} f &= (\sqcap^m_{k,\varepsilon} \hookrightarrow \square^m_{k,\varepsilon}) \mathbin{\hat{\odot}} (\partial\square^n \hookrightarrow \square^n), \\ g &= (\sqcap^m_{k,\varepsilon} \hookrightarrow \square^m_{k,\varepsilon}) \mathbin{\hat{\odot}} (\square^n \hookrightarrow \widetilde{\square}^n), \\ h &= (\square^m_{k,\varepsilon}{}' \hookrightarrow \square^m_{k,\varepsilon}{}'') \mathbin{\hat{\odot}} (\partial\square^n \hookrightarrow \square^n) \end{aligned}$$

are all comical.

Proof Since the case $n = 0$ is trivial, we will assume otherwise.

Consider a face of $\square^{m+n}$ whose normal form $\partial_{k_1,\varepsilon_1} \cdots \partial_{k_c,\varepsilon_c}$ does not involve $\partial_{k-1,\varepsilon}$, $\partial_{k,0}$, $\partial_{k,1}$ or $\partial_{k+1,\varepsilon}$. Then clearly any terminal segment of this normal form does not involve any of these four ∂'s. This observation implies that the second isomorphism of Proposition 1.17 may be lifted to the following commutative square:

$$\begin{tikzcd} \sqcap^{m+n}_{k,\varepsilon} \arrow[r] \arrow[d, hook] & (\square^m_{k,\varepsilon} \odot \partial\square^n) \cup (\sqcap^m_{k,\varepsilon} \odot \square^n) \arrow[d, hook, "f"] \\ \square^{m+n}_{k,\varepsilon} \arrow[r] & \square^m_{k,\varepsilon} \odot \square^n \end{tikzcd}$$

Observe that this is a pushout square on the underlying cubical set level.

In the case $\odot = \otimes$, it is in fact a pushout square in cSet^+. To see this, it suffices to check that the marking on $\square^m_{k,\varepsilon} \otimes \square^n$ agrees with that described in Remark 2.6. This is indeed the case since f is regular by Lemma 2.17(1) and the only marked nondegenerate cube in $\mathrm{cod}(f) \setminus \mathrm{dom}(f)$ is the $(m+n)$-cube, which is the image of a marked cube under the lower horizontal map. Thus f is indeed comical.

Now we consider the case $\odot = \circledast$. If $m = 1$ then we can simply repeat the above argument since we have natural isomorphisms

$$\sqcap^1_{1,\varepsilon} \circledast (-) \cong \sqcap^1_{1,\varepsilon} \otimes (-) \quad \text{and} \quad \square^1_{1,\varepsilon} \circledast (-) \cong \square^1_{1,\varepsilon} \otimes (-)$$

by Remark 2.14. So assume $m \ge 2$. Then there is an extra marked nondegenerate cube in $\mathrm{cod}(f) \setminus \mathrm{dom}(f)$, namely $\partial_{k,\varepsilon}$. But then it is straightforward to check that $\partial_{k-1,\varepsilon}$, $\partial_{k,1-\varepsilon}$ and $\partial_{k+1,\varepsilon}$ are also marked (whenever they exist). So f can be written as a pushout of the open box inclusion $\sqcap^{m+n}_{k,\varepsilon} \hookrightarrow \square^{m+n}_{k,\varepsilon}$ followed by a pushout of the comical marking extension $\square^{m+n}_{k,\varepsilon}{}' \hookrightarrow \square^{m+n}_{k,\varepsilon}{}''$.

The map g is entire by Lemma 2.17(2). Similarly to the above argument, one can deduce the existence of the following commutative square of entire monomorphisms:

$$\begin{tikzcd}
\square^{m+n'}_{k,\varepsilon} \arrow[r] \arrow[d, hook] & (\square^m_{k,\varepsilon} \odot \square^n) \cup (\sqcap^m_{k,\varepsilon} \odot \widetilde{\square}^n) \arrow[d, hook, "g"] \\
\square^{m+n''}_{k,\varepsilon} \arrow[r] & \square^m_{k,\varepsilon} \odot \widetilde{\square}^n
\end{tikzcd}$$

One can check that, in the case where $\odot = \circledast$ and $m \geq 2$, the map g is in fact invertible. Otherwise, the only cube in $\operatorname{cod}(g)$ that is not marked in $\operatorname{dom}(g)$ is $\partial_{k,\varepsilon}$ and it is the image of a marked cube under the lower horizontal map. This shows that the above square is a pushout. Hence g is a comical marking extension.

Similarly, one can check that the following square is a pushout:

$$\begin{tikzcd}
\square^{m+n'}_{k,\varepsilon} \arrow[r] \arrow[d, hook] & (\square^m_{k,\varepsilon}{}'' \odot \partial\square^n) \cup (\square^m_{k,\varepsilon}{}' \odot \square^n) \arrow[d, hook, "h"] \\
\square^{m+n''}_{k,\varepsilon} \arrow[r] & \square^m_{k,\varepsilon}{}'' \odot \square^n
\end{tikzcd}$$

Therefore h is a comical marking extension. □

Proof of Theorem 3.3 We apply the Cisinski–Olschok theory, ie Theorem 1.4 with $\mathcal{K} = \mathsf{cSet}^+$ and I the set of boundary inclusions and markers. The set S consists of the comical open box inclusions and the comical marking extensions in (1), and it additionally contains all Rezk maps in (2). For our cylinder functor C, we can use either $\widetilde{\square}^1 \otimes (-)$ or $\widetilde{\square}^1 \circledast (-)$ as they are equal by Remark 2.14. This produces a model structure on cSet^+ in which the cofibrations are the monomorphisms and $\Lambda(\mathsf{cSet}^+, I, C, S)$ is a pseudogenerating set of trivial cofibrations.

It remains to prove that the set S is in fact pseudogenerating, and moreover the model structure is monoidal with respect to either of the Gray tensor products. By duality and Proposition 1.5 it suffices to show that

- $f \mathbin{\hat{\odot}} g$ is in the cellular closure of I whenever $f, g \in I$, and
- $f \mathbin{\hat{\odot}} g$ is in the cellular closure of S whenever $f \in S$ and $g \in I$.

The first clause essentially follows from the unmarked version (Corollary 1.18). We now treat the second clause.

There are three kinds of maps in S, namely

(A) comical box inclusions,

(B) elementary comical marking extensions, and

(C) Rezk maps,

and two kinds of maps in I, namely

(a) boundary inclusions, and

(b) markers.

The case (Ca$\otimes$) is a straightforward consequence of the associativity of $\hat{\otimes}$ and Proposition 2.18. The case (Ca$\circledast$) then follows by Lemma 2.17(4). In the cases (Bb) and (Cb), the map $f \mathbin{\hat{\odot}} g$ is invertible by Lemma 2.17(3). The remaining cases are treated in Lemma 3.5. $\square$

There are also n-trivial versions of these model structures.

Theorem 3.6 (model structure for n-trivial comical sets) *The category* cSet^+ *carries two families of model structures*:

(1) *The* ***n-trivial comical model structure*** *characterized by the following properties*:
- *The cofibrations are the monomorphisms.*
- *The set of*
  - *comical open box inclusions*,
  - *elementary comical marking extensions of dimension* $\le n+1$, *and*
  - *markers of dimension* $> n$

  forms a pseudogenerating set of trivial cofibrations.

(2) *The* ***saturated n-trivial comical model structure*** *characterized by the following properties*:
- *The cofibrations are the monomorphisms.*
- *The set of*
  - *comical open box inclusions*,
  - *elementary comical marking extensions of dimension* $\le n+1$,
  - *markers of dimension* $> n$, *and*
  - *Rezk maps*

  form a pseudogenerating set of trivial cofibrations.

Proof Analogous to the proof of Theorem 3.3. Note that the Leibniz Gray tensor product of the m-marker with any monomorphism is in the cellular closure of the m'-markers with $m' \ge m$. $\square$

Proposition 3.7 *The functor* $\tau_n \colon \mathsf{cSet}^+ \to \mathsf{cSet}^+$ *is a left Quillen functor from the n-trivial comical model structure (resp. saturated n-trivial comical model structure) to the comical model structure (resp. saturated comical model structure).*

Lemma 3.8 *For any* $n \ge 0$, $1 \le k \le n+2$ *and* $\varepsilon \in \{0,1\}$, *the map* $\tau_n \sqcap^{n+2}_{k,\varepsilon} \hookrightarrow \tau_n \square^{n+2}_{k,\varepsilon} = \square^{n+2\,\prime\prime}_{k,\varepsilon}$ *is comical.*

Proof Recall the defining pushout square of the comical marking extension:

$$\begin{tikzcd} \sqcap^{n+2}_{k,\varepsilon} \arrow[r] \arrow[d] \arrow[dr, phantom, "\text{p.o.}"] & \tau_n \sqcap^{n+2}_{k,\varepsilon} \arrow[d] \arrow[ddr, bend left] & \\ \square^{n+2}_{k,\varepsilon} \arrow[r] \arrow[drr, bend right] & \cdot \arrow[dr, dashed] & \\ & & \tau_n \square^{n+2}_{k,\varepsilon} \end{tikzcd}$$

This diagram exhibits the desired result. □

Proof of Proposition 3.7 That τ_n preserves cofibrations is obvious. Thus it suffices to check (by [Joyal and Tierney 2007, Lemma 7.14]) that τ_n sends each member $f : X \to Y$ of the pseudogenerating set to a trivial cofibration.

Unless f is the open box inclusion $\sqcap^m_{k,\varepsilon} \hookrightarrow \square^m_{k,\varepsilon}$ with $m \geq n+2$, any marked cube in Y of dimension $> n$ admits a (not necessarily marked) preimage in X. In these cases, the naturality square for the unit

$$\begin{tikzcd} X \arrow[r] \arrow[d, "f"'] & \tau_n X \arrow[d, "\tau_n f"] \\ Y \arrow[r] & \tau_n Y \end{tikzcd}$$

is a pushout in cSet^+ by Remark 2.6, so $\tau_n f$ is a trivial cofibration.

So assume that f is the open box inclusion $\sqcap^m_{k,\varepsilon} \hookrightarrow \square^m_{k,\varepsilon}$ with $m \geq n+2$. Then we have the following commutative square:

$$\begin{tikzcd} \tau_{m-2} \sqcap^m_{k,\varepsilon} \arrow[r] \arrow[d, "\tau_{m-2} f"'] & \tau_n \sqcap^m_{k,\varepsilon} \arrow[d, "\tau_n f = \tau_n(\tau_{m-2} f)"] \\ \tau_{m-2} \square^m_{k,\varepsilon} \arrow[r] & \tau_n \square^m_{k,\varepsilon} \end{tikzcd}$$

Observe that the left vertical map is comical by Lemma 3.8 (with suitable substitution), and moreover it satisfies the condition on f described in the previous paragraph. Thus this square is a pushout, exhibiting $\tau_n f$ as a comical map. This completes the proof. □

As we mentioned earlier, our definition of comical box inclusion uses the marking described in [Steiner 2006], where Steiner characterizes the nerves of strict ω-categories. Phrased in the language of comical sets, his characterization implies the following result:

Theorem 3.9 (cf [Steiner 2006, Theorem 3.16]) *The cubical nerve of a globular ω-category, or equivalently the underlying cubical set of a cubical ω-category with connections, is a comical set.* □

This is analogous to the statement that the simplicial nerve of a strict ω-category is a complicial set [Verity 2008a].

We conclude this section with the following observation, which will be useful in Section 7.

Proposition 3.10 *For any $n \geq 2$, $1 \leq k \leq n$ and $\varepsilon \in \{0,1\}$, the comical box inclusion $\sqcap^n_{k,\varepsilon} \hookrightarrow \square^n_{k,\varepsilon}$ may be written as*

$$\begin{cases} (\sqcap^2_{1,\varepsilon} \hookrightarrow \square^2_{1,\varepsilon}) \hat{\otimes} (\partial\square^{n-2} \hookrightarrow \square^{n-2}) & \text{if } k = 1, \\ (\partial\square^{k-2} \hookrightarrow \square^{k-2}) \hat{\otimes} (\sqcap^3_{2,\varepsilon} \hookrightarrow \square^3_{2,\varepsilon}) \hat{\otimes} (\partial\square^{n-k-1} \hookrightarrow \square^{n-k-1}) & \text{if } 1 < k < n, \\ (\partial\square^{n-2} \hookrightarrow \square^{n-2}) \hat{\otimes} (\sqcap^2_{2,\varepsilon} \hookrightarrow \square^2_{2,\varepsilon}) & \text{if } k = n. \end{cases}$$

Proof It is easy to check that the underlying cubical maps match, and also the markings on the codomains match. Now observe that these Leibniz Gray tensor products are regular by Lemma 2.17(1). □

4 Homotopy 1-categories of comical sets

Suppose we are given two 1-cubes $f, g\colon x \to y$ in a comical set X. Then a marked 2-cube satisfying any one of the following boundary conditions may be reasonably regarded as a homotopy $f \sim g$:

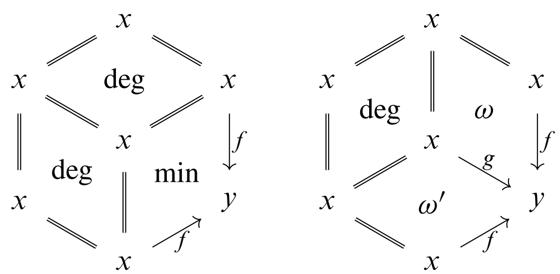
$$\begin{tikzcd} x \arrow[r, "f"] \arrow[d, "g"'] \arrow[dr, phantom, "\phi"] & y \arrow[d, equal] \\ y \arrow[r, equal] & y \end{tikzcd} \quad \begin{tikzcd} x \arrow[r, "g"] \arrow[d, "f"'] \arrow[dr, phantom, "\phi'"] & y \arrow[d, equal] \\ y \arrow[r, equal] & y \end{tikzcd} \quad \begin{tikzcd} x \arrow[r, equal] \arrow[d, "f"'] \arrow[dr, phantom, "\chi"] & x \arrow[d, "g"] \\ y \arrow[r, equal] & y \end{tikzcd} \quad \begin{tikzcd} x \arrow[r, equal] \arrow[d, "g"'] \arrow[dr, phantom, "\chi'"] & x \arrow[d, "f"] \\ y \arrow[r, equal] & y \end{tikzcd}$$

$$\begin{tikzcd} x \arrow[r, "f"] \arrow[d, equal] \arrow[dr, phantom, "\psi"] & y \arrow[d, equal] \\ x \arrow[r, "g"'] & y \end{tikzcd} \quad \begin{tikzcd} x \arrow[r, "g"] \arrow[d, equal] \arrow[dr, phantom, "\psi'"] & y \arrow[d, equal] \\ x \arrow[r, "f"'] & y \end{tikzcd} \quad \begin{tikzcd} x \arrow[r, equal] \arrow[d, equal] \arrow[dr, phantom, "\omega"] & x \arrow[d, "f"] \\ x \arrow[r, "g"'] & y \end{tikzcd} \quad \begin{tikzcd} x \arrow[r, equal] \arrow[d, equal] \arrow[dr, phantom, "\omega'"] & x \arrow[d, "g"] \\ x \arrow[r, "f"'] & y \end{tikzcd}$$

Here equalities indicate degenerate (and hence marked) 1-cubes.

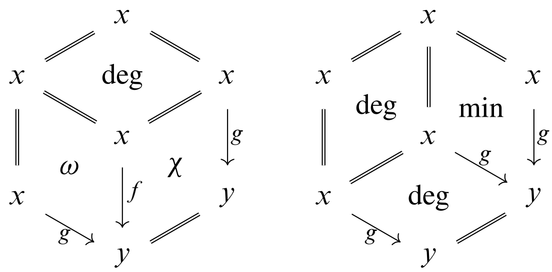
Proposition 4.1 *If any one of the above boundary conditions admits a marked solution in the comical set X then so do the others.*

Proof Consider the following picture:



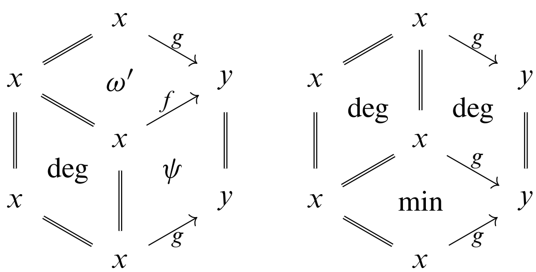
(Here the faces labeled "deg" are fully degenerate on the 0-cube x, and the face labeled "min" is the min-connection on f.) If we have a marked 2-cube ω satisfying the boundary condition specified above, then this picture may be interpreted as a map $\tau_1 \sqcap^3_{3,1} \to X$, which may be extended to $\square^3_{3,1}{}''$ by Lemma 3.8, yielding a marked 2-cube ω'. Conversely, if we are given ω' then this picture specifies a map $\tau_1 \sqcap^3_{1,1} \to X$ and extending it to $\square^3_{1,1}{}''$ yields a marked 2-cube ω.

Similarly, the following picture shows that χ exists if and only if ω does:

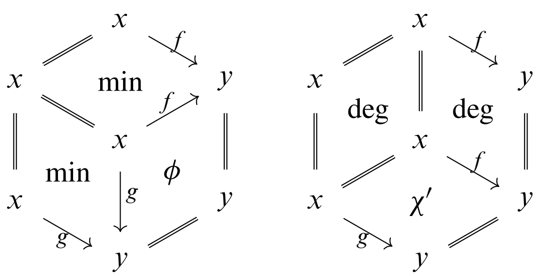


Dually, χ' exists if and only if ω' does.

The following picture shows that ψ exists if and only if ω' does (dually, ψ' exists if and only if ω does):



Finally the following picture shows that ϕ exists if and only if χ' does (dually, ϕ' exists if and only if χ does):



This completes the proof. $\square$

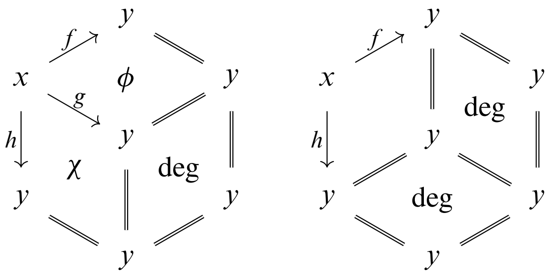
Definition 4.2 We say two 1-cubes f, g in a comical set X are *homotopic*, and write $f \sim g$ if any one of the above marked 2-cubes exists in X.

Proposition 4.3 *For any pair of 0-cubes x, y in a comical set X, the homotopy relation is an equivalence relation on the set of all 1-cubes $x \to y$.*

Proof The reflexivity of $\sim$ is obvious, and its symmetry follows from Proposition 4.1. For transitivity, suppose we are given two homotopies:

$$\begin{tikzcd} x \arrow[r, "f"] \arrow[d, "g"'] \arrow[dr, phantom, "\phi"] & y \arrow[d, equal] \\ y \arrow[r, equal] & y \end{tikzcd} \qquad \begin{tikzcd} x \arrow[r, "g"] \arrow[d, "h"'] \arrow[dr, phantom, "\chi"] & y \arrow[d, equal] \\ y \arrow[r, equal] & y \end{tikzcd}$$

Then the following picture specifies a map $\tau_1 \sqcap^3_{2,0} \to X$:



This map extends to ${\square^3_{2,0}}''$, which in particular yields a homotopy $f \sim h$. $\square$

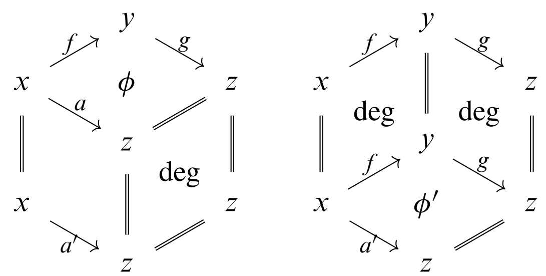
Now consider a "composable" pair of 1-cubes $f\colon x \to y$ and $g\colon y \to z$ in a comical set X. We may "compose" f and g by filling any one of the open boxes $\sqcap^2_{1,0}$, $\sqcap^2_{1,1}$, $\sqcap^2_{2,0}$, $\sqcap^2_{2,1}$ as follows:

$$\begin{tikzcd} x \arrow[r, "f"] \arrow[d, "a"'] \arrow[dr, phantom, "\phi"] & y \arrow[d, "g"] \\ z \arrow[r, equal] & z \end{tikzcd} \quad \begin{tikzcd} x \arrow[r, equal] \arrow[d, "f"'] \arrow[dr, phantom, "\chi"] & x \arrow[d, "b"] \\ y \arrow[r, "g"'] & z \end{tikzcd} \quad \begin{tikzcd} x \arrow[r, "c"] \arrow[d, "f"'] \arrow[dr, phantom, "\psi"] & z \arrow[d, equal] \\ y \arrow[r, "g"'] & z \end{tikzcd} \quad \begin{tikzcd} x \arrow[r, "f"] \arrow[d, equal] \arrow[dr, phantom, "\omega"] & y \arrow[d, "g"] \\ x \arrow[r, "d"'] & z \end{tikzcd}$$

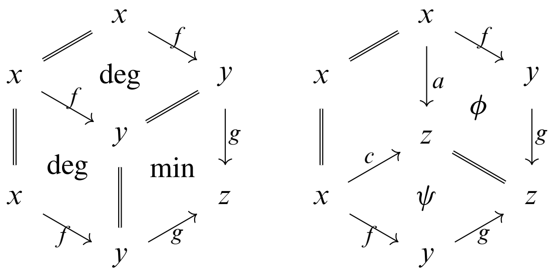
We will temporarily call such a a $(1,0)$-*composite* of f and g, and similarly call b, c and d $(1,1)$-, $(2,0)$- and $(2,1)$-composites of f and g respectively.

Proposition 4.4 *Any two composites of f and g are homotopic to each other.*

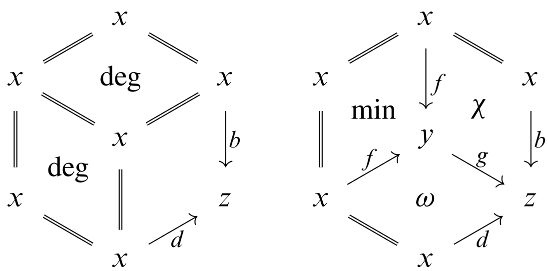
Proof First, consider two $(1,0)$-composites a and a', witnessed by 2-cubes ϕ and ϕ', respectively. Then the following picture specifies a map $\tau_1 \sqcap^3_{1,0} \to X$:



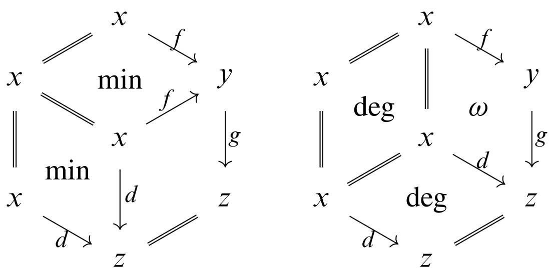
This map extends to ${\square^3_{2,0}}''$, which yields a homotopy $a \sim a'$. Similarly, for any given a, b, c, d as above, a homotopy $a \sim c$ can be obtained by filling the following open box:



and a homotopy $b \sim d$ can be obtained using



Finally, we can turn the $(2,1)$-composite d into a $(1,0)$-composite using the following open box:



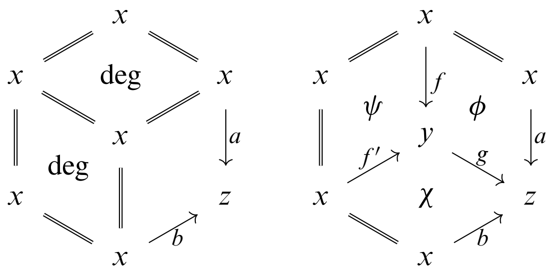
This completes the proof. $\square$

Proposition 4.5 *Let* $f, f' \colon x \to y$ *and* $g, g' \colon y \to z$ *be* 1*-cubes in a comical set* X *such that* $f \sim f'$ *and* $g \sim g'$*. Then any composite of* f *and* g *is homotopic to any composite of* f' *and* g'*.*

Proof Choose witnesses of the following forms for compositions and a homotopy $f \sim f'$:

$$
\begin{tikzcd}
x \arrow[r, equal] \arrow[d, "f"'] \arrow[dr, phantom, "\phi"] & x \arrow[d, "a"] \\
y \arrow[r, "g"'] & z
\end{tikzcd}
\qquad
\begin{tikzcd}
x \arrow[r, "f'"] \arrow[d, equal] \arrow[dr, phantom, "\chi"] & y \arrow[d, "g"] \\
x \arrow[r, "b"'] & z
\end{tikzcd}
\qquad
\begin{tikzcd}
x \arrow[r, equal] \arrow[d, equal] \arrow[dr, phantom, "\psi"] & x \arrow[d, "f"] \\
x \arrow[r, "f'"'] & y
\end{tikzcd}
$$

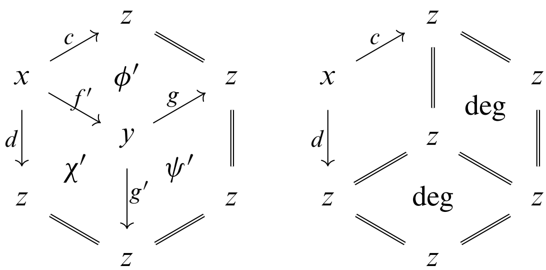
Then extending the following map $\tau_1 \sqcap^3_{2,1} \to X$ to ${\square^3_{2,1}}''$ yields a homotopy $a \sim b$:



Similarly, we may combine marked 2-cubes of the forms

$$\begin{tikzcd} x \arrow[r, "c"] \arrow[d, "f'"'] \arrow[dr, phantom, "\phi'"] & z \arrow[d, equal] \\ y \arrow[r, "g"'] & z \end{tikzcd} \qquad \begin{tikzcd} x \arrow[r, "f'"] \arrow[d, "d"'] \arrow[dr, phantom, "\chi'"] & y \arrow[d, "g'"] \\ x \arrow[r, equal] & z \end{tikzcd} \qquad \begin{tikzcd} y \arrow[r, "g"] \arrow[d, "g'"'] \arrow[dr, phantom, "\psi'"] & z \arrow[d, equal] \\ z \arrow[r, equal] & z \end{tikzcd}$$

into a map $\tau_1 \sqcap^3_{2,0} \to X$ as follows:



Extending this map to $\square^3_{2,0}{}''$ yields a homotopy $c \sim d$. The desired result now follows by Propositions 4.3 and 4.4. $\square$

Definition 4.6 We define the *homotopy* 1*-category* $\mathrm{ho}_1\, X$ of a comical set X to be the category of 0-cubes and homotopy classes of 1-cubes in X.

Proposition 4.7 *For a comical set X, $\mathrm{ho}_1\, X$ is indeed a* 1*-category.*

Proof Proposition 4.5 implies that we have a well-defined composition operation on $\mathrm{ho}_1\, X$. For any 0-cube x in X, we claim that the homotopy class containing $x\sigma_1$ is the identity at x. Indeed for any $f : x \to y$, the degenerate 2-cube

$$\begin{tikzcd} x \arrow[r, equal, "x\sigma_1"] \arrow[d, "f"'] \arrow[dr, phantom, "f\sigma_1"] & x \arrow[d, "f"] \\ y \arrow[r, equal, "y\sigma_1"'] & y \end{tikzcd}$$

exhibits f as a $(1,0)$-composite of $x\sigma_1$ and f, and also as a $(1,1)$-composite of f and $y\sigma_1$.

For associativity, suppose we are given 1-cubes $f: x \to y$, $g: y \to z$ and $h: z \to w$ in X. Compose these 1-cubes as follows:

$$
\begin{tikzcd}
x \arrow[r, "f"] \arrow[d, equal] \arrow[dr, phantom, "\phi"] & y \arrow[d, "g"] \\
x \arrow[r, "a"'] & z
\end{tikzcd}
\qquad
\begin{tikzcd}
x \arrow[r, "a"] \arrow[d, equal] \arrow[dr, phantom, "\chi"] & z \arrow[d, "h"] \\
x \arrow[r, "b"'] & w
\end{tikzcd}
\qquad
\begin{tikzcd}
y \arrow[r, "g"] \arrow[d, equal] \arrow[dr, phantom, "\psi"] & z \arrow[d, "h"] \\
y \arrow[r, "c"'] & w
\end{tikzcd}
$$

Then we may combine them into a map $\tau_1 \sqcap^3_{3,1} \to X$:

$$
\begin{tikzcd}
 & y \arrow[dr, "g"] & \\
x \arrow[ur, "f"] \arrow[dr, equal] \arrow[dd, equal] & \phi & z \arrow[dd, "h"] \\
\mathrm{deg} & x \arrow[ur, "a"] \arrow[dd, equal] & \chi \\
x \arrow[dr, equal] & & w \\
 & x \arrow[ur, "b"'] &
\end{tikzcd}
\qquad
\begin{tikzcd}
 & y \arrow[dr, "g"] \arrow[dd, equal] & \\
x \arrow[ur, "f"] \arrow[dd, equal] & \mathrm{deg} \quad \psi & z \arrow[dd, "h"] \\
 & y \arrow[dr, "c"] & \\
x \arrow[ur, "f"] \arrow[dr, equal] & & w \\
 & x \arrow[ur, "b"'] &
\end{tikzcd}
$$

Extending this map to $\square^{3}_{3,1}{}''$ then yields a marked 2-cube that witnesses the desired associativity. $\square$

The following proposition is straightforward to verify.

Proposition 4.8 *The assignment $X \mapsto \mathrm{ho}_1\, X$ extends to a functor from the category of comical sets to* Cat. *Moreover there is a natural isomorphism* $\mathrm{ho}_1(X^{\mathrm{op}}) \cong (\mathrm{ho}_1\, X)^{\mathrm{op}}$. $\square$

5 Triangulation

In this section, we upgrade the triangulation adjunction described in Section 1.2 to a marked version. We start by recalling the basic combinatorics of simplicial cubes, which can be found in [Verity 2007, Section 5]. (Note however that our indexing is reversed from Verity's.)

Given an r-simplex ϕ in the simplicial set $(\Delta^1)^n$, we can define a function

$$\{1, \dots, n\} \to \{1, \dots, r, \pm\infty\}$$

by declaring

$$i \mapsto \begin{cases} +\infty & \text{if } \pi_i \circ \phi(r) = 0, \\ p & \text{if } \pi_i \circ \phi(p-1) = 0 \text{ and } \pi_i \circ \phi(p) = 1, \\ -\infty & \text{if } \pi_i \circ \phi(0) = 1. \end{cases}$$

If we regard ϕ as an r-step walk on the n-cube with the p^{th} step connecting $\phi(p-1)$ to $\phi(p)$, the above function takes i to the unique p such that the p^{th} step moves in the i^{th} direction; it takes the value $+\infty$ if we never move in the i^{th} direction, and the value $-\infty$ if we have already moved in that direction before we start.

Conversely, any function $\{1, \dots, n\} \to \{1, \dots, r, \pm\infty\}$ determines a unique r-simplex in $(\Delta^1)^n$, so we will identify the r-simplices and these functions.

Remark 5.1 In what follows, we sometimes write such expressions as $p \pm k$ for $p \in \{1, \dots, r, \pm\infty\}$ and finite k. These expressions are to be interpreted as p when $p \in \{\pm\infty\}$. We will never consider expressions involving more than one $\pm\infty$.

Definition 5.2 We will write $\iota = \iota_n \colon \{1, \dots, n\} \to \{1, \dots, n, \pm\infty\}$ for the inclusion regarded as an n-simplex in $(\Delta^1)^n$.

We will think of any set of the form $\{1, \dots, r, \pm\infty\}$ as a linearly ordered set

$$-\infty < 1 < \cdots < r < +\infty.$$

Note however that simplices $\phi \colon \{1, \dots, n\} \to \{1, \dots, r, \pm\infty\}$ are not necessarily order-preserving.

Proposition 5.3 *Under this identification, a simplicial operator $\alpha \colon [q] \to [r]$ sends an r-simplex ϕ to the q-simplex $\phi\alpha$ given by*

$$(\phi\alpha)(i) = \begin{cases} +\infty & \text{if } \phi(i) > \alpha(q), \\ p & \text{if } \alpha(p-1) < \phi(i) \le \alpha(p), \\ -\infty & \text{if } \phi(i) \le \alpha(0). \end{cases}$$

Example 5.4 Again, let us think of an r-simplex ϕ as an r-step walk. Then the last face $\phi\partial_r$ of ϕ moves in the i^{th} direction at exactly the same step as ϕ does except that it does not have an r^{th} step. This agrees with the following formula obtained using Proposition 5.3:

$$(\phi\partial_r)(i) = \begin{cases} +\infty & \text{if } \phi(i) = r, \\ \phi(i) & \text{otherwise.} \end{cases}$$

On the other hand, taking the 0^{th} face decreases the index of each step by 1, so we have

$$(\phi\partial_0)(i) = \begin{cases} -\infty & \text{if } \phi(i) = 1, \\ \phi(i) - 1 & \text{otherwise.} \end{cases}$$

For $0 < k < n$, taking the k^{th} step merges the k^{th} and the $(k+1)^{\text{st}}$ steps, so it does not affect the endpoints of the whole walk. When regarded as a function, this means that $\phi\partial_k$ takes the values $\pm\infty$ on exactly the same inputs as ϕ does. However, some of the indices are shifted:

$$(\phi\partial_k)(i) = \begin{cases} \phi(i) - 1 & \text{if } k < \phi(i) \le r, \\ \phi(i) & \text{otherwise.} \end{cases}$$

For any $0 \le m \le r$, taking the last face $(r-m)$ times yields $\amalg_1^{m,r-m}$, so we have

$$(\phi \amalg_1^{m,r-m})(i) = \begin{cases} +\infty & \text{if } \phi(i) > m, \\ \phi(i) & \text{if } \phi(i) \le m. \end{cases}$$

Similarly, since taking the 0^{th} face m times yields $\amalg_2^{m,r-m}$, we have

$$(\phi \amalg_2^{m,r-m})(i) = \begin{cases} \phi(i) - m & \text{if } \phi(i) > m, \\ -\infty & \text{if } \phi(i) \le m. \end{cases}$$

It is easy to verify the following proposition using Proposition 5.3.

Proposition 5.5 *An r-simplex ϕ in $(\Delta^1)^n$ is nondegenerate if and only if $\phi^{-1}(p) \neq \varnothing$ for each $1 \le p \le r$. More precisely, ϕ is degenerate at $p-1$ if and only if $\phi^{-1}(p) = \varnothing$.*

Now we upgrade the *codomain* of the triangulation functor to a marked version. More precisely, we first consider the functor $\square \to \mathsf{PreComp}$ associated (in the sense of Theorem 1.11) to the cubical monoid Δ^1, where $\mathsf{PreComp}$ is considered to be monoidal with respect to the Gray tensor product $\otimes^{\text{pre}}$. Its object part is thus given by $[1]^n \mapsto (\Delta^1)^{\otimes^{\text{pre}} n}$. This functor induces a strong monoidal left adjoint $T : \mathsf{cSet} \to \mathsf{PreComp}$ with right adjoint U. We first show that $(\Delta^1)^{\otimes^{\text{pre}} n} = (\Delta^1)^{\otimes n}$.

Proposition 5.6 *An r-simplex $\phi\colon \{1, \dots, n\} \to \{1, \dots, r, \pm\infty\}$ in $(\Delta^1)^{\otimes n}$ is unmarked if and only if there exist*

$$1 \le i_1 < \dots < i_r \le n$$

such that $\phi(i_p) = p$ for all $1 \le p \le r$. In particular, the only unmarked n-simplex in $(\Delta^1)^{\otimes n}$ is ι_n.

Proof This is proved in [Verity 2007, Observation 27] (with opposite indexing from ours). $\square$

Using the above characterization, we can indeed prove the following.

Proposition 5.7 *The marked simplicial set $(\Delta^1)^{\otimes n}$ is precomplicial for any $n \ge 0$.*

Proof Suppose for the sake of contradiction that there exists a map $\Delta^r_k{}' \to (\Delta^1)^{\otimes n}$ that cannot be extended to $\Delta^r_k{}''$. Regard this map as an r-simplex $\phi\colon \{1, \dots, n\} \to \{1, \dots, r, \pm\infty\}$. Then $\phi\partial_k$ is unmarked, so Proposition 5.6 implies that there exist

$$1 \le i_1 \le \dots \le i_{r-1} \le n$$

such that

- $\phi(i_p) = p$ for $1 \le p \le k-1$,
- $\phi(i_k) = k$ or $\phi(i_k) = k+1$, and
- $\phi(i_p) = p+1$ for $k+1 \le p \le r-1$.

But if $\phi(i_k) = k$ then the same sequence $i_1, \dots, i_{r-1}$ witnesses that $\phi\partial_{k+1}$ is unmarked, and similarly if $\phi(i_k) = k+1$ then $\phi\partial_{k-1}$ is unmarked. In either case, it contradicts with our assumption that ϕ corresponds to a map $\Delta^r_k{}' \to (\Delta^1)^{\otimes n}$. $\square$

Next we would like to upgrade the *domain* of T to a marked version too, by sending the marked n-cube to $\tau_{n-1}((\Delta^1)^{\otimes n})$. Proposition 5.6 implies that this marked simplicial set may be obtained from $(\Delta^1)^{\otimes n}$ by marking ι_n. The following lemma shows that it is indeed an object in $\mathsf{PreComp}$.

Lemma 5.8 *The marked simplicial set $\tau_{n-1}((\Delta^1)^{\otimes n})$ is precomplicial for any $n \geq 1$.*

Proof Suppose for the sake of contradiction that we are given a map $\phi\colon \Delta^{m\prime}_k \to \tau_{n-1}((\Delta^1)^{\otimes n})$ that cannot be extended to $\Delta^{m\prime\prime}_k$. Then ϕ must not factor through the precomplicial set $(\Delta^1)^{\otimes n}$, so ϕ sends at least one of the marked, nondegenerate simplices in $\Delta^{m\prime}_k$ to ι_n. Since all simplices in $\Delta^{m\prime}_k$ of dimension $> m$ are degenerate, it follows that $m \geq n$. On the other hand, we cannot have $m > n$ since $\phi\partial_k$ is unmarked in the $(n-1)$-trivial marked simplicial set $\tau_{n-1}((\Delta^1)^{\otimes n})$. Thus we must have $m = n$ and $\phi = \iota_n$. But at least one of ∂_{k-1} and ∂_{k+1} is a well-defined face of $\Delta^{m\prime}_k$, and it can be easily checked using Proposition 5.6 that ϕ sends this face to an unmarked simplex. This is the desired contradiction. $\square$

Thus we have defined the object part of $T : \square^+ \to \mathsf{PreComp}$, but we still need to define its value on the generating morphisms φ^n, ζ^n_i and $\xi^n_{i,\varepsilon}$, and verify the cocubical identities. The maps $T\varphi^n\colon T[1]^n \to T[1]^n_e$ are identity on the underlying simplicial sets and add the additional marking on ι_n. To define $T\zeta^n_i$ (resp. $T\xi^n_{i,\varepsilon}$), notice that we must have $T\zeta^n_i T\varphi^n = T\sigma^n_i$ (resp. $T\xi^n_{i\varepsilon} T\varphi^n = T\gamma^n_{i,\varepsilon}$). Since $T\sigma^n_i$ (resp. $T\gamma^n_{i,\varepsilon}$) sends the n-simplex ι_n to a degenerate (and hence marked) one, it must factor through $T\varphi^n$. Moreover, since $T\varphi^n$ is (entire and hence) an epimorphism, this factorization is unique, yielding a unique possible choice for $T\zeta^n_i$ (resp. $T\xi^n_{i,\varepsilon}$). Finally, to see that this definition satisfies the additional identities, we note that these involving φ are clear, whereas the remaining ones can be reduced to the usual cubical identities by precomposing with φ and using the fact that it is an epimorphism.

Hence we obtain a left adjoint functor T from structurally marked cubical sets to precomplicial sets. Moreover, the right adjoint U takes values in marked cubical sets, because the map $\square^n \to \widetilde{\square}^n$ is carried by T to an epimorphism. Thus by restricting the domain of T, we have constructed an adjunction $T \dashv U$ between marked cubical sets and precomplicial sets. In the remainder of the paper, we show that T is strong monoidal with respect to either version of the Gray tensor products and moreover left Quillen with respect to suitable model structures. We will make use of the following observation.

Proposition 5.9 *There are isomorphisms $T(X^{\mathrm{op}}) \cong (TX)^{\mathrm{op}}$ natural in $X \in \mathsf{cSet}^+$.*

Proof Since both $X \mapsto T(X^{\mathrm{op}})$ and $X \mapsto T(X)^{\mathrm{op}}$ are cocontinuous, it suffices to verify the assertion for $X = \square^n$ for $n \geq 0$ and $X = \widetilde{\square}^n$ for $n \geq 1$. The component at each $\square^n$ is simply given by

$$T((\square^n)^{\mathrm{op}}) = T(\square^n) = (\Delta^1)^{\otimes n} = ((\Delta^1)^{\mathrm{op}})^{\otimes n} \cong ((\Delta^1)^{\otimes n})^{\mathrm{op}} = T(\square^n)^{\mathrm{op}},$$

where the isomorphism is induced by the antimonoidality of $(-)^{\mathrm{op}}\colon \mathsf{sSet}^+ \to \mathsf{sSet}^+$ [Verity 2008a, Lemma 131], and the component at each $\widetilde{\square}^n$ is then obtained by applying τ_{n-1}.

It remains to check that these components are natural. Since the forgetful functor $\mathsf{sSet}^+ \to \mathsf{sSet}$ is faithful, we may instead check the naturality of the whiskering:

$$\begin{tikzcd} \square^+ \arrow[r, bend left, "T((-)^{\mathrm{op}})"] \arrow[r, bend right, "T(-)^{\mathrm{op}}"'] \arrow[r, phantom, "\Downarrow"] & \mathsf{sSet}^+ \arrow[r] & \mathsf{sSet} \end{tikzcd}$$

Now it is straightforward to check that this (potentially unnatural) transformation may also be written as

$$\begin{tikzcd} \square^+ \arrow[r] & \square \arrow[r, bend left, "T((-)^{\mathrm{op}})"] \arrow[r, bend right, "T(-)^{\mathrm{op}}"'] \arrow[r, phantom, "\Downarrow"] & \mathsf{sSet} \end{tikzcd}$$

where the first factor forgets the marking and the second factor (potentially unnatural transformation) is the unmarked analogue of our desired natural isomorphism. But in the unmarked case, we know that both $X \mapsto T(X^{\mathrm{op}})$ and $X \mapsto T(X)^{\mathrm{op}}$ are antimonoidal, and moreover it is straightforward to manually check that the naturality of its restriction to the full subcategory spanned by $\square^0$, $\square^1$ and $\square^2$. Thus the desired naturality follows from Theorem 1.11. $\square$

6 Triangulating Gray tensor product

We now prove that the triangulation functor is strong monoidal with respect to either version of the Gray tensor product. We begin by describing a proof strategy that will be used in both the lax and the pseudo cases.

6.1 Proof strategy

The proof typically reduces to showing an entire inclusion $A \hookrightarrow B$ to be a complicial marking extension where A and B are certain entire supersets of $(\Delta^1)^{\otimes N}$. (The integer N will be of the form $N = m + n$ in the actual proofs, but this is irrelevant in this subsection.) There are three kinds of simplices of interest, namely those that are

(i) marked in $(\Delta^1)^{\otimes N}$,

(ii) marked in A but not in $(\Delta^1)^{\otimes N}$, and

(iii) marked in B but not in A.

The simplices of type (i) are characterized by Proposition 5.6. The first step of the proof will be to (define suitable A, B and) characterize simplices of type (ii) and (iii).

Before proceeding, we need the following definitions.

Definition 6.1 For any r-simplex $\phi\colon \{1,\dots,N\} \to \{1,\dots,r,\pm\infty\}$ in $(\Delta^1)^N$, define

$$\mathcal{D}(\phi) = |\phi^{-1}(\{1,\dots,r\})| - r \quad \text{and} \quad \mathcal{O}(\phi) = \{(i,j) \in \{1,\dots,N\}^2 \mid i < j, \phi(i) < \phi(j)\}.$$

The integer $\mathcal{D}(\phi)$ measures how "diagonal" ϕ is, and the set $\mathcal{O}(\phi)$ measures how "in order" ϕ is.

We complicially extend the marking on A to those simplices ϕ of type (iii) by nested induction on $\mathcal{D}(\phi)$ and $|\mathcal{O}(\phi)|$. More precisely, consider the lexicographical ordering on $\mathbb{Z} \times \mathbb{N}$ so that $(u_1,v_1) \le (u_2,v_2)$ if and only if

- $u_1 < u_2$, or
- $u_1 = u_2$ and $v_1 \le v_2$.

For each $(u,v) \in \mathbb{Z} \times \mathbb{N}$, let $A(u,v)$ denote the marked simplicial set obtained from A by marking those simplices ϕ such that ϕ is marked in B and $(\mathcal{D}(\phi), |\mathcal{O}(\phi)|) < (u,v)$. Then

- $A(u_1,v_1)$ is an entire subset of $A(u_2,v_2)$ for any $(u_1,v_1) \le (u_2,v_2)$,
- $\operatorname{colim}_v A(u,v) = A(u+1,0)$ for any $u \in \mathbb{Z}$,
- $\operatorname{colim}_{u,v} A(u,v) = B$, and
- $A(0,0) = A$ (by Proposition 5.5).

Now we assume the following.

Assumption 1 *Any marked simplex ϕ in B with $\mathcal{D}(\phi) = 0$ is marked in A.*

Then we may upgrade the last bulleted item to $A(1,0) = A$. Thus to prove that $A \to B$ is a complicial marking extension, it suffices to exhibit the map $A(u,v) \to A(u,v+1)$ as a complicial marking extension for each $(u,v) \ge (1,0)$.

So fix $(u,v) \ge (1,0)$ and suppose that we are given an r-simplex ϕ of type (iii) with $(\mathcal{D}(\phi), |\mathcal{O}(\phi)|) = (u,v)$. Then in particular $\mathcal{D}(\phi) \ge 1$. So by the pigeonhole principle, we can choose $1 \le p_\phi \le r$ such that $|\phi^{-1}(p_\phi)| \ge 2$. Let $i_\phi = \min \phi^{-1}(p_\phi)$. Let $\widetilde{\phi}$ be the $(r+1)$-simplex given by

$$\widetilde{\phi}(i) = \begin{cases} \phi(i) & \text{if } \phi(i) \le p_\phi \text{ and } i \ne i_\phi, \\ \phi(i) + 1 & \text{if } \phi(i) > p_\phi \text{ or } i = i_\phi. \end{cases}$$

Observe that we have $\widetilde{\phi}\partial_{p_\phi} = \phi$. We wish to show that this simplex $\widetilde{\phi}$ extends to $\Delta^{r+1'}_{p_\phi}$:

$$\begin{tikzcd} \Delta^{r+1} \arrow[r, "\widetilde{\phi}"] \arrow[d, hook] & A(u,v) \\ \Delta^{r+1'}_{p_\phi} \arrow[ur, dashed, "\exists"'] & \end{tikzcd}$$

Assuming this fact, we can deduce that we have a pushout square

$$\begin{tikzcd}
\coprod \Delta^{r+1'}_{p_\phi} \arrow[r] \arrow[d, hook] & A(u,v) \arrow[d] \\
\coprod \Delta^{r+1''}_{p_\phi} \arrow[r] & A(u,v+1)
\end{tikzcd}$$

where the coproducts are taken over all r-simplices ϕ of type (iii) with $(\mathcal{D}(\phi), |\mathcal{O}(\phi)|) = (u, v)$ for various r, and the horizontal maps are induced by $\widetilde{\phi}$.

The following lemma implies that $\widetilde{\phi}$ at least extends to $\Delta^{r+1}_{p_\phi}$.

Lemma 6.2 *Let $\alpha\colon [q] \to [r+1]$ be a face operator with $\{p_\phi, p_\phi \pm 1\} \subset \operatorname{im}\alpha$. Then $\widetilde{\phi}\alpha$ is marked in $(\Delta^1)^{\otimes N}$.*

Proof Let $p \in [q]$ be the necessarily unique element with $\alpha(p) = p_\phi$. Then we must have $\alpha(p-1) = p_\phi - 1$ and $\alpha(p+1) = p_\phi + 1$. Now one can check using Proposition 5.3 and the minimality of i_ϕ that $(\widetilde{\phi}\alpha)^{-1}(p+1) = \{i_\phi\}$ and moreover any $1 \le i \le N$ satisfying $(\widetilde{\phi}\alpha)(i) = p$ must also satisfy $i > i_\phi$. Thus $\widetilde{\phi}\alpha$ is marked in $(\Delta^1)^{\otimes N}$ by Proposition 5.6. $\square$

Therefore it remains to prove that the faces $\chi = \widetilde{\phi}\partial_{p_\phi - 1}$ and $\psi = \widetilde{\phi}\partial_{p_\phi+1}$ are marked in $A(u, v)$. First, we describe these simplices explicitly.

Lemma 6.3 *The simplex χ is given by*

$$\chi(i) = \begin{cases} -\infty & \text{if } \phi(i) = p_\phi \text{ and } i \neq i_\phi, \\ \phi(i) & \text{otherwise.} \end{cases}$$

if $p_\phi = 1$, and

$$\chi(i) = \begin{cases} p_\phi - 1 & \text{if } \phi(i) = p_\phi \text{ and } i \neq i_\phi, \\ \phi(i) & \text{otherwise} \end{cases}$$

if $p_\phi \ge 2$. The simplex ψ is given by

$$\psi(i) = \begin{cases} +\infty & \text{if } i = i_\phi, \\ \phi(i) & \text{otherwise} \end{cases}$$

if $p_\phi = r$ and

$$\psi(i) = \begin{cases} p_\phi + 1 & \text{if } i = i_\phi, \\ \phi(i) & \text{otherwise} \end{cases}$$

if $p_\phi < r$.

Proof This is a routine application of Proposition 5.3. $\square$

These explicit descriptions allow us to prove the following.

Lemma 6.4 *The simplices χ and ψ satisfy*

$$(\mathcal{D}(\chi), |\mathcal{O}(\chi)|) < (\mathcal{D}(\phi), |\mathcal{O}(\phi)|) \quad \text{and} \quad (\mathcal{D}(\psi), |\mathcal{O}(\psi)|) < (\mathcal{D}(\phi), |\mathcal{O}(\phi)|).$$

Proof If $p_\phi = 1$ then clearly $\mathcal{D}(\chi) < \mathcal{D}(\phi)$.

Suppose $p_\phi \geq 2$. Then we have $\mathcal{D}(\chi) = \mathcal{D}(\phi)$. We claim that $\mathcal{O}(\chi)$ is a proper subset of $\mathcal{O}(\phi)$. Indeed, it can be seen from Lemma 6.3 that if a pair (i, j) satisfies $\phi(i) \geq \phi(j)$ and $\chi(i) < \chi(j)$ then we must have $\phi(i) = \phi(j) = p_\phi$ and $i \neq i_\phi = j$. But then the minimality of i_ϕ implies $i > j$, and this shows that there is no pair (i, j) in $\mathcal{O}(\chi) \setminus \mathcal{O}(\phi)$. On the other hand, since ϕ is unmarked in $(\Delta^1)^{\otimes N}$, there exist $1 \leq i_1 \leq \cdots \leq i_r \leq N$ such that $\phi(i_p) = p$ for $1 \leq p \leq r$. It is straightforward to check that the pair $(i_{p_\phi - 1}, \max \phi^{-1}(p_\phi))$ is then in $\mathcal{O}(\phi) \setminus \mathcal{O}(\chi)$. Therefore $\mathcal{O}(\chi)$ is a proper subset of $\mathcal{O}(\phi)$, and this proves the lexicographical inequality concerning χ.

The simplex ψ can be treated dually. □

The last missing piece of the proof (that $A \to B$ is a complicial marking extension) is the following.

Assumption 2 *The simplices χ and ψ are marked in B.*

This completes the proof strategy. (Whether Assumptions 1 and 2 hold depends on the exact definitions of A and B, so there is no general strategy for verifying them.)

6.2 Triangulating the lax Gray tensor product

The goal of this subsection is to prove the following theorem.

Theorem 6.5 *The adjunction $T \dashv U$ is monoidal with respect to the lax Gray tensor products. Equivalently, $T \colon (\mathsf{cSet}^+, \otimes) \to (\mathsf{PreComp}, \otimes^{\mathrm{pre}})$ is strong monoidal.*

Fix $m \geq 1$ and $n \geq 0$. Observe that

$$\begin{tikzcd} \coprod \square^{m+k} \arrow[r] \arrow[d, "\coprod \varphi"'] & \square^{m+n} \arrow[d] \\ \coprod \widetilde{\square}^{m+k} \arrow[r] & \widetilde{\square}^m \otimes \square^n \end{tikzcd}$$

is a pushout square in cSet^+, where the coproducts are taken over all face maps $[1]^k \to [1]^n$. This pushout is preserved by T, so the right square in

$$\begin{tikzcd} \coprod \Delta^{m+k} \arrow[r, "\coprod \iota_{m+k}"] \arrow[d] & \coprod (\Delta^1)^{\otimes(m+k)} \arrow[r] \arrow[d] & (\Delta^1)^{\otimes(m+n)} \arrow[d] \\ \coprod \widetilde{\Delta}^{m+k} \arrow[r] & \coprod \tau_{m+k-1}\big((\Delta^1)^{\otimes(m+k)}\big) \arrow[r] & T(\widetilde{\square}^m \otimes \square^n) \end{tikzcd}$$

is a pushout square in PreComp, where the upper horizontal map is induced by $\mathrm{id}_{(\Delta^1)^{\otimes m}} \otimes T(\phi)$ for various face maps $\phi\colon [1]^k \to [1]^n$. The left square is also a pushout by Proposition 5.6, so the pasted square is a pushout too. In this subsection, we define A to be the corresponding pushout in sSet^+ (and not in PreComp)

$$\begin{tikzcd} \coprod \Delta^{m+k} \arrow[r] \arrow[d] \arrow[dr, phantom, "\text{p.o.}"] & (\Delta^1)^{\otimes(m+n)} \arrow[d] \\ \coprod \widetilde{\Delta}^{m+k} \arrow[r] & A \end{tikzcd}$$

so that its precomplicial reflection A^{pre} is precisely $T(\widetilde{\square}^m \otimes \square^n)$.

Now we give combinatorial characterizations of marked simplices in A and in $T(\widetilde{\square}^m) \otimes T(\square^n)$.

Lemma 6.6 *An r-simplex ϕ is marked in A but not in $(\Delta^1)^{\otimes(m+n)}$ if and only if $r \geq m$, $\phi(i) = i$ for all $1 \leq i \leq m$ and the restriction*

$$\phi^{-1}(\{1, \dots, r\}) \to \{1, \dots, r\}$$

of ϕ is an isomorphism of linearly ordered sets.

Proof Compute the colimit. □

Lemma 6.7 *Let ϕ be an unmarked r-simplex in $(\Delta^1)^{\otimes(m+n)}$. Then ϕ is marked in $T(\widetilde{\square}^m) \otimes T(\square^n)$ if and only if:*

(1) $r \geq m$;

(2) *$\phi(i) = i$ for all $1 \leq i \leq m$; and*

(3) *there does NOT exist a sequence $m < i_m < i_{m+1} < \cdots < i_r \leq m+n$ such that $\phi(i_p) = p$ for all $m \leq p \leq r$.*

Proof Write $\phi = (\phi_1, \phi_2)$ for ϕ regarded as a simplex in the product simplicial set $(\Delta^1)^m \times (\Delta^1)^n$. That is, $\phi_1\colon \{1, \dots, m\} \to \{1, \dots, r, \pm\infty\}$ and $\phi_2\colon \{1, \dots, n\} \to \{1, \dots, r, \pm\infty\}$ are respectively given by $\phi_1(i) = \phi(i)$ and $\phi_2(i) = \phi(m+i)$.

It follows from the definitions of $\otimes$ and $T(\widetilde{\square}^m)$ that ϕ (which we are assuming to be unmarked in $(\Delta^1)^{\otimes(m+n)}$) is marked in $T(\widetilde{\square}^m) \otimes T(\square^n)$ if and only if

(a) $r \geq m$,

(b) (ϕ_1, ϕ_2) is q-cloven for all q except for $q = m$,

(c) $\phi_1 \amalg_1^{m,r-m} = \iota_m$, and

(d) $\phi_2 \amalg_2^{m,r-m}$ is unmarked in $(\Delta^1)^{\otimes n}$.

The clauses (a) and (c) here clearly correspond respectively to (1) and (2) in the lemma. Since we are assuming ϕ to be unmarked in $(\Delta^1)^{\otimes(m+n)}$, Proposition 5.6 implies that there exists a sequence $1 \le j_1 < \cdots < j_r \le m+n$ such that $\phi(j_p) = p$ for all $1 \le p \le r$. Note that the strict inequalities imply $j_{m+1} > m$. One can now check using Example 5.4 that $(\phi_2 \amalg_2^{m,r-m})(j_p - m) = p - m$ for all $m+1 \le p \le r$. Thus the clause (d) is in fact redundant by Proposition 5.6.

It remains to check that, assuming (a), (c) and (d), the clauses (3) and (b) are equivalent. Note that (ϕ_1, ϕ_2) is q-cloven for any $m < q \le r$ since the q-simplex $\phi_1 \amalg_1^{q,r-q}$ in the simplicial set $(\Delta^1)^m$ must be degenerate. For $0 \le q < m$, since we are assuming (c), $\phi_1 \amalg_1^{q,r-q} = \iota_q$ is unmarked in $T(\widetilde{\square}^m)$. So (b) is equivalent to $\phi_2 \amalg_2^{q,r-q}$ being marked in $(\Delta^1)^{\otimes n}$ for all $0 \le q < m$. By Example 5.4 and Proposition 5.6, this latter condition for fixed q is equivalent to the NON-existence of a sequence

$$m < i_{q+1} < \cdots < i_r \le m+n$$

such that $\phi(i_p) = p$ for all $q+1 \le p \le r$. Clearly the nonexistence for $q = m-1$, which is precisely (3), implies the nonexistence for all other values of q. This completes the proof. □

By combining Proposition 5.6 and Lemma 6.7, we obtain the following.

Lemma 6.8 *An r-simplex ϕ in $T(\widetilde{\square}^m) \otimes T(\square^n)$ with $r \ge m$ is unmarked if and only if there exist*

$$1 \le i_1 < \cdots < i_r \le m+n$$

such that $\phi(i_p) = p$ for all $1 \le p \le r$ and moreover $i_m > m$.

Proof Let ϕ be an unmarked r-simplex in $(\Delta^1)^{\otimes(m+n)}$ with $r \ge m$. Note that ϕ is unmarked in $T(\widetilde{\square}^m) \otimes T(\square^n)$ if and only if it violates either Lemma 6.7(2) or (3).

The "if" direction is easy since the existence of a sequence satisfying the condition stated in the lemma would immediately contradict Lemma 6.7(3).

For the "only if" direction, assume that ϕ is unmarked in $T(\widetilde{\square}^m) \otimes T(\square^n)$. Recall that by Proposition 5.6 there exist

$$1 \le j_1 < \cdots < j_r \le m+n$$

such that $\phi(j_p) = p$ for all $1 \le p \le r$. If $j_m > m$, then simply taking $i_p = j_p$ for all p would yield the desired sequence. So assume $j_m = m$. Then since the inequalities $j_1 < \cdots < j_m$ are strict, we must have $j_p = p$ for all $1 \le p \le m$. Thus ϕ cannot violate Lemma 6.7(2), so it must violate (3). That is, there exist $m < i_m < i_{m+1} < \cdots < i_r \le m+n$ such that $\phi(i_p) = p$ for all $m \le p \le r$. We then obtain the desired sequence by taking $i_p = j_p$ for $1 \le p \le m-1$. □

Lemma 6.9 *There is a complicial marking extension $A \to T(\widetilde{\square}^m) \otimes T(\square^n)$ that commutes with the evident inclusions of $(\Delta^1)^{\otimes(m+n)}$.*

Proof We apply the proof strategy from Section 6.1 with $B = T(\tilde{\square}^m) \otimes T(\square^n)$.

One can easily check using Lemmas 6.6 and 6.7 that any marked simplex ϕ in $T(\tilde{\square}^m) \otimes T(\square^n)$ with $\mathcal{D}(\phi) = 0$ must also be marked in A. This verifies Assumption 1.

To verify Assumption 2, let ϕ be an r-simplex that is marked in $T(\tilde{\square}^m) \otimes T(\square^n)$ but not in A. Then we necessarily have $r \geq m$ by Lemma 6.7.

Consider the simplex $\chi = \tilde{\phi}\partial_{p_\phi - 1}$. Suppose for contradiction that χ is unmarked in $T(\tilde{\square}^m) \otimes T(\square^n)$. Then by Lemma 6.8 there exist $1 \leq i_1 < \cdots < i_r \leq m + n$ such that $\chi(i_p) = p$ for all $1 \leq p \leq r$ and $i_m > m$.

- If $p_\phi = 1$, then we also have $\phi(i_p) = p$ for all p by Lemma 6.3, thus ϕ is unmarked in $T(\tilde{\square}^m) \otimes T(\square^n)$. This is the desired contradiction.
- Suppose $p_\phi \geq 2$. We claim that $\phi(i_p) = p$ holds for all p in this case too. According to Lemma 6.3, the only thing we must check is that $\phi(i_{p_\phi - 1}) = p_\phi - 1$ holds (as opposed to $\phi(i_{p_\phi - 1}) = p_\phi$). To see that this is indeed the case, observe that $\chi^{-1}(p_\phi) = \{i_\phi\}$ by Lemma 6.3. Thus we must have $i_{p_\phi} = i_\phi$. Since $i_{p_\phi - 1} < i_{p_\phi}$, the minimality of i_ϕ implies that $i_{p_\phi - 1} \notin \phi^{-1}(p_\phi)$. Therefore we have obtained the desired contradiction.

The simplex $\psi = \tilde{\phi}\partial_{p_\phi + 1}$ can be similarly checked to be marked in $T(\tilde{\square}^m) \otimes T(\square^n)$. This completes the proof. $\square$

Proof of Theorem 6.5 Since both $T(- \otimes -)$ and $T(-) \otimes^{\mathrm{pre}} T(-)$ preserve colimits in each variable, it suffices to check the existence of natural isomorphisms $T(X \otimes Y) \cong T(X) \otimes^{\mathrm{pre}} T(Y)$ for X, Y generic (possibly marked) cubes.

By construction of T, we have $T(\square^m \otimes \square^n) \cong T(\square^m) \otimes^{\mathrm{pre}} T(\square^n)$ for any $m, n \geq 0$.

For any $m \geq 1$ and $n \geq 0$, we may obtain an isomorphism $T(\tilde{\square}^m \otimes \square^n) \cong T(\tilde{\square}^m) \otimes^{\mathrm{pre}} T(\square^n)$ by reflecting the complicial marking extension of Lemma 6.9 into $\mathsf{PreComp}$. Dually, we have $T(\square^m \otimes \tilde{\square}^n) \cong T(\square^m) \otimes^{\mathrm{pre}} T(\tilde{\square}^n)$ for any $m \geq 0$ and $n \geq 1$.

Let $m, n \geq 1$. Observe that the left square below is a pushout in cSet^+ by Lemma 2.17(3):

$$\begin{tikzcd}
\square^m \otimes \square^n \arrow[r] \arrow[d] & \tilde{\square}^m \otimes \square^n \arrow[d] \\
\square^m \otimes \tilde{\square}^n \arrow[r] & \tilde{\square}^m \otimes \tilde{\square}^n
\end{tikzcd}
\qquad
\begin{tikzcd}
T(\square^m \otimes \square^n) \arrow[r] \arrow[d] & T(\tilde{\square}^m \otimes \square^n) \arrow[d] \\
T(\square^m \otimes \tilde{\square}^n) \arrow[r] & T(\tilde{\square}^m \otimes \tilde{\square}^n)
\end{tikzcd}$$

Since T is cocontinuous, it follows that the right square is a pushout in $\mathsf{PreComp}$. On the other hand, since both $T(\square^m) \to T(\tilde{\square}^m)$ and $T(\square^n) \to T(\tilde{\square}^n)$ are entire, the square below is a pushout in $\mathsf{PreComp}$

by [Verity 2008a, Lemma 140]:

$$\begin{tikzcd} T(\square^m) \otimes^{\mathrm{pre}} T(\square^n) \arrow[r] \arrow[d] & T(\widetilde{\square}^m) \otimes^{\mathrm{pre}} T(\square^n) \arrow[d] \\ T(\square^m) \otimes^{\mathrm{pre}} T(\widetilde{\square}^n) \arrow[r] & T(\widetilde{\square}^m) \otimes^{\mathrm{pre}} T(\widetilde{\square}^n) \end{tikzcd}$$

Thus by comparing the two pushout squares in $\mathsf{PreComp}$, we obtain $T(\widetilde{\square}^m \otimes \widetilde{\square}^n) \cong T(\widetilde{\square}^m) \otimes^{\mathrm{pre}} T(\widetilde{\square}^n)$. The naturality of these isomorphisms is evident, and this completes the proof. $\square$

6.3 Triangulating the pseudo Gray tensor product

The goal of this subsection is to prove the following theorem.

Theorem 6.10 *The adjunction $T \dashv U$ is monoidal with respect to the pseudo Gray tensor products. Equivalently*, $T\colon (\mathsf{cSet}^+, \circledast) \to (\mathsf{PreComp}, \circledast)$ *is strong monoidal.*

Fix $m, n \geq 1$. By Remark 2.14, the square

$$\begin{tikzcd} \coprod \square^{k+\ell} \arrow[r] \arrow[d] & \square^{m+n} \arrow[d] \\ \coprod \widetilde{\square}^{k+\ell} \arrow[r] & \square^m \circledast \square^n \end{tikzcd}$$

is a pushout in cSet^+, where the coproducts are taken over all pairs of face maps $\square^k \to \square^m$ and $\square^\ell \to \square^n$ such that $k, \ell \geq 1$. This pushout is preserved by T, so the right square in

$$\begin{tikzcd} \coprod \Delta^{k+\ell} \arrow[r, "\coprod \iota_{k+\ell}"] \arrow[d] & \coprod (\Delta^1)^{\otimes(k+\ell)} \arrow[r] \arrow[d] & (\Delta^1)^{\otimes(m+n)} \arrow[d] \\ \coprod \widetilde{\Delta}^{k+\ell} \arrow[r] & \coprod \tau_{k+\ell-1}\big((\Delta^1)^{\otimes(k+\ell)}\big) \arrow[r] & T(\square^m \circledast \square^n) \end{tikzcd}$$

is a pushout square in $\mathsf{PreComp}$. The left square is also a pushout by Proposition 5.6, so the pasted square is a pushout too. In this subsection, we define A to be the corresponding pushout in sSet^+ (and not in $\mathsf{PreComp}$):

$$\begin{tikzcd} \coprod \Delta^{k+\ell} \arrow[r] \arrow[d] \arrow[dr, phantom, "\text{p.o.}"] & (\Delta^1)^{\otimes(m+n)} \arrow[d] \\ \coprod \widetilde{\Delta}^{k+\ell} \arrow[r] & A \end{tikzcd}$$

so that its precomplicial reflection A^{pre} is precisely $T(\square^m \circledast \square^n)$.

Lemma 6.11 *An r-simplex ϕ is marked in A but not in $(\Delta^1)^{\otimes(m+n)}$ if and only if the restriction*

$$\phi^{-1}(\{1, \dots, r\}) \to \{1, \dots, r\}$$

of ϕ is an isomorphism of linearly ordered sets and moreover $\phi^{-1}(\{1, \dots, r\})$ intersects both $\{1, \dots, m\}$ and $\{m+1, \dots, m+n\}$.

Proof Compute the colimit. $\square$

Lemma 6.12 *An r-simplex ϕ in $T(\square^m) \circledast T(\square^n)$ is unmarked if and only if there exist either*

$$1 \le i_1 < \cdots < i_r \le m \quad \text{or} \quad m+1 \le i_1 < \cdots < i_r \le m+n$$

such that $\phi(i_p) = p$ for all $1 \le p \le r$.

Proof Since $\circledast$ is the categorical product on $\mathsf{PreComp}$, ϕ is marked if and only if both $\pi_1(\phi)$ and $\pi_2(\phi)$ are marked. Equivalently, ϕ is unmarked if and only if either $\pi_1(\phi)$ or $\pi_2(\phi)$ is unmarked. Thus the assertion follows from Proposition 5.6. $\square$

Lemma 6.13 *There is a complicial marking extension $A \to T(\square^m) \circledast T(\square^n)$ that commutes with the evident inclusions of $(\Delta^1)^{\otimes(m+n)}$.*

Proof We apply the proof strategy from Section 6.1 with $B = T(\square^m) \circledast T(\square^n)$.

One can easily check using Proposition 5.6 and Lemmas 6.11 and 6.12 that any marked simplex ϕ in $T(\square^m) \circledast T(\square^n)$ with $\mathcal{D}(\phi) = 0$ must also be marked in A. This verifies Assumption 1.

To see that Assumption 2 holds for χ, suppose for contradiction that χ is unmarked in $T(\square^m) \circledast T(\square^n)$. By Lemma 6.12, this unmarked-ness is witnessed by a sequence $i_1, \dots, i_r$, but then the same sequence can be checked to witness that ϕ is unmarked in $T(\square^m) \circledast T(\square^n)$. The details are similar to the corresponding part in the proof of Lemma 6.9. Assumption 2 for ψ can be checked similarly. $\square$

Let $m \ge 1$ and $n \ge 0$. Observe that the square below is a pushout in cSet^+:

$$\begin{tikzcd}
\coprod \square^m \arrow[r] \arrow[d] & \square^m \circledast \square^n \arrow[d] \\
\coprod \widetilde{\square}^m \arrow[r] & \widetilde{\square}^m \circledast \square^n
\end{tikzcd}$$

where the coproducts are taken over all $[1]^0 \to [1]^n$. This pushout is preserved by T, so the right square in

$$\begin{tikzcd}
\coprod \Delta^m \arrow[r, "\coprod \iota_m"] \arrow[d] & \coprod T(\square^m) \arrow[r] \arrow[d] & T(\square^m \circledast \square^n) \arrow[d] \\
\coprod \widetilde{\Delta}^m \arrow[r] & \coprod T(\widetilde{\square}^m) \arrow[r] & T(\widetilde{\square}^m \circledast \square^n)
\end{tikzcd}$$

is a pushout in PreComp. The left square is also a pushout by Proposition 5.6, so the pasted square is a pushout too. Let A' denote the "corresponding" pushout in sSet^+ (and not in PreComp):

$$\begin{tikzcd} \coprod \Delta^m \arrow[r] \arrow[d] \arrow[dr, phantom, "\text{p.o.}"] & A \arrow[d] \\ \coprod \widetilde{\Delta}^m \arrow[r] & A' \end{tikzcd}$$

so that its precomplicial reflection $(A')^{\mathrm{pre}}$ is precisely $T(\widetilde{\square}^m \circledast \square^n)$.

Lemma 6.14 *The marked simplicial set A' is obtained from A by marking those m-simplices ϕ such that $\phi(i) = i$ for $1 \le i \le m$ and $\phi(i) \in \{\pm\infty\}$ for $i > m$.*

The unmarked simplices in $B' = T(\widetilde{\square}^m) \circledast T(\square^n)$ admit a characterization similar to Lemma 6.12.

Lemma 6.15 *An r-simplex in $T(\widetilde{\square}^m) \circledast T(\square^n)$ with $r \neq m$ is unmarked if and only if it is unmarked in $T(\square^m) \circledast T(\square^n)$. An m-simplex ϕ in $T(\widetilde{\square}^m) \circledast T(\square^n)$ is unmarked if and only if there exist*

$$m+1 \le i_1 < \cdots < i_m \le m+n$$

such that $\phi(i_p) = p$ for all $1 \le p \le m$.

Lemma 6.16 *There is a complicial marking extension $A' \to T(\widetilde{\square}^m) \circledast T(\square^n)$ that commutes with the evident inclusions of $(\Delta^1)^{\otimes(m+n)}$.*

Proof We first check that Assumption 1 holds. Let ϕ be a marked r-simplex in $T(\widetilde{\square}^m) \circledast T(\square^n)$ with $\mathcal{D}(\phi) = 0$. Note that, if ϕ is marked in $T(\square^m) \circledast T(\square^n)$ then we already know that it is marked in A (and so in A' too). So suppose otherwise. Then we can see from Lemmas 6.12 and 6.15 that we must have $r = m$ and a sequence

$$1 \le i_1 < \cdots < i_m \le m$$

such that $\phi(i_p) = p$ for all $1 \le p \le m$. It is then easy to check that ϕ is one of the extra marked simplices described in Lemma 6.14.

For Assumption 2, we assume for contradiction that χ or ψ is unmarked, obtain a sequence using Lemma 6.15, and deduce that ϕ is unmarked. The details are similar to those in the proofs of Lemmas 6.9 and 6.13. $\square$

Proof of Theorem 6.10 Since both $T(-\circledast-)$ and $T(-)\circledast T(-)$ preserve colimits in each variable, it suffices to check the existence of natural isomorphisms $T(X \circledast Y) \cong T(X) \circledast T(Y)$ for X, Y generic (possibly marked) cubes.

For the appropriate values of m and n, we may obtain isomorphisms

$$\begin{aligned} T(\square^m \circledast \square^n) &\cong T(\square^m) \circledast T(\square^n), \\ T(\widetilde{\square}^m \circledast \square^n) &\cong T(\widetilde{\square}^m) \circledast T(\square^n), \\ T(\square^m \circledast \widetilde{\square}^n) &\cong T(\square^m) \circledast T(\widetilde{\square}^n) \end{aligned}$$

by reflecting to $\mathsf{PreComp}$ the complicial marking extensions of Lemmas 6.13 and 6.16 and the dual of the latter respectively.

Let $m, n \geq 1$. Observe that the left square below is a pushout in cSet^+ by Lemma 2.17(3):

$$\begin{tikzcd} \square^m \circledast \square^n \arrow[r] \arrow[d] & \widetilde{\square}^m \circledast \square^n \arrow[d] \\ \square^m \circledast \widetilde{\square}^n \arrow[r] & \widetilde{\square}^m \circledast \widetilde{\square}^n \end{tikzcd} \qquad \begin{tikzcd} T(\square^m \circledast \square^n) \arrow[r] \arrow[d] & T(\widetilde{\square}^m \circledast \square^n) \arrow[d] \\ T(\square^m \circledast \widetilde{\square}^n) \arrow[r] & T(\widetilde{\square}^m \circledast \widetilde{\square}^n) \end{tikzcd}$$

Since T is cocontinuous, it follows that the right square is a pushout in $\mathsf{PreComp}$. On the other hand, since both $T(\square^m) \to T(\widetilde{\square}^m)$ and $T(\square^n) \to T(\widetilde{\square}^n)$ are entire, the square below is a pushout in $\mathsf{PreComp}$ by Proposition 1.28:

$$\begin{tikzcd} T(\square^m) \circledast T(\square^n) \arrow[r] \arrow[d] & T(\widetilde{\square}^m) \circledast T(\square^n) \arrow[d] \\ T(\square^m) \circledast T(\widetilde{\square}^n) \arrow[r] & T(\widetilde{\square}^m) \circledast T(\widetilde{\square}^n) \end{tikzcd}$$

Thus by comparing the two pushout squares in $\mathsf{PreComp}$, we obtain $T(\widetilde{\square}^m \circledast \widetilde{\square}^n) \cong T(\widetilde{\square}^m) \circledast T(\widetilde{\square}^n)$. The naturality of these isomorphisms is evident, and this completes the proof. $\square$

7 Triangulating model structures

The main theorem of our final section is the following.

Theorem 7.1 *The adjunction $T \dashv U$ is a Quillen adjunction with respect to the comical model structure on cSet^+ and the complicial model structure on $\mathsf{PreComp}$.*

In the following proof, we denote a nondegenerate r-simplex $\phi\colon \{1,2,3\} \to \{1,\dots,r,\pm\infty\}$ in the simplicial set $(\Delta^1)^{\times 3}$ by the sequence $\phi(1)\phi(2)\phi(3)$, omitting the letter ∞. For instance, $21-$ denotes the 2-simplex ϕ given by $\phi(1) = 2$, $\phi(2) = 1$ and $\phi(3) = -\infty$. Note that since ϕ is assumed to be nondegenerate, the dimension of ϕ can be recovered as the maximum integer appearing in the sequence.

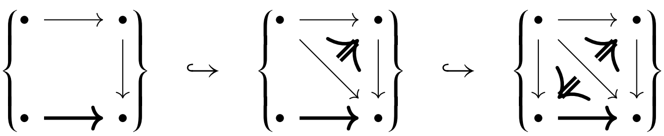
Proof We first show that T preserves cofibrations. It suffices to prove that T sends the boundary inclusions and the markers in cSet^+ to monomorphisms in sSet^+. Clearly T sends the boundary inclusions $\partial\square^0 \hookrightarrow \square^0$ and $\partial\square^1 \hookrightarrow \square^1$ to (maps that are isomorphic to) the boundary inclusions $\partial\Delta^0 \hookrightarrow \Delta^0$ and $\partial\Delta^1 \hookrightarrow \Delta^1$ respectively. For any $n \geq 2$, we have

$$T(\partial\square^n \hookrightarrow \square^n) \cong T((\partial\square^1 \hookrightarrow \square^1)^{\widehat{\otimes} n}) \cong (T(\partial\square^1 \hookrightarrow \square^1))^{\widehat{\otimes} n} \cong (\partial\Delta^1 \hookrightarrow \Delta^1)^{\widehat{\otimes} n}$$

by Proposition 1.17 and Theorem 6.5, and the last map is clearly a monomorphism. Also, T sends the marker $\square^n \hookrightarrow \widetilde{\square}^n$ to the monomorphism $(\Delta^1)^{\otimes n} \to \tau_{n-1}((\Delta^1)^{\otimes n})$ by definition. This shows that T preserves cofibrations.

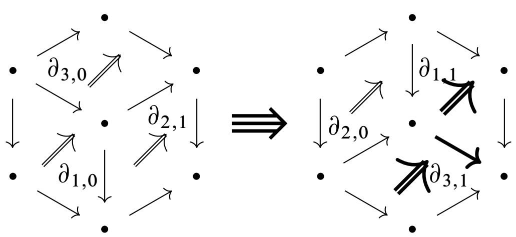
Next we show that T sends the open box inclusions to trivial cofibrations. We will check this "by hand" on the boxes of dimension ≤ 3. This will imply the general case since the higher-dimensional box inclusions are generated by these low-dimensional ones in the sense of Proposition 3.10, T is strong monoidal with respect to the lax Gray tensor products (Theorem 6.5), and the Leibniz Gray tensor product of a complicial horn inclusion and a monomorphism (in $\mathsf{PreComp}$) may be obtained as a composite of pushouts of complicial horn inclusions [Verity 2008b, Lemma 72].

Clearly T sends $\sqcap^1_{1,\varepsilon} \hookrightarrow \square^1_{1,\varepsilon}$ to the trivial cofibration $\Lambda^1_{1-\varepsilon} \hookrightarrow \Delta^1_{1-\varepsilon}$. Consider the open box inclusion $\sqcap^2_{1,0} \hookrightarrow \square^2_{1,0}$. Its image under T may be written as a pushout of the horn inclusion $\Lambda^2_1 \hookrightarrow \Delta^2_1$ followed by a pushout of $\Lambda^2_2 \hookrightarrow \Delta^2_1$. The following pictures (in which thick arrows indicate marked simplices) depict this factorization:



The box inclusions $\sqcap^2_{k,\varepsilon} \hookrightarrow \square^2_{k,\varepsilon}$ for other values of k and/or ε can be treated similarly.

Now consider the open box inclusion $\sqcap^3_{2,0} \hookrightarrow \square^3_{2,0}$. Observe that the only marked, nondegenerate cubes in $\square^3_{2,0}$ are id, $\partial_{1,1}$, $\partial_{3,1}$ and $\partial_{3,1}\partial_{1,1}$:



Let B' denote the marked simplicial set obtained from $(\Delta^1)^{\otimes 3}$ by marking the 3-simplex $\iota_3 = 123$, the 2-simplices $12-$ and -12 and the 1-simplex $-1-$. Then the precomplicial reflection $(B')^{\mathrm{pre}}$ is precisely $T(\square^3_{2,0})$. Observe that the 3-simplex 231 specifies a map $\Delta^3_1 \to B'$. So if we mark its first face 121, the resulting object B has the same precomplicial reflection as B'. We will adopt B (rather than B') as our

s	pushout of	interior	missing face
1	$\Lambda^2_1 \hookrightarrow \Delta^2_1$	211	111
2	$\Lambda^2_1 \hookrightarrow \Delta^2_1$	2+1	1+1
3	$\Lambda^3_2 \hookrightarrow \Delta^3_2$	312	212
4	$\Lambda^3_1 \hookrightarrow \Delta^3_1$	213	112
5	$\Lambda^3_2 \hookrightarrow \Delta^3_2$	123	122
6	$\tau_1\Lambda^3_2 \hookrightarrow \Delta^{3''}_2$	321	221
7	$\tau_1\Lambda^3_1 \hookrightarrow \Delta^{3''}_1$	231	121
8	$\Lambda^3_3 \hookrightarrow \Delta^3_3$	132	1+2

Table 1: Inclusions $A^{s-1} \hookrightarrow A^s$.

"model" for $T(\square^3_{2,0})$. For the open box, we define A to be the regular subset of B (or B') consisting of those simplices ϕ such that:

- $\phi(1) \in \{\pm\infty\}$,
- $\phi(2) = -\infty$, or
- $\phi(3) \in \{\pm\infty\}$,

so that the precomplicial reflection A^{pre} is $T(\sqcap^3_{2,0})$. Then we have a sequence of inclusions

$$A = A^0 \hookrightarrow A^1 \hookrightarrow \cdots \hookrightarrow A^8 = B,$$

where $A^{s-1} \hookrightarrow A^s$ is the pushout of a suitable trivial cofibration as indicated in Table 1. This table is to be interpreted as saying, for example, that the inclusion $A^0 \hookrightarrow A^1$ fits into the pushout square

$$\begin{tikzcd} \Lambda^2_1 \arrow[r] \arrow[d, hook] & A^0 \arrow[d, hook] \\ \Delta^2_1 \arrow[r] & A^1 \end{tikzcd}$$

in sSet^+, where the composite $\Delta^2_1 \to A^1 \hookrightarrow B$ corresponds to the simplex $\phi = 211$, and the face $\phi\partial_1$ corresponding to the missing face in the horn is 111. One can check that every nondegenerate face in $B \setminus A$ appears exactly once in Table 1, and moreover it is marked if and only if it appears either in the "interior" column or in the sixth or seventh row. It is also straightforward to verify using Proposition 5.6 that, for each $1 \le s \le 8$, the marked simplicial set A^{s-1} indeed contains enough (marked) simplices to support a map from the domain in the "pushout of" column. By reflecting everything to $\mathsf{PreComp}$, we can deduce that T sends the box inclusion $\sqcap^3_{2,0} \hookrightarrow \square^3_{2,0}$ to a trivial cofibration. The case $\sqcap^3_{2,1} \hookrightarrow \square^3_{2,1}$ is dual, and the other 3-dimensional boxes can be treated using Proposition 3.10.

It remains to prove that $T(\square^n_{k,\varepsilon}{}' \hookrightarrow \square^n_{k,\varepsilon}{}'')$ is a trivial cofibration for any n, k, ε. We show that this map is in fact invertible.

Consider the following commutative diagram in $\mathsf{PreComp}$:

$$\begin{tikzcd}
\Delta^{n-1} \arrow[r, "\iota_{n-1}"] \arrow[d, hook] & T(\square^{n-1}) \arrow[r, "T(\partial_{k,\varepsilon})"] \arrow[d, hook] & T(\square^n_{k,\varepsilon}{}') \arrow[d, hook] \\
\widetilde{\Delta}^{n-1} \arrow[r, "\iota_{n-1}"] & T(\widetilde{\square}^{n-1}) \arrow[r, "T(\partial_{k,\varepsilon})"] & T(\square^n_{k,\varepsilon}{}'')
\end{tikzcd}$$

The left square is a pushout by the definition of T, and the right square is a pushout because it is the image of a pushout square under T. Thus, to show that $T(\square^n_{k,\varepsilon}{}' \hookrightarrow \square^n_{k,\varepsilon}{}'')$ is invertible, it suffices to prove that the $(n-1)$-simplex in $T(\square^n_{k,\varepsilon}{}')$ corresponding to the top row above is marked. Note that, by unwinding the above argument, one can express $T(\sqcap^n_{k,\varepsilon} \hookrightarrow \square^n_{k,\varepsilon})$ as a composite

$$T(\sqcap^n_{k,\varepsilon}) = X^0 \hookrightarrow X^1 \hookrightarrow \cdots \hookrightarrow X^N = T(\square^n_{k,\varepsilon}),$$

where each map $X^{s-1} \hookrightarrow X^s$ is a pushout (in $\mathsf{PreComp}$) of the precomplicial reflection of a complicial horn inclusion. We show by induction on s that all $(n-1)$-simplices contained in X^s are marked in $T\big(\square^n_{k,\varepsilon}{}'\big)$.

For the base case, write $\square^n_{k,\varepsilon}{}'$ as a pushout

$$\begin{tikzcd}
\coprod \square^m \arrow[r] \arrow[d] & \square^n \arrow[d] \\
\coprod \widetilde{\square}^m \arrow[r] & \square^n_{k,\varepsilon}{}'
\end{tikzcd}$$

where the coproducts are taken over all marked faces of $\square^n_{k,\varepsilon}{}'$, which in particular include all faces $\partial_{\ell,\eta}$ of codimension 1 with $(\ell,\eta) \neq (k,\varepsilon)$. By applying T to this pushout square, we can deduce that any $(n-1)$-simplex of the form

$$\Delta^{n-1} \xrightarrow{\iota_{n-1}} (\Delta^1)^{\otimes(n-1)} = T(\square^{n-1}) \xrightarrow{T(\partial_{\ell,\eta})} T(\square^n)$$

is marked in $T(\square^n_{k,\varepsilon}{}')$. By combining this observation with Proposition 5.6, one can deduce that any $(n-1)$-simplex contained in X^0 is marked in $T(\square^n_{k,\varepsilon}{}')$. For the inductive step, suppose that all $(n-1)$-simplices contained in X^{s-1} are marked in $T(\square^n_{k,\varepsilon}{}')$. Suppose further that X^s contains a nondegenerate $(n-1)$-simplex ϕ that X^{s-1} does not contain (for otherwise we are done). Then $X^{s-1} \hookrightarrow X^s$ fits into either a pushout square of the form

$$\begin{tikzcd}
(\Lambda^{n-1}_\ell)^{\mathrm{pre}} \arrow[r] \arrow[d, hook] & X^{s-1} \arrow[d, hook] \\
(\Delta^{n-1}_\ell)^{\mathrm{pre}} \arrow[r, "\phi"] & X^s
\end{tikzcd}$$

or one of the form

$$\begin{tikzcd} (\Lambda^n_\ell)^{\mathrm{pre}} \arrow[r] \arrow[d, hook] & X^{s-1} \arrow[d, hook] \\ (\Delta^n_\ell)^{\mathrm{pre}} \arrow[r, "\chi"] & X^s \end{tikzcd}$$

with $\chi\partial_\ell = \phi$. In the former case, ϕ is marked in X^s and hence in $T(\Box^n_{k,\varepsilon}{}')$ since the unique nondegenerate $(n-1)$-simplex in Δ^{n-1}_ℓ is marked. In the latter case, the inductive hypothesis implies that χ extends to the marked simplicial set $\Delta^n_\ell{}'$. Since X^s is a precomplicial set, it follows that $\phi = \chi\partial_\ell$ is marked in X^s and hence in $T(\Box^n_{k,\varepsilon}{}')$. This completes the proof. $\Box$

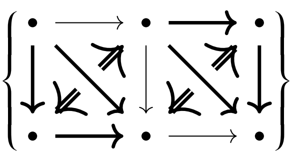
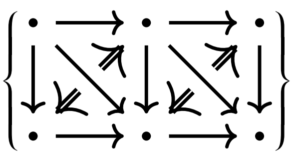
The saturated and n-trivial versions can be proved analogously.

Theorem 7.2 *The adjunction $T \dashv U$ is a Quillen adjunction when* cSet^+ *and* $\mathsf{PreComp}$ *are respectively equipped with*:

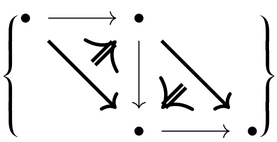
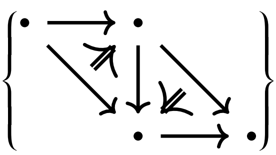
- *the saturated comical model structure and the saturated complicial model structure*,
- *the n-trivial comical model structure and the n-trivial complicial model structure for some $0 \le n < \infty$, or*
- *the saturated n-trivial comical model structure and the saturated n-trivial complicial model structure for some $0 \le n < \infty$.*

Proof The proof is analogous to that of Theorem 7.1. For the n-trivial versions, observe that T sends the (cubical) m-marker to a pushout of the (simplicial) m-marker.

For the saturated versions, we only check that T sends the basic Rezk maps to trivial cofibrations. (That the higher Rezk maps are also sent to trivial cofibrations then follows from Theorems 1.37 and 6.5.) Note that T sends all four basic Rezk maps to the same map (up to isomorphism). This unique image, which we denote by $TL \hookrightarrow TL'$, may be visualized as

$TL =$  and $TL' =$ .

Let A (resp. A') be the regular subset of TL (resp. TL') consisting of the middle two nondegenerate 2-simplices so that they look like

$A =$  and $A' =$ .

Clearly $TL \hookrightarrow TL'$ is a pushout of its restriction $A \hookrightarrow A'$, so it suffices to show that the latter is a trivial cofibration.

Observe that A is isomorphic to the regular subset of Δ^3_{eq} consisting of ∂_0 and ∂_3. One can check that the inclusion $A \hookrightarrow \Delta^3_{\mathrm{eq}}$ may be written as the composite of a pushout of $\Lambda^2_1 \hookrightarrow \Delta^2_1$ (attaching ∂_1) and a pushout of $\Lambda^3_2 \hookrightarrow \Delta^3_2$. Hence $A \hookrightarrow \Delta^3_{\mathrm{eq}}$ is complicial. Similarly, $A' \hookrightarrow (\Delta^3)^\sharp$ is the composite of pushouts of two complicial horn inclusions and one elementary complicial marking extension (marking the 1-simplex $\{0,3\}$), so it is complicial too. Since the square

$$\begin{tikzcd} A \arrow[r, hook] \arrow[d, hook] & A' \arrow[d, hook] \\ \Delta^3_{\mathrm{eq}} \arrow[r, hook] & (\Delta^3)^\sharp \end{tikzcd}$$

commutes, the desired conclusion now follows by the 2-out-of-3 property. □

References

[Aitchison 1986] **I R Aitchison**, *The geometry of oriented cubes*, research report 86–0082, Macquarie University (1986) arXiv 1008.1714

[Al-Agl et al. 2002] **F A Al-Agl**, **R Brown**, **R Steiner**, *Multiple categories: the equivalence of a globular and a cubical approach*, Adv. Math. 170 (2002) 71–118 MR Zbl

[Ara 2014] **D Ara**, *Higher quasi-categories vs higher Rezk spaces*, J. K-Theory 14 (2014) 701–749 MR Zbl

[Barwick 2005] **C Barwick**, (∞, n)-*Cat as a closed model category*, PhD thesis, University of Pennsylvania (2005) Available at `https://www.proquest.com/docview/305445747`

[Cisinski 2006] **D-C Cisinski**, *Les préfaisceaux comme modèles des types d'homotopie*, Astérisque 308, Soc. Math. France, Paris (2006) MR Zbl

[Crans 1995] **S E Crans**, *On combinatorial models for higher-dimensional homotopies*, PhD thesis, Utrecht University (1995)

[Doherty et al. 2023] **B Doherty**, **K Kapulkin**, **Y Maehara**, *Equivalence of cubical and simplicial approaches to (∞,n)-categories*, Adv. Math. 416 (2023) art. id. 108902 MR Zbl

[Doherty et al. 2024] **B Doherty**, **K Kapulkin**, **Z Lindsey**, **C Sattler**, *Cubical models of $(\infty, 1)$-categories*, Mem. Amer. Math. Soc. 1484, Amer. Math. Soc., Providence, RI (2024) MR Zbl

[Gaitsgory and Rozenblyum 2017a] **D Gaitsgory**, **N Rozenblyum**, *A study in derived algebraic geometry, I: Correspondences and duality*, Math. Surv. Monogr. 221, Amer. Math. Soc., Providence, RI (2017) MR Zbl

[Gaitsgory and Rozenblyum 2017b] **D Gaitsgory**, **N Rozenblyum**, *A study in derived algebraic geometry, II: Deformations, Lie theory and formal geometry*, Math. Surv. Monogr. 221, Amer. Math. Soc., Providence, RI (2017) MR Zbl

[Grandis and Mauri 2003] **M Grandis**, **L Mauri**, *Cubical sets and their site*, Theory Appl. Categ. 11 (2003) 185–211 MR Zbl

[Gray 1974] **J W Gray**, *Formal category theory: adjointness for 2-categories*, Lecture Notes in Math. 391, Springer (1974) MR Zbl

[Henry 2020] **S Henry**, *Weak model categories in classical and constructive mathematics*, Theory Appl. Categ. 35 (2020) art. id. 24 MR Zbl

[Jardine 2006] **J F Jardine**, *Categorical homotopy theory*, Homology Homotopy Appl. 8 (2006) 71–144 MR Zbl

[Joyal and Tierney 2007] **A Joyal**, **M Tierney**, *Quasi-categories vs Segal spaces*, from "Categories in algebra, geometry and mathematical physics", Contemp. Math. 431, Amer. Math. Soc., Providence, RI (2007) 277–326 MR Zbl

[Kapulkin and Voevodsky 2020] **K Kapulkin**, **V Voevodsky**, *A cubical approach to straightening*, J. Topol. 13 (2020) 1682–1700 MR Zbl

[Kapulkin et al. 2019] **K Kapulkin**, **Z Lindsey**, **L Z Wong**, *A co-reflection of cubical sets into simplicial sets with applications to model structures*, New York J. Math. 25 (2019) 627–641 MR Zbl

[Loubaton 2022] **F Loubaton**, *The complicial model of (∞, ω)-categories*, preprint (2022) arXiv 2207.08504

[Lurie 2009] **J Lurie**, *On the classification of topological field theories*, from "Current developments in mathematics, 2008", International, Somerville, MA (2009) 129–280 MR Zbl

[Maehara 2021] **Y Maehara**, *The Gray tensor product for 2-quasi-categories*, Adv. Math. 377 (2021) art. id. 107461 MR Zbl

[Maltsiniotis 2009] **G Maltsiniotis**, *La catégorie cubique avec connexions est une catégorie test stricte*, Homology Homotopy Appl. 11 (2009) 309–326 MR Zbl

[Olschok 2009] **M Olschok**, *On constructions of left determined model structures*, PhD thesis, Masaryk University (2009) Available at `https://is.muni.cz/th/xhch5/diss.pdf`

[Ozornova and Rovelli 2020] **V Ozornova**, **M Rovelli**, *Model structures for (∞, n)-categories on (pre)stratified simplicial sets and prestratified simplicial spaces*, Algebr. Geom. Topol. 20 (2020) 1543–1600 MR Zbl

[Rezk 2010] **C Rezk**, *A Cartesian presentation of weak n-categories*, Geom. Topol. 14 (2010) 521–571 MR Zbl

[Riehl and Verity 2020] **E Riehl**, **D Verity**, *(More) elements of ∞-category theory* (2020) Original draft (with additional material) of "Elements of ∞-category theory", Cambridge Univ. Presss (2022) Available at `https://elements-book.github.io/more-elements.pdf`

[Simpson 2012] **C Simpson**, *Homotopy theory of higher categories*, New Math. Monogr. 19, Cambridge Univ. Press (2012) MR Zbl

[Steiner 1993] **R Steiner**, *The algebra of directed complexes*, Appl. Categ. Structures 1 (1993) 247–284 MR Zbl

[Steiner 2006] **R Steiner**, *Thin fillers in the cubical nerves of omega-categories*, Theory Appl. Categ. 16 (2006) art. id. 8 MR Zbl

[Street 1987] **R Street**, *The algebra of oriented simplexes*, J. Pure Appl. Algebra 49 (1987) 283–335 MR Zbl

[Street 1991] **R Street**, *Parity complexes*, Cahiers Topologie Géom. Différentielle Catég. 32 (1991) 315–343 MR Zbl

[Tonks 1992] **A P Tonks**, *Cubical groups which are Kan*, J. Pure Appl. Algebra 81 (1992) 83–87 MR Zbl

[Verity 2007] **D Verity**, *Weak complicial sets, II: Nerves of complicial Gray-categories*, from "Categories in algebra, geometry and mathematical physics", Contemp. Math. 431, Amer. Math. Soc., Providence, RI (2007) 441–467 MR Zbl

[Verity 2008a] **D Verity**, *Complicial sets characterising the simplicial nerves of strict ω-categories*, Mem. Amer. Math. Soc. 905, Amer. Math. Soc., Providence, RI (2008) MR Zbl

[Verity 2008b] **D Verity**, *Weak complicial sets, I: Basic homotopy theory*, Adv. Math. 219 (2008) 1081–1149 MR Zbl

Department of Mathematics, Johns Hopkins University
Baltimore, MD, United States

Department of Mathematics, University of Western Ontario
London, ON, Canada

Research Institute for Mathematical Sciences, Kyoto University
Kyoto, Japan

tcampio1@jh.edu, kkapulki@uwo.ca, ymaehar@kurims.kyoto-u.ac.jp

Proposed: Mark Behrens
Seconded: Leonid Polterovich, Ulrike Tillmann

Received: 1 June 2020
Revised: 28 February 2024

Geometry & Topology **29**:3 (2025) 1171–1235
DOI: 10.2140/gt.2025.29.1171
Published: 31 May 2025

msp

Rank-one Hilbert geometries

MITUL ISLAM

We develop a notion of rank-one properly convex domains (or Hilbert geometries) in real projective space. This is in the spirit of rank-one nonpositively curved Riemannian manifolds and CAT(0) spaces. We define rank-one isometries for Hilbert geometries and characterize them as being equivalent to contracting elements (in the sense of geometric group theory). We prove that if a discrete subgroup of automorphisms of a Hilbert geometry contains a rank-one isometry, then the subgroup is either virtually cyclic or acylindrically hyperbolic. This leads to several applications like infinite dimensionality of the space of quasimorphisms, counting results for conjugacy classes and genericity results for rank-one isometries.

20F65, 20F67, 53A20; 53C60, 57N16

1 Introduction

A *properly convex domain* in $\mathbb{P}(\mathbb{R}^{d+1})$ is an open subset $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ such that $\overline{\Omega}$ is a bounded convex domain in an affine chart. Any such domain Ω carries a canonical distance function d_Ω, called the *Hilbert metric* on Ω, defined using projective cross-ratios; see Section 3. Then Ω equipped with its Hilbert metric constitutes a *Hilbert geometry*. A motivating example is given by the open projective disk $\Omega_2 := \{[x : y : 1] \in \mathbb{P}(\mathbb{R}^3) \mid x^2 + y^2 < 1\}$, a properly convex domain in $\mathbb{P}(\mathbb{R}^3)$. In fact, $(\Omega_2, \mathrm{d}_{\Omega_2})$ is the projective model of the 2-dimensional real hyperbolic space $\mathbb{H}^2$.

For a properly convex domain Ω, the group $\mathrm{Aut}(\Omega) := \{g \in \mathrm{PGL}_{d+1}(\mathbb{R}) \mid g\Omega = \Omega\}$ acts properly and isometrically on $(\Omega, \mathrm{d}_\Omega)$. If $\Gamma \leq \mathrm{Aut}(\Omega)$ is a discrete subgroup, then the quotient space Ω/Γ is "locally modeled" on $(\Omega, \mathrm{d}_\Omega)$. These are the main objects that we study in this paper. We make the following definition.

Definition 1.1 We say that Ω is a *Hilbert geometry* if $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is a properly convex domain. Further, we say that a pair (Ω, Γ) is a *Hilbert geometry* if $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is a properly convex domain and $\Gamma \leq \mathrm{Aut}(\Omega)$ is a discrete subgroup. A Hilbert geometry (Ω, Γ) is *divisible* if $\Gamma \leq \mathrm{Aut}(\Omega)$ acts cocompactly on Ω.

Example Consider the projective model Ω_2 of $\mathbb{H}^2$. Here $\mathrm{Aut}(\Omega_2) = \mathrm{PO}(2,1)$. If $\Gamma \leq \mathrm{PO}(2,1)$ is any discrete subgroup, then (Ω_2, Γ) is a Hilbert geometry and (Ω_2, Γ) is divisible when Γ is a uniform lattice.

The boundary of a Hilbert geometry Ω, denoted by $\partial\Omega$, is the topological boundary of Ω as a subset of $\mathbb{P}(\mathbb{R}^{d+1})$. The regularity of $\partial\Omega$ strongly influences the geometric properties of $(\Omega, \mathrm{d}_\Omega)$. For instance, consider the class of *strictly convex* Hilbert geometries, ie Hilbert geometries Ω such that $\partial\Omega$ does not contain any nontrivial projective line segments. Benoist [8] showed that strictly convex divisible Hilbert geometries (Ω, Γ) have C^1 boundaries and behave like compact Riemannian manifolds of negative curvature (more precisely, Γ is Gromov hyperbolic and the geodesic flow is Anosov). This analogy between strictly convex Hilbert geometries and Riemannian negative curvature was subsequently studied by many authors with much success; see Benoist [12] or Marquis [43] for a survey.

On the other hand, the *nonstrictly convex* Hilbert geometries (ie when $\partial\Omega$ contains nontrivial projective line segments) have remained elusive. There are only a few examples (see Section 3.4) and, until recently, only a limited number of results. Taking a cue from the strictly convex case, one hopes to liken nonstrictly convex Hilbert geometries to Riemannian nonpositive curvature, or more generally, CAT(0) spaces. This will be our guiding principle in this paper. But we remark that the similarity with CAT(0) geometry is superficial. In fact, an old theorem of Kelly and Straus [41] states that Ω is CAT(0) if and only if Ω is the projective model of the real hyperbolic space. Thus, one needs to use very different tools and techniques for working with Hilbert geometries as compared to CAT(0) spaces.

Our target in this paper is to classify Hilbert geometries into two broad classes: "rank one" and "higher rank". The motivation for this classification comes from the success of the rank rigidity theorem for

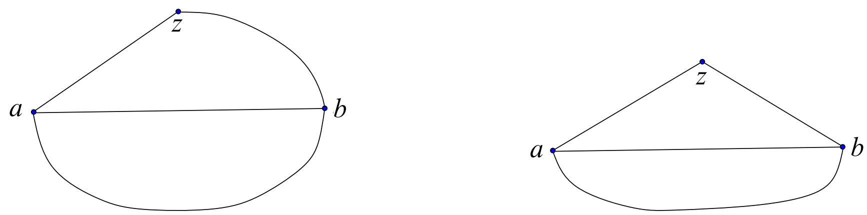


Figure 1: In the left figure, (a,b) is a *rank-one geodesic* while in the right figure, (a,b) is contained in a *half triangle in* Ω. However, Proposition 6.5 will show that neither of these can be a closed rank-one geodesic (ie a *rank-one axis*) in Ω/Γ.

nonpositively curved Riemannian manifolds; see Ballmann [3] and Burns and Spatzier [22]. Roughly, this theorem states that there is a dichotomy for irreducible compact Riemannian manifolds of nonpositive curvature: either the manifold is "rank one", or it is a higher rank Riemannian locally symmetric space. Similar rank rigidity theorems have been proven in other "nonpositive curvature" settings (see Caprace and Sageev [23] and Ricks [47]) and conjectured for CAT(0) spaces. We remark that the usual definition of rank for Riemannian manifolds uses Jacobi fields and will not be useful for Hilbert geometries. This is because the geodesic flow on a generic nonstrictly convex Hilbert geometry is only C^0.

We introduce a notion of rank-one geodesics in $(\Omega, \mathrm{d}_\Omega)$ using projective geometry. Consider an open projective line segment $(a,b) \subset \Omega$ with $a, b \in \partial\Omega$. Then (a,b) is a bi-infinite geodesic for the Hilbert metric d_Ω. We will say that (a,b) is a *rank-one geodesic* provided it is not contained in a *half triangle in* Ω, ie either $(a,c) \subset \Omega$ or $(c,b) \subset \Omega$ for any $c \in \partial\Omega$; see Figure 1 and Definitions 6.1 and 6.2. The notion of a half triangle in Hilbert geometry is analogous to the notion of a half flat in CAT(0) geometry; see Ballmann [4, Section III.3]. Our above definition of a rank-one geodesic is motivated by an analogous characterization of rank-one geodesics in CAT(0) geometry. In a CAT(0) geodesic metric space, a rank-one geodesic does not bound a half flat.

We will say that an isometry $\gamma \in \mathrm{Aut}(\Omega)$ is a *rank-one isometry* if γ acts by a translation along a rank-one geodesic $\ell \subset \Omega$; see Definition 6.3. We remark that acting by a translation along a rank-one geodesic (ie having a *rank-one axis*) is much more special than simply translating along any projective geodesic (ie having an *axis*); see Remark 6.4. Our definition of rank-one isometry is again analogous to a characterization of rank-one isometries in CAT(0) geometry; see Ballmann [2] and Ballmann and Brin [5]. A rank-one isometry $\gamma \in \mathrm{Aut}(\Omega)$ has several properties reminiscent of hyperbolic isometries in $\mathrm{Isom}(\mathbb{H}^2)$: γ is biproximal, has exactly two fixed points $\gamma^\pm$ in $\overline{\Omega}$, has a unique axis $(\gamma^+, \gamma^-) \subset \Omega$, both fixed points are "visible" (ie $(\gamma^+, z) \cup (z, \gamma^-)$ for any $z \in \partial\Omega - \{\gamma^+, \gamma^-\}$), and γ has the so-called north–south dynamics on $\partial\Omega$; see Proposition 6.5 and Corollary 6.7. In the case where (Ω, Γ) is divisible, it is quite easy to detect a rank-one isometry: if $\gamma \in \mathrm{Aut}(\Omega)$ has an axis and is biproximal, then γ is a rank-one isometry; see Proposition 6.8.

We further the analogy between rank-one isometries in Hilbert geometry and CAT(0) geometry by proving that rank-one isometries (in our sense above) are contracting elements in the sense of Sisto [49]. Sisto introduced the notion of contracting elements to capture the essence of "negative curvature" in groups; see Section 9. He proved in [49, Proposition 3.14] that if Λ acts properly by isometries on a proper CAT(0) space X, then an element of Λ is contracting if and only if it is rank one (in the sense of CAT(0) geometry). Our first main result in the paper is an analogue of this result for Hilbert geometries. If Ω is a Hilbert geometry, let $\mathcal{PS}^{\Omega} := \{[x, y] \mid x, y \in \Omega\}$, where $[x, y]$ is a projective line segment joining x and y.

Theorem 1.2 (see Part III) *If Ω is a Hilbert geometry, then $\gamma \in \mathrm{Aut}(\Omega)$ is a contracting element for $(\Omega, \mathcal{PS}^{\Omega})$ if and only if γ is a rank-one isometry.*

In the light of these analogies, one naturally expects that the presence of many rank-one isometries would induce interesting "negative curvature"-like behavior. To formalize this, we now introduce the notion of rank-one Hilbert geometries. An example to keep in mind is (Ω_2, Γ) where $\Omega_2 \subset \mathbb{P}(\mathbb{R}^3)$ is the projective model of $\mathbb{H}^2$ and $\Gamma \leq \mathrm{PO}(2,1)$ is an infinite discrete subgroup.

Definition 1.3 A *rank-one Hilbert geometry* is a pair (Ω, Γ) where $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is a Hilbert geometry and Γ is a discrete subgroup of $\mathrm{Aut}(\Omega)$ that contains a rank-one isometry.

Morally, if a rank-one group Γ as in Definition 1.3 is not virtually cyclic (ie does not contain a finite-index cyclic subgroup), then it contains many rank-one isometries and we expect the group Γ to appear quite "hyperbolic". But of course we cannot expect such a group Γ to always be Gromov hyperbolic — there are many examples to the contrary; see Section 3.4. The main result of this paper is to identify the notion of hyperbolicity that rank-one groups satisfy. We prove that a rank-one group is either virtually cyclic or an acylindrically hyperbolic group; see Theorem 1.4 below.

The notion of acylindrically hyperbolic groups, introduced by Osin in [45], is a generalization of the notion of nonelementary Gromov hyperbolic groups. Roughly speaking, a group is acylindrically hyperbolic if it admits a nonelementary action on a (possibly nonproper) Gromov hyperbolic metric space with all but finitely many elements acting "hyperbolically"; see Definition 12.1. This family includes many important classes of groups: mapping class groups of most finite-type surfaces, rank-one CAT(0) groups that are not virtually abelian, relatively hyperbolic groups that are not virtually cyclic and have proper peripheral subgroups, and outer automorphism groups of free groups on at least two generators; see [45, Appendix]. We prove the following.

Theorem 1.4 (see Section 12) *If (Ω, Γ) is a rank-one Hilbert geometry, then either Γ is virtually cyclic or Γ is an acylindrically hyperbolic group.*

The acylindrical hyperbolicity of rank-one Hilbert geometries (Ω, Γ) have several applications. We defer this discussion until the section Applications below. Instead we first mention some of the precursors to our above result. In all those previous results, the conclusion is that the group under consideration is either Gromov hyperbolic or relatively hyperbolic — both of which are acylindrically hyperbolic groups; see the previous paragraph. Benoist [8] showed that if (Ω, Γ) is divisible and $\partial\Omega$ is strictly convex, then

Γ is Gromov hyperbolic. If instead Ω/Γ is noncompact but has finite volume, then Cooper, Long and Tillmann [25, Theorem 0.15] showed that Γ is relatively hyperbolic (with respect to the cusp subgroups). More generally, if Ω/Γ is geometrically finite, then Crampon and Marquis [27, Theorem 1.8] proved that Γ is relatively hyperbolic. However, they require that $\partial\Omega$ is C^1, and not just strictly convex.

Outside the strictly convex setting, Islam and Zimmer [39] have recently shown that if Γ acts cocompactly on Ω, then Γ is relatively hyperbolic (with peripheral subgroups free abelian of rank at least two) if and only if the set of properly embedded simplices in Ω of dimension at least two (see Section 3.3) forms an "isolated family". The results in [39] hold in greater generality — whenever Γ acts convex cocompactly on Ω; see Definition A.1 and Section A.6 for further discussion. We can interpret the above Theorem 1.4 as a generalization of these aforementioned results in the general setup of (possibly nonstrictly convex) Hilbert geometries. Theorem 1.4 characterizes the existence of rank-one isometries in Γ as a key factor that underpins the presence of these various weak forms of hyperbolicity for the group Γ.

Zariski density and rank one

Before moving on to contrasting rank one against "higher rank" Hilbert geometries, we indulge in a short discussion about Hilbert geometries (Ω, Γ) where $\Gamma \le \mathrm{Aut}(\Omega)$ is Zariski dense in $\mathrm{SL}_{d+1}(\mathbb{R})$, ie it is "large" in an algebraic sense. For such groups Γ, one can define a notion of proximal limit set $\Lambda_\Gamma^{G/Q} \subset \mathbb{P}(\mathbb{R}^{d+1})$ (see Definition 8.3 and Remark 8.4) that is independent of the properly convex domain Ω. In Section 8, we prove that if $x, y \in \Lambda_\Gamma^{G/Q} \cap \partial\Omega$ are such that $(x, y) \subset \Omega$ is a rank-one geodesic, then the set of rank-one isometries in Γ form a Zariski dense set in $\mathrm{SL}_{d+1}(\mathbb{R})$ and (x, y) can be approximated by the axes of rank-one isometries (Lemma 8.10). In particular, a Hilbert geometry (Ω, Γ) with Γ Zariski dense in $\mathrm{SL}_{d+1}(\mathbb{R})$ is rank one if and only if Ω contains a rank-one geodesic (x, y) with $x, y \in \Lambda_\Gamma^{G/Q} \cap \partial\Omega$; see the answer to Question 8.1.

Rank one versus higher rank

The class of rank-one Hilbert geometries is quite rich. Besides the strictly convex Hilbert geometries, there are several examples of nonstrictly convex divisible Hilbert geometries which are rank one, eg the 3-manifold groups constructed in Benoist [10] from projective reflection groups. For more examples, see Section 3.4. In Appendix A, we discuss this further and also generalize the notion of rank one for convex cocompact actions.

On the other hand, there are several examples of Hilbert geometries that are not rank one, or in other words, have "higher rank". Projective simplices of dimension at least two and symmetric domains of real rank at least two are the key examples of "higher rank" Hilbert geometries; see Sections 3.3 and 3.4. The former are examples of reducible "higher rank" while the latter are examples of irreducible "higher rank" domains; see Definition 3.6. At this point, it is natural to ask whether these are all the "higher rank" divisible Hilbert geometries, akin to the case of Riemannian nonpositive curvature. Recently, A Zimmer [50] has proven that this is indeed the case.

We will now briefly discuss Zimmer's result for context. Zimmer calls Ω a *higher rank Hilbert geometry* if any $(p,q) \subset \Omega$ is contained in a properly embedded projective simplex S in Ω of dimension at least two; see also Section 3.4. Under some assumptions, he proves that his notion of higher rank is exactly complementary to our notion of rank one. We remark that Zimmer does not develop a theory of rank-one geometries. He focuses only on higher rank geometries and proves that an irreducible divisible Hilbert geometry (Ω,Γ) is higher rank (in his sense) if and only if it does not satisfy the notion of rank one (in the sense introduced in this paper).

Theorem 1.5 (part of [50, Theorem 1.4]) *Suppose* (Ω,Γ) *is a divisible Hilbert geometry and* Ω *is irreducible. Then the following are equivalent*:

- (i) Ω *has higher rank* (*in the sense of Zimmer* [50, Definition 1.1]).
- (ii) Ω *is a symmetric domain of real rank at least two.*
- (iii) Ω *does not contain any rank-one geodesics* (*in the sense of this paper*, *Definition 6.2*).
- (iv) Γ *does not contain any rank-one isometries* (*in the sense of this paper*, *Definition 6.3*).

Applications

We now return to our discussion about rank-one geometries. There is a sizeable literature exploring different properties of acylindrically hyperbolic groups. By virtue of Theorem 1.4, we can use these to establish several interesting results about rank-one Hilbert geometries. We remark that in the ensuing discussion, we usually do not require the additional assumption of divisibility.

1.1 Second bounded cohomology and quasimorphisms

A quasimorphism of a group G is a function $f: G \to \mathbb{R}$ such that $\sup_{g,h\in G} |f(gh)-f(g)-f(h)|$ is finite. We say that two quasimorphisms are equivalent if they differ by a bounded function or a homomorphism of G into $\mathbb{R}$. The set of all equivalence classes of quasimorphisms of G constitute $\widetilde{\mathrm{QH}}(G)$, which is a $\mathbb{R}$-vector space. More generally, if $\rho\colon G \to \mathcal{U}(E)$ is a unitary representation of G on a complete normed $\mathbb{R}$-vector space $(E, \|\cdot\|)$, then we can define $\widetilde{\mathrm{QC}}(G;\rho)$; see Section 13.

Bestvina and Fujiwara proved in [17] that if M is a compact nonpositively curved Riemannian manifold, then — under some mild assumptions — $\widetilde{\mathrm{QH}}(\pi_1(M))$ is infinite-dimensional if and only if M is a rank-one Riemannian manifold. We prove a similar cohomological characterization of rank-one Hilbert geometries.

Theorem 1.6 (see Section 13) *If* (Ω,Γ) *is a rank-one Hilbert geometry*, Γ *is torsion-free and* Γ *is not virtually cyclic*, *then*

- (i) $\dim(\widetilde{\mathrm{QH}}(\Gamma)) = \infty$, *and*
- (ii) *if* $p \in (1,\infty)$ *and* $\rho^p_{\mathrm{reg}}\colon \Gamma \to \mathcal{U}(\ell^p(\Gamma))$ *is the regular representation*, *then* $\dim(\widetilde{\mathrm{QC}}(\Gamma;\rho^p_{\mathrm{reg}})) = \infty$.

We prove a more general Theorem 13.1. On the other hand, if $\Gamma \leq G$ is a lattice in a higher-rank simple Lie group G, then a result of Burger and Monod [21, Theorem 21] implies that $\widetilde{\mathrm{QH}}(\Gamma) = 0$. Then Theorem 1.6 and the rank rigidity result (Theorem 1.5) implies:

Corollary 1.7 (see Section 13) *If (Ω, Γ) is a divisible Hilbert geometry and Ω is irreducible, then $\dim(\widetilde{\mathrm{QH}}(\Gamma)) = \infty$ if and only if (Ω, Γ) is a rank-one Hilbert geometry. Otherwise $\dim(\widetilde{\mathrm{QH}}(\Gamma)) = 0$.*

1.2 Counting of conjugacy classes

For $g \in \mathrm{Aut}(\Omega)$, define the translation length $\tau_\Omega(g) := \inf_{x\in\Omega} \mathrm{d}_\Omega(x, gx)$ (see also Section 3.8) and the stable translation length

$$\tau_\Omega^{\mathrm{stable}}(g) := \lim_{n\to\infty} \frac{\mathrm{d}_\Omega(x, g^n x)}{n}.$$

Note that $\tau_\Omega^{\mathrm{stable}}(g)$ is independent of the basepoint $x \in \Omega$. Now suppose that (Ω, Γ) is a rank-one Hilbert geometry. For $g \in \Gamma$, let $[c_g]$ denote the conjugacy class of g in Γ. Both τ_Ω and $\tau_\Omega^{\mathrm{stable}}$ are well-defined on the set of conjugacy classes in Γ. Then for $t > 0$, define

$$\mathscr{C}(t) := \#\{[c_g] \mid g \in \Gamma,\ \tau_\Omega([c_g]) \leq t\} \quad \text{and} \quad \mathscr{C}^{\mathrm{stable}}(t) := \#\{[c_g] \mid g \in \Gamma, \tau_\Omega^{\mathrm{stable}}([c_g]) \leq t\}.$$

Here $\mathscr{C}(t)$ (resp. $\mathscr{C}^{\mathrm{stable}}(t)$) counts the number of conjugacy classes in Γ whose translation length (resp. stable translation length) is at most t. For divisible rank-one Hilbert geometries, we prove an asymptotic growth formula for $\mathscr{C}(t)$ and $\mathscr{C}^{\mathrm{stable}}(t)$. To state our result, we will require the critical exponent of Γ, which is defined as

$$\omega_\Gamma := \limsup_{n\to\infty} \frac{\log \#\{g \in \Gamma \mid \mathrm{d}_\Omega(x, gx) \leq n\}}{n}$$

for some (and hence any) basepoint $x \in \Omega$.

Theorem 1.8 (see Section 14) *Suppose (Ω, Γ) is a divisible rank-one Hilbert geometry and Γ is not virtually cyclic. Then there exists a constant D' such that if $t \geq 1$,*

$$\frac{1}{D'} \frac{\exp(t\omega_\Gamma)}{t} \ \leq\ \mathscr{C}(t)\ \leq\ D' \frac{\exp(t\omega_\Gamma)}{t}.$$

The function $\mathscr{C}^{\mathrm{stable}}(t)$ also satisfies a similar growth formula as above.

Remark 1.9 (i) An element $g \in \Gamma$ is called primitive if $g \neq h^n$ for any $h \in \Gamma$ and $|n| \geq 2$. If $\mathscr{C}_{\mathrm{Prim}}(t)$ is the number of conjugacy classes $[c_g]$ of primitive elements in Γ such that $\tau_\Omega([c_g]) \leq t$, then $\mathscr{C}_{\mathrm{Prim}}(t)$ satisfies a similar growth formula as $\mathscr{C}(t)$.

(ii) In [18, Proposition 1.5], Blayac establishes finer counting results for (a related notion of) rank-one Hilbert geometries using very different techniques.

Counting of conjugacy classes is often connected to counting of closed geodesics. However, this connection is subtle for Hilbert geometries, since there could be elements in Γ that do not act by a translation along any projective line in Ω (ie do not have an axis, see Example 5.11(B)).

1.3 Genericity and random walks

Suppose (Ω, Γ) is a rank-one Hilbert geometry, Γ is not virtually cyclic and Γ is finitely generated. If S is a finite symmetric generating set of Γ, let $W_n(S)$ be the set of words of length n in the elements of S. A simple random walk on Γ (with support S) is a sequence of Γ-valued random variables $\{X_n\}_{n\in\mathbb{N}}$ with laws μ_n defined by: if $n \geq 1$ and $g \in \Gamma$,

$$\mu_n(\{g\}) = \frac{\#\{w \in W_n(S) \mid w \text{ represents } g\}}{\#W_n(S)}.$$

Using results in [49], we show that rank-one isometries in Γ are exponentially generic from the viewpoint of simple random walks. This roughly means that the probability that a long word, written down by randomly choosing generators of Γ, is not a rank-one isometry is small. In particular, this probability decays exponentially in the length of the word.

Proposition 1.10 (see Section 15) *Suppose (Ω, Γ) is a rank-one Hilbert geometry, Γ is not virtually cyclic and Γ is finitely generated. Then the rank-one isometries in Γ are exponentially generic: if $\{X_n\}_{n\in\mathbb{N}}$ is a simple random walk on Γ, then there exists a constant $C \geq 1$ such that for all $n \geq 1$,*

$$\mathbb{P}[X_n \text{ is not a rank-one isometry}] \leq Ce^{-n/C}.$$

1.4 More consequences of acylindrical hyperbolicity

Proposition 1.11 (see Section 15) *If (Ω, Γ) is a rank-one Hilbert geometry and Γ is not virtually cyclic, then:*

(i) *Γ is SQ-universal, ie every countable group embeds in a quotient of Γ.*

(ii) *If Γ is the Baumslag–Solitar group $\mathrm{BS}(m,n)$, then $m = n = 0$ and Γ is the free group on two generators.*

1.5 Morse geodesics and Morse boundary

Roughly speaking, the Morse geodesics [26] in a geodesic metric space identify the "hyperbolic directions". As a corollary to Theorem 1.2, we prove that the axis of a rank-one isometry is a Morse geodesic.

Proposition 1.12 (see Section 15) *If Ω is a Hilbert geometry and $\gamma \in \mathrm{Aut}(\Omega)$ is a rank-one isometry, then the axis ℓ_γ of γ is $\mathcal{K}$-Morse for some Morse gauge $\mathcal{K}\colon [1,\infty) \times [0,\infty) \to [0,\infty)$, ie if α is a (λ, ε)-quasigeodesic with endpoints on ℓ_γ, then $\alpha \subset \mathcal{N}_{\mathcal{K}(\lambda,\varepsilon)}(\ell_\gamma)$.*

In [26], Cordes introduced a notion of Morse boundary for proper geodesic metric spaces. In the cases of proper CAT(0) spaces and hyperbolic metric spaces, the Morse boundary coincides with the contracting boundary and the Gromov boundary respectively. Theorem 1.2 and Proposition 1.12 implies that the Morse boundary $\partial_M \Omega$ of a rank-one Hilbert geometry (Ω, Γ) is nonempty. This inspires the following question (that we will not answer in this paper).

Question 1.13 *Describe the Morse boundary $\partial_M \Omega$ of a rank-one Hilbert geometry (Ω, Γ).*

Recent developments

Since this paper first appeared on arXiv, there have been tremendous new developments in the field. We mention a few of them. Blayac [18] has developed the Patterson–Sullivan theory for rank-one Hilbert geometries. In [20], Blayac and Viaggi constructed examples of divisible rank-one Hilbert geometries (Ω, Γ) in every dimension $d \geq 3$ where Γ is not Gromov hyperbolic. In their examples, Γ is Zariski dense in $\mathrm{SL}_{d+1}(\mathbb{R})$ and relatively hyperbolic with peripheral subgroups isomorphic to $\mathbb{Z} \times H$, where H is possibly nonabelian [20, Theorem 1.3].

Outline of the paper

We discuss the preliminaries in Part I. Section 4 of Part I is of particular interest as it discusses the geometry of ω-limit sets of automorphisms in $\mathrm{Aut}(\Omega)$. Part II develops the notion of rank-one Hilbert geometries. We define rank-one isometries and study their geometric properties in Sections 6 and 7. Our main tools here are the lemmas proven in Section 5. In Section 8, we study rank-one groups which are Zariski dense.

In Part III, we prove Theorem 1.2 (in Sections 10 and 11) and Theorem 1.4 (in Section 12). Part IV discusses several applications of our results, like computing the dimension of the space of quasimorphisms, counting of conjugacy classes and genericity from the viewpoint of random walks. We discuss generalizations, examples and nonexamples of rank-one Hilbert geometries in Appendix A. In Appendix B, we discuss the equivalence of two notions of contracting elements.

Acknowledgements

I would like to heartily thank my advisor Ralf Spatzier. This work has benefited greatly from my conversations with Andrew Zimmer, Harrison Bray, Samantha Pinella, Thang Nguyen, Ludovic Marquis and Beatrice Pozzetti. I also thank the anonymous reviewers for their comments, corrections and constructive criticism. I was partially supported by NSF grant 1607260 during the course of this work.

Part I Preliminaries

2 Notation

We set the following notation as standard for the rest of the paper.

(i) If $v \in \mathbb{R}^{d+1} \setminus \{0\}$, let $[v]$ or $\pi(v)$ denote its image in $\mathbb{P}(\mathbb{R}^{d+1})$. Conversely, if $u \in \mathbb{P}(\mathbb{R}^{d+1})$, we will use $\widetilde{u}$ to denote a lift of u (ie $\pi(\widetilde{u}) = u$).

(ii) If $A \in \mathrm{GL}_{d+1}(\mathbb{R})$, let $[A]$ denote its image in $\mathrm{PGL}_{d+1}(\mathbb{R})$, while $\widetilde{B} \in \mathrm{GL}_{d+1}(\mathbb{R})$ will denote a lift of $B \in \mathrm{PGL}_{d+1}(\mathbb{R})$.

(iii) If $W \leq \mathbb{R}^{d+1}$ is a nonzero linear subspace, $\mathbb{P}(W)$ denotes its projectivization.

(iv) If $g \in \mathrm{GL}_{d+1}(\mathbb{R})$, the eigenvalues of g (over $\mathbb{C}$) are denoted by $\lambda_1(g), \dots, \lambda_{d+1}(g)$. We index them in the nonincreasing order of their absolute values, ie $|\lambda_1(g)| \geq \cdots \geq |\lambda_{d+1}(g)|$. Let $\lambda_{\max}(g) := |\lambda_1(g)|$ and $\lambda_{\min}(g) := |\lambda_{d+1}(g)|$.

(v) If $g \in \mathrm{PGL}_{d+1}(\mathbb{R})$ and $1 \leq i \neq j \leq d+1$, define $\left|\frac{\lambda_i}{\lambda_j}(g)\right| := \left|\frac{\lambda_i(\tilde{g})}{\lambda_j(\tilde{g})}\right|$ for some (hence any) lift $\tilde{g} \in \mathrm{GL}_{d+1}(\mathbb{R})$ of g.

3 Hilbert geometries

3.1 Properly convex domains

An open set $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is called a *properly convex domain* if there exists a codimension one subspace $H \leq \mathbb{R}^{d+1}$ such that $\overline{\Omega}$ is a bounded (Euclidean) convex domain in the affine chart $\mathbb{A} := \mathbb{P}(\mathbb{R}^{d+1}) \setminus \mathbb{P}(H)$.

Remark 3.1 If $L \leq \mathbb{R}^{d+1}$ is a 2-dimensional linear subspace, then $\mathbb{P}(L) \not\subset \overline{\Omega}$ for any properly convex domain Ω. This elementary observation that a properly convex domain cannot contain an entire projective line will be used in many of the proofs in this paper.

For a nonempty set $X \subset \mathbb{R}^{d+1}$, let $\mathrm{Span}(X)$ denote the linear span of X. If $X' \subset \mathbb{P}(\mathbb{R}^{d+1})$ is nonempty, define

$$\mathrm{Span}(X') := \mathrm{Span}(\{\tilde{x} \in \mathbb{R}^{d+1} \mid \pi(\tilde{x}) \in X'\}).$$

Suppose Ω is a properly convex domain. If $x, y \in \overline{\Omega}$, let $[x, y]$ denote the unique closed connected subset of $\mathbb{P}(\mathrm{Span}\{x, y\}) \cap \overline{\Omega}$ that joins x and y. We will call $[x, y]$ *the projective line segment* between x and y (note that the notion of a projective line segment depends on Ω, but we assume that Ω will be clear from context and suppress it for brevity). We introduce the notation $(x, y) := [x, y] \setminus \{x, y\}$, $[x, y) := [x, y] \setminus \{y\}$ and $(x, y] := [x, y] \setminus \{x\}$. We will call (x, y) an *open projective line segment*. When we say that $[x, y]$ (or (x, y)) is *nontrivial*, we mean $x \neq y$.

We have a notion of convexity and convex hull in a properly convex domain Ω. A set $Y \subset \overline{\Omega}$ is *convex* if $[y_1, y_2] \subset Y$ for all $y_1, y_2 \in Y$. If $X \subset \overline{\Omega}$ is a nonempty set, then $\mathrm{ConvHull}_{\overline{\Omega}}(X)$ is the smallest closed convex subset of $\overline{\Omega}$ that contains X. We define

$$\mathrm{ConvHull}_{\Omega}(X) := \Omega \cap \mathrm{ConvHull}_{\overline{\Omega}}(X).$$

3.2 Hilbert metric

Suppose Ω is a properly convex domain and $\mathbb{A}$ is an affine chart that contains $\overline{\Omega}$ as a compact subset. We equip $\mathbb{A}$ with the Euclidean norm $|\cdot|$. If $x, y \in \Omega$, then there exist $a, b \in \partial\Omega$ such that

$$\mathbb{P}(\mathrm{Span}(\{x, y\})) \cap \overline{\Omega} = [a, b],$$

where the four points appear in the order a, x, y, b. The cross-ratio of these four points is given by

$$[a, x, y, b] := \frac{|b-x||y-a|}{|b-y||x-a|}.$$

The *Hilbert metric* on Ω is defined by

$$\mathrm{d}_\Omega(x, y) := \tfrac{1}{2}\log([a, x, y, b]).$$

Observation 3.2 *If $\Omega' \subset \Omega$ are properly convex domains and $x, y \in \Omega'$, then $\mathrm{d}_\Omega(x, y) \le \mathrm{d}_{\Omega'}(x, y)$.*

If $x, y \in \Omega$, then $[x, y]$ is a geodesic in $(\Omega, \mathrm{d}_\Omega)$ joining x and y. In order to emphasize this fact, we will often refer to the projective line segment $[x, y]$ as the *projective geodesic segment* between x and y. If $(x, y) \subset \Omega$ with $x, y \in \partial\Omega$, then (x, y) is a bi-infinite geodesic in $(\Omega, \mathrm{d}_\Omega)$ and we will call it a (*bi-infinite*) *projective geodesic*.

The space $(\Omega, \mathrm{d}_\Omega)$ is a proper, complete and geodesic metric space and we will call Ω a *Hilbert geometry*, see Definition 1.1. The group $\mathrm{Aut}(\Omega) := \{g \in \mathrm{PGL}_{d+1}(\mathbb{R}) \mid g\Omega = \Omega\}$ acts properly and isometrically on $(\Omega, \mathrm{d}_\Omega)$. However, the projective geodesic may not be the unique geodesic between points in $(\Omega, \mathrm{d}_\Omega)$. Consider, for example, the two-dimensional simplex $T_2 := \mathbb{P}(\mathbb{R}^+ e_1 \oplus \mathbb{R}^+ e_2 \oplus \mathbb{R}^+ e_3)$ with its Hilbert metric d_{T_2}. Then generic points $x \neq y \in T_2$ have uncountably many geodesics (for the Hilbert metric d_{T_2}) joining them [36, Proposition 2].

Definition 3.3 For a Hilbert geometry Ω, the preimage $\pi^{-1}(\Omega) := \{v \in \mathbb{R}^{d+1} \mid \pi(v) \in \Omega\}$ has two connected components. The *cone above* (*or over*) Ω, denoted by $\widetilde{\Omega}$, is a connected component of $\pi^{-1}(\Omega)$.

Then $\pi^{-1}(\Omega) = \widetilde{\Omega} \sqcup (-\widetilde{\Omega})$. If $g \in \mathrm{Aut}(\Omega)$, then there is a lift $\widetilde{g} \in \mathrm{GL}_{d+1}(\mathbb{R})$ of g that preserves $\widetilde{\Omega}$, ie $\widetilde{g} \cdot \widetilde{\Omega} = \widetilde{\Omega}$. Indeed, if a lift $\widetilde{g}$ does not preserve $\widetilde{\Omega}$, then $\widetilde{g} \cdot \widetilde{\Omega} = -\widetilde{\Omega}$ and hence $-\widetilde{g}$ preserves $\widetilde{\Omega}$. We have the following elementary observation about such lifts of automorphisms.

Observation 3.4 *Suppose $\widetilde{\Omega}$ is a cone above Ω and $\widetilde{g} \in \mathrm{GL}_{d+1}(\mathbb{R})$ preserves $\widetilde{\Omega}$. If $\widetilde{a} \in \widetilde{\Omega}$ satisfies $\widetilde{g} \cdot \widetilde{a} = \lambda \cdot \widetilde{a}$, then $\lambda > 0$.*

Proof Clearly $\lambda \neq 0$. As $\widetilde{g}$ preserves $\widetilde{\Omega}$, $\widetilde{g} \cdot \widetilde{a} \in \widetilde{\Omega}$ which implies that $\lambda \cdot \widetilde{a} \in \widetilde{\Omega}$. Now if $\lambda < 0$, then $\widetilde{a} \in (-\widetilde{\Omega})$. But then $\widetilde{a} \in \widetilde{\Omega} \cap (-\widetilde{\Omega}) = \varnothing$, a contradiction. □

3.3 Projective simplices

The standard k-dimensional projective simplex in $\mathbb{P}(\mathbb{R}^{d+1})$ is

$$T_k := \{[x_1 : \cdots : x_{k+1} : 0 : \cdots : 0] \in \mathbb{P}(\mathbb{R}^{d+1}) \mid x_1, \ldots, x_{k+1} > 0\}.$$

A k-dimensional projective simplex is a subset of $\mathbb{P}(\mathbb{R}^{d+1})$ of the form gT_k for some $g \in \mathrm{PGL}_{d+1}(\mathbb{R})$. If $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is a properly convex domain and $S \subset \Omega$ is a projective simplex, then we say that S is a *properly embedded simplex in* Ω if and only if $\partial S \subset \partial\Omega$.

The Hilbert metric d_{T_k} on T_k is given by

$$\mathrm{d}_{T_k}([x_1:\cdots:x_{k+1}:0:\cdots:0],[y_1:\cdots:y_{k+1}:0:\cdots:0]) = \max_{1\le i,j\le k+1} \frac{1}{2}\left|\log \frac{x_i y_j}{x_j y_i}\right|.$$

Then (T_k, d_{T_k}) is quasi-isometric to the real Euclidean space of dimension k. For a more elaborate discussion, see [39, Section 5], [44] or [36]. The group $\mathrm{Aut}(T_d)$ is generated by the group of permutation matrices in $\mathrm{PGL}_{d+1}(\mathbb{R})$ and the group $\{[\mathrm{diag}(\lambda_1,\ldots,\lambda_{d+1})] \in \mathrm{PGL}_{d+1}(\mathbb{R}) \mid \lambda_1,\ldots,\lambda_{d+1} > 0\}$.

Lemma 3.5 *Suppose* $g := [\mathrm{diag}(\lambda_1,\ldots,\lambda_{d+1})] \in \mathrm{Aut}(T_d)$ *where* $\lambda_i > 0$ *for all* $i = 1,\ldots,d+1$. *Let* $\lambda_{\max} := \max_{1\le i\le d+1} \lambda_i$ *and* $\lambda_{\min} := \min_{1\le i\le d+1}\lambda_i$. *Then* $\mathrm{d}_{T_d}(x, gx) = \frac{1}{2}\log(\lambda_{\max}/\lambda_{\min})$ *for any* $x \in T_d$.

Proof Fix $x = [x_1:\cdots:x_{d+1}] \in T_d$. Using the formula for d_{T_d},

$$\mathrm{d}_{T_d}(x, gx) = \max_{1\le i,j\le d+1} \frac{1}{2}\left|\log \frac{x_i\lambda_j x_j}{x_j\lambda_i x_i}\right| = \max_{1\le i,j\le d+1} \frac{1}{2}\left|\log \frac{\lambda_j}{\lambda_i}\right| = \frac{1}{2}\log\frac{\lambda_{\max}}{\lambda_{\min}}. \qquad \square$$

3.4 Examples of Hilbert geometries

The projective open ball $\Omega_d := \left\{[x_1:\cdots:x_d:1] \mid \sum_{i=1}^d x_i^2 < 1\right\}$ in $\mathbb{P}(\mathbb{R}^{d+1})$ is the simplest example of a divisible strictly convex Hilbert geometry. In fact Ω_d with its Hilbert metric is isometric to $\mathbb{H}^d$ and is well-known as the Beltrami–Klein model of real hyperbolic space. Moreover, $\mathrm{Aut}(\Omega_d) = \mathrm{PO}(d,1)$. There are several examples of divisible strictly convex Hilbert geometries that are not isometric to $\mathbb{H}^d$: in dimension 4, there is a construction due to Benoist [11, Proposition 3.1], while Kapovich [40] constructed examples in all dimensions above 4.

Among nonstrictly convex (divisible) Hilbert geometries, the simplest example is the standard d-simplex T_d; see Section 3.3. But this example is reducible, a term which we now define. Recall that a convex cone in $\mathbb{R}^{d+1}$ is a set $C \subset \mathbb{R}^{d+1}$ such that $r_1v_1 + r_2v_2 \in C$ whenever $v_1, v_2 \in C$ and $r_1, r_2 > 0$.

Definition 3.6 A properly convex domain $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is *reducible* if there exist convex cones $C_1 \subset \mathbb{R}^{d_1}$ and $C_2 \subset \mathbb{R}^{d_2}$ with $d_1, d_2 \ge 1$ such that $\Omega = \mathbb{P}(C_1 \oplus C_2)$. Otherwise, Ω is *irreducible*.

An irreducible nonstrictly convex (divisible) example is Pos_d with $d \ge 3$, the set of positive-definite real symmetric $d\times d$ matrices of unit trace. It is a properly convex domain in $\mathbb{R}^{d(d+1)/2}$ and is a projective model for the symmetric space of $\mathrm{SL}_d(\mathbb{R})$. The notion of symmetric domains generalize Pos_d. A symmetric domain Ω is a properly convex domain such that: for each $x \in \Omega$, there exists an order two isometry $s_x \in \mathrm{Aut}(\Omega)$ where x is the unique fixed point of s_x in Ω. Symmetric domains of real rank at least two are real projective analogues of higher rank Riemannian symmetric spaces of nonpositive curvature; see [12; 50] for details. As one might expect, symmetric domains are very special in the theory of properly convex domains. The rank rigidity theorem (Theorem 1.5) mentioned in the introduction is also a result in this spirit. Benoist proved the following result.

Theorem 3.7 [12, Theorem 5.2] *Suppose $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is an irreducible properly convex domain that is not a symmetric domain. If $\Gamma \leq \mathrm{SL}_{d+1}(\mathbb{R})$ is a discrete subgroup that acts cocompactly on Ω, then Γ is Zariski dense in $\mathrm{SL}_{d+1}(\mathbb{R})$.*

Besides simplices and the symmetric domains of real rank at least two, only a few examples of divisible nonstrictly convex Hilbert geometries are known. These are low-dimensional examples; see for instance [10; 24], which rely on Coxeter group constructions, or [1], which uses "cusp-doubling" construction for certain three manifolds.

3.5 Closest-point projection for the Hilbert metric

Suppose Ω is a Hilbert geometry. If $r > 0$, we set

$$\mathcal{B}_\Omega(x,r) := \{y \in \Omega \mid \mathrm{d}_\Omega(x,y) < r\}.$$

Lemma 3.8 [25, Lemma 1.7] *$\mathcal{B}_\Omega(x,r)$ is a relatively compact and convex set.*

If $C \subset \Omega$ is a closed convex set, we define the closest-point projection on C as: if $x \in \Omega$,

$$\Pi_C(x) := C \cap \overline{\mathcal{B}_\Omega(x, \mathrm{d}_\Omega(x,C))}.$$

As the intersection of two closed convex sets is again a closed convex set, Lemma 3.8 immediately implies the following.

Observation 3.9 *Suppose Ω is a Hilbert geometry, C is a closed convex set and $x \in \Omega$. Then $\Pi_C(x)$ is a compact convex set.*

Corollary 3.10 *Suppose $\sigma : \mathbb{R} \to (\Omega, \mathrm{d}_\Omega)$ is a unit-speed parametrization of the bi-infinite projective geodesic $\sigma(\mathbb{R})$ with $\sigma(\pm\infty) \in \partial\Omega$. If $x \in \Omega$, then there exist $T_x^-, T_x^+ \in \mathbb{R}$ with $T_x^- \leq T_x^+$ such that*

$$\Pi_{\sigma(\mathbb{R})}(x) = [\sigma(T_x^-), \sigma(T_x^+)].$$

Proof Any compact convex subset of the bi-infinite projective geodesic $\sigma(\mathbb{R})$ is of the form $[\sigma(T), \sigma(T')]$ with $T \leq T'$. □

3.6 Faces of properly convex domains

Suppose Ω is a Hilbert geometry. We define the relation $\sim_\Omega$: if $p, q \in \overline{\Omega}$, then $p \sim_\Omega q$ if and only if there exists an open projective line segment in $\overline{\Omega}$ that contains both p and q. The relation $\sim_\Omega$ is an equivalence relation; see [27, Section 3.3]. The equivalence class of $p \in \overline{\Omega}$ is called the (open) face of p and is denoted by $F_\Omega(p)$.

Proposition 3.11 *Suppose $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is a Hilbert geometry.*

(i) *If $x \in \partial\Omega$, then $F_\Omega(x) \subset \partial\Omega$.*

(ii) *$F_\Omega(x) = \Omega$ if and only if $x \in \Omega$.*

(iii) *If $x, y \in \partial\Omega$, then either $[x, y] \subset \partial\Omega$ or $(x, y) \subset \Omega$.*

(iv) *Suppose $[x, y] \subset \partial\Omega$, $a \in F_\Omega(x)$ and $b \in F_\Omega(y)$. Then $[a, b] \subset \partial\Omega$.*

Proof For part (i), note that if $y \in \Omega$, then $[y, x]$ cannot be extended beyond x in $\overline{\Omega}$. Thus $F_\Omega(x) \subset \overline{\Omega} - \Omega = \partial\Omega$. Part (ii) follows from part (i).

(iii) If $[x, y] \not\subset \partial\Omega$, choose any $z \in (x, y) \cap \Omega$. So $F_\Omega(z) = \Omega$. Then $(x, y) \subset F_\Omega(z) \subset \Omega$.

(iv) It suffices to prove this for $b = y$, ie to prove that $[a, y] \subset \partial\Omega$. Suppose, for a contradiction, that $(a, y) \subset \Omega$. Then $a \neq x$. Pick $a' \in F_\Omega(x)$ such that $x \in (a, a')$. As $(a, y) \subset \Omega$, pick $w \in (a, y)$. Then $(w, a') \subset \Omega$. Thus $\mathrm{ConvHull}_\Omega\{a', y, a\}$ is a nonempty set, as it contains $(w, a') \subset \Omega$. Hence, $\mathrm{ConvHull}_\Omega\{a', y, a\} \subset \Omega$. This implies that a, a' and y span a 2-simplex in $\overline{\Omega}$ and the interior of this 2-simplex is contained in Ω. As $x \in (a, a')$, (x, y) is contained in the interior of this 2-simplex. Thus $(x, y) \subset \Omega$, a contradiction. $\square$

3.7 Distance estimates

Proposition 3.12 [38, Proposition 5.2] *Suppose Ω is a Hilbert geometry and $\{x_n\}$ and $\{y_n\}$ are sequences in Ω such that $x := \lim_{n\to\infty} x_n$ and $y := \lim_{n\to\infty} y_n$ exist in $\overline{\Omega}$. If*

$$\liminf_{n\to\infty} \mathrm{d}_\Omega(x_n, y_n) < \infty,$$

then $y \in F_\Omega(x)$ and

$$\mathrm{d}_{F_\Omega(x)}(y, x) \le \liminf_{n\to\infty} \mathrm{d}_\Omega(x_n, y_n).$$

Note that if $\liminf_{n\to\infty} \mathrm{d}_\Omega(x_n, y_n) = 0$, then the above proposition implies that $y = x$.

Next we need the notion of Hausdorff distance. If (X, d) is a metric space, then the Hausdorff distance between $A, B \subset X$ is defined by

$$d^{\mathrm{Haus}}(A, B) = \max\Big\{\sup_{a\in A} d(a, B), \sup_{b\in B} d(b, A)\Big\}.$$

Proposition 3.13 [38, Proposition 5.3], [25, Lemma 1.8] *Suppose Ω is a Hilbert geometry and $x_1, x_2, y_1, y_2 \in \overline{\Omega}$ satisfy $F_\Omega(x_1) = F_\Omega(x_2)$ and $F_\Omega(y_1) = F_\Omega(y_2)$. If $(x_1, y_1) \subset \Omega$, then*

$$\mathrm{d}_\Omega^{\mathrm{Haus}}((x_1, y_1), (x_2, y_2)) \le \max\{\mathrm{d}_{F_\Omega(x_1)}(x_1, x_2), \mathrm{d}_{F_\Omega(x_2)}(y_1, y_2)\}.$$

In particular, if $x_i, y_i \in \Omega$, then $\mathrm{d}_\Omega^{\mathrm{Haus}}([x_1, y_1], [x_2, y_2]) \le \max\{\mathrm{d}_\Omega(x_1, x_2), \mathrm{d}_\Omega(y_1, y_2)\}$.

3.8 Translation length

Suppose Ω is a Hilbert geometry and let $g \in \mathrm{Aut}(\Omega)$. Its *translation length* is defined by $\tau_\Omega(g) := \inf_{x\in\Omega} \mathrm{d}_\Omega(x, gx)$.

Remark 3.14 [38, Observation 7.2] Suppose $g \in \mathrm{Aut}(\Omega)$ and $W \leq \mathbb{R}^{d+1}$ is a g-invariant subspace of dimension ≥ 2 such that $\Omega \cap \mathbb{P}(W)$ is nonempty. Then $\tau_\Omega(g) \leq \tau_{\Omega\cap\mathbb{P}(W)}(g|_W)$.

Proposition 3.15 [25, Proposition 2.1] *If* $g \in \mathrm{Aut}(\Omega)$, *then* $\tau_\Omega(g) = \frac{1}{2}\log\left|\frac{\lambda_1}{\lambda_{d+1}}(g)\right|$.

This differs from the formula in [25] by a factor of $\frac{1}{2}$. This is because our definition of Hilbert metric has the factor of $\frac{1}{2}$. We further remark that if $\tilde{g} \in \mathrm{GL}_{d+1}(\mathbb{R})$ is any lift of g, then $\tau_\Omega(g) = \frac{1}{2}\log(\lambda_{\max}(\tilde{g})/\lambda_{\min}(\tilde{g}))$.

3.9 Minimal translation sets

Suppose Ω is a Hilbert geometry and $\Gamma \leq \mathrm{Aut}(\Omega)$. If $H \leq \Gamma$ is a subgroup, then the minimal translation set of H in Ω is

$$\mathrm{Min}_\Omega(H) := \bigcap_{h\in H} \{x \in \Omega \mid \mathrm{d}_\Omega(x, h\cdot x) = \tau_\Omega(h)\}.$$

Example 3.16 (i) If $g = [\mathrm{diag}(\lambda_1, \dots, \lambda_{d+1})] \in \mathrm{Aut}(T_d)$ with $\lambda_i > 0$ for all $i = 1, \dots, d+1$, then Lemma 3.5 implies that $T_d = \mathrm{Min}_{T_d}(\langle g\rangle)$.

(ii) The minimal translation set could be empty; eg if u is a parabolic isometry in $\mathrm{PO}(2,1)$, then $\tau_{\mathbb{H}^2}(u) = 0$ and the minimal translation set of $\langle u\rangle$ is empty.

We will need the following result connecting eigenspaces with minimal translation sets.

Lemma 3.17 *Suppose* $a, b, c \in \partial\Omega$ *are three distinct fixed points of* $g \in \mathrm{Aut}(\Omega)$*.Then*

$$\mathrm{ConvHull}_\Omega\{a, b, c\} \subset \mathrm{Min}_\Omega(\langle g\rangle).$$

Proof Without loss of generality, we can assume that $\mathrm{ConvHull}_\Omega\{a, b, c\}$ is nonempty since the result is obviously true otherwise. Let $T := \mathrm{ConvHull}_\Omega\{a, b, c\}$. Fix a cone $\tilde{\Omega}$ over Ω and a lift $\tilde{g}$ of g that preserves $\tilde{\Omega}$. Let $V = \mathrm{Span}\{a, b, c\}$ and $g_0 := \tilde{g}|_V$. Let $\tilde{a}, \tilde{b}, \tilde{c} \in \tilde{\Omega}$ be lifts of a, b and c respectively, and fix the basis $\{\tilde{a}, \tilde{b}, \tilde{c}\}$ of V. In this basis, $g_0 = \mathrm{diag}(t_1, t_2, t_3)$. By Observation 3.4, we can assume that $t_1, t_2, t_3 > 0$. Since T is a 2-simplex in $\mathbb{P}(V)$, Lemma 3.5 implies that

$$d_T(x, g_0 x) = \frac{1}{2}\log\frac{\max\{t_1, t_2, t_3\}}{\min\{t_1, t_2, t_3\}}$$

for any $x \in T$.

Suppose $\lambda_1(\tilde{g}), \dots, \lambda_{d+1}(\tilde{g})$ are the eigenvalues of $\tilde{g}$, indexed in the nonincreasing order of their modulus. Then

$$|\lambda_1(\tilde{g})| \geq \max\{t_1, t_2, t_3\} \geq \min\{t_1, t_2, t_3\} \geq |\lambda_{d+1}(\tilde{g})|.$$

By Proposition 3.15,

$$\tau_\Omega(g) = \frac{1}{2}\log\left|\frac{\lambda_1(\tilde{g})}{\lambda_{d+1}(\tilde{g})}\right|.$$

Then $d_T(x, g_0 x) \leq \tau_\Omega(g)$ for any $x \in T$.

As $\Omega \cap \mathbb{P}(V) \supset T$, Observation 3.2 implies that $d_T(y', y) \geq d_{\Omega\cap\mathbb{P}(V)}(y', y)$ for any $y', y \in T$. Then $d_T(x, g_0 x) \geq \tau_{\Omega\cap\mathbb{P}(V)}(g_0)$ for any $x \in T$. Then Remark 3.14 implies that $d_T(x, g_0 x) \geq \tau_\Omega(g)$. Thus $d_T(x, g_0 x) = \tau_\Omega(g)$ for any $x \in T$. Hence $T \subset \mathrm{Min}_\Omega(g) = \mathrm{Min}_\Omega(\langle g\rangle)$. □

The next result concerns translation length and minimal translation sets of compact subgroups. This is essentially a restatement of [43, Lemma 2.1], which shows that every compact subgroup of $\mathrm{Aut}(\Omega)$ has a fixed point in Ω.

Lemma 3.18 [43, Lemma 2.1] *Suppose Ω is a Hilbert geometry and $H \leq \mathrm{Aut}(\Omega)$ is a compact subgroup. Then $\tau_\Omega(h) = 0$ for all $h \in H$ and $\mathrm{Min}_\Omega(H) = \{x \in \Omega \mid H \cdot x = x\} \neq \varnothing$.*

3.10 Centralizers

Suppose Ω is a Hilbert geometry and $\Gamma \leq \mathrm{Aut}(\Omega)$. If $H \leq \Gamma$ is a subgroup, the centralizer of H in Γ is

$$C_\Gamma(H) := \bigcap_{h\in H} \{g \in \Gamma \mid ghg^{-1} = h\}.$$

We will need to following result on cocompactness of centralizer subgroups.

Theorem 3.19 [38, Theorem 1.10] *Suppose Ω is a Hilbert geometry, $\mathscr{C} \subset \Omega$ is a closed convex subset, and $\Gamma \leq \mathrm{Aut}(\Omega)$ is a discrete subgroup that acts cocompactly on $\mathscr{C}$. If $A \leq \Gamma$ is an abelian subgroup, then $C_\Gamma(A)$ acts cocompactly on $\mathrm{ConvHull}_\Omega(\mathrm{Min}_{\mathscr{C}}(A))$, where*

$$\mathrm{Min}_{\mathscr{C}}(A) := \mathscr{C} \cap \mathrm{Min}_\Omega(A).$$

3.11 Proximality

We call $g \in \mathrm{GL}_{d+1}(\mathbb{R})$ proximal if g has a unique eigenvalue of maximum modulus and the multiplicity of this eigenvalue is 1, or equivalently if

$$|\lambda_1(g)| > |\lambda_2(g)|.$$

We will say that $g \in \mathrm{GL}_{d+1}(\mathbb{R})$ is biproximal if both g and g^{-1} are proximal, ie $|\lambda_1(g)| > |\lambda_2(g)|$ and $|\lambda_d(g)| > |\lambda_{d+1}(g)|$. Note that the notion of proximality is invariant under scaling a matrix by nonzero real numbers. Then $\gamma \in \mathrm{PGL}_{d+1}(\mathbb{R})$ is proximal (resp. biproximal) if some (hence any) lift of γ is proximal (resp. biproximal).

4 Dynamics of automorphisms

4.1 ω-limit sets of automorphisms

Let $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ be a Hilbert geometry and let $\gamma \in \mathrm{Aut}(\Omega)$ with

$$\tau_\Omega(\gamma) = \frac{1}{2}\log\left|\frac{\lambda_1}{\lambda_{d+1}}(\gamma)\right| > 0.$$

Recall that for any $X \subset \Omega$, $\overline{X}$ denotes the closure of X in $\overline{\Omega}$. We define the ω-limit set of γ as

$$\omega(\gamma, \Omega) := \bigcup_{x \in \Omega} (\overline{\{\gamma^n x \mid n \in \mathbb{N}\}} \cap \partial\Omega).$$

Thus, $\omega(\gamma, \Omega)$ is the union of all accumulation points in $\partial\Omega$ of all $\{\gamma^n \mid n \in \mathbb{N}\}$-orbits in Ω.

Example 4.1 Let $\Omega = T_2$ and $\gamma = [\mathrm{diag}(1, 2, 2)]$. Then for any $x = [x_1 : x_2 : x_3] \in T_2$, $\lim_{n\to\infty} \gamma^n x = [0 : x_2 : x_3]$. Thus

$$\omega(\gamma, T_2) = \{[0 : x_2 : x_3] \in \mathbb{P}(\mathbb{R}^3) \mid x_2, x_3 > 0\}.$$

Thus $\omega(\gamma, T_2)$ is the open projective line segment $(\pi(e_2), \pi(e_3)) \subset \partial T_2$. Also note that $\omega(\gamma, T_2) = E_\gamma^+ - \{\pi(e_2), \pi(e_3)\}$, where $E_\gamma^+ = \mathbb{P}(\mathrm{Span}\{e_2, e_3\}) \cap \partial\Omega$. Here $\mathbb{P}(\mathrm{Span}\{e_2, e_3\})$ is the projectivization of the direct sum of the eigenspace of γ that correspond to eigenvalues of maximum modulus. This observation that $\omega(\gamma, T_2) \subset E_\gamma^+$ holds more generally, as we will see in Proposition 4.9.

Remark 4.2 We now compare the notion of ω-limit set with that of the full orbital limit set introduced in [30]. Given an infinite discrete subgroup $H \leq \mathrm{Aut}(\Omega)$, the full orbital limit set of H is defined in [30] as

$$\mathcal{L}_\Omega^{\mathrm{orb}}(H) := \bigcup_{x \in \Omega} (\overline{H \cdot x} \cap \partial\Omega).$$

If $\gamma \in \mathrm{Aut}(\Omega)$ and $\tau_\Omega(\gamma) > 0$, then $\{\gamma^n \mid n \in \mathbb{N}\}$ is an infinite discrete subsemigroup of $\mathrm{Aut}(\Omega)$. Then $\omega(\gamma, \Omega)$ can be interpreted as the full orbital limit set of the subsemigroup $\{\gamma^n \mid n \in \mathbb{N}\}$.

4.2 Geometry of ω-limit sets of automorphisms

For the rest of this subsection, fix a Hilbert geometry $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ and $\gamma \in \mathrm{Aut}(\Omega)$ with $\tau_\Omega(\gamma) > 0$. Fix a lift $\widetilde{\gamma}$ of γ. Our goal here is Proposition 4.9 — a description of $\omega(\gamma, \Omega)$ using the real Jordan decomposition of $\widetilde{\gamma}$. We first give an intuitive idea. Suppose $c_1, \dots, c_q$ are all the eigenvalues (with repetitions) of $\widetilde{\gamma}$ of modulus $\lambda_{\max}(\widetilde{\gamma})$. If $c_1, \dots, c_q \in \mathbb{R}$, then $\omega(\gamma, \Omega)$ is contained in the projective subspace spanned by the eigenvectors corresponding to the eigenvalues of modulus $\lambda_{\max}(\widetilde{\gamma})$. In the notation of Definition 4.3 below, this subspace is precisely $\mathbb{P}(E_{\widetilde{\gamma}})$. Now suppose that among the c_i, there is a complex conjugate pair of eigenvalues $\mu, \overline{\mu} \in \mathbb{C} - \mathbb{R}$. Then, in the above subspace, we need to replace the eigenvectors for μ and $\overline{\mu}$ with a 2-dimensional γ-invariant projective real subspace on which γ acts by a rotation (E_μ in the notation of Definition 4.3). This is the key intuition behind Proposition 4.9. The references for this section are [42, II.1] and [25, Section 2]).

Now we start the formal discussion. First we introduce some notation. If $\mu \in \mathbb{R}$, define

$$J_\mu := \begin{pmatrix} \mu & 1 & 0 & \dots & 0 \\ 0 & \mu & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & 1 \\ 0 & \dots & 0 & \dots & \mu \end{pmatrix}.$$

If $\mu = \alpha + i\beta \in \mathbb{C} - \mathbb{R}$, define

$$J_\mu := \begin{pmatrix} R(\mu) & \mathrm{Id}_2 & 0 & \dots & 0 \\ 0 & R(\mu) & \mathrm{Id}_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & \mathrm{Id}_2 \\ 0 & \dots & 0 & \dots & R(\mu) \end{pmatrix},$$

where Id_2 is the 2×2 identity matrix and

$$R(\mu) := \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}.$$

Consider the real Jordan decomposition of $\widetilde{\gamma}$. This says that there is a $\widetilde{\gamma}$ invariant decomposition $\mathbb{R}^{d+1} = V_{\mu_1} \oplus \dots \oplus V_{\mu_n}$ into real vector subspaces and, with an appropriate choice of basis for V_{μ_j},

$$\widetilde{\gamma} = \begin{pmatrix} J_{\mu_1} & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & J_{\mu_n} \end{pmatrix}.$$

We remark that $V_\mu = V_{\overline{\mu}}$, as conjugate pairs of eigenvalues correspond to the same invariant subspace in the real Jordan decomposition. Without loss of generality, we can assume that $\mu_1, \dots, \mu_l \in \mathbb{R}$ and $\mu_{l+1}, \dots, \mu_n \in \mathbb{C} - \mathbb{R}$. Then $\mu_1, \dots, \mu_l, \mu_{l+1}, \overline{\mu}_{l+1}, \dots, \mu_n, \overline{\mu}_n$ are eigenvalues (possibly with repetitions) of $\widetilde{\gamma}$ over $\mathbb{C}$ and the multiplicity of μ_i is determined by the Jordan block J_{μ_i}.

Note that $\lambda_{\max}(\widetilde{\gamma})$ and $\lambda_{\min}(\widetilde{\gamma})$ are the maximum and the minimum of the set $\{|\mu_i| \mid 1 \le i \le n\}$. Now we focus on the eigenvalues of maximum modulus. By reindexing the μ_i, we now assume that $\mu_1, \dots, \mu_m$ are precisely those μ_i that satisfy $|\mu_i| = \lambda_{\max}(\widetilde{\gamma})$. We further assume that among them, $\mu_1, \dots, \mu_k \in \mathbb{R}$ and $\mu_{k+1}, \dots, \mu_m \in \mathbb{C} - \mathbb{R}$. Then $\mu_1, \dots, \mu_k, \mu_{k+1}, \overline{\mu}_{k+1}, \dots, \mu_m, \overline{\mu}_m$ are eigenvalues (possibly with repetitions) of $\widetilde{\gamma}$ of modulus $\lambda_{\max}(\widetilde{\gamma})$ and their multiplicities are determined by the Jordan block structure of $\widetilde{\gamma}$.

Definition 4.3 If $\mu_j \in \mathbb{R}$, let E_{μ_j} be the eigenvector for $\widetilde{\gamma}$ in V_{μ_j} with eigenvalue μ_j. If $\mu_j \in \mathbb{C} - \mathbb{R}$, let E_{μ_j} be the two-dimensional $\widetilde{\gamma}$-invariant subspace of V_{μ_j} such that $\widetilde{\gamma}|_{E_{\mu_j}}$ is conjugated to $R(\mu_j)$. Define

$$E_{\widetilde{\gamma}} := \bigoplus_{1 \le j \le m} E_{\mu_j} = \bigoplus_{|\mu_j| = \lambda_{\max}(\widetilde{\gamma})} E_{\mu_j}.$$

We also define

$$L_{\tilde{\gamma}} := \bigoplus_{|\mu_j| = \lambda_{\max}(\tilde{\gamma})} V_{\mu_j} \quad \text{and} \quad K_{\tilde{\gamma}} := \bigoplus_{|\mu_j| < \lambda_{\max}(\tilde{\gamma})} V_{\mu_j}.$$

Then $\tilde{\gamma}|_{E_{\tilde{\gamma}}}$ is conjugated in $\mathrm{GL}(E_{\tilde{\gamma}})$ to

$$\lambda_{\max}(\tilde{\gamma}) \cdot \begin{pmatrix} M_k & 0 & \dots & 0 \\ 0 & R\left(\frac{\mu_{k+1}}{\lambda_{\max}(\tilde{\gamma})}\right) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & R\left(\frac{\mu_m}{\lambda_{\max}(\tilde{\gamma})}\right) \end{pmatrix}, \tag{1}$$

where M_k is a $k \times k$ diagonal matrix with each diagonal entry ± 1. Thus $\langle \tilde{\gamma}|_{E_{\tilde{\gamma}}} \rangle$ is conjugated in $\mathrm{GL}(E_{\tilde{\gamma}})$ to a cyclic subgroup of $\{\pm \mathrm{Id}\}^k \times O(E_{\mu_{k+1}}) \times \cdots \times O(E_{\mu_m}) < O(E_{\tilde{\gamma}})$. Here, $O(W)$ denotes the group of orthogonal transformations preserving a linear subspace $W \subset \mathbb{R}^d$.

Claim 4.3.1 *There exists a sequence $\{m_k\}$ in $\mathbb{N}$ with $m_k \to \infty$ such that*

$$\lim_{k \to \infty} \frac{1}{\lambda_{\max}(\tilde{\gamma})^{m_k}} (\tilde{\gamma}|_{E_{\tilde{\gamma}}})^{m_k} = \mathrm{Id}|_{E_{\tilde{\gamma}}}.$$

Proof Let $k_\gamma := (1/\lambda_{\max}(\tilde{\gamma}))\tilde{\gamma}|_{E_{\tilde{\gamma}}}$ and $\mathcal{K} := \overline{\langle k_\gamma \rangle}$. By equation (1), there exists $h \in \mathrm{GL}(E_{\tilde{\gamma}})$ such that $h \cdot \mathcal{K} \cdot h^{-1}$ is a compact subgroup of $\{\pm \mathrm{Id}\}^k \times O(E_{\mu_{k+1}}) \times \cdots \times O(E_{\mu_m})$. Thus $\mathcal{K}' := h \cdot \mathcal{K} \cdot h^{-1}$ is a Lie subgroup of $O(E_{\tilde{\gamma}})$. Hence either Id is an isolated point of $\mathcal{K}'$ or there exists a neighborhood U of Id in $O(E_{\tilde{\gamma}})$ such that $U \cap \mathcal{K}' \subset \mathcal{K}'$.

In the latter case, it is obvious that there exists a monotonic sequence of integers $\{m_p\}$ such that $\lim_{p \to \infty} k_\gamma^{m_p} = \mathrm{Id}|_{E_{\tilde{\gamma}}}$. Up to passing to a subsequence, we can assume that $m_p \to \infty$ or $m_p \to -\infty$. If $m_p \to \infty$, the claim is proved. Otherwise, choose the sequence $-m_p$.

In the former case (ie when Id is an isolated point of $\mathcal{K}'$), it implies $(k_\gamma)^s = \mathrm{Id}|_{E_{\tilde{\gamma}}}$ for some $s \in \mathbb{N}$. Then $m_p := sp$ proves the claim. □

We will now discuss the dynamics of $(\tilde{\gamma})^n$ on $\mathbb{P}(\mathbb{R}^{d+1})$. The results are quite standard and the proofs are fairly elementary computations using Jordan blocks; see [42, II.1] or [25, Lemma 2.5] for instance.

Observation 4.4 (i) *For a generic point $v \in V_\mu$, all accumulation points of*

$$\left\{ \frac{1}{|\mu|^n} (\tilde{\gamma}|_{V_\mu})^n v \;\middle|\; n \in \mathbb{N} \right\}$$

lie in E_μ.

(ii) *Let $W = V_{\mu_1} \oplus V_{\mu_2}$ and $|\mu_1| > |\mu_2|$. Then, for any $w \in W - V_{\mu_2}$, all accumulation points of $\{(1/|\mu_1|^n)(\tilde{\gamma}|_W)^n w \mid n \in \mathbb{N}\}$ lie in E_{μ_1}.*

(iii) *Let $W' = V_\mu \oplus V_{\mu'}$ with $|\mu| = |\mu'|$. Then, for a generic point $w' \in W'$, all accumulation points of $\{(1/|\mu|^n)(\tilde{\gamma}|_W)^n w' \mid n \in \mathbb{N}\}$ lie in E_μ if $\dim V_\mu > \dim V_{\mu'}$. If $\dim V_\mu = \dim V_{\mu'}$, then the accumulation points lie in $E_\mu \oplus E_{\mu'}$.*

Recall the notation from Definition 4.3. Then the above observations imply the following result.

Fact 4.5 *If $w \in \mathbb{P}(\mathbb{R}^{d+1}) \setminus \mathbb{P}(K_{\tilde{\gamma}})$, then the accumulation points of $\{\gamma^n w \mid n > 0\}$ lie in $\mathbb{P}(E_{\tilde{\gamma}})$. In particular, if $w' \in \mathbb{P}(L_{\tilde{\gamma}})$, then all accumulation points of $\{\gamma^n w' \mid n > 0\}$ also lie in $\mathbb{P}(E_{\tilde{\gamma}})$.*

Remark 4.6 In fact, a finer conclusion is possible in Fact 4.5. Following [25], call a real Jordan subspace V_{μ_i} *most powerful* if $|\mu_i| = \lambda_{\max}(\tilde{\gamma})$ and $\dim(V_{\mu_i}) = \max\{\dim(V_{\mu_j}) \mid |\mu_j| = \lambda_{\max}(\tilde{\gamma})\}$. Let $F_{\tilde{\gamma}}$ be the direct sum of the E_{μ_j} that correspond to the most powerful Jordan subspaces V_{μ_j}. Then, $F_{\tilde{\gamma}} \subset E_{\tilde{\gamma}}$. For any $w \in \mathbb{P}(\mathbb{R}^{d+1}) \setminus \mathbb{P}(K_{\tilde{\gamma}})$ as above, the accumulation points of $\{\gamma^n w \mid n > 0\}$ actually lie in $\mathbb{P}(F_{\tilde{\gamma}})$; see part (iii) of Observation 4.4 above or [25, Proposition 2.5]. We record this finer conclusion for completeness, but we will not need it in this paper.

Claim 4.6.1 $\quad \Omega \cap \mathbb{P}(K_{\tilde{\gamma}}) = \varnothing, \quad \mathbb{P}(E_{\tilde{\gamma}}) \cap \overline{\Omega} \subset \partial\Omega \quad$ *and* $\quad \omega(\gamma, \Omega) \subset \mathbb{P}(E_{\tilde{\gamma}}) \cap \partial\Omega.$

Proof We first note that $\Omega \cap \mathbb{P}(K_{\tilde{\gamma}}) = \varnothing$. Otherwise, Remark 3.14 implies that

$$\tau_{\Omega \cap \mathbb{P}(K_{\tilde{\gamma}})}(\gamma) = \log\left(\frac{\lambda_{\max}(\tilde{\gamma}|_{K_{\tilde{\gamma}}})}{\lambda_{\min}(\tilde{\gamma}|_{K_{\tilde{\gamma}}})}\right) < \log\left(\frac{\lambda_{\max}(\tilde{\gamma})}{\lambda_{\min}(\tilde{\gamma})}\right) = \tau_\Omega(\gamma),$$

a contradiction. Suppose, if possible, that $\mathbb{P}(E_{\tilde{\gamma}}) \cap \Omega$ is nonempty. Then $\tau_\Omega(\gamma) \le \tau_{\mathbb{P}(E_{\tilde{\gamma}}) \cap \Omega}(\gamma|_{E_{\tilde{\gamma}}}) = 0$ by Remark 3.14, a contradiction. Finally, $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1}) \setminus \mathbb{P}(K_{\tilde{\gamma}})$ since $\Omega \cap \mathbb{P}(K_{\tilde{\gamma}}) = \varnothing$. Then Fact 4.5 implies that $\omega(\gamma, \Omega) \subset \mathbb{P}(E_{\tilde{\gamma}})$. Moreover, $\omega(\gamma, \Omega) \subset \partial\Omega$ by definition. □

Note that these subspaces in Definition 4.3 as well as the discussion above are independent of the lift $\tilde{\gamma}$ of γ that we fix. Thus we introduce the following definitions.

Definition 4.7 If $\gamma \in \mathrm{Aut}(\Omega)$, fix some (hence any) lift $\tilde{\gamma} \in \mathrm{GL}_{d+1}(\mathbb{R})$ of γ that preserves the cone $\widetilde{\Omega}$ above Ω, and define

$$E_\gamma^+ := \mathbb{P}(E_{\tilde{\gamma}}), \quad L_\gamma^+ := \mathbb{P}(L_{\tilde{\gamma}}) \quad \text{and} \quad K_\gamma^+ := \mathbb{P}(K_{\tilde{\gamma}}),$$

where the subspaces $E_{\tilde{\gamma}}$, $L_{\tilde{\gamma}}$ and $K_{\tilde{\gamma}}$ are as in Definition 4.3. We also define

$$E_\gamma^- := E_{\gamma^{-1}}^+, \quad L_\gamma^- := L_{\gamma^{-1}}^+ \quad \text{and} \quad K_\gamma^- := K_{\gamma^{-1}}^+.$$

Remark 4.8 A linear subspace $V \subset \mathbb{R}^{d+1}$ is a real Jordan subspace for $\widetilde{\gamma}$ with eigenvalue μ if and only if V is a real Jordan subspace for $\widetilde{\gamma}^{-1}$ with eigenvalue μ^{-1}. Indeed, this follows because $\ker(\widetilde{\gamma}-\mu\,\mathrm{Id})^k = \ker(\widetilde{\gamma}^{-1}-\mu^{-1}\,\mathrm{Id})^k$ for any $k \in \mathbb{N}$. Thus, if the V_μ are the real Jordan subspaces for $\widetilde{\gamma}$ as above, then

$$E_{\widetilde{\gamma}^{-1}} = \bigoplus_{|\mu|=\lambda_{\min}(\widetilde{\gamma})} E_\mu, \quad L_{\widetilde{\gamma}^{-1}} = \bigoplus_{|\mu|=\lambda_{\min}(\widetilde{g})} V_\mu \quad \text{and} \quad K_{\widetilde{\gamma}^{-1}} = \bigoplus_{|\mu|>\lambda_{\min}(\widetilde{g})} V_\mu.$$

The key upshot of the discussion in this subsection is the following proposition.

Proposition 4.9 *If Ω is a Hilbert geometry, $\gamma \in \mathrm{Aut}(\Omega)$ and $\tau_\Omega(\gamma) > 0$, then*

(i) $\omega(\gamma,\Omega) \subset E_\gamma^+$,

(ii) *the action of γ on E_γ^+ is conjugated into the projective orthogonal group $\mathrm{PO}(E_\gamma^+)$, and*

(iii) *there exists a sequence of positive integers $\{m_k\}$ with $m_k \to \infty$ such that*

$$\lim_{k\to\infty} (\gamma|_{E_\gamma^+})^{m_k} = \mathrm{Id}|_{E_\gamma^+}.$$

Remark 4.10 A similar proposition is true if we replace γ by γ^{-1} and E_γ^+ by E_γ^-. Moreover, it is possible that $\omega(\gamma,\Omega) \subsetneq E_\gamma^+ \subset \partial\Omega$; see Example 4.1. We finally remark that a finer conclusion is possible here: $\omega(\gamma,\Omega) \subset \mathbb{P}(F_{\widetilde{\gamma}}) \subset E_\gamma^+$, where $F_{\widetilde{\gamma}}$ is as defined in Remark 4.6. We will not need this finer conclusion, but we record it for completeness.

4.3 ω-limit sets and faces in a properly convex domain

We continue our discussion about ω-limit sets from the previous subsection. Our goal now is to prove a result about the faces $F_\Omega(x)$ for $x \in E_\gamma^\pm$. This result will be used in Section 11. Before formulating the precise result, we give an illustrative example.

Example 4.11 Let $g = \mathrm{diag}(\lambda_1,\lambda_2,\lambda_3)$ where $\lambda_1 > \lambda_2 > \lambda_3 > 0$, and let g preserve a properly convex domain $\Omega \subset \mathbb{P}(\mathbb{R}^3)$. Suppose $\pi(e_3) \in \partial\Omega$ and let $F := F_\Omega(\pi(e_3))$. We will show that $\pi(e_2) \notin F$. Suppose, on the contrary, that $\pi(e_2) \in F$. Then $I_t := [\pi(e_3 - te_2), \pi(e_3 + te_2)] \subset F$ for some $t > 0$. It is an elementary observation that I_t gets expanded by the action of g and $\overline{\bigcup_{k=1}^\infty g^k I_t} = \mathbb{P}(\mathrm{Span}\{e_2,e_3\})$. Thus $\mathbb{P}(\mathrm{Span}\{e_2,e_3\}) \subset \overline{F} \subset \overline{\Omega}$, which contradicts that Ω is a properly convex domain. Thus $\pi(e_2) \notin F$. By similar reasoning $\pi(e_1) \notin F$.

The takeaway from this example should be the following: since $\pi(e_3)$ is an eigenvector corresponding to an eigenvalue of modulus $\lambda_{\min}(g)$, the corresponding face $F_\Omega(\pi(e_3))$ cannot intersect any eigenspace whose eigenvalue has modulus greater than $\lambda_{\min}(g)$. The above philosophy works even if we replace eigenspaces by Jordan blocks, and is the key idea behind the next result.

We now state the precise version of the result. Recall the notation L_γ^- from the previous section (see Definition 4.7 and Remark 4.8): for any $\gamma \in \mathrm{Aut}(\Omega)$,

$$L_\gamma^- = \mathbb{P}\Big(\bigoplus_{|\mu|=\lambda_{\min}(\widetilde{\gamma})} V_\mu\Big).$$

As in the previous subsection, $\widetilde{\gamma}$ is some (hence any) lift of γ and V_μ is the real Jordan subspace of $\widetilde{\gamma}$ for the eigenvalue μ. Thus L_γ^- is the direct sum of all the Jordan subspaces corresponding to the eigenvalues of $\widetilde{\gamma}$ of minimum absolute value.

Lemma 4.12 *Suppose Ω is a Hilbert geometry and $\gamma \in \mathrm{Aut}(\Omega)$ with $\tau_\Omega(\gamma) > 0$. If $y \in E_\gamma^-$, then $F_\Omega(y) \subset L_\gamma^-$.*

Proof Suppose, for contradiction, that $v \in F_\Omega(y) - L_\gamma^-$. Fix a lift $\widetilde{\gamma}$ of γ. As $y \in E_\gamma^-$, Proposition 4.9(iii) implies we can find a sequence $\{d_k\}$ of positive integers with $d_k \to \infty$ such that $(\gamma|_{E_\gamma^-})^{d_k} \to \mathrm{Id}|_{E_\gamma^-}$.

Up to passing to a subsequence of $\{d_k\}$, we can assume that $\gamma^{d_k} v \to v_\infty \in \overline{\Omega}$. As $v \notin L_\gamma^-$, Observation 4.4 part (ii) implies that there exists $c > \lambda_{\min}(\widetilde{\gamma})$ such that the accumulation points of $\{(\widetilde{\gamma}/c)^{d_k} v \mid k \geq 1\}$ do not lie in L_γ^-. Thus $v_\infty \notin L_\gamma^-$ and $\lim_{k\to\infty}(c/\lambda_{\min})^{d_k} = \infty$. We can then fix lifts $\widetilde{y}$, $\widetilde{v}$ and $\widetilde{v}_\infty$ such that

$$\left(\frac{\widetilde{\gamma}}{\lambda_{\min}(\widetilde{\gamma})}\right)^{d_k} \widetilde{y} \to \widetilde{y} \quad \text{and} \quad \left(\frac{\widetilde{\gamma}}{c}\right)^{d_k} \widetilde{v} \to \widetilde{v}_\infty.$$

We claim that

$$\mathbb{P}(\mathrm{Span}\{y, v_\infty\}) \subset \overline{\Omega}.$$

To prove this claim, it suffices to show that $\pi(\widetilde{y} + t\widetilde{v}_\infty) \in \overline{\Omega}$ for any real number $t \neq 0$. Fix $0 \neq t \in \mathbb{R}$. Define

$$s_k := t \cdot \frac{\lambda_{\min}^{d_k}}{c^{d_k} + t\lambda_{\min}^{d_k}}.$$

Then $s_k \to 0$ as $k \to \infty$. In fact, for k large enough, s_k belongs to $(0,1)$ or $(-1,0)$ accordingly as $t > 0$ or $t < 0$. Set

$$w_k := \pi((1-s_k)\widetilde{y} + s_k\widetilde{v}) = \pi\left(\widetilde{y} + \frac{s_k}{1-s_k}\widetilde{v}\right) = \pi\left(\widetilde{y} + t\frac{\lambda_{\min}^{d_k}}{c^{d_k}}\widetilde{v}\right),$$

since $s_k/(1-s_k) = t(\lambda_{\min}^{d_k}/c^{d_k})$. Then $w_k \in \mathbb{P}(\mathrm{Span}\{y,v\})$ and $\lim_{k\to\infty} w_k = y$. Thus, for k large enough, $w_k \in F_\Omega(y) \cap \mathbb{P}(\mathrm{Span}\{y,v\})$ because $v \in F_\Omega(y)$. Moreover, w_k lies on opposite sides of y in $F_\Omega(y) \cap \mathbb{P}(\mathrm{Span}\{y,v\})$ accordingly as $t > 0$ or $t < 0$. Thus the following computation will show that γ^{d_k} expands small neighborhoods of y in $\mathbb{P}(\mathrm{Span}\{y,v\}) \cap F_\Omega(y)$ to large subintervals of the projective line $\mathbb{P}(\mathrm{Span}\{y, v_\infty\})$. More precisely, we observe that

$$\begin{aligned}\lim_{k\to\infty} \gamma^{d_k} w_k &= \lim_{k\to\infty} \pi\left((1-s_k)\frac{\widetilde{\gamma}^{d_k}}{\lambda_{\min}(\widetilde{\gamma})^{d_k}}\widetilde{y} + s_k\frac{\widetilde{\gamma}^{d_k}}{\lambda_{\min}^{d_k}}\widetilde{v}\right) = \lim_{k\to\infty} \pi\left(\frac{\widetilde{\gamma}^{d_k}}{\lambda_{\min}(\widetilde{\gamma})^{d_k}}\widetilde{y} + \frac{s_k}{1-s_k}\frac{c^{d_k}}{\lambda_{\min}^{d_k}}\frac{\widetilde{\gamma}^{d_k}}{c^{d_k}}\widetilde{v}\right) \\ &= \lim_{k\to\infty} \pi\left(\frac{\widetilde{\gamma}^{d_k}}{\lambda_{\min}(\widetilde{\gamma})^{d_k}}\widetilde{y} + t\frac{\widetilde{\gamma}^{d_k}}{c^{d_k}}\widetilde{v}\right) = \pi(\widetilde{y} + t\widetilde{v}_\infty).\end{aligned}$$

Thus $\pi(\tilde{y}+t\tilde{v}_\infty)\in\overline{\Omega}$, since $w_k\in F_\Omega(y)$ for k large enough. Since $t\neq 0$ is arbitrary, $\mathbb{P}(\mathrm{Span}\{y,v_\infty\})\subset\overline{\Omega}$. This proves the claim.

However, if $\overline{\Omega}$ contains the nontrivial projective line $\mathbb{P}(\mathrm{Span}\{y,v_\infty\})$, then Ω cannot be properly convex. This is a contradiction. $\square$

Corollary 4.13 *Suppose $\Omega\subset\mathbb{P}(\mathbb{R}^{d+1})$ is a Hilbert geometry and $\gamma\in\mathrm{Aut}(\Omega)$ with $\tau_\Omega(\gamma)>0$.*

(i) *If $y\in E_\gamma^-$, then $F_\Omega(y)\cap E_\gamma^+=\varnothing$.*

(ii) *If $y\in E_\gamma^-$, $z\in F_\Omega(y)$ and $\{i_k\}$ is a sequence in $\mathbb{Z}$ such that $z_\infty:=\lim_{k\to\infty}\gamma^{i_k}z$ exists, then $z_\infty\in E_\gamma^-$.*

Proof By Lemma 4.12, $F_\Omega(y)\subset L_\gamma^-$. Since $\tau_\Omega(\gamma)>0$, $L_\gamma^-\cap E_\gamma^+$ is empty by definition and this proves the first part. For the second part, note that $z\in F_\Omega(y)$ implies that $z\in L_\gamma^-$. On $\mathrm{Span}(L_\gamma^-)$, all eigenvalues of $\tilde{\gamma}$ have the same modulus $\lambda_{\min}(\tilde{\gamma})$. Then Observation 4.4(iii) implies that all accumulation points of $\{\gamma^n z\mid n\in\mathbb{N}\}$ lie in E_γ^-. By similar reasoning, all accumulation points of $\{\gamma^{-n}z\mid n\in\mathbb{N}\}$ also lie in E_γ^-. This proves the second part. $\square$

Remark 4.14 Analogues of Lemma 4.12 and Corollary 4.13 hold for $F_\Omega(x)$ where $x\in E_\gamma^+$. One has to replace γ with γ^{-1} to obtain the analogous results.

Part II Rank-one Hilbert geometries

5 Axis of isometries

Definition 5.1 Suppose $\Omega\subset\mathbb{P}(\mathbb{R}^{d+1})$ is a Hilbert geometry and $g\in\mathrm{Aut}(\Omega)$. An *axis* of g is a nontrivial projective line segment $\ell_g:=\mathbb{P}(V_g)\cap\Omega$ where $V_g\leq\mathbb{R}^{d+1}$ is a two-dimensional g-invariant linear subspace.

We will show that if g has an axis and $\tau_\Omega(g)>0$, then g acts by a translation along its axis ℓ_g and the endpoints of ℓ_g correspond to eigenvectors with eigenvalues of maximum and minimum modulus respectively. Recall the notation $E_g^+,E_g^-\subset\mathbb{P}(\mathbb{R}^{d+1})$ from Definition 4.7.

Lemma 5.2 *Suppose that $\Omega\subset\mathbb{P}(\mathbb{R}^{d+1})$ is a Hilbert geometry, and that $g\in\mathrm{Aut}(\Omega)$ with $\tau_\Omega(g)>0$ and g has an axis $\ell_g=\mathbb{P}(V_g)\cap\Omega$. If $\tilde{g}$ is a lift of g in $\mathrm{GL}_{d+1}(\mathbb{R})$, then*

(i) *$\tilde{g}|_{V_g}$ has two distinct eigenvalues $\lambda_+>\lambda_-$,*

(ii) *there exist $\tilde{g_+},\tilde{g_-}\in\mathbb{R}^{d+1}$ such that $\tilde{g}\cdot\tilde{g_\pm}=\lambda_\pm\cdot\tilde{g_\pm}$ and $\ell_g=(g_+,g_-)$, where $g_\pm=\pi(\tilde{g_\pm})$,*

(iii) *$|\lambda_+|=\lambda_{\max}(\tilde{g})$, $|\lambda_-|=\lambda_{\min}(\tilde{g})$ and $\tau_\Omega(g)=\log(|\lambda_+/\lambda_-|)>0$,*

(iv) *$g_+\in E_g^+$ and $g_-\in E_g^-$.*

Remark 5.3 If the lift $\widetilde{g}$ preserves the cone $\widetilde{\Omega}$ above Ω and $\widetilde{g_\pm} \in \widetilde{\Omega}$, then $\lambda_\pm > 0$; see Observation 3.4. Then $\lambda_+ = \lambda_{\max}(\widetilde{g})$ and $\lambda_- = \lambda_{\min}(\widetilde{g})$.

Proof Let $\ell_g = (a,b)$. Note that g preserves the set $\{a,b\} = \ell_g \cap \partial\Omega$. Fix any lift $\widetilde{g}$ of g. In the basis $\{\mathbb{R}\cdot a, \mathbb{R}\cdot b\}$ of V_g, there exist $c_1, c_2 \neq 0$ such that $\widetilde{g}|_{V_g}$ is either

$$\begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & c_1 \\ c_2 & 0 \end{pmatrix}.$$

In the latter case, both eigenvalues of $\widetilde{g}|_{V_g}$ have the same modulus and $\tau_{\Omega\cap\mathbb{P}(V_g)}(g|_{\mathbb{P}(V_g)}) = 0$, by Proposition 3.15. But then Remark 3.14 implies

$$0 \le \tau_\Omega(g) \le \tau_{\Omega\cap\mathbb{P}(V_g)}(g|_{\mathbb{P}(V_g)}) = 0,$$

a contradiction. Thus we are in the former case and g is diagonalizable with eigenvalues c_1 and c_2. Note that $c_1 \neq c_2$, since otherwise the same reasoning as above implies that $\tau_\Omega(g) = 0$. Then set $\lambda_+ := \max\{c_1, c_2\}$ and $\lambda_- := \min\{c_1, c_2\}$ and this proves part (i). For part (ii), let $\widetilde{g_\pm}$ be the eigenvectors of $\widetilde{g}$ in V_g with eigenvalues $\lambda_\pm$. Then note that by previous discussion, the set $\{\pi(\widetilde{g_+}), \pi(\widetilde{g_-})\}$ equals the set $\{a,b\}$. Thus $\ell_g = (\pi(\widetilde{g_+}), \pi(\widetilde{g_-}))$ and $\widetilde{g}|_{V_g} = \operatorname{diag}(\lambda_+, \lambda_-)$ in this basis.

For part (iii), first note that Remark 3.14 implies $\tau_\Omega(g) \le \tau_{\Omega\cap\mathbb{P}(V_g)}(g|_{\Omega\cap\mathbb{P}(V_g)})$. Proposition 3.15 then implies that $\log(\lambda_{\max}/\lambda_{\min})(\widetilde{g}) \le \log|\lambda_+/\lambda_-|$. Since $|\lambda_+| \le \lambda_{\max}(\widetilde{g})$ and $|\lambda_-| \ge \lambda_{\min}(\widetilde{g})$, we get $|\lambda_+/\lambda_-| \le (\lambda_{\max}/\lambda_{\min})(\widetilde{g})$. Thus

$$\left|\frac{\lambda_+}{\lambda_-}\right| = \frac{\lambda_{\max}}{\lambda_{\min}}(\widetilde{g}).$$

Then $|\lambda_+| = |\lambda_-|\cdot(\lambda_{\max}/\lambda_{\min})(\widetilde{g}) \ge \lambda_{\max}(\widetilde{g})$, implying $|\lambda_+| = \lambda_{\max}(\widetilde{g})$. Similarly, $|\lambda_-| = \lambda_{\min}(\widetilde{g})$. This proves part (iii). Then part (iv) follows by definition of E_g^+ and E_g^-. □

Corollary 5.4 *Suppose $g \in \operatorname{Aut}(\Omega)$ with $\tau_\Omega(g) > 0$ and g has an axis. If $\#(E_g^+) = \#(E_g^-) = 1$, then g has a unique axis given by $(E_g^+, E_g^-) \subset \Omega$. In particular, if g is biproximal* (see *Section 3.11*) *and has an axis, then the axis of g is unique.*

Proof Immediate from Lemma 5.2 parts (ii) and (iv), and the hypothesis that $\#(E_g^+) = \#(E_g^-) = 1$. For the "In particular" part, it suffices to note that if g is biproximal, then $\#(E_g^+) = \#(E_g^-) = 1$. □

Remark 5.5 Although biproximality of g implies $\#(E_g^+) = \#(E_g^-) = 1$, its converse fails in general. For example, consider

$$g = \begin{pmatrix} 0.25 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

However, we will show in Lemma 5.17 that if g has an axis, then g is biproximal if and only if $\#(E_g^+) = \#(E_g^-) = 1$.

An isometry $g \in \mathrm{Aut}(\Omega)$ may not have an axis; see Example 5.11 part (B) below. Hence we introduce the notion of a pseudoaxis.

Definition 5.6 Suppose $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is a Hilbert geometry and $g \in \mathrm{Aut}(\Omega)$. A *pseudoaxis* of g is a nontrivial projective line segment $\sigma_g := \mathbb{P}(W_g) \cap \overline{\Omega}$, where $W_g \leq \mathbb{R}^{d+1}$ is a two-dimensional g-invariant linear subspace such that $\mathbb{P}(W_g) \cap \Omega = \varnothing$.

Observation 5.7 *If $\tau_\Omega(g) > 0$, then g has either an axis or a pseudoaxis.*

This observation is immediate from the following result of Benoist (also see [43, Proposition 2.2]). Here $\widetilde{\Omega}$ denotes a cone above Ω; see Definition 3.3 and the remark that follows.

Proposition 5.8 [9, Lemma 3.2] *Suppose Ω is a Hilbert geometry, $g \in \mathrm{Aut}(\Omega)$ and $\tau_\Omega(g) > 0$. Let $\widetilde{g}$ be a lift of g that preserves $\widetilde{\Omega}$. Then $\widetilde{g}$ has a real positive eigenvalue that equals $\lambda_{\max}(\widetilde{g})$ and there exists v such that $\widetilde{g} \cdot v = \lambda_{\max}(\widetilde{g}) \cdot v$ and $\pi(v) \in \overline{\Omega}$. A similar result holds if we replace $\lambda_{\max}(\widetilde{g})$ by $\lambda_{\min}(\widetilde{g})$.*

Remark 5.9 If $\widetilde{g}$ is an arbitrary element of $\mathrm{GL}_{d+1}(\mathbb{R})$, then $\lambda_{\max}(\widetilde{g})$ doesn't have to be an eigenvalue of $\widetilde{g}$. In fact, $\widetilde{g}$ may only have complex nonreal eigenvalues of modulus $\lambda_{\max}(\widetilde{g})$. So the key point of the above proposition is that preserving the cone $\widetilde{\Omega}$ above Ω imposes a strong restriction, namely that $\widetilde{g}$ has a positive real eigenvalue that equals $\lambda_{\max}(\widetilde{g})$.

However, the proposition does not imply anything about the number (or nature) of the other eigenvalues whose modulus is $\lambda_{\max}(\widetilde{g})$. In Example 5.11 part (A), the matrix g_2^{-1} has a repeated eigenvalue $1/\lambda_2$ of maximum modulus. Moreover, $\widetilde{g}$ can have complex eigenvalues of modulus $\lambda_{\max}(\widetilde{g})$; see Example 5.12.

We will now discuss a few examples to illustrate the notions introduced. An isometry may have a unique axis, infinitely many axes, or no axes at all. An isometry can have pseudoaxes without having an axis, and vice versa.

Example 5.10 (unique axis, no pseudoaxes) Consider the Hilbert geometry $\Omega := \{[x : y : 1] \mid x^2 + y^2 < 1\}$ in $\mathbb{P}(\mathbb{R}^3)$. It is the projective model of $\mathbb{H}^2$ and $\mathrm{Aut}(\Omega) = \mathrm{PO}(2,1)$. If $g \in \mathrm{SO}(2,1)$ has $\tau_\Omega([g]) > 0$ (ie g is a hyperbolic isometry in $\mathrm{Isom}(\mathbb{H}^2)$), then $[g]$ has a unique axis.

Example 5.11 Consider the two-dimensional simplex $T_2 := \{[x_1 : x_2 : x_3] \mid x_1, x_2, x_3 > 0\}$.

- (A) (uncountably many axes, several pseudoaxes) Let $g_2 := [\mathrm{diag}(\lambda_1, \lambda_2, \lambda_2)]$, where $\lambda_1 > \lambda_2 > 0$. For $0 < t < 1$, let $Q_t := ([e_1], [te_2 + (1-t)e_3])$. Then $\{Q_t\}_{t \in (0,1)}$ is an uncountable family of axes of g_2. There are three pseudoaxes: $[e_1, e_2]$, $[e_2, e_3]$ and $[e_1, e_3]$.
- (B) (several pseudoaxes, no axis) Let $g_1 := [\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)]$, where $\lambda_1 > \lambda_2 > \lambda_3 > 0$. The pseudoaxes of g_1 are $[e_1, e_2]$, $[e_2, e_3]$ and $[e_1, e_3]$. But g_1 does not have an axis.

Example 5.12 Let $\Omega_2 \subset \mathbb{P}(\mathbb{R}^3)$ be the projective disk model of $\mathbb{H}^2$ and fix a cone $\widetilde{\Omega}_2$ over Ω_2. Define $\Omega_* := \{[v : x] \in \mathbb{P}(\mathbb{R}^4) \mid v \in \widetilde{\Omega}_2, x > 0\}$, ie $\Omega_* \subset \mathbb{P}(\mathbb{R}^4)$ is the properly convex domain obtained by the join of Ω_2 with a point. Let

$$g := \begin{bmatrix} \lambda A & 0 \\ 0 & \dfrac{1}{\lambda^3} \end{bmatrix} \in \mathrm{Aut}(\Omega_*), \quad \text{where } \lambda > 1 \text{ and } A = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \in \mathrm{SO}(2,1).$$

Then g has three eigenvalues $\lambda, \lambda e^{\pm i\theta}$ of maximum modulus.

Note that g has an axis $\ell_g := (\pi(e_3), \pi(e_4)) \subset \Omega_*$. The action of g is by a translation along ℓ_g and a rotation (by angle θ) around ℓ_g. The axis ℓ_g is contained in properly embedded triangles in Ω_*.

5.1 Three key lemmas

We conclude this section by establishing three lemmas that will be used in the next section. The first one is a consequence of Lemma 5.2.

Lemma 5.13 *Suppose $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is a Hilbert geometry, $g \in \mathrm{Aut}(\Omega)$ with $\tau_\Omega(g) > 0$, and a, b are fixed points of g with $a \in E_g^+$ and $b \in E_g^-$. If c is a fixed point of g such that $c \in \overline{\Omega} - (E_g^+ \cup E_g^-)$, then $[a,c] \cup [b,c] \subset \partial\Omega$.*

Proof First observe that $c \in \partial\Omega$. Otherwise, $\tau_\Omega(g) = d_\Omega(c, gc) = 0$, a contradiction. Suppose $(a,c) \subset \Omega$. Then (a,c) is an axis of g with $a \in E_g^+$. Lemma 5.2 then implies that $c \in E_g^-$, a contradiction. Thus $[a,c] \subset \partial\Omega$. Similar reasoning implies that $[c,b] \subset \partial\Omega$. □

The next lemma shows that if $g \in \mathrm{Aut}(\Omega)$, $\tau_\Omega(g) > 0$, g has an axis (a,b) and $\#(E_g^+) > 1$, then $F_\Omega(a)$ contains a nontrivial projective line segment in $\partial\Omega$. Before formulating the precise result and its proof, let us give an intuitive explanation of the main idea. Suppose $u \neq a \in E_g^+$ and let ξ be a point in $(a,b) \subset \Omega$. As Ω is open, we can find a point $\xi' \in \Omega \cap \mathbb{P}(\mathrm{Span}\{\xi, u\})$ that is distinct from ξ. Then, up to extracting a suitable subsequence of $\{g^n\}$, $g^{n_k}\xi \to a$ while $g^{n_k}\xi' \to a'$. As $\xi \neq \xi'$ and $u, a \in E_g^+$, one can check that $a' \neq a$; see the proof below. Then a property of the Hilbert metric (Proposition 3.12) implies that $a' \in F_\Omega(a)$. This is the gist of the proof below.

Lemma 5.14 *Suppose $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is a Hilbert geometry, $g \in \mathrm{Aut}(\Omega)$ with $\tau_\Omega(g) > 0$, and g has an axis (a,b), where $a \in E_g^+$ and $b \in E_g^-$. If $u \in E_g^+ \setminus \{a\}$, then there exist $x_u^- \neq x_u^+ \in \partial\Omega$ such that $a \in (x_u^-, x_u^+)$ and*

$$F_\Omega(a) \cap \mathbb{P}(\mathrm{Span}\{a,u\}) = (x_u^-, x_u^+).$$

Remark 5.15 Suppose we have the same setup as Lemma 5.14. By similar reasoning, if $v \in E_g^- \setminus \{b\}$, then $F_\Omega(b) \cap \mathbb{P}(\mathrm{Span}\{b,v\}) = (x_v^-, x_v^+)$, where $x_v^- \neq x_v^+$.

Proof Let us fix a cone $\widetilde{\Omega}$ over Ω. Then, we fix lifts $\tilde{g}, \tilde{a}, \tilde{b}, \tilde{u}$ of g, a, b, u such that $\tilde{a}, \tilde{b}, \tilde{u} \in \widetilde{\Omega}$ and $\tilde{g} \cdot \widetilde{\Omega} = \widetilde{\Omega}$. Note that $\tilde{a}$ is an eigenvector of $\tilde{g}$ corresponding to the eigenvalue $\lambda_{\max}(\tilde{g})$ or $-\lambda_{\max}(\tilde{g})$. Since $\tilde{g}$ preserves $\widetilde{\Omega}$, Observation 3.4 implies that $\tilde{g} \cdot \tilde{a} = \lambda_{\max}(\tilde{g}) \cdot \tilde{a}$. Similarly, $\tilde{g} \cdot \tilde{b} = \lambda_{\min}(\tilde{g}) \cdot \tilde{b}$. Since $u \in E_g^+$, Proposition 4.9(iii) implies that there exists an unbounded sequence of positive integers $\{m_k\}$ such that

$$\left(\frac{\tilde{g}}{\lambda_{\max}(\tilde{g})}\right)^{m_k} \tilde{u} = \tilde{u}.$$

For $t \in \mathbb{R}$, let $\tilde{p}_t := \frac{1}{2}(\tilde{a} + \tilde{b}) + t\tilde{u}$ and $p_t := \pi(\tilde{p}_t)$. Since (a, b) is an axis, $p_0 \in \Omega$. Then, as Ω is an open set, there exists $\varepsilon_0 > 0$ such that $\tilde{p}_t \in \widetilde{\Omega}$ for all $t \in (-\varepsilon_0, \varepsilon_0)$. Fix $t \in (-\varepsilon_0, \varepsilon_0)$. Then

$$\begin{aligned}\lim_{k\to\infty} g^{m_k} p_t &= \lim_{k\to\infty} \pi\left(\left(\frac{\tilde{g}}{\lambda_{\max}(\tilde{g})}\right)^{m_k} \tilde{p}_t\right) = \lim_{k\to\infty} \pi\left(\frac{\tilde{a}}{2} + \left(\frac{\lambda_{\min}(\tilde{g})}{\lambda_{\max}(\tilde{g})}\right)^{m_k} \frac{\tilde{b}}{2} + t\left(\frac{\tilde{g}}{\lambda_{\max}(\tilde{g})}\right)^{m_k} \tilde{u}\right) \\ &= \pi(\tilde{a} + 2t\tilde{u}) \in \overline{\Omega}.\end{aligned}$$

Then, $\lim_{k\to\infty} g^{m_k} p_0 = a$, and $\lim_{k\to\infty} g^{m_k} p_t \neq a$ whenever $t \neq 0$. By Proposition 3.12,

$$\lim_{k\to\infty} g^{m_k} p_t \in F_\Omega(a)$$

because $\lim_{k\to\infty} d_\Omega(g^{m_k} p_0, g^{m_k} p_t) = d_\Omega(p_0, p_t)$. Thus there exist $x_u^+ \neq x_u^- \in \partial\Omega$ such that

$$F_\Omega(a) \cap \mathbb{P}(\mathrm{Span}\{a, u\}) = (x_u^-, x_u^+). \qquad \square$$

The next lemma shows that if $\gamma \in \mathrm{Aut}(\Omega)$ has an axis and $\#(E_\gamma^-) = 1$, then γ^{-1} is a proximal element in $\mathrm{PGL}_{d+1}(\mathbb{R})$; see Section 3.11. Before stating the precise version of the result, we give an illustrative example to explain the main idea behind it.

Example 5.16 Let $\mu > \lambda > 0$. Suppose that

$$g = \begin{pmatrix} \mu & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

preserves $\Omega \subset \mathbb{P}(\mathbb{R}^3)$ and $\pi(e_1), \pi(e_2) \in \partial\Omega$. Here g satisfies $\#(E_g^-) = 1$ but g^{-1} is not proximal. The main takeaway from this example will be that such a matrix g cannot have an axis in Ω, ie the only candidate for an axis, namely $(\pi(e_1), \pi(e_2))$, cannot lie in Ω.

To proceed, we will first explain that $\pi(e_3)$ cannot lie in $\overline{\Omega}$. For this, first note that

$$g^{\pm k} \pi(e_3) = \pi(k\lambda^{k-1} e_2 + \lambda^k e_3).$$

Hence $g^{\pm k} \pi(e_3) \to \pi(e_2)$ as $k \to \infty$, but they approach $\pi(e_2)$ from "opposite directions" in the projective line $\mathbb{P}(\mathrm{Span}\{e_2, e_3\})$. That is, $g^{\pm k}$ "wraps" $[g^{-1}\pi(e_3), g\pi(e_3)]$ around $\mathbb{P}(\mathrm{Span}\{e_2, e_3\})$. Then, $\pi(e_3) \in \overline{\Omega}$ will imply that $\mathbb{P}(\mathrm{Span}\{e_2, e_3\}) \subset \overline{\Omega}$, which is a contradiction as Ω is a properly convex domain. Thus $\pi(e_3) \notin \overline{\Omega}$.

Now we revisit our basic proposition: that $(\pi(e_1), \pi(e_2))$ cannot lie in Ω. Suppose this is false and $(\pi(e_1), \pi(e_2)) \subset \Omega$. Since $g^k(\pi(e_1), \pi(e_3)) \to (\pi(e_1), \pi(e_2))$, we can find $\pi(y_k) \in \Omega \cap (\pi(e_1), \pi(e_3))$ such that $g^k \pi(y_k)$ converges to the midpoint of $(\pi(e_1), \pi(e_2))$. Now unless $\pi(y_k) \to \pi(e_3)$, one can use the action of g to show that $g^k \pi(y_k) \to \pi(e_1)$, a contradiction; see the computation in equation (5). Thus $\pi(y_k) \to \pi(e_3)$ and hence $\pi(e_3) \in \overline{\Omega}$. This contradicts the previous paragraph.

The argument discussed above is the gist of the proof below. We now precisely formulate and prove our result.

Lemma 5.17 *Suppose $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is a Hilbert geometry, $\gamma \in \mathrm{Aut}(\Omega)$ with $\tau_\Omega(\gamma) > 0$ and γ has an axis. If $\#(E_\gamma^-) = 1$, then*

$$\left| \frac{\lambda_d}{\lambda_{d+1}}(\gamma) \right| > 1.$$

Remark 5.18 Similar reasoning with γ replaced by γ^{-1} implies that if $\#(E_\gamma^+) = 1$, then

$$\left| \frac{\lambda_1}{\lambda_2}(\gamma) \right| > 1.$$

Proof Suppose the axis of γ is (a, b) with $a \in E_\gamma^+$ and $b \in E_\gamma^-$. Let us fix $\widetilde{\Omega}$, a cone above Ω. Fix lifts $\widetilde{\gamma}$, $\widetilde{a}$ and $\widetilde{b}$ where $\widetilde{a}, \widetilde{b} \in \widetilde{\Omega}$ and $\widetilde{\gamma} \cdot \widetilde{\Omega} = \widetilde{\Omega}$. Set $\lambda_{\max} := \lambda_{\max}(\widetilde{\gamma})$ and $\lambda_{\min} := \lambda_{\min}(\widetilde{\gamma})$. Since $b \in E_\gamma^-$ is a fixed point and $\widetilde{b} \in \widetilde{\Omega}$, Observation 3.4 implies that $\widetilde{\gamma} \cdot \widetilde{b} = \lambda_{\min} \cdot \widetilde{b}$. Similarly, $\widetilde{\gamma} \cdot \widetilde{a} = \lambda_{\max} \cdot \widetilde{a}$.

Since $\#(E_\gamma^-) = 1$, there is a one-dimensional eigenspace of $\widetilde{\gamma}$ (namely $\mathbb{R}\widetilde{b}$) and a single Jordan block $J_{\min}$ corresponding to eigenvalues of modulus $\lambda_{\min}$ (immediate from the definition, see Definition 4.7). Thus, in order to prove $|(\lambda_d/\lambda_{d+1})(\gamma)| > 1$, it is enough to show that the Jordan block $J_{\min}$ has size 1. Suppose this is false. Then there exists $\widetilde{w} \in \mathbb{R}^{d+1}$ such that if $k \in \mathbb{Z}$, then

$$\widetilde{\gamma}^k \widetilde{w} = k \lambda_{\min}^{k-1} \widetilde{b} + \lambda_{\min}^k \widetilde{w}. \tag{2}$$

Setting $w := \pi(\widetilde{w})$, $\lim_{k\to\infty} \gamma^k w = b$. Since $\gamma^k a = a$ for all k, $\lim_{k\to\infty} \gamma^k [a, w] = [a, b]$. Fix $p \in (a, b) \subset \Omega$. Then there exist $y_k \in (a, w)$ such that

$$\lim_{k\to\infty} \gamma^k y_k = p. \tag{3}$$

Since $p \in \Omega$ and Ω is open, $\gamma^k y_k \in \Omega$ for k large enough. Thus, up to truncating finitely many terms of the sequence $\{y_k\}$, we can assume that for $k \geq 1$,

$$y_k \in (a, w) \cap \Omega.$$

We can fix lifts $\widetilde{y}_k$ of y_k in $\widetilde{\Omega}$ such that

$$\widetilde{y}_k = c_k \widetilde{a} + d_k \widetilde{w}, \tag{4}$$

where $c_k, d_k \in [0, 1]$. Thus, up to passing to a subsequence, we can assume that $c_\infty := \lim_{k\to\infty} c_k$ and $d_\infty := \lim_{k\to\infty} d_k$ exist. Then $\widetilde{y}_\infty := \lim_{k\to\infty} \widetilde{y}_k$ exists and we set

$$y_\infty := \pi(\widetilde{y}_\infty) = \pi(c_\infty \widetilde{a} + d_\infty \widetilde{w}).$$

We now claim that $y_\infty = \pi(\widetilde{w}) = w$. If this is not true, then $c_\infty \neq 0$. Then, there exists $k_0 \in \mathbb{N}$ such that $c_k > c_\infty/2$ for all $k > k_0$, and $\lim_{k\to\infty}(d_k/c_k) = d_\infty/c_\infty$ exists in $\mathbb{R}$. Then using equation (3) followed by (4) and (2),

$$(5)\quad p = \lim_{k\to\infty} \gamma^k y_k = \lim_{k\to\infty} \pi\left(\frac{\widetilde{\gamma}^k \widetilde{y}_k}{c_k \lambda_{\max}^k}\right) = \lim_{k\to\infty} \pi\left(\widetilde{a} + \frac{d_k}{c_k}\left(\frac{k}{\lambda_{\max}}\left(\frac{\lambda_{\min}}{\lambda_{\max}}\right)^{k-1}\widetilde{b} + \left(\frac{\lambda_{\min}}{\lambda_{\max}}\right)^k \widetilde{w}\right)\right)$$
$$= \pi(\widetilde{a}) = a.$$

This is a contradiction since $p \in \Omega$ while $a \in \partial\Omega$. Thus $y_\infty = w$.

Since $y_k \in \Omega$ for $k \geq 1$, $w = y_\infty$ implies that $w \in \overline{\Omega}$. Then for all $k \in \mathbb{Z}$,

$$(6)\qquad [w, \gamma^k w] \subset \overline{\Omega}.$$

We now show that this implies $\mathbb{P}(\mathrm{Span}\{w, b\}) \subset \overline{\Omega}$. For $t > 0$, let

$$\mathcal{H}_t := \{\pi(\widetilde{w} + r\widetilde{b}) \mid -t \leq r \leq t\}.$$

Then $\overline{\bigcup_{t>0} \mathcal{H}_t} = \mathbb{P}(\mathrm{Span}\{b, w\})$. Now observe that if $k \in \mathbb{Z}$, then equation (2) implies that

$$\gamma^k w = \pi\left(\frac{\widetilde{\gamma}^k \widetilde{w}}{\lambda_{\min}^k}\right) = \pi\left(\widetilde{w} + \frac{k}{\lambda_{\min}}\widetilde{b}\right).$$

Then, for every $t > 0$, there exists $k_t \in \mathbb{N}$ such that $\mathcal{H}_t \subset [\gamma^{-(k_t-1)} w, w] \cup [w, \gamma^{k_t} w]$. Then, by equation (6), $\mathcal{H}_t \subset \overline{\Omega}$ for any $t > 0$. Thus $\mathbb{P}(\mathrm{Span}\{w, b\}) = \overline{\bigcup_{t>0} \mathcal{H}_t} \subset \overline{\Omega}$. This is a contradiction as Ω is a properly convex domain and hence $\overline{\Omega}$ cannot contain a projective line. □

6 Rank-one isometries: definition and properties

In this section, we introduce the notion of rank-one isometries for Hilbert geometries. Our definition is analogous to the definition of rank-one isometries for CAT(0) spaces [2; 5]. The notion of half triangles that we introduce is analogous to the notion of half flats used in the CAT(0) setting. Refer to Figure 1 for the next two definitions.

Definition 6.1 Suppose $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is a Hilbert geometry. Then the points $x, z, y \in \partial\Omega$ form a *half triangle in* Ω if

$$[x, z] \cup [y, z] \subset \partial\Omega \quad \text{and} \quad (x, y) \subset \Omega.$$

For $x, y \in \partial\Omega$, we will say that the projective geodesic $(x, y) \subset \Omega$ is *not contained in any half triangle in* Ω if for any $z \in \partial\Omega$, either $(x, z) \subset \Omega$ or $(z, y) \subset \Omega$.

Definition 6.2 Suppose $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is a Hilbert geometry and $a, b \in \partial\Omega$. The projective geodesic (a, b) is a *rank-one geodesic* provided $(a, b) \subset \Omega$ is not contained in any half triangle in Ω.

We now define rank-one isometries for Hilbert geometries. An isometry in $\mathrm{Aut}(\Omega)$ is rank one if it acts by a translation along a rank-one geodesic; see Figure 1.

Definition 6.3 Suppose $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is a Hilbert geometry.

(i) An element $\gamma \in \mathrm{Aut}(\Omega)$ is a *rank-one isometry* if
  (a) $\tau_\Omega(\gamma) = \log\left|\frac{\lambda_1}{\lambda_{d+1}}(\gamma)\right| > 0$,
  (b) γ has an axis,
  (c) none of the axes ℓ_γ of γ are contained in a half triangle in Ω.

(ii) A bi-infinite projective geodesic $\ell \subset \Omega$ is a *rank-one axis* if ℓ is the axis of a rank-one isometry in $\mathrm{Aut}(\Omega)$.

Remark 6.4 The prototypical example of a rank-one isometry is a hyperbolic isometry $[\mathrm{diag}(\lambda, 1, 1/\lambda)]$ with $\lambda > 1$ in $\mathrm{Isom}(\mathbb{H}^2)$; see Example 5.10. On the other hand, any element in $\mathrm{Aut}(T_d)$, where T_d is a d-dimensional simplex, is a nonexample. In fact, this nonexample highlights the necessity of the half triangle condition in the definition of a rank-one isometry, as we now explain. Recall Example 5.11 part (A). In that example, $g_2 = [\mathrm{diag}(\lambda_1, \lambda_2, \lambda_2)]$ has an axis Q_t for each $0 < t < 1$ and $\tau_{T_2}(g_2) > 0$. But all of these axes are contained in the projective triangle T_2 (and hence a half triangle). For another nonexample, see Example 5.12.

Recall Definition 1.3: a *rank-one Hilbert geometry* is a pair (Ω, Γ) where Ω is a Hilbert geometry and $\Gamma \le \mathrm{Aut}(\Omega)$ is a discrete subgroup that contains a rank-one isometry. In Appendix A, we discuss examples and also a generalization for convex cocompact groups.

We will now establish some key geometric and dynamical properties of rank-one isometries. The essence here is that translating along a rank-one axis is much more special than translating along any axis, and Proposition 6.5 could be interpreted as strengthening Lemma 5.2 under the rank one assumption. Our results are reminiscent of Ballmann's results in rank-one Riemannian nonpositive curvature [2; 4]. Recall the notation $E_g^{\pm}$ from Definition 4.7 and the notion of proximality from Section 3.11.

Proposition 6.5 *Suppose Ω is a Hilbert geometry and $\gamma \in \mathrm{Aut}(\Omega)$ is a rank-one isometry with an axis $\ell_\gamma = (a, b)$, where $a \in E_\gamma^+$ and $b \in E_\gamma^-$. Then*

(i) *γ is biproximal,*

(ii) *ℓ_γ is the unique axis of γ in Ω,*

(iii) *the only fixed points of γ in $\overline{\Omega}$ are a and b,*

(iv) *if $z' \in \partial\Omega \setminus \{a, b\}$, then $(a, z') \cup (b, z') \subset \Omega$ (see Figure 1),*

(v) *if $z \in \partial\Omega \setminus \{a, b\}$, then neither (a, z) nor (b, z) is contained in a half triangle in Ω.*

Remark 6.6 If γ is a rank-one isometry, then the above proposition shows that $\#(E_\gamma^\pm) = 1$ and we will henceforth use the notation $\gamma^\pm := E_\gamma^\pm$. We will call γ^+ the *attracting fixed point of* γ and γ^- the *repelling fixed point of* γ. We choose this terminology because γ has *north–south dynamics* on $\partial\Omega$; see Corollary 6.7.

Proof Let us fix $\widetilde{\Omega}$, a cone over Ω. For the rest of this proof, fix lifts $\widetilde{\gamma}$, $\widetilde{a}$ and $\widetilde{b}$, where $\widetilde{a}, \widetilde{b} \in \widetilde{\Omega}$ and $\widetilde{\gamma} \cdot \widetilde{\Omega} = \widetilde{\Omega}$. Set $\lambda_{\max} := \lambda_{\max}(\widetilde{\gamma})$ and $\lambda_{\min} := \lambda_{\min}(\widetilde{\gamma})$. Since $a \in E_\gamma^+$ is a fixed point of γ, the lift $\widetilde{a}$ is an eigenvector of $\widetilde{\gamma}$ corresponding to the eigenvalue $\lambda_{\max}$ or $-\lambda_{\max}$. By Observation 3.4,

$$\widetilde{\gamma} \cdot \widetilde{a} = \lambda_{\max} \cdot \widetilde{a}.$$

Similarly, $\widetilde{\gamma} \cdot \widetilde{b} = \lambda_{\min} \cdot \widetilde{b}$.

(i) By the hypothesis, $\#(E_\gamma^\pm) \geq 1$. In order to prove that γ is biproximal, we first prove that:

Claim 6.6.1

$$\#(E_\gamma^+) = \#(E_\gamma^-) = 1.$$

Proof It suffices to prove the claim for E_γ^+, since the same arguments with γ replaced by γ^{-1} implies the result for E_γ^-. Now suppose the claim is false and there exists $u \in E_\gamma^+ \setminus \{a\}$. Then Lemma 5.14 implies that there exist $z^-, z^+ \in \partial\Omega$ such that $a \in (z^-, z^+)$ and

$$F_\Omega(a) \cap \mathbb{P}(\mathrm{Span}\{a, u\}) = (z^-, z^+).$$

Then, $\mathcal{I}_z := [z_-, z_+]$ is the maximal projective line segment in $\partial\Omega$ containing both z_- and z_+.

Since γ is a rank-one isometry, its axis (a, b) cannot be contained in a half triangle in Ω. But $[a, z_+] \subset \partial\Omega$, which implies that $(z_+, b) \subset \Omega$. Similarly, $(z_-, b) \subset \Omega$. Choose $x_\pm \in (z_\pm, b) \cap \Omega$. By Proposition 4.9 part (iii), there exists a sequence $\{m_k\}$ of positive integers with $m_k \to \infty$ such that

$$\lim_{k\to\infty} (\gamma|_{E_\gamma^+})^{m_k} = \mathrm{Id}_{E_\gamma^+}.$$

Since $z_+ \in \mathbb{P}(\mathrm{Span}\{a, u\})$, it follows that $z_+ \in E_\gamma^+$. Fix a lift $\widetilde{z_+} \in \widetilde{\Omega}$ of z_+. Then

$$\lim_{k\to\infty} \left(\frac{\widetilde{\gamma}}{\lambda_{\max}}\right)^{m_k} \widetilde{z_+} = \widetilde{z_+}.$$

On the other hand,

$$\lim_{k\to\infty} \left(\frac{\widetilde{\gamma}}{\lambda_{\max}}\right)^{m_k} \widetilde{b} = \lim_{k\to\infty} \left(\frac{\lambda_{\min}}{\lambda_{\max}}\right)^{m_k} \widetilde{b} = 0,$$

as $\lambda_{\max} > \lambda_{\min}$. Then, since $x_+ \in (z_+, b)$,

$$\lim_{k\to\infty} \gamma^{m_k} x_+ = z_+.$$

Similarly,

$$\lim_{k\to\infty} \gamma^{m_k} x_- = z_-.$$

Since $\lim_{k\to\infty} \mathrm{d}_\Omega(\gamma^{m_k} x_+, \gamma^{m_k} x_-) = \mathrm{d}_\Omega(x_+, x_-)$, Proposition 3.12 implies that $z_+ \in F_\Omega(z_-)$. Thus there is an open projective line segment in $\partial\Omega$ containing both z_+ and z_-. This contradicts the maximality of $\mathcal{I}_z$ and finishes the proof of Claim 6.6.1. □

By the above claim, $\#(E_\gamma^+) = \#(E_\gamma^-) = 1$, where $\tau_\Omega(\gamma) > 0$ and γ has an axis (a,b). Then Lemma 5.17 implies that γ is biproximal.

(ii) This follows from biproximality of γ and Corollary 5.4.

(iii) Suppose c is a fixed point of γ in $\partial\Omega$ that is distinct from both a and b. By part (i) of this proposition, γ is biproximal. Thus $c \notin E_\gamma^+ \cup E_{\gamma^-}$. Then, by Lemma 5.13, $[a,c] \subset \partial\Omega$ and $[b,c] \subset \partial\Omega$. Thus, the axis $\ell_\gamma = (a,b)$ of γ is contained in a half triangle, contradicting that γ is a rank-one isometry.

(iv) Let $v \in \partial\Omega \setminus \{a,b\}$. Then $v \notin \mathbb{P}(\mathrm{Span}\{a,b\})$ as $(a,b) \subset \Omega$. Suppose $[a,v] \subset \partial\Omega$. Since γ is biproximal, there exists a γ-invariant decomposition of $\mathbb{R}^{d+1}$ given by

$$\mathbb{R}^{d+1} = \mathbb{R}\tilde{a} \oplus \mathbb{R}\tilde{b} \oplus \tilde{E}.$$

Choose any lift $\tilde{v}$ of v in $\tilde{\Omega}$. Then $\tilde{v}$ decomposes as

$$\tilde{v} = c_1\tilde{a} + c_2\tilde{b} + \tilde{v}_0,$$

where $c_1, c_2 \in \mathbb{R}$ and $\tilde{v}_0 \neq 0$. If $c_2 \neq 0$, then $\lim_{n\to\infty} \gamma^{-n} v = b$, that is, $\lim_{n\to\infty} \gamma^{-n}[a,v] = [a,b]$. Since $[a,v] \subset \partial\Omega$ by assumption, $[a,b] \subset \partial\Omega$. This is a contradiction since $(a,b) \subset \Omega$. Thus, $c_2 = 0$.

Set $\lambda_{\tilde{E}} := \lambda_{\max}(\tilde{\gamma}|_{\tilde{E}})$. Since γ is biproximal, $\lambda_{\tilde{E}} < \lambda_{\max}$. Then, for every $n > 0$,

$$\left(\frac{\tilde{\gamma}}{\lambda_{\tilde{E}}}\right)^{-n} \tilde{v} = c_1 \left(\frac{\lambda_{\max}}{\lambda_{\tilde{E}}}\right)^{-n} \tilde{a} + \left(\frac{\tilde{\gamma}|_{\tilde{E}}}{\lambda_{\tilde{E}}}\right)^{-n} \tilde{v}_0.$$

Up to passing to a subsequence, we can assume that $v_\infty := \lim_{n\to\infty} \gamma^{-n} v$ exists. Note that $v_\infty \in \overline{\Omega} \cap \mathbb{P}(\tilde{E})$. But $\overline{\Omega} \cap \mathbb{P}(\tilde{E})$ is a γ-invariant nonempty convex compact subset of $\mathbb{R}^{d-1}$ and Brouwer's fixed point theorem implies that γ has a fixed point in $\overline{\Omega} \cap \mathbb{P}(\tilde{E})$. But $\overline{\Omega} \cap \mathbb{P}(\tilde{E}) \cap \{a,b\} = \varnothing$. This contradicts part (iii). Hence, $(a,v) \subset \Omega$. Similarly we can show that $(b,v) \subset \Omega$.

(v) This is a consequence of part (iv). □

Corollary 6.7 *Suppose $\gamma \in \mathrm{Aut}(\Omega)$ is a rank-one isometry. Then γ has* ***north–south dynamics*** *on $\partial\Omega$, that is,*

$$(\gamma|_{\overline{\Omega} - \{\gamma^\mp\}})^{\pm n} \to \gamma^\pm \quad \text{as } n \to \infty,$$

and the convergence is locally uniform on compact subsets of $\overline{\Omega} - \{\gamma^\mp\}$.

Proof The proof is very similar to part (iv) of Proposition 6.5. By the above proposition, γ is biproximal. Thus there exists a γ-invariant decomposition $\mathbb{R}^{d+1} = \mathbb{R}\gamma^+ \oplus H_\gamma \oplus \mathbb{R}\gamma^-$, where $\gamma^\pm = E_\gamma^\pm$. Moreover, γ^n converges to the constant map γ^+ locally uniformly on compact subsets of $\mathbb{P}(\mathbb{R}^{d+1}) - \mathbb{P}(H_\gamma \oplus \mathbb{R}\gamma^-)$ as $n \to \infty$.

We claim that $\mathbb{P}(H_\gamma \oplus \mathbb{R}\cdot\gamma^-) \cap \overline{\Omega} = \{\gamma^-\}$. If the claim is false, pick $v \in \mathbb{P}(H_\gamma \oplus \mathbb{R}\cdot\gamma^-) \cap \overline{\Omega}$ such that $v \neq \gamma^-$. Up to passing to a subsequence, we can assume that $v_\infty = \lim_{n\to\infty} \gamma^n v$ exists in $\overline{\Omega}$.

Since $v \in \mathbb{P}(H_\gamma \oplus \mathbb{R}\gamma^-) - \{\gamma^-\}$, similar reasoning as in part (iv) implies that $v_\infty \in \mathbb{P}(H_\gamma)$. Thus $v_\infty \in \overline{\Omega} \cap \mathbb{P}(H_\gamma)$. Again, as in part (iv), Brouwer's fixed point theorem will imply the existence of a fixed point of γ in $\overline{\Omega} \cap \mathbb{P}(H_\gamma)$ which is distinct from $\gamma^\pm$. This contradicts Proposition 6.5 part (iii). This finishes the proof of the claim.

By the claim and the first paragraph of the proof, γ^n converges to the constant map γ^+ locally uniformly on compact subsets of $\overline{\Omega} - \{\gamma^-\}$ as $n \to \infty$. The proof for γ^{-n} follows by similar reasoning. □

Now we prove a simpler characterization of rank-one isometries for cocompact actions.

Proposition 6.8 *Suppose Ω is a Hilbert geometry, $\Gamma \leq \mathrm{Aut}(\Omega)$ is a discrete subgroup that acts cocompactly on Ω and $\gamma \in \Gamma$, where $\tau_\Omega(\gamma) > 0$. If $\gamma \in \Gamma$ has an axis, then the following are equivalent*:

(i) *γ is biproximal.*

(ii) *None of the axes of γ are contained in a half triangle in Ω.*

(iii) *γ is a rank-one isometry.*

Proof Note that (ii) $\iff$ (iii) is by definition (see Definition 6.3), and (iii) $\implies$ (i) is Proposition 6.5 part (i). We will prove (i) $\implies$ (ii) under the assumption that Ω/Γ is compact.

Let (a, b) be the axis of γ with $a \in E_\gamma^+$ and $b \in E_\gamma^-$. We first show that γ has no other fixed points in $\partial\Omega$ except a and b. If this is not true, let v be such a fixed point of γ. Since γ is biproximal, $v \notin E_\gamma^+ \cup E_\gamma^-$. Then Lemma 5.13 implies that

$$[a, v] \cup [v, b] \subset \partial\Omega. \tag{7}$$

Since $(a, b) \subset \Omega$, $\mathrm{ConvHull}_\Omega\{a, v, b\}$ is a nonempty set.

Let $A_\gamma := \langle\gamma\rangle$. Recall the notation $\mathrm{Min}_\Omega(A_\gamma) = \bigcap_{h \in A_\gamma}\{x \in \Omega \mid \mathrm{d}_\Omega(x, h \cdot x) = \tau_\Omega(h)\}$ from Section 3.9. Lemma 3.17 implies that

$$\mathrm{ConvHull}_\Omega\{a, v, b\} \subset \mathrm{Min}_\Omega(A_\gamma). \tag{8}$$

The group Γ acts cocompactly on Ω. Then, Theorem 3.19 implies that $C_\Gamma(A_\gamma)$ acts cocompactly on $\mathrm{ConvHull}_\Omega(\mathrm{Min}_\Omega(A_\gamma))$. Fix $p \in (a, b)$ and choose $v_n \in [p, v)$ such that $\lim_{n\to\infty} v_n = v$. By equation (8), $v_n \in \mathrm{Min}_\Omega(A_\gamma)$. Then there exists $h_n \in C_\Gamma(A_\gamma)$ such that $q := \lim_{n\to\infty} h_n v_n$ exists in Ω. Thus $\lim_{n\to\infty} \mathrm{d}_\Omega(h_n^{-1}q, v_n) = 0$. Then Proposition 3.12 implies that, up to passing to a subsequence,

$$\lim_{n\to\infty} h_n^{-1} q = \lim_{n\to\infty} v_n = v.$$

Pick a point $q' \in (a, b)$. Up to passing to a subsequence, $v' := \lim_{n\to\infty} h_n^{-1}q'$ exists in $\overline{\Omega}$. Since $\lim_{n\to\infty} \mathrm{d}_\Omega(h_n^{-1}q, h_n^{-1}q') = \mathrm{d}_\Omega(q, q')$, Proposition 3.12 implies that $v \in F_\Omega(v')$. Now we show that $v' \in \{a, b\}$. Since $h_n \in C_\Gamma(A_\gamma)$, $h_n(a, b)$ is an axis of γ. As γ is biproximal and has an axis, Corollary 5.4 implies that $h_n(a, b) = (a, b)$. Then, since $q' \in (a, b)$, we get $v' = \lim_{n\to\infty} h_n^{-1}q' \in \{a, b\}$. Hence

$$v \in F_\Omega(a) \cup F_\Omega(b).$$

If possible, let $v \in F_\Omega(a)$. By equation (7), $[a,v] \cup [v,b] \subset \partial\Omega$. Now, by Proposition 3.11 part (iv), $v \in F_\Omega(a)$ and $[v,b] \subset \partial\Omega$ implies that $[a,b] \subset \partial\Omega$. This is a contradiction as $(a,b) \subset \Omega$. Thus, $v \notin F_\Omega(a)$. So v must be in $F_\Omega(b)$. By a similar reasoning, we now observe that $v \notin F_\Omega(b)$. Thus we have a contradiction.

So we have shown that if γ has an axis (a,b) and is biproximal, then γ has no fixed points in $\partial\Omega$ other than a and b. Then the proof of part (iv) of Proposition 6.5 goes through verbatim. Thus $(a,z) \cup (z,b) \subset \Omega$ for all $z \in \partial\Omega \setminus \{a,b\}$, that is, the axis (a,b) is not contained in any half triangle in $\partial\Omega$. This finishes the proof. □

7 Rank-one axis and thin triangles

In this section, we prove that any projective geodesic triangle in Ω with one of its sides on a rank-one axis ℓ is $\mathfrak{D}_\ell$-thin for some constant $\mathfrak{D}_\ell$.

Proposition 7.1 *Suppose Ω is a Hilbert geometry. If $\ell \subset \Omega$ is a rank-one axis, then there exists a constant $\mathfrak{D}_\ell \geq 0$ such that if $\Delta(x,y,z) := [x,y] \cup [y,z] \cup [z,x]$ is a projective geodesic triangle in Ω with $[y,z] \subset \ell$, then $\Delta(x,y,z)$ is $\mathfrak{D}_\ell$-thin.*

Remark 7.2 The thinness constant $\mathfrak{D}_\ell$ in the above theorem depends only on the axis ℓ (and not on the rank-one isometry that has ℓ as its axis).

But first let us introduce some relevant definitions and technical results that we will need.

7.1 Thin triangles

Definition 7.3 Suppose (X,d) is a geodesic metric space.

(i) A geodesic triangle T with vertices y_1, y_2, y_3 is a union of geodesics $\sigma_1 \cup \sigma_2 \cup \sigma_3$ where σ_i is a geodesic joining y_i and y_{i+1}, where the indices $i \in \{1,2,3\}$ are counted modulo 3.

(ii) A geodesic triangle $T := \sigma_1 \cup \sigma_2 \cup \sigma_3$ is called D-thin for some $D \geq 0$ if

$$\sigma_i \subset \{x \in X \mid d(x, \sigma_{i-1} \cup \sigma_{i+1}) < D\},$$

where the indices $i \in \{1,2,3\}$ are counted modulo 3.

The following is an elementary observation about thin triangles that we use later in the paper.

Observation 7.4 *Suppose (X,d) is a geodesic metric space and $T := \sigma_1 \cup \sigma_2 \cup \sigma_3$ is a geodesic triangle with vertices y_1, y_2, y_3, and each σ_i is a continuous geodesic path joining y_i and y_{i+1} (the indices $i \in \{1,2,3\}$ are counted modulo 3). If T is D-thin, then there exist $x_i \in \sigma_i$ for $i = 1,2,3$ such that* $\max\{d(x_1,x_2), d(x_1,x_3)\} \leq D$.

Proof By slight abuse of notation, let $\sigma_1\colon [0,b] \to X$ denote the continuous parametrization of the geodesic path σ_1 for some $b \geq 0$. Without loss of generality, we assume that $\sigma_1(0) = y_1$. Since T is D-thin,

$$\sigma_1([0,b]) \subset \{x \in X \mid d(x, \sigma_2 \cup \sigma_3) < D\}. \tag{9}$$

Note that $d(\sigma_1(0), \sigma_3) = 0$ as $y_1 \in \sigma_1 \cap \sigma_3$. Let $E := \{t \in [0,b] \mid d(\sigma_1(t), \sigma_3) < D\}$. Then $0 \in E$ and $s_0 := \sup E$ exists. We can find a sequence $\{t_n\}$ in E such that $t_n \to s_0$. Then, by continuity of σ_1,

$$d(\sigma_1(s_0), \sigma_3) = \lim_{t_n \to s_0} d(\sigma_1(t_n), \sigma_3) \leq D.$$

Now note that $d(\sigma_1(s_0), \sigma_2) \leq D$. Indeed, if $t > s_0$, then $d(\sigma_1(t), \sigma_3) \geq D$ by definition of s_0. Then equation (9) implies that $d(\sigma_1(t), \sigma_2) < D$. By continuity of σ_1,

$$d(\sigma_1(s_0), \sigma_2) = \lim_{t \to s_0^+} d(\sigma_1(t), \sigma_2) \leq D.$$

Then set $x_1 := \sigma_1(s_0)$ and for $i = 2, 3$, let $x_i \in \sigma_i$ be such that $d(x_1, x_i) = d(x_1, \sigma_i)$. □

Suppose $(\Omega, \mathrm{d}_\Omega)$ is a Hilbert geometry. Then there are some special geodesic triangles in Ω, namely the ones whose edges are projective geodesic segments.

Definition 7.5 If $v_1, v_2, v_3 \in \Omega$, a projective geodesic triangle (with vertices v_1, v_2 and v_3) is

$$\Delta(v_1, v_2, v_3) := [v_1, v_2] \cup [v_2, v_3] \cup [v_3, v_1].$$

We will say that $\Delta(v_1, v_2, v_3)$ is D-thin if it is D-thin in the sense of Definition 7.3. There is a simple criterion to determine whether a projective geodesic triangle is D-thin. This proof comes from [39], and we include it here for the convenience of the reader.

Lemma 7.6 *Suppose Ω is a Hilbert geometry, $R \geq 0$ and $\Delta(x, y, z)$ is projective geodesic triangle such that $[y, z] \subset \mathcal{N}_R([x, y] \cup [x, z])$. Then $\Delta(x, y, z)$ is $(2R)$-thin.*

Proof Since $[y, z] \subset \mathcal{N}_R([x, y] \cup [x, z])$, there exist $m_{yz} \in [y, z]$, $m_{xy} \in [x, y]$ and $m_{xz} \in [x, z]$ such that $\mathrm{d}_\Omega(m_{yz}, m_{xy}) \leq R$ and $\mathrm{d}_\Omega(m_{yz}, m_{xz}) \leq R$. Indeed, the existence of three such points follows by a similar reasoning as in the proof of Observation 7.4. Then, by Proposition 3.13,

$$\begin{aligned} \mathrm{d}_\Omega^{\mathrm{Haus}}([y, m_{yz}], [y, m_{xy}]) &\leq R, \\ \mathrm{d}_\Omega^{\mathrm{Haus}}([z, m_{yz}], [z, m_{xz}]) &\leq R, \\ \mathrm{d}_\Omega^{\mathrm{Haus}}([x, m_{xy}], [x, m_{xz}]) &\leq 2R. \end{aligned}$$

Hence, $\Delta(x, y, z)$ is $(2R)$-thin. □

7.2 Proof of Proposition 7.1

Now we prove Proposition 7.1. Fix a Hilbert geometry Ω and a rank-one axis $\ell \subset \Omega$. The remark following Proposition 7.1 will be a consequence of the proof — the proof only uses the fact that there is some rank-one isometry γ that acts along ℓ by a translation; it does not rely on γ in any other manner. Lemma 7.6 reduces Proposition 7.1 to the following.

Proposition 7.7 *If $\ell \subset \Omega$ is a rank-one axis, then there exists a constant $\mathcal{B}_\ell$ with the following property: if $\Delta(x, y, z)$ is an projective geodesic triangle in Ω with $[y, z] \subset \ell$, then $[y, z] \subset \mathcal{N}_{\mathcal{B}_\ell}([x, y] \cup [x, z])$. Moreover, this constant $\mathcal{B}_\ell$ depends only on the rank-one axis ℓ (and not on the rank-one isometry whose axis is ℓ).*

Proof The "moreover" statement will again follow from the proof since the proof is independent of the choice of the rank-one isometry which has ℓ as its axis. Now we begin the proof of the first part.

If the claim is false, then for each $n \geq 0$, there exists a projective geodesic triangle $\Delta(a_n, b_n, c_n) \subset \Omega$ with $[a_n, b_n] \subset \ell$, $c_n \in \Omega$ and $e_n \in (a_n, b_n)$ such that

$$\mathrm{d}_\Omega(e_n, [c_n, a_n] \cup [c_n, b_n]) \geq n.$$

Since ℓ is a rank-one axis, there exists a rank-one isometry γ' whose axis is ℓ. Thus, translating $\Delta(a_n, b_n, c_n)$ by elements in $\langle \gamma' \rangle$ and passing to a subsequence, we can assume that $e := \lim_{n\to\infty} e_n$ exists and $e \in \ell$. Up to passing to a subsequence, we can assume that $a := \lim_{n\to\infty} a_n$, $b := \lim_{n\to\infty} b_n$ and $c := \lim_{n\to\infty} c_n$ exist. Observe that

$$\mathrm{d}_\Omega(e, [a, c] \cup [c, b]) = \lim_{n\to\infty} \mathrm{d}_\Omega(e_n, [a_n, c_n] \cup [c_n, b_n]) \geq \lim_{n\to\infty} n = \infty.$$

Thus $[a, c] \cup [c, b] \subset \partial\Omega$. But $(a, b) \subset \Omega$ since $e \in (a, b) \cap \Omega$. Thus a, c and b form a half triangle in Ω. But since $[a_n, b_n] \subset \ell$, $[a, b] \subset \overline{\ell}$. Since $a, b \in \partial\Omega$ and $\ell \subset \Omega$, we get $\overline{\ell} = [a, b]$. Thus $\ell = (a, b)$. So the rank-one axis ℓ is contained in a half triangle in Ω, a contradiction. □

8 Rank-one Hilbert geometry: Zariski density and limit sets

Recall the definition of rank-one geodesics from Definition 6.2. In this section we would like to address the following question.

Question 8.1 *Suppose (Ω, Γ) is a Hilbert geometry and Ω contains a rank-one geodesic. When does this imply that (Ω, Γ) is a rank-one Hilbert geometry?*

It is a natural question that aims to understand how the geometry of a properly convex domain influences the algebraic properties of a "large" group acting on it. Under certain assumptions, Zimmer answers this question in [50].

Proposition 8.2 *Suppose Ω is an irreducible Hilbert geometry and $\Gamma \leq \mathrm{Aut}(\Omega)$ acts cocompactly on Ω. Then (Ω, Γ) is a rank-one Hilbert geometry if and only if Ω contains a rank-one geodesic.*

This is immediate from Theorem 1.5. So our main goal in this section is to answer Question 8.1 without the assumptions of irreducibility or cocompactness as above. Instead, we will work with groups that satisfy the following assumption.

Assumption $\Gamma \leq \mathrm{SL}_{d+1}(\mathbb{R})$ is a discrete Zariski dense subgroup of $\mathrm{SL}_{d+1}(\mathbb{R})$ and there exists a properly convex domain $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ such that $\Gamma \cdot \Omega = \Omega$.

In this assumption, Zariski density may be interpreted as an assurance that the group Γ is "large". We will work with $\mathrm{SL}_{d+1}(\mathbb{R})$ in this section instead of $\mathrm{PGL}_{d+1}(\mathbb{R})$. Indeed, given $\Gamma \leq \mathrm{PGL}_{d+1}(\mathbb{R})$, we can pass to a subgroup of index at most 2 and assume that $\Gamma \leq \mathrm{SL}_{d+1}(\mathbb{R})$. In Section 8.2, we will formulate a hypothesis on the proximal limit set $\Lambda_\Gamma^{G/Q}$ (see Definition 8.3), that we call Hypothesis $(\star)$, and use it to provide an answer to Question 8.1.

Notation For the rest of this section, let $G := \mathrm{SL}_{d+1}(\mathbb{R})$, let $P \leq G$ be the subgroup of upper-triangular matrices and Q be the stabilizer in G of $[e_1] = [1 : 0 : \dots : 0] \in \mathbb{P}(\mathbb{R}^{d+1})$. Fix the standard inner product on $\mathbb{R}^{d+1}$ and let $K := \mathrm{SO}(d+1)$.

Let ε_i be the evaluation of the i^{th} diagonal entry of a diagonal matrix. Take $\Delta := \{\varepsilon_i - \varepsilon_{i+1} \mid 1 \leq i \leq d\}$ to be the set of positive simple roots. For any $\theta = \{\varepsilon_{i_1} - \varepsilon_{i_1+1}, \dots, \varepsilon_{i_k} - \varepsilon_{i_k+1}\} \subset \Delta$, let P_θ denote the subgroup of block upper-triangular matrices in G with square diagonal blocks of sizes $i_1, i_2 - i_1, \dots, i_k - i_{k-1}$, $d - i_k$, respectively. In particular, $P_\Delta = P$ and G/P is the full flag variety, while $P_{\{\varepsilon_1 - \varepsilon_2\}} = Q$ and $G/Q \cong \mathbb{P}(\mathbb{R}^{d+1})$.

8.1 Limit sets in flag varieties

We will require the notion of limit sets of discrete subgroups of G in flag varieties, in particular G/P and G/Q. This has been defined and studied by various authors in different degrees of generality: Guivarch [35] (for subgroups of $\mathrm{SL}_{d+1}(\mathbb{R})$ acting proximally and strongly irreducibly on $\mathbb{R}^d$), Benoist [6] (for Zariski dense subgroups of reductive groups) and Guéritaud, Guichard, Kassel and Wienhard [34] (for arbitrary subgroups of reductive groups). We use the definition from [34, Section 5.1]

First recall the notion of singular value decomposition (or more generally, Cartan decomposition in G): for any $g \in G$, there exist $k_1, k_2 \in \mathrm{SO}(d+1)$ and $A_g = \mathrm{diag}(a_1, \dots, a_{d+1})$ with $a_1 \geq \dots \geq a_{d+1} > 0$ such that

$$g = k_1 A_g k_2.$$

The Cartan decomposition defines the Cartan projection $\mu(g) := \mathrm{diag}(\log(a_1), \dots, \log(a_{d+1}))$. It maps G into the space of trace-zero diagonal matrices of size $(d+1) \times (d+1)$.

Let $\theta \subset \Delta$. If $g \in G$ has a singular value decomposition $g = k_1 A_g k_2$, define $E_\theta\colon G \to G/P_\theta$ by

$$E_\theta(g) := k_1 \cdot eP_\theta.$$

The map E_θ does not depend on the choices of k_1 and k_2, provided $\alpha(\mu(g)) > 0$ for all $\alpha \in \theta$; see [34, Section 5.1].

Definition 8.3 [34, Definition 5.1] Suppose Γ_0 is a discrete subgroup of G. The limit set $\Lambda_{\Gamma_0}^{G/P_\theta}$ of Γ_0 in G/P_θ is defined to be the set of all accumulation points of sequences $\{E_\theta(\gamma_n)\}_{n\in\mathbb{N}}$ where $\{\gamma_n\}_{n\in\mathbb{N}}$ is any sequence in Γ_0 such that $\alpha(\mu(\gamma_n)) \to \infty$ for all $\alpha \in \theta$.

Remark 8.4 Suppose Γ_0 is Zariski dense in G.

(i) Then $\Lambda_{\Gamma_0}^{G/P_\theta}$ is nonempty and is the closure of the set of attracting fixed points of proximal elements in G/P_θ; see [6] and [34, Section 5.1]. Here, an element $g \in G$ is called *proximal*[1] in G/P_θ provided $\alpha(\mu(g)) > 0$ for all $\alpha \in \theta$. Moreover, g is proximal in G/P_θ if and only if g has a unique attracting fixed point[2] in G/P_θ; see [34, Definition 2.25].

(ii) Suppose $\theta = \{\varepsilon_1 - \varepsilon_2\}$ so that $P_\theta = Q$. Then $\Lambda_{\Gamma_0}^{G/Q}$ is the unique minimal closed Γ-invariant subset of G/Q; see [13, Lemma 4.2]. This may not be true for arbitrary choices of θ; see [13, Remark 4.4].

Lemma 8.5 *Suppose $\Gamma \le \mathrm{SL}_{d+1}(\mathbb{R})$ satisfies the assumption. Then $\Lambda_\Gamma^{G/Q} \neq \varnothing$ and $\Lambda_\Gamma^{G/Q} \subset \partial\Omega$ is the unique minimal closed Γ-invariant subset of $\partial\Omega$.*

Proof Note that $\partial\Omega$ is a closed Γ-invariant set and the unique attracting fixed point of any proximal element (in G/Q) of Γ lies in $\partial\Omega$. The lemma then follows from Remark 8.4 above. $\square$

If we do not assume Zariski density, then we may still have nonempty limit set (in an appropriate G/P_θ) but with some unusual properties. The following is such an example.

Example 8.6 Consider the discrete subgroup $\Gamma' := \left\{\mathrm{diag}(2^{m_1}, \dots, 2^{m_{d+1}}) \mid \sum_{i=1}^{d+1} m_i = 0\right\}$ of $\mathrm{Aut}(T_d)$ and d-dimensional torus T_d/Γ'. Although Γ' is not Zariski dense in $\mathrm{SL}_{d+1}(\mathbb{R})$, the proximal limit set in $\mathbb{P}(\mathbb{R}^{d+1})$ is nonempty and in fact $\Lambda_{\Gamma'}^{G/Q} = \{[e_1], \dots, [e_{d+1}]\}$. Thus $\Lambda_{\Gamma'}^{G/Q}$ is a proper subset of ∂T_d. Note that (T_d, Γ') is not a rank-one Hilbert geometry; see Remark 6.4.

In the light of Lemma 8.5 and this example, it is natural to ask when does $\Lambda_\Gamma^{G/Q}$ equal $\partial\Omega$.

Remark 8.7 In general, if Γ only satisfies the assumption, then $\Lambda_\Gamma^{G/Q}$ can be a proper subset of $\partial\Omega$. For example, let $\Gamma \le \mathrm{PO}(2,1)$ be a Zariski dense convex cocompact Kleinian group. Then $\Lambda_\Gamma^{G/Q} = \overline{\Gamma \cdot x} \cap \partial\mathbb{H}^2$, where $x \in \mathbb{H}^2$. Unless Γ is cocompact, $\Lambda_\Gamma^{G/Q} \neq \partial\mathbb{H}^2$. However, under the additional cocompactness assumption, we often have equality. Blayac [19, Theorem 1.3] has recently shown that if (Ω, Γ) is a divisible rank-one Hilbert geometry, then $\Lambda_\Gamma^{G/Q} = \partial\Omega$.

[1] This coincides with the notion of proximality discussed in Section 3.11 when $\theta = \{\varepsilon_1 - \varepsilon_2\}$.

[2] A fixed point $x \in X$ of a smooth map $f : X \to X$ is attracting if $\|Df_x\| < 1$.

8.2 Hypothesis ($\star$) and an answer to Question 8.1

We now introduce a special hypothesis under which we can answer Question 8.1.

Hypothesis ($\star$) *Suppose $\Gamma \le G$ is a discrete subgroup and $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is a properly convex domain such that $\Gamma \cdot \Omega = \Omega$. We will say that (Ω, Γ) satisfies Hypothesis ($\star$) if there exists a rank-one geodesic $(a', b') \subset \Omega$ with its endpoints $a', b' \in \Lambda_\Gamma^{G/Q} \cap \partial\Omega$.*

We will show that for any Zariski dense discrete subgroup Γ, this hypothesis is equivalent to the rank one property. One implication is easy and does not require Zariski density.

Lemma 8.8 *Suppose $\Gamma \le G$ is a discrete subgroup that preserves a properly convex domain Ω and (Ω, Γ) is a rank-one Hilbert geometry. Then (Ω, Γ) satisfies Hypothesis ($\star$).*

Remark 8.9 In this lemma, we do not assume that Γ is Zariski dense in G.

Proof Since (Ω, Γ) is a rank-one Hilbert geometry, we can find a rank-one isometry $\gamma \in \Gamma$. Let $\gamma^\pm \in \partial\Omega$ be the attracting and the repelling fixed points of γ. Then $\gamma^\pm \in \Lambda_\Gamma^{G/Q}$ by definition of $\Lambda_\Gamma^{G/Q}$. Also (γ^+, γ^-) is the axis of γ and hence a rank-one geodesic; see Definition 6.3 and Proposition 6.5. □

Next we will seek a converse of the above lemma and this will require the Zariski density assumption on Γ. But first we recall the notion of loxodromic elements. We will call $g \in G$ *loxodromic* if

$$|\lambda_1(g)| > \cdots > |\lambda_{d+1}(g)|.$$

If g is loxodromic, then it has unique attracting fixed points in both G/Q and G/P. We will denote by $a_g^\pm \in G/Q$ (resp. $\mathbb{X}_g^\pm \in G/P$) the unique attracting fixed point of $g^{\pm 1}$ in G/Q (resp. G/P). With this notation, $\Pi_{PQ}(\mathbb{X}_g^\pm) = a_g^\pm$, where

$$\Pi_{PQ} \colon G/P \to G/Q$$

is the natural smooth projection map. Also recall that if $g \in \mathrm{Aut}(\Omega)$ is a rank-one isometry, then we denote by $g^\pm$ the attracting and the repelling fixed points of g; see Remark 6.6.

Lemma 8.10 *Suppose $\Gamma \le G$ satisfies the assumption. If there exists a rank-one geodesic $(a', b') \subset \Omega$ with its endpoints $a', b' \in \Lambda_\Gamma^{G/Q} \cap \partial\Omega$, then*

- (i) *there exist rank-one isometries $\{g_n\}$ in Γ such that $\lim_{n\to\infty} g_n^+ = a'$ and $\lim_{n\to\infty} g_n^- = b'$,*
- (ii) *(Ω, Γ) is a rank-one Hilbert geometry,*
- (iii) *the set of rank-one isometries in Γ is Zariski dense in G,*
- (iv) *$\Lambda_\Gamma^{G/Q} = \overline{\{\gamma^+ \mid \gamma \text{ is a rank-one isometry}\}}$.*

Corollary 8.11 *Suppose $\Gamma \leq G$ satisfies the assumption, and (Ω, Γ) satisfies Hypothesis $(\star)$. Then (Ω, Γ) is a rank-one Hilbert geometry.*

Proof of Lemma 8.10 The key idea of this proof is in [46] and it relies on results of Benoist [7]. Before starting the proof, we informally outline the main idea. The key technical point is to find a sequence $\{g_n\}$ of biproximal elements in Γ such that $a_{g_n^+} \to a'$ and $a_{g_n}^- \to b'$. A direct way to find such a $\{g_n\}$ is: using Zariski density, find a pair $g, h \in \Gamma$ of transversally biproximal elements [13, Chapter 7] such that a_g^+ and a_h^- are arbitrarily close to a' and b', respectively. Then, for large enough n, $g^n h^n$ is a biproximal element whose attracting and repelling fixed points are close to a' and b'. However, in this proof, we will take a more indirect approach by passing to the limit set in G/P and using a result of Benoist. We rely on [7, Lemma 2.6(c)]: given two distinct points $\mathbb{X}_+, \mathbb{X}_- \in \Lambda_\Gamma^{G/P}$, there exist loxodromic elements $g_n \in \Gamma$ such that $\mathbb{X}_{g_n}^{\pm} \to \mathbb{X}_{\pm}$. Once we have this sequence $\{g_n\}$, Claim 8.11.1 implies that all but finitely many of them are rank-one isometries.

Now we begin the formal proof. Equip G/P and G/Q with K-invariant Riemannian metrics and denote the corresponding Riemannian distance functions by d_P and d_Q respectively. We remark that this specific choice of Riemannian metrics will be insignificant as G/P and G/Q are compact manifolds. Let Γ_{lox} be the set of loxodromic elements in Γ. Since Γ is Zariski dense in G, Remark 8.4 implies that $\Pi_{PQ}(\Lambda_\Gamma^{G/P}) = \Lambda_\Gamma^{G/Q}$. Then pick $\mathbb{X}_a, \mathbb{X}_b \in \Lambda_\Gamma^{G/P}$ such that $\Pi_{PQ}(\mathbb{X}_a) = a'$ and $\Pi_{PQ}(\mathbb{X}_b) = b'$. For any $\varepsilon > 0$,

$$\Gamma_\varepsilon := \{g \in \Gamma_{\mathrm{lox}} \mid \mathrm{d}_P(\mathbb{X}_g^+, \mathbb{X}_a) < \varepsilon, \mathrm{d}_P(\mathbb{X}_g^-, \mathbb{X}_b) < \varepsilon\}$$

is Zariski dense in G; see [7, Lemma 2.6(c)].

For any $g \in \Gamma_{\mathrm{lox}}$, $a_g^{\pm} = \Pi_{PQ}(\mathbb{X}_g^{\pm})$ and $a_g^{\pm} \in \partial\Omega$. Moreover, Π_{PQ} is continuous and $(a', b') \subset \Omega$. Thus there exists ε' such that if $\varepsilon \in (0, \varepsilon')$, then $(a_g^+, a_g^-) \subset \Omega$ for any $g \in \Gamma_\varepsilon$. In fact $(a_g^+, a_g^-) \subset \Omega$ is the unique axis in Ω for any such $g \in \Gamma_\varepsilon$. Indeed, the uniqueness follows from Corollary 5.4 because g has an axis $(a_g^+, a_g^-) \subset \Omega$, g is loxodromic and $\tau_\Omega(g) > 0$. We now claim that:

Claim 8.11.1 *If $\varepsilon \in (0, \varepsilon')$ is small enough, then g is a rank-one isometry for all $g \in \Gamma_\varepsilon$.*

Proof Suppose the claim is false. Then there exist a sequence $\{\varepsilon_n\}$ in $(0, \varepsilon')$ with $\varepsilon_n \to 0$ and $g_n \in \Gamma_{\varepsilon_n}$ such that g_n is not a rank-one isometry. Then $\mathbb{X}_{g_n}^+ \to \mathbb{X}_a$ and $\mathbb{X}_{g_n}^- \to \mathbb{X}_b$. Since Π_{PQ} is continuous, $a_{g_n}^+ \to a'$ and $a_{g_n}^- \to b'$.

By the paragraph before the claim, each g_n has a unique axis $(a_{g_n}^+, a_{g_n}^-) \subset \Omega$. Moreover, $(a_{g_n}^+, a_{g_n}^-) \to (a', b')$. But since g_n is not a rank-one isometry by assumption, this implies that there exists $\{c_n\}$ with $c_n \in \partial\Omega - \{a_{g_n}^+, a_{g_n}^-\}$ such that

$$[a_{g_n}^+, c_n] \cup [c_n, a_{g_n}^-] \subset \partial\Omega.$$

Up to passing to a subsequence, we can assume that $c_n \to c$ in $\partial\Omega$. Then $[a', c] \cup [c, b'] \subset \partial\Omega$ while $(a', b') \subset \Omega$. Thus (a', b') cannot be a rank-one geodesic and we have a contradiction. This finishes the proof of this claim. □

Now we finish the proof of the lemma. Let us choose an $\varepsilon \in (0, \varepsilon')$ as in the above claim.

(i) The result follows by choosing $g_n \in \Gamma_{\varepsilon/n}$ for all $n \geq 1$.

(ii) This follows from (i), since there is at least one rank-one isometry in Γ.

(iii) The set Γ_ε is a subset of the set of rank-one isometries of Γ and Γ_ε is Zariski dense.

(iv) By Lemma 8.5, $\Lambda_\Gamma^{G/Q} \subset \partial\Omega$ is a minimal, closed Γ-invariant set which contains the unique attracting fixed points of all proximal elements. Since a rank-one isometry is necessarily proximal,

$$\overline{\{\gamma^+ \mid \gamma \text{ is a rank-one isometry}\}} \subset \Lambda_\Gamma^{G/Q}.$$

Since $\overline{\{\gamma^+ \mid \gamma \text{ is a rank-one isometry}\}}$ is a closed Γ-invariant set, the equality then follows from minimality of $\Lambda_\Gamma^{G/Q}$. □

We now observe that Hypothesis $(\star)$ gives:

Answer to Question 8.1 (see Lemma 8.8 and Corollary 8.11) If $\Gamma \leq \mathrm{SL}_{d+1}(\mathbb{R})$ is a discrete Zariski dense subgroup of $\mathrm{SL}_{d+1}(\mathbb{R})$ that preserves a properly convex domain Ω, then (Ω, Γ) *is a rank-one Hilbert geometry* if and only if Ω *contains a rank-one geodesic* $(a', b') \subset \Omega$ *with* $a', b' \in \Lambda_\Gamma^{G/Q} \cap \partial\Omega$.

We finish the section with an example where Hypothesis $(\star)$ fails. Recall Example 8.6. In that case, Γ' preserves the standard d-simplex T_d, T_d/Γ' is homeomorphic to a d-torus, and T_d does not contain any rank-one geodesics. Thus (T_d, Γ') does not satisfy Hypothesis $(\star)$. However, in this example, the group Γ' is not Zariski dense in $\mathrm{SL}_{d+1}(\mathbb{R})$, and one may wonder if that is the reason why Hypothesis $(\star)$ fails. So we ask the following question.

Question 8.12 *Suppose $\Gamma \leq G$ is a discrete subgroup that preserves a properly convex domain Ω. If Γ is Zariski dense in G, then does (Ω, Γ) satisfy Hypothesis $(\star)$?*

To the best of the author's knowledge, the answer to this question is not known unless one makes other assumptions, eg say Ω/Γ is compact and Ω is irreducible. Under these assumptions, Remark 8.7 and Theorem 1.5 together provide an answer.

Part III Contracting elements in Hilbert geometry

In this part of the paper, we prove our main results: Theorems 1.2 and 1.4. The outline of this part of the paper is as follows. We recall the notion of contracting elements in Section 9. The proof of Theorem 1.2 is split into two sections: Sections 10 and 11. In Section 12, we introduce the notion of acylindrically hyperbolic groups and prove Theorem 1.4.

9 Contracting elements: definition and properties

Suppose $K \geq 1$ and $C \geq 0$. A function $F: (X, \mathrm{d}_X) \to (Y, \mathrm{d}_Y)$ is called a (K, C)-quasi-isometric embedding if for any $x_1, x_2 \in X$,

$$\frac{1}{K}\,\mathrm{d}_X(x_1, x_2) - C \leq \mathrm{d}_Y(F(x_1), F(x_2)) \leq K\,\mathrm{d}_X(x_1, x_2) + C.$$

Fix a proper geodesic metric space (X, d) and a group G that acts properly and by isometries on X. If $K \geq 1$ and $C \geq 0$, then a (K, C)-path in (X, d) is a set $F(\mathbb{R})$ where $F: (\mathbb{R}, |\cdot|) \to (X, \mathrm{d})$ is a (K, C)-quasi-isometric embedding. A subpath of the path $F(\mathbb{R})$ is $F(I)$ where $I \subset \mathbb{R}$ is an interval, possibly unbounded.

Definition 9.1 Let $K \geq 1$ and $C \geq 0$. Let $\mathcal{PS}$ be a collection of (K, C)-paths in X. Then:

(i) $\mathcal{PS}$ is called a *path system* on X if
   (a) any subpath of a path in $\mathcal{PS}$ is also in $\mathcal{PS}$, and
   (b) any pair of points in X can be connected by a path in $\mathcal{PS}$.

(ii) $\mathcal{PS}$ is called a *geodesic path system* if all paths in $\mathcal{PS}$ are geodesics in (X, d).

(iii) If G preserves $\mathcal{PS}$, then $(X, \mathcal{PS})$ is called a *path system for the group* G.

Definition 9.2 (contracting subsets [49]) If $\mathcal{PS}$ is a path system on X, then $\mathcal{A} \subset X$ is said to be *$\mathcal{PS}$-contracting* (with constant C) if there exists a map $\pi_{\mathcal{A}}: X \to \mathcal{A}$ such that

(i) if $x \in \mathcal{A}$, then $\mathrm{d}(x, \pi_{\mathcal{A}}(x)) \leq C$,

(ii) if $x, y \in X$ and $\mathrm{d}(\pi_{\mathcal{A}}(x), \pi_{\mathcal{A}}(y)) \geq C$, then for any path $\sigma \in \mathcal{PS}$ joining x and y,

$$\mathrm{d}(\sigma, \pi_{\mathcal{A}}(x)) \leq C \quad \text{and} \quad \mathrm{d}(\sigma, \pi_{\mathcal{A}}(y)) \leq C.$$

A prototypical example of a contracting subset is a bi-infinite geodesic in $\mathbb{H}^2$ (with the map $\pi_{\mathcal{A}}$ being the closest point projection on the geodesic). Generally speaking, one should think of the projection map $\pi_{\mathcal{A}}$ as an analogue of the closest-point projection. In fact, the following lemma makes this analogy concrete in the context of geodesic path systems. We will use the notation

$$\rho_{\mathcal{A}}(x) := \{a \in \mathcal{A} \mid \mathrm{d}(x, a) = \mathrm{d}(x, \mathcal{A})\}$$

for the set-valued closest-point projection map on $\mathcal{A}$.

Lemma 9.3 *Suppose $\mathcal{PS}$ is a geodesic path system and $\mathcal{A} \subset X$ is $\mathcal{PS}$-contracting (with constant C) with the projection map $\pi_{\mathcal{A}}: X \to \mathcal{A}$. Then $\sup_{a \in \rho_{\mathcal{A}}(x)} \mathrm{d}(\pi_{\mathcal{A}}(x), a) \leq 2C$ for all $x \in X$.*

Proof Suppose there exist $x \in X$ and $a \in \rho_{\mathcal{A}}(x)$ such that

$$\mathrm{d}(\pi_{\mathcal{A}}(x), a) > 2C.$$

Since $\mathscr{A}$ is $\mathcal{PS}$-contracting and $a \in \mathscr{A}$, one gets that $\mathrm{d}(\pi_{\mathscr{A}}(a), a) \leq C$. Then

$$\mathrm{d}(\pi_{\mathscr{A}}(x), \pi_{\mathscr{A}}(a) \geq \mathrm{d}(\pi_{\mathscr{A}}(x), a) - \mathrm{d}(\pi_{\mathscr{A}}(a), a) > C.$$

Let $\sigma_{x,a}$ be a geodesic path in $\mathcal{PS}$ joining x and a. Since $\mathscr{A}$ is $\mathcal{PS}$-contracting, there exists $z \in \sigma_{x,a}$ such that $\mathrm{d}(z, \pi_{\mathscr{A}}(x)) \leq C$. As $z \in \sigma_{x,a}$, $\mathrm{d}(a,z) = \mathrm{d}(a,x) - \mathrm{d}(z,x)$. As $a \in \rho_{\mathscr{A}}(x)$, $\mathrm{d}(x,a) \leq \mathrm{d}(\pi_{\mathscr{A}}(x), x)$. Then

$$\mathrm{d}(a,z) \leq \mathrm{d}(\pi_{\mathscr{A}}(x), x) - \mathrm{d}(x,z) \leq \mathrm{d}(\pi_{\mathscr{A}}(x), z) + \mathrm{d}(z,x) - \mathrm{d}(z,x) \leq C.$$

Then $\mathrm{d}(\pi_{\mathscr{A}}(x), a) \leq \mathrm{d}(\pi_{\mathscr{A}}(x), z) + \mathrm{d}(z,a) \leq 2C$, a contradiction. □

Using the notion of contracting subsets, one introduces the notion of contracting group elements. A prototypical example of a contracting element is

$$g = \begin{pmatrix} \lambda & 0 \\ 0 & \dfrac{1}{\lambda} \end{pmatrix}$$

for some $\lambda > 1$, that acts on $\mathbb{H}^2$ by a translation along a bi-infinite geodesic in $\mathbb{H}^2$.

Definition 9.4 (contracting elements [49]) If $(X, \mathcal{PS})$ is a path system for G, then $g \in G$ is a *contracting element* for $(X, \mathcal{PS})$ provided for some (hence any) $x_0 \in X$,

(i) g has infinite order and $\langle g \rangle \cdot x_0$ is a quasi-isometric embedding of $\mathbb{Z}$ in X, and

(ii) there exists $\mathscr{A} \subset X$ containing x_0 that is $\langle g \rangle$-invariant, $\mathcal{PS}$-contracting and has cobounded $\langle g \rangle$-action.

Remark 9.5 If $g \in G$ is a contracting element and $\mathcal{PS}$ is a geodesic path system, then $\pi_{\mathscr{A}}$ is coarsely $\langle g \rangle$-equivariant. This is immediate from Lemma 9.3 since $\pi_{\mathscr{A}}$ is coarsely equivalent to $\rho_{\mathscr{A}}$ and $\rho_{\mathscr{A}}$ is clearly $\langle g \rangle$-equivariant.

In the definition of a contracting element, the set $\mathscr{A}$ is not necessarily a $\langle g \rangle$-orbit in X. We will now explain that we can always replace $\mathscr{A}$ by a $\langle g \rangle$-orbit. Moreover, we also show g has positive translation length for its action on X. We remark that the following observation does not require that $\mathcal{PS}$ is a geodesic path system.

Observation 9.6 *Suppose $(X, \mathcal{PS})$ is a path system for G and $g \in G$ is a contracting element for $(X, \mathcal{PS})$. Then*

(i) *$\tau_X(g) := \inf_{x \in X} \mathrm{d}(x, gx)$ is positive, and*

(ii) *for any $x_0 \in \mathscr{A}$, $\mathscr{A}_{\min}(x_0) := \langle g \rangle x_0$ is the minimal $\mathcal{PS}$-contracting, $\langle g \rangle$-invariant subset of X containing x_0 with a cobounded $\langle g \rangle$-action.*

Proof (i) Recall the definition of stable translation length

$$\tau_X^{\text{stable}}(g) := \lim_{n\to\infty} \frac{\mathrm{d}(x, g^n x)}{n}.$$

Then $\tau_X(g) \geq \tau_X^{\text{stable}}(g)$ and it suffices to show $\tau_X^{\text{stable}}(g) > 0$. Fix any $x_0 \in X$. Since g is contracting, $\langle g\rangle x_0$ is a quasigeodesic, that is, there exists $K \geq 1$ and $C \geq 0$ such that $\mathrm{d}(x_0, g^n x_0) \geq (1/K)|n| - C$ for every $n \in \mathbb{Z}$. Then, $\tau_X^{\text{stable}}(g) \geq 1/K > 0$.

(ii) Let $\mathcal{A}$ be $\mathcal{PS}$-contracting with constant $C_{\mathcal{A}}$ and the map $\pi_{\mathcal{A}}\colon X \to \mathcal{A}$. Fix any $x_0 \in \mathcal{A}$ and set $R_{\mathcal{A}} := \operatorname{diam}(\mathcal{A}/\langle g\rangle)$, $C_0 := C_{\mathcal{A}} + 2R_{\mathcal{A}}$ and $\mathcal{A}_{\min}(x_0) := \langle g\rangle x_0$.

Since $\mathcal{A}_{\min}(x_0) \subset \mathcal{A}$, if $x \in X$, then there exists $m \in \mathbb{Z}$ such that $\mathrm{d}(\pi_{\mathcal{A}}(x), g^m x_0) \leq R_{\mathcal{A}}$. Define $\pi_{\min}\colon X \to \mathcal{A}_{\min}(x_0)$ by setting $\pi_{\min}(x) = g^m x_0$. Then, if $x \in \mathcal{A}_{\min}(x_0)$, $\pi_{\min}(x) = x$. If $x, y \in X$ and $\mathrm{d}(\pi_{\min}(x), \pi_{\min}(y)) \geq C_0$, then $\mathrm{d}(\pi_{\mathcal{A}}(x), \pi_{\mathcal{A}}(y)) \geq C_{\mathcal{A}}$. Thus, if $\sigma \in \mathcal{PS}$ is a path from x to y, $\mathrm{d}(\pi_{\mathcal{A}}(x), \sigma) \leq C_{\mathcal{A}}$ and $\mathrm{d}(\pi_{\mathcal{A}}(y), \sigma) \leq C_{\mathcal{A}}$. Hence,

$$\mathrm{d}(\pi_{\min}(x), \sigma) \leq C_0 \quad \text{and} \quad \mathrm{d}(\pi_{\min}(y), \sigma) \leq C_0. \qquad \square$$

There are many other notions of contracting subsets in geometric group theory. We will require one such notion in Section 14 for proving our Theorem 1.8. We will call this notion *contraction in the sense of BF* — it was introduced by Bestvina and Fujiwara for CAT(0) spaces [17] and by Gekhtman and Yang [32] in general. We defer all further discussion about this to Appendix B and only remark that in our case, this notion of contraction will be equivalent to Definition 9.2.

Remark 9.7 If Ω is a Hilbert geometry, we use the geodesic path system $\mathcal{PS}^{\Omega} := \{[x, y] \mid x, y \in \Omega\}$ given by projective geodesics. We use the $\mathcal{PS}$-contracting notion everywhere in the paper except in Section 14 (where we use *contraction in the sense of BF*, see Definition B.2). Proposition 9.8 below implies that these two notions of contraction are equivalent in our setup. Hence, in the rest of the paper, we will use the term contracting subset (and element) without additional labels.

Proposition 9.8 *Suppose $(X, \mathcal{PS})$ is a geodesic path system. Then*:

(i) *$\mathcal{A} \subset X$ is $\mathcal{PS}$-contracting if and only if $\mathcal{A}$ is contracting in the sense of BF.*

(ii) *If G preserves $\mathcal{PS}$, then $g \in G$ is a contracting element for $(X, \mathcal{PS})$ if and only if $g \in G$ is a contracting element in the sense of BF.*

Proof See Appendix B. $\square$

10 Rank-one isometries are contracting

In this section, we prove one implication in Theorem 1.2. Fix a Hilbert geometry Ω and let $\mathcal{PS}^{\Omega} := \{[x, y] \mid x, y \in \Omega\}$.

Theorem 10.1 *If $\gamma \in \mathrm{Aut}(\Omega)$ is a rank-one isometry, then γ is a contracting element for $(\Omega, \mathcal{PS}^{\Omega})$.*

The key step will be part (ii) of Lemma 10.3, which shows that a rank-one axis is $\mathcal{PS}^{\Omega}$-contracting. First, we construct suitable projection maps on a rank-one axis. Recall the notion of closest-point projection on closed convex subsets, particularly Corollary 3.10.

Definition 10.2 Suppose Ω is a Hilbert geometry, ℓ is a bi-infinite projective geodesic in Ω and $\sigma : \mathbb{R} \to \ell$ is its unit-speed parametrization. Then $\Pi_\ell(x) = [\sigma(T_x^-), \sigma(T_x^+)]$ for $T_x^-, T_x^+ \in \mathbb{R}$. We define the projection map $\pi_\ell : \Omega \to \ell$ as

$$\pi_\ell(x) := \sigma\left(\frac{T_x^- + T_x^+}{2}\right).$$

We now establish some properties of the map π_ℓ when ℓ is a rank-one axis.

Lemma 10.3 *If $\ell \subset \Omega$ is a rank-one axis, then there exists $\mathcal{C}_\ell \geq 0$ such that:*

(i) *If $x \in \Omega$ and $z \in \ell$, then there exists $p_{xz} \in [x, z]$ such that*

$$\mathrm{d}_\Omega(\pi_\ell(x), p_{xz}) \leq 3\mathcal{C}_\ell.$$

(ii) *The geodesic ℓ is $\mathcal{PS}^{\Omega}$-contracting with constant $\mathcal{C}_\ell$ and the map π_ℓ.*

Proof (i) Let $x \in \Omega$ and $z \in \ell$. Choose any $C_\ell \geq D_\ell$, where D_ℓ is the constant from Proposition 7.1. Proposition 7.1 implies that $\Delta(x, \pi_\ell(x), z)$ is $\mathcal{D}_\ell$-thin. By Observation 7.4, there exists $p \in [x, \pi_\ell(x)]$, $q \in [\pi_\ell(x), z]$ and $r \in [z, x]$ such that

$$\mathrm{d}_\Omega(q, p) \leq \mathcal{D}_\ell \quad \text{and} \quad \mathrm{d}_\Omega(q, r) \leq \mathcal{D}_\ell.$$

Then

$$\mathrm{d}_\Omega(\pi_\ell(x), p) = \mathrm{d}_\Omega(\pi_\ell(x), x) - \mathrm{d}_\Omega(p, x) \leq \mathrm{d}_\Omega(q, x) - \mathrm{d}_\Omega(p, x) \leq \mathrm{d}_\Omega(p, q) \leq \mathcal{D}_\ell.$$

Thus

$$\mathrm{d}_\Omega(\pi_\ell(x), q) \leq \mathrm{d}_\Omega(\pi_\ell(x), p) + \mathrm{d}_\Omega(q, p) \leq 2\mathcal{D}_\ell.$$

Set $p_{xz} := r$. Then

$$\mathrm{d}_\Omega(\pi_\ell(x), p_{xz}) \leq \mathrm{d}_\Omega(\pi_\ell(x), q) + \mathrm{d}_\Omega(q, r) \leq 3\mathcal{D}_\ell \leq 3C_\ell.$$

(ii) Set $\pi := \pi_\ell$ for ease of notation. Let us label the endpoints of ℓ so that $\ell := (a, b)$. Observe that we only need to verify (ii) in Definition 9.2. Suppose, for a contradiction, that it is not satisfied. Then, for every $n \in \mathbb{N}$, there exist $x_n, y_n \in \Omega$ such that

$$\mathrm{d}_\Omega(\pi(x_n), \pi(y_n)) \geq n \quad \text{and} \quad \mathrm{d}_\Omega([x_n, y_n], \pi(x_n)) \geq n.$$

Since ℓ is a rank-one axis, fix a rank-one isometry γ whose axis is ℓ. Then $\gamma \circ \pi = \pi \circ \gamma$. Hence, up to translating x_n and y_n using elements in $\langle \gamma \rangle$, we can assume that $\alpha := \lim_{n \to \infty} \pi(x_n)$ exists in $\ell \subset \Omega$.

Up to passing to a subsequence, we can further assume that the following limits exist in $\overline{\Omega}$:

$$x := \lim_{n\to\infty} x_n, \quad y := \lim_{n\to\infty} y_n, \quad \beta := \lim_{n\to\infty} \pi(y_n).$$

Then $\lim_{n\to\infty}[x_n, y_n] = [x, y]$. We will now show that

$$[x, y] \subset \partial\Omega. \tag{10}$$

This follows from the following estimate:

$$\begin{aligned} d_\Omega(\alpha, [x, y]) = \lim_{n\to\infty} d_\Omega(\alpha, [x_n, y_n]) &\geq \lim_{n\to\infty} \big(d_\Omega(\pi(x_n), [x_n, y_n]) - d_\Omega(\pi(x_n), \alpha)\big) \\ &\geq \lim_{n\to\infty} \big(n - d_\Omega(\pi(x_n), \alpha)\big) = \infty. \end{aligned}$$

We also observe that

$$\begin{aligned} d_\Omega(\alpha, \beta) = \lim_{n\to\infty} d_\Omega(\alpha, \pi(y_n)) &\geq \lim_{n\to\infty} \big(d_\Omega(\pi(x_n), \pi(y_n)) - d_\Omega(\pi(x_n), \alpha)\big) \\ &\geq \lim_{n\to\infty} \big(n - d_\Omega(\pi(x_n), \alpha)\big) = \infty. \end{aligned}$$

Thus $\beta \in \partial\Omega$. However, since $\beta \in \overline{\ell} = [a, b]$, $\beta \in \{a, b\}$. Thus, up to switching the labels of the endpoints of ℓ, we can assume that

$$\beta = b. \tag{11}$$

Claim 10.3.1 $$x = y = b.$$

Proof We first show that $y = b$. Since $y_n \in \Omega$ and $\alpha \in \ell$, part (i) of Lemma 10.3 implies that there exists $p_n \in [y_n, \alpha]$ such that

$$d_\Omega(p_n, \pi(y_n)) \leq 3\mathscr{C}_\ell.$$

Up to passing to a subsequence, we can assume that $p := \lim_{n\to\infty} p_n$ exists in $\overline{\Omega}$. Then by Proposition 3.12, $p \in F_\Omega(\beta)$. By equation (11), $\beta = b$, which implies $p \in F_\Omega(b)$. Since b is an endpoint of the rank-one axis ℓ, part (iv) of Proposition 6.5 implies that $F_\Omega(b) = b$. Thus $p = b$. On the other hand, since $p_n \in [y_n, \alpha]$, we have $p \in [y, \alpha]$. Since $p = b$, we have $p \in \partial\Omega$. Thus,

$$p \in [\alpha, y] \cap \partial\Omega = \{y\}.$$

Hence,

$$y = p = b.$$

We now show that $x = b$. By equation (10), $[x, y] \subset \partial\Omega$. But since $y = b$, this contradicts part (iv) of Proposition 6.5 unless $x = y$. Hence $x = y = b$. This concludes the proof of Claim 10.3.1. □

Consider the points $x_n \in \Omega$ and $\pi(y_n) \in \ell$. By part (i) of Lemma 10.3, there exists $q_n \in [x_n, \pi(y_n)]$ such that $d_\Omega(\pi(x_n), q_n) \leq 3\mathscr{C}_\ell$. Up to passing to a subsequence, we can assume that $q := \lim_{n\to\infty} q_n$ exists in $\overline{\Omega}$. Then by Proposition 3.12, $q \in F_\Omega(\alpha) = \Omega$. Thus $\lim_{n\to\infty}[x_n, \pi(y_n)]$ is a projective line segment containing q and hence intersects Ω. However, $\lim_{n\to\infty}[x_n, \pi(y_n)] = [x, \beta] = \{b\} \subset \partial\Omega$. This is a contradiction. □

We will now apply Lemma 10.3 to prove Theorem 10.1. Suppose $\gamma \in \mathrm{Aut}(\Omega)$ is a rank-one isometry. Then $\tau_\Omega(\gamma) > 0$, which implies that γ has infinite order. By part (ii) of Proposition 6.5, γ has a unique axis ℓ_γ along which γ acts by a translation. Fix $x_0 \in \ell_\gamma$. As $\langle\gamma\rangle$ acts cocompactly on ℓ_γ, $\langle\gamma\rangle \cdot x_0$ is a quasi-isometric embedding of $\mathbb{Z}$ in Ω. Part (ii) of Lemma 10.3 implies that ℓ_γ is a $\mathcal{PS}^\Omega$-contracting set. Thus γ is a contracting element for $(\Omega, \mathcal{PS}^\Omega)$; see Definition 9.4.

11 Contracting isometries are rank one

In this section, we prove the other implication of Theorem 1.2. Fix a Hilbert geometry Ω and let $\mathcal{PS}^\Omega := \{[x, y] \mid x, y \in \Omega\}$.

Theorem 11.1 *If $\gamma \in \mathrm{Aut}(\Omega)$ is a contracting element for $(\Omega, \mathcal{PS}^\Omega)$, then γ is a rank-one isometry.*

We begin by recalling a result of Sisto which says that contracting elements are "Morse" in the following sense.

Proposition 11.2 [49, Lemma 2.8] *If $\mathcal{PS}$ is a path system on (X, d) and $\mathcal{A} \subset X$ is $\mathcal{PS}$-contracting with constant C, then there exists a constant $M = M(C)$ such that if θ is a (C, C)-quasigeodesic with endpoints in $\mathcal{A}$, then $\theta \subset \mathcal{N}_M(\mathcal{A}) := \{x \in X \mid \mathrm{d}(x, \mathcal{A}) < M\}$.*

We use this Morse property to show that a contracting element has at least one axis and none of its axes are contained in half triangles in Ω. The first step is the following lemma. Recall the notation E_γ^+, E_γ^- from Definition 4.7.

Lemma 11.3 *Suppose Ω is a Hilbert geometry and $\gamma \in \mathrm{Aut}(\Omega)$ is a contracting element for $(\Omega, \mathcal{PS}^\Omega)$. If there exist $x_0 \in \Omega$ and two unbounded sequences of positive integers $\{n_k\}_{k\in\mathbb{N}}$ and $\{m_k\}_{k\in\mathbb{N}}$ such that*

$$p := \lim_{k\to\infty} \gamma^{n_k} x_0 \quad \text{belongs to } E_\gamma^+ \qquad \text{and} \qquad q := \lim_{k\to\infty} \gamma^{-m_k} x_0 \quad \text{belongs to } E_\gamma^-,$$

then

(i) *$(p, q) \subset \Omega$, and*

(ii) *(p, q) is not contained in any half triangle in Ω.*

Proof Since γ is a contracting element, Observation 9.6 implies that $\tau_\Omega(\gamma) > 0$. Thus $p \neq q$.

(i) Suppose this is false. Then $[p, q] \subset \partial\Omega$. Choose any $r \in (p, q)$. Set $L_k := [\gamma^{-m_k} x_0, \gamma^{n_k} x_0]$. Then $L_\infty := \lim_{k\to\infty} L_k = [q, p]$. Thus we can choose $r_k \in L_k$ such that $\lim_{k\to\infty} r_k = r$.

Since γ is a contracting element, part (ii) of Observation 9.6 implies that $\mathcal{A}_{\min}(x_0) := \langle\gamma\rangle x_0$ is $\mathcal{PS}^\Omega$-contracting. Since the L_k are geodesics with endpoints in $\mathcal{A}_{\min}(x_0)$, Proposition 11.2 implies that there exists a constant M such that for all $k \geq 1$, $L_k \subset \mathcal{N}_M(\mathcal{A}_{\min}(x_0))$. Thus for every $k \geq 1$, there exists $\gamma^{t_k} x_0 \in \mathcal{A}_{\min}(x_0)$ such that

$$\mathrm{d}_\Omega(r_k, \gamma^{t_k} x_0) \leq M. \tag{12}$$

Up to passing to a subsequence, we can assume that

$$t := \lim_{k\to\infty} \gamma^{t_k} x_0$$

exists in $\overline{\Omega}$. Since r_k leaves every compact subset of Ω, $\{t_k\}$ is an unbounded sequence. Then by Proposition 4.9 part (i), $t \in (E_\gamma^+ \sqcup E_\gamma^-)$. On the other hand, by Proposition 3.12 and equation (12),

$$t \in F_\Omega(r) \subset \partial\Omega. \tag{13}$$

We now analyze the two possibilities:

Case 1 If possible, suppose $t \in E_\gamma^-$. Then consider the sequence $\{\gamma^{n_k} r\}_{k\in\mathbb{N}}$. Up to passing to a subsequence, we can assume that $r_\infty := \lim_{k\to\infty} \gamma^{n_k} r$ exists in $\partial\Omega$. Since $p \in E_\gamma^+$, $q \in E_\gamma^-$ and $r \in (p,q)$ with $n_k > 0$, Observation 4.4 part (ii) implies that

$$r_\infty = \lim_{k\to\infty} \gamma^{n_k} r \in E_\gamma^+. \tag{14}$$

To sum up, we have $r \in F_\Omega(t)$, where $t \in E_\gamma^-$ and $r_\infty = \lim_{k\to\infty} \gamma^{n_k} r$; see (13) and (14). Now we apply part (ii) of Corollary 4.13 with t, r and $\{n_k\}$ taking the role of y, z and $\{i_k\}$ respectively. Then the conclusion is that $r_\infty \in E_\gamma^-$. This contradicts equation (14).

Case 2 If possible, suppose $t \in E_\gamma^+$. We can repeat the same arguments as in Case 1 by considering the sequence $\{\gamma^{-m_k} r\}_{k\in\mathbb{N}}$, and arrive at a contradiction — we need a version of Corollary 4.13 with γ replaced by γ^{-1}; see Remark 4.14.

The contradiction to both of these possibilities finishes the proof of (i).

(ii) By part (i), $(p,q) \subset \Omega$. Suppose there exists $z \in \partial\Omega$ such that p, z, q form a half triangle in Ω. Choose any sequence of points $z_k \in [x_0, z] \cap \Omega$ such that $\lim_{k\to\infty} z_k = z$. Since γ is contracting, part (ii) of Observation 9.6 implies that $\mathcal{A}_{\min}(x_0) = \langle\gamma\rangle x_0$ is $\mathcal{PS}^\Omega$-contracting (with constant, say C). Thus there exists a projection $\pi : \Omega \to \mathcal{A}_{\min}(x_0)$ that satisfies Definition 9.2. We will analyze the sequence $\pi(z_k)$. Since $\pi(z_k) \in \mathcal{A}_{\min}(x_0)$, there exists a sequence of integers $\{i_k\}$ such that $\pi(z_k) = \gamma^{i_k} x_0$. Up to passing to a subsequence, we can assume that the following limit exists in $\overline{\Omega}$:

$$w := \lim_{k\to\infty} \pi(z_k) = \lim_{k\to\infty} \gamma^{i_k} x_0. \tag{15}$$

Claim 11.3.1 *It holds that $w \in \partial\Omega$ and $w \in (E_\gamma^+ \sqcup E_\gamma^-)$.*

Proof Recall that Γ acts properly discontinuously on Ω. Moreover, $\omega(\gamma, \Omega) \cup \omega(\gamma^{-1}, \Omega) \subset E_\gamma^+ \cup E_\gamma^-$; see Proposition 4.9. Thus it suffices to show that $\{i_k\}$ is an unbounded sequence. Suppose, on the contrary, that $\{i_k\}$ is a bounded sequence. Then $w \in \Omega$ and $\lim_{k\to\infty} \mathrm{d}_\Omega(w, \pi(z_k)) = 0$; see (15).

Recall that $\{n_k\}$ is the sequence such that $\gamma^{n_k} x_0 \to p \in \partial\Omega$. We will prove this claim by comparing $\gamma^{i_k} x_0 (= \pi(z_k))$ with $\gamma^{n_k} x_0$. We claim that $\mathrm{d}_\Omega(\pi(z_k), \pi(\gamma^{n_k} x_0)) \to \infty$ as $k \to \infty$. To prove this subclaim, first note that (i) of Definition 9.2 implies that

$$\mathrm{d}_\Omega(\gamma^{n_k} x_0, \pi(\gamma^{n_k} x_0)) \le C$$

because $\gamma^{n_k} x_0 \in \mathscr{A}_{\min}(x_0)$. The subclaim then follows from the equation

$$\begin{aligned}\lim_{k\to\infty} \mathrm{d}_\Omega(\pi(z_k), \pi(\gamma^{n_k} x_0)) &\ge \lim_{k\to\infty} \bigl(\mathrm{d}_\Omega(w, \gamma^{n_k} x_0) - \mathrm{d}_\Omega(w, \pi(z_k)) - \mathrm{d}_\Omega(\gamma^{n_k} x_0, \pi(\gamma^{n_k} x_0))\bigr) \\ &\ge \liminf_{k\to\infty} \mathrm{d}_\Omega(w, \gamma^{n_k} x_0) - C = \infty.\end{aligned}$$

The above equation then implies that for k large enough, $\mathrm{d}_\Omega(\pi(z_k), \pi(\gamma^{n_k} x_0)) \ge C$. Since π is a projection into a $\mathcal{PS}^\Omega$-contracting set, condition (ii) of Definition 9.2 implies that

$$\mathrm{d}_\Omega(\pi(z_k), [z_k, \gamma^{n_k} x_0]) \le C.$$

Thus

$$\mathrm{d}_\Omega(w, [z, p]) \le \lim_{k\to\infty} \mathrm{d}_\Omega(\pi(z_k), [z_k, \gamma^{n_k} x_0]) \le C.$$

Then $[z, p] \cap \Omega \ne \varnothing$. But since p, z, q form a half triangle, $[z, p] \subset \partial\Omega$. This is a contradiction and it concludes the proof of this claim. □

Claim 11.3.2 $\quad w \in F_\Omega(z).$

Proof First observe that for k large enough,

$$\mathrm{d}_\Omega(\pi(z_k), \pi(x_0)) \ge C.$$

Indeed, this follows because $\pi(x_0) \in \Omega$ while $w = \lim_{k\to\infty} \pi(z_k) \in \partial\Omega$. Again, as π is a projection into a $\mathcal{PS}^\Omega$-contracting set, we have

$$\mathrm{d}_\Omega(\pi(z_k), [x_0, z_k]) \le C.$$

Choose $\eta_k \in [x_0, z_k]$ such that $\mathrm{d}_\Omega(\pi(z_k), \eta_k) \le C$. Up to passing to a subsequence, we can assume that $\eta := \lim_{k\to\infty} \eta_k$ exists. By Proposition 3.12, $\eta \in F_\Omega(w)$. Since $w \in \partial\Omega$, $\eta \in \partial\Omega$ (Proposition 3.11(i)). But $\eta \in [x_0, z]$, which intersects $\partial\Omega$ at exactly one point, namely z. Thus, $\eta = z$ implying $z \in F_\Omega(w)$, or equivalently, $w \in F_\Omega(z)$. This concludes the proof of Claim 11.3.2. □

Since p, z, q form a half triangle, $[p, z] \cup [z, q] \subset \partial\Omega$. By Claim 11.3.2, $w \in F_\Omega(z)$. Then part (iv) of Proposition 3.11 implies that

$$[p, w] \cup [q, w] \subset \partial\Omega. \tag{16}$$

Recall from Claim 11.3.1 that $\{i_k\}$ is an unbounded sequence and that $w = \lim_{k\to\infty} \gamma^{i_k} x_0$ lies in $E_\gamma^+ \sqcup E_\gamma^-$. We will now show that (16) contradicts this.

Suppose, up to passing to a subsequence, that $\{i_k\}$ is a sequence of positive integers. Then $w \in E_\gamma^+$. Since $\lim_{k\to\infty} \gamma^{i_k} x_0 = w \in E_\gamma^+$ and $\lim_{k\to\infty} \gamma^{-m_k} x_0 = q \in E_\gamma^-$, then part (i) of Lemma 11.3 implies that $(w, q) \subset \Omega$. This contradicts (16). On the other hand, if we suppose that $\{i_k\}$ is a sequence of negative integers, then $w \in E_\gamma^-$. Then, by a similar reasoning, $(p, w) \subset \Omega$ which again contradicts (16). These contradictions show that p, z, q cannot form a half triangle. □

We now prove Theorem 11.1 using the above lemma. Let $\gamma \in \mathrm{Aut}(\Omega)$ be a contracting element for $(\Omega, \mathcal{PS}^\Omega)$. By part (i) of Observation 9.6, $\tau_\Omega(\gamma) > 0$. The following will imply that γ is a rank-one isometry.

- **γ has an axis** By Proposition 5.8, there exists $(a, b) \subset \overline{\Omega}$ with a, b fixed points of γ such that $a \in E_\gamma^+$ and $b \in E_\gamma^-$. We will show that $(a, b) \subset \Omega$; hence it is an axis of γ.

Fix $x_0 \in \Omega$. Proposition 4.9 part (i) implies $\{\gamma^n x_0 \mid n \in \mathbb{N}\}$ has an accumulation point p in E_γ^+ and $\{\gamma^{-n} x_0 \mid n \in \mathbb{N}\}$ has accumulation point q in E_γ^-. By part (i) of Lemma 11.3, $(p, q) \subset \Omega$.

Note that $E_\gamma^+ \cap \overline{\Omega} \subset \partial\Omega$; see Claim 4.6.1. Thus $[a, p] \subset \partial\Omega$ as $a, p \in E_\gamma^+ \cap \partial\Omega$. Similarly, $[b, q] \subset \partial\Omega$. By part (ii) of Lemma 11.3, $(p, q) \subset \Omega$ is not contained in any half triangle in Ω. Since $[b, q] \subset \partial\Omega$, this implies that $(p, b) \subset \Omega$.

We will use $(p, b) \subset \Omega$ to derive that $(a, b) \subset \Omega$. First note that since $b \in E_\gamma^-$ is the endpoint of a pseudoaxis, b is a fixed point of γ. Thus $\lim_{k\to\infty} \gamma^{-k} y' = b \in E_\gamma^-$ for any $y' \in (p, b)$. We then note that p is an "almost-fixed" point of γ, ie there exists $\{n_k\}$ with $n_k \to \infty$ such that $\lim_{k\to\infty} \gamma^{n_k} p = p \in E_\gamma^+$. Indeed, Proposition 4.9 part (iii) implies that there exists a sequence of positive integers $\{n_k\}$ with $n_k \to \infty$ such that $\lim_{k\to\infty} \gamma|_{E_\gamma^+}^{n_k} = \mathrm{Id}_{E_\gamma^+}$, ie $\lim_{k\to\infty} \gamma^{n_k} p = p$. Now pick $y_0 \in (p, b) \in \Omega$. The above discussion implies that $\lim_{k\to\infty} \gamma^{n_k} y_0 = p \in E_\gamma^+$ while $\lim_{k\to\infty} \gamma^{-k} y_0 = b \in E_\gamma^-$. Then, by part (ii) of Lemma 11.3, $(p, b) \subset \Omega$ cannot be contained in a half triangle in Ω. But we know that $[a, p] \subset \partial\Omega$. Thus, $(a, b) \subset \Omega$.

- **None of the axes of γ are contained in a half triangle in Ω** Let $(a', b') \subset \Omega$ be any axis of γ with $a' \in E_\gamma^+$ and $b' \in E_\gamma^-$. If $z_0 \in (a', b')$, then $\lim_{k\to\infty} \gamma^k z_0 = a'$ and $\lim_{k\to\infty} \gamma^{-k} z_0 = b'$. Then, by part (ii) of Lemma 11.3, (a', b') cannot be contained in a half triangle in Ω.

12 Acylindrical hyperbolicity: proof of Theorem 1.4

Acylindrically hyperbolic groups are a generalization of nonelementary Gromov hyperbolic groups with many interesting examples, like mapping class groups of most finite-type surfaces, rank-one CAT(0) groups that are not virtually cyclic, outer automorphisms of free groups on at least two generators and relatively hyperbolic groups with proper peripheral subgroups that are not virtually cyclic [45, Appendix]. In this section, we will add a new class of examples by showing that discrete groups acting on Hilbert geometries with at least one rank-one isometry are either virtually cyclic or acylindrically hyperbolic.

12.1 Acylindrically hyperbolic groups

We first recall some basic definitions about Gromov hyperbolic metric spaces (not necessarily proper) and we refer to [33] for details. A geodesic metric space (Y, d_Y) is called *Gromov hyperbolic* if there exists $\delta \geq 0$ such that every geodesic triangle in Y is δ-thin (recall Definition 7.3). If (Y, d_Y) is Gromov hyperbolic, let ∂Y denote the *boundary of* Y defined via equivalence classes of sequences in Y "convergent at infinity"; see [33, Section 1.8]. We remark that this definition of ∂Y does not require that Y is a proper metric space.

If G acts isometrically on a Gromov hyperbolic space (Y, d_Y), let $\Lambda_G(Y) \subset \partial Y$ denote the *limit set* of the G-action (ie $\Lambda_G(Y)$ is the set of accumulation points in ∂Y of any G orbit in Y). The action is called *nonelementary* if $\#(\Lambda_G(Y)) = \infty$; see [45] for details.

Finally we define the notion of acylindrical actions on a metric space (not necessarily Gromov hyperbolic). An isometric action of a group G on a metric space (Y, d_Y) is called *acylindrical* if, for every $\varepsilon > 0$, there exists $R_\varepsilon, N_\varepsilon > 0$ such that if $x, y \in Y$ with $\mathrm{d}_Y(x, y) \geq R_\varepsilon$, then

$$\#\{g \in G \mid \mathrm{d}_Y(x, gx) \leq \varepsilon \text{ and } \mathrm{d}_Y(y, gy) \leq \varepsilon\} \leq N_\varepsilon.$$

Definition 12.1 A group G is called *acylindrically hyperbolic* if it admits an isometric nonelementary acylindrical action on a (possibly nonproper) Gromov hyperbolic metric space (Y, d_Y).

A motivating example of acylindrically hyperbolic groups is a nonelementary Gromov hyperbolic group. Indeed, if H is a finitely generated nonelementary Gromov hyperbolic group, then it has a nonelementary acylindrical action on its Cayley graph which is a Gromov hyperbolic metric space. More generally if H is a finitely generated nonelementary relatively hyperbolic group with proper peripheral subgroups, then H has an acylindrical action on its coned-off Cayley graph. Another interesting example is the mapping class group of a closed hyperbolic surface. It acts acylindrically and nonelementarily on the curve graph of the surface, which is a (nonproper) Gromov hyperbolic space.

Although Definition 12.1 of acylindrically hyperbolic groups is perhaps the cleanest to state, a characterization of acylindrically hyperbolic groups using contracting elements will be particularly well-suited for our purpose. We state such a characterization now, which follows directly from work of Osin and Sisto; a proof is included because we could not find a result stated in this form.

Theorem 12.2 [45; 49] *Suppose G has a proper isometric action on a geodesic metric space (X, d), and suppose that $(X, \mathcal{PS})$ is a path system for G and $g \in G$ is a contracting element for $(X, \mathcal{PS})$. Then either G is virtually cyclic, or G is acylindrically hyperbolic.*

Sketch of proof In [45], Osin introduces several characterizations of acylindrically hyperbolic groups that are equivalent to Definition 12.1. The one that we will use (Proposition 12.3) requires the notion of *hyperbolically embedded subgroups*. Results due to Osin and Sisto (see Propositions 12.3 and 12.4) will allow us to use this notion without defining it precisely. See [45, Definition 2.8] or [49, Definition 4.6].

Proposition 12.3 (Osin [45, Theorem 1.2 and Definition 1.3] and Remark 12.5) *A group G is acylindrically hyperbolic if G contains a proper infinite hyperbolically embedded subgroup.*

So in order to prove that a group is acylindrically hyperbolic, it suffices to produce a proper infinite hyperbolically embedded subgroup. For this, we rely on a result of Sisto.

Proposition 12.4 [49, Theorem 4.7] *Suppose $g \in G$ is a contracting element for $(X, \mathcal{PS})$ and $\mathcal{A} \subset X$ is $\langle g\rangle$-invariant, $\mathcal{PS}$-contracting and has cobounded $\langle g\rangle$-action. Then*

$$E(g) := \{h \in G \mid d^{\mathrm{Haus}}(\pi_{\mathcal{A}}(h\mathcal{A}), \mathcal{A}) < \infty\}$$

is a hyperbolically embedded subgroup of G which is infinite and contains $\langle g\rangle$ as a finite-index subgroup, ie $E(g)$ is virtually cyclic.

Now let us summarize how these results give us our desired conclusion. Suppose $g \in G$ is a contracting element. By Proposition 12.4, $E(g)$ is an infinite hyperbolically embedded subgroup of G which is virtually cyclic. Now note that if G is virtually cyclic, there is nothing to prove. So suppose that G is not virtually cyclic. Then $E(g) \subsetneq G$ as $E(g)$ is virtually cyclic. Thus $E(g)$ is a proper infinite hyperbolically embedded subgroup and Proposition 12.3 implies that G is an acylindrically hyperbolic group. See the following remark for further comments on the proof. □

Remark 12.5 Recall the alternate definition of an acylindrically hyperbolic group from Proposition 12.3. A subgroup $H \le G$ is *proper infinite* if $H \subsetneq G$ and H is infinite. Such proper infinite hyperbolically embedded subgroups are sometimes called *nondegenerate* hyperbolically embedded subgroups in the terminology of [45; 29]. Notably, the existence of one such nondegenerate hyperbolically embedded subgroup $H \le G$ implies the existence of nonabelian free subgroups in G; see [45, Lemma 5.12] or [29, Theorem 6.14]. Thus G is not virtually cyclic and contains infinitely many "independent loxodromic" elements [45; 29]. Roughly speaking, this is akin to producing nonabelian free subgroups in any nonelementary Gromov hyperbolic group.

12.2 Proof of Theorem 1.4

We first recall the theorem.

Theorem 1.4 *If (Ω, Γ) is a rank-one Hilbert geometry, then either Γ is virtually cyclic or Γ is an acylindrically hyperbolic group.*

The proof of Theorem 1.4 will be immediate from Theorem 12.2, thanks to the well-developed machinery of acylindrically hyperbolic groups due to the work of many authors; see for instance [45; 29; 49; 16]. In case the proof seems a bit opaque to a reader, we will first give an informal sketch of the underlying idea before providing a formal proof.

Our result Theorem 1.2 implies that rank-one isometries in $\mathrm{Aut}(\Omega)$ are contracting elements for $(\Omega, \mathcal{PS}^{\Omega})$. Thus a rank-one Hilbert geometry (Ω, Γ) contains contracting elements by definition. Now it is possible that Γ is virtually cyclic in which case Γ, up to passing to a finite-index subgroup, is generated by a single rank-one isometry. But if Γ is not virtually cyclic, then there will be infinitely many rank-one isometries $\gamma_1, \gamma_2, \ldots$ which are "independent loxodromics", ie there exists an abstract Gromov hyperbolic space X on which each γ_i acts "loxodromically" with exactly two distinct fixed points $\gamma_i^{\pm}$ and the sets $\{\gamma_i^{\pm}\}$ and $\{\gamma_j^{\pm}\}$ are pairwise disjoint whenever $i \neq j$. This last conclusion follows from results in [29] and [49] that we referred to in Remark 12.5. These infinitely many independent rank-one isometries γ_i generate nonabelian free subgroups of Γ, and the γ_i lie in distinct hyperbolically embedded subgroups $E(\gamma_i)$; see Proposition 12.4.

Now let us give the formal proof.

Proof of Theorem 1.4 Since (Ω, Γ) is a rank-one Hilbert geometry, Γ contains a rank-one isometry. Then Theorem 1.2 implies that Γ contains a contracting element for $(\Omega, \mathcal{PS}^{\Omega})$. The result follows from Theorem 12.2. $\square$

Remark 12.6 By Theorem 1.4, a rank-one Hilbert geometry (Ω, Γ) where Γ is not virtually cyclic gives an example of an acylindrically hyperbolic group Γ. A natural question is: what is an example of a Gromov hyperbolic metric space X on which Γ acts acylindrically and nonelementarily? Is there a way to understand this space X in terms of the Hilbert geometry Ω?

It seems that one might be able to apply the projection complex construction in [16] (see also [14; 49]) to construct such a space X from the Hilbert geometry Ω. Roughly speaking, this will be a metric space obtained by collecting all rank-one axes in Ω and adding edges between them depending on diameters of images of some projection maps. We do not pursue this direction in this paper and this remark is mostly speculative in nature.

Part IV Applications

13 Second bounded cohomology and quasimorphisms

13.1 Definitions

We first introduce some definitions following [15, Section 1]. Suppose G is a group, $(E, \|\cdot\|)$ is a complete normed $\mathbb{R}$-vector space and $\rho\colon G \to \mathcal{U}(E)$ is a unitary representation. Let $C(G, E)$ be the space of all functions from G to E.

A function $F \in C(G, E)$ is called a quasicocycle if

$$\Delta(F) := \sup_{g,g' \in G} \|F(gg') - F(g) - \rho(g)F(g')\| < \infty.$$

Let V be the vector subspace of $C(G,E)$ that consists of all quasicocyles. Let V_0 be the subspace of V generated by bounded functions and the set

$$\{F\colon G \to E \mid F(gg') = F(g) + \rho(g)F(g') \text{ for all } g, g' \in G\}.$$

Define

$$\widetilde{\mathrm{QC}}(G;\rho) := V/V_0.$$

If ρ is the trivial representation $\rho_{\mathrm{triv}}\colon G \to \mathbb{R}$, then V is the space of quasimorphisms of G while V_0 is the space generated by bounded functions and group homomorphisms from G to $\mathbb{R}$. In this case, $\widetilde{\mathrm{QC}}(G;\rho_{\mathrm{triv}})$ recovers a classical object called the space of "nontrivial" quasimorphisms of G, usually denoted by $\widetilde{\mathrm{QH}}(G)$; see the definitions preceding Theorem 1.6.

Group cohomology of G (twisted by the representation ρ) affords an interesting interpretation of $\widetilde{\mathrm{QC}}(G;\rho)$. If F is a quasicocycle, then $dF(g,g') := F(gg') - F(g) - \rho(g)F(g')$ defines a class in the second bounded cohomology group $H^2_b(G;\rho)$. This class dF is trivial in the ordinary cohomology group $H^2(G;\rho)$. On the other hand, the class dF is nontrivial in $H^2_b(G;\rho)$ whenever F is nontrivial in V/V_0. Thus $\widetilde{\mathrm{QC}}(G;\rho)$ is the kernel of the comparison map $H^2_b(G,\rho) \to H^2(G;\rho)$. For a more detailed discussion, we refer the reader to [15, Section 1] or [31].

13.2 Results

Infinite dimensionality of $\widetilde{\mathrm{QH}}(G)$ and $\widetilde{\mathrm{QC}}(G;\rho)$ is often related to geometric phenomena. For example, [17] shows that a compact irreducible nonpositively curved Riemannian manifold M is (Riemannian) rank one if and only if $\dim(\widetilde{\mathrm{QH}}(\pi_1(M))) = \infty$. Now, in the same spirit as in Riemannian nonpositive curvature, we prove a cohomological characterization of rank-one Hilbert geometries. We will only consider unitary representations on uniformly convex Banach spaces,[3] eg $\mathbb{R}$ or $\ell^p(G)$, where G is a discrete group and $1 < p < \infty$.

Theorem 13.1 *Suppose that (Ω,Γ) is a rank-one Hilbert geometry, Γ is torsion-free and ρ is any unitary representation of Γ on a uniformly convex Banach space $E \neq 0$. Then either Γ is virtually cyclic or* $\dim(\widetilde{\mathrm{QC}}(\Gamma;\rho)) = \infty$.

The proof follows directly from the following general result about acylindrically hyperbolic groups.

Theorem 13.2 [15, Corollary 1.2] *If G is an acylindrically hyperbolic group, $E \neq 0$ is a uniformly convex Banach space, $\rho\colon G \to \mathfrak{U}(E)$ is a unitary representation and any maximal finite normal subgroup of G has a nonzero fixed vector, then* $\dim(\widetilde{\mathrm{QC}}(G;\rho)) = \infty$. □

Proof of Theorem 13.1 If Γ is not virtually cyclic, then Theorem 1.4 implies that Γ is an acylindrically hyperbolic group. Since Γ is torsion-free, there are no finite normal subgroups. The claim then follows from Theorem 13.2. □

[3] A Banach space E is uniformly convex if for any $\varepsilon' > 0$, there exists $\delta' > 0$ such that if $u, v \in E$, $\|u\| \le 1$, $\|v\| \le 1$, $\|u - v\| \ge \varepsilon'$, then $\|(u+v)/2\| \le 1 - \delta'$.

We will now apply Theorem 13.1 to two specific choices of ρ and E to get Theorem 1.6. For the first, $\rho = \rho_{\mathrm{triv}}$ and $E = \mathbb{R}$, in which case $\widetilde{\mathrm{QC}}(\Gamma; \rho) = \widetilde{\mathrm{QH}}(\Gamma)$, the space of nontrivial quasimorphisms. For the second, $E = \ell^p(\Gamma)$ with $1 < p < \infty$ and $\rho = \rho^p_{\mathrm{reg}}$ is the regular representation, ie $\rho^p_{\mathrm{reg}}(\gamma) f(x) = f(\gamma^{-1} x)$ for any $f \in \ell^p(\Gamma)$ and $x \in \Gamma$.

Theorem 1.6 *If (Ω, Γ) is a rank-one Hilbert geometry, Γ is torsion-free and Γ is not virtually cyclic, then $\dim(\widetilde{\mathrm{QH}}(\Gamma)) = \infty$ and $\dim(\widetilde{\mathrm{QC}}(\Gamma; \rho^p_{\mathrm{reg}})) = \infty$ if $1 < p < \infty$.*

Proof Immediate from Theorem 13.1 and the fact that $\mathbb{R}$ and $\ell^p(\Gamma)$ with $1 < p < \infty$ are uniformly convex Banach spaces; see [15, Section 3]. □

Corollary 1.7 *If (Ω, Γ) is a divisible Hilbert geometry and Ω is irreducible, then $\dim(\widetilde{\mathrm{QH}}(\Gamma)) = \infty$ if and only if (Ω, Γ) is a rank-one Hilbert geometry. Otherwise, $\dim(\widetilde{\mathrm{QH}}(\Gamma)) = 0$.*

Proof If (Ω, Γ) is a rank-one Hilbert geometry, then Theorem 1.6 implies that $\dim(\widetilde{\mathrm{QH}}(\Gamma)) = \infty$. If (Ω, Γ) is not rank one, then Theorem 1.5 implies that $\mathrm{Aut}(\Omega)$ is locally isomorphic to a simple Lie group of real rank at least two, ie Ω is an irreducible symmetric domain of rank at least two. Thus Γ is isomorphic to a uniform lattice in a higher-rank simple Lie group, which implies that $\dim(\widetilde{\mathrm{QH}}(\Gamma)) = 0$ [21, Theorem 21]. □

14 Counting of conjugacy classes

Suppose (Ω, Γ) is a rank-one Hilbert geometry. Recall the notions of translation length and stable translation length of a conjugacy class in Γ; see Section 1.2. We now introduce the notion of pointed length for a conjugacy class $[c_g]$ of $g \in \Gamma$; see [32]. Fix a basepoint $p \in \Omega$. The pointed length of $[c_g]$ is

$$\mathscr{L}_p([c_g]) := \inf_{g' \in [c_g]} \mathrm{d}_\Omega(p, g' p).$$

We first show that

$$\tau_\Omega([c_g]) = \tau_\Omega^{\mathrm{stable}}([c_g]).$$

Indeed, triangle inequality implies $\tau_\Omega^{\mathrm{stable}}(g) \leq \tau_\Omega(g)$. On the other hand, by Proposition 3.15,

$$\tau_\Omega^{\mathrm{stable}}(g) \geq \lim_{n \to \infty} \frac{\tau_\Omega(g^n)}{n} = \frac{1}{n} \log \frac{\lambda_{\max}(\tilde{g}^n)}{\lambda_{\min}(\tilde{g}^n)} = \log \frac{\lambda_{\max}(\tilde{g})}{\lambda_{\min}(\tilde{g})} = \tau_\Omega(g).$$

Next, we show that if Ω / Γ is compact and $R := \mathrm{diam}(\Omega / \Gamma)$, then

$$\tau_\Omega([c_g]) \leq \mathscr{L}_p([c_g]) \leq \tau_\Omega([c_g]) + 2R.$$

Clearly $\tau_\Omega([c_g]) \leq \mathscr{L}_p([c_g])$. On the other hand, if $x \in \Omega$ then there exists $h_x \in \Gamma$ such that $\mathrm{d}_\Omega(x, h_x p) \leq R$. Then

$$\mathscr{L}_p([c_g]) \leq \mathrm{d}_\Omega(p, h_x^{-1} g h_x p) \leq 2\, \mathrm{d}_\Omega(h_x p, x) + \mathrm{d}_\Omega(x, gx) \leq 2R + \mathrm{d}_\Omega(x, gx).$$

Thus, $\mathscr{L}_p([c_g]) \leq \tau_\Omega([c_g]) + 2R$.

Now let us consider the following counting functions for conjugacy classes in Γ:

$$\mathscr{C}(t) := \#\{[c_g] \mid g \in \Gamma, \tau_\Omega([c_g]) \le t\},$$
$$\mathscr{C}^{\mathrm{stable}}(t) := \#\{[c_g] \mid g \in \Gamma, \tau_\Omega^{\mathrm{stable}}([c_g]) \le t\},$$
$$\mathscr{C}^{\mathscr{L}_p}(t) := \#\{[c_g] \mid g \in \Gamma, \mathscr{L}_p([c_g]) \le t\}.$$

Based on the above discussion,

$$\mathscr{C}(t) = \mathscr{C}^{\mathrm{stable}}(t). \tag{17}$$

If Ω/Γ is compact and $R = \mathrm{diam}(\Omega/\Gamma)$, then

$$\mathscr{C}^{\mathscr{L}_p}(t) \le \mathscr{C}(t) \le \mathscr{C}^{\mathscr{L}_p}(t + 2R). \tag{18}$$

We now prove asymptotic growth formula for these functions. It is a direct consequence of the Main Theorem in [32]. Recall that the critical exponent of Γ (see Section 1.2) is defined by

$$\omega_\Gamma := \limsup_{n\to\infty} \frac{\log \#\{g \in \Gamma \mid \mathrm{d}_\Omega(x, gx) \le n\}}{n}.$$

Theorem 1.8 *Suppose (Ω, Γ) is a divisible rank-one Hilbert geometry and Γ is not virtually cyclic. Then there exists a constant D' such that for all $t \ge 1$,*

$$\frac{1}{D'}\frac{\exp(t\omega_\Gamma)}{t} \le \mathscr{C}(t) \le D'\frac{\exp(t\omega_\Gamma)}{t}. \tag{19}$$

The functions $\mathscr{C}^{\mathrm{stable}}(t)$, $\mathscr{C}^{\mathscr{L}_p}(t)$ and $\mathscr{C}_{\mathrm{Prim}}(t)$ (see Remark 1.9) also satisfy similar growth formulas.

Proof Part (1) of the Main Theorem in [32] implies that if Γ is a nonelementary group with a cocompact action (more generally, statistically convex cocompact action) on a geodesic metric space and Γ contains a contracting element (in the sense of BF, see Appendix B), then $\mathscr{C}^{\mathscr{L}_p}(t)$ satisfies the growth formula in (19). If (Ω, Γ) is as above, then it satisfies all of these conditions; see Theorem 1.2 and Remark 9.7. Then $\mathscr{C}^{\mathscr{L}_p}(t)$ satisfies equation (19). By equations (17) and (18), $\mathscr{C}(t)$ and $\mathscr{C}^{\mathrm{stable}}(t)$ also satisfy equation (19).

For proving Remark 1.9 part (ii), set

$$\mathscr{C}^{\mathscr{L}_p}_{\mathrm{Prim}}(t) := \#\{[c_g] \mid g \in \Gamma \text{ is primitive}, \mathscr{L}_p([c_g]) \le t\}.$$

Part (1) of the Main Theorem in [32] implies that the $\mathscr{C}^{\mathscr{L}_p}_{\mathrm{Prim}}(t)$ satisfies a similar growth formula as (19). Since $\tau_\Omega([c_g]) \le \mathscr{L}_p([c_g]) \le \tau_\Omega([c_g]) + 2R$, this implies the result for $\mathscr{C}_{\mathrm{Prim}}(t)$. □

15 Proofs of Propositions 1.10, 1.11 and 1.12

For the proofs in this section, recall the following implication of Theorem 1.4: if (Ω, Γ) is a rank-one Hilbert geometry and Γ is not virtually cyclic, then Γ is an acylindrically hyperbolic group.

Proposition 1.10 *If* (Ω, Γ) *is a rank-one Hilbert geometry*, Γ *is not virtually cyclic and* Γ *is finitely generated*, *then the rank-one isometries in* Γ *are exponentially generic*: *if* $(X_n)_{n\in\mathbb{N}}$ *is a simple random walk on* Γ, *then there exists a constant* $C \geq 1$ *such that for all* $n \geq 1$,

$$\mathbb{P}[X_n \text{ is not a rank-one isometry}] \leq C e^{-n/C}.$$

Proof Under the hypotheses, Γ is an acylindrically hyperbolic group. The result then follows from [49, Theorem 1.6]. □

Proposition 1.11 *If* (Ω, Γ) *is a rank-one Hilbert geometry and* Γ *is not virtually cyclic*, *then*:

(i) Γ *is SQ-universal*, *ie every countable group embeds in a quotient of* Γ.

(ii) *If* Γ *is the Baumslag-Solitar group* $\mathrm{BS}(m,n)$, *then* $m = n = 0$ *and* Γ *is the free group on two generators.*

Proof Under the hypotheses, Γ is an acylindrically hyperbolic group. Then SQ-universality follows from [45, Theorem 8.1]. The second part follows from [45, Example 7.4], where Osin proves that $\mathrm{BS}(m,n)$ is acylindrically hyperbolic if and only if $m = n = 0$. But $\mathrm{BS}(0,0) = F_2$. □

Proposition 1.12 *If* Ω *is a Hilbert geometry and* $\gamma \in \mathrm{Aut}(\Omega)$ *is a rank-one isometry*, *then the axis* ℓ_γ *of* γ *is* $\mathcal{K}$*-Morse for some Morse gauge* $\mathcal{K}\colon [1,\infty) \times [0,\infty) \to [0,\infty)$, *ie if* α *is a* (λ, ε)*-quasigeodesic with endpoints on* ℓ_γ, *then* $\alpha \subset \mathcal{N}_{\mathcal{K}(\lambda,\varepsilon)}(\ell_\gamma)$.

Proof Since γ is a rank-one isometry, Theorem 1.2 implies that γ is a contracting element for $(\Omega, \mathcal{PS}^{\Omega})$. Then the axis of ℓ_γ of γ is $\mathcal{PS}^{\Omega}$-contracting. Thus [49, Lemma 2.8] (Proposition 11.2 in this paper) implies that ℓ_γ is a Morse geodesic. □

Appendix A Rank-one Hilbert geometries: generalization, examples and nonexamples

This section is devoted to the discussion of examples and nonexamples of rank-one Hilbert geometries (see Definition 1.3) and generalizing the notion of rank one to convex cocompact actions.

A.1 Strictly convex examples

If Ω is a strictly convex Hilbert geometry, then $\partial\Omega$ does not contain any line segments. Thus, if $g \in \mathrm{Aut}(\Omega)$ with $\tau_\Omega(g) > 0$, then g is a rank-one isometry (as it has an axis and there are no half triangles in Ω). Then, all the strictly convex divisible examples in Section 3.4 are rank one.

A.2 Non-strictly-convex examples

Suppose (Ω, Γ) is a divisible Hilbert geometry where Γ is infinite and not virtually abelian. Assume that Γ is a relatively hyperbolic group with respect to a finite collection of free abelian subgroups of rank at least two. Then we claim that (Ω,Γ) *is a divisible rank-one Hilbert geometry*. The proof of this follows from Remark A.3(B) and Proposition A.4; see below. This claim implies that the divisible nonstrictly convex examples discussed in Section 3.4, that are neither simplices nor symmetric domains of rank at least two, are all examples of rank-one Hilbert geometries.

A.3 Nonexamples

The d-simplices T_d for $d \geq 2$ are clearly nonexamples of rank-one Hilbert geometries. If Ω is an irreducible symmetric domain of rank at least two and $\Gamma \leq \mathrm{Aut}(\Omega)$ acts cocompactly on Ω, then (Ω,Γ) *cannot be a rank-one Hilbert geometry*; see Theorem 1.5.

A.4 Generalization of rank one to convex cocompact actions

The notion of convex cocompact actions on Hilbert geometries [30] generalizes divisible Hilbert geometries. Suppose Ω is a Hilbert geometry and $\Gamma \leq \mathrm{Aut}(\Omega)$ is a discrete subgroup. The *full orbital limit set* is defined as $\mathcal{L}^{\mathrm{orb}}_\Omega(\Gamma) := \bigcup_{x\in\Omega}(\overline{\Gamma\cdot x}\cap\partial\Omega)$ and let $\mathscr{C}^c_\Omega(\Gamma) := \mathrm{ConvHull}_\Omega(\mathcal{L}^{\mathrm{orb}}_\Omega(\Gamma))$.

Definition A.1 An infinite discrete group $\Gamma \leq \mathrm{Aut}(\Omega)$ is *convex cocompact* if $\mathscr{C}^c_\Omega(\Gamma) \neq \varnothing$ and $\mathscr{C}^c_\Omega(\Gamma)/\Gamma$ is compact.

The *ideal boundary* of $\mathscr{C}^c_\Omega(\Gamma)$ is given by $\partial_{\mathrm{i}}\mathscr{C}^c_\Omega(\Gamma) := \partial\Omega \cap \overline{\mathscr{C}^c_\Omega(\Gamma)}$. For convex cocompact groups, $\partial_{\mathrm{i}}\mathscr{C}^c_\Omega(\Gamma)$ is the only part of $\partial\Omega$ "visible" to the group acting on Ω. Thus it is natural to modify the notion of rank-one isometries by considering half triangles in $\mathscr{C}^c_\Omega(\Gamma)$ instead of Ω. We say that the projective geodesic $(a,b) \subset \mathscr{C}^c_\Omega(\Gamma)$ is *not contained in any half triangle in* $\mathscr{C}^c_\Omega(\Gamma)$ if either $(a,z) \subset \mathscr{C}^c_\Omega(\Gamma)$ or $(z,b) \subset \mathscr{C}^c_\Omega(\Gamma)$ for any $z \in \partial_{\mathrm{i}}\mathscr{C}^c_\Omega(\Gamma)$.

Definition A.2 Suppose $\Omega \subset \mathbb{P}(\mathbb{R}^{d+1})$ is a Hilbert geometry and $\Gamma \leq \mathrm{Aut}(\Omega)$ is a convex cocompact group.

(i) An element $\gamma \in \Gamma$ is a *convex cocompact rank-one isometry* if
   (a) $\log|(\lambda_1/\lambda_{d+1})(\gamma)| > 0$ and γ has an axis (see Definition 5.1),
   (b) none of the axes ℓ_γ of γ are contained in a half triangle in $\mathscr{C}^c_\Omega(\Gamma)$.

(ii) We say that Γ is a *rank-one convex cocompact group* if Γ contains a convex cocompact rank-one isometry.

Remark A.3 (A) The notion of a *convex cocompact rank-one isometry* differs from the notion of a *rank-one isometry* (see Definition 6.3) only in condition (i)(b): for convex cocompact actions, we consider half triangles in $\mathscr{C}^c_\Omega(\Gamma)$ instead of Ω.

(B) If Γ acts cocompactly on Ω, then (Ω,Γ) is a *divisible rank-one Hilbert geometry* if and only if Γ is a *rank-one convex cocompact group*. This is because divisibility implies $\mathscr{C}^c_\Omega(\Gamma) = \Omega$.

If $\Gamma \leq \mathrm{Aut}(\Omega)$ is a convex cocompact group and $\gamma \in \Gamma$ is a convex cocompact rank-one isometry, then the analogues of Propositions 6.5, 6.7 and 6.8 hold. But now we need to replace Ω with $\mathcal{C}_\Omega^c(\Gamma)$ and $\partial\Omega$ with $\partial_{\mathrm{i}}\mathcal{C}_\Omega^c(\Gamma)$. In particular, we have that $\gamma \in \Gamma$ is a convex cocompact rank-one isometry if and only if γ is biproximal and has an axis.

We sketch the proof ideas of these analogues; see [37] for details. Observe that if $|(\lambda_1/\lambda_{d+1})(\gamma)| > 0$, then $E_\gamma^\pm \cap \partial\Omega = E_\gamma^\pm \cap \partial_{\mathrm{i}}\mathcal{C}_\Omega^c(\Gamma)$. Recall that if $x \in \partial_{\mathrm{i}}\mathcal{C}_\Omega^c(\Gamma)$, then

$$F_{\mathcal{C}_\Omega^c(\Gamma)}(x) := \{x\} \cup \{y \in \overline{\mathcal{C}_\Omega^c(\Gamma)} \mid \text{an open projective line segment in } \overline{\mathcal{C}_\Omega^c(\Gamma)} \text{ contains } x \text{ and } y\}.$$

Convex cocompact groups have a special property: if $x \in \partial_{\mathrm{i}}\mathcal{C}_\Omega^c(\Gamma)$, then $F_{\mathcal{C}_\Omega^c(\Gamma)}(x) = F_\Omega(x)$; see [30, Corollary 4.13]. Using these properties, one can now see that the proofs in Section 6 go through verbatim after replacing Ω by $\mathcal{C}_\Omega^c(\Gamma)$ and $\partial\Omega$ by $\partial_{\mathrm{i}}\mathcal{C}_\Omega^c(\Gamma)$. Thus the analogues of Propositions 6.5, 6.7 and 6.8 hold; also see [37].

A.5 Convex cocompact examples: hyperbolic groups

Suppose $\Gamma \leq \mathrm{Aut}(\Omega)$ is a convex cocompact group that is word hyperbolic. We claim that Γ *is a rank-one convex cocompact group*. Indeed, [30, Theorem 1.15] implies that word hyperbolicity of Γ is equivalent to the property that $\partial_{\mathrm{i}}\mathcal{C}_\Omega^c(\Gamma)$ does not contain any nontrivial projective line segments. Then there are no half triangles in $\mathcal{C}_\Omega^c(\Gamma)$. Moreover, any infinite-order element γ has an axis [30, Corollary 7.4]. Thus every such γ is a convex cocompact rank-one isometry and the claim follows.

A.6 Convex cocompact examples: relatively hyperbolic groups

Proposition A.4 *Suppose $\Gamma \leq \mathrm{Aut}(\Omega)$ is a convex cocompact group that is relatively hyperbolic with respect to $\{A_1, A_2, \dots, A_m\}$, where each A_i is a virtually free abelian group of rank at least two. Then Γ is either a rank-one convex cocompact group or a virtually abelian group.*

This proposition shows that the divisible examples of Section A.2 and their convex cocompact deformations produce relatively hyperbolic examples that are rank-one convex cocompact. We will spend the rest of this subsection proving this proposition. We will rely on results from [39].

Proof Let $\mathcal{S}_\Gamma$ be the collection of all maximal properly embedded simplices in $\mathcal{C}_\Omega^c(\Gamma)$ of dimension at least two. Since Γ is relatively hyperbolic with respect to virtually abelian subgroups of rank at least two, [39, Theorem 1.7] implies that $(\mathcal{C}_\Omega^c(\Gamma), \mathrm{d}_\Omega)$ is a *Hilbert geometry with isolated simplices*, ie $\mathcal{S}_\Gamma$ is closed and discrete in the local Hausdorff topology induced by d_Ω. In this case, [39, Theorem 1.18] implies that for each $i \in \{1, \dots, m\}$, we can assume $A_i = \mathrm{Stab}_\Gamma(S_i)$, where S_i is a maximal properly embedded simplex in $\mathcal{C}_\Omega^c(\Gamma)$ of dimension ≥ 2 and $\mathcal{S}_\Gamma = \bigsqcup_{i=1}^m \Gamma \cdot S_i$. We will require the following result regarding simplices in $\mathcal{S}_\Gamma$.

Proposition A.5 [39, Theorem 1.8] *Suppose Γ and $\mathcal{S}_\Gamma$ are as above. Then*:

(i) *If $[x,y] \subset \partial_{\mathrm{i}}\mathcal{C}^c_\Omega(\Gamma)$ with $x \neq y$, then there exists $S \in \mathcal{S}_\Gamma$ such that $[x,y] \subset \partial S$.*

(ii) *If $S_1 \neq S_2 \in \mathcal{S}_\Gamma$, then $\#(S_1 \cap S_2) \leq 1$ and $\partial S_1 \cap \partial S_2 = \varnothing$.*

Since Γ is relatively hyperbolic with respect to $\{A_1, A_2, \dots, A_m\}$, [28, Lemma 2.3] implies either

- **Case 1** Γ is virtually gA_ig^{-1} for some $g \in \Gamma$ and $1 \leq i \leq m$, or
- **Case 2** there exists $\gamma \in \Gamma$ such that $\gamma \notin \bigcup_{g\in\Gamma}\bigcup_{i=1}^m gA_ig^{-1} = \bigcup_{S\in\mathcal{S}_\Gamma} \mathrm{Stab}_\Gamma(S)$.

In Case 1, Γ is a virtually abelian group. So we can now assume that we are in Case 2.

Claim *If γ is as in Case 2, then γ is a convex cocompact rank-one isometry.*

From this claim, Proposition A.4 is immediate. □

All that remains is to prove this claim.

Proof of claim As Γ is a convex cocompact group, $\log|(\lambda_1/\lambda_{d+1})(\gamma)| = \tau_{\mathcal{C}^c_\Omega(\Gamma)}(\gamma) > 0$. We first show that γ has an axis in $\mathcal{C}^c_\Omega(\Gamma)$. Let

$$\mathcal{C}^+ := \overline{E^+_\gamma \cap \mathcal{C}^c_\Omega(\Gamma)} \quad \text{and} \quad \mathcal{C}^- := \overline{E^-_\gamma \cap \mathcal{C}^c_\Omega(\Gamma)}.$$

Then $\mathcal{C}^+$ and $\mathcal{C}^-$ are disjoint, nonempty, compact, convex, γ-invariant subsets of $\mathbb{R}^d$. Then the Brouwer fixed point theorem implies the existence of distinct fixed points $\gamma^\pm$ of γ in $\mathcal{C}^\pm$. If $[\gamma^+,\gamma^-] \subset \partial_{\mathrm{i}}\mathcal{C}^c_\Omega(\Gamma)$, Proposition A.5 implies that there exists $S \in \mathcal{S}_\Gamma$ such that $[\gamma^+,\gamma^-] \subset \partial S$. Then $\partial(\gamma S) \cap \partial S \supset [\gamma^+,\gamma^-]$ and Proposition A.5 implies that $\gamma S = S$. Thus, $\gamma \in \mathrm{Stab}_\Gamma(S)$. This contradiction implies that $(\gamma^+,\gamma^-) \subset \mathcal{C}^c_\Omega(\Gamma)$ and is an axis of γ.

Suppose $A_\gamma := [A^+_\gamma, A^-_\gamma]$ is an axis of γ contained in a half triangle in $\mathcal{C}^c_\Omega(\Gamma)$: $[A^+_\gamma, z] \cup [z, A^-_\gamma] \subset \partial_{\mathrm{i}}\mathcal{C}^c_\Omega(\Gamma)$. Then, by Proposition A.5, there exist $S^\pm \in \mathcal{S}_\Gamma$ such that $[z, A^\pm_\gamma] \subset \partial S^\pm$. Since $z \in \partial S^+ \cap \partial S^-$, Proposition A.5 implies that $S := S^+ = S^-$ and $A_\gamma \subset S$. Since γ acts by a translation along A_γ, $\gamma S \cap S \supset A_\gamma$ which implies $\#(\gamma S \cap S) = \infty$. Then by Proposition A.5, $\gamma S = S$. Thus $\gamma \in \mathrm{Stab}_\Gamma(S)$, a contradiction. Thus A_γ is not contained in any half triangle in $\mathcal{C}^c_\Omega(\Gamma)$. This proves the claim. □

Appendix B Contracting elements

Fix a proper geodesic metric space (X, d) and a group G that acts properly isometrically on X. If $x \in X$ and $R > 0$, let $B(x,R) := \{y \in X \mid \mathrm{d}(x,y) < R\}$. If $\mathcal{A} \subset X$ and $x \in X$, let the closest-point projection onto $\mathcal{A}$ be defined by $\rho_{\mathcal{A}}(x) := \{y \in \mathcal{A} \mid \mathrm{d}(x,y) = \mathrm{d}(x,\mathcal{A})\}$. We let $\mathcal{N}_r(\mathcal{A}) := \{y \in X \mid \mathrm{d}(y,\mathcal{A}) < r\}$ and $\overline{\mathcal{N}_r}(\mathcal{A}) := \{y \in X \mid \mathrm{d}(y,\mathcal{A}) \leq r\}$) denote the open and the closed r-neighborhoods of $\mathcal{A}$, respectively.

In [17], Bestvina and Fujiwara introduced the following notion of contracting subsets.

Definition B.1 A set $\mathcal{A} \subset X$ is B-contracting if there exists a constant B such that if $x \in X$, $R > 0$ and $B(x,R) \cap \mathcal{A} = \varnothing$, then $\mathrm{diam}\big(\rho_{\mathcal{A}}(B(x,R))\big) \leq B$.

We will, however, use a related but stronger notion of contracting subsets introduced in [32].

Definition B.2 [32] Fix a geodesic path system $\mathcal{PS}$ on X; see Definition 9.1. A set $\mathcal{A} \subset X$ is a *contracting subset in the sense of BF* if there exists a constant C such that if $\sigma \subset X$ is a geodesic in $\mathcal{PS}$ for which $\mathrm{d}(\sigma, \mathcal{A}) > C$, then

$$\operatorname{diam}(\rho_{\mathcal{A}}(\sigma)) \le C.$$

Suppose G preserves $\mathcal{PS}$. Then $g \in G$ is a *contracting element in the sense of BF* if for any $x_0 \in X$,

(i) g has infinite order and $\langle g\rangle \cdot x_0$ is a quasi-isometric embedding of $\mathbb{Z}$ in X, and

(ii) $\langle g\rangle \cdot x_0$ is a contracting subset in the sense of BF.

Remark B.3 If X is a proper CAT(0) geodesic metric space, Bestvina and Fujiwara [17, Corollary 3.4] prove that Definitions B.1 and B.2 are equivalent. In fact, they prove this equivalence for any metric space that satisfies their axioms DD and FT; see [17]. However, it is unclear whether Definitions B.1 and B.2 are equivalent in complete generality. We will discuss this in Proposition B.6 below. Proposition B.6 suggests that it is unlikely that these definitions are equivalent in general.

We will now prove Proposition 9.8, that contraction in the sense of BF is equivalent to Sisto's notion (see Definitions 9.2 and 9.4), in the context of a geodesic path system. Before starting the proof, we record the following immediate consequence of Definition B.2.

Lemma B.4 *Suppose $A \subset X$ is* ***contracting in the sense of BF*** *with constant C. Let $\sigma_{x,x'} \in \mathcal{PS}$ be a geodesic joining x and x' such that $\sigma_{x,x'} \cap \overline{\mathcal{N}_{2C}(A)} = \{x'\}$. Then $\sup_{a \in \rho_A(x)} \mathrm{d}(a, x') \le 3C$.*

Proof Since $\mathrm{d}(\sigma_{x,x'}, A) \ge 2C > C$, it follows that $\mathrm{d}(y, y') \le C$ for any $y \in \rho_A(x)$ and any $y' \in \rho_A(x')$. But $\mathrm{d}(a', x') \le 2C$ for any $a' \in \rho_A(x')$. Hence the conclusion. □

Proposition B.5 (Proposition 9.8) *Suppose $(X, \mathcal{PS})$ is a geodesic path system. Then*:

(i) *$\mathcal{A} \subset X$ is $\mathcal{PS}$-contracting if and only if $\mathcal{A}$ is contracting in the sense of BF.*

(ii) *If G preserves $\mathcal{PS}$, then $g \in G$ is a contracting element for $(X, \mathcal{PS})$ if and only if $g \in G$ is a contracting element in the sense of BF.*

Proof It suffices to prove only part (i), as (ii) then follows from definitions. We will use $\sigma_{p,q}$ to denote a geodesic path in $\mathcal{PS}$ joining p and q. We now start the proof of (i).

($\Longrightarrow$) Suppose $\mathcal{A}$ is contracting in the sense of BF. Define a projection map $\pi\colon X \to \mathcal{A}$ by choosing $\pi(x) \in \rho_{\mathcal{A}}(x)$ for each $x \in X$. We remark that such a map π is coarsely unique, ie any other such map π' has the property that $\mathrm{d}(\pi(x), \pi'(x)) \le 2C$ for any $x \in X$. Indeed, for any x such that $\mathrm{d}(x, \mathcal{A}) > C$, we have that $\operatorname{diam}(\rho_{\mathcal{A}}(x)) \le C$ as $\{x\}$ is a geodesic in $\mathcal{PS}$. On the other hand, if $\mathrm{d}(x, \mathcal{A}) \le C$, then $\sup_{a \in \rho_{\mathcal{A}}(x)} \mathrm{d}(x, a) \le C$ and hence $\operatorname{diam}(\rho_{\mathcal{A}}(x)) \le 2C$. Hence the remark.

We now show that π satisfies Definition 9.2 with constant $3C$. Clearly, if $x \in \mathscr{A}$, then $\mathrm{d}(x, \pi(x)) = 0$. Now suppose that $x, y \in X$ is such that $\mathrm{d}(\pi(x), \pi(y)) \geq 3C$. Then $\operatorname{diam}(\rho_{\mathscr{A}}(\sigma_{x,y})) \geq 3C > C$. Then $d(\sigma_{x,y}, \mathscr{A}) \leq C$ and thus $\sigma_{x,y} \cap \overline{\mathscr{N}_{2C}(\mathscr{A})} \neq \varnothing$. Let x' be the first point along $\sigma_{x,y}$ that intersects $\overline{\mathscr{N}_{2C}(\mathscr{A})}$ (assume that $\sigma_{x,y}$ is continuously parametrized in the direction from x to y). If $x = x'$, then $\mathrm{d}(\pi(x), x') \leq 2C$. Otherwise apply Lemma B.4 to $\sigma_{x,x'} \subset \sigma_{x,y}$ to see that $\mathrm{d}(\pi(x), x') \leq 3C$. Similarly, if y' is the last point along $\sigma_{x,y}$ where $\sigma_{x,y}$ intersects $\overline{\mathscr{N}_{2C}(A)}$, then $\mathrm{d}(y', \pi(y)) \leq 3C$. Thus π is a contracting projection with constant $3C$.

($\Longleftarrow$) Suppose $\pi\colon X \to \mathscr{A}$ is a contracting projection with constant C. By Lemma 9.3, it follows that $\sup_{a \in \rho_{\mathscr{A}}(x)} \mathrm{d}(a, \pi(x)) \leq 2C$ for any $x \in X$. Let $\sigma_{x,y} \in \mathscr{PS}$ be such that $\mathrm{d}(\sigma_{x,y}, \mathscr{A}) > 5C$. If possible, let there exist $a_1 \in \rho_{\mathscr{A}}(x)$ and $b_1 \in \rho_{\mathscr{A}}(y)$ such that $\mathrm{d}(a_1, b_1) > 5C$. Then

$$\mathrm{d}(\pi(x), \pi(y)) \geq \mathrm{d}(a_1, b_1) - \mathrm{d}(a_1, \pi(x)) - \mathrm{d}(b_1, \pi(y)) > C.$$

Then, $\sigma_{x,y}$ must intersect $\mathscr{N}_C(\mathscr{A})$, a contradiction. □

B.1 Comparison between Definitions B.1 and B.2

To discuss the relationship between B.1 and B.2 for a general metric space, we need the following condition (♦). We will say that $A \subset X$ *satisfies* (♦) *if there exists a constant C such that for any $x \in X$, $z \in A$ and $a \in \rho_A(x)$,*

$$\mathrm{d}(x, z) \geq \mathrm{d}(x, a) + \mathrm{d}(a, z) - C.$$

We will now show:

Proposition B.6 *Fix a proper geodesic metric space X and a geodesic path system $\mathscr{PS}$ on X. Then $A \subset X$ is contracting in the sense of BF if and only if it satisfies (♦) and Definition B.1.*

The implication ($\Longrightarrow$) follows from [48, Lemma 2.10]. Note that in [48], condition (♦) is called (AP1) while Definition B.1 is called (AP2). This direction is then immediate from [48, Lemma 2.10]. The proof of the converse ($\Longleftarrow$) follows from the next two lemmas. For $p, q \in X$, we will denote by $\sigma_{p,q}$ a geodesic in $\mathscr{PS}$ joining p and q.

Lemma B.7 *Suppose $A \subset X$ satisfies Definition B.2. Then A satisfies (♦).*

Proof Fix any $x \in X$, $z \in A$ and $a \in \rho_A(x)$. It suffices to only consider the case when $\mathrm{d}(x, A) > 2C$. Let $x' \in \sigma_{x,z}$ be the first point along $\sigma_{x,z}$ that intersects $\overline{\mathscr{N}_{2C}(A)}$ (assume that $\sigma_{x,z}$ is continuously parametrized in the direction from x to z). By Lemma B.4, $\mathrm{d}(x', a) \leq 3C$. Then $\mathrm{d}(x, a) - \mathrm{d}(x, x') \leq \mathrm{d}(x', a) \leq 3C$ and $\mathrm{d}(z, a) - \mathrm{d}(z, x') \leq \mathrm{d}(x', a) \leq 3C$. Since $x' \in \sigma_{x,z}$, it follows that $\mathrm{d}(x, z) = \mathrm{d}(x, x') + \mathrm{d}(x', z)$. Thus

$$\mathrm{d}(x, z) - \mathrm{d}(x, a) - \mathrm{d}(a, z) = (\mathrm{d}(x, x') - \mathrm{d}(x, a)) + (\mathrm{d}(x', z) - \mathrm{d}(a, z)) \geq -6C.$$ □

Lemma B.8 *Suppose $\mathscr{A} \subset X$ satisfies Definition B.2. Then $\mathscr{A}$ also satisfies Definition B.1.*

Proof Proposition B.5 implies that $\mathscr{A}$ is $\mathscr{PS}$-contracting. Then there exists a projection map $\pi_{\mathscr{A}}\colon X \to \mathscr{A}$ with constant C satisfying Definition 9.2. Suppose $x \in X$ and $0 < R < d(x, \mathscr{A})$. We claim that

$\mathrm{diam}(\rho_{\mathcal{A}}(B(x,R))) \le 20C$. By Lemma 9.3, it suffices to prove that $\mathrm{d}(\pi_{\mathcal{A}}(x), \pi_{\mathcal{A}}(y)) \le 8C$ for any $y \in B(x,R)$.

Fix $y \in B(x,R)$ and let $\sigma_{x,y} \in \mathcal{PS}$. Without loss of generality, we can assume that $\mathrm{d}(\pi_{\mathcal{A}}(x), \pi_{\mathcal{A}}(y)) \ge C$. Then there exists $x_1 \in \sigma_{x,y}$ such that $\mathrm{d}(x_1, \pi_{\mathcal{A}}(x)) \le C$. Then

$$|\mathrm{d}(x,x_1) - \mathrm{d}(x,\mathcal{A})| \le \mathrm{d}(x_1, \pi_{\mathcal{A}}(x)) + \sup_{a \in \rho_{\mathcal{A}}(x)} \mathrm{d}(a, \pi_{\mathcal{A}}(x)) \le 3C.$$

Thus, $\mathrm{d}(y,x_1) = \mathrm{d}(y,x) - \mathrm{d}(x,x_1) \le \mathrm{d}(y,x) - \mathrm{d}(x,\mathcal{A}) + 3C$. As $\mathrm{d}(y,x) < \mathrm{d}(x,\mathcal{A})$, we get that $\mathrm{d}(y,x_1) \le 3C$. Then

$$\mathrm{d}(y, \pi_{\mathcal{A}}(y)) \le \mathrm{d}(y, \pi_{\mathcal{A}}(x)) \le \mathrm{d}(y,x_1) + \mathrm{d}(x_1, \pi_{\mathcal{A}}(x)) \le 4C,$$

which implies that

$$\mathrm{d}(\pi_{\mathcal{A}}(y), \pi_{\mathcal{A}}(x)) \le \mathrm{d}(\pi_{\mathcal{A}}(y), y) + \mathrm{d}(y,x_1) + \mathrm{d}(x_1, \pi_{\mathcal{A}}(x)) \le 8C. \qquad \square$$

References

[1] **S A Ballas**, **J Danciger**, **G-S Lee**, *Convex projective structures on nonhyperbolic three-manifolds*, Geom. Topol. 22 (2018) 1593–1646 MR Zbl

[2] **W Ballmann**, *Axial isometries of manifolds of nonpositive curvature*, Math. Ann. 259 (1982) 131–144 MR Zbl

[3] **W Ballmann**, *Nonpositively curved manifolds of higher rank*, Ann. of Math. 122 (1985) 597–609 MR Zbl

[4] **W Ballmann**, *Lectures on spaces of nonpositive curvature*, DMV Seminar 25, Birkhäuser, Basel (1995) MR Zbl

[5] **W Ballmann**, **M Brin**, *Orbihedra of nonpositive curvature*, Inst. Hautes Études Sci. Publ. Math. 82 (1995) 169–209 MR Zbl

[6] **Y Benoist**, *Propriétés asymptotiques des groupes linéaires*, Geom. Funct. Anal. 7 (1997) 1–47 MR Zbl

[7] **Y Benoist**, *Automorphismes des cônes convexes*, Invent. Math. 141 (2000) 149–193 MR Zbl

[8] **Y Benoist**, *Convexes divisibles, I*, from "Algebraic groups and arithmetic" (S G Dani, G Prasad, editors), Tata Inst. Fund. Res., Mumbai (2004) 339–374 MR Zbl

[9] **Y Benoist**, *Convexes divisibles, III*, Ann. Sci. École Norm. Sup. 38 (2005) 793–832 MR Zbl

[10] **Y Benoist**, *Convexes divisibles, IV: Structure du bord en dimension* 3, Invent. Math. 164 (2006) 249–278 MR Zbl

[11] **Y Benoist**, *Convexes hyperboliques et quasiisométries*, Geom. Dedicata 122 (2006) 109–134 MR Zbl

[12] **Y Benoist**, *A survey on divisible convex sets*, from "Geometry, analysis and topology of discrete groups" (L Ji, K Liu, L Yang, S-T Yau, editors), Adv. Lect. Math. 6, International, Somerville, MA (2008) 1–18 MR Zbl

[13] **Y Benoist**, **J-F Quint**, *Random walks on reductive groups*, Ergebnisse der Math. 62, Springer (2016) MR Zbl

[14] **M Bestvina**, **K Bromberg**, **K Fujiwara**, *Constructing group actions on quasi-trees and applications to mapping class groups*, Publ. Math. Inst. Hautes Études Sci. 122 (2015) 1–64 MR Zbl

[15] **M Bestvina**, **K Bromberg**, **K Fujiwara**, *Bounded cohomology with coefficients in uniformly convex Banach spaces*, Comment. Math. Helv. 91 (2016) 203–218 MR Zbl

[16] **M Bestvina**, **K Bromberg**, **K Fujiwara**, **A Sisto**, *Acylindrical actions on projection complexes*, Enseign. Math. 65 (2019) 1–32 MR Zbl

[17] **M Bestvina**, **K Fujiwara**, *A characterization of higher rank symmetric spaces via bounded cohomology*, Geom. Funct. Anal. 19 (2009) 11–40 MR Zbl

[18] **P-L Blayac**, *Patterson–Sullivan densities in convex projective geometry*, preprint (2021) arXiv 2106.08089

[19] **P-L Blayac**, *The boundary of rank-one divisible convex sets*, Bull. Soc. Math. France 152 (2024) 1–17 MR Zbl

[20] **P-L Blayac**, **G Viaggi**, *Divisible convex sets with properly embedded cones*, preprint (2023) arXiv 2302.07177

[21] **M Burger**, **N Monod**, *Continuous bounded cohomology and applications to rigidity theory*, Geom. Funct. Anal. 12 (2002) 219–280 MR Zbl

[22] **K Burns**, **R Spatzier**, *Manifolds of nonpositive curvature and their buildings*, Inst. Hautes Études Sci. Publ. Math. 65 (1987) 35–59 MR Zbl

[23] **P-E Caprace**, **M Sageev**, *Rank rigidity for* CAT(0) *cube complexes*, Geom. Funct. Anal. 21 (2011) 851–891 MR Zbl

[24] **S Choi**, **G-S Lee**, **L Marquis**, *Convex projective generalized Dehn filling*, Ann. Sci. École Norm. Sup. 53 (2020) 217–266 MR Zbl

[25] **D Cooper**, **D D Long**, **S Tillmann**, *On convex projective manifolds and cusps*, Adv. Math. 277 (2015) 181–251 MR Zbl

[26] **M Cordes**, *Morse boundaries of proper geodesic metric spaces*, Groups Geom. Dyn. 11 (2017) 1281–1306 MR Zbl

[27] **M Crampon**, **L Marquis**, *Finitude géométrique en géométrie de Hilbert*, Ann. Inst. Fourier (Grenoble) 64 (2014) 2299–2377 MR Zbl

[28] **F Dahmani**, **V Guirardel**, *Recognizing a relatively hyperbolic group by its Dehn fillings*, Duke Math. J. 167 (2018) 2189–2241 MR Zbl

[29] **F Dahmani**, **V Guirardel**, **D Osin**, *Hyperbolically embedded subgroups and rotating families in groups acting on hyperbolic spaces*, Mem. Amer. Math. Soc. 1156, Amer. Math. Soc., Providence, RI (2017) MR Zbl

[30] **J Danciger**, **F Guéritaud**, **F Kassel**, *Convex cocompact actions in real projective geometry*, preprint (2017) arXiv 1704.08711

[31] **R Frigerio**, *Bounded cohomology of discrete groups*, Math. Surv. Monogr. 227, Amer. Math. Soc., Providence, RI (2017) MR Zbl

[32] **I Gekhtman**, **W-y Yang**, *Counting conjugacy classes in groups with contracting elements*, J. Topol. 15 (2022) 620–665 MR Zbl

[33] **M Gromov**, *Hyperbolic groups*, from "Essays in group theory", Math. Sci. Res. Inst. Publ. 8, Springer (1987) 75–263 MR Zbl

[34] **F Guéritaud**, **O Guichard**, **F Kassel**, **A Wienhard**, *Anosov representations and proper actions*, Geom. Topol. 21 (2017) 485–584 MR Zbl

[35] **Y Guivarc'h**, *Produits de matrices aléatoires et applications aux propriétés géométriques des sous-groupes du groupe linéaire*, Ergodic Theory Dynam. Systems 10 (1990) 483–512 MR Zbl

[36] **P de la Harpe**, *On Hilbert's metric for simplices*, from "Geometric group theory, I" (G A Niblo, M A Roller, editors), Lond. Math. Soc. Lect. Note Ser. 181, Cambridge Univ. Press (1993) 97–119 MR Zbl

[37] **M Islam**, *Rank one phenomena in convex projective geometry*, PhD thesis, University of Michigan (2021) Available at `https://www.proquest.com/docview/2593621239`

[38] **M Islam**, **A Zimmer**, *A flat torus theorem for convex co-compact actions of projective linear groups*, J. Lond. Math. Soc. 103 (2021) 470–489 MR Zbl

[39] **M Islam**, **A Zimmer**, *Convex cocompact actions of relatively hyperbolic groups*, Geom. Topol. 27 (2023) 417–511 MR Zbl

[40] **M Kapovich**, *Convex projective structures on Gromov–Thurston manifolds*, Geom. Topol. 11 (2007) 1777–1830 MR Zbl

[41] **P Kelly**, **E Straus**, *Curvature in Hilbert geometries*, Pacific J. Math. 8 (1958) 119–125 MR Zbl

[42] **G A Margulis**, *Discrete subgroups of semisimple Lie groups*, Ergebnisse der Math. 17, Springer (1991) MR Zbl

[43] **L Marquis**, *Around groups in Hilbert geometry*, from "Handbook of Hilbert geometry" (A Papadopoulos, M Troyanov, editors), IRMA Lect. Math. Theor. Phys. 22, Eur. Math. Soc., Zürich (2014) 207–261 MR

[44] **R D Nussbaum**, *Hilbert's projective metric and iterated nonlinear maps*, Mem. Amer. Math. Soc. 391, Amer. Math. Soc., Providence, RI (1988) MR Zbl

[45] **D Osin**, *Acylindrically hyperbolic groups*, Trans. Amer. Math. Soc. 368 (2016) 851–888 MR Zbl

[46] **S Pinella**, *Hilbert domains, conics, and rigidity*, PhD thesis, University of Michigan (2020) Available at `https://www.proquest.com/docview/2481054689`

[47] **R Ricks**, *A rank rigidity result for* CAT(0) *spaces with one-dimensional Tits boundaries*, Forum Math. 31 (2019) 1317–1330 MR Zbl

[48] **A Sisto**, *Projections and relative hyperbolicity*, Enseign. Math. 59 (2013) 165–181 MR Zbl

[49] **A Sisto**, *Contracting elements and random walks*, J. Reine Angew. Math. 742 (2018) 79–114 MR Zbl

[50] **A Zimmer**, *A higher-rank rigidity theorem for convex real projective manifolds*, Geom. Topol. 27 (2023) 2899–2936 MR Zbl

Department of Mathematics, University of Michigan
Ann Arbor, MI, United States

Current address: *Max Planck Institute for Mathematics in the Sciences*
Leipzig, Germany

`mitul.islam@gmail.com`

Proposed: Mladen Bestvina
Seconded: Benson Farb, David Fisher

Received: 2 July 2021
Revised: 26 November 2023

msp

Geometry & Topology **29**:3 (2025) 1237–1281
DOI: 10.2140/gt.2025.29.1237
Published: 31 May 2025

Random unitary representations of surface groups II: The large n limit

MICHAEL MAGEE

Let Σ_g be a closed surface of genus $g \geq 2$ and Γ_g denote the fundamental group of Σ_g. We establish a generalization of Voiculescu's theorem on the asymptotic $*$-freeness of Haar unitary matrices from free groups to Γ_g. We prove that, for a random representation of Γ_g into $\mathsf{SU}(n)$, with law given by the volume form arising from the Atiyah–Bott–Goldman symplectic form on moduli space, the expected value of the trace of a fixed nonidentity element of Γ_g is bounded as $n \to \infty$. The proof involves an interplay between Dehn's work on the word problem in Γ_g and classical invariant theory.

14H60, 22D10, 46L54; 20C30, 20C35, 32G15, 70S15

To Oona

1 Introduction

In a foundational series of papers, Voiculescu [1985; 1986; 1987; 1990; 1991] developed a robust theory of noncommuting random variables that became known as *free probability*. One of the initial landmarks of this theory is the following result. Let $\boldsymbol{F}_r$ denote the noncommutative free group of rank r. Let $\mathsf{U}(n)$ denote the group of $n \times n$ complex unitary matrices. For any $w \in \boldsymbol{F}_r$ we obtain a *word map* $w\colon \mathsf{U}(n)^r \to \mathsf{U}(n)$ by substituting matrices for generators of $\boldsymbol{F}_r$. Let $\mu^{\mathrm{Haar}}_{\mathsf{U}(n)^r}$ denote the probability Haar measure on $\mathsf{U}(n)^r$ and $\mathrm{Tr}\colon \mathsf{U}(n) \to \mathbb{C}$ the standard trace. Any integral over a compact group will be done with respect to the probability Haar measure, denoted by $d\mu$.

A simplified version of Voiculescu's result [1991, Theorem 3.8] can be formulated as follows:[1]

Theorem 1.1 (Voiculescu) *For any nonidentity $w \in \boldsymbol{F}_r$, as $n \to \infty$,*

$$\int_{\mathsf{U}(n)^r} \operatorname{Tr}(w(x))\, d\mu(x) = o_w(n). \tag{1-1}$$

We describe the interpretation of Theorem 1.1 as convergence of noncommutative random variables in a moment. Before this, we explain the main result of the current paper.

Another way to think about the integral (1-1), which invites generalization, is to identify $\mathsf{U}(n)^r$ with $\operatorname{Hom}(\boldsymbol{F}_r, \mathsf{U}(n))$ and Haar measure as a natural probability measure on this *representation variety*. Now it is natural to ask whether there are other infinite discrete groups G besides $\boldsymbol{F}_r$ such that $\operatorname{Hom}(G, \mathsf{U}(n))$ has a natural measure, and whether similar phenomena as in Theorem 1.1 may hold. *The main point of this paper is to establish the analog of Theorem 1.1 when $\boldsymbol{F}_r$ is replaced by the fundamental group of a compact surface of genus at least* 2.

We now explain this generalization of Theorem 1.1; for technical reasons it superficially looks slightly different, as follows:

(1) The integral (1-1) is equal to 0 if $w \notin [\boldsymbol{F}_r, \boldsymbol{F}_r]$, the commutator subgroup of $\boldsymbol{F}_r$ [Magee and Puder 2015, Claim 3.1], and, if $w \in [\boldsymbol{F}_r, \boldsymbol{F}_r]$, the value of (1-1) is, for $n \geq n_0(w)$, the same as the corresponding integral over $\mathsf{SU}(n)^r \leq \mathsf{U}(n)^r$, where $\mathsf{SU}(n)$ is the subgroup of determinant one matrices [Magee 2022, Proposition 3.1]. So in all cases of interest we can replace $\mathsf{U}(n)$ by $\mathsf{SU}(n)$ in (1-1).

(2) Since $\operatorname{Tr} \circ w$ is invariant under the diagonal conjugation action of $\mathsf{SU}(n)$ on $\operatorname{Hom}(\boldsymbol{F}_r, \mathsf{SU}(n)) \cong \mathsf{SU}(n)^r$, the integral $\int_{\mathsf{SU}(n)^r} \operatorname{Tr}(w(x))\, d\mu(x)$ can be written as one over $\operatorname{Hom}(\boldsymbol{F}_r, \mathsf{SU}(n))/\mathsf{PSU}(n)$. Here $\mathsf{PSU}(n)$ is $\mathsf{SU}(n)$ modulo its center.

For $g \geq 2$ let Σ_g denote a closed topological surface of genus g. We let Γ_g denote the fundamental group of Σ_g with explicit presentation

$$\Gamma_g = \langle a_1, b_1, \ldots, a_g, b_g \mid [a_1, b_1] \cdots [a_g, b_g] \rangle.$$

The most natural measure on $\operatorname{Hom}(\Gamma_g, \mathsf{SU}(n))/\mathsf{PSU}(n)$ to replace the measure induced by Haar measure on $\operatorname{Hom}(\boldsymbol{F}_r, \mathsf{SU}(n))/\mathsf{PSU}(n)$ is called the Atiyah–Bott–Goldman measure. The definition of this measure involves removing singular parts of $\operatorname{Hom}(\Gamma_g, \mathsf{SU}(n))/\mathsf{PSU}(n)$. Indeed, let $\operatorname{Hom}(\Gamma_g, \mathsf{SU}(n))^{\mathrm{irr}}$ denote the collection of homomorphisms that are irreducible as linear representations. Then

$$\mathcal{M}_{g,n} := \operatorname{Hom}(\Gamma_g, \mathsf{SU}(n))^{\mathrm{irr}}/\mathsf{PSU}(n)$$

[1] Voiculescu's result [1991, Theorem 3.8] is more general than what we state here, also involving a deterministic sequence of unitary matrices.

is a smooth manifold [Goldman 1984]. Moreover there is a symplectic form $\omega_{g,n}$ on $\mathcal{M}_{g,n}$, called the Atiyah–Bott–Goldman form after [Atiyah and Bott 1983; Goldman 1984]. This symplectic form gives, in the usual way, a volume form on $\mathcal{M}_{g,n}$ denoted by $\mathrm{Vol}_{\mathcal{M}_{g,n}}$. For many more details, see [Goldman 1984] or our prequel paper [Magee 2022, Section 2.7].

For any $\gamma \in \Gamma$, we obtain a function $\mathrm{Tr}_\gamma \colon \mathrm{Hom}(\Gamma_g, \mathsf{SU}(n)) \to \mathbb{C}$ defined by

$$\mathrm{Tr}_\gamma(\phi) := \mathrm{Tr}(\phi(\gamma)).$$

This function descends to a function $\mathrm{Tr}_\gamma \colon \mathcal{M}_{g,n} \to \mathbb{C}$. We are interested in the expected value

$$\mathbb{E}_{g,n}[\mathrm{Tr}_\gamma] := \frac{\int_{\mathcal{M}_{g,n}} \mathrm{Tr}_\gamma \, d\,\mathrm{Vol}_{\mathcal{M}_{g,n}}}{\int_{\mathcal{M}_{g,n}} d\,\mathrm{Vol}_{\mathcal{M}_{g,n}}}.$$

The main theorem of this paper is the following:

Theorem 1.2 *Let $g \geq 2$. If $\gamma \in \Gamma_g$ is not the identity, then $\mathbb{E}_{g,n}[\mathrm{Tr}_\gamma] = O_\gamma(1)$ as $n \to \infty$.*

The noncommutative probabilistic consequences of Theorem 1.2 will be discussed in the next section.

1.1 Noncommutative probability

We follow [Voiculescu et al. 1992]. A *noncommutative probability space* is a pair $(\mathcal{B}, \tau)$ where $\mathcal{B}$ is a complex unital algebra and τ is a linear functional on $\mathcal{B}$ such that $\tau(1) = 1$. Let $\mathbb{C}\langle x_1, \dots, x_r\rangle$ denote the free noncommutative unital algebra in indeterminates $x_1, \dots, x_r$. A *random variable* in $(\mathcal{B}, \tau)$ is an element of $\mathcal{B}$. If $(X_1, \dots, X_r) \in \mathcal{B}^r$ are random variables in $(\mathcal{B}, \tau)$, their *joint distribution* is defined to be the linear functional

$$\tilde{\tau} \colon \mathbb{C}\langle x_1, \dots, x_r\rangle \to \mathbb{C}$$

given by $\tilde{\tau}(z) := \tau(\Phi(z))$, where $\Phi \colon \mathbb{C}\langle x_1, \dots, x_r\rangle \to \mathcal{B}$ is the linear map defined by $\Phi(x_i) = X_i$. For a linear functional $\tilde{\tau}_\infty \colon \mathbb{C}\langle x_1, \dots, x_r\rangle \to \mathbb{C}$ with $\tilde{\tau}_\infty(1) = 1$, we say that a sequence of random variables $(X_1^{(n)}, \dots, X_r^{(n)}) \in (\mathcal{B}_n, \tau_n)$ converges in distribution as $n \to \infty$ to $\tilde{\tau}_\infty$ if $\tilde{\tau}_n$ converges pointwise to $\tilde{\tau}_\infty$ on $\mathbb{C}\langle x_1, \dots, x_r\rangle$.

A very concrete example of this phenomenon is as follows. The function

$$\tau_n \colon \boldsymbol{F}_r \to \mathbb{C}, \quad \tau_n(w) := \frac{1}{n} \int_{\mathsf{U}(n)^r} \mathrm{Tr}(w(x)) \, d\mu(x)$$

extends to a linear functional τ_n on the algebra $\mathbb{C}[\boldsymbol{F}_r]$ with $\tau_n(\mathrm{id}) = 1$. From this point of view, Theorem 1.1 implies the following statement:

Theorem 1.3 (Voiculescu) *Let $r \geq 0$ and $X_1, \dots, X_r$ denote fixed generators of $\boldsymbol{F}_r$, and $\overline{X}_1, \dots. \overline{X}_r$ denote their inverses, ie $\overline{X}_i = X_i^{-1}$. The random variables $X_1, \dots, X_r, \overline{X}_1, \dots. \overline{X}_r$ in the noncommutative probability spaces $(\mathbb{C}[\boldsymbol{F}_r], \tau_n)$ converge as $n \to \infty$ to a limiting distribution*

$$\tilde{\tau}_\infty \colon \mathbb{C}\langle x_1, \dots, x_r, \bar{x}_1, \dots, \bar{x}_r\rangle \to \mathbb{C}$$

that is completely determined by (1-1). Indeed, if w is any monomial in $x_1, \dots, x_r, \bar{x}_1, \dots, \bar{x}_r$, then $\tilde{\tau}_\infty(w) = 1$ if and only if, after identifying $\bar{x}_i$ with x_i^{-1}, w reduces to the identity in $\boldsymbol{F}_r = \langle x_1, \dots, x_r\rangle$, and $\tilde{\tau}_\infty(w) = 0$ otherwise.

In the language of [Voiculescu 1991], in the limiting noncommutative probability space

$$(\mathbb{C}\langle x_1, \dots, x_r, \bar{x}_1, \dots, \bar{x}_r\rangle, \tilde{\tau}_\infty),$$

the subalgebras

$$\mathscr{A}_1 := \mathbb{C}\langle x_1, \bar{x}_1\rangle, \quad \dots, \quad \mathscr{A}_r := \mathbb{C}\langle x_r, \bar{x}_r\rangle$$

are a *free family of subalgebras*: if $a_j \in \mathscr{A}_{i_j}$ for $j \in [q]$ with $i_1 \neq i_2 \neq \cdots \neq i_q$, and $\tilde{\tau}_\infty(a_j) = 0$ for $j \in [q]$, then

$$\tilde{\tau}_\infty(a_1 a_2 \cdots a_q) = 0.$$

Accordingly [Voiculescu 1991, Theorem 3.8], if $\{u_j(n) : 1 \leq j \leq r\}$ are independent Haar-random elements of $\mathsf{U}(n)$, the family $\{\{u_j(n), u_j^*(n)\} : 1 \leq j \leq r\}$ of sets of random variables are *asymptotically free*.

Because Γ_g is not free, asymptotic freeness does not correctly capture the asymptotic behavior of the expected values $\mathbb{E}_{g,n}[\mathrm{Tr}_\gamma]$; however, an analog of Theorem 1.3 is implied by Theorem 1.2. For $\gamma \in \Gamma_g$ let

$$\tau_{g,n}(\gamma) := \frac{1}{n}\mathbb{E}_{g,n}[\mathrm{Tr}_\gamma].$$

Corollary 1.4 *Let $g \geq 2, a_1, b_1, \dots, a_g, b_g$ denote the previously fixed generators of Γ_g, and $\bar{a}_1, \bar{b}_1, \dots, \bar{a}_g, \bar{b}_g$ denote their inverses. The random variables $a_1, b_1, \dots, a_g, b_g, \bar{a}_1, \bar{b}_1, \dots, \bar{a}_g, \bar{b}_g$ in the noncommutative probability spaces $(\mathbb{C}[\Gamma_g], \tau_{g,n})$ converge in distribution as $n \to \infty$ to a limiting distribution*

$$\tilde{\tau}_{g,\infty} \colon \mathbb{C}\langle x_1, \dots, x_g, y_1, \dots, y_g, \bar{x}_1, \dots, \bar{x}_g, \bar{y}_1, \dots, \bar{y}_g\rangle \to \mathbb{C},$$

where x_i (resp. $y_i, \bar{x}_i, \bar{y}_i$) corresponds to a_i (resp. $b_i, \bar{a}_i, \bar{b}_i$). This can be described explicitly as follows. If w is any monomial in $x_1, \dots, x_g, y_1, \dots, y_g, \bar{x}_1, \dots, \bar{x}_g, \bar{y}_1, \dots, \bar{y}_g$, then $\tilde{\tau}_{g,\infty}(w) = 1$ if and only if w maps to the identity under the map

$$\mathbb{C}\langle x_1, \dots, x_g, y_1, \dots, y_g, \bar{x}_1, \dots, \bar{x}_g, \bar{y}_1, \dots, \bar{y}_g\rangle \to \mathbb{C}[\Gamma_g]$$

obtained by identifying x_i, y_i, $\bar{x}_i$ and $\bar{y}_i$ with the corresponding elements of Γ_g. If w does not map to the identity under this map, then $\tilde{\tau}_{g,\infty}(w) = 0$.

Notice that the estimate given in Theorem 1.2 is stronger than needed to establish Corollary 1.4.

1.2 Related works and further questions

The most closely related existing result to Theorem 1.2 is [Magee and Puder 2023, Theorem 1.2], which establishes Theorem 1.2 when the family of groups $\mathsf{SU}(n)$ is replaced by the family of symmetric groups S_n, and Tr is replaced by the character fix given by the number of fixed points of a permutation. In this case, the result is phrased in terms of integrating over $\mathrm{Hom}(\Gamma_g, S_n)$ with respect to the uniform probability measure. The corresponding result for $\mathrm{Hom}(\boldsymbol{F}_r, S_n)$ was proved much longer ago [Nica 1994].

The problem of integrating geometric functions like Tr_γ over $\mathcal{M}_{g,n}$ is also connected to the work of Mirzakhani, since, as Goldman [1984, Section 2] explains, the Atiyah–Bott–Goldman symplectic form generalizes the Weil–Petersson symplectic form on the Teichmüller space of genus g Riemann surfaces. Mirzakhani [2007] developed a method for integrating geometric functions on moduli spaces of Riemann surfaces with respect to the Weil–Petersson volume form. Although there is certainly a similarity between [loc. cit.] and the current work, here the emphasis is on $n \to \infty$, whereas [loc. cit.] caters to the regime $g \to \infty$; the target group playing the role of $\mathsf{SU}(n)$ is always $\mathsf{PSL}(2, \mathbb{R})$.

We now take the opportunity to mention some questions that Theorem 1.2 leads to. Voiculescu [1991] is able to boost Theorem 1.1 from a convergence in distribution result to a result on convergence in probability; that is, for any $\epsilon > 0$ and fixed $w \in \boldsymbol{F}_r$, the Haar measure of the set

$$\{\phi \in \mathrm{Hom}(\boldsymbol{F}_r, \mathsf{U}(n)) : |\mathrm{Tr}(\phi(w))| \le \epsilon n\}$$

tends to one as $n \to \infty$ [Voiculescu 1991, Theorem 3.9]. To do this, Voiculescu uses that the family of measure spaces $\bigl(\mathrm{Hom}(\boldsymbol{F}_r, \mathsf{U}(n)), \mu\bigr)$ form a *Levy family* in the sense of [Gromov and Milman 1983]. This latter fact relies on an estimate for the first nonzero eigenvalue of the Laplacian on $\mathrm{Hom}(\boldsymbol{F}_r, \mathsf{U}(n))$. It is interesting to ask whether a similar phenomenon holds for the family of measure spaces $(\mathcal{M}_{g,n}, \mu_{g,n}^{\mathrm{ABG}})$, where $\mu_{g,n}^{\mathrm{ABG}}$ is the probability measure corresponding to $\mathrm{Vol}_{\mathcal{M}_{g,n}}$. The fact that $\mathcal{M}_{g,n}$ is noncompact seems to be a significant complication in answering this question using isoperimetric inequalities.

On the other hand, as pointed out to us by a referee, the results of this paper can very likely be extended to give bounds on the variance

$$\mathbb{E}_{g,n}[|\mathrm{Tr}_\gamma|^2]$$

that can be used to improve Theorem 1.2 to the result that, for $\gamma \ne \mathrm{id}$, the normalized traces Tr_γ/n converge in probability to zero as $n \to \infty$. To avoid adding complications to this paper, this will be pursued elsewhere.

In the prequel to this paper [Magee 2022], we proved that, for any fixed $\gamma \in \Gamma_g$, there is an infinite sequence of rational numbers $a_{-1}(\gamma), a_0(\gamma), a_1(\gamma), \ldots \in \mathbb{Q}$ such that, for any $M \in \mathbb{N}$,

$$\mathbb{E}_{g,n}[\mathrm{Tr}_\gamma] = a_{-1}(\gamma)n + a_0(\gamma) + \frac{a_1(\gamma)}{n} + \cdots + \frac{a_{M-1}(\gamma)}{n^{M-1}} + O_{\gamma,M}\left(\frac{1}{n^M}\right) \tag{1-2}$$

as $n \to \infty$. Theorem 1.2 implies that $a_{-1}(\gamma) = 0$ if $\gamma \neq \mathrm{id}$. It is also interesting to understand the other coefficients of this series. This has been accomplished when Γ_g is replaced by $\boldsymbol{F}_r$ in [Magee and Puder 2019], where in fact it is proved that

$$\mathbb{E}_{\boldsymbol{F}_r,n}[\mathrm{Tr}_w] := \int_{\mathsf{U}(n)^r} \mathrm{Tr}(w(x))\, d\mu(x)$$

is given by a *rational* function of n and, in particular, can be expanded as in (1-2). The corresponding coefficients of the Laurent series of $\mathbb{E}_{\boldsymbol{F}_r,n}[\mathrm{Tr}_w]$ are explained in terms of Euler characteristics of subgroups of mapping class groups. One corollary is that, as $n \to \infty$,

$$\mathbb{E}_{\boldsymbol{F}_r,n}[\mathrm{Tr}_w] = O\left(\frac{1}{n^{2\,\mathrm{cl}(w)-1}}\right), \tag{1-3}$$

where $\mathrm{cl}(w)$ is the *commutator length of* w: the minimal number of commutators that w can be written as a product of, or ∞ if $w \notin [\boldsymbol{F}_r, \boldsymbol{F}_r]$. We guess that an estimate like (1-3) should hold for $\mathbb{E}_{g,n}[\mathrm{Tr}_\gamma]$, where commutator length in $\boldsymbol{F}_r$ is replaced by commutator length in Γ_g.

Another strengthening of Theorem 1.1 is the *strong asymptotic freeness* of Haar unitaries. This states that, for any complex linear combination

$$\sum_w a_w w \in \mathbb{C}[\boldsymbol{F}_r],$$

almost surely with respect to Haar random $\phi \in \mathrm{Hom}(\boldsymbol{F}_r, \mathsf{U}(n))$ as $n \to \infty$, we have

$$\left\| \sum_w a_w \phi(w) \right\| \to \left\| \sum_w a_w w \right\|_{\mathrm{Op}(\ell^2(\boldsymbol{F}_r))},$$

where the left-hand side is the operator norm on $\mathbb{C}^n$ with standard Hermitian inner product and the norm on the right-hand side is the operator norm in the regular representation of $\boldsymbol{F}_r$. This result was proved in [Collins and Male 2014]. It is probably very hard to extend this result to Γ_g; the proof of Collins and Male relies on seminal work of Haagerup and Thorbjørnsen [2005] in a way that does not obviously extend to Γ_g.

We finally mention that the expected values $\mathbb{E}_{g,n}[\mathrm{Tr}_\gamma]$ arise as a limiting form of expected values of Wilson loops in 2D Yang–Mills theory, when the coupling constant is set to zero. This will not be discussed in detail here; we refer the reader instead to the introduction of [Magee 2022]. Here we just mention the recent works [Lemoine 2022; Dahlqvist and Lemoine 2023], which make progress on related problems in the Yang–Mills setting.

1.3 Overview of paper

Here we explain the structure of the paper.

In Sections 2.1–2.5, we give some general background to the paper not depending on [Magee 2022]. In Section 2.6, we import results that we proved in the prequel and that are needed here.

At the beginning of Section 3, we state the key result (Theorem 3.1) of the remainder of the paper. To motivate things, Section 3.1 contains a discussion of why the most straightforward approach does not work, and also a discussion of what will follow instead. In the remainder of Section 3, we explain how to augment the Weingarten calculus to arrive to a formula for the key quantity $\mathcal{J}_n(w,\mu,\nu)$ (defined in Proposition 2.9) in combinatorial terms that are "good" for the next part of the argument.

Indeed, in Section 4.1 we explain how each combinatorial datum we encountered in our formula for $\mathcal{J}_n(w,\mu,\nu)$ can be used to build a decorated surface. In Corollary 4.5 we obtain a bound on $\mathcal{J}_n(w,\mu,\nu)$ in terms of the Euler characteristics of some of the surfaces that previously arose. We may restrict to certain surfaces of simplified form by performing two surgery arguments explained in Section 4.2. Given that now we have reduced estimating $\mathcal{J}_n(w,\mu,\nu)$ to estimating Euler characteristics of certain surfaces, in Section 4.3 we formulate a topological result (Proposition 4.8) which suffices to prove Theorem 3.1. Proposition 4.8 is proved in Section 4.5 using arguments related to Dehn's algorithm and the work of Birman and Series. The necessary additional background for this proof is given in Section 4.4.

In Section 5, we show how Theorem 3.1, in conjunction with the results of [Magee 2022], proves Theorem 1.2.

1.4 Notation

We write $\mathbb{N}$ for the natural numbers $\{1,2,3,\dots\}$ and $\mathbb{N}_0 := \mathbb{N}\cup\{0\}$. We write $[n] := \{1,\dots,n\}$ for $n\in\mathbb{N}$ and $[k,l] := \{k,k+1,\dots,l\}$ for $k,l\in\mathbb{N}$. If A and B are two sets, we write $A\backslash B$ for the elements of A not in B. If H is a group and $h_1,h_2\in H$, we write $[h_1,h_2] := h_1h_2h_1^{-1}h_2^{-1}$. We let id denote the identity element of a group. We let $[H,H]$ be the subgroup of H generated by elements of the form $[h_1,h_2]$; this is called the commutator subgroup of H. If V is a complex vector space, for $q\in\mathbb{N}_0$ we let

$$V^{\otimes q} := \underbrace{V\otimes V\otimes\cdots\otimes V}_{q}.$$

We use Vinogradov notation as follows. If f and h are functions of $n\in\mathbb{N}$, we write $f \ll h$ to mean that there are constants $n_0\ge 0$ and $C_0\ge 0$ such that, for $n\ge n_0$, $|f(n)|\le C_0h(n)$. We write $f=O(h)$ to mean $f\ll h$. We write $f\asymp h$ to mean both $f\ll h$ and $h\ll f$. If in any of these statements the implied constants depend on additional parameters, we add these parameters as subscripts to $\ll$, O or $\asymp$. Throughout the paper we view the genus g as fixed and so any implied constant may depend on g.

In this paper, Tr denotes the standard (unnormalized) trace on square complex matrices.

Acknowledgments We thank Benoît Collins, Antoine Dahlqvist, Doron Puder, Sanjaye Ramgoolam, Calum Shearer and Henry Wilton for valuable discussions about this work. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 949143).

2 Background

2.1 Representation theory of symmetric groups

Let S_k denote the symmetric group of permutations of $[k] := \{1, \dots, k\}$, and $\mathbb{C}[S_k]$ denote its group algebra. The group S_0 is by definition the group with one element.

If we refer to $S_l \le S_k$ with $l \le k$, we always view S_l as the subgroup of permutations that fix every element of $[l+1, k] := \{l+1, \dots, k\}$. We write $S'_r \le S_k$ for the subgroup of permutations that fix every element of $[k-r]$. As a consequence, we obtain fixed inclusions $\mathbb{C}[S_l] \subset \mathbb{C}[S_k]$ for l and k as above. When we write $S_l \times S_{k-l} \le S_k$, the first factor is S_l and the second factor is S'_{k-l}.

A *Young diagram* λ is a left-aligned contiguous collection of identical square boxes in the plane such that the number of boxes in each row is nonincreasing from top to bottom. We write λ_i for the number of boxes in the i^{th} row of λ and say $\lambda \vdash k$ if λ has k boxes. We write $\ell(\lambda)$ for the number of rows of λ. For each $\lambda \vdash k$, there is a *Young subgroup*

$$S_\lambda := S_{\lambda_1} \times S_{\lambda_2} \times \cdots \times S_{\lambda_{\ell(\lambda)}} \le S_k,$$

where the factors are subgroups in the obvious way, according to the increasing order of $[k]$.

The equivalence classes of irreducible representations of S_k are in one-to-one correspondence with Young diagrams $\lambda \vdash k$. Given λ, the construction of the corresponding irreducible representation V^λ can be done, for example, using Young symmetrizers as in [Fulton and Harris 1991, Lecture 4]. We write χ_λ for the character of S_k associated to V^λ and $d_\lambda := \chi_\lambda(\mathrm{id}) = \dim V^\lambda$. Given $\lambda \vdash k$, the element

$$\mathfrak{p}_\lambda := \frac{d_\lambda}{k!} \sum_{\sigma \in S_k} \chi_\lambda(\sigma)\sigma \in \mathbb{C}[S_k]$$

is a central idempotent in $\mathbb{C}[S_k]$.

If G is a compact group, (ρ, W) is an irreducible representation of G, and (π, V) is any finite-dimensional representation of G, the (ρ, W)-isotypic subspace of (π, V) is the invariant subspace of V spanned by all irreducible direct summands of (π, V) that are isomorphic to (ρ, W). When ρ and π can be inferred from W and V, we call this simply the *W-isotypic subspace of V*. If $H \le G$ is a subgroup and (ρ, W) is an irreducible representation of H, then the W-isotypic subspace of V for H is the W-isotypic subspace of the restriction of (π, V) to H.

If (π, V) is any finite-dimensional unitary representation of S_k, and $\lambda \vdash k$, then V is also a module for $\mathbb{C}[S_k]$ by linear extension of π and $\pi(\mathfrak{p}_\lambda)$ is the orthogonal projection onto the V^λ-isotypic subspace of V.

For any compact group G, we write $(\mathrm{triv}_G, \mathbb{C})$ for the trivial representation of G. The following lemma can be deduced for example by combining Young's rule [Fulton and Harris 1991, Corollary 4.39] with Frobenius reciprocity.

Lemma 2.1 *Let $k \in \mathbb{N}_0$ and $\lambda \vdash k$. The space of vectors in V^λ fixed by S_λ is one-dimensional.*

2.2 Representation theory of $\mathsf{U}(n)$ and $\mathsf{SU}(n)$

Every irreducible representation of $\mathsf{U}(n)$ restricts to an irreducible representation of $\mathsf{SU}(n)$, and all equivalence classes of irreducible representations of $\mathsf{SU}(n)$ arise in this way. The equivalence classes of irreducible representations of $\mathsf{U}(n)$ are parametrized by dominant weights, which can be thought of as nonincreasing sequences

$$\Lambda = (\Lambda_1, \dots, \Lambda_n) \in \mathbb{Z}^n,$$

also known as *signatures*. We write W^Λ for the irreducible representation of $\mathsf{U}(n)$ corresponding to the signature Λ. Two irreducible representations of $\mathsf{U}(n)$ restrict to the same one of $\mathsf{SU}(n)$ if and only if their signatures differ by a constant vector. Let $\mathbb{T}(n)$ denote the maximal torus of $\mathsf{U}(n)$ consisting of diagonal matrices. Any matrix of $\mathbb{T}(n)$ has the form $\operatorname{diag}(\exp(i\theta_1), \dots, \exp(i\theta_n))$, where all $\theta_j \in \mathbb{R}$. Associated to the signature Λ is the character ξ_Λ of $\mathbb{T}(n)$ given by

$$\xi_\Lambda\bigl(\operatorname{diag}(\exp(i\theta_1), \dots, \exp(i\theta_n))\bigr) := \exp\biggl(i\biggl(\sum_{j=1}^{n} \Lambda_j\theta_j\biggr)\biggr).$$

The highest weight theory says among other things that the ξ_Λ-isotypic subspace of W^Λ for $\mathbb{T}(n)$ is one-dimensional. Any vector in this subspace is called a *highest weight vector* of W^Λ.

Given $k, l \in \mathbb{N}_0$ and fixed Young diagrams $\mu \vdash k$ and $\nu \vdash l$, we define a family of representations of $\mathsf{U}(n)$ as follows. For $n \geq \ell(\mu) + \ell(\nu)$, define

$$\Lambda_{\mu,\nu}(n) := (\mu_1, \mu_2, \dots, \mu_{\ell(\mu)}, \underbrace{0, \dots, 0}_{n-\ell(\mu)-\ell(\nu)}, -\nu_{\ell(\nu)}, -\nu_{\ell(\nu)-1}, \dots, -\nu_1).$$

We let $(\rho_n^{\mu,\nu}, W_n^{\mu,\nu})$ denote the irreducible representation of $\mathsf{U}(n)$ corresponding to $\Lambda_{\mu,\nu}(n)$ when $n \geq \ell(\mu) + \ell(\nu)$. We let $D_{\mu,\nu}(n) := \dim W_n^{\mu,\nu}$ and $s_{\mu,\nu}(g) := \operatorname{Tr}(\rho_n^{\mu,\nu}(g))$ for $g \in \mathsf{U}(n)$. If $\mu \vdash k$ and $\nu \vdash l$, then, as $n \to \infty$,

$$D_{\mu,\nu}(n) \asymp n^{k+l} \tag{2-1}$$

by [Magee 2022, Corollary 2.3] (alternatively [Enomoto and Izumi 2016, Lemma 3.5]).

We now present a version of Schur–Weyl duality for mixed tensors due to Koike [1989]. The very definition of $\mathsf{U}(n)$ makes $\mathbb{C}^n$ into a unitary representation of $\mathsf{U}(n)$ for the standard Hermitian inner product. We let $\{e_1, \dots, e_n\}$ denote the standard basis of $\mathbb{C}^n$. If (ρ, W) is any finite-dimensional representation of $\mathsf{U}(n)$, we write $(\rho^\vee, W^\vee)$ for the dual representation, where $W^\vee$ is the space of complex linear functionals on W. The vector space $(\mathbb{C}^n)^\vee$ has a dual basis $\{\check{e}_1, \dots, \check{e}_n\}$ given by $\check{e}_j(v) := \langle v, e_j\rangle$. Throughout the paper we frequently use certain canonical isomorphisms, eg

$$((\mathbb{C}^n)^{\otimes p})^\vee \cong ((\mathbb{C}^n)^\vee)^{\otimes p}, \quad \operatorname{End}(W) \cong W \otimes W^\vee,$$

to change points of view on representations; if we use noncanonical isomorphisms, we point them out.

Let $\mathcal{T}_n^{k,l} := (\mathbb{C}^n)^{\otimes k} \otimes ((\mathbb{C}^n)^\vee)^{\otimes l}$, with the convention that $(\mathbb{C}^n)^{\otimes 0} := \mathbb{C}$. With the natural inner product induced by that on $\mathbb{C}^n$, this is a unitary representation of $\mathsf{U}(n)$ under the diagonal action and also a unitary representation of $S_k \times S_l$, where S_k acts by permuting the indices of $(\mathbb{C}^n)^{\otimes k}$ and S_l acts by permuting the indices of $((\mathbb{C}^n)^\vee)^{\otimes l}$. We write $\pi_n^{k,l} : \mathsf{U}(n) \to \mathrm{End}[\mathcal{T}_n^{k,l}]$ and $\rho_n^{k,l} : \mathbb{C}[S_k \times S_l] \to \mathrm{End}[\mathcal{T}_n^{k,l}]$ for these representations. The actions of $\mathsf{U}(n)$ and $S_k \times S_l$ on $\mathcal{T}_n^{k,l}$ commute. We use the notation, for $I = (i_1, \dots, i_k) \in [n]^k$ and $J = (j_1, \dots, j_l) \in [n]^l$,

$$e_I := e_{i_1} \otimes \cdots \otimes e_{i_k} \in (\mathbb{C}^n)^{\otimes k}, \quad \check{e}_J := \check{e}_{j_1} \otimes \cdots \otimes \check{e}_{j_l} \in ((\mathbb{C}^n)^\vee)^{\otimes l}, \quad e_I^J := e_I \otimes \check{e}_J \in \mathcal{T}_n^{k,l}.$$

We write $I \sqcup J$ for the concatenation $(i_1, \dots, i_k, j_1, \dots, j_l)$.

For $k, l \geq 1$, let $\dot{\mathcal{T}}_n^{k,l}$ denote the intersection of the kernels of the mixed contractions $c_{pq} : \mathcal{T}_n^{k,l} \to \mathcal{T}^{k-1,l-1}$ for $p \in [k]$ and $q \in [l]$ given by

$$\begin{aligned} c_{pq}(e_{i_1} \otimes \cdots \otimes e_{i_k} \otimes \check{e}_{j_1} \otimes \cdots \otimes \check{e}_{j_l}) \\ := \delta_{i_p j_q} e_{i_1} \otimes \cdots \otimes e_{i_{p-1}} \otimes e_{i_{p+1}} \otimes \cdots \otimes e_{i_k} \otimes \check{e}_{j_1} \otimes \cdots \otimes \check{e}_{j_{q-1}} \otimes \check{e}_{j_{q+1}} \otimes \cdots \otimes \check{e}_{j_l}, \end{aligned} \tag{2-2}$$

where $\delta_{i_p j_q}$ is the Kronecker delta. If $k = 1$ or $l = 1$, then the definition is extended in the natural way, interpreting an empty tensor of e_i or $\check{e}_i$ as 1. If either $k = 0$ or $l = 0$, then $\dot{\mathcal{T}}_n^{k,l} = \mathcal{T}_n^{k,l}$ by convention. The space $\dot{\mathcal{T}}_n^{k,l}$ is an invariant subspace under $\mathsf{U}(n) \times S_k \times S_l$ and hence a unitary subrepresentation of $\mathcal{T}_n^{k,l}$. On $\dot{\mathcal{T}}_n^{k,l}$ there is an analog of Schur–Weyl duality due to Koike.

Theorem 2.2 [Koike 1989, Theorem 1.1] *There is an isomorphism of unitary representations of* $\mathsf{U}(n) \times S_k \times S_l$

$$\dot{\mathcal{T}}_n^{k,l} \cong \bigoplus_{\substack{\mu \vdash k, \nu \vdash l \\ \ell(\mu)+\ell(\nu) \leq n}} W_n^{\mu,\nu} \otimes V^\mu \otimes V^\nu. \tag{2-3}$$

Next we explain how to construct $\mathsf{U}(n)$-subrepresentations of $\dot{\mathcal{T}}_n^{k,l}$ isomorphic to $W_n^{\mu,\nu}$. Suppose that $\xi \in \dot{\mathcal{T}}_n^{k,l}$ is a nonzero vector such that, under the isomorphism (2-3),

$$\xi \cong w \otimes v \tag{2-4}$$

for $w \in W_n^{\mu,\nu}$ and $v \in V^\mu \otimes V^\nu$. Then $\mathsf{U}(n) \cdot \xi$ linearly spans a $\mathsf{U}(n)$-subrepresentation of $\dot{\mathcal{T}}_n^{k,l}$ isomorphic to $W_n^{\mu,\nu}$. The following argument to construct such a vector ξ, given $\mu \vdash k$ and $\nu \vdash l$, appears implicitly in [Koike 1989] and is elaborated in [Benkart et al. 1994]. For $n \geq \ell(\mu) + \ell(\nu)$, let

$$\tilde{\theta}_{\mu,\nu}^n := e_1^{\otimes \mu_1} \otimes \cdots \otimes e_{\ell(\mu)}^{\otimes \mu_{\ell(\mu)}} \otimes (\check{e}_n)^{\otimes \nu_1} \otimes \cdots \otimes (\check{e}_{n-\ell(\nu)+1})^{\otimes \nu_{\ell(\nu)}}. \tag{2-5}$$

This vector is in the $\xi_{\mu,\nu}$-isotypic subspace of $\dot{\mathcal{T}}_n^{k,l}$ for the maximal torus $\mathbb{T}(n)$ of $\mathsf{U}(n)$, where $\xi_{\mu,\nu}$ is the character of $\mathbb{T}(n)$ corresponding to the highest weight in $W_n^{\mu,\nu}$.

Let $\mathfrak{p}_\mu \in \mathbb{C}[S_k]$ and $\mathfrak{p}_\nu \in \mathbb{C}[S_l]$ be the projections defined in Section 2.1. Let $\rho_n^k : S_k \to \mathrm{End}(\mathcal{T}_n^{k,l})$ denote the representation of S_k described above and $\hat{\rho}_n^l : S_l \to \mathrm{End}(\mathcal{T}_n^{k,l})$ that of S_l. Clearly these two representations commute. Now let

$$\theta^n_{\mu,\nu} := \rho^k_n(\mathfrak{p}_\mu)\hat{\rho}^l_n(\mathfrak{p}_\nu)\tilde{\theta}^n_{\mu,\nu} \in \dot{\mathcal{T}}^{k,l}_n. \tag{2-6}$$

Now this is in the same isotypic subspace for $\mathbb{T}(n)$ as before since $S_k \times S_l$ commutes with $\mathsf{U}(n)$. Moreover, it is in the subspace of $\dot{\mathcal{T}}^{k,l}_n$ corresponding to $W^{\mu,\nu}_n \otimes V^\mu \otimes V^\nu$ under the isomorphism (2-3). The intersection of the two subspaces of $\dot{\mathcal{T}}^{k,l}_n$ just discussed corresponds via (2-3) to $\mathbb{C}w \otimes V^\mu \otimes V^\nu$, where w is a highest weight vector in $W^{\mu,\nu}_n$, and hence $\theta^n_{\mu,\nu}$ takes the form of (2-4), as we desired.

Of course, we also want to know $\theta^n_{\mu,\nu} \neq 0$.

Lemma 2.3 *Suppose that $k, l \in \mathbb{N}_0$, $\mu \vdash k$, $\nu \vdash l$, and $\theta^n_{\mu,\nu}$ is as in (2-6) for $n \geq \ell(\mu)+\ell(\nu)$. We have*

$$\|\theta^n_{\mu,\nu}\|^2 = \frac{d_\mu d_\nu}{[S_k : S_\mu][S_l : S_\nu]}.$$

Proof Recall the definition of Young subgroups S_μ and S_ν from Section 2.1. Letting $\tilde{\theta} = \tilde{\theta}^n_{\mu,\nu}$ (as in (2-5)) and $\theta = \theta^n_{\mu,\nu}$, we have

$$\begin{aligned}\theta = \rho^k_n(\mathfrak{p}_\mu)\hat{\rho}^l_n(\mathfrak{p}_\nu)\tilde{\theta} &= \frac{d_\mu d_\nu}{k!\,l!} \sum_{\sigma=(\sigma_1,\sigma_2)\in S_k\times S_l} \chi_\mu(\sigma_1)\chi_\nu(\sigma_2)\rho^k_n(\sigma_1)\hat{\rho}^l_n(\sigma_2)\tilde{\theta} \\ &= \frac{d_\mu d_\nu}{k!\,l!} \sum_{\substack{[\sigma_1]\in S_k/S_\mu \\ [\sigma_2]\in S_l/S_\nu}} \Bigl(\sum_{\tau_1\in S_\mu}\chi_\mu(\sigma_1\tau_1)\Bigr)\Bigl(\sum_{\tau_2\in S_\nu}\chi_\nu(\sigma_2\tau_2)\Bigr)\rho^k_n(\sigma_1)\hat{\rho}^l_n(\sigma_2)\tilde{\theta}.\end{aligned}$$

The second equality used that $\tilde{\theta}$ is invariant under $S_\mu \times S_\nu$.

By Lemma 2.1, there is a one-dimensional subspace of invariant vectors for S_μ in V^μ. If $v_\mu \in V^\mu$ is a unit vector in this space, then

$$\sum_{\tau_1\in S_\mu}\chi_\mu(\sigma_1\tau_1) = |S_\mu|\langle\sigma_1 v_\mu, v_\mu\rangle. \tag{2-7}$$

Since the vectors $\rho^k_n(\sigma_1)\hat{\rho}^l_n(\sigma_2)\tilde{\theta}$ for $[\sigma_1] \in S_k/S_\mu$ and $[\sigma_2] \in S_l/S_\nu$ are orthogonal unit vectors, this gives

$$\begin{aligned}\|\theta\|^2 &= \Bigl(\frac{d_\mu d_\nu}{k!\,l!}\Bigr)^2 \sum_{\substack{[\sigma_1]\in S_k/S_\mu \\ [\sigma_2]\in S_l/S_\nu}} \Bigl(\sum_{\tau_1\in S_\mu}\chi_\mu(\sigma_1\tau_1)\Bigr)^2\Bigl(\sum_{\tau_2\in S_\nu}\chi_\nu(\sigma_2\tau_2)\Bigr)^2 \\ &= \Bigl(\frac{d_\mu d_\nu}{k!\,l!}\Bigr)^2 |S_\mu|^2|S_\nu|^2 \sum_{\substack{[\sigma_1]\in S_k/S_\mu \\ [\sigma_2]\in S_l/S_\nu}} |\langle\sigma_1 v_\mu, v_\mu\rangle|^2|\langle\sigma_2 v_\nu, v_\nu\rangle|^2 \qquad \text{(by (2-7))} \\ &= \Bigl(\frac{d_\mu d_\nu}{k!\,l!}\Bigr)^2 |S_\mu||S_\nu| \sum_{\substack{\sigma_1\in S_k \\ \sigma_2\in S_l}} |\langle\sigma_1 v_\mu, v_\mu\rangle|^2|\langle\sigma_2 v_\nu, v_\nu\rangle|^2 = \frac{d_\mu d_\nu}{[S_k : S_\mu][S_l : S_\nu]}.\end{aligned}$$

The last inequality used the orthogonality relations for matrix coefficients. □

Recall that we write $\pi_n^{k,l}:\mathsf{U}(n)\to\mathrm{End}(\mathcal{T}_n^{k,l})$ for the diagonal representation of $\mathsf{U}(n)$ on $\mathcal{T}_n^{k,l}$. Lemma 2.3 implies that $\theta^n_{\mu,\nu}$ is a nonzero vector. By the remarks following (2-6), it is of the pure tensor form $w\otimes v$ under the Schur–Weyl isomorphism (2-3), with $w\in W_n^{\mu,\nu}$, and hence we obtain the following corollary:

Corollary 2.4 *Suppose $n\ge \ell(\mu)+\ell(\nu)$. The subspace*

$$W_n(\theta^n_{\mu,\nu}) := \mathrm{span}\{\pi_n^{k,l}(u)\theta^n_{\mu,\nu} : u\in\mathsf{U}(n)\}\subset \dot{\mathcal{T}}_n^{k,l}$$

is, under $\pi_n^{k,l}$, a $\mathsf{U}(n)$-subrepresentation of $\dot{\mathcal{T}}_n^{k,l}$ isomorphic to $W_n^{\mu,\nu}$.

2.3 The Weingarten calculus

The Weingarten calculus is a method based on Schur–Weyl duality that allows one to calculate integrals of products of matrix coefficients in the defining representation of $\mathsf{U}(n)$ in terms of sums over permutations. It was discovered initially by Weingarten [1978], and developed further in [Xu 1997; Collins 2003; Collins and Śniady 2006].

We present two formulations of the Weingarten calculus. Given $k\in\mathbb{N}$ and $n\in\mathbb{N}$, the *Weingarten function* with parameters n and k is the element[2] of $\mathbb{C}[S_k]$ [Collins and Śniady 2006, equation (9)]

$$\mathrm{Wg}_{n,k} := \frac{1}{(k!)^2}\sum_{\substack{\lambda\vdash k\\ \ell(\lambda)\le n}}\frac{d_\lambda^2}{D_\lambda(n)}\sum_{\sigma\in S_k}\chi_\lambda(\sigma)\sigma. \tag{2-8}$$

We write $\mathrm{Wg}_{n,k}(\sigma)$ for the coefficient of σ in (2-8). The following theorem was proved by Collins and Śniady [2006, Corollary 2.4]:

Theorem 2.5 *For $k\in\mathbb{N}$ and for $i_1,i_1',j_k,j_k',\dots,i_k,i_k',j_k,j_k'\in[n]$,*

$$\begin{aligned}\int_{u\in\mathsf{U}(n)} u_{i_1j_1}\cdots u_{i_kj_k}\bar u_{i_1'j_1'}\cdots\bar u_{i_k'j_k'}\,d\mu(u)\qquad\qquad\\ =\sum_{\sigma,\tau\in S_k}\delta_{i_1i'_{\sigma(1)}}\cdots\delta_{i_ki'_{\sigma(k)}}\delta_{j_1j'_{\tau(1)}}\cdots\delta_{j_kj'_{\tau(k)}}\mathrm{Wg}_{n,k}(\tau\sigma^{-1}),\end{aligned}\tag{2-9}$$

where δ_{pq} is the Kronecker delta function.

It is sometimes more flexible to reformulate Theorem 2.5 in terms of projections. Here $u\in\mathsf{U}(n)$ acts on $A\in\mathrm{End}((\mathbb{C}^n)^{\otimes k})$ by $A\mapsto \pi_n^k(u)A\pi_n^k(u^{-1})$, where $\pi_n^k:\mathsf{U}(n)\to\mathrm{End}((\mathbb{C}^n)^{\otimes k})$ is the diagonal action. Write $P_{n,k}$ for the orthogonal projection in $\mathrm{End}((\mathbb{C}^n)^{\otimes k})$ onto the $\mathsf{U}(n)$-invariant vectors. The following proposition is due to [Collins and Śniady 2006, Proposition 2.3]:

[2] Although not relevant here, classically the Weingarten function arises as the multiplicative inverse of $\sum_{\sigma\in S_k} n^{\#\mathrm{cycles}(\sigma)}\sigma$ in $\mathbb{C}[S_k]$ whenever $n\ge k$.

Proposition 2.6 (Collins and Śniady) *Let $n, k \in \mathbb{N}$. Suppose $A \in \mathrm{End}((\mathbb{C}^n)^{\otimes k})$. Then*

$$P_{n,k}[A] = \rho_n^k(\Phi[A] \cdot \mathrm{Wg}_{n,k}),$$

where

$$\Phi[A] := \sum_{\sigma \in S_k} \mathrm{Tr}(A\rho_n^k(\sigma^{-1}))\sigma.$$

Later we will need the following bound for the Weingarten function due to [Collins and Śniady 2006, Proposition 2.6]. For a permutation σ, let $|\sigma|$ denote the minimum number of transpositions that σ can be written as a product of.

Proposition 2.7 *For any fixed $\sigma \in S_k$, $\mathrm{Wg}_{n,k}(\sigma) \ll_k n^{-k-|\sigma|}$ as $n \to \infty$.*

2.4 Free groups and surface groups

Let $\boldsymbol{F}_{2g} := \langle a_1, b_1, \dots, a_g, b_g \rangle$ be the free group on $2g$ generators $a_1, b_1, \dots, a_g, b_g$ and $R_g := [a_1, b_1] \cdots [a_g, b_g] \in \boldsymbol{F}_{2g}$. There is a quotient map $\boldsymbol{F}_{2g} \to \Gamma_g$ given by reduction modulo R_g. We say that $w \in \boldsymbol{F}_{2g}$ *represents the conjugacy class of* $\gamma \in \Gamma_g$ if the projection of w to Γ_g is in the conjugacy class of γ in Γ_g.

Given $w \in \boldsymbol{F}_{2g}$, we view w as a combinatorial word in $a_1, a_1^{-1}, b_1, b_1^{-1}, \dots, a_g, a_g^{-1}, b_g, b_g^{-1}$ by writing it in reduced (shortest) form; ie a_1 does not follow a_1^{-1} etc. We say that w is *cyclically reduced* if the first letter of its reduced word is not the inverse of the last letter. The length $|w|$ of $w \in \boldsymbol{F}_{2g}$ is the length of its reduced form word. We say $w \in \boldsymbol{F}_{2g}$ is a *shortest element* representing the conjugacy class of $\gamma \in \Gamma_g$ if it has minimal length among all elements representing the conjugacy class of γ. If w is a shortest element representing some conjugacy class in Γ_g, then w is cyclically reduced.

For any group H, the commutator subgroup $[H, H] \le H$ is the subgroup generated by all elements of the form $[h_1, h_2] := h_1 h_2 h_1^{-1} h_2^{-1}$ with $h_1, h_2 \in H$. If $\gamma \in [\Gamma_g, \Gamma_g]$ and w represents the conjugacy class of γ, then $w \in [\boldsymbol{F}_{2g}, \boldsymbol{F}_{2g}]$ (see [Magee 2022, Section 2.6]).

2.5 Witten zeta functions

Witten zeta functions appeared first in [Witten 1991] and were named by Zagier [1994]. The *Witten zeta function of* $\mathsf{SU}(n)$ is defined, for s in a half-plane of convergence, by

$$\zeta(s; n) := \sum_{(\rho, W) \in \widehat{\mathsf{SU}(n)}} \frac{1}{(\dim W)^s}, \tag{2-10}$$

where $\widehat{\mathsf{SU}(n)}$ denotes the equivalence classes of irreducible representations of $\mathsf{SU}(n)$. Indeed, the series (2-10) converges for $\mathrm{Re}(s) > 2/n$ by [Larsen and Lubotzky 2008, Theorem 5.1] (see also [Häsä and Stasinski 2019, Section 2]). Also relevant to this work is a result of Guralnick, Larsen and Manack [Guralnick et al. 2012, Theorem 2 and equation (7)], which states, for fixed $s > 0$,

$$\lim_{n \to \infty} \zeta(s; n) = 1. \tag{2-11}$$

2.6 Results of the prequel paper

By [Magee 2022, Proposition 1.5], if $\gamma \notin [\Gamma_g, \Gamma_g]$, then $\mathbb{E}_{g,n}[\mathrm{Tr}_\gamma] = 0$ for $n \geq n_0(\gamma)$. This proves Theorem 1.2 in this case. *Hence, in the rest of the paper we need only consider* $\gamma \in [\Gamma_g, \Gamma_g]$ *and hence* $w \in [\boldsymbol{F}_{2g}, \boldsymbol{F}_{2g}]$ *if* $w \in \boldsymbol{F}_{2g}$ *represents the conjugacy class of* γ.

For each $w \in \boldsymbol{F}_{2g}$, we have a *word map* $w : \mathsf{U}(n)^{2g} \to \mathsf{U}(n)$ obtained by substituting matrices for the generators of $\boldsymbol{F}_{2g}$. For example, if $u_1, v_1, \dots, u_g, v_g \in \mathsf{U}(n)$ then $R_g(u_1, v_1, \dots, u_g, v_g) = [u_1, v_1] \cdots [u_g, v_g]$. We begin with the following result from [Magee 2022, Corollary 1.8]:

Proposition 2.8 *Suppose that* $g \geq 2$, $\gamma \in \Gamma_g$, *and* $w \in \boldsymbol{F}_{2g}$ *represents the conjugacy class of* γ. *For any* $B \in \mathbb{N}$, *we have, as* $n \to \infty$,

$$\mathbb{E}_{g,n}[\mathrm{Tr}_\gamma] = \zeta(2g-2; n)^{-1} \sum_{\substack{\mu,\nu \text{ Young diagrams} \\ \ell(\mu),\ell(\nu) \leq B \\ \mu_1,\nu_1 \leq B^2}} D_{\mu,\nu}(n) \mathcal{I}_n(w, \mu, \nu) + O_{B,w,g}(n^{|w|} n^{-2\log B}), \tag{2-12}$$

where

$$\mathcal{I}_n(w, \mu, \nu) := \int_{\mathsf{SU}(n)^{2g}} \mathrm{Tr}(w(x)) \overline{s_{\mu,\nu}(R_g(x))}\, d\mu(x). \tag{2-13}$$

Notice that, for $n \geq 2B$, the right-hand side of (2-12) makes sense, ie $D_{\mu,\nu}, s_{\mu,\nu}$ are well defined. We also have the following proposition, which follows from [Magee 2022, Proposition 3.1] together with $\bar{s}_{\mu,\nu} = s_{\nu,\mu}$:

Proposition 2.9 *Let* $w \in [\boldsymbol{F}_{2g}, \boldsymbol{F}_{2g}]$. *Then, for any fixed* μ, ν *and* $n \geq \ell(\mu) + \ell(\nu)$,

$$\mathcal{I}_n(w, \mu, \nu) = \mathcal{J}_n(w, \nu, \mu) := \int_{\mathsf{U}(n)^{2g}} \mathrm{Tr}(w(x)) s_{\nu,\mu}(R_g(x))\, d\mu(x).$$

This is convenient as it will allow us to use the Weingarten calculus directly as it is presented in Section 2.3 for $\mathsf{U}(n)$ rather than $\mathsf{SU}(n)$. By using Proposition 2.9, taking a representative $w \in \boldsymbol{F}_{2g}$ of the conjugacy class of γ and taking B such that $|w| - 2\log B \leq -1$ in Proposition 2.8, we obtain the following result, from which we begin the new arguments of this paper:

Corollary 2.10 *Let* $\gamma \in [\Gamma_g, \Gamma_g]$ *and* $w \in [\boldsymbol{F}_{2g}, \boldsymbol{F}_{2g}]$ *be a representative of the conjugacy class of* $\gamma \in \Gamma$. *Then there exists a* ***finite*** *set* $\widetilde{\Omega}$ *of pairs* (μ, ν) *of Young diagrams such that*

$$\mathbb{E}_{g,n}[\mathrm{Tr}_\gamma] = \zeta(2g-2; n)^{-1} \sum_{(\mu,\nu) \in \widetilde{\Omega}} D_{\mu,\nu}(n) \mathcal{J}_n(w, \nu, \mu) + O_{w,g}\Big(\frac{1}{n}\Big).$$

As we know $\lim_{n\to\infty} \zeta(2g-2, n) = 1$ by (2-11), we have now reduced the proof of Theorem 1.2 to establishing suitable bounds for the integrals $\mathcal{J}_n(w, \mu, \nu)$, where we can view μ and ν as *fixed* Young diagrams since $\widetilde{\Omega}$ is finite.

3 Combinatorial integration

3.1 Setup and motivation

The main result of the rest of the paper is the following:

Theorem 3.1 *Let $\gamma \in \Gamma_g$ with $\gamma \neq \mathrm{id}$. Let $w \in \boldsymbol{F}_{2g}$ be a shortest element representing the conjugacy class of γ. For each $k, l \in \mathbb{N}_0$, there is a constant $C(w,k,l) > 0$ such that, for any $\mu \vdash k$, $\nu \vdash l$*

$$|D_{\mu,\nu}(n)\mathcal{J}_n(w,\mu,\nu)| \leq C(w,k,l)$$

for all $n \in \mathbb{N}$.

Accordingly, since we know the large n behavior of $D_{\mu,\nu}(n)$ from (2-1), in this section we wish to estimate

$$\mathcal{J}_n(w,\mu,\nu) = \int_{\mathsf{U}(n)^{2g}} \mathrm{Tr}(w(x)) s_{\mu,\nu}(R_g(x))\, d\mu(x)$$

for fixed $\mu \vdash k$, $\nu \vdash l$.

What doesn't work We begin by discussing why the most straightforward approach to this problem leads to serious complications. It is possible to approach the problem by writing $s_{\mu,\nu}(h)$ as a fixed finite linear combination of functions

$$p_{\mu'}(h) p_{\nu'}(h^{-1}),$$

where $p_{\mu'}(h)$ (resp. $p_{\nu'}(h^{-1})$) is a power sum symmetric polynomial of the eigenvalues of h (resp. h^{-1} or $\bar{h}$). See for example [Magee 2022, Section 3.3] for one way to do this. The coefficients of this expansion are fixed, but not transparent, since they involve Littlewood–Richardson coefficients. In any case, this approach leads to writing $\mathcal{J}_n(w,\mu,\nu)$ as a finite linear combination of integrals of the form

$$\int_{\mathsf{U}(n)^{2g}} \mathrm{Tr}(w(x))\mathrm{Tr}(R_g(x)^{k_1}) \cdots \mathrm{Tr}(R_g(x)^{k_p})\mathrm{Tr}(R_g(x)^{-l_1}) \cdots \mathrm{Tr}(R_g(x)^{-l_q})\, d\mu(x), \tag{3-1}$$

where $\sum k_j = |\mu|$ and $\sum l_j = |\nu|$.

Magee and Puder [2019] give a full asymptotic expansion for (3-1) as $n \to \infty$. However, these estimates are not sufficient for the current paper and, to motivate the rest of this section, we explain briefly the issues involved. However, this discussion is not needed to understand the arguments that we will make to prove Theorem 3.1.

The main result of [Magee and Puder 2019] gives a full "genus" expansion of (3-1) in terms of surfaces and maps on surfaces dictated by $w \in \boldsymbol{F}_{2g}$. Roughly speaking, every term in this expansion comes from a homotopy class of map f from an orientable surface Σ_f to $\bigvee_{i=1}^{2g} S^1$; to contribute to (3-1) the surface Σ_f has one boundary component that maps to w at the level of the fundamental groups, p boundary components that map respectively to $R_g^{k_1}, \ldots, R_g^{k_p}$ at the level of fundamental groups, and q

boundary components that map respectively to $R_g^{-l_1}, \dots, R_g^{-l_q}$ at the level of fundamental groups. The contribution of the pair (f, Σ_f) to (3-1) is of the form $c(f, \Sigma_f) n^{\chi(\Sigma_f)}$; the coefficient $c(f, \Sigma_f)$ is an Euler characteristic of a symmetry group of (f, Σ_f) and is not easy to calculate in general. However, one could still hope to get decay of (3-1) by controlling the possible $\chi(\Sigma_f)$ that could appear.

There are two issues with this. The first one is that, if w is not the shortest element representing the conjugacy class of γ, then we get bounds that are not helpful. For a very simple example, let $w = R_g^l$ and $\gamma = \mathrm{id}_{\Gamma_g}$, and consider the potential contribution from $p = 0$, $q = 1$ and $l_1 = l$. Then, for any ν with $|\nu| = l$, there is contribution to $\mathcal{J}_n(w, \varnothing, \nu)$ that is a multiple of

$$\int_{\mathsf{U}(n)^{2g}} \mathrm{Tr}(R_g(x)^l) \mathrm{Tr}(R_g(x)^{-l}) \, d\mu(x).$$

Here, in the theory of [Magee and Puder 2019], there is a (Σ_f, f) that is an annulus, one boundary component corresponding to $w = R_g^l$ and one corresponding to R_g^{-l}, so we can only bound the corresponding contribution to $D_{\varnothing,\nu}(n) \mathcal{J}_n(w, \varnothing, \nu)$ by using [Magee and Puder 2019] on the order of $D_{\varnothing,\nu}(n) \asymp n^l$. On the other hand, any approach that works to establish Theorem 3.1 (for $\gamma \neq \mathrm{id}$) should extend to show that, when $\gamma = \mathrm{id}$, $D_{\varnothing,\nu}(n) \mathcal{J}_n(w, \varnothing, \nu) \ll n$ as $\mathbb{E}_{g,n}[\mathrm{Tr}_{\mathrm{id}}] = n$.

Indeed, this phenomenon extends to words of the form $w_0 R_g^l$ and more generally to words that are not shortest representatives of some conjugacy class in Γ_g. It means that, even if we use something similar in spirit to [Magee and Puder 2019], to prove Theorem 3.1 we must incorporate the theory of shortest representative words. This indeed takes place in Sections 4.3–4.5; the topological result proved there hinges on this theory.

The second issue is a little more subtle and only appears for "mixed" representations, ie both $\mu, \nu \neq \varnothing$. In this case, suppose w is a shortest element representing some conjugacy class in Γ_g and $w \in [\boldsymbol{F}_{2g}, \boldsymbol{F}_{2g}]$. This means that there is a pair (f_0, Σ_{f_0}) where Σ_{f_0} has one boundary component that maps to w at the level of the fundamental groups. Let us take $\mu, \nu = (k), (k)$, ie each Young diagram has one row of k boxes. This means we get a potential contribution to $D_{\mu,\nu}(n) \mathcal{J}_n(w, \mu, \nu)$ that is a constant multiple of

$$D_{(k),(k)}(n) \int_{\mathsf{U}(n)^{2g}} \mathrm{Tr}(w(x)) \mathrm{Tr}(R_g(x)^k) \mathrm{Tr}(R_g(x)^{-k}) \, d\mu(x). \tag{3-2}$$

Now, for every $k \in \mathbb{N}$, there is (f, Σ_f) contributing to (3-2) with one component that is (f_0, Σ_{f_0}) and the other an annulus with boundary components corresponding to R_g^k and R_g^{-k}. Since the annulus has Euler characteristic 0, and $D_{(k),(k)} \asymp n^{2k}$, the order of this contribution to $D_{(k),(k)}(n) \mathcal{J}_n(w, (k), (k))$ is potentially $\gg n^{2k} n^{\chi(\Sigma_{f_0})}$. For large enough k, the exponent here is arbitrarily large, which is clearly catastrophic. In reality, this contribution must cancel with some other contribution, but we do not know how to see these cancellations.

This ends the discussion of the difficulties of the most straightforward approach to the problems of this paper.

What does work To bypass the previous issues we produce a refined version of the Weingarten calculus that leads to a restricted set of surfaces, for instance not including the ones causing the problem above as well as all generalizations of this issue.

The basic approach is the following. Instead of trying to deal with a complicated formula for $s_{\mu,\nu}(R_2(x))$ (as above), we instead use the copy $W_n(\theta^n_{\mu,\nu})$ of $W_n^{\mu,\nu}$ in $\dot{\mathcal{T}}_n^{k,l}$ that we found in Corollary 2.4. In Section 3.3, we compute the orthogonal projection $\mathfrak{q}_\theta$ from $\mathcal{T}_n^{k,l}$ (note: not $\dot{\mathcal{T}}_n^{k,l}$) onto $W_n(\theta^n_{\mu,\nu})$ (Proposition 3.2). In the formula we obtain, we give bounds on the coefficients appearing therein (Lemma 3.3). In addition, we remember that $q_\theta \in \mathrm{End}(\dot{\mathcal{T}}_n^{k,l})$; this fact is not obvious from our formula but turns out to be vital going forward.

The calculation of $\mathfrak{q}_\theta$ is extra to, but in the same spirit as, the vanilla Weingarten calculus, which is why we claim to have refined the Weingarten calculus here.

In the expression for $\mathcal{I}_n(w,\mu,\nu)$, we now write

$$s_{\mu,\nu}(R_2(x)) = \mathrm{Tr}_{\mathcal{T}_n^{k,l}}(A\mathfrak{q}_\theta B\mathfrak{q}_\theta A^{-1}\mathfrak{q}_\theta B^{-1}\mathfrak{q}_\theta C\mathfrak{q}_\theta D\mathfrak{q}_\theta C^{-1}\mathfrak{q}_\theta D^{-1}\mathfrak{q}_\theta),$$

where A, B, C and D are the images of the generators of Γ_2 under x. Then the entire integral of $\mathrm{Tr}(w(x))s_{\mu,\nu}(R_2(x))$ is done using the usual Weingarten calculus. The fact that $q_\theta \in \mathrm{End}(\dot{\mathcal{T}}_n^{k,l})$ intervenes at a critical point to show that certain contributions from the classical Weingarten calculus cancel and lead to restrictions on the nonzero contributions. Precisely, the restriction we obtain is summarized in the *forbidden matching* property below (Section 3.4) and property (P4) (Section 4.3).

3.2 Proof of Theorem 3.1 when $k = l = 0$

Here we give a proof of Theorem 3.1 when $k = l = 0$. This will allow us to bypass the slightly confusing issue of using the Weingarten function $\mathrm{Wg}_{n,k+l}$ when $k + l = 0$ in Section 3.3.

If $k = l = 0$, then the only possible $\mu \vdash k$ and $\nu \vdash l$ are empty Young diagrams $\mu = \nu = \varnothing$, and $W_n^{\varnothing,\varnothing}$ is the trivial representation of $\mathsf{U}(n)$, so $D_{\varnothing,\varnothing}(n) = 1$ for all $n \geq 1$ and $s_{\varnothing,\varnothing}(h) = 1$ for all $h \in \mathsf{U}(n)$. We then have

$$D_{\varnothing,\varnothing}(n)\mathcal{I}_n(w,\varnothing,\varnothing) = \mathcal{I}_n(w,\varnothing,\varnothing) = \int_{\mathsf{U}(n)^{2g}} \mathrm{Tr}(w(x))\,d\mu(x). \tag{3-3}$$

If $w \in \boldsymbol{F}_{2g}$ is a cyclically shortest word representing the conjugacy class of $\gamma \in \Gamma_g$ with $\gamma \neq \mathrm{id}$, then $w \neq \mathrm{id}$. It then follows from (1-1) that $D_{\varnothing,\varnothing}(n)\mathcal{I}_n(w,\varnothing,\varnothing) = o_w(n)$ as $n \to \infty$, but, in fact, (3-3) is given by a rational function of n for $n \geq n_0(w)$ by a straightforward application of the Weingarten calculus [Magee and Puder 2019]. This implies $D_{\varnothing,\varnothing}(n)\mathcal{I}_n(w,\varnothing,\varnothing) = O_w(1)$ as $n \to \infty$, as required.

This proves Theorem 3.1 when $k = l = 0$. Hence, in the rest of Section 3, we can assume $k + l > 0$.

3.3 A projection formula

Here we develop an integral calculus that is more powerful than the usual Weingarten calculus and allows us to directly tackle $\mathcal{J}_n(w,\mu,\nu)$ without writing it in terms of integrals as in (3-1). The key point is that our method leads to the *forbidden matchings* property of Section 3.4 and property (P4) of Section 4.3.

We now view k, l, $\mu \vdash k$ and $\nu \vdash l$ as fixed, assume $k+l>0$ and $n \geq \ell(\mu)+\ell(\nu)$, and write $\theta = \theta^n_{\mu,\nu}$ as in (2-6), suppressing the dependence on n. Let $W_n(\theta)$ be defined as in Corollary 2.4. Thus $W_n(\theta)$ is an irreducible summand of $\dot{\mathcal{T}}^{k,l}_n$ isomorphic to $W^{\mu,\nu}_n$ for the group $\mathsf{U}(n)$.

In the remainder of the paper we drop the dependence of our notation on n whenever it adds clarity.

Our first task is to compute the orthogonal projection $\mathfrak{q}_\theta$ onto $W(\theta)$. Let P_θ denote the orthogonal projection in $\mathcal{T}^{k,l}_n$ onto θ. We also view P_θ as an element of $\mathrm{End}(\dot{\mathcal{T}}^{k,l}_n)$ by restriction.

Under the canonical isomorphism $\mathrm{End}(\dot{\mathcal{T}}^{k,l}_n) \cong \dot{\mathcal{T}}^{k,l}_n \otimes (\dot{\mathcal{T}}^{k,l}_n)^\vee$, we have $P_\theta \cong (\theta \otimes \theta^\vee)/\|\theta\|^2$, and also, from (2-6),

$$P_\theta = \frac{1}{\|\theta\|^2}\rho^k(\mathfrak{p}_\mu)\hat{\rho}^l(\mathfrak{p}_\nu)[\tilde{\theta}_{\mu,\nu} \otimes \tilde{\theta}^\vee_{\mu,\nu}]\rho^k(\mathfrak{p}_\mu)\hat{\rho}^l(\mathfrak{p}_\nu); \tag{3-4}$$

here the inner square bracket is interpreted as an element of $\mathrm{End}(\dot{\mathcal{T}}^{k,l}_n)$. By Schur's lemma, we have

$$\mathfrak{q}_\theta = D_{\mu,\nu}(n)\int_{h\in\mathsf{U}(n)} \pi(h)P_\theta\pi(h^{-1})\,d\mu(h) \tag{3-5}$$

since the right-hand side is an element of $\mathrm{End}(W(\theta)) \subset \mathrm{End}(\mathcal{T}^{k,l}_n)$ that commutes with $\pi^{k,l}(\mathsf{U}(n))$, so it is a multiple of $\mathfrak{q}_\theta$, and it has the correct trace.

On the other hand, we can view $\mathcal{T}^{k,l}_n \otimes (\dot{\mathcal{T}}^{k,l}_n)^\vee \cong \mathcal{T}^{k+l,k+l}_n$ by the canonical isomorphism

$$\mathcal{T}^{k,l}_n \otimes (\dot{\mathcal{T}}^{k,l}_n)^\vee \cong (\mathbb{C}^n)^{\otimes k} \otimes ((\mathbb{C}^n)^{\otimes l})^\vee \otimes ((\mathbb{C}^n)^{\otimes k})^\vee \otimes (\mathbb{C}^n)^{\otimes l}$$

followed by the fixed isomorphism

$$\varphi\colon e^J_I \otimes \check{e}^{J'}_{I'} \mapsto e_{I\sqcup J'} \otimes \check{e}_{I'\sqcup J}. \tag{3-6}$$

Finally, there is a canonical isomorphism $\mathcal{T}^{k+l,k+l}_n \cong \mathrm{End}((\mathbb{C}^n)^{\otimes k+l})$. So, combining these, we fix isomorphisms

$$\mathrm{End}(\mathcal{T}^{k,l}_n) \cong \dot{\mathcal{T}}^{k,l}_n \otimes (\dot{\mathcal{T}}^{k,l}_n)^\vee \xrightarrow[\varphi]{\sim} \mathcal{T}^{k+l,k+l}_n \cong \mathrm{End}((\mathbb{C}^n)^{\otimes k+l}). \tag{3-7}$$

We view the outer two isomorphisms as fixed identifications. These isomorphisms are of unitary representations of $\mathsf{U}(n)$ when everything is given its natural inner product. Moreover, for $\sigma = (\sigma_1,\sigma_2) \in S_k \times S_l$ and $\tau = (\tau_1,\tau_2) \in S_k \times S_l$, we have, for $A \in \mathrm{End}(\mathcal{T}^{k,l}_n)$,

$$\varphi[\rho^k(\sigma_1)\hat{\rho}^l(\sigma_2)A\rho^k(\tau_1)\hat{\rho}^l(\tau_2)] = \rho^{k+l}(\sigma_1,\tau_2^{-1})\varphi[A]\rho^{k+l}(\tau_1,\sigma_2^{-1}), \tag{3-8}$$

recalling that $\rho^{k+l}\colon \mathbb{C}[S_{k+l}] \to \mathrm{End}((\mathbb{C}^n)^{\otimes k+l})$ is the representation by permuting coordinates.

We now return to the calculation of $\mathfrak{q}_\theta$ in (3-5). We have

$$\mathfrak{q}_\theta = D_{\mu,\nu}(n)\varphi^{-1}[P_{n,k+l}[\varphi(P_\theta)]], \tag{3-9}$$

where $P_{n,k+l}$ is the projection onto the $\mathsf{U}(n)$-invariant vectors (by conjugation) in $\mathrm{End}((\mathbb{C}^n)^{\otimes k+l})$. This can now be done using the classical Weingarten calculus. By Proposition 2.6, we have

$$P_{n,k+l}[\varphi(P_\theta)] = \rho^{k+l}\big(\Phi[\varphi(P_\theta)]\cdot \mathrm{Wg}_{n,k+l}\big), \tag{3-10}$$

where

$$\Phi[\varphi(P_\theta)] = \sum_{\sigma\in S_{k+l}} \mathrm{Tr}(\varphi(P_\theta)\rho^{k+l}(\sigma^{-1}))\sigma.$$

By (3-8) and (3-4), and since eg $\chi_\mu(g) = \chi_\mu(g^{-1})$, we obtain

$$\begin{aligned}\varphi(P_\theta) &= \frac{1}{\|\theta\|^2}\varphi(\rho^k(\mathfrak{p}_\mu)\hat{\rho}^l(\mathfrak{p}_\nu)[\tilde{\theta}_{\mu,\nu}\otimes\tilde{\theta}^\vee_{\mu,\nu}]\rho^k(\mathfrak{p}_\mu)\hat{\rho}^l(\mathfrak{p}_\nu))\\ &= \frac{1}{\|\theta\|^2}\rho^{k+l}(\mathfrak{p}_{\mu\otimes\nu})\varphi(\tilde{\theta}_{\mu,\nu}\otimes\tilde{\theta}^\vee_{\mu,\nu})\rho^{k+l}(\mathfrak{p}_{\mu\otimes\nu}),\end{aligned}$$

where

$$\mathfrak{p}_{\mu\otimes\nu} := \frac{d_\mu d_\nu}{k!\,l!}\sum_{\sigma=(\sigma_1,\sigma_2)\in S_k\times S_l}\chi_\mu(\sigma_1)\chi_\nu(\sigma_2)\sigma \in \mathbb{C}[S_{k+l}].$$

Now, using that Φ is a $\mathbb{C}[S_{k+l}]$-bimodule morphism [Collins and Śniady 2006, Proposition 2.3 (1)], we obtain

$$\begin{aligned}\Phi[\varphi(P_\theta)] &= \frac{1}{\|\theta\|^2}\mathfrak{p}_{\mu\otimes\nu}\Phi[\varphi(\tilde{\theta}_{\mu,\nu}\otimes\tilde{\theta}^\vee_{\mu,\nu})]\mathfrak{p}_{\mu\otimes\nu}\\ &= \frac{1}{\|\theta\|^2}\mathfrak{p}_{\mu\otimes\nu}\bigg(\sum_{\sigma\in S_{k+l}}\mathrm{Tr}(\varphi(\tilde{\theta}_{\mu,\nu}\otimes\tilde{\theta}^\vee_{\mu,\nu})\rho^{k+l}(\sigma^{-1}))\sigma\bigg)\mathfrak{p}_{\mu\otimes\nu}.\end{aligned}$$

Now, $\mathrm{Tr}(\varphi(\tilde{\theta}_{\mu,\nu}\otimes\tilde{\theta}^\vee_{\mu,\nu})\rho^{k+l}(\sigma^{-1}))$ is equal to 1 if and only if σ is in $S_\mu\times S_\nu\le S_k\times S_l$, and is 0 otherwise. So we obtain

$$\Phi[\varphi(P_\theta)] = \frac{1}{\|\theta\|^2}\mathfrak{p}_{\mu\otimes\nu}\bigg(\sum_{\sigma\in S_\mu\times S_\nu}\sigma\bigg)\mathfrak{p}_{\mu\otimes\nu};$$

hence, from (3-10),

$$P_{n,k+l}[\varphi(P_\theta)] = \rho^{k+l}(z_\theta),$$

where

$$z_\theta := \sum_{\tau\in S_{k+l}} z_\theta(\tau)\tau := \frac{1}{\|\theta\|^2}\mathfrak{p}_{\mu\otimes\nu}\bigg(\sum_{\sigma\in S_\mu\times S_\nu}\sigma\bigg)\mathfrak{p}_{\mu\otimes\nu}\mathrm{Wg}_{n,k+l}\in\mathbb{C}[S_{k+l}]. \tag{3-11}$$

Therefore we obtain the following proposition:

Proposition 3.2
$$\mathfrak{q}_\theta = D_{\mu,\nu}(n)\varphi^{-1}[\rho^{k+l}(z_\theta)].$$

We can use the bound for the coefficients of $\mathrm{Wg}_{n,k+l}$ from Proposition 2.7 to infer a bound on the coefficients $z_\theta(\tau)$. For $\sigma \in S_{k+l}$, let $\|\sigma\|_{k,l}$ denote the minimum m for which

$$\sigma = \sigma_0 t_1 t_2 \cdots t_m,$$

where $\sigma_0 \in S_k \times S_l$ and $t_1, \dots, t_m$ are transpositions in S_{k+l}.

Lemma 3.3 *For all $\tau \in S_{k+l}$ and $\theta = \theta_{\mu,\nu}$ as above, $z_\theta(\tau) = O_{k,l}(n^{-k-l-\|\tau\|_{k,l}})$ as $n \to \infty$.*

Proof Referring to (3-11), as $n \to \infty$, $\|\theta\|^{-2} = O_{k,l}(1)$ by Lemma 2.3 and the coefficients of $\mathfrak{p}_{\mu\otimes\nu}\big(\sum_{\sigma\in S_\mu\times S_\nu}\sigma\big)\mathfrak{p}_{\mu\otimes\nu}$ are clearly $O_{k,l}(1)$, so z_θ has the form

$$\Big(\sum_{\sigma\in S_k\times S_l} A(\sigma)\sigma\Big)\mathrm{Wg}_{n,k+l},$$

where each $A(\sigma)$ is $O_{k,l}(1)$. This means

$$z_\theta(\tau) = \sum_{\substack{\sigma\in S_k\times S_l\\ \sigma'\in S_{k+l}\\ \sigma\sigma'=\tau}} A(\sigma)\mathrm{Wg}_{n,k+l}(\sigma').$$

The order of any of the finitely many summands above is $n^{-k-l-|\sigma'|}$ by Proposition 2.7, and the minimum possible value of $|\sigma'|$ is $\|\tau\|_{k,l}$. ◻

Before moving on, it is useful to explain the operator $\varphi^{-1}[\rho^{k+l}(\pi)]$ for $\pi \in S_{k+l}$. For $I = (i_1, \dots, i_{k+l})$, let $I'(I;\pi) := i_{\pi(1)}, \dots, i_{\pi(k)}$ and $J'(I;\pi) := i_{\pi(k+1)}, \dots, i_{\pi(k+l)}$. As an element of

$$(\mathbb{C}^n)^{\otimes k+l} \otimes ((\mathbb{C}^n)^\vee)^{\otimes k+l},$$

$\rho^{k+l}(\pi)$ is given by

$$\sum_{\substack{I=(i_1,\dots,i_k)\\ J=(j_{k+1},\dots,j_{k+l})}} e_{I'(I\sqcup J;\pi)\sqcup J'(I\sqcup J;\pi)} \otimes \check{e}_{I\sqcup J},$$

so, from (3-6),

$$\varphi^{-1}[\rho^{k+l}(\pi)] = \sum_{\substack{I=(i_1,\dots,i_k)\\ J=(j_{k+1},\dots,j_{k+l})}} e^J_{I'(I\sqcup J;\pi)} \otimes \check{e}_I^{J'(I\sqcup J;\pi)}. \tag{3-12}$$

3.4 A combinatorial integration formula

In this rest of Section 3, we assume $g = 2$. All proofs extend to $g \ge 3$. We write $\{a,b,c,d\}$ for the generators of $\boldsymbol{F}_4$ and $R := [a,b][c,d]$. Assume both γ and w are not the identity and $w \in [\boldsymbol{F}_4, \boldsymbol{F}_4]$ according to the remarks at the beginning of Section 2.6. We write w in the reduced form

$$w = f_1^{\epsilon_1} f_2^{\epsilon_2} \dots f_{|w|}^{\epsilon_{|w|}}, \quad \epsilon_u \in \{\pm1\},\ f_u \in \{a,b,c,d\}, \tag{3-13}$$

where, if $f_u = f_{u+1}$, then $\epsilon_u = \epsilon_{u+1}$. For $f \in \{a, b, c, d\}$, let p_f denote the number of occurrences of f^{+1} in (3-13). The expression (3-13) implies that, for $h := (h_a, h_b, h_c, h_d) \in \mathsf{U}(n)^4$,

$$\operatorname{Tr}(w(h)) = \sum_{i_j \in [n]} (h_{f_1}^{\epsilon_1})_{i_1 i_2} (h_{f_2}^{\epsilon_2})_{i_2 i_3} \cdots (h_{f_{|w|}}^{\epsilon_{|w|}})_{i_{|w|} i_1}. \tag{3-14}$$

Working with this expression will be cumbersome so we explain a diagrammatic way to think about (3-14). This will be the starting point for how we eventually understand $\mathcal{J}_n(w, \mu, \nu)$ in terms of decorated surfaces. We begin with a collection of intervals as follows:

$\boldsymbol{w}$-intervals and the $\boldsymbol{w}$-loop Firstly, for every $j \in [w]$ with $f_j = f$ as in (3-13) and $\epsilon_j = 1$, we take a copy of $[0, 1]$ and direct it from 0 to 1.

In our constructions, every interval will have two directions: the *intrinsic direction* (which is the direction from 0 to 1) and the *assigned direction.* In the case just discussed, these agree, but in general they will not.

We write $[0, 1]_{f,j,w}$ for such an interval and $\mathfrak{I}^+_{f,w}$ for the collection of these intervals.

For every $j \in [w]$ with $f_j = f$ as in (3-13) and $\epsilon_j = -1$, we take a copy of $[0, 1]$ and direct this interval from 1 to 0. We write $[0, 1]_{f^{-1},j,w}$ for such an interval and $\mathfrak{I}^-_{f,w}$ for the collection of these intervals.

All the intervals described above are called *w-intervals*. There are $|w|$ of these intervals in total.

$\boldsymbol{w}$-intermediate-intervals Between each $[0, 1]_{f_j^{\epsilon_j},j,w}$ and $[0, 1]_{f_{j+1}^{\epsilon_{j+1}},j+1,w}$ we add a new interval connecting $1_{f_j^{\epsilon_j},j,w}$ to $0_{f_{j+1}^{\epsilon_{j+1}},j+1,w}$, where the indices j run mod $|w|$. These intervals added are called *w-intermediate-intervals*. Note that these intervals together with the w-intervals now form a closed cycle that is paved by $2|w|$ intervals alternating between w-intervals and w-intermediate-intervals. Starting at $[0, 1]_{f_1^{\epsilon_1},1,w}$, reading the directions and f-labels of the w-intervals so that every w-interval is traversed from 0 to 1 spells out the word w. The resulting circle is called the w-loop and the previously defined orientation of this loop is now fixed. See Figure 1 for an illustration of the w-loop in a particular example.

We now view the indices i_j as an assignment

$$\boldsymbol{a}\colon \{\text{endpoints of } w\text{-intervals}\} \to [n],$$

$$\boldsymbol{a}(0_{f,j,w}) := i_j, \quad \boldsymbol{a}(1_{f,j,w}) = i_{j+1}, \quad \boldsymbol{a}(0_{f^{-1},j,w}) = i_j, \quad \boldsymbol{a}(1_{f^{-1},j,w}) = i_{j+1}.$$

The condition that $\boldsymbol{a}$ comes from a single collection of i_j is precisely that *if two endpoints of w-intervals are connected by a w-intermediate-interval, they are assigned the same value by $\boldsymbol{a}$*. Let $\mathcal{A}(w)$ denote the collection of such $\boldsymbol{a}$. If I is any copy of $[0, 1]$, we write 0_I for the copy of 0 and 1_I for the copy of 1 in I. We can now write

$$\operatorname{Tr}(w(h)) = \sum_{\boldsymbol{a} \in \mathcal{A}(w)} \prod_{f \in \{a,b,c,d\}} \Biggl(\prod_{\mathfrak{i} \in \mathfrak{I}^+_{f,w}} h_{\boldsymbol{a}(0_\mathfrak{i})\boldsymbol{a}(1_\mathfrak{i})} \Biggr) \Biggl(\prod_{\mathfrak{j} \in \mathfrak{I}^-_{f,w}} \bar{h}_{\boldsymbol{a}(1_\mathfrak{j})\boldsymbol{a}(0_\mathfrak{j})} \Biggr).$$

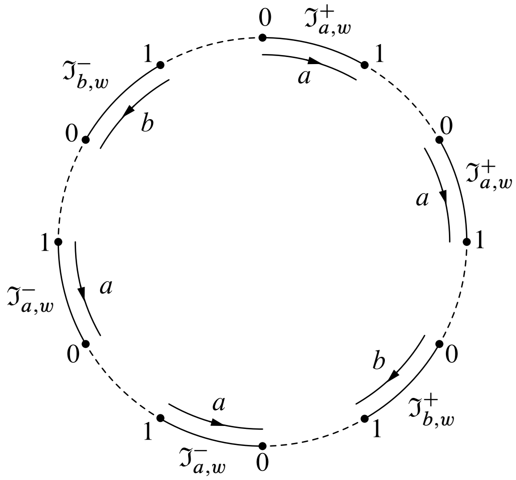


Figure 1: The w-loop for $w = a^2ba^{-2}b^{-1}$. The solid intervals are w-intervals and the dashed intervals are w-intermediate-intervals. We also label each interval by the set, eg $\mathfrak{I}^+_{a,w}$, to which they belong.

Now let v_p be an orthonormal basis for $W_n(\theta)$. We have

$$s_{\mu,\nu}(R_g(h_a,h_b,h_c,h_d)) = \sum_{p_i} \langle h_a v_{p_2}, v_{p_1}\rangle \langle h_b v_{p_3}, v_{p_2}\rangle \langle h_a^{-1} v_{p_4}, v_{p_3}\rangle \langle h_b^{-1} v_{p_5}, v_{p_4}\rangle \\ \cdot \langle h_c v_{p_6}, v_{p_5}\rangle \langle h_d v_{p_7}, v_{p_6}\rangle \langle h_c^{-1} v_{p_8}, v_{p_7}\rangle \langle h_d^{-1} v_{p_1}, v_{p_8}\rangle.$$

Here we have written eg $h_a v_{p_2}$ for $\pi_n^{k,l}(h_a)v_{p_2}$ to make things easier to read. Next we write each $v_p = \sum_{I,J} \beta^{J}_{pI} e^{J}_{I}$, where $\beta^{J}_{pI} := \langle v_p, e^{J}_{I}\rangle$. We then have

$$\begin{aligned}(3\text{-}15)\quad & \langle h_a v_{p_2}, v_{p_1}\rangle \langle h_b v_{p_3}, v_{p_2}\rangle \langle h_a^{-1} v_{p_4}, v_{p_3}\rangle \langle h_b^{-1} v_{p_5}, v_{p_4}\rangle \langle h_c v_{p_6}, v_{p_5}\rangle \langle h_d v_{p_7}, v_{p_6}\rangle \langle h_c^{-1} v_{p_8}, v_{p_7}\rangle \\ & \qquad \cdot \langle h_d^{-1} v_{p_1}, v_{p_8}\rangle \\ &= \sum_{\substack{\boldsymbol{r}_f,\boldsymbol{R}_f,\boldsymbol{V}_f,\boldsymbol{v}_f \\ \boldsymbol{U}_f,\boldsymbol{u}_f,\boldsymbol{s}_f,\boldsymbol{S}_f}} \beta^{\boldsymbol{V}_a}_{p_2\boldsymbol{s}_a} \bar\beta^{\boldsymbol{U}_a}_{p_1\boldsymbol{r}_a} \beta^{\boldsymbol{V}_b}_{p_3\boldsymbol{s}_b} \bar\beta^{\boldsymbol{U}_b}_{p_2\boldsymbol{r}_b} \beta^{\boldsymbol{u}_a}_{p_4\boldsymbol{R}_a} \bar\beta^{\boldsymbol{v}_a}_{p_3\boldsymbol{S}_a} \beta^{\boldsymbol{u}_b}_{p_5\boldsymbol{R}_b} \bar\beta^{\boldsymbol{v}_b}_{p_4\boldsymbol{S}_b} \beta^{\boldsymbol{V}_c}_{p_6\boldsymbol{s}_c} \bar\beta^{\boldsymbol{U}_c}_{p_5\boldsymbol{r}_c} \beta^{\boldsymbol{V}_d}_{p_7\boldsymbol{s}_d} \bar\beta^{\boldsymbol{U}_d}_{p_6\boldsymbol{r}_d} \beta^{\boldsymbol{u}_c}_{p_8\boldsymbol{R}_c} \\ & \qquad \cdot \bar\beta^{\boldsymbol{v}_c}_{p_7\boldsymbol{S}_c} \beta^{\boldsymbol{u}_d}_{p_1\boldsymbol{R}_d} \bar\beta^{\boldsymbol{v}_d}_{p_8\boldsymbol{S}_d} \langle h_a e^{\boldsymbol{V}_a}_{\boldsymbol{s}_a}, e^{\boldsymbol{U}_a}_{\boldsymbol{r}_a}\rangle \langle h_b e^{\boldsymbol{V}_b}_{\boldsymbol{s}_b}, e^{\boldsymbol{U}_b}_{\boldsymbol{r}_b}\rangle \langle h_a^{-1} e^{\boldsymbol{u}_a}_{\boldsymbol{R}_a}, e^{\boldsymbol{v}_a}_{\boldsymbol{S}_a}\rangle \langle h_b^{-1} e^{\boldsymbol{u}_b}_{\boldsymbol{R}_b}, e^{\boldsymbol{v}_b}_{\boldsymbol{S}_b}\rangle \\ & \qquad \cdot \langle h_c e^{\boldsymbol{V}_c}_{\boldsymbol{s}_c}, e^{\boldsymbol{U}_c}_{\boldsymbol{r}_c}\rangle \langle h_d e^{\boldsymbol{V}_d}_{\boldsymbol{s}_d}, e^{\boldsymbol{U}_d}_{\boldsymbol{r}_d}\rangle \langle h_c^{-1} e^{\boldsymbol{u}_c}_{\boldsymbol{R}_c}, e^{\boldsymbol{v}_c}_{\boldsymbol{S}_c}\rangle \langle h_d^{-1} e^{\boldsymbol{u}_d}_{\boldsymbol{R}_d}, e^{\boldsymbol{v}_d}_{\boldsymbol{S}_d}\rangle.\end{aligned}$$

We calculate

$$\begin{aligned}(3\text{-}16)\quad \langle h_f e^{\boldsymbol{V}_f}_{\boldsymbol{s}_f}, e^{\boldsymbol{U}_f}_{\boldsymbol{r}_f}\rangle \langle h_f^{-1} e^{\boldsymbol{u}_f}_{\boldsymbol{R}_f}, e^{\boldsymbol{v}_f}_{\boldsymbol{S}_f}\rangle &= \langle h_f e_{\boldsymbol{s}_f}, e_{\boldsymbol{r}_f}\rangle \overline{\langle h_f e_{\boldsymbol{V}_f}, e_{\boldsymbol{U}_f}\rangle}\, \overline{\langle h_f e_{\boldsymbol{S}_f}, e_{\boldsymbol{R}_f}\rangle} \langle h_f e_{\boldsymbol{v}_f}, e_{\boldsymbol{u}_f}\rangle \\ &= \langle h_f e_{\boldsymbol{s}_f \sqcup \boldsymbol{v}_f}, e_{\boldsymbol{r}_f \sqcup \boldsymbol{u}_f}\rangle \overline{\langle h_f e_{\boldsymbol{S}_f \sqcup \boldsymbol{V}_f}, e_{\boldsymbol{R}_f \sqcup \boldsymbol{U}_f}\rangle}.\end{aligned}$$

We now want a diagrammatic interpretation of (3-15) similarly to before. We make the following constructions:

***R*-intervals** For each $j \in [k]$ and $f \in \{a,b,c,d\}$, we make a copy of $[0,1]$, direct it from 0 to 1, label it by f, and also number it by j. We write $\mathfrak{I}^+_{f,R}$ for the collection of these intervals. These correspond to occurrences of f in R.

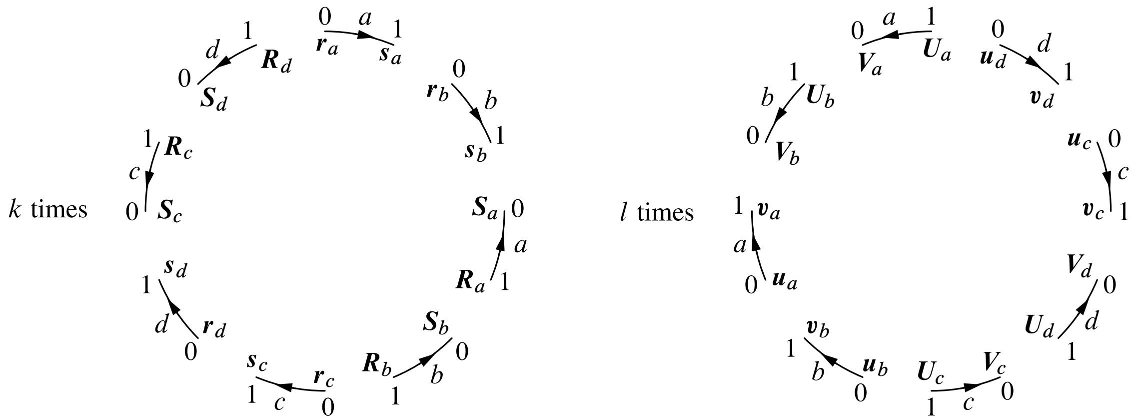


Figure 2: The R-intervals (left) and the R^{-1}-intervals (right). We have indicated their assigned direction and label (which f they correspond to). We have also, for each endpoint of an interval, indicated which index function, eg $\boldsymbol{r}_a$, has this endpoint in its domain.

For each $j \in [k]$ and $f \in \{a,b,c,d\}$, we make a copy of $[0,1]$, direct it from 1 to 0, label it by f, and also number it by j. We write $\mathfrak{I}^-_{f,R}$ for the collection of these intervals. These correspond to occurrences of f^{-1} in R.

(These two constructions of k intervals correspond to the presence of f and f^{-1} each exactly once in R.)

These intervals are called *R-intervals.* There are $8k$ R-intervals in total (for general g, there are $4gk$ of these intervals).

$\boldsymbol{R^{-1}}$-intervals For each $j \in [k+1,k+l]$ and $f \in \{a,b,c,d\}$, we make a copy of $[0,1]$, direct it from 0 to 1, label it by f, and also number it by j. We write $\mathfrak{I}^+_{f,R^{-1}}$ for the collection of these intervals. These correspond to occurrences of f in R^{-1}.

For each $j \in [k+1,k+l]$ and $f \in \{a,b,c,d\}$, we make a copy of $[0,1]$, direct it from 1 to 0, label it by f, and also number it by j. We write $\mathfrak{I}^-_{f,R^{-1}}$ for the collection of these intervals. These correspond to occurrences of f^{-1} in R^{-1}.

These intervals are called *R^{-1}-intervals.* There are $8l$ R^{-1}-intervals in total (for general g, there are $4gl$ of these intervals). See Figure 2 for an illustration of the R- and R^{-1}-intervals.

We now view (by identifying endpoints of intervals with the given numbers of intervals in $[k+l]$)

$$\begin{aligned}
\boldsymbol{r}_f &: \{0_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^+_{f,R}\} \to [n], & \boldsymbol{R}_f &: \{1_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^-_{f,R}\} \to [n],\\
\boldsymbol{s}_f &: \{1_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^+_{f,R}\} \to [n], & \boldsymbol{S}_f &: \{0_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^-_{f,R}\} \to [n],\\
\boldsymbol{U}_f &: \{1_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^-_{f,R^{-1}}\} \to [n], & \boldsymbol{u}_f &: \{0_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^+_{f,R^{-1}}\} \to [n],\\
\boldsymbol{V}_f &: \{0_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^-_{f,R^{-1}}\} \to [n], & \boldsymbol{v}_f &: \{1_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^+_{f,R^{-1}}\} \to [n].
\end{aligned}$$

We obtain, from (3-16),

$$\langle h_a e^{\boldsymbol{V}_a}_{\boldsymbol{s}_a}, e^{\boldsymbol{U}_a}_{\boldsymbol{r}_a}\rangle\langle h_b e^{\boldsymbol{V}_b}_{\boldsymbol{s}_b}, e^{\boldsymbol{U}_b}_{\boldsymbol{r}_b}\rangle\langle h_a^{-1} e^{\boldsymbol{u}_a}_{\boldsymbol{R}_a}, e^{\boldsymbol{v}_a}_{\boldsymbol{S}_a}\rangle\langle h_b^{-1} e^{\boldsymbol{u}_b}_{\boldsymbol{R}_b}, e^{\boldsymbol{v}_b}_{\boldsymbol{S}_b}\rangle \cdot \langle h_c e^{\boldsymbol{V}_c}_{\boldsymbol{s}_c}, e^{\boldsymbol{U}_c}_{\boldsymbol{r}_c}\rangle\langle h_d e^{\boldsymbol{V}_d}_{\boldsymbol{s}_d}, e^{\boldsymbol{U}_d}_{\boldsymbol{r}_d}\rangle\langle h_c^{-1} e^{\boldsymbol{u}_c}_{\boldsymbol{R}_c}, e^{\boldsymbol{v}_c}_{\boldsymbol{S}_c}\rangle\langle h_d^{-1} e^{\boldsymbol{u}_d}_{\boldsymbol{R}_d}, e^{\boldsymbol{v}_d}_{\boldsymbol{S}_d}\rangle$$

$$= \prod_f \prod_{\substack{\mathfrak{i}^+\in\mathfrak{I}^+_{f,R}\\ \mathfrak{i}^-\in\mathfrak{I}^-_{f,R}}} \prod_{\substack{j^+\in\mathfrak{I}^+_{f,R^{-1}}\\ j^-\in\mathfrak{I}^-_{f,R^{-1}}}} h_{\boldsymbol{r}_f(0_{\mathfrak{i}^+})\boldsymbol{s}_f(1_{\mathfrak{i}^+})} h_{\boldsymbol{u}_f(0_{\mathfrak{j}^+})\boldsymbol{v}_f(1_{\mathfrak{j}^+})} \bar{h}_{\boldsymbol{R}_f(1_{\mathfrak{i}^-})\boldsymbol{S}_f(0_{\mathfrak{i}^-})} \bar{h}_{\boldsymbol{U}_f(1_{\mathfrak{j}^-})\boldsymbol{V}_f(0_{\mathfrak{j}^-})}.$$

With this formalism, we obtain

$$\mathcal{J}_n(w,\mu,\nu) = \sum_{p_i} \sum_{\substack{\boldsymbol{r}_f,\boldsymbol{R}_f,\boldsymbol{V}_f,\boldsymbol{v}_f\\ \boldsymbol{U}_f,\boldsymbol{u}_f,\boldsymbol{s}_f,\boldsymbol{S}_f}} \sum_{\boldsymbol{a}\in\mathscr{A}(w)} \beta^{\boldsymbol{V}_a}_{p_2\boldsymbol{s}_a}\bar{\beta}^{\boldsymbol{U}_a}_{p_1\boldsymbol{r}_a}\beta^{\boldsymbol{V}_b}_{p_3\boldsymbol{s}_b}\bar{\beta}^{\boldsymbol{U}_b}_{p_2\boldsymbol{r}_b}\beta^{\boldsymbol{u}_a}_{p_4\boldsymbol{R}_a}\bar{\beta}^{\boldsymbol{v}_a}_{p_3\boldsymbol{S}_a}\beta^{\boldsymbol{u}_b}_{p_5\boldsymbol{R}_b}\bar{\beta}^{\boldsymbol{v}_b}_{p_4\boldsymbol{S}_b}\beta^{\boldsymbol{V}_c}_{p_6\boldsymbol{s}_c} \cdot \bar{\beta}^{\boldsymbol{U}_c}_{p_5\boldsymbol{r}_c}\beta^{\boldsymbol{V}_d}_{p_7\boldsymbol{s}_d}\bar{\beta}^{\boldsymbol{U}_d}_{p_6\boldsymbol{r}_d}\beta^{\boldsymbol{u}_c}_{p_8\boldsymbol{R}_c}\bar{\beta}^{\boldsymbol{v}_c}_{p_7\boldsymbol{S}_c}\beta^{\boldsymbol{u}_d}_{p_1\boldsymbol{R}_d}\bar{\beta}^{\boldsymbol{v}_d}_{p_8\boldsymbol{S}_d}$$
$$\cdot \prod_{f\in\{a,b,c,d\}} \int_{h\in\mathsf{U}(n)} \prod_{\substack{\mathfrak{i}\in\mathfrak{I}^+_{f,w},\mathfrak{j}\in\mathfrak{I}^-_{f,w}\\ \mathfrak{i}^+\in\mathfrak{I}^+_{f,R},\mathfrak{i}^-\in\mathfrak{I}^-_{f,R}\\ j^+\in\mathfrak{I}^+_{f,R^{-1}},j^-\in\mathfrak{I}^-_{f,R^{-1}}}} h_{\boldsymbol{r}_f(0_{\mathfrak{i}^+})\boldsymbol{s}_f(1_{\mathfrak{i}^+})} h_{\boldsymbol{u}_f(0_{\mathfrak{j}^+})\boldsymbol{v}_f(1_{\mathfrak{j}^+})} \bar{h}_{\boldsymbol{R}_f(1_{\mathfrak{i}^-})\boldsymbol{S}_f(0_{\mathfrak{i}^-})} \cdot \bar{h}_{\boldsymbol{U}_f(1_{\mathfrak{j}^-})\boldsymbol{V}_f(0_{\mathfrak{j}^-})}\,dh. \tag{3-17}$$

For each f, the integral in (3-17) can be done using the Weingarten calculus (Theorem 2.5). To do this, fix bijections for each $f\in\{a,b,c,d\}$

$$\mathfrak{I}^+_f := \mathfrak{I}^+_{f,R}\cup\mathfrak{I}^+_{f,R^{-1}}\cup\mathfrak{I}^+_{f,w}\cong[k+l+p_f],$$
$$\mathfrak{I}^-_f := \mathfrak{I}^-_{f,R}\cup\mathfrak{I}^-_{f,R^{-1}}\cup\mathfrak{I}^-_{f,w}\cong[k+l+p_f]$$

such that

$$\mathfrak{I}^+_{f,w}\cong[k+l+1,k+l+p_f],\quad \mathfrak{I}^-_{f,w}\cong[k+l+1,k+l+p_f]$$

and

$$\mathfrak{I}^+_{f,R}\cong[k],\quad \mathfrak{I}^-_{f,R}\cong[k],\quad \mathfrak{I}^+_{f,R^{-1}}\cong[k+1,k+l],\quad \mathfrak{I}^-_{f,R^{-1}}\cong[k+1,k+l] \tag{3-18}$$

correspond to the original numberings of $\mathfrak{I}^+_{f,R}$, $\mathfrak{I}^-_{f,R}$, $\mathfrak{I}^+_{f,R^{-1}}$ and $\mathfrak{I}^-_{f,R^{-1}}$.

Hence, if $\sigma_f,\tau_f\in S_{k+l+p_f}$ we view $\sigma_f,\tau_f:\mathfrak{I}^+_f\to\mathfrak{I}^-_f$ by the above fixed bijections. For each $f\in\{a,b,c,d\}$, we say $(\boldsymbol{a},\boldsymbol{r}_f,\boldsymbol{u}_f,\boldsymbol{R}_f,\boldsymbol{U}_f)\to\sigma_f$ if, for all $\mathfrak{i}\in\mathfrak{I}^+_f$ and $\mathfrak{i}'\in\mathfrak{I}^-_f$ with $\sigma_f(\mathfrak{i})=\mathfrak{i}'$, we have

$$[\boldsymbol{r}_f\sqcup\boldsymbol{u}_f\sqcup\boldsymbol{a}](0_{\mathfrak{i}}) = [\boldsymbol{R}_f\sqcup\boldsymbol{U}_f\sqcup\boldsymbol{a}](1_{\mathfrak{i}'});$$

here we wrote eg $[\boldsymbol{r}_f\sqcup\boldsymbol{u}_f\sqcup\boldsymbol{a}]$ for the function that $\boldsymbol{a}$, $\boldsymbol{r}_f$ and $\boldsymbol{u}_f$ induce on $\{0_{\mathfrak{i}}:\mathfrak{i}\in\mathfrak{I}^+_f\}$. Similarly, we say $(\boldsymbol{a},\boldsymbol{s}_f,\boldsymbol{v}_f,\boldsymbol{S}_f,\boldsymbol{V}_f)\to\tau_f$ if, for all $\mathfrak{i}\in\mathfrak{I}^+_f$, $\mathfrak{i}'\in\mathfrak{I}^-_f$ with $\tau_f(\mathfrak{i})=\mathfrak{i}'$, we have

$$[\boldsymbol{s}_f\sqcup\boldsymbol{v}_f\sqcup\boldsymbol{a}](1_{\mathfrak{i}}) = [\boldsymbol{S}_f\sqcup\boldsymbol{V}_f\sqcup\boldsymbol{a}](0_{\mathfrak{i}'}).$$

Theorem 2.5 translates to

$$\int_{h\in\mathsf{U}(n)} \prod_{\substack{\mathfrak{i}\in\mathfrak{I}^+_{f,w},\mathfrak{j}\in\mathfrak{I}^-_{f,w}\\ \mathfrak{i}^+\in\mathfrak{I}^+_{f,R},\mathfrak{i}^-\in\mathfrak{I}^-_{f,R}\\ j^+\in\mathfrak{I}^+_{f,R-1},j^-\in\mathfrak{I}^-_{f,R-1}}} h_{\boldsymbol{r}_f(0_{\mathfrak{i}^+})\boldsymbol{s}_f(1_{\mathfrak{i}^+})} h_{\boldsymbol{u}_f(0_{\mathfrak{j}^+})\boldsymbol{v}_f(1_{\mathfrak{j}^+})} \bar{h}_{\boldsymbol{R}_f(1_{\mathfrak{i}^-})\boldsymbol{S}_f(0_{\mathfrak{i}^-})} \bar{h}_{\boldsymbol{U}_f(1_{\mathfrak{j}^-})\boldsymbol{V}_f(0_{\mathfrak{j}^-})}\,dh$$
$$= \sum_{\sigma_f,\tau_f\in S_{k+l+p_f}} \mathrm{Wg}_{n,k+l+p_f}(\sigma_f\tau_f^{-1})\mathbb{1}\{(\boldsymbol{a},\boldsymbol{r}_f,\boldsymbol{u}_f,\boldsymbol{R}_f,\boldsymbol{U}_f)\to\sigma_f,(\boldsymbol{a},\boldsymbol{s}_f,\boldsymbol{v}_f,\boldsymbol{S}_f,\boldsymbol{V}_f)\to\tau_f\},$$

so putting this into (3-17) gives

$$\mathcal{I}_n(w,\mu,\nu) = \sum_{\sigma_f,\tau_f\in S_{k+l+p_f}} \Biggl(\prod_{f\in\{a,b,c,d\}} \mathrm{Wg}_{n,k+l+p_f}(\sigma_f\tau_f^{-1})\Biggr)$$
$$\cdot\sum_{p_i}\sum_{\substack{\boldsymbol{a}\in\mathcal{A}(w),\boldsymbol{r}_f,\boldsymbol{R}_f,\boldsymbol{V}_f,\boldsymbol{v}_f,\boldsymbol{U}_f,\boldsymbol{u}_f,\boldsymbol{s}_f,\boldsymbol{S}_f\\ (\boldsymbol{a},\boldsymbol{r}_f,\boldsymbol{u}_f,\boldsymbol{R}_f,\boldsymbol{U}_f)\to\sigma_f\\ (\boldsymbol{a},\boldsymbol{s}_f,\boldsymbol{v}_f,\boldsymbol{S}_f,\boldsymbol{V}_f)\to\tau_f}} \beta^{\boldsymbol{V}_a}_{p_2\boldsymbol{s}_a}\bar{\beta}^{\boldsymbol{U}_a}_{p_1\boldsymbol{r}_a}\beta^{\boldsymbol{V}_b}_{p_3\boldsymbol{s}_b}\bar{\beta}^{\boldsymbol{U}_b}_{p_2\boldsymbol{r}_b}\beta^{\boldsymbol{u}_a}_{p_4\boldsymbol{R}_a}\bar{\beta}^{\boldsymbol{v}_a}_{p_3\boldsymbol{S}_a}\beta^{\boldsymbol{u}_b}_{p_5\boldsymbol{R}_b}\bar{\beta}^{\boldsymbol{v}_b}_{p_4\boldsymbol{S}_b}$$
$$\cdot\beta^{\boldsymbol{V}_c}_{p_6\boldsymbol{s}_c}\bar{\beta}^{\boldsymbol{U}_c}_{p_5\boldsymbol{r}_c}\beta^{\boldsymbol{V}_d}_{p_7\boldsymbol{s}_d}\bar{\beta}^{\boldsymbol{U}_d}_{p_6\boldsymbol{r}_d}\beta^{\boldsymbol{u}_c}_{p_8\boldsymbol{R}_c}\bar{\beta}^{\boldsymbol{v}_c}_{p_7\boldsymbol{S}_c}\beta^{\boldsymbol{u}_d}_{p_1\boldsymbol{R}_d}\bar{\beta}^{\boldsymbol{v}_d}_{p_8\boldsymbol{S}_d}.$$

Here we make our main improvement over the classical Weingarten calculus. We introduce the following beneficial property that the σ_f and τ_f possibly have:

Forbidden matchings property For every $f\in\{a,b,c,d\}$, the following hold: neither σ_f nor τ_f map any element of $\mathfrak{I}^+_{f,R}$ to an element of $\mathfrak{I}^-_{f,R-1}$, or map an element of $\mathfrak{I}^+_{f,R-1}$ to an element of $\mathfrak{I}^-_{f,R}$.

We have the following key lemma:

Lemma 3.4 *If for some $f\in\{a,b,c,d\}$, σ_f and τ_f do **not** have the **forbidden matchings** property, then, for any choice of $p_1,\dots,p_8$,*

$$\sum_{\substack{\boldsymbol{a}\in\mathcal{A}(w),\boldsymbol{r}_f,\boldsymbol{R}_f,\boldsymbol{V}_f,\boldsymbol{v}_f,\boldsymbol{U}_f,\boldsymbol{u}_f,\boldsymbol{s}_f,\boldsymbol{S}_f\\ (\boldsymbol{a},\boldsymbol{r}_f,\boldsymbol{u}_f,\boldsymbol{R}_f,\boldsymbol{U}_f)\to\sigma_f\\ (\boldsymbol{a},\boldsymbol{s}_f,\boldsymbol{v}_f,\boldsymbol{S}_f,\boldsymbol{V}_f)\to\tau_f}} \beta^{\boldsymbol{V}_a}_{p_2\boldsymbol{s}_a}\bar{\beta}^{\boldsymbol{U}_a}_{p_1\boldsymbol{r}_a}\beta^{\boldsymbol{V}_b}_{p_3\boldsymbol{s}_b}\bar{\beta}^{\boldsymbol{U}_b}_{p_2\boldsymbol{r}_b}\beta^{\boldsymbol{u}_a}_{p_4\boldsymbol{R}_a}\bar{\beta}^{\boldsymbol{v}_a}_{p_3\boldsymbol{S}_a}\beta^{\boldsymbol{u}_b}_{p_5\boldsymbol{R}_b}\bar{\beta}^{\boldsymbol{v}_b}_{p_4\boldsymbol{S}_b}$$
$$\cdot\beta^{\boldsymbol{V}_c}_{p_6\boldsymbol{s}_c}\bar{\beta}^{\boldsymbol{U}_c}_{p_5\boldsymbol{r}_c}\beta^{\boldsymbol{V}_d}_{p_7\boldsymbol{s}_d}\bar{\beta}^{\boldsymbol{U}_d}_{p_6\boldsymbol{r}_d}\beta^{\boldsymbol{u}_c}_{p_8\boldsymbol{R}_c}\bar{\beta}^{\boldsymbol{v}_c}_{p_7\boldsymbol{S}_c}\beta^{\boldsymbol{u}_d}_{p_1\boldsymbol{R}_d}\bar{\beta}^{\boldsymbol{v}_d}_{p_8\boldsymbol{S}_d} = 0. \tag{3-19}$$

Proof Indeed, suppose σ_a matches an element $\mathfrak{i}\in\mathfrak{I}^+_{a,R}$ with $\mathfrak{j}\in\mathfrak{I}^-_{a,R-1}$; $\sigma_a(\mathfrak{i})=\mathfrak{j}$. With our given fixed bijections (3-18), $\mathfrak{i}$ corresponds to an element of $[k]$ and $\mathfrak{j}$ corresponds to an element of $[k+1,k+l]$. Without loss of generality in the argument suppose that $0_\mathfrak{i}$ corresponds to 1 and $0_\mathfrak{j}$ corresponds to $k+1$. The condition $\sigma_a(\mathfrak{i})=\mathfrak{j}$ and $(\boldsymbol{a},\boldsymbol{r}_a,\boldsymbol{u}_a,\boldsymbol{R}_a,\boldsymbol{U}_a)\to\sigma_f$ means that, as functions on $[k]$ and $[k+1,k+l]$, $\boldsymbol{r}_a(1)=\boldsymbol{U}_a(k+1)$. There are no other constraints on these values.

Then, for all variables in (3-19) fixed apart from $\boldsymbol{r}_a$ and $\boldsymbol{U}_a$, and all values of $\boldsymbol{r}_a$ and $\boldsymbol{U}_a$ fixed other than $\boldsymbol{r}_a(1)$ and $\boldsymbol{U}_a(k+1)$, the ensuing sum over $\boldsymbol{r}_a$ and $\boldsymbol{U}_a$ is

$$\sum_{\boldsymbol{r}_a(1)=\boldsymbol{U}_a(k+1)} \beta^{\boldsymbol{U}_a}_{p_2\boldsymbol{r}_a}.$$

But, recalling the contraction operators from (2-2), this sum is the coordinate of

$$e_{\boldsymbol{r}_a(2)} \otimes \cdots\cdots e_{\boldsymbol{r}_a(k)} \otimes \check{e}_{\boldsymbol{U}_a(k+2)} \otimes \cdots \otimes \check{e}_{\boldsymbol{U}_a(k+l)}$$

in $c_{1,1}(v_{p_2})$. But $c_{1,1}(v_{p_2}) = 0$ because $v_{p_2} \in \dot{\mathfrak{T}}_n^{k,l}$. $\square$

We henceforth write $\sum^*_{\sigma_f,\tau_f}$ *to mean the sum is restricted to* σ_f *and* τ_f *satisfying the* ***forbidden matchings*** *property.* Lemma 3.4 now implies

$$\begin{aligned}(3\text{-}20)\quad &\mathcal{J}_n(w,\mu,\nu)\\ &= \sum^*_{\sigma_f,\tau_f \in S_{k+l+p_f}} \Bigg(\prod_{f\in\{a,b,c,d\}} \mathrm{Wg}_{n,k+l+p_f}(\sigma_f\tau_f^{-1}) \Bigg) \\ &\quad \cdot \sum_{p_i} \sum_{\substack{\boldsymbol{a}\in\mathcal{A}(w), \boldsymbol{r}_f,\boldsymbol{R}_f,\boldsymbol{V}_f,\boldsymbol{v}_f,\boldsymbol{U}_f,\boldsymbol{u}_f,\boldsymbol{s}_f,\boldsymbol{S}_f \\ (\boldsymbol{a},\boldsymbol{r}_f,\boldsymbol{u}_f,\boldsymbol{R}_f,\boldsymbol{U}_f)\to\sigma_f \\ (\boldsymbol{a},\boldsymbol{s}_f,\boldsymbol{v}_f,\boldsymbol{S}_f,\boldsymbol{V}_f)\to\tau_f}} \beta^{\boldsymbol{V}_a}_{p_2\boldsymbol{s}_a}\bar{\beta}^{\boldsymbol{U}_a}_{p_1\boldsymbol{r}_a}\beta^{\boldsymbol{V}_b}_{p_3\boldsymbol{s}_b}\bar{\beta}^{\boldsymbol{U}_b}_{p_2\boldsymbol{r}_b}\beta^{\boldsymbol{u}_a}_{p_4\boldsymbol{R}_a}\bar{\beta}^{\boldsymbol{v}_a}_{p_3\boldsymbol{S}_a}\beta^{\boldsymbol{u}_b}_{p_5\boldsymbol{R}_b}\bar{\beta}^{\boldsymbol{v}_b}_{p_4\boldsymbol{S}_b} \\ &\qquad\qquad \cdot \beta^{\boldsymbol{V}_c}_{p_6\boldsymbol{s}_c}\bar{\beta}^{\boldsymbol{U}_c}_{p_5\boldsymbol{r}_c}\beta^{\boldsymbol{V}_d}_{p_7\boldsymbol{s}_d}\bar{\beta}^{\boldsymbol{U}_d}_{p_6\boldsymbol{r}_d}\beta^{\boldsymbol{u}_c}_{p_8\boldsymbol{R}_c}\bar{\beta}^{\boldsymbol{v}_c}_{p_7\boldsymbol{S}_c}\beta^{\boldsymbol{u}_d}_{p_1\boldsymbol{R}_d}\bar{\beta}^{\boldsymbol{v}_d}_{p_8\boldsymbol{S}_d}.\end{aligned}$$

Moreover, we can significantly tidy up (3-20). For everything in (3-20) fixed except for eg p_2, the ensuing sum over p_2 is

$$\sum_{p_2} \beta^{\boldsymbol{V}_a}_{p_2\boldsymbol{s}_a}\bar{\beta}^{\boldsymbol{U}_b}_{p_2\boldsymbol{r}_b} = \sum_{p_2} \langle e^{\boldsymbol{U}_b}_{\boldsymbol{r}_b}, v_{p_2}\rangle\langle v_{p_2}, e^{\boldsymbol{V}_a}_{\boldsymbol{s}_a}\rangle = \langle \mathfrak{q}_\theta e^{\boldsymbol{U}_b}_{\boldsymbol{r}_b}, e^{\boldsymbol{V}_a}_{\boldsymbol{s}_a}\rangle.$$

Therefore, executing the sums over p_i in (3-20), we replace the sum over p_i and the product over β-terms by

$$\begin{aligned}(3\text{-}21)\quad &\langle \mathfrak{q}_\theta e^{\boldsymbol{U}_b}_{\boldsymbol{r}_b}, e^{\boldsymbol{V}_a}_{\boldsymbol{s}_a}\rangle\langle \mathfrak{q}_\theta e^{\boldsymbol{v}_a}_{\boldsymbol{S}_a}, e^{\boldsymbol{V}_b}_{\boldsymbol{s}_b}\rangle\langle \mathfrak{q}_\theta e^{\boldsymbol{v}_b}_{\boldsymbol{S}_b}, e^{\boldsymbol{u}_a}_{\boldsymbol{R}_a}\rangle\langle \mathfrak{q}_\theta e^{\boldsymbol{U}_c}_{\boldsymbol{r}_c}, e^{\boldsymbol{u}_b}_{\boldsymbol{R}_b}\rangle\langle \mathfrak{q}_\theta e^{\boldsymbol{U}_d}_{\boldsymbol{r}_d}, e^{\boldsymbol{V}_c}_{\boldsymbol{s}_c}\rangle\langle \mathfrak{q}_\theta e^{\boldsymbol{v}_c}_{\boldsymbol{S}_c}, e^{\boldsymbol{V}_d}_{\boldsymbol{s}_d}\rangle\langle \mathfrak{q}_\theta e^{\boldsymbol{v}_d}_{\boldsymbol{S}_d}, e^{\boldsymbol{u}_c}_{\boldsymbol{R}_c}\rangle \\ &\qquad \cdot \langle \mathfrak{q}_\theta e^{\boldsymbol{U}_a}_{\boldsymbol{r}_a}, e^{\boldsymbol{u}_d}_{\boldsymbol{R}_d}\rangle.\end{aligned}$$

By Proposition 3.2, we have eg

$$\langle \mathfrak{q}_\theta e^{\boldsymbol{U}_b}_{\boldsymbol{r}_b}, e^{\boldsymbol{V}_a}_{\boldsymbol{s}_a}\rangle = D_{\mu,\nu}(n) \sum_{\pi\in S_{k+l}} z_\theta(\pi)\langle \varphi^{-1}[\rho_n^{k,l}(\pi)] e^{\boldsymbol{U}_b}_{\boldsymbol{r}_b}, e^{\boldsymbol{V}_a}_{\boldsymbol{s}_a}\rangle.$$

Now recall from (3-12) that

$$\varphi^{-1}[\rho_n^{k+l}(\pi)] = \sum_{\substack{I=(i_1,\dots,i_k)\\ J=(j_{k+1},\dots,j_{k+l})}} e^J_{I'(I\sqcup J;\pi)} \otimes \check{e}_I^{J'(I\sqcup J;\pi)}.$$

This means that $\langle \varphi^{-1}[\rho_n^{k+l}(\pi)] e^{\boldsymbol{U}_b}_{\boldsymbol{r}_b}, e^{\boldsymbol{V}_a}_{\boldsymbol{s}_a}\rangle$ is equal to either 0 or 1 and $\langle \varphi^{-1}[\rho_n^{k+l}(\pi)] e^{\boldsymbol{U}_b}_{\boldsymbol{r}_b}, e^{\boldsymbol{V}_a}_{\boldsymbol{s}_a}\rangle = 1$ if and only if, letting (3-18) induce identifications

$$\{1_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^+_{a,R}\} \cong [k], \quad \{1_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^-_{b,R^{-1}}\} \cong [k+1,k+l],$$
$$\{0_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^+_{b,R}\} \cong [k], \quad \{0_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^-_{a,R^{-1}}\} \cong [k+1,k+l]$$

via their given indexing of intervals, we have $[\boldsymbol{s}_a \sqcup \boldsymbol{U}_b] \circ \pi = [\boldsymbol{r}_b \sqcup \boldsymbol{V}_a]$, where eg $\boldsymbol{s}_a \sqcup \boldsymbol{U}_b$ is the function either on endpoints of intervals or on $[k+l]$ induced by the union of $\boldsymbol{s}_a$ and $\boldsymbol{U}_b$. Hence, repeating this argument,

$$\begin{aligned}(3\text{-}21) = D_{\mu,\nu}(n)^8 \sum_{\pi_1,\ldots,\pi_8 \in S_{k+l}} \biggl(\prod_{i=1}^{8} z_\theta(\pi_i)\biggr) \mathbb{1}\bigl\{&[\boldsymbol{s}_a \sqcup \boldsymbol{U}_b]\circ\pi_1 = [\boldsymbol{r}_b \sqcup \boldsymbol{V}_a], [\boldsymbol{s}_b \sqcup \boldsymbol{v}_a]\circ\pi_2 = [\boldsymbol{S}_a \sqcup \boldsymbol{V}_b],\\ &[\boldsymbol{R}_a \sqcup \boldsymbol{v}_b]\circ\pi_3 = [\boldsymbol{S}_b \sqcup \boldsymbol{u}_a], [\boldsymbol{R}_b \sqcup \boldsymbol{U}_c]\circ\pi_4 = [\boldsymbol{r}_c \sqcup \boldsymbol{u}_b],\\ &[\boldsymbol{s}_c \sqcup \boldsymbol{U}_d]\circ\pi_5 = [\boldsymbol{r}_d \sqcup \boldsymbol{V}_c], [\boldsymbol{s}_d \sqcup \boldsymbol{v}_c]\circ\pi_6 = [\boldsymbol{S}_c \sqcup \boldsymbol{V}_d],\\ &[\boldsymbol{R}_c \sqcup \boldsymbol{v}_d]\circ\pi_7 = [\boldsymbol{S}_d \sqcup \boldsymbol{u}_c], [\boldsymbol{R}_d \sqcup \boldsymbol{U}_a]\circ\pi_8 = [\boldsymbol{r}_a \sqcup \boldsymbol{u}_d]\bigr\}.\end{aligned}$$

Putting all these arguments together gives

$$\begin{aligned}&\mathcal{I}_n(w,\mu,\nu)\\ &= D_{\mu,\nu}(n)^8 \sideset{}{^*}\sum_{\sigma_f,\tau_f \in S_{p_f+k+l}} \sum_{\pi_1,\ldots,\pi_8 \in S_{k+l}} \biggl(\prod_{f\in\{a,b,c,d\}} \mathrm{Wg}_{n,k+l+p_f}(\sigma_f\tau_f^{-1})\biggr)\biggl(\prod_{i=1}^{8} z_\theta(\pi_i)\biggr)\\ &\qquad \cdot \sum_{p_i} \sum_{\substack{\boldsymbol{a}\in\mathcal{A}(w),\boldsymbol{r}_f,\boldsymbol{R}_f,\boldsymbol{V}_f,\boldsymbol{v}_f,\boldsymbol{U}_f,\boldsymbol{u}_f,\boldsymbol{s}_f,\boldsymbol{S}_f\\ (\boldsymbol{a},\boldsymbol{r}_f,\boldsymbol{u}_f,\boldsymbol{R}_f,\boldsymbol{U}_f)\to\sigma_f\\ (\boldsymbol{a},\boldsymbol{s}_f,\boldsymbol{v}_f,\boldsymbol{S}_f,\boldsymbol{V}_f)\to\tau_f}} \mathbb{1}\bigl\{[\boldsymbol{s}_a \sqcup \boldsymbol{U}_b]\circ\pi_1 = [\boldsymbol{r}_b \sqcup \boldsymbol{V}_a], [\boldsymbol{s}_b \sqcup \boldsymbol{v}_a]\circ\pi_2 = [\boldsymbol{S}_a \sqcup \boldsymbol{V}_b],\\ &\qquad\qquad [\boldsymbol{R}_a \sqcup \boldsymbol{v}_b]\circ\pi_3 = [\boldsymbol{S}_b \sqcup \boldsymbol{u}_a], [\boldsymbol{R}_b \sqcup \boldsymbol{U}_c]\circ\pi_4 = [\boldsymbol{r}_c \sqcup \boldsymbol{u}_b],\\ &\qquad\qquad [\boldsymbol{s}_c \sqcup \boldsymbol{U}_d]\circ\pi_5 = [\boldsymbol{r}_d \sqcup \boldsymbol{V}_c], [\boldsymbol{s}_d \sqcup \boldsymbol{v}_c]\circ\pi_6 = [\boldsymbol{S}_c \sqcup \boldsymbol{V}_d],\\ &\qquad\qquad [\boldsymbol{R}_c \sqcup \boldsymbol{v}_d]\circ\pi_7 = [\boldsymbol{S}_d \sqcup \boldsymbol{u}_c], [\boldsymbol{R}_d \sqcup \boldsymbol{U}_a]\circ\pi_8 = [\boldsymbol{r}_a \sqcup \boldsymbol{u}_d]\bigr\}.\end{aligned}$$

This formula says that we can calculate $\mathcal{I}_n(w,\mu,\nu)$ by summing over some combinatorial data of matchings (the σ_f, τ_f and π_i) a quantity that we can understand well times a count of the number of indices that satisfy the prescribed matchings. To formalize this point of view we make the following definition:

Definition 3.5 A *matching datum* of the triple (w,k,l) is a pair $(\sigma_f,\tau_f) \in S_{k+l+p_f} \times S_{k+l+p_f}$ as above, satisfying the *forbidden matchings* property for each $f \in \{a,b,c,d\}$, together with $(\pi_1,\ldots,\pi_8) \in (S_{k+l})^8$. We write

$$\mathsf{MATCH}(w,k,l)$$

for the finite collection of all matching data for (w,k,l).

Given a matching datum $\{\sigma_f,\tau_f,\pi_i\}$, we write $\mathcal{N}(\{\sigma_f,\tau_f,\pi_i\})$ for the number of choices of $\boldsymbol{a} \in \mathcal{A}(w)$, $\boldsymbol{r}_f$, $\boldsymbol{R}_f$, $\boldsymbol{V}_f$, $\boldsymbol{v}_f$, $\boldsymbol{U}_f$, $\boldsymbol{u}_f$, $\boldsymbol{s}_f$ and $\boldsymbol{S}_f$ such that

$$\begin{aligned} &(\boldsymbol{a},\boldsymbol{r}_f,\boldsymbol{u}_f,\boldsymbol{R}_f,\boldsymbol{U}_f) \to \sigma_f, && (\boldsymbol{a},\boldsymbol{s}_f,\boldsymbol{v}_f,\boldsymbol{S}_f,\boldsymbol{V}_f) \to \tau_f,\\ &[\boldsymbol{s}_a \sqcup \boldsymbol{U}_b]\circ\pi_1 = [\boldsymbol{r}_b \sqcup \boldsymbol{V}_a], && [\boldsymbol{s}_b \sqcup \boldsymbol{v}_a]\circ\pi_2 = [\boldsymbol{S}_a \sqcup \boldsymbol{V}_b],\\ &[\boldsymbol{R}_a \sqcup \boldsymbol{v}_b]\circ\pi_3 = [\boldsymbol{S}_b \sqcup \boldsymbol{u}_a], && [\boldsymbol{R}_b \sqcup \boldsymbol{U}_c]\circ\pi_4 = [\boldsymbol{r}_c \sqcup \boldsymbol{u}_b],\\ &[\boldsymbol{s}_c \sqcup \boldsymbol{U}_d]\circ\pi_5 = [\boldsymbol{r}_d \sqcup \boldsymbol{V}_c], && [\boldsymbol{s}_d \sqcup \boldsymbol{v}_c]\circ\pi_6 = [\boldsymbol{S}_c \sqcup \boldsymbol{V}_d],\\ &[\boldsymbol{R}_c \sqcup \boldsymbol{v}_d]\circ\pi_7 = [\boldsymbol{S}_d \sqcup \boldsymbol{u}_c], && [\boldsymbol{R}_d \sqcup \boldsymbol{U}_a]\circ\pi_8 = [\boldsymbol{r}_a \sqcup \boldsymbol{u}_d].\end{aligned} \tag{3-22}$$

With this notation, we have proved the following theorem:

Theorem 3.6 *For $k+l>0$ with $\mu \vdash k$ and $\nu \vdash l$ and $w \in [\boldsymbol{F}_4, \boldsymbol{F}_4]$, we have*

$$\mathcal{I}_n(w,\mu,\nu) = D_{\mu,\nu}(n)^8 \sum_{\{\sigma_f,\tau_f,\pi_i\}\in \mathsf{MATCH}(w,k,l)} \Bigl(\prod_{i=1}^{8} z_\theta(\pi_i)\Bigr)\Bigl(\prod_{f\in\{a,b,c,d\}} \mathrm{Wg}_{n,k+l+p_f}(\sigma_f\tau_f^{-1})\Bigr) \cdot \mathcal{N}(\{\sigma_f,\tau_f,\pi_i\}). \tag{3-23}$$

We conclude this section by bounding the terms $z_\theta(\pi_i)$ and $\mathrm{Wg}_{n,k+l+p_f}(\sigma_f\tau_f^{-1})$ using Proposition 2.7 and Lemma 3.3, recalling also (2-1). Note that $\sum_{f\in\{a,b,c,d\}} p_f = \frac{1}{2}|w|$. This yields:

Corollary 3.7 *For $k+l>0$ with $\mu \vdash k$ and $\nu \vdash l$ and $w \in [\boldsymbol{F}_4, \boldsymbol{F}_4]$, we have*

$$\mathcal{I}_n(w,\mu,\nu) \ll_{k,l,w} n^{-4k-4l-|w|/2} \sum_{\{\sigma_f,\tau_f,\pi_i\}\in \mathsf{MATCH}(w,k,l)} n^{-\sum_f |\sigma_f\tau_f^{-1}| - \sum_{i=1}^{8} \|\pi_i\|_{k,l}} \mathcal{N}(\{\sigma_f,\tau_f,\pi_i\}). \tag{3-24}$$

We will proceed in the next section to understand all the quantities in (3-24) in topological terms by constructing a surface from each $\{\sigma_f,\tau_f,\pi_i\}$.

4 Topology

4.1 Construction of surfaces from matching data

We now show how a datum in $\mathsf{MATCH}(w,k,l)$ can be used to construct a surface such that the terms appearing in (3-23) can be bounded by topological features of the surface. This construction is similar to the constructions of [Magee and Puder 2019; 2015], but with the presence of additional π_i adding a new aspect. We continue to assume $g=2$ for simplicity. We can still assume that $\gamma \in [\Gamma_2,\Gamma_2]$ and hence $w \in [\boldsymbol{F}_4, \boldsymbol{F}_4]$.

Construction of the 1-skeleton

π-intervals The identifications of the previous section mean that we view

$$\begin{aligned}
\pi_1 &: \{0_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^+_{b,R} \cup \mathfrak{I}^-_{a,R-1}\} \to \{1_{\mathfrak{i}'} : \mathfrak{i}' \in \mathfrak{I}^+_{a,R} \cup \mathfrak{I}^-_{b,R-1}\},\\
\pi_2 &: \{0_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^-_{a,R} \cup \mathfrak{I}^-_{b,R-1}\} \to \{1_{\mathfrak{i}'} : \mathfrak{i}' \in \mathfrak{I}^+_{b,R} \cup \mathfrak{I}^+_{a,R-1}\},\\
\pi_3 &: \{0_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^-_{b,R} \cup \mathfrak{I}^+_{a,R-1}\} \to \{1_{\mathfrak{i}'} : \mathfrak{i}' \in \mathfrak{I}^-_{a,R} \cup \mathfrak{I}^+_{b,R-1}\},\\
\pi_4 &: \{0_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^+_{c,R} \cup \mathfrak{I}^+_{b,R-1}\} \to \{1_{\mathfrak{i}'} : \mathfrak{i}' \in \mathfrak{I}^-_{b,R} \cup \mathfrak{I}^-_{c,R-1}\},\\
\pi_5 &: \{0_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^+_{d,R} \cup \mathfrak{I}^-_{c,R-1}\} \to \{1_{\mathfrak{i}'} : \mathfrak{i}' \in \mathfrak{I}^+_{c,R} \cup \mathfrak{I}^-_{d,R-1}\},\\
\pi_6 &: \{0_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^-_{c,R} \cup \mathfrak{I}^-_{d,R-1}\} \to \{1_{\mathfrak{i}'} : \mathfrak{i}' \in \mathfrak{I}^+_{d,R} \cup \mathfrak{I}^+_{c,R-1}\},\\
\pi_7 &: \{0_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^-_{d,R} \cup \mathfrak{I}^+_{c,R-1}\} \to \{1_{\mathfrak{i}'} : \mathfrak{i}' \in \mathfrak{I}^-_{c,R} \cup \mathfrak{I}^+_{d,R-1}\},\\
\pi_8 &: \{0_{\mathfrak{i}} : \mathfrak{i} \in \mathfrak{I}^+_{a,R} \cup \mathfrak{I}^+_{d,R-1}\} \to \{1_{\mathfrak{i}'} : \mathfrak{i}' \in \mathfrak{I}^-_{d,R} \cup \mathfrak{I}^-_{a,R-1}\}.
\end{aligned} \tag{4-1}$$

We add an arc between any two interval endpoints that are mapped to one another by some π_i. All the intervals added here are called *π-intervals*. The purpose of this construction is that the conditions concerning π_i in (3-22) correspond to the fact that *two endpoints of intervals connected by a π-interval are assigned the same value in $[n]$ by the relevant functions out of $\boldsymbol{r}_f$, $\boldsymbol{R}_f$, $\boldsymbol{V}_f$, $\boldsymbol{v}_f$, $\boldsymbol{U}_f$, $\boldsymbol{u}_f$, $\boldsymbol{s}_f$ and $\boldsymbol{S}_f$* (*at most one of these functions has any given interval endpoint in its domain*).

The π-intervals together with the R-intervals and R^{-1}-intervals form a collection of loops, which we call $R^{\pm}$-π-loops.

σ-arcs and τ-arcs Recall from the previous sections that we view

$$\sigma_f, \tau_f : \mathfrak{I}_f^+ \to \mathfrak{I}_f^-.$$

We add an arc between each $0_{\mathfrak{i}}$ and $1_{\mathfrak{i}'}$ with $\sigma_f(\mathfrak{i}) = \mathfrak{i}'$ and between each $1_{\mathfrak{i}}$ and $0_{\mathfrak{i}'}$ with $\tau_f(\mathfrak{i}) = \mathfrak{i}'$. These arcs are called σ_f-arcs and τ_f-arcs, respectively. Any σ_f-arc (resp. τ_f-arc) is also called a σ-arc (resp. τ-arc). Notice even though an arc is formally the same as an interval, we distinguish these types of objects. The only arcs that exist are σ-arcs and τ-arcs. The purpose of this construction is that the conditions pertaining to σ_f and τ_f in (3-22) are equivalent to the fact that *two endpoints of intervals connected by a σ-arc or τ-arc are assigned the same value in $[n]$ by the relevant functions out of $\boldsymbol{a}$, $\boldsymbol{r}_f$, $\boldsymbol{R}_f$, $\boldsymbol{V}_f$, $\boldsymbol{v}_f$, $\boldsymbol{U}_f$, $\boldsymbol{u}_f$, $\boldsymbol{s}_f$ and $\boldsymbol{S}_f$.*

After adding these arcs, every endpoint of an interval has exactly one arc emanating from it. We have therefore now constructed a trivalent graph

$$G(\{\sigma_f, \tau_f, \pi_i\}).$$

Each vertex of the graph is an endpoint of two intervals and one arc. The number of vertices of this graph is twice the total number of w-intervals, R-intervals and R^{-1}-intervals, which is $2(|w| + 8(k+l))$. Therefore we have

$$\chi(G(\{\sigma_f, \tau_f, \pi_i\})) = -(|w| + 8(k+l)). \tag{4-2}$$

(For general g, we have $\chi(G(\{\sigma_f, \tau_f, \pi_i\})) = -(|w| + 4g(k+l))$.) Moreover, *the conditions in (3-22) are now interpreted purely in terms of the combinatorics of this graph.*

Gluing in discs There are two types of cycles in $G(\{\sigma_f, \tau_f, \pi_i\})$ that we wish to consider:

- Cycles that alternate between following either a w-intermediate-interval or a π-interval and then either a σ-arc or a τ-arc. These cycles are disjoint from one another, and every σ- or τ-arc is contained in exactly one such cycle. We call these cycles *type I cycles*. For every type I cycle, we glue a disc to $G(\{\sigma_f, \tau_f, \pi_i\})$ along its boundary, following the cycle. These discs will be called *type I discs*. (These are analogous to the o-discs of [Magee and Puder 2019].)

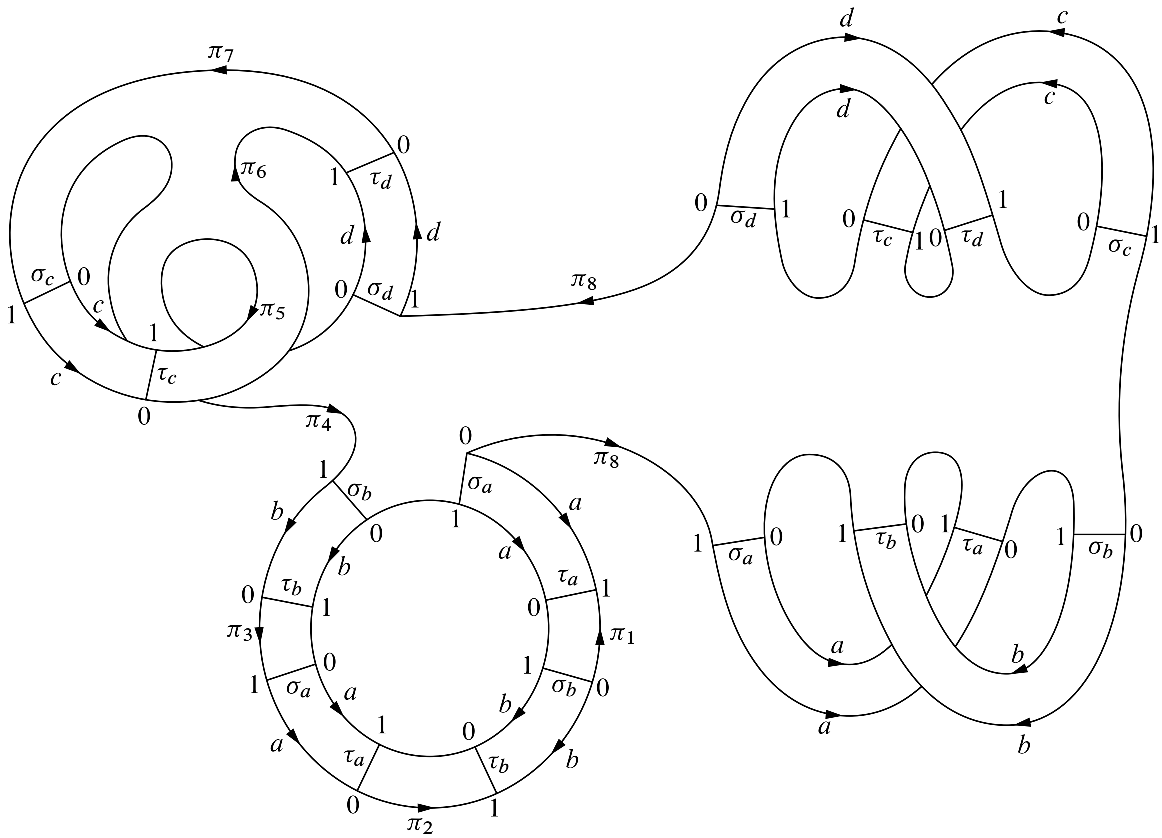


Figure 3: An example $\Sigma(\{\sigma_f, \tau_f, \pi_i\})$ for $w = ab^{-1}a^{-1}b$. The σ, τ and some of the π_i-arcs are labeled along with the numbers (0 or 1) of the points being matched in the w-intervals. Each w-interval is also labeled with its corresponding letter. Here $k = l = 1$; π_8 is a transposition and all other π_i are the identity. There is one resulting $R^{\pm}$-π-loop. In this example, for each $f \in \{a, b, c, d\}$, $\sigma_f = \tau_f$. This means that all type II discs are rectangles.

- Cycles that alternate between following either a w-interval, an R-interval or an R^{-1}-interval and then either a σ-arc or a τ-arc. Again, these cycles are disjoint, and every σ- or τ-arc is contained in exactly one such cycle. We call these cycles *type II cycles.* For every type II cycle, we glue a disc to $G(\{\sigma_f, \tau_f, \pi_i\})$ identifying the boundary of the disc with the cycle. These discs will be called *type II discs*. (These are similar to the z-discs of [Magee and Puder 2019].)

Because every interior of an interval meets exactly one of the glued-in discs, and every arc has two boundary segments of discs glued to it, the object resulting from gluing in these discs is a decorated topological surface, which we denote by

$$\Sigma(\{\sigma_f, \tau_f, \pi_i\}).$$

An example of this construction is depicted in Figure 3.

The boundary components of $\Sigma(\{\sigma_f,\tau_f,\pi_i\})$ consist of the w-loop and the $R^{\pm}$-π-loops. It is not hard to check that $\Sigma(\{\sigma_f,\tau_f,\pi_i\})$ is orientable with an orientation compatible with the fixed orientations of the boundary loops corresponding to traversing every w-interval or $R^{\pm1}$-interval from 0 to 1.

We view the given CW–complex structure and the assigned labelings and directions of the intervals that now pave $\partial\Sigma$ as part of the data of $\Sigma(\{\sigma_f,\tau_f,\pi_i\})$. The number of discs of $\Sigma(\{\sigma_f,\tau_f,\pi_i\})$ is connected to the quantities appearing in Theorem 3.6 as follows:

Lemma 4.1 $$\mathcal{N}(\{\sigma_f,\tau_f,\pi_i\}) = n^{\#\{\text{type I discs of }\Sigma(\{\sigma_f,\tau_f,\pi_i\})\}}.$$

Proof The constraints on the functions $\boldsymbol{a}$, $\boldsymbol{r}_f$, $\boldsymbol{R}_f$, $\boldsymbol{V}_f$, $\boldsymbol{v}_f$, $\boldsymbol{U}_f$, $\boldsymbol{u}_f$, $\boldsymbol{s}_f$ and $\boldsymbol{S}_f$ in (3-22) now correspond to the fact that, altogether, they assign the same value in $[n]$ to every interval endpoint in the same type I cycle, and there are no other constraints between them. □

The quantities $|\sigma_f\tau_f^{-1}|$ in (3-24) can also be related to $\Sigma(\{\sigma_f,\tau_f,\pi_i\})$ as follows:

Lemma 4.2 $$\prod_{f\in\{a,b,c,d\}} n^{-|\sigma_f\tau_f^{-1}|} = n^{-4(k+l)-|w|/2}n^{\#\{\text{type II discs of }\Sigma(\{\sigma_f,\tau_f,\pi_i\})\}}.$$

Proof Recalling the definition of $|\sigma_f\tau_f^{-1}|$ from Proposition 2.7, we can also write
$$|\sigma_f\tau_f^{-1}| = k+l+p_f-\#\{\text{cycles of }\sigma_f\tau_f^{-1}\}.$$
The cycles of $\{\sigma_f\tau_f^{-1} : f\in\{a,b,c,d\}\}$ are in one-to-one correspondence with the type II cycles of $\Sigma(\{\sigma_f,\tau_f,\pi_i\})$ and hence also the type II discs. Therefore,
$$\begin{aligned}\prod_{f\in\{a,b,c,d\}} n^{-|\sigma_f\tau_f^{-1}|} &= n^{-4(k+l)}n^{\sum_{f\in\{a,b,c,d\}}(-p_f+\#\{\text{cycles of }\sigma_f\tau_f^{-1}\})}\\ &= n^{-4(k+l)-|w|/2}n^{\#\{\text{type II discs of }\Sigma(\{\sigma_f,\tau_f,\pi_i\})\}}.\end{aligned}$$ □

We are now able to prove the following:

Theorem 4.3 *For $k+l>0$ with $\mu\vdash k$ and $\nu\vdash l$ and $w\in[\boldsymbol{F}_4,\boldsymbol{F}_4]$, we have*
$$\mathcal{I}_n(w,\mu,\nu) \ll_{w,k,l} \sum_{\{\sigma_f,\tau_f,\pi_i\}\in\mathsf{MATCH}(w,k,l)} n^{-\sum_{i=1}^{8}\|\pi_i\|_{k,l}}n^{\chi(\Sigma(\{\sigma_f,\tau_f,\pi_i\}))}.$$

Proof Combining Lemmas 4.1 and 4.2 with Corollary 3.7 gives
$$\mathcal{I}_n(w,\mu,\nu) \ll_{w,k,l} n^{-8k-8l-|w|} \sum_{\{\sigma_f,\tau_f,\pi_i\}\in\mathsf{MATCH}(w,k,l)} n^{-\sum_{i=1}^{8}\|\pi_i\|_{k,l}}n^{\#\{\text{discs of }\Sigma(\{\sigma_f,\tau_f,\pi_i\})\}}.$$
Then, from (4-2), we obtain
$$\begin{aligned}\mathcal{I}_n(w,\mu,\nu) &\ll_{w,k,l} \sum_{\{\sigma_f,\tau_f,\pi_i\}\in\mathsf{MATCH}(w,k,l)} n^{-\sum_{i=1}^{8}\|\pi_i\|_{k,l}}n^{\chi(G(\{\sigma_f,\tau_f,\pi_i\}))+\#\{\text{discs of }\Sigma(\{\sigma_f,\tau_f,\pi_i\})\}}\\ &= \sum_{\{\sigma_f,\tau_f,\pi_i\}\in\mathsf{MATCH}(w,k,l)} n^{-\sum_{i=1}^{8}\|\pi_i\|_{k,l}}n^{\chi(\Sigma(\{\sigma_f,\tau_f,\pi_i\}))}.\end{aligned}$$ □

4.2 Two simplifying surgeries

Theorem 4.3 suggests that we now bound

$$\chi(\Sigma(\{\sigma_f,\tau_f,\pi_i\}))-\sum_{i=1}^{8}\|\pi_i\|_{k,l}$$

for all $\{\sigma_f,\tau_f,\pi_i\}\in\mathsf{MATCH}(w,k,l)$. To do this, we make some observations that simplify the task. If C is a simple closed curve in a surface S, then *compressing S along C* means that we cut S along C and then glue discs to cap off any new boundary components created by the cut.

Suppose that we are given $\{\sigma_f,\tau_f,\pi_i\}\in\mathsf{MATCH}(w,k,l)$. Then $\{\sigma_f,\sigma_f,\pi_i\}$ is also in $\mathsf{MATCH}(w,k,l)$ (the *forbidden matching* property continues to hold). It is not hard to see that

$$\chi(\Sigma(\{\sigma_f,\sigma_f,\pi_i\}))\geq\chi(\Sigma(\{\sigma_f,\tau_f,\pi_i\})).$$

Indeed, the τ_f-arcs can be replaced by σ_f-parallel arcs inside the type II discs of $\Sigma(\{\sigma_f,\tau_f,\pi_i\})$. The resulting surface's arcs may not cut the surface into discs, but this can be fixed by (possibly repeatedly) compressing the surface along simple closed curves disjoint from the arcs, leaving the combinatorial data of the arcs unchanged but only potentially increasing the Euler characteristic.

It remains to deal with the sum $\sum_{i=1}^{8}\|\pi_i\|_{k,l}$.

Suppose again that an arbitrary $\{\sigma_f,\tau_f,\pi_i\}\in\mathsf{MATCH}(w,k,l)$ is given. For each $i\in[8]$, write

$$\pi_i=\pi_i^*\sigma_i,$$

where $\pi_i^*\in S_k\times S_l$, $\sigma_i=(\pi_i^*)^{-1}\pi_i\in S_{k+l}$ and $|\sigma_i|=\|\pi_i\|_{k,l}$. Let $X_0:=\Sigma(\{\sigma_f,\tau_f,\pi_i\})$.

Take $\Sigma(\{\sigma_f,\tau_f,\pi_i\})$ and add to it all the π_i^*-intervals that would have been added if π_i was replaced by π_i^* for each $i\in[8]$ in its construction. The resulting object X_1 is the decorated surface X_0 together with a collection of π_i^*-intervals with endpoints in the boundary of X_0, and interiors disjoint from X_0. This adds $8(k+l)$ edges to X_0 and hence

$$\chi(X_1)=\chi(\Sigma(\{\sigma_f,\tau_f,\pi_i\}))-8(k+l).$$

Now we consider all cycles that for any fixed $i\in[8]$, alternate between π_i-intervals and π_i^*-intervals. The number of these cycles is the total number of cycles of the permutations $\{(\pi_i^*)^{-1}\pi_i : i\in[8]\}$. On the other hand, the number of cycles of $(\pi_i^*)^{-1}\pi_i$ is

$$k+l-|(\pi_i^*)^{-1}\pi_i|=k+l-|\sigma_i|=k+l-\|\pi_i\|_{k,l}.$$

So in total there are $8(k+l)-\sum_i\|\pi_i\|_{k,l}$ of these cycles. For every such cycle, we glue a disc along its boundary to the cycle. The resulting object is denoted X_2. Now, X_2 is a topological surface, and we added $8(k+l)-\sum_i\|\pi_i\|_{k,l}$ discs to X_1 to form X_2, so

$$\chi(X_2)=\chi(X_1)+8(k+l)-\sum_i\|\pi_i\|_{k,l}=\chi(\Sigma(\{\sigma_f,\tau_f,\pi_i\}))-\sum_i\|\pi_i\|_{k,l}.$$

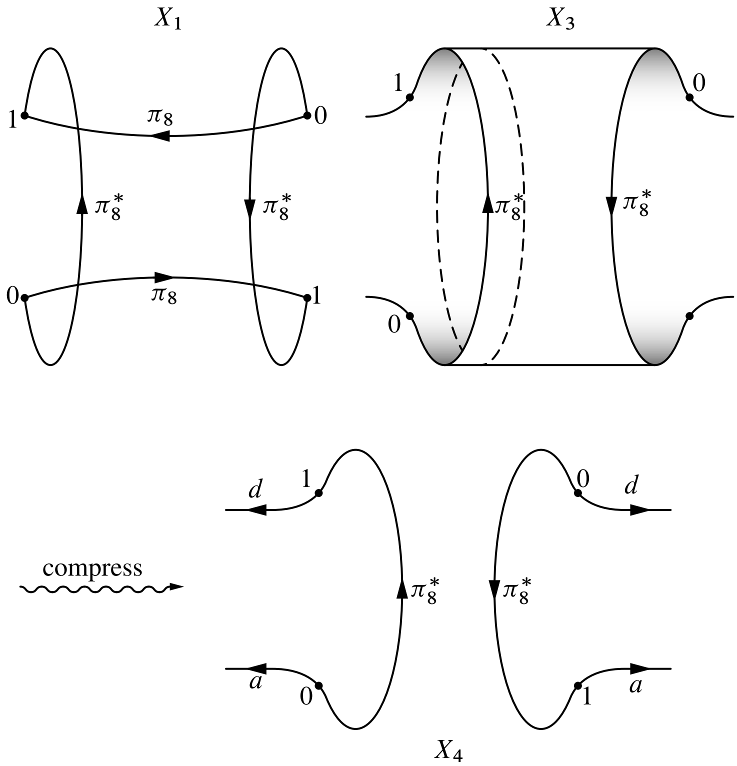


Figure 4: A local illustration of the second type of simplifying surgery, precisely in the context of Figure 3. The dashed simple closed curve in X_3 is disjoint from any arcs, and cutting along this curve and gluing in two discs yields X_4. Going back to Figure 3 again, the net effect of this surgery is to cut the left half from the right half.

Now "forget" all the original π_i-intervals from X_2 to form X_3. The surface X_3 is a decorated surface in the same sense as X_0, except the connected components of $X_3 - \{\text{arcs}\}$ may not be discs. Similarly to before, by sequentially compressing X_3 along nonnullhomotopic simple closed curves disjoint from arcs, if they exist, we obtain a new decorated surface X_4. See Figure 4 for an illustration of this surgery taking place. Moreover, and this is the main point, X_4 is the same as $\Sigma(\{\sigma_f, \tau_f, \pi_i^*\})$ in the sense that they are related by a decoration-respecting cellular homeomorphism. Compression can only increase the Euler characteristic, so we obtain

$$\chi(\Sigma(\{\sigma_f, \tau_f, \pi_i^*\})) \geq \chi(X_3) = \chi(X_2) = \chi(\Sigma(\{\sigma_f, \tau_f, \pi_i\})) - \sum_i \|\pi_i\|_{k,l}.$$

Combining these two arguments proves the following proposition:

Proposition 4.4 *For any given $\{\sigma_f, \tau_f, \pi_i\}$, there exist $\pi_i^* \in S_k \times S_l$ for $i \in [8]$ such that*

$$\chi(\Sigma(\{\sigma_f, \sigma_f, \pi_i^*\})) - \sum_{i=1}^{8} \|\pi_i^*\|_{k,l} = \chi(\Sigma(\{\sigma_f, \sigma_f, \pi_i^*\})) \geq \chi(\Sigma(\{\sigma_f, \tau_f, \pi_i\})) - \sum_{i=1}^{8} \|\pi_i\|_{k,l}.$$

This has the following immediate corollary when combined with Theorem 4.3. Let

$$\mathsf{MATCH}^*(w,k,l)$$

denote the subset of $\mathsf{MATCH}(w,k,l)$ consisting of $\{\sigma_f,\sigma_f,\pi_i\}$ (ie $\sigma_f=\tau_f$ for each $f\in\{a,b,c,d\}$) with $\pi_i\in S_k\times S_l$ for each $i\in[8]$.

Corollary 4.5 *For $k+l>0$ with $\mu\vdash k$ and $\nu\vdash l$ and $w\in[\boldsymbol{F}_4,\boldsymbol{F}_4]$, we have*

$$\mathcal{I}_n(w,\mu,\nu)\ll_{w,k,l} n^{\max_{\{\sigma_f,\sigma_f,\pi_i\}\in\mathsf{MATCH}^*(w,k,l)}\chi(\Sigma(\{\sigma_f,\sigma_f,\pi_i\}))}.$$

The benefit to having $\pi_i\in S_k\times S_l$ for $i\in[8]$ is the following. Suppose now that $\{\sigma_f,\sigma_f,\pi_i\}\in\mathsf{MATCH}^*(w,k,l)$. Recall that the boundary loops of $\Sigma(\{\sigma_f,\sigma_f,\pi_i\})$ consist of one w-loop and some number of $R^{\pm}$-π-loops. The condition that each $\pi_i\in S_k\times S_l$ means that no π-interval ever connects an endpoint of a R-interval with an endpoint of an R^{-1}-interval. So every boundary component of $\Sigma(\{\sigma_f,\sigma_f,\pi_i\})$ that is not the w-loop contains either only R-intervals or only R^{-1}-intervals, and, in fact, when following the boundary component and reading the directions and labels of the intervals according to traversing each from 0 to 1, reads out a positive power of R (in the former case of only R-intervals) or a negative power of R^{-1} (in the latter case of only R^{-1}-intervals). The sum of the positive powers of R in boundary loops is k, and the sum of the negative powers of R is $-l$. Knowing this boundary structure is extremely important for the arguments in the next sections.

4.3 A topological result that proves Theorem 3.1

Here, in the spirit of [Culler 1981], we explain another way to think about the surfaces $\Sigma(\{\sigma_f,\sigma_f,\pi_i\})$ for $\{\sigma_f,\sigma_f,\pi_i\}\in\mathsf{MATCH}^*(w,k,l)$ that is easier to work with than the construction we gave. At this point we also show how things work for general $g\geq 2$. An *arc* in a surface Σ is a properly embedded interval in Σ with endpoints in the boundary $\partial\Sigma$.

Definition 4.6 For $w\in\boldsymbol{F}_{2g}$, we define $\mathsf{surfaces}(w,k,l)$ to be the set of all decorated surfaces Σ^* as follows. A decorated surface $\Sigma^*\in\mathsf{surfaces}(w,k,l)$ is an oriented surface with boundary, with compatibly oriented boundary components, together with a collection of disjoint embedded arcs that cut Σ^* into topological discs. One boundary component is assigned to be a w-loop, and every other boundary component is assigned to be either a R-loop or an R^{-1}-loop. Each arc is assigned a transverse direction and a label in $\{a_1,b_1,\dots,a_g,b_g\}$. Every arc-endpoint in $\partial\Sigma^*$ inherits a transverse direction and label from the assigned direction and label of its arc. We require that Σ^* satisfy the following properties:

(P1) When one follows the w-loop according to its assigned orientation, and reads f when an f-labeled arc-endpoint is traversed in its given direction, and f^{-1} when an f-labeled arc-endpoint is traversed counter to its given direction, one reads a cyclic rotation of w in reduced form, depending on where one begins to read.

(P2) When one follows any R-loop according to its assigned orientation in the same way as before, one reads (a cyclic rotation) of some positive power of R_g in reduced form. The sum of these positive powers over all R-loops is k.

(P3) When one follows any R^{-1}-loop according to its assigned orientation in the same way as before, one reads (a cyclic rotation) of some negative power of R_g in reduced form. The sum of these negative powers over all R^{-1}-loops is $-l$.

(P4) No arc connects an R-loop to an R^{-1}-loop.

Given a surface $\Sigma(\{\sigma_f, \sigma_f, \pi_i\})$ with $\{\sigma_f, \sigma_f, \pi_i\} \in \mathsf{MATCH}^*(w,k,l)$, all the type II discs of the surface are rectangles. Hence, by collapsing each w-interval, R-interval and R^{-1}-interval to a point, and collapsing every type II rectangle to an arc, we obtain a CW–complex that is a surface with boundary, cut into discs by arcs. Every arc inherits a transverse direction and label from the compatible assigned directions and labels of the intervals in the boundary of its originating type II rectangle. We call this modified surface $\Sigma^* = \Sigma^*(\{\sigma_f, \pi_i\})$. It clearly satisfies (P1)–(P3) and (P4) follows from the *forbidden matchings* property. (Of course, when $g = 2$, we identify $\{a,b,c,d\}$ with $\{a_1,b_1,a_2,b_2\}$.) We also have $\chi(\Sigma(\{\sigma_f, \sigma_f, \pi_i\})) = \chi(\Sigma^*(\{\sigma_f, \pi_i\}))$. With Definition 4.6 and the remarks proceeding it, we can now state a further consequence of Corollary 4.5 as it extends to general $g \geq 2$:

Corollary 4.7 *For $k+l > 0$ with $\mu \vdash k$ and $\nu \vdash l$ and $w \in [\boldsymbol{F}_{2g}, \boldsymbol{F}_{2g}]$, as $n \to \infty$,*

$$\mathcal{I}_n(w,\mu,\nu) \ll_{w,k,l} n^{\max\{\chi(\Sigma^*):\Sigma^* \in \mathsf{surfaces}(w,k,l)\}}.$$

In order for Corollary 4.7 to give us strong enough results, it needs to be combined with the following nontrivial topological bound:

Proposition 4.8 *If $w \in [\boldsymbol{F}_{2g}, \boldsymbol{F}_{2g}]$ is a shortest element representing the conjugacy class of $\gamma \in \Gamma_g$, $w \neq \mathrm{id}$ and $\Sigma^* \in \mathsf{surfaces}(w,k,l)$, then $\chi(\Sigma^*) \leq -(k+l)$.*

Remark 4.9 Proposition 4.8 is by no means a trivial statement and one has to use that w is a shortest element representing the conjugacy class of some element of Γ_g. For example, if $w = R_g$, then w represents the conjugacy class of id_{Γ_g}, but for $k = 0$ and $l = 1$ there is an "obvious" annulus in $\mathsf{surfaces}(w,0,1)$. This has $\chi = 0 > -(k+l) = -1$. Proposition 4.8 also requires $w \neq \mathrm{id}$; if $w = \mathrm{id}$ then for $k = 0$ and $l = 1$ one can take a disc with no arcs as a valid element of $\mathsf{surfaces}(\mathrm{id},0,0)$. This has $\chi = 1 > -(k+l) = 0$. In fact this disc is ultimately responsible for $\mathbb{E}_{g,n}[\mathrm{Tr}_{\mathrm{id}}] = n$.

The proof of Proposition 4.8 is self-contained and given in Section 4.5. Before doing this, we prove Theorem 3.1.

Proof of Theorem 3.1 given Proposition 4.8 Since Theorem 3.1 was proved when $k = l = 0$ in Section 3.2, we can assume $k + l > 0$. Then combining Corollary 4.7 and Proposition 4.8 gives

$$\mathcal{J}_n(w, \mu, \nu) \ll_{w,k,l} n^{-(k+l)}.$$

On the other hand, $D_{\mu,\nu}(n) = O(n^{k+l})$ from (2-1). Therefore, $D_{\mu,\nu}(n)\mathcal{J}_n(w, \mu, \nu) \ll_{w,k,l} 1$. □

4.4 Work of Dehn and Birman–Series

As we mentioned in Section 3.1, to prove Proposition 4.8 we have to use the fact that $w \in [\boldsymbol{F}_{2g}, \boldsymbol{F}_{2g}]$ is a shortest element representing the conjugacy class of $\gamma \in \Gamma_g$. We use a combinatorial characterization of such words that stems from Dehn's algorithm [1912] for solving the problem of whether a given word represents the identity in Γ_g. The ideas of Dehn's algorithm were refined in [Birman and Series 1987]. Magee and Puder [2023] used Birman and Series' results (alongside other methods) to obtain the analog of Theorem 1.2 when the family of groups $\mathsf{SU}(n)$ is replaced by the family of symmetric groups S_n. Similar consequences of the work of Dehn, Birman and Series that we used in [loc. cit.] will be used here.

We now follow the language of [Magee and Puder 2023] to state the results we need in this paper. These results are simple and direct consequences of the work of Birman and Series.

We view the universal cover of Σ_g as a disc tiled by $4g$-gons that we call U. We assume every edge of this tiling is directed and labeled by some element of $\{a_1, b_1, \dots, a_g, b_g\}$ such that when we read counterclockwise along the boundary of any octagon we read the reduced cyclic word $[a_1, b_1] \cdots [a_g, b_g]$. By fixing a basepoint $u \in U$, we obtain a free cellular action of Γ_g on U that respects the labels and directions of edges and identifies the quotient $\Gamma_g \backslash U$ with Σ_g; this gives a description of Σ_g as a $4g$-gon with glued sides, as is typical.

Now suppose that $\gamma \in \Gamma$ is not the identity. The quotient $A_\gamma := \langle \gamma \rangle \backslash U$ of U by the cyclic group generated by γ is an open annulus tiled by infinitely many $4g$-gons. The edges of A_γ inherit directions and labels from those of the edges of U. The point $u \in U$ maps to some point, denoted by $x_0 \in A_\gamma$.

Now let $w \in \boldsymbol{F}_{2g}$ be an element that represents γ, and identify w with a combinatorial word by writing w in reduced form. Beginning at x_0, and following the path spelled out by w beginning at x_0, we obtain an oriented closed loop L_w in the one-skeleton of A_γ. If w is a shortest element representing the conjugacy class of γ, then this loop L_w must not have self-intersections. In this case, which we from now assume, L_w is therefore a topologically embedded circle in the annulus A_γ that is nonnullhomotopic and cuts A_γ into two annuli $A_\gamma^\pm$.

Every vertex of A_γ has $4g$ incident half-edges each of which has an orientation and direction given by the edge they are in. Going clockwise, the cyclic order of the half-edges incident at any vertex is:

a_1-outgoing, b_1-incoming, a_1-incoming, b_1-outgoing, $\dots$, a_g-outgoing, b_g-incoming, a_g-incoming, b_g-outgoing.

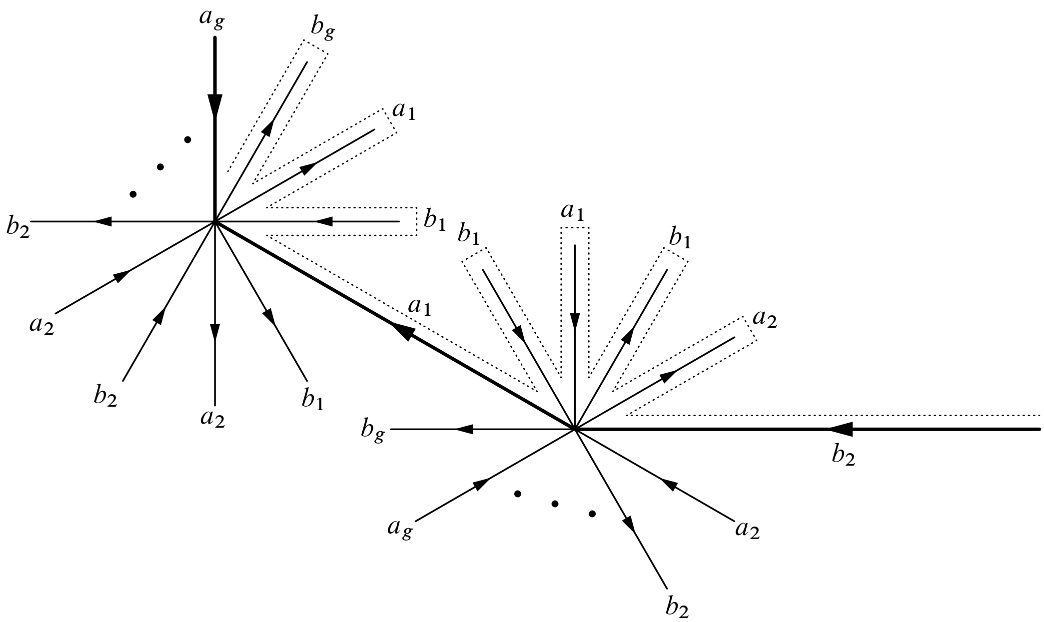


Figure 5: A piece P of $\widehat{L}_w$ in the case when the reduced form of w contains $a_g a_1^{-1} b_2^{-1}$ as a subword. The edges of L_w are in bold. The piece is indicated by the dotted lines. This piece P has $\mathfrak{e}(P) = 2$, $\mathfrak{he}(P) = 7$ and $\chi(P) = 1$. Note that a piece may also run along the other side of L_w.

We define $\widehat{L}_w$ to be the loop L_w with all incident half edges in A_γ attached. We call the new half-edges added *hanging half-edges.*

Moreover, we thicken up $\widehat{L}_w$ by viewing each edge of L_w as a rectangle, each hanging half-edge as a half-rectangle, and each vertex replaced by a disc. In other words, we take a small neighborhood of $\widehat{L}_w$ in A_γ. We now think of $\widehat{L}_w$ as the thickened version. This is a topological annulus, where the hanging half-edges have become stubs hanging off. A *piece of* $\widehat{L}_w$ is a contiguous collection of hanging half-rectangles and rectangle sides following edges of L_w in the boundary of $\widehat{L}_w$. Such a piece is in either A_γ^+ or A_γ^-. Given a piece P of $\widehat{L}_w$ we write $\mathfrak{e}(P)$ for the number of rectangle sides following edges of L_w, and $\mathfrak{he}(P)$ for the number of hanging-half edges in P. We say that a piece P has Euler characteristic $\chi(P) = 0$ if it follows an entire boundary component of $\widehat{L}_w$, and $\chi(P) = 1$ otherwise, as we view it as an interval running along the rectangle sides and around the sides of the hanging half-rectangles. See Figure 5 for an illustration of a piece of $\widehat{L}_w$.

Birman and Series [1987, Theorem 2.12(a)] prove that, if w is a shortest element representing the conjugacy class of $\gamma \in \Gamma_g$, then there are strong restrictions on the pieces of $\widehat{L}_w$ that can appear. This has the following consequence, which is given by[3] [Magee and Puder 2023, Proof of Lemma 5.18]:

[3]We stress that Lemma 4.10 is a straightforward consequence of Birman and Series' work, so, even though we cite [Magee and Puder 2023], this paper does not depend on that work in any significant way.

Lemma 4.10 *If w is a shortest element representing the conjugacy class of $\gamma \in \Gamma_g$, and both γ and hence w are nonidentity, then for any piece P of $\widehat{L}_w$, we have*

$$\mathfrak{e}(P) \le (2g-1)\mathfrak{he}(P) + 2g\chi(P).$$

Proof Since w is a shortest element representing some nonidentity conjugacy class in Γ_g, in the language of [Magee and Puder 2023], L_w is a boundary reduced tiled surface. Then the proof of [Magee and Puder 2023, Lemma 5.18] contains the result stated in the lemma. The basic idea of the proof is not complicated and goes back to [Dehn 1912]: if there are too many edges (ie $\mathfrak{e}(P)$ is large) then one can find a string of letters in the reduced word of w (eg $aba^{-1}b^{-1}c$) that can be shortened using the relator R (eg $aba^{-1}b^{-1}c = dcd^{-1}$). □

This inequality plays a crucial role in the next section.

4.5 Proof of Proposition 4.8

Suppose that $g \ge 2$ and $w \in [\boldsymbol{F}_{2g}, \boldsymbol{F}_{2g}]$ is a nonidentity shortest element representing the conjugacy class of $\gamma \in \Gamma_g$. In particular, w is cyclically reduced. We let $R = R_g$. Now fix $k, l \in \mathbb{N}_0$ and suppose $\Sigma^* \in \mathsf{surfaces}(w, k, l)$. The arcs of Σ^* are of three different types:

(WR) An arc with one endpoint in the w-loop and one endpoint in an R- or R^{-1}-loop.

(RR) An arc with both endpoints in R- or R^{-1}-loops. By property (P4), the endpoints of such an arc are both in R-loops or both in R^{-1}-loops.

(WW) An arc with both endpoints in the w-loop.

The boundary of any disc of Σ^* alternates between segments of $\partial\Sigma^*$ and arcs. A disc is a *pre-piece disc* if its boundary contains exactly one segment of the w-loop. A disc is called a *junction disc* if it is not a pre-piece disc. We say that a junction disc is *piece-adjacent* if it meets a WR-arc-side.

To be precise, we view all discs as open discs, and hence not containing any arcs. A disc meets certain arc-sides along its boundary; it is possible for a disc to meet both sides of the same arc and we view this scenario as the disc meeting two separate arc-sides. We say an arc-side has the same type WR/RR/WW as its corresponding arc.

Note that any pre-piece disc cannot meet any WW-arc-side: if it did, the disc could only meet this one arc-side together with one segment of the w-loop and this would contradict the fact that w is cyclically reduced since the arc matches a letter f with a cyclically adjacent letter f^{-1} of w. It is also clear that any pre-piece disc meets exactly two WR-arc-sides: the ones that emanate from the sole segment of the w-loop. So, in light of (P4), a pre-piece disc takes one of the forms shown in Figure 6.

We define a *piece of* Σ^* to be a connected component of

$$\{\text{pre-piece discs}\} \cup \{\text{WR-arcs}\}.$$

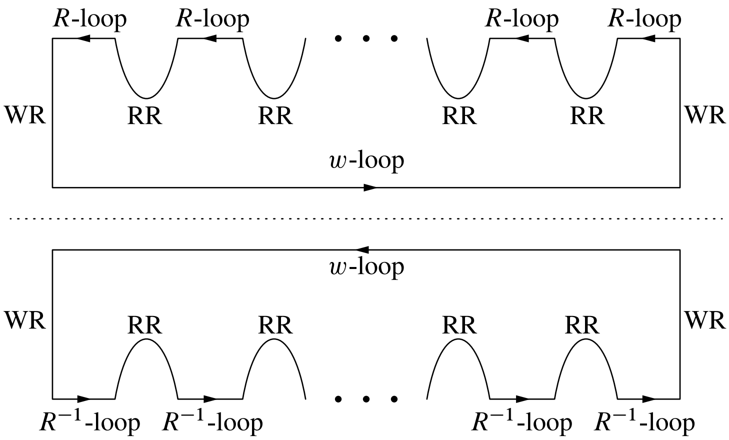


Figure 6: Possible forms of pre-piece discs. The number of R-loop segments or R^{-1}-loop segments is at least 1 and bounded given k and l. The arrows denote the orientations of the boundary loops.

A piece of Σ^* is therefore either a contiguous collection of pre-piece discs that meet only along WR-arcs, or a single WR-arc. If P is a piece of Σ^*, either $\chi(P) = 1$, or $\chi(P) = 0$, in which case P meets the entire w-loop and is the unique piece.

We now have *two* definitions of pieces: pieces of $\widehat{L}_w$ and pieces of Σ^*. These are, as the names suggest, closely related, and this is the key observation in the proof of Proposition 4.8. Indeed, the reader should carefully consider Figure 7, which leads to the following lemma. In analogy to pieces of $\widehat{L}_w$, if P is any piece of Σ^*, we write $\mathfrak{e}(P)$ for the number of WR-arcs in P, and $\mathfrak{he}(P)$ for the number of RR-arc-sides that meet P (this is zero if P is a single WR-arc).

Lemma 4.11 *If w is a shortest element representing the conjugacy class of $\gamma \in \Gamma_g$, $k, l \in \mathbb{N}_0$ and $\Sigma^* \in \mathsf{surfaces}(w, k, l)$, then, for any piece P of Σ^*, we have*

$$\mathfrak{e}(P) \le (2g-1)\mathfrak{he}(P) + 2g\chi(P).$$

Proof Given any piece P of Σ^*, it contains a consecutive (possibly cyclic) series of WR-arcs that correspond to a contiguous collection of edges in the loop L_w. The discs of P correspond to certain vertices of L_w; each of these vertices has two emanating half-edges belonging to the edges defined by WR-arcs of P. The piece P can either meet only R-loops or meet only R^{-1}-loops.

We define a piece P' of $\widehat{L}_w$ corresponding to P as follows. If P meets R-loops, then P' consists of rectangle sides along the edges of L_w corresponding to the WR-arcs of P together with all hanging half-edges at vertices corresponding to discs of P that are on the *left* of L_w as it is traversed in its assigned orientation (corresponding to reading w along L_w). If P' meets R^{-1}-loops, then P' is defined similarly with the modification that we include instead hanging half-edges on the *right* of L_w. Figure 7 together

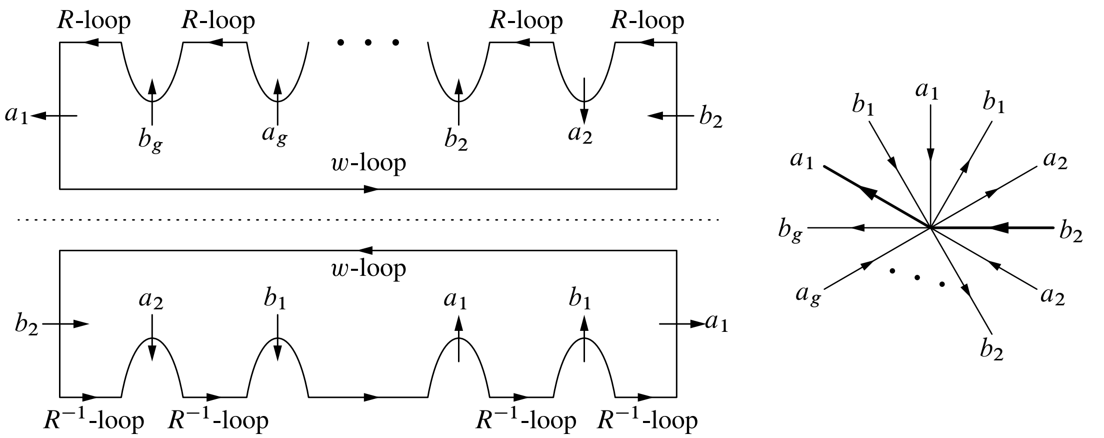


Figure 7: Given a segment of the w-loop corresponding to a juncture between letters $a_1^{-1}b_2^{-1}$ in w, if this segment is part of a pre-piece disc then some possible forms of that disc are shown above. This juncture between letters of w corresponds to a vertex in L_w. The right-hand illustration shows the neighborhood of this vertex in the annulus A_γ, where the bold lines correspond to half-edges of L_w. The right-hand picture actually almost determines the left-hand pictures. Indeed, given the a_1-arc on the top-left, the next arc has to be a b_g-arc with the given direction, since *only* b_g^{-1} cyclically precedes a_1 in R_g or any power of R_g. Then the next arc a_g with its direction is determined since *only* a_g cyclically precedes b_g in R_g. This continues until an arc labeled by b_2 and with an incoming direction is reached, as in the right arc of the top-left picture. At this point, the boundary of the disc may close up. (This is analogous to what happens in the bottom picture, where an analogous pattern occurs.) The only indeterminacy is that after reaching a b_2-arc with an incoming direction for the first time, the *entire pattern* shown in the right-hand picture may repeat any number of times, as long as k and l allow it. The upshot of this is that any pre-piece disc has at least as many incident RR-arc-sides as there are hanging half-edges on the corresponding side of L_w, at the corresponding vertex.

with its captioned discussion now shows that

$$\mathfrak{he}(P') \le \mathfrak{he}(P),$$

and $\mathfrak{e}(P) = \mathfrak{e}(P')$ by construction. We also have $\chi(P') = \chi(P)$. Therefore Lemma 4.10 applied to P' implies

$$\mathfrak{e}(P) = \mathfrak{e}(P') \le (2g-1)\mathfrak{he}(P') + 2g\chi(P') \le (2g-1)\mathfrak{he}(P) + 2g\chi(P). \qquad \square$$

Let N_{RR} be the number of RR-arcs, N_{WR} the number of WR-arcs, and N_{WW} the number of WW-arcs in Σ^*. In the following, we refer to discs of Σ^* simply as discs. Since there are $4g(k+l)$ incidences between arcs and R-loops or R^{-1}-loops, we have

$$2N_{\mathrm{RR}} + N_{\mathrm{WR}} = 4g(k+l). \tag{4-3}$$

Let Σ_1 be the surface formed by cutting Σ^* along all RR-arcs. We have

$$\chi(\Sigma_1) = \sum_{\text{discs } D} \bigl(1 - \tfrac{1}{2}d'(D)\bigr),$$

where $d'(D)$ is the number of arc-sides meeting D that are *not* of type RR. This formula holds because $d'(D)$ is the degree of the disc D in the dual graph G_1 of Σ_1, the right-hand side is easily seen to be $\chi(G_1) = V(G_1) - E(G_1)$, and, since Σ_1 deformation retracts to an obvious embedded copy of G_1, $\chi(G_1) = \chi(\Sigma_1)$. We partition the sum above according to $\chi(\Sigma_1) = S_0 + S_1 + S_2$, where

$$\begin{aligned}
S_0 &:= \sum_{\text{pre-piece discs } D} \left(1 - \tfrac{1}{2} d'(D)\right),\\
S_1 &:= \sum_{\text{piece-adjacent junction discs } D} \left(1 - \tfrac{1}{2} d'(D)\right),\\
S_2 &:= \sum_{\text{not piece-adjacent junction discs } D} \left(1 - \tfrac{1}{2} d'(D)\right).
\end{aligned}$$

Note first that a pre-piece disc has $d'(D) = 2$ (see Figure 6). Hence $S_0 = 0$. We deal with S_1 next. For a disc D of Σ^*, let $d_{\mathrm{WR}}(D)$ denote the number of WR-arc-sides meeting D. Note that a piece-adjacent junction disc D has $d_{\mathrm{WR}}(D) > 0$ by definition. We rewrite S_1 as

$$\begin{aligned}
S_1 &= \sum_{\text{piece-adjacent junction discs } D} \left(1 - \tfrac{1}{2} d'(D)\right) \frac{1}{d_{\mathrm{WR}}(D)} \sum_{\text{incidences between } D \text{ and WR-arc-sides}} 1 \\
&= \sum_{\text{pieces } P} \; \sum_{\text{incidences between } P \text{ and some junction disc } D \text{ along WR-arc}} Q(D),
\end{aligned} \tag{4-4}$$

where, for a piece-adjacent junction disc D,

$$Q(D) := \frac{1}{d_{\mathrm{WR}}(D)} \left(1 - \tfrac{1}{2} d'(D)\right).$$

Suppose that D is a piece-adjacent junction disc. By parity considerations, $d_{\mathrm{WR}}(D)$ is even. We estimate $Q(D)$ by splitting into two cases. If $d_{\mathrm{WR}}(D) = 2$ then $d'(D) \geq 3$, since, otherwise, D would meet only two WR-arc-sides and other RR-arc-sides, whence be a pre-piece disc and not be a junction disc. In this case,

$$Q(D) = \tfrac{1}{2}\left(1 - \tfrac{1}{2} d'(D)\right) \leq \tfrac{1}{2}\left(1 - \tfrac{3}{2}\right) = -\tfrac{1}{4}.$$

Otherwise, $d_{\mathrm{WR}}(D) \geq 4$ and, since $d'(D) \geq d_{\mathrm{WR}}(D)$, we have

$$Q(D) \leq \frac{1}{d_{\mathrm{WR}}(D)} \left(1 - \tfrac{1}{2} d_{\mathrm{WR}}(D)\right) = \frac{1}{d_{\mathrm{WR}}(D)} - \tfrac{1}{2} \leq \tfrac{1}{4} - \tfrac{1}{2} = -\tfrac{1}{4}.$$

So we have proved that, for all piece-adjacent junction discs D, $Q(D) \leq -\tfrac{1}{4}$. Putting this into (4-4) gives

$$\begin{aligned}
S_1 &\leq -\frac{1}{4} \sum_{\text{pieces } P} \; \sum_{\text{incidences between } P \text{ and some junction disc } D \text{ along WR-arc}} 1 \\
&= -\frac{1}{4} \sum_{\text{pieces } P} 2\chi(P) = -\frac{1}{2} \sum_{\text{pieces } P} \chi(P).
\end{aligned} \tag{4-5}$$

We now turn to S_2. *Here is the key moment where* $w \neq \mathrm{id}$ *is used.*[4] Since $w \neq \mathrm{id}$, any disc must meet an arc. Indeed, the only other possibility is that the boundary of the disc is an entire boundary loop that has no emanating arcs. This hypothetical boundary loop cannot be an R- or R^{-1}-loop, so it has to be the w-loop. But this would entail $w = \mathrm{id}$.

Hence any disc contributing to S_2 meets no WR-arc-side, but meets some arc-side. Therefore it meets only WW-arcs or only RR-arcs. Every disc D contributing to S_2 meeting only WW-arcs gives a nonpositive contribution since w is cyclically reduced, and hence $d'(D) \geq 2$. Every disc D contributing to S_2 meeting only RR-arcs, which we will call an *RR-disc*, has $d'(D) = 0$ and hence contributes 1 to S_2.

This shows

$$S_2 \leq \#\{\text{RR-discs}\}. \tag{4-6}$$

In total, combining $S_0 = 0$ with (4-5) and (4-6), we get

$$\chi(\Sigma_1) \leq \#\{\text{RR-discs}\} - \frac{1}{2} \sum_{\text{pieces } P \text{ of } \Sigma^*} \chi(P).$$

To obtain Σ^* from Σ_1 we have to glue all cut RR-arcs, of which there are N_{RR}. Each gluing decreases χ by 1, so

$$\chi(\Sigma^*) \leq \#\{\text{RR-discs}\} - N_{\mathrm{RR}} - \frac{1}{2} \sum_{\text{pieces } P \text{ of } \Sigma^*} \chi(P). \tag{4-7}$$

Using Lemma 4.11 with the above gives

$$\begin{aligned} \chi(\Sigma^*) &\leq \#\{\text{RR-discs}\} - N_{\mathrm{RR}} - \frac{1}{2} \sum_{\text{pieces } P \text{ of } \Sigma^*} \chi(P) \\ &\leq \#\{\text{RR-discs}\} - N_{\mathrm{RR}} - \frac{1}{4g} \sum_{\text{pieces } P \text{ of } \Sigma^*} \mathfrak{e}(P) + \frac{2g-1}{4g} \sum_{\text{pieces } P \text{ of } \Sigma^*} \mathfrak{he}(P) \\ &= \#\{\text{RR-discs}\} - N_{\mathrm{RR}} - \frac{N_{\mathrm{WR}}}{4g} + \frac{2g-1}{4g} \sum_{\text{pieces } P \text{ of } \Sigma^*} \mathfrak{he}(P). \end{aligned} \tag{4-8}$$

Let $\mathfrak{he}'(\Sigma^*)$ denote the total number of RR-arc-sides meeting RR-discs. Every RR-disc has to meet at least $4g$ arc-sides; this observation is similar to the reasoning in Figure 7. Therefore

$$\mathfrak{he}'(\Sigma^*) \geq 4g\#\{\text{RR-discs}\}. \tag{4-9}$$

Every RR-arc-side either meets a piece P and contributes to $\mathfrak{he}(P)$ or a disc meeting only RR-arc-sides and contributes to $\mathfrak{he}'(\Sigma^*)$. Hence

$$\mathfrak{he}'(\Sigma^*) + \sum_{\text{pieces } P \text{ of } \Sigma^*} \mathfrak{he}(P) = 2N_{\mathrm{RR}}. \tag{4-10}$$

[4] Although, technically, $w \neq \mathrm{id}$ was used to define L_w and pieces etc, if w is the identity, the proof of Proposition 4.8 could, a priori, circumvent these definitions.

Combining (4-3), (4-9) and (4-10) with (4-8) gives

$$\begin{aligned}
\chi(\Sigma^*) &\le \frac{\mathfrak{he}'(\Sigma^*)}{4g} - N_{\mathrm{RR}} - \frac{N_{\mathrm{WR}}}{4g} + \frac{2g-1}{4g} \sum_{\text{pieces } P \text{ of } \Sigma^*} \mathfrak{he}(P) && \text{(by (4-9))} \\
&= \frac{\mathfrak{he}'(\Sigma^*)}{4g} - N_{\mathrm{RR}} - \frac{N_{\mathrm{WR}}}{4g} + \frac{(2g-1)N_{\mathrm{RR}}}{2g} - \frac{2g-1}{4g}\mathfrak{he}'(\Sigma^*) && \text{(by (4-10))} \\
&= -\frac{1}{4g}(2N_{\mathrm{RR}} + N_{\mathrm{WR}}) - \frac{2g-2}{4g}\mathfrak{he}'(\Sigma^*) \\
&\le -\frac{1}{4g}(2N_{\mathrm{RR}} + N_{\mathrm{WR}}) = -\frac{4g(k+l)}{4g} && \text{(by (4-3))} \\
&= -(k+l).
\end{aligned}$$

This completes the proof of Proposition 4.8. □

5 Proof of the main theorem

Proof of Theorem 1.2 Assume $\gamma \in [\Gamma_g, \Gamma_g]$ is not the identity and that $w \in [\boldsymbol{F}_{2g}, \boldsymbol{F}_{2g}]$ is a shortest element representing the conjugacy class of γ, hence also not the identity. By Corollary 2.10, we have

$$\mathbb{E}_{g,n}[\mathrm{Tr}_\gamma] = \zeta(2g-2;n)^{-1} \sum_{(\mu,\nu)\in\widetilde{\Omega}} D_{\mu,\nu}(n)\mathcal{J}_n(w,\nu,\mu) + O_{w,g}\Big(\frac{1}{n}\Big),$$

where $\widetilde{\Omega}$ is a finite collection of pairs of Young diagrams. We know $\lim_{n\to\infty}\zeta(2g-2;n) = 1$ from (2-11) and, for each fixed (μ,ν), $D_{\mu,\nu}(n)\mathcal{J}_n(w,\nu,\mu) = D_{\nu,\mu}(n)\mathcal{J}_n(w,\nu,\mu) = O_{w,\mu,\nu}(1)$ by Theorem 3.1. Hence $\mathbb{E}_{g,n}[\mathrm{Tr}_\gamma] = O_\gamma(1)$ as $n \to \infty$, as required. □

References

[Atiyah and Bott 1983] **M F Atiyah**, **R Bott**, *The Yang–Mills equations over Riemann surfaces*, Philos. Trans. Roy. Soc. A 308 (1983) 523–615 MR Zbl

[Benkart et al. 1994] **G Benkart**, **M Chakrabarti**, **T Halverson**, **R Leduc**, **C Lee**, **J Stroomer**, *Tensor product representations of general linear groups and their connections with Brauer algebras*, J. Algebra 166 (1994) 529–567 MR Zbl

[Birman and Series 1987] **J S Birman**, **C Series**, *Dehn's algorithm revisited, with applications to simple curves on surfaces*, from "Combinatorial group theory and topology" (S M Gersten, J R Stallings, editors), Ann. of Math. Stud. 111, Princeton Univ. Press (1987) 451–478 MR Zbl

[Collins 2003] **B Collins**, *Moments and cumulants of polynomial random variables on unitary groups, the Itzykson–Zuber integral, and free probability*, Int. Math. Res. Not. 2003 (2003) 953–982 MR Zbl

[Collins and Male 2014] **B Collins**, **C Male**, *The strong asymptotic freeness of Haar and deterministic matrices*, Ann. Sci. École Norm. Sup. 47 (2014) 147–163 MR Zbl

[Collins and Śniady 2006] **B Collins**, **P Śniady**, *Integration with respect to the Haar measure on unitary, orthogonal and symplectic group*, Comm. Math. Phys. 264 (2006) 773–795 MR Zbl

[Culler 1981] **M Culler**, *Using surfaces to solve equations in free groups*, Topology 20 (1981) 133–145 MR Zbl

[Dahlqvist and Lemoine 2023] **A Dahlqvist**, **T Lemoine**, *Large N limit of Yang–Mills partition function and Wilson loops on compact surfaces*, Probab. Math. Phys. 4 (2023) 849–890 MR Zbl

[Dehn 1912] **M Dehn**, *Transformation der Kurven auf zweiseitigen Flächen*, Math. Ann. 72 (1912) 413–421 MR Zbl

[Enomoto and Izumi 2016] **T Enomoto**, **M Izumi**, *Indecomposable characters of infinite dimensional groups associated with operator algebras*, J. Math. Soc. Japan 68 (2016) 1231–1270 MR Zbl

[Fulton and Harris 1991] **W Fulton**, **J Harris**, *Representation theory: a first course*, Graduate Texts in Math. 129, Springer (1991) MR Zbl

[Goldman 1984] **W M Goldman**, *The symplectic nature of fundamental groups of surfaces*, Adv. Math. 54 (1984) 200–225 MR Zbl

[Gromov and Milman 1983] **M Gromov**, **V D Milman**, *A topological application of the isoperimetric inequality*, Amer. J. Math. 105 (1983) 843–854 MR Zbl

[Guralnick et al. 2012] **R Guralnick**, **M Larsen**, **C Manack**, *Low degree representations of simple Lie groups*, Proc. Amer. Math. Soc. 140 (2012) 1823–1834 MR Zbl

[Haagerup and Thorbjørnsen 2005] **U Haagerup**, **S Thorbjørnsen**, *A new application of random matrices: $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ is not a group*, Ann. of Math. 162 (2005) 711–775 MR Zbl

[Häsä and Stasinski 2019] **J Häsä**, **A Stasinski**, *Representation growth of compact linear groups*, Trans. Amer. Math. Soc. 372 (2019) 925–980 MR Zbl

[Koike 1989] **K Koike**, *On the decomposition of tensor products of the representations of the classical groups: by means of the universal characters*, Adv. Math. 74 (1989) 57–86 MR Zbl

[Larsen and Lubotzky 2008] **M Larsen**, **A Lubotzky**, *Representation growth of linear groups*, J. Eur. Math. Soc. 10 (2008) 351–390 MR Zbl

[Lemoine 2022] **T Lemoine**, *Large N behaviour of the two-dimensional Yang–Mills partition function*, Combin. Probab. Comput. 31 (2022) 144–165 MR Zbl

[Magee 2022] **M Magee**, *Random unitary representations of surface groups, I: Asymptotic expansions*, Comm. Math. Phys. 391 (2022) 119–171 MR Zbl

[Magee and Puder 2015] **M Magee**, **D Puder**, *Word measures on unitary groups*, preprint (2015) arXiv 1509.07374

[Magee and Puder 2019] **M Magee**, **D Puder**, *Matrix group integrals, surfaces, and mapping class groups, I: $\mathcal{U}(n)$*, Invent. Math. 218 (2019) 341–411 MR Zbl

[Magee and Puder 2023] **M Magee**, **D Puder**, *The asymptotic statistics of random covering surfaces*, Forum Math. Pi 11 (2023) art. id. e15 MR Zbl

[Mirzakhani 2007] **M Mirzakhani**, *Simple geodesics and Weil–Petersson volumes of moduli spaces of bordered Riemann surfaces*, Invent. Math. 167 (2007) 179–222 MR Zbl

[Nica 1994] **A Nica**, *On the number of cycles of given length of a free word in several random permutations*, Random Structures Algorithms 5 (1994) 703–730 MR Zbl

[Voiculescu 1985] **D Voiculescu**, *Symmetries of some reduced free product C^*-algebras*, from "Operator algebras and their connections with topology and ergodic theory" (H Araki, C C Moore, Ş Strătilă, D Voiculescu, editors), Lecture Notes in Math. 1132, Springer (1985) 556–588 MR Zbl

[Voiculescu 1986] **D Voiculescu**, *Addition of certain noncommuting random variables*, J. Funct. Anal. 66 (1986) 323–346 MR Zbl

[Voiculescu 1987] **D Voiculescu**, *Multiplication of certain noncommuting random variables*, J. Operator Theory 18 (1987) 223–235 MR Zbl

[Voiculescu 1990] **D Voiculescu**, *Noncommutative random variables and spectral problems in free product C^*-algebras*, Rocky Mountain J. Math. 20 (1990) 263–283 MR Zbl

[Voiculescu 1991] **D Voiculescu**, *Limit laws for random matrices and free products*, Invent. Math. 104 (1991) 201–220 MR Zbl

[Voiculescu et al. 1992] **D V Voiculescu**, **K J Dykema**, **A Nica**, *Free random variables*, CRM Monogr. Ser. 1, Amer. Math. Soc., Providence, RI (1992) MR Zbl

[Weingarten 1978] **D Weingarten**, *Asymptotic behavior of group integrals in the limit of infinite rank*, J. Math. Phys. 19 (1978) 999–1001 MR Zbl

[Witten 1991] **E Witten**, *On quantum gauge theories in two dimensions*, Comm. Math. Phys. 141 (1991) 153–209 MR Zbl

[Xu 1997] **F Xu**, *A random matrix model from two-dimensional Yang–Mills theory*, Comm. Math. Phys. 190 (1997) 287–307 MR Zbl

[Zagier 1994] **D Zagier**, *Values of zeta functions and their applications*, from "First European Congress of Mathematics, II" (A Joseph, F Mignot, F Murat, B Prum, R Rentschler, editors), Progr. Math. 120, Birkhäuser, Basel (1994) 497–512 MR Zbl

Department of Mathematical Sciences, Durham University
Durham, United Kingdom

michael.r.magee@durham.ac.uk

https://www.mmagee.net/

Proposed: David Fisher — Received: 2 December 2021
Seconded: John Lott, Leonid Polterovich — Revised: 3 January 2023

Geometry & Topology **29**:3 (2025) 1283–1344
DOI: 10.2140/gt.2025.29.1283
Published: 31 May 2025

Partial Okounkov bodies and Duistermaat–Heckman measures of non-Archimedean metrics

MINGCHEN XIA

Let X be a smooth complex projective variety. We construct partial Okounkov bodies associated with Hermitian big line bundles (L,ϕ) on X. We show that partial Okounkov bodies are universal invariants of the singularities of ϕ. As an application, we construct Duistermaat–Heckman measures associated with finite-energy metrics on the Berkovich analytification of an ample line bundle.

14M25, 32U05

1 Introduction

1.1 Background

Let X be an irreducible smooth projective variety of dimension n and L be a big holomorphic line bundle on X. Given any admissible flag $X = Y_0 \supseteq Y_1 \supseteq \cdots \supseteq Y_n$ on X (see Definition 2.7 for the precise definition), one can attach a natural convex body $\Delta(L)$ of dimension n to L, generalizing the classical Newton polytope construction in toric geometry. This construction was first considered by Okounkov [47; 48] and then extended by Lazarsfeld and Mustaţă [45] and Kaveh and Khovanskii [43]. The convex body $\Delta(L)$ is known as the *Okounkov body* or *Newton–Okounkov body* associated with L

(with respect to the given flag). We briefly recall its definition: given any nonzero $s \in \mathrm{H}^0(X, L^k)$, let $\nu_1(s)$ be the vanishing order of s along Y_1. Then s can be regarded as a section of $\mathrm{H}^0(X, L^k \otimes \mathcal{O}_X(-\nu_1(s)Y_1))$. It follows that $s_1 := s|_{Y_1}$ is a nonzero section of $L|_{Y_1}^k \otimes \mathcal{O}_X(-\nu_1(s)Y_1)|_{Y_1}$. We can then repeat the same procedure with s_1 and Y_2 in place of s and Y_1. Repeating this construction, we end up with $\nu(s) = (\nu_1(s), \dots, \nu_n(s)) \in \mathbb{N}^n$. In fact, ν extends naturally to a rank n valuation on $\mathbb{C}(X)$. Consider the semigroup

$$\Gamma(L) := \{(\nu(s), k) \in \mathbb{Z}^{n+1} \mid k \in \mathbb{N}, s \in \mathrm{H}^0(X, L^k)^\times\}.$$

Then $\Delta(L)$ is the intersection of the closed convex cone in $\mathbb{R}^{n+1}$ generated by $\Gamma(L)$ with the hyperplane $\{(x, 1) \mid x \in \mathbb{R}^n\}$. A key property of $\Delta(L)$ is that the Lebesgue volume of $\Delta(L)$ is proportional to the volume of the line bundle L:

$$\operatorname{vol} \Delta(L) = \frac{1}{n!}\langle L^n \rangle. \tag{1-1}$$

Here $\langle \bullet \rangle$ denotes the movable intersection product in the sense of Boucksom, Demailly, Păun and Peternell [12] and Boucksom, Favre and Jonsson [15].

In [45], Lazarsfeld and Mustaţă showed moreover that $\Delta(L)$ depends only on the numerical class of L. Conversely, it is shown by Jow [41] that the information of all Okounkov bodies with respect to various flags actually determines the numerical class of L. In other words, Okounkov bodies can be regarded as universal numerical invariants of big line bundles.

This paper concerns a similar problem. Assume that L is equipped with a singular plurisubharmonic (psh) metric ϕ. We will construct universal invariants of the singularity type of ϕ. We call these universal invariants the *partial Okounkov bodies* of (L, ϕ).

1.2 Main results

Let us explain more details about the construction of partial Okounkov bodies. Recall that any admissible flag on X induces a rank n valuation on $\mathbb{C}(X)$ with values in $\mathbb{Z}^n$. We will work more generally with such valuations, not necessarily coming from admissible flags on X. We define a set

$$\Gamma(L, \phi) := \{(\nu(s), k) \in \mathbb{Z}^{n+1} \mid k \in \mathbb{N}, s \in \mathrm{H}^0(X, L^k \otimes \mathcal{I}(k\phi))^\times\} \tag{1-2}$$

similar to $\Gamma(L)$. Here $\mathcal{I}(\bullet)$ denotes the multiplier ideal sheaf in the sense of Nadel. However, a key difference here is that $\Gamma(L, \phi)$ is not a semigroup in general. Thus, the constructions in both [45] and [43] break down. We will show that in this case, there is still a canonical construction of Okounkov bodies.

Before stating our main theorem, let us recall the definition of volume. The volume of (L, ϕ) is defined as

$$\operatorname{vol}(L, \phi) := \lim_{k \to \infty} \frac{1}{k^n} h^0(X, L^k \otimes \mathcal{I}(k\phi)).$$

The existence of this limit is proved in Darvas and Xia [33].

Theorem A *Let* (L,ϕ) *be as above. Assume that* $\mathrm{vol}(L,\phi)>0$*. Then there is a convex body* $\Delta(L,\phi)\subseteq\Delta(L)$ *associated with* (L,ϕ) *satisfying*

$$\mathrm{vol}\,\Delta(L,\phi)=\mathrm{vol}(L,\phi). \tag{1-3}$$

Moreover, $\Delta(L,\phi)$ *is continuous in* ϕ *if* $\int_X(\mathrm{dd}^{\mathrm{c}}\phi)^n>0$*.* (Here the set of ϕ is endowed with the d_S-pseudometric in the sense of Darvas, Di Nezza and Lu [29] and the set of convex bodies is endowed with the Hausdorff metric.)

Define

$$\Gamma_k:=\{k^{-1}\nu(s)\in\mathbb{R}^n \mid s\in\mathrm{H}^0(X,L^k\otimes\mathcal{I}(k\phi))^\times\}$$

and let Δ_k *denote the convex hull of* Γ_k*. Then*

$$\Delta_k\to\Delta(L,\phi) \tag{1-4}$$

with respect to the Hausdorff metric if $\mathrm{vol}(L,\phi)>0$*.*

Observe that the last assertion actually uniquely determines $\Delta(L,\phi)$, so $\Delta(L,\phi)$ can be regarded as canonically attached to the given data (X,L,ϕ,ν).

The convex body $\Delta(L,\phi)$ is called the *partial Okounkov body* of (L,ϕ) with respect to the given valuation. Here the word *partial* refers to the fact that the partial Okounkov bodies are contained in $\Delta(L)$. One should not confuse them with the notion of Okounkov bodies with respect to partial flags.

We will also extend the definition to the case $\mathrm{vol}(L,\phi)=0$ in Section 5.6, at the expense of losing continuity in ϕ.

Observe that (1-3) bears strong resemblance with (1-1). In fact, when ϕ has minimal singularities, $\Delta(L,\phi)=\Delta(L)$ and (1-3) just reduces to (1-1).

The second main result says that partial Okounkov bodies uniquely determine the $\mathcal{I}$-singularity type of ϕ.

Theorem B *Let* L *be a big line bundle on* X*. Let* ϕ *and* ϕ' *be two singular psh metrics on* L *with positive volumes. Then the following are equivalent*:

(1) $\phi\sim_{\mathcal{I}}\phi'$.

(2) $\Delta(L,\phi)=\Delta(L,\phi')$ *for all rank* n *valuations on* $\mathbb{C}(X)$ *taking values in* $\mathbb{Z}^n$*.*

Recall that $\phi\sim_{\mathcal{I}}\phi'$ means $\mathcal{I}(k\phi)=\mathcal{I}(k\phi')$ for all real $k>0$. This relation is studied in detail in Darvas and Xia [32; 33]. It captures a lot of important information about the singularity of a psh metric. Theorem B should be regarded as a metric analogue of Jow's theorem.

As a byproduct of our proof of Theorem B, we reprove a formula computing the generic Lelong numbers of currents of minimal singularities in $c_1(L)$, slightly generalizing Boucksom [9, Theorem 5.4]:

Theorem 1.1 (Corollary 5.25) *Let L be a big line bundle on X. Consider a current $T_{\min}$ of minimal singularity in $c_1(L)$. Then for any prime divisor E over X, we have*

$$\nu(T_{\min}, E) = \lim_{k\to\infty} \frac{1}{k} \operatorname{ord}_E \mathrm{H}^0(X, L^k). \tag{1-5}$$

Here $\nu(T_{\min}, E)$ denotes the generic Lelong number of $T_{\min}$ along E.

As a consequence, we find a new formula computing the multiplier ideal sheaf $\mathcal{I}(T_{\min})$ in Corollary 5.26.

The third main result is an analogue of Witt Nyström [52]. Given any continuous metric ψ on L, one can naturally construct a convex function $c[\psi]$ on $\operatorname{Int}\Delta(L)$, known as the *Chebyshev transform* of ψ. The main property of $c[\psi]$ is that given another continuous metric ψ' on L, we have

$$\int_{\Delta(L)} (c[\psi] - c[\psi'])\,\mathrm{d}\lambda = \operatorname{vol}(\psi, \psi'), \tag{1-6}$$

where $\operatorname{vol}(\psi, \psi')$ is the relative volume as studied in Berman and Boucksom [4] and Berman, Boucksom and Witt Nyström [6] and $\mathrm{d}\lambda$ is the Lebesgue measure on $\mathbb{R}^n$. In our setup, we also associate a convex function $c_{[\phi]}[\psi]\colon \operatorname{Int}\Delta(L,\phi) \to \mathbb{R}$. Moreover:

Theorem C *Assume that the valuation ν is induced by an admissible flag on X. Let ψ and ψ' be two continuous metrics on L. Then*

$$\int_{\Delta(L,\phi)} (c_{[\phi]}[\psi] - c_{[\phi]}[\psi'])\,\mathrm{d}\lambda = -\mathcal{E}^{\theta}_{[\phi]}(\psi) + \mathcal{E}^{\theta}_{[\phi]}(\psi'), \tag{1-7}$$

where $\mathcal{E}^{\theta}_{[\phi]}$ is the partial equilibrium energy functional defined in (6-1).

Theorems A, B and C together give convex-geometric interpretations of the main results of [32; 33]. These results also provide us with a convex-geometric approach to the study of psh singularities.

As an application of our theory, we prove a generalization of Boucksom–Chen theorem (Theorem 7.9). Recall that the Boucksom–Chen theorem [11] says that given a *multiplicative* filtration $\mathscr{F}$ on the section ring $R(X, L)$, one can naturally associate a probability measure on $\mathbb{R}$, known as the *Duistermaat–Heckman measure*. Moreover, the Duistermaat–Heckman measure is the weak limit of a sequence of discrete measures μ_k associated with the filtration $\mathscr{F}$ on $\mathrm{H}^0(X, L^k)$. We show that this construction can be generalized to all $\mathcal{I}$-model test curves, not necessarily coming from filtrations. Here we only prove the generalized Boucksom–Chen theorem for filtrations on the full graded linear series, which suffices for our purpose. It is, however, easy to see that the techniques apply to more general situations.

More generally, we introduce the notion of an Okounkov test curve (Definition 7.2) and generalize Duistermaat–Heckman measures to this setting.

When L is ample, this construction allows us to associate a Radon measure $\mathrm{DH}(\eta)$ on $\mathbb{R}$ with each element η in the non-Archimedean space $\mathcal{E}^1(L^{\mathrm{an}})$ in the sense of Boucksom and Jonsson [17]; see Definition 7.13. The space $\mathcal{E}^1(L^{\mathrm{an}})$ can be seen as the completion of the space of test configurations.

Theorem 1.2 *The Duistermaat–Heckman measure construction of test configurations as in Witt Nyström [51] admits a unique continuous extension* $\mathrm{DH}\colon \mathcal{E}^1(L^{\mathrm{an}}) \to \mathcal{M}(\mathbb{R})$*. Here* $\mathcal{M}(\mathbb{R})$ *is the space of Radon measures on* $\mathbb{R}$*.*

The Duistermaat–Heckman measure of a non-Archimedean metric is also constructed by Inoue [40] using a different method. See Remark 7.17 for more details.

In Theorem 7.16, we will furthermore prove that $\mathrm{DH}(\eta)$ contains a lot of interesting information of η.

In the last section, we interpret the partial Okounkov bodies in the toric setting. We prove the following results:

Theorem 1.3 *Let X be a smooth toric variety of dimension n and (L,ϕ) be a toric invariant Hermitian big line bundle on X with positive volume. Fix a toric invariant admissible flag on X. Recall that upon choosing a toric invariant rational section of L, ϕ can be identified with a convex function $\phi_{\mathbb{R}}$ on $\mathbb{R}^n$. Then the partial Okounkov body $\Delta(L,\phi)$ is naturally identified with the closure of the image of $\nabla\phi_{\mathbb{R}}$.*

Theorem D *Let (L_i,ϕ_i) for $i=1,\dots,n$ be toric invariant Hermitian big line bundles on X of positive volumes. If the toric invariant flag $(Y_\bullet)$ satisfies the additional condition that Y_n is not contained in the polar locus of any ϕ_i, then*

$$\int_X \mathrm{dd}^{\mathrm{c}}\phi_1\wedge\cdots\wedge\mathrm{dd}^{\mathrm{c}}\phi_n = n!\,\mathrm{vol}(\Delta(L_1,\phi_1),\dots,\Delta(L_n,\phi_n)).$$

It is of interest to generalize Theorem D to the nontoric setting as well. As shown by Example 8.5, the nontoric generalization has to involve all valuations instead of just one.

Lastly, let us mention that our generalization of the Boucksom–Chen theorem has important consequences in Archimedean pluripotential theory as well. When applied to *generalized deformation to the normal cone* in the sense of Xia [55], it gives a number of interesting equidistribution results of the jumping numbers of multiplier ideal sheaves. As a detailed investigation would lead us too far away, we do not include these results in this paper.

1.3 Strategy of the proofs

We will sketch the proof of these theorems.

Proof of Theorem A In general, the graded linear space

$$W(L,\phi) := \bigoplus_{k=0}^{\infty} \mathrm{H}^0(X, L^k\otimes\mathcal{I}(k\phi))$$

is not an algebra and similarly $\Gamma(L,\phi)$ as defined in (1-2) is not a semigroup. Thus, one cannot directly apply the theory of graded linear series or the theory of semigroups as in Lazarsfeld and Mustaţă [45] and Kaveh and Khovanskii [43].

A key observation here is that although $\Gamma(L,\phi)$ is not a semigroup, it is not too far away from being one.

To make this precise, we introduce a pseudometric d on the space $\widehat{S}$ of subsets of $\mathbb{Z}^{n+1}$ lying in a suitable strictly convex cone:

$$d(S,S') := \varlimsup_{k\to\infty} k^{-n}(|S_k|+|S'_k|-2|S_k\cap S'_k|).$$

Let $\sim$ be the equivalence relation defined by d. The classical Okounkov body construction associates with each semigroup a convex body. As we will prove later, this map factorizes through the $\sim$-equivalence classes, and it extends continuously to an *almost semigroup*, namely an object in $\widehat{S}$ which can be approximated by certain *nice* semigroups with respect to d.

In order to define the Okounkov body of (L,ϕ), we will actually show that $\Gamma(L,\phi)$ is an almost semigroup and we could simply define

$$\Delta(L,\phi) := \Delta(\Gamma(L,\phi)).$$

The proof follows the same pattern as the proof in [33]. We proceed by approximations. We first consider the case where ϕ has analytic singularities. In this case, after taking a suitable resolution, we can easily see that $W(L,\phi)$ can be approximated by graded linear series both from above and from below. In the case of a singular ϕ with $\mathrm{dd^c}\phi$ being a Kähler current, we make use of analytic approximations as in Demailly, Peternell and Schneider [36] and Cao [19]. More precisely, take a quasi-equisingular approximation ϕ^j of ϕ. Based on the convergence theorems proved in [33], we can show that $\Gamma(L,\phi^j)$ converges to $\Gamma(L,\phi)$ with respect to the pseudometric d, which enables us to conclude in this case. Finally, in the general case, a trick discovered in [29] and [33] enables us to reduce to the previous case.

Along the lines of the proof, we actually find that $\Gamma(L,\phi)$ satisfies a stronger property (1-4). This property is essential to the proof of Theorem B; we call it the *Hausdorff convergence property*.

Proof of Theorem B Recall that in the classical setting, we can read information about the asymptotic base loci of L from the Okounkov body $\Delta(L)$ directly; see [21]. In our setup, the analogue says that the Okounkov body $\Delta(L,\phi)$ gives information about the generic Lelong numbers of ϕ. We will prove a qualitative version of Theorem B:

Theorem 1.4 *Let E be a prime divisor over X. Let $\pi: Z\to X$ be a birational model of X such that E is a divisor on Z. Take an admissible flag $(Y_\bullet)$ on Z with $Y_1=E$, then*

$$\nu(\phi,E) = \min_{x\in\Delta(\pi^*L,\pi^*\phi)} x_1.$$

Here $\nu(\phi,E)$ is the generic Lelong number of ϕ along E, defined as the minimum of the Lelong numbers $\nu(\pi^*\phi,x)$ for all $x\in E$. The proof of Theorem 1.4 again follows the same pattern as in the proof of Theorem A. With some efforts, we can reduce the problem to the case where ϕ has analytic singularities along some normal crossing $\mathbb{Q}$-divisor on X and $\mathrm{dd^c}\phi$ is a Kähler current. In this case, the desired result follows from a result proved in [55].

Proofs of Theorems C and D The proofs roughly follow the same pattern as above. Namely, we first handle the case of analytic singularities and then conclude the general case by suitable approximations. We will not repeat the details here.

As explained above, our approach to general psh singularities requires a number of approximations; this motivates the study of the metric geometry of the space of psh singularity types. We prove the continuity of mixed masses under d_S-approximations:

Theorem 1.5 (Theorem 4.2) *Let θ_i for $i = 1, \dots, n$ be smooth closed real $(1,1)$-forms representing big classes on a connected compact Kähler manifold X of dimension n. Let $\varphi_i^k, \varphi_i \in \mathrm{PSH}(X, \theta_i)$ for $i = 1, \dots, n$ and $k \in \mathbb{N}$. Assume that $\varphi_i^k \xrightarrow{d_{S,\theta_i}} \varphi_i$ for all i as $k \to \infty$. Then*

$$\lim_{k\to\infty} \int_X \theta_{1,\varphi_1^k} \wedge \dots \wedge \theta_{n,\varphi_n^k} = \int_X \theta_{1,\varphi_1} \wedge \dots \wedge \theta_{n,\varphi_n}. \tag{1-8}$$

Here the Monge–Ampère operators are taken in the nonpluripolar sense.

This theorem and its various consequences are indispensable in all of our proofs. They are of independent interest as well.

1.4 Structure of the paper

In Section 2, we collect a few preliminaries. In Section 3, we study the Okounkov bodies of almost semigroups. In Section 4, we further develop the theory of d_S-pseudometrics on the space of singularity types initiated in [29]. In Section 5, we define partial Okounkov bodies associated with Hermitian pseudoeffective line bundles and prove a number of properties. In Section 6, we define and study Chebyshev transforms of continuous metrics. In Section 7, we generalize the theory of Boucksom–Chen and study the non-Archimedean Duistermaat–Heckman measures. In Section 8, we give an explicit description of partial Okounkov bodies construction in terms of the moment polytope in the toric situation.

1.5 Conventions

In this paper, Monge–Ampère operators θ_φ^n refer to the nonpluripolar product in the sense of Boucksom, Eyssidieux, Guedj and Zeriahi [13]. The group $\mathbb{Z}^n$ is always endowed with the lexicographic order. A line bundle always refers to a holomorphic line bundle. We do not distinguish a line bundle and the associated invertible sheaf. When talking about a birational modification (resolution) $\pi : Y \to X$, we always assume that Y is smooth and π is projective. We follow the convention that $\mathrm{dd^c} = (\mathrm{i}/2\pi)\partial\bar\partial$.

Acknowledgements I benefited a lot from discussions with David Witt Nyström, Chen Jiang, Sébastien Boucksom,Yi Yao, Jian Xiao, Tamás Darvas, Kewei Zhang, Longke Tang and Rémi Reboulet. I would like to thank especially Yi Yao for explaining his computations in the toric setting to me, Chen Jiang for providing Example 8.5 and Yaxiong Liu for pointing out a number of typos in the arXiv version.

I also want to express my gratitude to the referees for numerous valuable comments helping to improve this paper in many aspects.

The author is supported by Knut och Alice Wallenbergs Stiftelse grant KAW 2021.0231.

2 Preliminaries

2.1 Hausdorff metric of convex bodies

In this section, we recall the theory of Hausdorff metrics on the set of convex bodies following [50, Section 1.8]. Fix $n \in \mathbb{N}$. Recall that a convex body in $\mathbb{R}^n$ is a nonempty compact convex subset of $\mathbb{R}^n$, which may have empty interior. Let $\mathcal{K}_n$ denote the set of convex bodies in $\mathbb{R}^n$. We will fix the Lebesgue measure $\mathrm{d}\lambda$ on $\mathbb{R}^n$, normalized so that the unit cube has volume 1.

Recall the definition of the Hausdorff metric between $K_1, K_2 \in \mathcal{K}_n$:

$$d_n(K_1, K_2) := \max\left\{ \sup_{x_1 \in K_1} \inf_{x_2 \in K_2} |x_1 - x_2|, \sup_{x_2 \in K_2} \inf_{x_1 \in K_1} |x_1 - x_2| \right\}.$$

We extend d_n to an extended metric on $\mathcal{K}_n \cup \{\varnothing\}$ by setting

$$d_n(K, \varnothing) = \infty \quad \text{for all } K \in \mathcal{K}_n.$$

Theorem 2.1 *The metric space $(\mathcal{K}_n, d_n)$ is complete.*

Theorem 2.2 (Blaschke selection) *Every bounded sequence in $\mathcal{K}_n$ has a convergent subsequence.*

Theorem 2.3 *The Lebesgue volume* $\mathrm{vol}\colon \mathcal{K}_n \to \mathbb{R}_{\geq 0}$ *is continuous.*

Theorem 2.4 *Let $K_i, K \in \mathcal{K}_n$ for $i \in \mathbb{N}$. Then $K_i \xrightarrow{d_n} K$ if and only if the following conditions hold:*

(1) *Each point $x \in K$ is the limit of a sequence $x_i \in K_i$.*

(2) *The limit of any convergent sequence $(x_{i_j})_{j \in \mathbb{N}}$ with $x_{i_j} \in K_{i_j}$ lies in K, where i_j is a subsequence of $1, 2, \dots$.*

The proofs of all these results can be found in [50, Section 1.8].

Lemma 2.5 *Let $K_0, K_1 \in \mathcal{K}_n$. Assume that $K_0 \subseteq K_1$ and*

$$\mathrm{vol}\, K_0 = \mathrm{vol}\, K_1 > 0.$$

Then $K_0 = K_1$.

Proof In fact, if $K_1 \neq K_0$, then $K_1 \setminus K_0$ is a nonempty open subset of K_1. As $\mathrm{vol}\, K_1 > 0$, $(K_1 \setminus K_0) \cap \mathrm{Int}\, K_1 \neq \varnothing$. Thus, $\mathrm{vol}\, K_1 > \mathrm{vol}\, K_0$, which is a contradiction. $\square$

Let $K \in \mathcal{K}_n$ be a convex body with positive volume. For $\delta > 0$ small enough, let

$$K^\delta := \{x \in K \mid d(x, \partial K) \geq \delta\}.$$

Then $K_\delta \in \mathcal{K}_n$ for δ small enough.

Lemma 2.6 *Let $K \in \mathcal{K}_n$ be a convex body with positive volume and $K' \in \mathcal{K}_n$. Assume that for some large enough $k \in \mathbb{Z}_{>0}$, K' contains $K \cap (k^{-1}\mathbb{Z})^n$, then $K' \supseteq K^{n^{1/2}k^{-1}}$.*

Proof Let $x \in K^{n^{1/2}k^{-1}}$. By assumption, the closed ball B with center x and radius $n^{1/2}k^{-1}$ is contained in K. Observe that x can be written as a convex combination of points in $B \cap (k^{-1}\mathbb{Z})^n$, which are contained in K' by assumption. It follows that $x \in K'$. □

Given a sequence of convex bodies K_i ($i \in \mathbb{N}$), we set

$$\varliminf_{i\to\infty} K_i = \overline{\bigcup_{i=0}^{\infty} \bigcap_{j\geq i} K_j}.$$

Suppose K is the limit of a subsequence of K_i, we have

$$\varliminf_{i\to\infty} K_i \subseteq K. \tag{2-1}$$

This is a simple consequence of Theorem 2.4.

2.2 Admissible flags and valuations

Let X be an irreducible normal projective variety of dimension n.

Definition 2.7 An *admissible flag* $(Y_\bullet)$ on X is a flag of subvarieties

$$X = Y_0 \supseteq Y_1 \supseteq \cdots \supseteq Y_n$$

such that Y_i is irreducible of codimension i and smooth at the point Y_n.

Given any admissible flag $(Y_\bullet)$, we can define a rank n valuation $\nu_{(Y_\bullet)}\colon \mathbb{C}(X)^\times \to \mathbb{Z}^n$ as in [45]. Here we consider $\mathbb{Z}^n$ as a totally ordered abelian group with the lexicographic order. We recall the definition: let $s \in \mathbb{C}(X)^\times$. Let $\nu_1(s) = \operatorname{ord}_{Y_1} s$. After localization around Y_n, we can take a local defining equation t^1 of Y_1; set $s_1 = (s(t^1)^{-\nu_1(s)})|_{Y_1}$. Then $s_1 \in \mathbb{C}(Y_1)$. We can repeat this construction with Y_2 in place of Y_1 to get $\nu_2(s)$ and s_2. Repeating this construction n times, we get $\nu_{(Y_\bullet)}(s) = \nu(s) = (\nu_1(s), \nu_2(s), \dots, \nu_n(s)) \in \mathbb{Z}^n$. It is easy to verify that ν is indeed a rank n valuation.

Remark 2.8 Conversely, by a theorem of Abhyankar, any valuation of $\mathbb{C}(X)$ with Noetherian valuation ring of rank n is equivalent to a valuation taking value in $\mathbb{Z}^n$; see [38, Chapter 0, Theorem 6.5.2]. As shown in [23, Theorem 2.9], any such valuation is equivalent to (but not necessarily equal to) a valuation induced by an admissible flag on a birational modification of X. Here two valuations ν and ν' with value in $\mathbb{Z}^n$ are equivalent if one can find a matrix G of the form $\mathrm{I} + N$, where N is strictly upper triangular with integral entries, such that $\nu' = G\nu$.

2.3 Model potentials and $\mathcal{I}$-model potentials

Let X be a connected compact Kähler manifold of dimension n and θ be a smooth closed real $(1,1)$-form representing a $(1,1)$-cohomology class $[\theta]$. Define $V_\theta := \sup\{\varphi \in \mathrm{PSH}(X,\theta) \mid \varphi \le 0\}$. For any two $\varphi, \psi \in \mathrm{PSH}(X,\theta)$, we say φ is *more singular* than ψ and write $[\varphi] \preceq [\psi]$ if there is a constant C such that $\varphi \le \psi + C$. When $\varphi \preceq \psi$ and $\psi \preceq \varphi$, we say that they have the same *singularity type*. We write $\theta_\varphi = \theta + \mathrm{dd^c}\varphi$.

Definition 2.9 Let $\varphi \in \mathrm{PSH}(X,\theta)$. Define

$$C^\theta[\varphi] := \sup{*}\left\{\psi \in \mathrm{PSH}(X,\theta) \;\middle|\; [\varphi] \preceq [\psi], \psi \le 0, \int_X \theta_\varphi^k \wedge \theta_{V_\theta}^{n-k} = \int_X \theta_\psi^k \wedge \theta_{V_\theta}^{n-k} \text{ for all } k\right\}. \tag{2-2}$$

If $C^\theta[\varphi] = \varphi$, we say φ is a *model potential*. We omit θ from the notation if there is no risk of confusion.

Here and in the sequel the Monge–Ampère type operators are taken in the nonpluripolar sense [13].

Proposition 2.10 [29, Proposition 2.6] *For any $\varphi \in \mathrm{PSH}(X,\theta)$, $C^\theta[\varphi]$ is a model potential in $\mathrm{PSH}(X,\theta)$. When $\int_X \theta_\varphi^n > 0$ we have*

$$C^\theta[\varphi] = P^\theta[\varphi],$$

where

$$P^\theta[\varphi] := \sup{*}\{\psi \in \mathrm{PSH}(X,\theta) \mid [\psi] \preceq [\varphi], \psi \le 0\}. \tag{2-3}$$

In general, we only have

$$C^\theta[\varphi] = \lim_{\epsilon \to 0+} P^\theta[(1-\epsilon)\varphi + \epsilon V_\theta]. \tag{2-4}$$

We omit θ from the notation $P^\theta[\varphi]$ if there is no risk of confusion.

Definition 2.11 A *birational model* of X is a projective birational morphism $\pi: Y \to X$ from a *smooth* projective variety Y to X.

Recall that $\mathcal{I}(\varphi)$ denotes the multiplier ideal sheaf of a qpsh function φ on X in the sense of Nadel, namely the coherent subsheaf of $\mathcal{O}_X$ consisting of functions f such that $|f|^2\exp(-\varphi)$ is locally integrable.

Definition 2.12 Let φ, ψ be two quasi-psh functions, we say $\varphi \preceq_{\mathcal{I}} \psi$ if the following equivalent conditions are satisfied:

1. $\mathcal{I}(k\varphi) \subseteq \mathcal{I}(k\psi)$ for all real $k > 0$.
2. $\mathcal{I}(k\varphi) \subseteq \mathcal{I}(k\psi)$ for all integer $k > 0$.
3. For any birational model $\pi: Y \to X$ and any $y \in Y$, we have $\nu(\pi^*\varphi, y) \ge \nu(\pi^*\psi, y)$.

The equivalence between (1) and (3) is just [32, Corollary 2.16]. The equivalence between (2) and (3) follows from [32, Proposition 2.14].

We say $\varphi \sim_{\mathcal{I}} \psi$ if $\varphi \preceq_{\mathcal{I}} \psi$ and $\psi \preceq_{\mathcal{I}} \varphi$.

Given any $\varphi \in \mathrm{PSH}(X,\theta)$, we define

$$P^{\theta}[\varphi]_{\mathcal{I}} := \sup\{\psi \in \mathrm{PSH}(X,\theta) \mid \psi \preceq_{\mathcal{I}} \varphi, \psi \le 0\}.$$

We omit θ when there is no risk of confusion. We say φ is $\mathcal{I}$*-model* if $\varphi = P[\varphi]_{\mathcal{I}}$.

It is shown in [32, Proposition 2.18] that $P[\varphi]_{\mathcal{I}} \in \mathrm{PSH}(X,\theta)$ and $\varphi \sim_{\mathcal{I}} P[\varphi]_{\mathcal{I}}$. Moreover, $P[\varphi]_{\mathcal{I}}$ is always $\mathcal{I}$-model. We can also talk about the $\sim_{\mathcal{I}}$ relation of two psh metric on L in the obvious manner.

Typical model potentials are not $\mathcal{I}$-model; however, the converse is true:

Proposition 2.13 *If $\psi \in \mathrm{PSH}(X,\theta)$ is an $\mathcal{I}$-model potential then it is model.*

Proof We need to show that $\psi \sim_{\mathcal{I}} C[\psi]$. Let $\pi : Z \to X$ be a birational modification. Let $z \in Z$. As $\psi \le C[\psi] + C$ for some constant C, it suffices to show that

$$\nu(C[\psi], z) \ge \nu(\psi, z).$$

Here $\nu(\psi, z)$ denotes the Lelong number of $\pi^*\psi$ at z. By (2-4) and the upper semicontinuity of Lelong numbers (see [39, page 73, Exercise 2.7]), we find

$$\nu(C[\psi], z) \ge \lim_{\epsilon\to 0+} \nu(P[(1-\epsilon)\psi + \epsilon V_\theta], z) = \lim_{\epsilon\to 0+} \nu((1-\epsilon)\psi + \epsilon V_\theta, z) = \nu(\psi, z).$$

We conclude our assertion. □

2.4 Potentials with analytic singularities

Definition 2.14 A quasi-plurisubharmonic function (quasi-psh) φ on X is said to have *analytic singularities* if for each $x \in X$, there is a neighborhood $U_x \subseteq X$ of x with respect to the Euclidean topology, such that on U_x,

$$\varphi = c \log\Bigl(\sum_{j=1}^{N_x} |f_j|^2\Bigr) + \psi, \tag{2-5}$$

where $c \in \mathbb{Q}_{\ge 0}$, the f_j are analytic functions on U_x, $N_x \in \mathbb{Z}_{>0}$ is an integer depending on x, $\psi \in L^\infty(U_x)$.

Definition 2.15 Let D be an effective normal crossing $\mathbb{R}$-divisor on X. Let $D = \sum_i a_i D_i$ with D_i being prime divisors and $a_i \in \mathbb{R}_{>0}$. We say that a quasi-psh function φ has *analytic singularities along D* if locally, in the Euclidean topology,

$$\varphi = \sum_i a_i \log|s_i|^2 + \psi,$$

where s_i is a local holomorphic function defining D_i, ψ is a bounded function.

In the sequel, when we talk about a normal crossing divisor, we always assume that it is effective.

Note that a potential with analytic singularities along a normal crossing $\mathbb{Q}$-divisor has analytic singularities in the sense of Definition 2.14.

For any quasi-psh function φ on X with analytic singularities, there is always a birational model $\pi : Y \to X$ such that $\pi^*\varphi$ has analytic singularities along a normal crossing $\mathbb{Q}$-divisor on Y. See [46, Lemma 2.3.19] for example. We remind the readers that in [46], the definition of analytic singularities differs slightly from ours: they require the remainder ψ to be smooth instead of just bounded. However, the proof of [46, Lemma 2.3.19] works *verbatim* with our definition.

2.5 Quasi-equisingular approximations

We recall the concept of quasi-equisingular approximations in the sense of [19; 36].

Let X be a connected compact Kähler manifold of dimension n and θ (resp. θ_i for $i = 1, \dots, n$) be a smooth real $(1,1)$-form representing a pseudoeffective $(1,1)$-cohomology class $[\theta]$ (resp. $[\theta_i]$). Take a Kähler form ω on X.

Definition 2.16 Let $\varphi \in \mathrm{PSH}(X, \theta)$. Define a *quasi-equisingular approximation* to be a sequence $\varphi^j \in \mathrm{PSH}(X, \theta + \epsilon_j \omega)$ with $\epsilon_j \to 0$ such that

1. $\varphi^j \to \varphi$ in L^1,
2. φ^j has analytic singularities,
3. $\varphi^{j+1} \le \varphi^j$,
4. For any $\delta > 0$, $k > 0$, there is $j_0 > 0$ such that for $j \ge j_0$,
$$\mathcal{I}(k(1+\delta)\varphi^j) \subseteq \mathcal{I}(k\varphi) \subseteq \mathcal{I}(k\varphi^j).$$

The existence of a quasi-equisingular approximation follows from the arguments in [19; 35; 36].

2.6 Volumes of Hermitian pseudoeffective line bundles

Let X be a smooth irreducible projective variety of dimension n.

Definition 2.17 A *Hermitian pseudoeffective* (*psef*) *line bundle* on X is a pair (L, ϕ), where L is a pseudoeffective line bundle on X and ϕ is a psh metric on L.

When L is big, we say (L, ϕ) is a Hermitian big line bundle.

Let (L, ϕ) be a Hermitian psef line bundle on X. In this section, we recall the main results in [32; 33] concerning the volume of (L, ϕ).

Definition 2.18 The *volume* of (L,ϕ) is defined as

$$\operatorname{vol}(L,\phi) := \lim_{k\to\infty} \frac{1}{k^n} h^0(X, L^k \otimes \mathcal{I}(k\phi)).$$

The existence of the limit follows from [33, Theorem 1.1].

We take a smooth Hermitian metric h on L. Set $\theta = c_1(L,h)$. Then we can identify ϕ with a θ-psh function φ, namely $\phi = h\exp(-\varphi)$.

Theorem 2.19 [33, Theorem 1.1] *Under the above assumptions*,

$$\operatorname{vol}(L,\phi) = \frac{1}{n!}\int_X \theta^n_{P[\varphi]_{\mathcal{I}}}.$$

We argue that vol deserves the name *volume* by proving that it satisfies the Brunn–Minkowski inequality.

Corollary 2.20 *Let (L,ϕ) and (L,ϕ') be two Hermitian psef line bundles on X. Then*

$$\operatorname{vol}(L+L',\phi+\phi')^{1/n} \geq \operatorname{vol}(L,\phi)^{1/n} + \operatorname{vol}(L',\phi')^{1/n}. \tag{2-6}$$

Proof Fix a smooth Hermitian metric h' on L' with $\theta' = c_1(L',h')$. We identify ϕ' with $\varphi' \in \operatorname{PSH}(X,\theta')$. By Theorem 2.19, (2-6) is equivalent to

$$\left(\int_X (\theta+\theta'+\mathrm{dd^c} P^{\theta+\theta'}[\varphi+\varphi']_{\mathcal{I}})^n\right)^{1/n} \geq \left(\int_X \theta^n_{P^\theta[\varphi]_{\mathcal{I}}}\right)^{1/n} + \left(\int_X \theta'^n_{P^{\theta'}[\varphi']_{\mathcal{I}}}\right)^{1/n}.$$

Observe that

$$P^{\theta+\theta'}[\varphi+\varphi']_{\mathcal{I}} \geq P^\theta[\varphi]_{\mathcal{I}} + P^{\theta'}[\varphi']_{\mathcal{I}}.$$

Thus, by the monotonicity theorem of [53], it suffices to show that

$$\left(\int_X (\theta+\theta'+\mathrm{dd^c} P^{\theta}[\varphi]_{\mathcal{I}}+\mathrm{dd^c} P^{\theta'}[\varphi']_{\mathcal{I}})^n\right)^{1/n} \geq \left(\int_X \theta^n_{P^\theta[\varphi]_{\mathcal{I}}}\right)^{1/n} + \left(\int_X \theta'^n_{P^{\theta'}[\varphi']_{\mathcal{I}}}\right)^{1/n}.$$

This follows from [28, Theorem 6.1]. □

2.7 Non-Archimedean pluripotential theory

In this section, we briefly recall the notion of Berkovich analytification of a smooth complex projective variety and the pluripotential theory in the sense of Boucksom and Jonsson [17] on it.

For simplicity, we assume that X is a connected smooth projective variety of dimension n and L is an ample line bundle on X.

The set of real valuations on $\mathbb{C}(X)$ trivial on $\mathbb{C}$ is denoted by X^{val}. This set can be defined in the same way for nonsmooth varieties as well.

The center of a valuation v is the scheme-theoretic point $c = c(v)$ of X such that $v \geq 0$ on $\mathcal{O}_{X,c}$ and $v > 0$ on the maximal ideal $\mathfrak{m}_{X,c}$ of $\mathcal{O}_{X,c}$. The center exists and is unique.

Let X^{an} denote the Berkovich analytification X^{an} of X with respect to the trivial valuation on $\mathbb{C}$. As a set, X^{an} is the set of semivaluations on X, in other words, real-valued valuations v on irreducible reduced subvarieties Y in X that is trivial on $\mathbb{C}$. We call Y the *support* of the semivaluation v. In other words,

$$X^{\mathrm{an}} = \coprod_{Y} Y^{\mathrm{val}}.$$

The Berkovich space X^{an} admits a natural topology, called the Berkovich topology and a sheaf of analytic functions. The natural morphism of ringed spaces $X^{\mathrm{an}} \to X$ allows us to pullback L to an invertible sheaf L^{an} on X^{an}. See [2] for more details.

In [17], Boucksom and Jonsson developed a pluripotential theory with respect to $(X^{\mathrm{an}}, L^{\mathrm{an}})$, similar to its complex counterpart. In particular, there is a natural notion of plurisubharmonic metrics on L^{an}. In [17, Section 7.1], Boucksom and Jonsson defined the notion of energy pairings $(\varphi_0, \dots, \varphi_n)$ between $n+1$ plurisubharmonic metrics $\varphi_0, \dots, \varphi_n$ on L^{an}. One can then define the space $\mathcal{E}^1(L^{\mathrm{an}})$ of finite-energy metrics as the space of plurisubharmonic functions φ on L^{an} such that

$$E(\varphi) := \frac{1}{n+1}(\varphi, \dots, \varphi) > -\infty.$$

Note that our definition of E differs from the definition of [17] by a multiple $1/V$. We will explain the relation between the non-Archimedean pluripotential theory and the complex pluripotential theory in Section 7.4.

3 The Okounkov bodies of almost semigroups

Fix an integer $n \geq 0$. Fix a closed convex cone $C \subseteq \mathbb{R}^n \times \mathbb{R}_{\geq 0}$ such that $C \cap \{x_{n+1} = 0\} = \{0\}$. Here x_{n+1} is the last coordinate of $\mathbb{R}^{n+1}$.

3.1 Generality on semigroups

Write $\hat{\mathcal{S}}(C)$ for the set of subsets of $C \cap \mathbb{Z}^{n+1}$ and $\mathcal{S}(C)$ for the set of subsemigroups $S \subseteq C \cap \mathbb{Z}^{n+1}$. For each $k \in \mathbb{N}$ and $S \in \hat{\mathcal{S}}(C)$, we write

$$S_k := \{x \in \mathbb{Z}^n \mid (x, k) \in S\}.$$

Note that S_k is a finite set by our assumption on C.

We introduce a pseudometric on $\hat{\mathcal{S}}(C)$ as follows:

$$d(S, S') := \varlimsup_{k\to\infty} k^{-n}(|S_k| + |S'_k| - 2|(S \cap S')_k|).$$

Here $|\bullet|$ denotes the cardinality of a finite set.

Lemma 3.1 *The above-defined d is a pseudometric on $\hat{\mathcal{S}}(C)$.*

Proof Only the triangle inequality needs to be argued. Take $S, S', S'' \in \hat{\mathcal{S}}(C)$. We claim that for any $k \in \mathbb{N}$,

$$|S_k| + |S'_k| - 2|S_k \cap S'_k| + |S''_k| + |S'_k| - 2|S''_k \cap S'_k| \geq |S_k| + |S''_k| - 2|S_k \cap S''_k|.$$

From this the triangle inequality follows. To argue the claim, we rearrange it to the form

$$|S'_k| - |S_k \cap S'_k| \geq |S'_k \cap S''_k| - |S_k \cap S''_k|,$$

which is obvious. □

Given $S, S' \in \hat{\mathcal{S}}(C)$, we say S is equivalent to S' and write $S \sim S'$ if $d(S, S') = 0$. This is an equivalence relation by Lemma 3.1.

Lemma 3.2 *Given $S, S', S'' \in \hat{\mathcal{S}}(C)$, we have*

$$d(S \cap S'', S' \cap S'') \leq d(S, S').$$

In particular, if $S^i, S'^i \in \hat{\mathcal{S}}(C)$ $(i \in \mathbb{N})$ and $S^i \to S$, $S'^i \to S'$, then

$$S^i \cap S'^i \to S \cap S'.$$

Proof Observe that for any $k \in \mathbb{N}$,

$$|S_k \cap S''_k| - |S_k \cap S'_k \cap S''_k| \leq |S_k| - |S_k \cap S'_k|.$$

The same holds if we interchange S with S'. It follows that

$$|S_k \cap S''_k| + |S'_k \cap S''_k| - 2|S_k \cap S'_k \cap S''_k| \leq |S_k| + |S'_k| - 2|S_k \cap S'_k|.$$

The first assertion follows.

Next we compute

$$d(S^i \cap S'^i, S \cap S') \leq d(S^i \cap S'^i, S^i \cap S') + d(S^i \cap S', S \cap S') \leq d(S'^i, S') + d(S^i, S)$$

and the second assertion follows. □

The volume of $S \in \mathcal{S}(C)$ is defined as

$$\operatorname{vol} S := \lim_{k\to\infty} (ka)^{-n} |S_{ka}| = \varlimsup_{k\to\infty} k^{-n} |S_k|,$$

where a is a sufficiently divisible positive integer. The existence of the limit and its independence from a both follow from the more precise result [43, Theorem 2].

Lemma 3.3 *Let $S, S' \in \mathcal{S}(C)$, then*

$$|\operatorname{vol} S - \operatorname{vol} S'| \leq d(S, S').$$

Proof By definition, we have

$$d(S, S') \geq \operatorname{vol} S + \operatorname{vol} S' - 2\operatorname{vol}(S \cap S').$$

It follows that $\operatorname{vol} S - \operatorname{vol} S' \leq d(S, S')$ and $\operatorname{vol} S' - \operatorname{vol} S \leq d(S, S')$. □

We define $\bar{\mathcal{S}}(C)$ as the closure of $\mathcal{S}(C)$ in $\hat{\mathcal{S}}(C)$ with respect to the topology defined by the pseudometric d. By Lemma 3.3, $\mathrm{vol}\colon \mathcal{S}(C) \to \mathbb{R}$ admits a unique 1-Lipschitz extension to

$$\mathrm{vol}\colon \bar{\mathcal{S}}(C) \to \mathbb{R}. \tag{3-1}$$

Lemma 3.4 *Suppose that $S, S' \in \bar{\mathcal{S}}(C)$ and $S \subseteq S'$. Then*

$$\mathrm{vol}\, S \leq \mathrm{vol}\, S'.$$

Proof Take sequences S^j and S'^j in $\mathcal{S}(C)$ such that $S^j \to S$ and $S'^j \to S'$. By Lemma 3.2, after replacing S^j by $S^j \cap S'^j$, we may assume that $S^j \subseteq S'^j$ for each j. Then our assertion follows easily. □

3.2 Okounkov bodies of semigroups

Given $S \in \hat{\mathcal{S}}(C)$, we will write $C(S) \subseteq C$ for the closed convex cone generated by $S \cup \{0\}$. Moreover, for each $k \in \mathbb{Z}_{>0}$, we define

$$\Delta_k(S) := \mathrm{Conv}\{k^{-1}x \in \mathbb{R}^n \mid x \in S_k\} \subseteq \mathbb{R}^n.$$

Here Conv denotes the convex hull.

Definition 3.5 Let $\mathcal{S}'(C)$ be the subset of $\mathcal{S}(C)$ consisting of semigroups S such that S generates $\mathbb{Z}^{n+1}$ (as an abelian group).

Note that for any $S \in \mathcal{S}'(C)$, the cone $C(S)$ has full dimension (ie the topological interior is nonempty). Given a full-dimensional subcone $C' \subseteq C$, it is clear that $C' \cap \mathbb{Z}^{n+1} \in \mathcal{S}'(C)$.

This class behaves well under intersections:

Lemma 3.6 *Let $S, S' \in \mathcal{S}'(C)$. Assume that $\mathrm{vol}(S \cap S') > 0$. Then $S \cap S' \in \mathcal{S}'(C)$.*

The lemma obviously fails if $\mathrm{vol}(S \cap S') = 0$.

Proof We first observe that the cone $C(S) \cap C(S')$ has full dimension since otherwise $\mathrm{vol}(S \cap S') = 0$. Take a full-dimensional subcone C' in $C(S) \cap C(S')$ such that C' intersects the boundary of $C(S) \cap C(S')$ only at 0. It follows from [43, Theorem 1] that there is an integer $N > 0$ such that for any $x \in \mathbb{Z}^{n+1} \cap C'$ with Euclidean norm no less than N lies in $S \cap S'$. Therefore, $S \cap S' \in \mathcal{S}'(C)$. □

We recall the following definition from [43].

Definition 3.7 Given $S \in \mathcal{S}'(C)$, its *Okounkov body* is defined as

$$\Delta(S) := \{x \in \mathbb{R}^n \mid (x, 1) \in C(S)\}.$$

Theorem 3.8 *For each $S \in \mathcal{S}'(C)$, we have*

$$\operatorname{vol} S = \lim_{k\to\infty} k^{-n}|S_k| = \operatorname{vol}\Delta(S) > 0. \tag{3-2}$$

Moreover, as $k \to \infty$,

$$\Delta_k(S) \xrightarrow{d_n} \Delta(S). \tag{3-3}$$

This is essentially proved in [52, Lemma 4.8], which itself follows from a theorem of Khovanskii [44]. We remind the readers that (3-2) fails for a general $W \in \mathcal{S}(C)$; see [43, Theorem 2].

Proof The equalities (3-2) follow from the general theorem [43, Theorem 2].

It remains to prove (3-3). By the argument of [52, Lemma 4.8], for any compact set $K \subseteq \operatorname{Int}\Delta(S)$, there is $k_0 > 0$ such that for any $k \geq k_0$, $\alpha \in K \cap (k^{-1}\mathbb{Z})^n$ implies that $\alpha \in \Delta_k(S)$.

In particular, taking $K = \Delta(S)^\delta$ for any $\delta > 0$ and applying Lemma 2.6, we find

$$d_n(\Delta(S), \Delta_k(S)) \leq n^{1/2}k^{-1} + \delta$$

when k is large enough. This implies (3-3). □

Corollary 3.9 *Let $S, S' \in \mathcal{S}'(C)$. Assume that $\operatorname{vol}(S \cap S') > 0$. Then we have*

$$d(S, S') = \operatorname{vol} S + \operatorname{vol} S' - 2\operatorname{vol}(S \cap S').$$

Proof This is a direct consequence of Lemma 3.6 and (3-2). □

Lemma 3.10 *Given $S \in \mathcal{S}'(C)$, we have $S \sim \operatorname{Reg}(S)$.*

Recall that the regularization $\operatorname{Reg}(S)$ of S is defined as $C(S) \cap \mathbb{Z}^{n+1}$.

Proof Since S and $\operatorname{Reg}(S)$ have the same Okounkov body, we have $\operatorname{vol} S = \operatorname{vol}\operatorname{Reg}(S)$ by Theorem 3.8. By Corollary 3.9 again,

$$d(\operatorname{Reg}(S), S) = \operatorname{vol}\operatorname{Reg}(S) - \operatorname{vol} S = 0. \qquad \square$$

Lemma 3.11 *Let $S, S' \in \mathcal{S}'(C)$. Assume that $d(S, S') = 0$, then $\Delta(S) = \Delta(S')$.*

Proof Observe that $\operatorname{vol}(S \cap S') > 0$, as otherwise

$$d(S, S') \geq \operatorname{vol} S + \operatorname{vol} S' > 0,$$

which is a contradiction.

It follows from Lemma 3.6 that $S \cap S' \in \mathcal{S}'(C)$. It suffices to show that $\Delta(S) = \Delta(S \cap S')$. In fact, suppose that this holds, since $\operatorname{vol}\Delta(S') = \operatorname{vol} S' = \operatorname{vol} S = \operatorname{vol}\Delta(S)$, the inclusion $\Delta(S') \supseteq \Delta(S \cap S') = \Delta(S)$ is an equality.

By Lemma 3.2, we can therefore replace S' by $S \cap S'$ and assume that $S \supseteq S'$. Then clearly $\Delta(S) \supseteq \Delta(S')$. By (3-2),

$$\operatorname{vol} \Delta(S) = \operatorname{vol} \Delta(S').$$

Thus, $\Delta(S) = \Delta(S')$ by Lemma 2.5. □

Lemma 3.12 *Suppose that $S^i \in \mathcal{S}'(C)$ is a decreasing sequence such that*

$$\lim_{i\to\infty} \operatorname{vol} S^i > 0.$$

Then there is $S \in \mathcal{S}'(C)$ such that $S^i \to S$.

In general, one cannot simply take $S = \bigcap_i S^i$. For example, consider the sequence $S^i = S^1 \cap \{x_{n+1} \geq i\}$.

Proof By Lemma 3.10, we may replace S^i by its regularization and assume that $S^i = C(S^i) \cap \mathbb{Z}^{n+1}$. We define

$$S = \left(\bigcap_{i=1}^{\infty} C(S^i)\right) \cap \mathbb{Z}^{n+1}.$$

Since $\bigcap_{i=1}^{\infty} C(S^i)$ is a full-dimensional cone by assumption, we have $S \in \mathcal{S}'(C)$. By Corollary 3.9 and Theorem 3.8, we can compute the distance

$$d(S, S^i) = \operatorname{vol} S^i - \operatorname{vol} S = \operatorname{vol} \Delta(S^i) - \operatorname{vol} \Delta(S),$$

which tends to 0 by construction. □

3.3 Okounkov bodies of almost semigroups

Definition 3.13 We define $\overline{\mathcal{S}'(C)}_{>0}$ as elements in the closure of $\mathcal{S}'(C)$ in $\hat{\mathcal{S}}(C)$ with positive volume. An element in $\overline{\mathcal{S}'(C)}_{>0}$ is called an *almost semigroup* in C.

Recall that the volume here is defined in (3-1).

Our goal is to prove the following theorem:

Theorem 3.14 *The Okounkov body map $\Delta\colon \mathcal{S}'(C) \to \mathcal{K}_n$, as defined in Definition 3.7, admits a unique continuous extension*

$$\Delta\colon \overline{\mathcal{S}'(C)}_{>0} \to \mathcal{K}_n. \tag{3-4}$$

Moreover, for any $S \in \overline{\mathcal{S}'(C)}_{>0}$, we have

$$\operatorname{vol} S = \operatorname{vol} \Delta(S). \tag{3-5}$$

Proof The uniqueness of the extension is clear as long as it exists. Moreover, (3-5) follows easily from Theorem 3.8 and Theorem 2.3 by continuity. It remains to argue the existence of the continuous extension. We first construct an extension and prove its continuity.

Step 1 We construct the desired map (3-4). Let $S \in \overline{\mathcal{S}'(C)}_{>0}$. We wish to construct a convex body $\Delta(S) \in \mathcal{K}_n$.

Let $S^i \in \mathcal{S}'(C)$ be a sequence that converges to S such that

$$d(S^i, S^{i+1}) \le 2^{-i}.$$

For each $i, j \ge 0$, we introduce

$$S^{i,j} = S^i \cap S^{i+1} \cdots \cap S^{i+j}.$$

Then, by Lemma 3.2,

$$d(S^{i,j}, S^{i,j+1}) \le 2^{-i-j}.$$

Take $i_0 > 0$ large enough that, for $i \ge i_0$, $\operatorname{vol} S^i > 2^{-1} \operatorname{vol} S$ and $2^{2-i} < \operatorname{vol} S$ and hence

$$\operatorname{vol} S^i - \operatorname{vol} S^{i,j} \le d(S^{i,0}, S^{i,1}) + d(S^{i,1}, S^{i,2}) + \cdots + d(S^{i,j-1}, S^{i,j}) \le 2^{1-i}.$$

It follows that $\operatorname{vol} S^{i,j} > 2^{-1} \operatorname{vol} S - 2^{1-i} > 0$ whenever $i \ge i_0$. In particular, by Lemma 3.6, $S^{i,j} \in \mathcal{S}'(C)$ for $i \ge i_0$.

By Lemma 3.12, for $i \ge i_0$, there exists $T^i \in \mathcal{S}'(C)$ such that $S^{i,j} \to T^i$ as $j \to \infty$. Moreover,

$$d(T^i, S) = \lim_{j\to\infty} d(S^{i,j}, S) \le \lim_{j\to\infty} d(S^{i,j}, S^i) + d(S^i, S) \le 2^{1-i} + d(S^i, S).$$

Therefore, $T^i \to S$. We then define

$$\Delta(S) := \overline{\bigcup_{i=i_0}^{\infty} \Delta(T^i)}.$$

In other words, we have defined

$$\Delta(S) := \varliminf_{i\to\infty} \Delta(S^i).$$

This is an honest limit: if Δ is the limit of a subsequence of $\Delta(S^i)$, then $\Delta(S) \subseteq \Delta$ by (2-1). Comparing the volumes, we find that equality holds. So by Theorem 2.2,

$$\Delta(S) = \lim_{i\to\infty} \Delta(S^i). \tag{3-6}$$

Next we claim that $\Delta(S)$ as defined above does not depend on the choice of the sequence S^i. In fact, suppose that $S'^i \in \mathcal{S}'(C)$ is another sequence satisfying the same conditions as S^i. The same holds for $R^i := S^{i+1} \cap S'^{i+1}$. It follows that

$$\lim_{i\to\infty} \Delta(R^i) \subseteq \lim_{i\to\infty} \Delta(S^i).$$

Comparing the volumes, we find that equality holds. The same is true with S'^i in place of S^i. So we conclude that $\Delta(S)$ as in (3-6) does not depend on the choices we made.

Step 2 It remains to prove the continuity of Δ defined in Step 1. Suppose that $S^i \in \overline{\mathcal{S}'(C)}_{>0}$ is a sequence with limit $S \in \overline{\mathcal{S}'(C)}_{>0}$. We want to show that

$$\Delta(S^i) \xrightarrow{d_n} \Delta(S). \tag{3-7}$$

We first reduce to the case where $S^i \in \mathcal{S}'(C)$. By (3-6), for each i, we can choose $T^i \in \mathcal{S}'(C)$ such that $d(S^i, T^i) < 2^{-i}$ and $d_n(\Delta(S^i), \Delta(T^i)) < 2^{-i}$. If we have shown $\Delta(T^i) \xrightarrow{d_n} \Delta(S)$, then (3-7) follows immediately.

Next we reduce to the case where $d(S^i, S^{i+1}) \le 2^{-i}$. In fact, thanks to Theorem 2.2, in order to prove (3-7), it suffices to show that each subsequence of $\Delta(S^i)$ admits a subsequence that converges to $\Delta(S)$. Hence, we easily reduce to the required case.

After these reductions, (3-7) is nothing but (3-6). □

Corollary 3.15 *Suppose that $S, S' \in \overline{\mathcal{S}'(C)}_{>0}$ with $S \subseteq S'$, then*

$$\Delta(S) \subseteq \Delta(S'). \tag{3-8}$$

Proof Let $S^j, S'^j \in \mathcal{S}'(C)$ be elements such that $S^j \to S$ and $S'^j \to S'$. Then it follows from Lemma 3.2 that $S^j \cap S'^j \to S$. Since vol is continuous, for large j, $S^j \cap S'^j$ has positive volume and hence lies in $\mathcal{S}'(C)$ by Lemma 3.6. We may therefore replace S^j by $S^j \cap S'^j$ and assume that $S^j \subseteq S'^j$. Hence (3-8) follows from the continuity of Δ proved in Theorem 3.14. □

Remark 3.16 As the readers can easily verify, the construction of Δ is independent of the choice of C in the following sense: Suppose that C' is another cone satisfying the same assumptions as C and $C' \supseteq C$, then the Okounkov body map $\Delta\colon \overline{\mathcal{S}'(C')}_{>0} \to \mathcal{K}_n$ is an extension of the corresponding map (3-4). We will constantly use this fact without further explanations.

4 The metric on the space of singularity types

Let X be a connected compact Kähler manifold of dimension n and θ (resp. θ_i for $i = 1, \dots, n$) be a smooth real $(1,1)$-form representing a big $(1,1)$-cohomology class $[\theta]$ (resp. $[\theta_i]$). Let ω be a Kähler form on X.

In this section, we develop further the metric geometry on the space of singularity types of quasi-psh functions, initiated in [29] and studied further in [32].

As explained in [29, Section 3], one can introduce a pseudometric d_S on the set of singularity types of functions in $\mathrm{PSH}(X,\theta)$. In particular, d_S lifts to a pseudometric on $\mathrm{PSH}(X,\theta)$ as well. We do not recall the precise definition, as the following double inequality from [29, Proposition 3.5] will be enough for us. For any $\varphi, \psi \in \mathrm{PSH}(X,\theta)$ we have

$$d_S(\varphi,\psi) \le \sum_{i=0}^{n}\left(2\int_X \theta^i_{\max\{\varphi,\psi\}} \wedge \theta^{n-i}_{V_\theta} - \int_X \theta^i_\varphi \wedge \theta^{n-i}_{V_\theta} - \int_X \theta^i_\psi \wedge \theta^{n-i}_{V_\theta}\right) \le C_0 d_S(\varphi,\psi), \tag{4-1}$$

where $C_0 > 1$ is a constant depending only on n. In addition, $d_S(\varphi,\psi) = 0$ if and only if

$$C[\varphi] = C[\psi].$$

When there is a risk of confusion, we write $d_{S,\theta}$ instead of d_S.

Lemma 4.1 *Let $\varphi_i \in \mathrm{PSH}(X,\theta_i)$ for $i = 1,\dots,n$. Then*

$$\int_X \theta_{1,\varphi_1} \wedge \cdots \wedge \theta_{n,\varphi_n} = \int_X \theta_{1,C[\varphi_1]} \wedge \cdots \wedge \theta_{n,C[\varphi_n]}.$$

Proof From the definition (2-2), we have $[u] \preceq [C[u]]$, the $\le$ direction is obvious. For the reverse direction, recall that $C[\varphi_i] = \lim_{\epsilon\to 0+} P[(1-\epsilon)\varphi_i + \epsilon V_{\theta_i}]$. Thus, for $\epsilon \in (0,1)$,

$$\int_X \theta_{1,C[\varphi_1]} \wedge \cdots \wedge \theta_{n,C[\varphi_n]} \ge (1-\epsilon)^n \int_X \theta_{1,\varphi_1} \wedge \cdots \wedge \theta_{n,\varphi_n}.$$

Letting $\epsilon \to 0+$, we conclude. □

Theorem 4.2 *Let $\varphi_i^k, \varphi_i \in \mathrm{PSH}(X,\theta_i)$ for $i = 1,\dots,n$ and $k \in \mathbb{N}$. Assume that $\varphi_i^k \xrightarrow{d_{S,\theta_i}} \varphi_i$ as $k \to \infty$. Then*

$$\lim_{k\to\infty} \int_X \theta_{1,\varphi_1^k} \wedge \cdots \wedge \theta_{n,\varphi_n^k} = \int_X \theta_{1,\varphi_1} \wedge \cdots \wedge \theta_{n,\varphi_n}. \tag{4-2}$$

Proof By Lemma 4.1 and [29, Theorem 3.3], we may assume that φ_i^k and φ^k are model potentials.

Step 1 We assume that there is a constant $\delta > 0$ such that for all i and k,

$$\int_X \theta^n_{i,\varphi_i^k} > \delta.$$

In order to show (4-2), it suffices to prove that any subsequence of $\int_X \theta_{1,\varphi_1^k} \wedge \cdots \wedge \theta_{n,\varphi_n^k}$ has a converging subsequence with limit $\int_X \theta_{1,\varphi_1} \wedge \cdots \wedge \theta_{n,\varphi_n}$. Thus, by [29, Theorem 5.6], we may assume that for each fixed i, φ_i^k is either increasing or decreasing. We may assume that for $i \le i_0$, the sequence is decreasing and for $i > i_0$, the sequence is increasing.

Recall that in (4-2) the $\ge$ inequality always holds [26, Theorem 2.3], it suffices to prove

$$\varlimsup_{k\to\infty} \int_X \theta_{1,\varphi_1^k} \wedge \cdots \wedge \theta_{n,\varphi_n^k} \le \int_X \theta_{1,\varphi_1} \wedge \cdots \wedge \theta_{n,\varphi_n}. \tag{4-3}$$

By Witt Nyström's monotonicity theorem [53; 26], in order to prove (4-3), we may assume that for $j > i_0$, the sequences φ_j^k are constant. Thus, we are reduced to the case where for all i, the φ_i^k are decreasing.

In this case, for each i we may take an increasing sequence $b_i^k > 1$, tending to ∞, such that

$$(b_i^k)^n \int_X \theta^n_{i,\varphi_i} > ((b_i^k)^n - 1) \int_X \theta^n_{i,\varphi_i^k}.$$

Let ψ_i^k be the maximal θ_i-psh function such that

$$(b_i^k)^{-1}\psi_i^k + (1-(b_i^k)^{-1})\varphi_i^k \le \varphi_i,$$

whose existence is guaranteed by [29, Lemma 4.3].

Then by Witt Nyström's monotonicity theorem [53; 26] again,

$$\prod_{i=1}^{n}(1-(b_i^k)^{-1})\int_X \theta_{1,\varphi_1^k}\wedge\cdots\wedge\theta_{n,\varphi_n^k}\le\int_X \theta_{1,\varphi_1}\wedge\cdots\wedge\theta_{n,\varphi_n}.$$

Letting $k\to\infty$, we conclude (4-3).

Step 2 Now we deal with the general case.

We claim that if $t\in(0,1]$, then $(1-t)\varphi_i^k+tV_{\theta_i}\xrightarrow{d_S}(1-t)\varphi_i+tV_{\theta_i}$ as $k\to\infty$ for each i. From this and Step 1, we find that for $t_i\in(0,1]$,

$$\lim_{k\to\infty}\int_X \theta_{1,(1-t_1)\varphi_1^k+t_1V_{\theta_1}}\wedge\cdots\wedge\theta_{n,(1-t_n)\varphi_n^k+t_nV_{\theta_n}}=\int_X \theta_{1,(1-t_1)\varphi_1+t_1V_{\theta_1}}\wedge\cdots\wedge\theta_{n,(1-t_n)\varphi_n+t_nV_{\theta_n}}.$$

Thus, (4-2) follows, after letting $t_i\searrow 0$.

It remains to prove the claim. For simplicity, we suppress the i indices momentarily. We need to argue that

$$2\int_X \theta^j_{\max\{(1-t)\varphi^k+tV_\theta,(1-t)\varphi+tV_\theta\}}\wedge\theta^{n-j}_{V_\theta}-\int_X \theta^j_{(1-t)\varphi^k+tV_\theta}\wedge\theta^{n-j}_{V_\theta}-\int_X \theta^j_{(1-t)\varphi+tV_\theta}\wedge\theta^{n-j}_{V_\theta}\to 0.$$

Note that the above expression is a linear combination of terms of the following type:

$$2\int_X \theta^r_{\max\{\varphi^k,\varphi\}}\wedge\theta^{n-r}_{V_\theta}-\int_X \theta^r_{\varphi^k}\wedge\theta^{n-r}_{V_\theta}-\int_X \theta^r_{\varphi^k}\wedge\theta^{n-r}_{V_\theta}.$$

Thanks to (4-1), all these expressions tend to 0 as $k\to\infty$ since $\varphi^k\xrightarrow{d_S}\varphi$, which proves our claim. □

Corollary 4.3 *Let $\varphi^k,\varphi\in\mathrm{PSH}(X,\theta)$ for $k\in\mathbb{N}$. Let ω be a Kähler form on X. Assume that $\varphi^k\xrightarrow{d_{S,\theta}}\varphi$. Then $\varphi^k\xrightarrow{d_{S,\theta+\omega}}\varphi$.*

Proof It suffices to show that for each $j=0,\dots,n$, we have

$$2\int_X(\theta+\omega)^j_{\max\{\varphi^k,\varphi\}}\wedge(\theta+\omega)^{n-j}_{V_{\theta+\omega}}-\int_X(\theta+\omega)^j_{\varphi^k}\wedge(\theta+\omega)^{n-j}_{V_{\theta+\omega}}-\int_X(\theta+\omega)^j_{\varphi}\wedge(\theta+\omega)^{n-j}_{V_{\theta+\omega}}\to 0$$

as $k\to\infty$. Note that this quantity is a linear combination of terms of the form

$$2\int_X\theta^r_{\max\{\varphi^k,\varphi\}}\wedge\omega^{j-r}\wedge(\theta+\omega)^{n-j}_{V_{\theta+\omega}}-\int_X\theta^r_{\varphi^k}\wedge\omega^{j-r}\wedge(\theta+\omega)^{n-j}_{V_{\theta+\omega}}-\int_X\theta^r_{\varphi}\wedge\omega^{j-r}\wedge(\theta+\omega)^{n-j}_{V_{\theta+\omega}},$$

where $r=0,\dots,j$. By Theorem 4.2, it suffices to show that $\max\{\varphi,\varphi^k\}\xrightarrow{d_S}\varphi$. But this follows from [29, Proposition 3.5]. □

Corollary 4.4 *Let $\varphi\in\mathrm{PSH}(X,\theta)$ be an $\mathcal{I}$-model potential of positive mass. Let ω be a Kähler form on X. Then $P^{\theta+\omega}[\varphi]$ is $\mathcal{I}$-model.*

Proof By [33, Theorem 3.8], we may take a sequence φ^j with analytic singularities such that $\varphi^j\xrightarrow{d_{S,\theta}}\varphi$. Then $\varphi^j\xrightarrow{d_{S,\theta+\omega}}\varphi$ by Corollary 4.3. Thus, $P^{\theta+\omega}[\varphi]$ is $\mathcal{I}$-model. □

Corollary 4.5 *Let $\varphi^j, \varphi \in \mathrm{PSH}(X, \theta_1)$ and $\psi^j, \psi \in \mathrm{PSH}(X, \theta_2)$ for $j \in \mathbb{N}$. Assume that $\varphi^j \xrightarrow{d_{S,\theta_1}} \varphi$, $\psi^j \xrightarrow{d_{S,\theta_2}} \psi$. Then*

$$\varphi^j + \psi^j \xrightarrow{d_{S,\theta_1+\theta_2}} \varphi + \psi.$$

Proof Let $\theta = \theta_1 + \theta_2$. It suffices to show that for each $r = 0, \dots, n$,

$$2\int_X \theta^r_{\max\{\varphi^j+\psi^j,\varphi+\psi\}} \wedge \theta^{n-r}_{V_\theta} - \int_X \theta^r_{\varphi^j+\psi^j} \wedge \theta^{n-r}_{V_\theta} - \int_X \theta^r_{\varphi+\psi} \wedge \theta^{n-r}_{V_\theta} \to 0.$$

Observe that

$$\max\{\varphi^j + \psi^j, \varphi + \psi\} \le \max\{\varphi^j, \varphi\} + \max\{\psi^j, \psi\}.$$

Thus, it suffices to show that

$$2\int_X \theta^r_{\max\{\varphi^j,\varphi\}+\max\{\psi^j,\psi\}} \wedge \theta^{n-r}_{V_\theta} - \int_X \theta^r_{\varphi^j+\psi^j} \wedge \theta^{n-r}_{V_\theta} - \int_X \theta^r_{\varphi+\psi} \wedge \theta^{n-r}_{V_\theta} \to 0.$$

The left-hand side is a linear combination of

$$2\int_X \theta^a_{1,\max\{\varphi^j,\varphi\}} \wedge \theta^{r-a}_{2,\max\{\psi^j,\psi\}} \wedge \theta^{n-r}_{V_\theta} - \int_X \theta^a_{1,\varphi^j} \wedge \theta^{r-a}_{2,\psi^j} \wedge \theta^{n-r}_{V_\theta} - \int_X \theta^a_{1,\varphi} \wedge \theta^{r-a}_{2,\psi} \wedge \theta^{n-r}_{V_\theta}$$

with $a = 0, \dots, r$. Observe that $\max\{\varphi^j, \varphi\} \xrightarrow{d_S} \varphi$ and $\max\{\psi^j, \psi\} \xrightarrow{d_S} \psi$ by [29, Proposition 3.5], each term tends to 0 by Theorem 4.2. □

Finally, we prove the continuity of $P[\bullet]_{\mathcal{I}}$.

Theorem 4.6 *The map $\mathrm{PSH}(X, \theta)_{>0} \to \mathrm{PSH}(X, \theta)_{>0}$ given by $\varphi \mapsto P[\varphi]_{\mathcal{I}}$ is continuous with respect to the d_S-pseudometric.*

Here $\mathrm{PSH}(X, \theta)_{>0}$ denotes the subset of $\mathrm{PSH}(X, \theta)$ consisting of φ with $\int_X \theta^n_\varphi > 0$.

Proof Let $\varphi_i, \varphi \in \mathrm{PSH}(X, \theta)_{>0}$, with $\varphi_i \xrightarrow{d_S} \varphi$. We want to show that

$$P[\varphi_i]_{\mathcal{I}} \xrightarrow{d_S} P[\varphi]_{\mathcal{I}}. \tag{4-4}$$

We may assume that the φ_i and φ are all model potentials by [29, Theorem 3.3]. By [29, Theorem 5.6], we may assume that φ_i is either increasing or decreasing. These cases follow from [32, Lemma 2.21] and [29, Proposition 4.8, Lemma 4.1]. □

5 Partial Okounkov bodies

Let X be an irreducible smooth complex projective variety of dimension n and L be a big line bundle on X. Take a singular psh metric ϕ on L. We assume that $\mathrm{vol}(L, \phi) > 0$. Let h be a smooth Hermitian metric on L. Let $\theta = c_1(L, h)$. Then we can identify ϕ with a function $\varphi \in \mathrm{PSH}(X, \theta)$. We will use interchangeably the notations (θ, φ) and (L, ϕ).

For each $k \geq 0$,

$$W_k(\theta,\varphi) := \mathrm{H}^0(X, L^k \otimes \mathcal{I}(k\varphi)) \quad \text{and} \quad W(\theta,\varphi) := \bigoplus_{k=0}^{\infty} W_k(\theta,\varphi).$$

We omit (θ,φ) from our notations when there is no risk of confusion. Observe that $W_k(\theta,\varphi) \neq 0$ when k is large enough, as follows from Theorem 2.19.

Fix a rank n valuation $\nu\colon \mathbb{C}(X)^\times \to \mathbb{Z}^n$. We will write

$$\Gamma_{\nu,k}(\theta,\varphi) = \{k^{-1}\nu(s) \mid s \in W_k(\theta,\varphi)^\times\} \quad \text{for } k \geq 1,$$
$$\Gamma_\nu(\theta,\varphi) = \{(\nu(s),k) \mid k \in \mathbb{N}, s \in W_k(\theta,\varphi)^\times\}.$$

In [45], Lazarsfeld and Mustaţă only considered the case where ν is induced by an admissible flag, but thanks to Remark 2.8, their results can be easily extended to the current setup. We will use these results without further comments.

5.1 Construction of partial Okounkov bodies

Our goal in this section is to show that $\Gamma_\nu(\theta,\varphi) \in \overline{\mathcal{S}'(\Delta_\nu(L))}_{>0}$, namely it is an almost semigroup. Then we shall define

$$\Delta_\nu(\theta,\varphi) := \Delta(\Gamma_\nu(\theta,\varphi)) \tag{5-1}$$

using the theory of Okounkov bodies of almost semigroups developed in Section 3.3. Moreover, we have

$$\operatorname{vol} \Delta_\nu(\theta,\varphi) = \frac{1}{n!}\int_X \theta^n_{P[\varphi]_{\mathcal{I}}}. \tag{5-2}$$

5.1.1 The case of analytic singularities Assume that φ has analytic singularities and θ_φ is a Kähler current.

For any rational $\epsilon \geq 0$, we define

$$W_k^\epsilon = W_k^\epsilon(\theta,\varphi) := \{s \in \mathrm{H}^0(X, L^k) \mid |s|^2_{h^k} \mathrm{e}^{-k(1-\epsilon)\varphi} \text{ is bounded}\}. \tag{5-3}$$

Then $W^\epsilon := \bigoplus_{k=0}^\infty W_k^\epsilon$ has the property that

$$\Gamma_\nu(W^\epsilon) := \{(\nu(s),k) \mid k \in \mathbb{N}, s \in W_k^{\epsilon,\times}\} \in \mathcal{S}'(\Delta_\nu(L)). \tag{5-4}$$

To see this, we may assume that φ has analytic singularities along a $\mathbb{Q}$-divisor E. Then (5-4) follows from the fact that $L-(1-\epsilon)E$ is big, proved in [55, Lemma 2.4]; cf [45, Lemma 2.2].

For any $\epsilon \in \mathbb{Q}_{>0}$, we have that

$$W_k^0 \subseteq W_k \subseteq W_k^\epsilon \tag{5-5}$$

for k large enough depending on ϵ. The first inclusion is of course trivial. The second inclusion is widely known among experts. A detailed proof can be found in [33, Remark 2.9].

Let $\pi : Y \to X$ be a resolution such that $\pi^*\varphi$ has analytic singularities along a normal crossing $\mathbb{Q}$-divisor E. Then we have a natural identification for sufficiently divisible k,

$$W_k^\epsilon \cong \mathrm{H}^0(Y, \pi^* L^k \otimes \mathcal{O}_Y(-(1-\epsilon)kE)).$$

On the other hand,

$$W_k^0 \cong \mathrm{H}^0(Y, \pi^* L^k \otimes \mathcal{O}_Y(-kE)) \subseteq \mathrm{H}^0(Y, \pi^* L^k).$$

We compute the volumes

$$\text{(5-6)} \qquad \operatorname{vol} \Gamma_\nu(W^\epsilon) = \frac{1}{n!} \int_X \theta^n_{(1-\epsilon)\varphi} \quad \text{and} \quad \operatorname{vol} \Gamma_\nu(W^0) = \frac{1}{n!} \int_X \theta^n_\varphi.$$

It follows that $\Gamma_\nu(W^\epsilon) \to \Gamma_\nu(W^0)$ and $\Gamma_\nu(\theta, \varphi)$ is equivalent to $\Gamma_\nu(W^0)$. In particular, we get that $\Gamma_\nu(\theta, \varphi) \in \overline{\mathcal{S}'(\Delta_\nu(L))}_{>0}$, (5-1) makes sense and (5-2) holds.

Remark 5.1 It follows from the proof that if $W^0(\theta, \varphi)$ is defined as in (5-3) and (5-4):

$$W_k^0(\theta, \varphi) := \{s \in \mathrm{H}^0(X, L^k) \mid |s|^2_{h^k} \mathrm{e}^{-k\varphi} \text{ is bounded}\},$$

then

$$\text{(5-7)} \qquad \Delta(\Gamma_\nu(W^0(\theta, \varphi))) = \Delta_\nu(\theta, \varphi).$$

If we assume furthermore that $\pi^*\varphi$ has analytic singularity along some normal crossing $\mathbb{Q}$-divisor E on Y, then $\Delta_\nu(\theta, \varphi)$ is just the translation of $\Delta_\nu(\pi^* L - E)$ by $\nu(E)$.

5.1.2 The case of Kähler currents Now assume that θ_φ is Kähler current. Let $\varphi^j \in \mathrm{PSH}(X, \theta)$ be a quasi-equisingular approximation of φ. Then $\varphi^j \xrightarrow{d_S} P[\varphi]_{\mathcal{I}}$ by [33, Proposition 3.3].

In this case, we claim that

$$\text{(5-8)} \qquad \Gamma_\nu(\theta, \varphi^j) \to \Gamma_\nu(\theta, \varphi).$$

In fact, by Theorem 2.19, we have

$$\begin{aligned} d(\Gamma_\nu(\theta, \varphi^j), \Gamma_\nu(\theta, \varphi)) &= \varlimsup_{k\to\infty} k^{-n} \big(h^0(X, L^k \otimes \mathcal{I}(k\varphi^j)) - h^0(X, L^k \otimes \mathcal{I}(k\varphi)) \big) \\ &= \lim_{k\to\infty} k^{-n} h^0(X, L^k \otimes \mathcal{I}(k\varphi^j)) - \lim_{k\to\infty} k^{-n} h^0(X, L^k \otimes \mathcal{I}(k\varphi)) \\ &= \frac{1}{n!} \int_X \theta^n_{\varphi^j} - \frac{1}{n!} \int_X \theta^n_{P[\varphi]_{\mathcal{I}}}. \end{aligned}$$

Letting $j \to \infty$, we conclude (5-8) by Theorem 4.2.

Thus, $\Gamma_\nu(\theta, \varphi) \in \overline{\mathcal{S}'(\Delta_\nu(L))}_{>0}$ and (5-1) makes sense. By Theorem 3.14, we find that

$$\Delta_\nu(\theta, \varphi) = \bigcap_{j=0}^{\infty} \Delta_\nu(\theta, \varphi^j).$$

In particular, (5-2) holds.

5.1.3 General case Now we consider general φ with the assumption that $\int_X \theta^n_{P[\varphi]_{\mathcal{I}}} > 0$. We may replace φ with $P[\varphi]_{\mathcal{I}}$ and then assume that the nonpluripolar mass of φ is positive. Take a potential $\psi \in \mathrm{PSH}(X,\theta)$ such that $\psi \le \varphi$ and θ_ψ is a Kähler current. The existence of ψ is proved in [33, Proposition 3.6]. For each $\epsilon \in \mathbb{Q} \cap (0,1]$, let $\varphi_\epsilon = (1-\epsilon)\varphi + \epsilon\psi$. Then we have $W(\theta,\varphi_\epsilon) \subseteq W(\theta,\varphi)$. By (5-2),

$$\operatorname{vol} \Delta_\nu(\theta,\varphi_\epsilon) = \frac{1}{n!}\int_X \theta^n_{P[\varphi_\epsilon]_{\mathcal{I}}}.$$

We claim that

$$\Gamma_\nu(\theta,\varphi_\epsilon) \to \Gamma_\nu(\theta,\varphi).$$

In fact, this follows from the simple computation

$$\begin{aligned} d\big(\Gamma_\nu(\theta,\varphi_\epsilon),\Gamma_\nu(\theta,\varphi)\big) &= \varlimsup_{k\to\infty} k^{-n}\big(h^0(X,L^k\otimes\mathcal{I}(k\varphi)) - h^0(X,L^k\otimes\mathcal{I}(k\varphi_\epsilon))\big) \\ &= \lim_{k\to\infty} k^{-n}h^0(X,L^k\otimes\mathcal{I}(k\varphi)) - \lim_{k\to\infty} k^{-n}h^0(X,L^k\otimes\mathcal{I}(k\varphi_\epsilon)) \\ &= \frac{1}{n!}\int_X \theta^n_\varphi - \frac{1}{n!}\int_X \theta^n_{P[\varphi_\epsilon]_{\mathcal{I}}}. \end{aligned}$$

By [33, Proposition 2.7], as ϵ decreases to 0, $P[\varphi_\epsilon]_{\mathcal{I}}$ increases to $P[\varphi]_{\mathcal{I}} = \varphi$ a.e., which implies the d_S-convergence by [29, Lemma 4.1]. Therefore, the right-hand side of the above equation converges to 0 by Theorem 4.2. Our claim is proved. It follows that $\Gamma_\nu(\theta,\varphi) \in \overline{\mathcal{S}'(\Delta_\nu(L))}_{>0}$ and (5-1) makes sense. By Theorem 3.14,

$$\Delta_\nu(\theta,\varphi) = \overline{\bigcup_{\epsilon>0} \Delta_\nu(\theta,\varphi_\epsilon)}.$$

It remains to verify (5-2):

$$\operatorname{vol} \Delta_\nu(\theta,\varphi) = \frac{1}{n!}\lim_{\epsilon\to 0+}\int_X \theta^n_{P[\varphi_\epsilon]_{\mathcal{I}}} = \frac{1}{n!}\int_X \theta^n_{P[\varphi]_{\mathcal{I}}}.$$

Definition 5.2 Assume that $\varphi \in \mathrm{PSH}(X,\theta)$, where $\int_X \theta^n_{P[\varphi]_{\mathcal{I}}} > 0$. We call $\Delta_\nu(\theta,\varphi)$ the *partial Okounkov body* of (L,ϕ) or of (θ,φ) with respect to ν. When ν is induced by an admissible flag $(Y_\bullet)$ on X (see Definition 2.7), we also say that $\Delta_\nu(\theta,\varphi)$ the *partial Okounkov body* of (L,ϕ) or of (θ,φ) with respect to $(Y_\bullet)$. In this case, we also write $\Delta_{Y_\bullet}$ instead of Δ_ν.

We use interchangeably the notations $\Delta_\nu(\theta,\varphi)$ and $\Delta_\nu(L,\phi)$. When there is no risk of confusion, we write Δ instead of Δ_ν or $\Delta_{Y_\bullet}$.

Remark 5.3 We have assumed X to be smooth only for simplicity. All of our constructions work equally well when X is normal or merely unibranch, based on the pluripotential theory in these settings developed in [54].

Remark 5.4 In the transcendental setting, a theory of Okounkov bodies was recently established in [31] based on the work of [37]. The proof of the volume identity of transcendental Okounkov bodies relies on the technique of partial Okounkov bodies developed in this paper. The transcendental analogue of the partial Okounkov bodies is constructed in a forthcoming joint paper with T Darvas.

5.2 Basic properties of partial Okounkov bodies

We first show that $\Delta(\theta,\varphi)$ does not depend on the explicit choices of L, h and φ, it just depends on $\mathrm{dd}^{\mathrm{c}}\phi$.

Lemma 5.5 *Let L' be another big line bundle on X. Let h' be a smooth Hermitian metric on L' with $c_1(L,h)=c_1(L',h')$. Then $\Delta(\theta,\varphi)$ defined with respect to (L,h) is the same as the one defined with respect to (L',h').*

Proof From our construction, we may assume that θ_φ is a Kähler current and φ has analytic singularities. After taking a birational resolution, it suffices to deal with the case where φ has analytic singularities along normal crossing $\mathbb{Q}$-divisors E. By rescaling, we may also assume that E is a divisor. By Remark 5.1, we further reduce to the case without the singular potential ϕ.

In this case, the assertion is proved in [45, Proposition 4.1]. □

Lemma 5.6 *Let h' be another smooth Hermitian metric on L. Set $\theta'=c_1(L,h')$. Write $\mathrm{dd}^{\mathrm{c}} f=\theta-\theta'$. Let $\varphi'=\varphi+f\in\mathrm{PSH}(X,\theta')$. Then*

$$\Delta(\theta,\varphi)=\Delta(\theta',\varphi'). \tag{5-9}$$

Proof This is obvious as $W(\theta,\varphi)=W(\theta',\varphi')$. □

Corollary 5.7 *The partial Okounkov body $\Delta(L,\phi)$ depends only on $\mathrm{dd}^{\mathrm{c}}\phi$, not on the explicit choices of L, ϕ and h.*

Thanks to this result, given a closed positive $(1,1)$-current $T\in c_1(L)$ on X with $\int_X T^n>0$, we can define $\Delta(T)$ as $\Delta(\theta,\varphi)$ if $T=\theta+\mathrm{dd}^{\mathrm{c}}\varphi$ for some $\varphi\in\mathrm{PSH}(X,\theta)$.

Proof This is a direct consequence of Lemmas 5.5 and 5.6. □

Let $\mathrm{PSH}(X,\theta)_{>0}$ denote the subset of $\mathrm{PSH}(X,\theta)$ consisting of potentials φ such that $\int_X \theta_\varphi^n>0$.

Proposition 5.8 *Let $\varphi,\psi\in\mathrm{PSH}(X,\theta)_{>0}$. Assume that $\varphi\preceq_{\mathcal{I}}\psi$. Then*

$$\Delta(\theta,\varphi)\subseteq\Delta(\theta,\psi). \tag{5-10}$$

In particular, as by definition, $\Delta(\theta,V_\theta)=\Delta(L)$, we have

$$\Delta(\theta,\varphi)\subseteq\Delta(L).$$

Proof This follows from Corollary 3.15. □

Theorem 5.9 *The Okounkov body map*

$$\Delta(\theta,\bullet)\colon(\mathrm{PSH}(X,\theta)_{>0},d_S)\to(\mathcal{K}_n,d_n)$$

is continuous.

Remark 5.10 On the other hand, it is of interest to understand the dependence of $\Delta(\theta,\bullet)$ on ν as well. For some preliminary results and anticipations in the usual Okounkov body setting, see [1]. In particular, see [1, Conjecture 10.1] for a concrete continuity conjecture.

Proof Let $\varphi_j \to \varphi$ be a d_S-convergent sequence in $\mathrm{PSH}(X,\theta)_{>0}$. We want to show that

$$\Delta(\theta,\varphi_j) \xrightarrow{d_n} \Delta(\theta,\varphi). \tag{5-11}$$

By Proposition 5.8 and [29, Theorem 3.3], we may assume that all the φ_j and φ are model potentials. By Theorem 2.2 and [29, Theorem 5.6], we may assume that φ_j is either decreasing or increasing. By Theorem 4.6, we may further assume that the φ_j are $\mathcal{I}$-model. In both cases, we claim that $\Gamma_\nu(\theta,\varphi_j) \to \Gamma_\nu(\theta,\varphi)$. In fact, we can compute their distance as

$$d\big(\Gamma_\nu(\theta,\varphi_j),\Gamma_\nu(\theta,\varphi)\big) = \varlimsup_{k\to\infty} k^{-n}|h^0(X,L^k\otimes\mathcal{I}(k\varphi_j)) - h^0(X,L^k\otimes\mathcal{I}(k\varphi))| = \frac{1}{n!}\left|\int_X \theta_{\varphi_j}^n - \int_X \theta_\varphi^n\right|,$$

where we applied Theorem 2.19 at the last step. Then Theorem 4.2 implies our claim. Hence, (5-11) follows from Theorem 3.14. □

Although $W(\theta,\varphi)$ and $\Gamma_\nu(\theta,\varphi)$ are not birationally invariant, we could still show that the Okounkov body is.

Proposition 5.11 *Let $\pi : Y \to X$ be a birational resolution. Let (L,ϕ) be a Hermitian big line bundle on X with positive volume. Then*

$$\Delta(\pi^*L,\pi^*\phi) = \Delta(L,\phi).$$

Here we are using the same valuation ν on the function field $\mathbb{C}(Y) = \mathbb{C}(X)$ of Y.

Proof By Definition 2.12(3), $P_\theta[\bullet]_{\mathcal{I}}$ commutes with birational pullbacks, we may assume that φ is $\mathcal{I}$-model. By [33, Theorem 3.8], we can find a sequence $\varphi^j \in \mathrm{PSH}(X,\theta)$ with analytic singularities such that $\varphi^j \xrightarrow{d_S} \varphi$. It follows from (4-1) that $\pi^*\varphi^j \xrightarrow{d_S} \pi^*\varphi$. By Theorem 5.9, we may then reduce to the case where φ has analytic singularities. In this case, up to replacing Y by a further sequences of blowups, we may assume that $\pi^*\varphi$ has analytic singularities along a normal crossing $\mathbb{Q}$-divisor D. It suffices to apply Remark 5.1. □

Next we prove the Brunn–Minkowski inequality.

Proposition 5.12 *Let (L,ϕ), (L',ϕ') be Hermitian big line bundles on X of positive volumes. Then*

$$(\operatorname{vol}\Delta(L+L',\phi+\phi'))^{1/n} \ge (\operatorname{vol}\Delta(L,\phi))^{1/n} + (\operatorname{vol}\Delta(L',\phi'))^{1/n}.$$

Proof This follows from Corollary 2.20. □

Proposition 5.13 *Let (L', ϕ') be another Hermitian big line bundle on X with positive volume. Then*

$$\Delta(L,\phi)+\Delta(L',\phi')\subseteq\Delta(L\otimes L',\phi\otimes\phi').$$

Proof Take a smooth metric h' on L', and let $\theta'=c_1(L',h')$. We identify ϕ' with $\varphi'\in\mathrm{PSH}(X,\theta')$. Then we need to show

$$\Delta(\theta,\varphi)+\Delta(\theta',\varphi')\subseteq\Delta(\theta+\theta',\varphi+\varphi'). \tag{5-12}$$

By [33, Theorem 3.8], we can find $\varphi^j\in\mathrm{PSH}(X,\theta)$ and $\varphi'^j\in\mathrm{PSH}(X,\theta')$ such that

(1) φ^j and φ'^j both have analytic singularities and have positive masses,

(2) $\varphi^j\xrightarrow{d_S}\varphi$ and $\varphi'^j\xrightarrow{d_S}\varphi'$.

Then $\varphi^j+\varphi'^j\in\mathrm{PSH}(X,\theta+\theta')$ and $\varphi^j+\varphi'^j\xrightarrow{d_S}\varphi+\varphi'$ by Corollary 4.5. Thus, by Theorem 5.9, we may assume that φ and ψ both have analytic singularities. Taking a birational resolution, we may further assume that they have analytic singularities along some normal crossing divisors. By Remark 5.1, we reduce to the case without singularities, in which case the result is well-known; see for example [45, Proof of Corollary 4.12]. □

Theorem 5.14 *Let $\varphi,\psi\in\mathrm{PSH}(X,\theta)_{>0}$. Then for any $t\in(0,1)$,*

$$\Delta(\theta,t\varphi+(1-t)\psi)\supseteq t\Delta(\theta,\varphi)+(1-t)\Delta(\theta,\psi). \tag{5-13}$$

Proof We may assume that t is rational as a consequence of Theorem 5.9. Similarly, by [33, Theorem 3.8], we could reduce to the case where both φ and ψ have analytic singularities. Taking a resolution, we may assume that φ (resp. ψ) has analytic singularities along a normal crossing $\mathbb{Q}$-divisor E (resp. E'). In this case, let $N>0$ be an integer such that Nt is an integer. Then for any $s\in W^0_k(\theta,\varphi)$ and $r\in W^0_k(\theta,\psi)$, we have

$$(s^tr^{1-t})^N\in W^0_{Nk}(\theta,t\varphi+(1-t)\psi).$$

By Theorem 3.8, (5-13) follows. □

Proposition 5.15 *For any integer $a>0$,*

$$\Delta(a\theta,a\varphi)=a\Delta(\theta,\varphi).$$

Proof By Theorem 5.9, it suffices to treat the case where φ has analytic singularities. Taking a birational resolution, we may assume that φ has analytic singularities along a normal crossing $\mathbb{Q}$-divisor E. By Remark 5.1, we reduce to the case without the singularity φ, which is already proved in [45]. □

In particular, if T is a closed positive $(1,1)$-current on X with $\int_X T^n>0$ and such that the cohomology class of T lies in the Néron–Severi group with rational coefficients, then we can define $\Delta(T)$ as $a^{-1}\Delta(aT)$ for a sufficiently divisible positive integer a.

We also need the following perturbation. Let A be an ample line bundle on X. Fix a smooth Hermitian metric h_A on A such that $\omega := c_1(A, h_A)$ is a Kähler form on X. Then for any $\delta \in \mathbb{Q}_{>0}$, we can define

$$\Delta(\theta + \delta\omega, \varphi) := \Delta(\theta + \delta\omega + \mathrm{dd^c}\varphi) = C^{-1}\Delta(C\theta + C\delta\omega, C\varphi),$$

where $C \in \mathbb{N}_{>0}$ is any integer so that $C\delta \in \mathbb{N}$.

Proposition 5.16 *Under the above assumptions, as $\delta \in \mathbb{Q}_{>0}$ decreases to 0, $\Delta(\theta + \delta\omega, \varphi)$ is decreasing under inclusion with Hausdorff limit $\Delta(\theta, \varphi)$.*

Proof Let $0 \le \delta < \delta'$ be two rational numbers. Take $C \in \mathbb{N}_{>0}$ divisible enough, so that $C\delta$ and $C\delta'$ are both integers. Then by Proposition 5.13,

$$\Delta(C\theta + C\delta\omega, C\varphi) \subseteq \Delta(C\theta + C\delta'\omega, C\varphi).$$

It follows that

$$\Delta(\theta + \delta\omega, \varphi) \subseteq \Delta(\theta + \delta'\omega, \varphi).$$

On the other hand,

$$\operatorname{vol}\Delta(\theta + \delta\omega, \varphi) = \frac{1}{n!}\int_X (\theta + \delta\omega)^n_{P^{\theta+\delta\omega}[\varphi]_{\mathcal{I}}} = \frac{1}{n!}\int_X (\theta + \delta\omega)^n_{P^{\theta}[\varphi]_{\mathcal{I}}},$$

where we applied Corollary 4.4. As $\delta \to 0+$, the right-hand side converges to

$$\operatorname{vol}\Delta(\theta, \varphi) = \frac{1}{n!}\int_X \theta^n_{P^{\theta}[\varphi]_{\mathcal{I}}}.$$

It follows that

$$\Delta(\theta, \varphi) = \bigcap_{\delta \in \mathbb{Q}_{>0}} \Delta(\theta + \delta\omega, \varphi). \qquad \square$$

5.3 The Hausdorff convergence property of partial Okounkov bodies

For each $k \in \mathbb{Z}_{>0}$, we introduce

$$\Delta_k(\theta, \varphi) := \operatorname{Conv}\{k^{-1}\nu(f) \mid f \in \mathrm{H}^0(X, L^k \otimes \mathcal{I}(k\varphi))^\times\} \subseteq \mathbb{R}^n.$$

Here Conv denotes the convex hull. The convex hull is a polytope if it is nonempty by [45, Lemma 1.4]. For large enough $\Delta_k(\theta, \varphi)$ is nonempty thanks to Theorem 2.19.

For later use, we introduce a twisted version as well. If T is a holomorphic line bundle on X, we introduce

$$\Delta_{k,T}(\theta, \varphi) := \operatorname{Conv}\{k^{-1}\nu(f) \mid f \in \mathrm{H}^0(X, T \otimes L^k \otimes \mathcal{I}(k\varphi))^\times\} \subseteq \mathbb{R}^n.$$

We also write

$$\Delta_{k,T}(L) := \operatorname{Conv}\{k^{-1}\nu(f) \mid f \in \mathrm{H}^0(X, T \otimes L^k)^\times\} \subseteq \mathbb{R}^n,$$
$$\Delta_k(L) := \operatorname{Conv}\{k^{-1}\nu(f) \mid f \in \mathrm{H}^0(X, L^k)^\times\} \subseteq \mathbb{R}^n.$$

We write $\mathcal{I}_\infty(\varphi) = \mathcal{I}_\infty(\phi)$ for the ideal sheaf on X locally consisting of holomorphic functions f such that $|f|_\phi$ is locally bounded.

The main result is the following:

Theorem 5.17 (Hausdorff convergence property) *Let T be a holomorphic line bundle on X. As $k \to \infty$, we have $\Delta_{k,T}(\theta,\varphi) \xrightarrow{d_n} \Delta(\theta,\varphi)$.*

Although we are only interested in the untwisted case, the proof given below requires twisted case.

We first extend Theorem 3.8 to the twisted case.

Proposition 5.18 *For any holomorphic line bundle T on X,*

$$\Delta_{k,T}(L) \xrightarrow{d_n} \Delta(L) \quad \text{as } k \to \infty.$$

Proof As L is big, we can take $k_0 \in \mathbb{Z}_{>0}$ so that

(1) $T^{-1} \otimes L^{k_0}$ admits a nonzero global holomorphic section s_0,

(2) $T \otimes L^{k_0}$ admits a nonzero global holomorphic section s_1.

Then for $k \in \mathbb{Z}_{>k_0}$, we have injective linear maps

$$\mathrm{H}^0(X, L^{k-k_0}) \xrightarrow{\times s_1} \mathrm{H}^0(X, T \otimes L^k) \xrightarrow{\times s_0} \mathrm{H}^0(X, L^{k+k_0}).$$

It follows that

$$(k-k_0)\Delta_{k-k_0}(L) + \nu(s_1) \subseteq k\Delta_{k,T}(L) \subseteq (k+k_0)\Delta_{k+k_0}(L) - \nu(s_0).$$

By Theorem 3.8, we conclude. □

Lemma 5.19 *Let T be a holomorphic line bundle on X. Assume that φ has analytic singularities and θ_φ is a Kähler current. Then as $k \to \infty$,*

$$\Delta_{k,T}(\theta,\varphi) \xrightarrow{d_n} \Delta(\theta,\varphi).$$

Proof Up to replacing X by a birational model and twisting T accordingly, we may assume that φ has analytic singularities along a normal crossing $\mathbb{Q}$-divisor D; cf Proposition 5.11. Take $\epsilon \in (0,1) \cap \mathbb{Q}$. In this case, as in (5-5), for large enough $k \in \mathbb{Z}_{>0}$ we have

$$\mathrm{H}^0(X, T \otimes L^k \otimes \mathcal{I}_\infty(k\varphi)) \subseteq \mathrm{H}^0(X, T \otimes L^k \otimes \mathcal{I}(k\varphi)) \subseteq \mathrm{H}^0(X, T \otimes L^k \otimes \mathcal{I}_\infty(k(1-\epsilon)\varphi)).$$

Take an integer $N \in \mathbb{Z}_{>0}$ so that ND is a divisor and $N\epsilon$ is an integer.

Let Δ' be the limit of a subsequence of $(\Delta_{k,T}(\theta,\varphi))_k$, say the sequence defined by the indices $k_1, k_2, \ldots$. We want to show that $\Delta' = \Delta(\theta,\varphi)$.

There exists $t \in \{0,1,\ldots,N-1\}$ such that $k_i \equiv t$ modulo N for infinitely many i, up to replacing k_i by a subsequence, we may assume that $k_i \equiv t$ modulo N for all i. Write $k_i = Ng_i + t$. Then

$$\begin{aligned}\mathrm{H}^0(X, T \otimes L^{-N+t} \otimes L^{N(g_i+1)} \otimes \mathcal{I}_\infty(N(g_i+1)\varphi)) &\subseteq \mathrm{H}^0(X, T \otimes L^{k_i} \otimes \mathcal{I}(k_i\varphi)) \\ &\subseteq \mathrm{H}^0(X, T \otimes L^t \otimes L^{Ng_i} \otimes \mathcal{I}_\infty(g_iN(1-\epsilon)\varphi)).\end{aligned}$$

So

$$(g_i+1)\Delta_{g_i+1,T\otimes L^{-N+t}}(NL-ND)+N(g_i+1)\nu(D)$$
$$\subseteq (Ng_i+t)\Delta_{k,T}(\theta,\varphi)\subseteq g_i\Delta_{g_i,T\otimes L^t}(NL-N(1-\epsilon)D)+Ng_i(1-\epsilon)\nu(D).$$

Letting $i\to\infty$, by Proposition 5.18,

$$\Delta(L-D)+\nu(D)\subseteq\Delta'\subseteq\Delta(L-(1-\epsilon)D)+(1-\epsilon)\nu(D).$$

Letting $\epsilon\to 0+$, we find that

$$\Delta(L-D)+\nu(D)=\Delta'.$$

It follows from Theorem 2.2 that

$$\Delta_{k,T}(\theta,\varphi)\xrightarrow{d_n}\Delta(L-D)+\nu(D)=\Delta(\theta,\varphi)\quad\text{as } k\to\infty.$$ □

Lemma 5.20 *Assume that θ_φ is a Kähler current. Then as $\beta\to 0+$ with $\beta\in\mathbb{Q}$, we have*

$$\Delta((1-\beta)\theta,\varphi)\to\Delta(\theta,\varphi).$$

Proof By Proposition 5.13, we have

$$\Delta((1-\beta)\theta,\varphi)+\beta\Delta(L)\subseteq\Delta(\theta,\varphi).$$

In particular, if Δ' is a limit of a subsequence of $(\Delta((1-\beta)\theta,\varphi))_\beta$, then

$$\Delta'\subseteq\Delta(\theta,\varphi).$$

But

$$\operatorname{vol}\Delta'=\lim_{\beta\to 0+}\Delta((1-\beta)\theta,\varphi)=\lim_{\beta\to 0+}\int_X((1-\beta)\theta+\mathrm{dd^c}P^{(1-\beta)\theta}[\varphi]_{\mathcal{I}})^n=\int_X(\theta+\mathrm{dd^c}P^{\theta}[\varphi]_{\mathcal{I}})^n,$$

where the last step follows easily from [56, Theorem 0.6]. It follows that $\Delta'=\Delta(\theta,\varphi)$. We conclude by Theorem 2.2. □

Proof of Theorem 5.17 Fix a Kähler form $\omega\geq\theta$ on X.

Step 1 We first handle the case where θ_φ is a Kähler current, say $\theta_\varphi\geq\beta_0\omega$ for some $\beta_0\in(0,1)$.

Take a decreasing quasi-equisingular approximation φ_j of φ. Up to replacing β_0 by $\beta_0/2$, we may assume that $\theta_{\varphi_j}\geq\beta_0\omega$ for all $j\geq 1$.

Let Δ' be a limit of a subsequence of $(\Delta_{k,T}(\theta,\varphi))_k$. Let us say the indices of the subsequence are $k_1<k_2<\cdots$. By Theorem 2.2, it suffices to show that $\Delta'=\Delta(\theta,\varphi)$.

As $[\varphi]\preceq[\varphi_j]$ for each $j\geq 1$, we have $\Delta'\subseteq\Delta(\theta,\varphi_j)$ by Lemma 5.19. Letting $j\to\infty$, we find

$$\Delta'\subseteq\Delta(\theta,\varphi).$$

In particular, it suffices to prove that

$$\operatorname{vol}\Delta'\geq\operatorname{vol}\Delta(\theta,\varphi).$$

Take $\beta \in (0,\beta_0)\cap\mathbb{Q}$. Write $\beta = p/q$ with $p,q\in\mathbb{Z}_{>0}$. Observe that for any $j\geq 1$,

$$\theta_{\varphi_j} \geq \beta\omega \geq \beta\theta.$$

Namely, $\varphi_j \in \mathrm{PSH}(X,(1-\beta)\theta)$. Similarly, $\varphi\in\mathrm{PSH}(X,(1-\beta)\theta)$. By Lemma 5.20, it suffices to argue that

$$\operatorname{vol}\Delta' \geq \operatorname{vol}\Delta((1-\beta)\theta,\varphi). \tag{5-14}$$

For this purpose, we are free to replace the k_i by a subsequence, so we may assume that $k_i\equiv a$ modulo q for all $i\geq 1$, where $a\in\{0,1,\dots,q-1\}$. We write $k_i = g_iq+a$. Observe that for each $i\geq 1$,

$$\mathrm{H}^0(X, T\otimes L^{k_i}\otimes\mathcal{I}(k_i\varphi)) \supseteq \mathrm{H}^0(X, T\otimes L^{-q+a}\otimes L^{g_iq+q}\otimes\mathcal{I}((g_iq+q)\varphi)).$$

Up to replacing T by $T\otimes L^{-q+a}$, we may therefore assume that $a=0$.

By [33, Lemma 4.2], we can find $k'\in\mathbb{Z}_{>0}$ such that for all $k\geq k'$, there is a $v_{\beta,k}\in\mathrm{PSH}(X,\theta)$ such that

(1) $P[\varphi]_{\mathcal{I}} \geq (1-\beta)\varphi_k + \beta v_{\beta,k}$,

(2) $v_{\beta,k}$ has positive mass.

Fix $k\geq k'$. It suffices to show that

$$\Delta((1-\beta)\theta,\varphi_k) + v' \subseteq \Delta' \tag{5-15}$$

for some $v'\in\mathbb{R}^n$. In fact, if this is true, we have

$$\operatorname{vol}\Delta' \geq \operatorname{vol}\Delta((1-\beta)\theta,\varphi_k).$$

Letting $k\to\infty$ and applying Theorem 5.9, we conclude (5-14).

It remains to prove (5-15). We will fix $k\geq k'$. Let $\pi: Y\to X$ be a log resolution of the singularities of φ_k. By the proof of [33, Proposition 4.3], there is $j_0=j_0(\beta,k)\in\mathbb{Z}_{>0}$ such that for any $j\geq j_0$, we can find a nonzero section $s_j\in\mathrm{H}^0(Y,\pi^*L^{pj}\otimes\mathcal{I}(jp\pi^*v_{\beta,k}))$ such that we get an injective linear map

$$\mathrm{H}^0(Y,\pi^*T\otimes K_{Y/X}\otimes\pi^*L^{(q-p)j}\otimes\mathcal{I}(jq\pi^*\varphi_k)) \xrightarrow{\times s_j} \mathrm{H}^0(X, T\otimes L^{jq}\otimes\mathcal{I}(jq\varphi)).$$

In particular, when $j=k_i$ for some i large enough, we then find

$$\Delta_{k_i,\pi^*T\otimes K_{Y/X}}((1-\beta)q\pi^*\theta, q\pi^*\varphi_k) + (k_i)^{-1}\nu(s_{k_i}) \subseteq q\Delta_{k_i,T}(\theta,\varphi).$$

We observe that $(k_i)^{-1}\nu(s_{k_i})$ is bounded as both convex bodies appearing in this equation are bounded when i varies. Then by Lemma 5.19, there is a vector $v'\in\mathbb{R}^n$ such that

$$\Delta((1-\beta)\pi^*\theta,\pi^*\varphi_k) + v' \subseteq \Delta'.$$

By Proposition 5.11, we find (5-15).

Step 2 Next we handle the general case.

Let Δ' be the limit of a subsequence of $(\Delta_{k,T}(\theta,\varphi))_k$, say the subsequence with indices $k_1<k_2<\cdots$. By Theorem 2.2, it suffices to prove that $\Delta'=\Delta(\theta,\varphi)$.

Take $\psi \in \mathrm{PSH}(X,\theta)$ such that

(1) θ_ψ is a Kähler current,

(2) $\psi \leq \varphi$.

The existence of ψ is proved in [33, Proposition 3.6].

Then for any $\epsilon \in \mathbb{Q} \cap (0,1)$,

$$\Delta_{k,T}(\theta,\varphi) \supseteq \Delta_{k,T}(\theta,(1-\epsilon)\varphi+\epsilon\psi)$$

for all k. It follows from Step 1 that

$$\Delta' \supseteq \Delta(\theta,(1-\epsilon)\varphi+\epsilon\psi).$$

Letting $\epsilon \to 0$ and applying Theorem 5.9, we have $\Delta' \supseteq \Delta(\theta,\varphi)$. It remains to establish that

$$\operatorname{vol}\Delta' \leq \operatorname{vol}\Delta(\theta,\varphi). \tag{5-16}$$

For this purpose, we are free to replace $k_1 < k_2 < \cdots$ by a subsequence. Fix $q > 0$, we may then assume that $k_i \equiv a$ modulo q for all $i \geq 1$ for some $a \in \{0,1,\dots,q-1\}$. We write $k_i = g_i q + a$. Observe that

$$\mathrm{H}^0(X, T \otimes L^{k_i} \otimes \mathcal{I}(k_i\varphi)) \subseteq \mathrm{H}^0(X, T \otimes L^a \otimes L^{g_i q} \otimes \mathcal{I}(g_i q\varphi)).$$

Up to replacing T by $T \otimes L^a$, we may assume that $a = 0$.

Take a very ample line bundle H on X and fix a Kähler form $\omega \in c_1(H)$, and take a nonzero section $s \in \mathrm{H}^0(X,H)$.

We have an injective linear map

$$\mathrm{H}^0(X, T \otimes L^{jq} \otimes \mathcal{I}(jq\varphi)) \xrightarrow{\times s^j} \mathrm{H}^0(X, T \otimes H^j \otimes L^{jq} \otimes \mathcal{I}(jq\varphi))$$

for each $j \geq 1$. In particular, for each $i \geq 1$,

$$k_i \Delta_{k_i,T}(q\theta,q\varphi) + k_i \nu(s) \subseteq k_i \Delta_{k_i,T}(\omega+q\theta,q\varphi).$$

Letting $i \to \infty$, by Step 1, we have

$$q\Delta' + \nu(s) \subseteq \Delta(\omega+q\theta,q\varphi).$$

So

$$\operatorname{vol}\Delta' \leq \operatorname{vol}\Delta(q^{-1}\omega+\theta,\varphi) = \int_X (q^{-1}\omega+\theta+\mathrm{dd^c} P^{q^{-1}\omega+\theta}[\varphi]_{\mathcal{I}})^n.$$

By Corollary 4.4,

$$\operatorname{vol}\Delta' \leq \int_X (q^{-1}\omega+\theta+\mathrm{dd^c} P^{\theta}[\varphi]_{\mathcal{I}})^n.$$

Letting $q \to \infty$, we conclude (5-16). □

Theorem 5.21 *The Okounkov body $\Delta(L,\phi)$ is independent of the choice of a very general flag in a family of admissible flags.*

Proof By Theorem 5.17, it suffices to show that $\Delta_k(W(\theta,\varphi))$ is independent of the choice of a very general flag. For this purpose, we may assume that $k=1$.

Let T be an irreducible component of the moduli space of admissible flags. Let

$$X\times T=\mathcal{Y}_0\supseteq\cdots\supseteq\mathcal{Y}_n$$

be the universal flag. The Hermitian line bundle (L,ϕ) pulls back to $(\mathcal{L},\Phi)$ on $X\times T$. We denote quantities at the fiber at $t\in T$ by a subindex t.

We claim that for each $\sigma\in\mathbb{N}^n$, there is a proper Zariski closed set $\Sigma\subsetneq T$, so that

$$\dim \mathrm{H}^0(X_t,L_t\otimes\mathcal{I}(\phi_t))^{\geq\sigma}$$

are constants for $t\in T\setminus\Sigma$, where $\mathrm{H}^0(X_t,L_t\otimes\mathcal{I}(\phi_t))^{\geq\sigma}$ is the space of sections in $\mathrm{H}^0(X_t,L_t\otimes\mathcal{I}(\phi_t))$ with valuations no less than σ.

Let $\mathcal{L}^{\geq\sigma}$ be the coherent subsheaf of $\mathcal{L}$ introduced in [45, Remark 1.6]. After possibly shrinking T, we may guarantee that $\mathcal{L}^{\geq\sigma}\otimes\mathcal{I}(\Phi)$ is flat over T. By further shrinking T, we may guarantee that

$$t\mapsto\dim\mathrm{H}^0\big(X_t,(\mathcal{L}^{\geq\sigma}\otimes\mathcal{I}(\Phi))|_{X_t}\big)$$

is constant. Observe that

$$(\mathcal{L}^{\geq\sigma}\otimes\mathcal{I}(\Phi))|_{X_t}\cong L_t^{\geq\sigma}\otimes\mathcal{I}(\phi).$$

Thus, our claim follows.

From this claim, it follows that the images of $\Gamma_k(W(L,\phi))$ are independent of the choice of a very general flag $(Y_\bullet)$ as [45, Proof of Theorem 5.1]. Thus, $\Delta(W(L,\phi))$ is independent of the choice of a very general flag. □

5.4 Recover Lelong numbers from partial Okounkov bodies

Lemma 5.22 *Let $\varphi\in\mathrm{PSH}(X,\theta)$ be such that θ_φ is a Kähler current. Let φ^j be a quasi-equisingular approximation of φ. Then $\nu(\varphi^j,E)\to\nu(\varphi,E)$ for any prime divisor E over X.*

This result is essentially [55, Lemma 2.2], proved under slightly different assumptions. We reproduce the argument for the convenience of the readers.

Proof Fix $k\in\mathbb{Z}_{>0}$, $\delta\in\mathbb{Q}_{>0}$, take $j_0>0$, so that when $j>j_0$, $\mathcal{I}((1+\delta)k\varphi^j)\subseteq\mathcal{I}(k\varphi)$. When $j>j_0$, we get

$$\frac{1}{k}\operatorname{ord}_E(\mathcal{I}(k\varphi))\leq\frac{1}{k}\operatorname{ord}_E(\mathcal{I}((1+\delta)k\varphi^j)).$$

By Fekete's lemma,

$$\nu(\varphi^j,E)=\sup_{k\in\mathbb{Z}_{>0}}\frac{1}{k}\operatorname{ord}_E(\mathcal{I}(k\varphi^j)).$$

So
$$\frac{1}{k}\operatorname{ord}_E(\mathcal{I}(k\varphi)) \le (1+\delta)\nu(\varphi^j, E).$$
Take sup with respect to $k \in \mathbb{Z}_{>0}$, we get
$$\nu(\varphi, E) \le (1+\delta)\nu(\varphi^j, E).$$
Letting $j \to \infty$ and then $\delta \to 0+$, we get
$$\nu(\varphi, E) \le \lim_{j\to\infty} \nu(\varphi^j, E).$$
The reverse inequality is trivial. □

Theorem 5.23 *Let E be a prime divisor on X. Let $(Y_\bullet)$ be an admissible flag with $E = Y_1$. Then*

$$\nu(\varphi, E) = \min_{x\in\Delta(\theta,\varphi)} x_1. \tag{5-17}$$

Here x_1 denotes the first component of x. The generic Lelong number $\nu(\varphi, E)$ means the minimum of $\nu(\varphi, x)$ for various $x \in E$.

Proof We first reduce to the case where θ_φ is a Kähler current. Let $\psi \le \varphi$, θ_ψ is a Kähler current. Then by (5-17) applied to $\varphi_\epsilon := (1-\epsilon)\varphi + \epsilon\psi$, we have
$$\nu(\varphi_\epsilon, E) = \min_{x\in\Delta(\theta,\varphi_\epsilon)} x_1.$$
Letting $\epsilon \to 0+$ using Theorem 5.9, we conclude (5-17).

Similarly, taking a quasi-equisingular approximation of φ and applying Lemma 5.22, we easily reduce to the case where φ also has analytic singularities. Replacing X by a birational model, we may assume that φ has analytic singularities along a simple normal crossing $\mathbb{Q}$-divisor F. Perturbing L by an ample $\mathbb{Q}$-line bundle by Proposition 5.16, we may assume that θ_φ is a Kähler current. Finally, by rescaling, we may assume that F is a divisor and L is a line bundle and $L - F$ is ample by [55, Lemma 2.4]. In fact, since θ_φ is a Kähler current, the same holds for $\theta_\varphi - \epsilon\omega$, where ω is a Hodge form lying in $c_1(A)$ for some ample line bundle A on X and $\epsilon > 0$ is a small enough rational number. By [55, Lemma 2.4], we deduce that $L - F - \epsilon A$ is nef and big and hence $L - F$ is ample.

By Theorem 5.17, we know that
$$\min_{x\in\Delta(\theta,\varphi)} x_1 = \lim_{k\to\infty} \min_{x\in\Delta_k(\theta,\varphi)} x_1.$$
By definition,
$$\min_{x\in\Delta_k(\theta,\varphi)} x_1 = k^{-1}\operatorname{ord}_E \mathrm{H}^0(X, L^k \otimes \mathcal{I}(k\varphi)).$$
It remains to show that
$$\lim_{k\to\infty} k^{-1}\operatorname{ord}_E \mathrm{H}^0(X, L^k \otimes \mathcal{I}(k\varphi)) = \lim_{k\to\infty} k^{-1}\operatorname{ord}_E \mathcal{I}(k\varphi). \tag{5-18}$$

The $\geq$ direction is trivial, we prove the converse. Observe that

$$\mathrm{H}^0(X, L^k \otimes \mathcal{I}(k\varphi)) = \mathrm{H}^0(X, L^k \otimes \mathcal{O}_X(-kF)), \quad \mathcal{I}(k\varphi) = \mathcal{O}(-kF).$$

As $L - F$ is ample, for large enough k, we have

$$\mathrm{ord}_E \mathrm{H}^0(X, L^k \otimes \mathcal{O}_X(-kF)) = \mathrm{ord}_E(kF).$$

Thus, (5-18) is clear. □

Corollary 5.24 *Let $\varphi, \psi \in \mathrm{PSH}(X, \theta)_{>0}$. If*

$$\Delta(\pi^*\theta, \pi^*\varphi) \subseteq \Delta(\pi^*\theta, \pi^*\psi)$$

for all birational models $\pi : Y \to X$ and all admissible flags on Y, then $\varphi \preceq_{\mathcal{I}} \psi$.

Proof In view of Theorem 5.23, the assumption implies the following: for any prime divisor E over X, we have $\nu(\varphi, E) \geq \nu(\psi, E)$. This implies $\varphi \preceq_{\mathcal{I}} \psi$: take a birational model $\pi : Y \to X$ and $y \in Y$, we need to show that $\nu(\pi^*\varphi, y) \geq \nu(\pi^*\psi, y)$. Let E be the exceptional divisor of the blowup of Y at $\{y\}$. As explained in [8, Corollaire 1.1.8], we have $\nu(\pi^*\varphi, y) = \nu(\varphi, E)$ and $\nu(\pi^*\psi, y) = \nu(\psi, E)$. Our assertion follows. □

In particular, Theorem B is proved. This corollary is similar to [41]. It suggests that $\Delta(\theta, \varphi)$ is a universal invariant of the singularities of φ.

Corollary 5.24 has a reminiscence of [14]: in order to understand plurisubharmonic singularities, we need to consider all birational models of our variety at the same time.

Theorem 5.23 can be regarded as a generalization of the following (slightly generalized form of the) classical result proved by Boucksom; see [9, Theorem 5.4].

Corollary 5.25 *Let E be a prime divisor over X. Then*

$$\nu(V_\theta, E) = \lim_{k\to\infty} \frac{1}{k} \mathrm{ord}_E \mathrm{H}^0(X, L^k). \tag{5-19}$$

Proof This follows from Theorem 5.23 and the fact that $\Delta(\theta, V_\theta) = \Delta(L)$. □

We write

$$\mathrm{ord}_E \|L\| := \lim_{k\to\infty} \frac{1}{k} \mathrm{ord}_E \mathrm{H}^0(X, L^k).$$

Corollary 5.26 *We have*

$$\mathcal{I}(V_\theta) = \{f \in \mathcal{O}_X \mid \exists \epsilon > 0 \text{ such that } \mathrm{ord}_E(f) \geq (1+\epsilon)\,\mathrm{ord}_E \|L\| - A_X(E)\ \forall \text{ primes } E \text{ over } X\},$$

where $A_X(E)$ is the log discrepancy of E over X.

Proof This follows from [10, Corollary 10.17] and Corollary 5.25. □

5.5 Okounkov bodies induced by filtrations

Assume that L is ample.

Definition 5.27 A *multiplicative filtration* on $R(X,L)$ is a decreasing, left continuous, multiplicative $\mathbb{R}$-filtration $\mathscr{F}^\bullet$ on the ring $R(X,L)$, which is linearly bounded in the sense that there is $C>0$ such that

$$\mathscr{F}^{-k\lambda}\mathrm{H}^0(X,L^k)=\mathrm{H}^0(X,L^k) \quad \text{and} \quad \mathscr{F}^{k\lambda}\mathrm{H}^0(X,L^k)=0 \quad \text{when } \lambda>C.$$

A multiplicative filtration $\mathscr{F}$ is called a *multiplicative* $\mathbb{Z}$*-filtration* if $\mathscr{F}^\lambda=\mathscr{F}^{\lfloor\lambda\rfloor}$ for any $\lambda\in\mathbb{R}$.

A multiplicative $\mathbb{Z}$-filtration $\mathscr{F}$ is called *finitely generated* if the bigraded algebra

$$\bigoplus_{\lambda\in\mathbb{Z},\,k\in\mathbb{Z}_{\geq 0}}\mathscr{F}^\lambda\mathrm{H}^0(X,L^k)$$

is finitely generated over $\mathbb{C}$.

Let $\mathscr{F}^\bullet$ be a multiplicative filtration on $R(X,L)$. Then we can associate a test curve $\psi_\bullet$ as in [49; 55]:

$$\psi_\tau := \sup_{k\in\mathbb{Z}_{>0}}{}^{*}\, k^{-1}\sup{}^{*}\Big\{\log|s|^2_{h^k} \;\Big|\; s\in\mathscr{F}^{k\tau}\mathrm{H}^0(X,L^k),\ \sup_X|s|_{h^k}\leq 1\Big\}. \tag{5-20}$$

Here $\sup^*$ denotes the upper-semicontinuous regularized supremum. By [32, Theorem 3.11], ψ_τ is $\mathcal{I}$-model or $-\infty$ for each $\tau\in\mathbb{R}$.

Theorem 5.28 *Let $\mathscr{F}^\bullet$ be a finitely generated multiplicative $\mathbb{Z}$-filtration on $R(X,L)$. Let $\psi_\bullet$ be the test curve associated with $\mathscr{F}$. For any $\tau<\tau^+$,*

$$\Delta\Big(\bigoplus_{k=0}^{\infty}\mathscr{F}^{k\tau}\mathrm{H}^0(X,L^k)\Big)=\Delta(\theta,\psi_\tau).$$

Proof Observe that $\mathscr{F}^{k\tau}\mathrm{H}^0(X,L^k)\subseteq\mathrm{H}^0(X,L^k\otimes\mathcal{I}(k\psi_\tau))$ for any $k\in\mathbb{N}$. Thus, by Corollary 3.15,

$$\Delta\Big(\bigoplus_{k=0}^{\infty}\mathscr{F}^{k\tau}\mathrm{H}^0(X,L^k)\Big)\subseteq\Delta(\theta,\psi_\tau).$$

On the other hand, the two sides have the same volume by [55, Lemma 4.5]. Thus, equality holds. □

5.6 Limit partial Okounkov bodies

Let $\varphi\in\mathrm{PSH}(X,\theta)$, not necessarily of positive volume. Take an ample effective divisor H on X and a Kähler form $\omega\in c_1(H)$. Then we just set

$$\Delta(\theta,\varphi):=\bigcap_{\epsilon\in\mathbb{Q}_{>0}}\Delta(\theta+\epsilon\omega,\varphi).$$

Clearly, this definition does not depend on the choice of H and ω. As in [22], we cannot expect $\Delta(\theta,\varphi)$ to be continuous along decreasing sequences of φ. Note that Theorem 5.23, Corollary 5.24 and Proposition 5.8 extend to this setup without changes.

Conjecture 5.29 *Under the above assumptions,*

$$\dim \Delta(\theta,\varphi) = \mathrm{nd}(\theta,\varphi).$$

For the definition of the analytic numerical dimension $\mathrm{nd}(\theta,\varphi)$, we refer to [19, Definition 4].

We expect this conjecture to follow from the arguments in [22] together with the numerical criterion of [19].

6 Chebyshev transform

Let X be an irreducible smooth complex projective variety of dimension n and L be a big line bundle on X. Let h be a fixed smooth Hermitian metric on L and $\theta = c_1(L,h)$. Consider a singular positive Hermitian metric ϕ on L corresponding to $\varphi \in \mathrm{PSH}(X,\theta)$. Assume that $\int_X \theta^n_{P[\varphi]_{\mathcal{I}}} > 0$.

Let $v \in C^0(X)$ corresponding to a continuous metric $h\mathrm{e}^{-v/2}$ on L. We do not distinguish v and $h\mathrm{e}^{-v/2}$. Fix a valuation $\nu = (\nu_1,\dots,\nu_n)\colon \mathbb{C}(X)^\times \to \mathbb{Z}^n$ of rank n. Assume that ν is defined by an admissible flag $(Y_\bullet)$ on X.

The whole section is devoted to the proof of Theorem C. Our results are direct extensions of the results of Witt Nyström [52]. The latter is motivated by [57].

6.1 Equilibrium energy

Let $\mathcal{E}^\infty(X,\theta;P[\varphi]_{\mathcal{I}})$ denote the set of $\psi \in \mathrm{PSH}(X,\theta)$ such that ψ and $P[\varphi]_{\mathcal{I}}$ have the same singularity types.

Let $E^\theta_{[\varphi]}\colon \mathcal{E}^\infty(X,\theta;P[\varphi]_{\mathcal{I}}) \to \mathbb{R}$ be the relative Monge–Ampère energy:

$$E^\theta_{[\varphi]}(\psi) := \frac{1}{n+1}\sum_{i=0}^n \int_X (\psi - P[\varphi]_{\mathcal{I}})\,\theta^i_\psi \wedge \theta^{n-i}_{P[\varphi]_{\mathcal{I}}}.$$

Define the equilibrium energy $\mathcal{E}^\theta_{[\varphi]}\colon C^0(X) \to \mathbb{R}$:

$$\mathcal{E}^\theta_{[\varphi]}(v) := E^\theta_{[\varphi]}(P[\varphi]_{\mathcal{I}}(v)). \tag{6-1}$$

Here

$$P[\varphi]_{\mathcal{I}}(v) = \sup{*}\{\eta \in \mathrm{PSH}(X,\theta) \mid \eta \le v, \eta \preceq_{\mathcal{I}} \varphi\}.$$

Note that this definition is different from the energy defined in [33], so we choose a different notation.

Theorem 6.1 *The Gateaux differential of $\mathcal{E}^\theta_{[\varphi]}$ at $v \in C^0(X)$ is given by $\theta^n_{P[\varphi]_{\mathcal{I}}(v)}$. In other words, for any $f \in C^0(X)$,*

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathcal{E}^\theta_{[\varphi]}(v+tf) = \int_X f\,\theta^n_{P[\varphi]_{\mathcal{I}}(v)}. \tag{6-2}$$

Proof This is not exactly [33, Proposition 5.10] because we are using $P[\bullet]_{\mathcal{I}}$ projections instead of $P[\bullet]$ projections, but the proofs are identical. □

The metric $he^{-v/2}$ induces an L^∞-type norm $\|\bullet\|_{L^\infty(kv)}$ on $\mathrm{H}^0(X, L^k \otimes \mathcal{I}(k\varphi))$:

$$\|s\|_{L^\infty(kv)} := \sup_X |s|_{h^k} \mathrm{e}^{-kv/2}.$$

In particular, $\det \|\bullet\|_{L^\infty(kv)}$ is a Hermitian metric on $\det \mathrm{H}^0(X, L^k \otimes \mathcal{I}(k\varphi))$.

Theorem 6.2 *Let $v, v' \in C^0(X)$. Then*

$$\lim_{k\to\infty} \frac{n!}{k^{n+1}} \log\left(\frac{\det\|\bullet\|_{L^\infty(kv)}}{\det\|\bullet\|_{L^\infty(kv')}}\right) = \mathcal{E}^\theta_{[\varphi]}(v) - \mathcal{E}^\theta_{[\varphi]}(v'). \tag{6-3}$$

Remark 6.3 When $\varphi = V_\theta$, the left-hand side of (6-3) is known as the *relative volume* between the two metrics $he^{-v/2}$ and $he^{-v'/2}$. They are studied in detail in [6].

This theorem partially generalizes [4, Theorem A]. We remind the readers that our conventions of multiplier ideal sheaves are different from those in [4] and [6], which explains the difference between our coefficients and theirs.

For the definition of the Bernstein–Markov property, see [4, Definition 2.3].

Proof We may assume that $v' = 0$. Let ν be a smooth volume form on X. Then recall that ν satisfies the Bernstein–Markov property with respect to tv for all $t \in [0, 1]$; see [4, Theorem 2.4]. We may replace the L^∞-norm on the left-hand side with the $L^2(\nu)$-norm by [33, Lemma 6.5]. We recall the definition of the partial Bergman kernel:

$$\begin{aligned} B^k_{tv,\varphi,\nu}(x) &:= \sup\left\{ |s|^2_{h^k} \mathrm{e}^{-kv}(x) \;\middle|\; \int_X |s|^2_{h^k} \mathrm{e}^{-tv} \le 1, s \in \mathrm{H}^0(X, L^k \otimes \mathcal{I}(k\varphi)) \right\}, \\ \beta^k_{tv,\varphi,\nu} &:= \frac{n!}{k^n} B^k_{tv,\varphi,\nu}\, \mathrm{d}\nu, \end{aligned}$$

where $k \in \mathbb{Z}_{>0}$.

By [33, Theorem 1.2],

$$\beta^k_{tv,\varphi,\nu} \rightharpoonup \theta^n_{P_X[\varphi]_{\mathcal{I}}(tv)}$$

as $k \to \infty$ for all $t \in [0, 1]$. By the dominated convergence theorem,

$$\lim_{k\to\infty} \int_0^1 \int_X v\, \beta^k_{tv,\varphi,\nu}\, \mathrm{d}t = \int_0^1 \int_X v\, \theta^n_{P_X[\varphi]_{\mathcal{I}}(tv)}\, \mathrm{d}t,$$

and (6-3) follows. □

Proposition 6.4 *Let $\varphi \in \mathrm{PSH}(X, \theta)$ such that θ_φ is a Kähler current. Let $(\varphi^j)_{j\in\mathbb{N}}$ be a quasi-equisingular approximation of φ. Then*

$$\lim_{j\to\infty} \mathcal{E}^\theta_{[\varphi^j]}(v) = \mathcal{E}^\theta_{[\varphi]}(v). \tag{6-4}$$

Proof By Theorem 6.1, for $j \in \mathbb{N}$,

$$\mathcal{E}^{\theta}_{[\varphi^j]}(v) = \int_0^1 \int_X v\, \theta^n_{P[\varphi^j]_{\mathcal{I}}(tv)}\, \mathrm{d}t \quad \text{and} \quad \mathcal{E}^{\theta}_{[\varphi]}(v) = \int_0^1 \int_X v\, \theta^n_{P[\varphi]_{\mathcal{I}}(tv)}\, \mathrm{d}t.$$

It follows from [33, Proposition 3.3] and [26, Theorem 1.2] that as $j \to \infty$,

$$\theta^n_{P[\varphi^j]_{\mathcal{I}}(tv)} \rightharpoonup \theta^n_{P[\varphi]_{\mathcal{I}}(tv)}.$$

By the dominated convergence theorem, (6-4) follows. □

Proposition 6.5 *Let $\varphi, \psi \in \mathrm{PSH}(X, \theta)$. Assume that $\psi \le \varphi$. Set $\varphi_\epsilon = (1-\epsilon)\varphi + \epsilon\psi$ for any $\epsilon \in [0,1]$. Then*

$$\lim_{\epsilon \to 0+} \mathcal{E}^{\theta}_{[\varphi_\epsilon]}(v) = \mathcal{E}^{\theta}_{[\varphi]}(v). \tag{6-5}$$

Proof The proof is similar to that of Proposition 6.4. We just replace [33, Proposition 3.3] by [33, Proposition 2.7]. □

We finally recall a technical lemma.

Lemma 6.6 [52, Corollary 3.4] *Let $C \subseteq \mathbb{R}^{n+1}$ be an open convex cone. Let F be a subadditive function on $C \cap \mathbb{Z}^{n+1}$ defined outside a compact set. Then for any sequence $\alpha_k \in C \cap \mathbb{Z}^{n+1}$ tending to infinity such that $\alpha_k/|\alpha_k|$ converges to some point $p \in C$. Then the limit*

$$c[F](p) := \lim_{k\to\infty} \frac{F(\alpha_k)}{|\alpha_k|}$$

exists and depends only on p and F. Moreover, $c[F]$ is a convex function on $C \cap \{x_{n+1} = 1\}$.

Here $|\alpha_k|$ denotes the absolute value of the last component of α_k.

Recall that a real-valued function F defined on a semigroup Γ is said to be *subadditive* if for any $x, y \in \Gamma$, $F(x+y) \le F(x) + F(y)$.

6.2 The case of analytic singularities

Assume that φ has analytic singularities.

Let $\pi : Y \to X$ be a resolution such that $\pi^*\varphi$ has analytic singularity along a normal crossing $\mathbb{Q}$-divisor E. We define as before

$$W^0_k = \mathrm{H}^0(Y, \pi^* L^k \otimes \mathcal{O}_Y(-kE)) \subseteq \mathrm{H}^0(X, L^k).$$

Fix $a \in \Gamma_k(W^0)$. Let p be the center of ν on X. Let $z = (z_1, \dots, z_n)$ be a regular sequence in $\mathcal{O}_{X,p}$ such that $(Y_i)_x$ is the zero locus of $z_1, \dots, z_i$. Fix a local trivialization of L near p. Define

$$A^a_k := \{s \in W^0_k \mid \nu(s) \ge ka,\ s = z^{ka} + \text{higher-order terms near } p\}.$$

Define

$$F[v](ka,k) = \inf_{s\in A_{a,k}} \log |s|_{L^\infty(kv)}.$$

Recall the following two lemmas proved in [52, Lemmas 5.3 and 5.4].

Lemma 6.7 *$F[v]$ is subadditive on $\Gamma(W^0)$.*

Lemma 6.8 *There is a $C>0$ such that for any $(ka,k)\in\Gamma(W^0)$,*

$$F[v](ka,k) \geq C|(ka,k)|.$$

Proof It suffices to apply [52, Lemma 5.4]. □

Let $c_{[\varphi]}[v]\colon \operatorname{Int}\Delta(\theta,\varphi)\to\mathbb{R}$ be the convex function $c[F[v]]$ defined by Lemma 6.6.

Theorem 6.9 *We have*

$$\int_{\Delta(W(\theta,\varphi))} (c_{[\varphi]}[v]-c_{[\varphi]}[0])\,\mathrm{d}\lambda = -\mathcal{E}^{\theta}_{[\varphi]}(v).$$

Proof The proof follows *verbatim* from that of [52, Theorem 6.2], taking into account Theorem 6.2. □

Observe that

$$\sup_{\operatorname{Int}\Delta(W(\theta,\varphi))} |c_{[\varphi]}[v]-c_{[\varphi]}[0]| \leq \tfrac{1}{2}\|v\|_{C^0(X)}. \tag{6-6}$$

The following result is obvious:

Lemma 6.10 *Let $\varphi,\varphi'\in \operatorname{PSH}(X,\theta)$ be potentials with analytic singularities. If $[\varphi]\preceq[\varphi']$, then*

$$c_{[\varphi]}[v] \geq c_{[\varphi']}[v]$$

when restricted to $\operatorname{Int}\Delta(\theta,\varphi)$.

6.3 The case of Kähler currents

Assume that θ_φ is a Kähler current. Let φ^j be a quasi-equisingular approximation of φ. Then $c_{[\varphi^j]}[v]$ restricted to $\operatorname{Int}\Delta(W(\theta,\varphi))$ is an increasing sequence. Thus, we can define $c_{[\varphi]}[v]\colon \operatorname{Int}\Delta(\theta,\varphi)\to\mathbb{R}\cup\{\infty\}$ by

$$c_{[\varphi]}[v] := \lim_{j\to\infty} c_{[\varphi^j]}[v].$$

Lemma 6.11 *Let $s\in W_k(\theta,\varphi)$, locally written as z^{ka} plus higher-order terms near p. Then*

$$c_{[\varphi]}[v](a) \leq k^{-1}\log\|s\|_{L^\infty(kv)}.$$

Proof This follows from the corresponding result for the φ^j. □

By convexity, $c_{[\varphi]}[v]$ takes finite values.

It follows that (6-6) still holds in this case. By the dominated convergence theorem, Proposition 6.4 and the previous case we find

$$\int_{\Delta(\theta,\varphi)} (c_{[\varphi]}[v]-c_{[\varphi]}[0])\,\mathrm{d}\lambda = -\mathcal{E}^{\theta}_{[\varphi]}(v).$$

It follows from Lemma 6.10 that our definition of $c_{[\varphi]}(v)$ is independent of the choice of φ^j.

Lemma 6.12 *Let $\varphi,\varphi' \in \mathrm{PSH}(X,\theta)$ be potentials such that θ_φ and $\theta_{\varphi'}$ are both Kähler currents. If $[\varphi] \preceq_{\mathcal{I}} [\varphi']$, then*

$$c_{[\varphi]}[v] \geq c_{[\varphi']}[v]$$

when restricted to $\mathrm{Int}\,\Delta(\theta,\varphi)$.

Proof This follows from Lemma 6.10. □

6.4 General case

Let $\varphi \in \mathrm{PSH}(X,\theta)$ such that $\int_X \theta^n_{P[\varphi]_{\mathcal{I}}} > 0$. We may replace φ with $P[\varphi]_{\mathcal{I}}$ and therefore assume that the nonpluripolar mass of φ is positive.

Let $\eta \in \mathrm{PSH}(X,\theta)$ be a potential so that θ_η is a Kähler current and $\eta \preceq \varphi$. The existence of such η is guaranteed by [33, Proposition 3.6]. Define $\varphi_\epsilon := (1-\epsilon)\varphi + \epsilon\eta$. Then we define

$$c_{[\varphi]}[v]\colon \mathrm{Int}\,\Delta(\theta,\varphi) \to \mathbb{R}\cup\{-\infty\}, \quad c_{[\varphi]}[v] := \lim_{\epsilon\to 0+} c_{[\varphi_\epsilon]}[v].$$

This is a decreasing limit by Lemma 6.12. On the other hand, $c_{[\varphi]}[v] \geq c_{[V_\theta]}[v]$, the latter is finite by [52]. Thus, $c_{[\varphi]}[v]$ is real-valued. Inequality (6-6) extends to this situation. By the dominated convergence theorem and Proposition 6.5 again,

$$\int_{\Delta(\theta,\varphi)} (c_{[\varphi]}[v]-c_{[\varphi]}[0])\,\mathrm{d}\lambda = -\mathcal{E}^{\theta}_{[\varphi]}(v).$$

We do not know if $c_{[\varphi]}[v]$ is independent of the choice of η.

7 A generalization of Boucksom–Chen theorem

In this section, let X be an irreducible smooth projective variety of dimension n. Let L be a big line bundle on X. Take a smooth Hermitian metric h on L with $\theta = c_1(L,h)$.

Fix a rank n valuation $\nu\colon \mathbb{C}(X)^\times \to \mathbb{Z}^n$.

7.1 The theory of test curves

Let $V = \langle L^n \rangle$.

Definition 7.1 We define a *test curve* (*of finite energy*) with respect to (X, θ) to be a map $\psi = \psi_\bullet \colon \mathbb{R} \to \mathrm{PSH}(X, \theta) \cup \{-\infty\}$ such that

(1) $\psi_\bullet$ is concave in $\bullet$,

(2) ψ_τ is a model potential or $-\infty$ for any τ,

(3) ψ is usc as a function in the $\mathbb{R}$-variable,

(4) $\lim_{\tau \to -\infty} \psi_\tau = V_\theta$ in L^1,

(5) $\psi_\tau = -\infty$ for τ large enough,

(6) ψ satisfies

$$\mathbf{E}(\psi_\bullet) := \tau^+ V + \int_{-\infty}^{\tau^+} \left(\int_X \theta_{\psi_\tau}^n - V \right) \mathrm{d}\tau > -\infty. \tag{7-1}$$

Here $\tau^+ := \inf\{\tau \in \mathbb{R} \mid \psi_\tau = -\infty\}$. The set of test curves of finite energy with respect to (X, θ) is denoted by $\mathcal{TC}^1(X, \theta)$. We say ψ is *normalized* if $\tau^+ = 0$. The test curve is called *bounded* if $\psi_\tau = V_\theta$ for τ small enough. Let $\tau^- := \sup\{\tau \in \mathbb{R} \mid \psi_\tau = V_\theta\}$ in this case. The set of bounded test curves is denoted by $\mathcal{TC}^\infty(X, \theta)$.

We say a test curve is $\mathcal{I}$*-model* if ψ_τ is $\mathcal{I}$-model for each $\tau < \tau^+$. The set of $\mathcal{I}$-model test curves is denoted by $\mathcal{TC}^1_{\mathcal{I}}(X, \theta)$.

7.2 Okounkov test curves

Let $\Delta \in \mathcal{K}^n$. Assume that $V = n! \operatorname{vol} \Delta > 0$.

Definition 7.2 An *Okounkov test curve* relative to Δ is an assignment $(\Delta_\tau)_{\tau \le \tau^+}$ for $\tau^+ \in \mathbb{R}$ such that:

(1) Δ_τ is a decreasing assignment of convex bodies in $\mathbb{R}^n$ for $\tau \le \tau^+$.

(2) Δ_τ converges to Δ as $\tau \to -\infty$ with respect to the Hausdorff metric (cf Section 2.1).

(3) Δ_τ is concave in the τ variable.

(4) The energy is finite:

$$\mathbf{E}(\Delta_\bullet) := \tau^+ V + V \int_{-\infty}^{\tau^+} \left(\frac{n!}{V} \operatorname{vol} \Delta_\tau - 1 \right) \mathrm{d}\tau > -\infty.$$

(5) Continuity holds at τ^+:

$$\Delta_{\tau^+} = \bigcap_{\tau < \tau^+} \Delta_\tau.$$

Proposition 7.3 *Any Okounkov test curve $(\Delta_\tau)_{\tau \le \tau^+}$ relative to Δ is continuous for $\tau < \tau^+$.*

Proof We first claim that $\operatorname{vol}\Delta_{\tau'} > 0$ for all $\tau' < \tau^+$. By condition (2) and Theorem 2.3, we know that $\operatorname{vol}\Delta_{\tau''} > 0$ when τ'' is small enough. Fix one such τ''. Any $\tau' < \tau^+$ can be written as a convex combination of τ^+ and τ'', thus $\Delta_{\tau'}$ has positive volume by condition (3).

Next we claim that $\operatorname{vol}\Delta_\tau$ is continuous for $\tau < \tau^+$. In fact, by condition (3) and the Minkowski inequality, we know that $\log\operatorname{vol}\Delta_\tau$ is concave for $\tau < \tau^+$. The continuity follows.

Next we show that
$$\Delta_\tau = \bigcap_{\tau'<\tau} \Delta_{\tau'}.$$
The $\supseteq$ direction is obvious. By the continuity of the volume, both sides have the same volume and the volume is positive, hence, equality holds by Lemma 2.5.

Similarly, we have
$$\Delta_\tau = \overline{\bigcup_{\tau'>\tau} \Delta_{\tau'}}.$$
The continuity of Δ_τ at $\tau < \tau^+$ is proved. □

Definition 7.4 A *test function* on Δ is a function $F\colon \Delta \to [-\infty, \infty)$ such that:

(1) F is concave.

(2) F is finite on $\operatorname{Int}\Delta$.

(3) F is usc.

(4) The energy is finite:

$$\mathbf{E}(F) := n! \int_\Delta F \,\mathrm{d}\lambda > -\infty. \tag{7-2}$$

Let $\tau^+ = \sup_\Delta F$. Then
$$\mathbf{E}(F) = \tau^+ V + V \int_{-\infty}^{\tau^+} \left(\frac{n!}{V} \operatorname{vol}\{F \geq \tau\} - 1 \right) \mathrm{d}\tau. \tag{7-3}$$

Let $\Delta_\bullet$ be an Okounkov test curve relative to Δ. We define the *Legendre transform* of $\Delta_\bullet$ as
$$G[\Delta_\bullet]\colon \Delta \to [-\infty, \infty), \quad a \mapsto \sup\{\tau < \tau^+ \mid a \in \Delta_\tau\}.$$
Conversely, a test function F on Δ, set $\tau^+ = \sup_\Delta F$. We define the *inverse Legendre transform* of F as
$$\Delta[F]\colon (-\infty, \tau^+] \to \mathcal{K}_n, \quad \Delta[F]_\tau = \{F \geq \tau\}.$$

Theorem 7.5 *The Legendre transform and inverse Legendre transform are inverse to each other, defining a bijection between the set of Okounkov test curves relative to Δ and test functions on Δ. Moreover, if $\Delta_\bullet$ is an Okounkov test curve relative to Δ, then*
$$\mathbf{E}(\Delta_\bullet) = \mathbf{E}\left(G[\Delta_\bullet]\right). \tag{7-4}$$

Proof Let $\Delta_\bullet$ be an Okounkov test curve relative to Δ. We prove that $G[\Delta_\bullet]$ is a test function on Δ.

Firstly $G[\Delta_\bullet]$ is concave by condition (1) and condition (3) in Definition 7.2. More precisely, take $a, b \in \Delta$. We want to prove that for any $t \in (0,1)$,

$$G[\Delta_\bullet](ta+(1-t)b) \geq tG[\Delta_\bullet](a)+(1-t)G[\Delta_\bullet](b). \tag{7-5}$$

There is nothing to prove if $G[\Delta_\bullet](a)$ or $G[\Delta_\bullet](b)$ is $-\infty$. So we assume that both are finite. Take $\epsilon > 0$, then $a \in \Delta_{G[\Delta_\bullet](a)-\epsilon}$ and $b \in \Delta_{G[\Delta_\bullet](b)-\epsilon}$. Thus,

$$ta+(1-t)b \in t\Delta_{G[\Delta_\bullet](a)-\epsilon}+(1-t)\Delta_{G[\Delta_\bullet](b)-\epsilon} \subseteq \Delta_{tG[\Delta_\bullet](a)+(1-t)G[\Delta_\bullet](b)-\epsilon}.$$

We deduce that

$$G[\Delta_\bullet](ta+(1-t)b) \geq tG[\Delta_\bullet](a)+(1-t)G[\Delta_\bullet](b)-\epsilon.$$

Since $\epsilon > 0$ is arbitrary, (7-5) follows.

Next $G[\Delta_\bullet]$ is finite on $\operatorname{Int}\Delta$ by condition (2). In fact, as Δ_τ is increasing and converges to Δ as $\tau \to -\infty$, we have

$$\Delta = \overline{\bigcup_\tau \Delta_\tau}.$$

Hence, by [50, Theorem 1.1.15] and the assumption that $\operatorname{vol}\Delta > 0$, $\bigcup_\tau \Delta_\tau$ contains $\operatorname{Int}\Delta$.

Thirdly, we show that $G[\Delta_\bullet]$ is usc. Let $a_i \in \Delta$ with $a_i \to a \in \Delta$. Define $\tau_i = G[\Delta_\bullet](a_i)$. Let $\tau = \overline{\lim}_i \tau_i$. We need to show that

$$G[\Delta_\bullet](a) \geq \tau. \tag{7-6}$$

There is nothing to prove if $\tau = -\infty$. We assume that it is not this case. Up to subtracting a subsequence we may assume that $\tau_i \to \tau$. In particular, we can assume that $\tau_i \neq -\infty$ for all i. Fix $\epsilon > 0$, then $a_i \in \Delta_{\tau_i-\epsilon}$. Observe that $\Delta_{\tau_i-\epsilon} \xrightarrow{d_n} \Delta_{\tau-\epsilon}$. By Theorem 2.4 it follows that $a \in \Delta_{\tau-\epsilon}$. Thus, (7-6) follows since $\epsilon > 0$ is arbitrary.

Finally, (7-4) follows from (7-3), and it follows that $\mathrm{E}(G[\Delta_\bullet]) > -\infty$.

Conversely, if $F\colon \Delta \to [-\infty,\infty)$ is a test function on Δ. Let $\Delta[F]$ be the inverse Legendre transform of F. Then one can similarly show that $\Delta[F]$ is an Okounkov test curve.

Firstly, for each $\tau < \tau^+ := \sup_\Delta F$, $\Delta[F](\tau)$ is a convex body as F is concave and usc. Moreover, $\Delta[F]_\tau$ is clearly decreasing in τ. Hence, $\Delta[F]_{\tau^+}$ is also a convex body.

Secondly, for each $a \in \Delta$, we can write $a = \lim_i a_i$ with $a_i \in \operatorname{Int}\Delta$. By assumption, F is finite at a_i. Thus,

$$a \in \overline{\{F > -\infty\}} = \overline{\bigcup_\tau \Delta[F]_\tau}.$$

By Theorem 2.4, $\Delta[F]_\tau \xrightarrow{d_n} \Delta$ as $\tau \to -\infty$.

Thirdly, $\Delta[F]$ is concave. To see, take $\tau, \tau' \leq \tau^+$, we need to prove that for any $t \in (0,1)$,

$$\Delta[F]_{t\tau+(1-t)\tau'} \supseteq t\Delta[F]_\tau + (1-t)\Delta[F]_{\tau'}. \tag{7-7}$$

Let $a \in \Delta[F]_\tau$ and $b \in \Delta[F]_{\tau'}$. We have $F(a) \geq \tau$ and $F(b) \geq \tau'$. F is concave, so $F(ta+(1-t)b) \geq t\tau + (1-t)\tau'$. Thus,

$$ta + (1-t)b \in \Delta[F]_{t\tau+(1-t)\tau'}$$

and (7-7) follows.

Fourthly, (7-2) follows immediately from (7-3).

Finally, we show that $\Delta[F]_\bullet$ is continuous at τ^+. This amounts to

$$\{F \geq \tau^+\} = \bigcap_{\tau<\tau^+} \{F \geq \tau\},$$

which is obvious.

To see that these two operations are inverse to each other, observe that by definition for any Okounkov test curve $\Delta_\bullet$, any $a \in \Delta$ and any $\tau \leq \tau^+$, one has $G[\Delta_\bullet](a) \geq \tau$ if and only if $a \in \Delta_{\tau-\epsilon}$ for any $\epsilon > 0$. By Proposition 7.3, this happens if and only if $a \in \Delta_\tau$, that is,

$$\{G[\Delta_\bullet] \geq \tau\} = \Delta_\tau.$$

Conversely, for any test function $F\colon \Delta \to [-\infty, \infty)$, any $\tau \leq \tau^+$, by definition,

$$\{F \geq \tau\} = \Delta[F]_\tau. \qquad \square$$

Definition 7.6 Let $\Delta_\bullet$ be an Okounkov test curve relative to Δ. We define the *Duistermaat–Heckman measure* $\mathrm{DH}(\Delta_\bullet)$ as

$$\mathrm{DH}(\Delta_\bullet) := G[\Delta_\bullet]_*(\mathrm{d}\lambda).$$

It is a Radon measure on $\mathbb{R}$.

Observe that

$$\int_{\mathbb{R}} \mathrm{DH}(\Delta_\bullet) = \operatorname{vol} \Delta. \tag{7-8}$$

7.3 Boucksom–Chen theorem

Let $\psi_\bullet \in \mathcal{TC}^1_{\mathcal{I}}(X,\theta)$. Let $\tau^+ = \inf\{\tau \in \mathbb{R} \mid \psi_\tau = -\infty\}$.

Lemma 7.7 *The curve*

$$\Delta[\psi_\bullet]_\tau := \begin{cases} \Delta(\theta, \psi_\tau) & \text{if } \tau < \tau^+, \\ \bigcap_{\tau'<\tau^+} \Delta[\psi_\bullet]_{\tau'} & \text{if } \tau = \tau^+, \end{cases}$$

is an Okounkov test curve relative to $\Delta(L)$. Moreover,

$$\mathbf{E}(\psi_\bullet) = \mathbf{E}(\Delta[\psi_\bullet]_\bullet). \tag{7-9}$$

Proof We verify the conditions in Definition 7.2. condition (1) follows from Proposition 5.8. Condition (2) follows from the fact that

$$\lim_{\tau\to-\infty} \operatorname{vol} \Delta_\tau = \operatorname{vol} \Delta.$$

Condition (3) follows from Theorem 5.14 and Proposition 5.8. Condition (4) is a translation of (7-1). Condition (5) is obvious.

Finally, (7-9) follows from (7-1) and (1-3). □

Definition 7.8 Let $\psi_\bullet \in \mathcal{TC}^1_{\mathcal{I}}(X,\theta)$. Define the *Duistermaat–Heckman measure* of $\psi_\bullet$ as

$$\mathrm{DH}(\psi_\bullet) := \mathrm{DH}(\Delta[\psi_\bullet]_\bullet).$$

We write

$$G[\psi_\bullet] = G[\Delta[\psi_\bullet]].$$

Then

$$\mathrm{DH}(\psi_\bullet) = G[\psi_\bullet]_*(\mathrm{d}\lambda).$$

Now consider the (not necessarily multiplicative) filtration

$$\mathscr{F}^k_\tau \mathrm{H}^0(X,L^k) := \begin{cases} \mathrm{H}^0(X, L^k \otimes \mathcal{I}(k\psi_\tau)) & \text{if } \tau < \tau^+, \\ 0 & \text{if } \tau \geq \tau^+. \end{cases}$$

Let $e_j(\mathrm{H}^0(X,L^k),\mathscr{F}^k)$ be the jumping numbers of $\mathscr{F}^k$ listed in decreasing order. In other words,

$$e_j(\mathrm{H}^0(X,L^k),\mathscr{F}^k) := \sup\{\tau \in \mathbb{R} \mid \dim \mathscr{F}^k_\tau \mathrm{H}^0(X,L^k) \geq j\}.$$

Let

$$\mu_k := \frac{1}{k^n} \sum_{j=1}^{h^0(X,L^k)} \delta_{e_j(\mathrm{H}^0(X,L^k),\mathscr{F}^k)}.$$

Theorem 7.9 *Let $\psi_\bullet \in \mathcal{TC}^1_{\mathcal{I}}(X,\theta)$. Then as $k \to \infty$, the measure μ_k converges weakly to* $\mathrm{DH}(\psi_\bullet)$.

As explained in [49; 32; 55], $\mathcal{TC}^1_{\mathcal{I}}(X,\theta)$ is the completion of the space of filtrations, so this theorem indeed generalizes [11, Theorem A], in the case of full-graded linear series.

Proof It suffices to show the convergence holds as distributions. By our definition, μ_k is the distributional derivative of the function

$$h_k(\tau) := \begin{cases} k^{-n} h^0(X, L^k \otimes \mathcal{I}(k\psi_\tau)) & \text{if } \tau < \tau^+, \\ 0 & \text{if } \tau \geq \tau^+. \end{cases}$$

On the other hand, $\mathrm{DH}(\psi_\bullet)$ is the distributional derivative of $h(\tau) := \mathrm{vol}\{G[\Delta[\psi_\bullet]_\bullet] \geq \tau\} = \mathrm{vol}\,\Delta_\tau$ by the Fubini–Tonelli theorem.

By Theorem 2.19, $h_k(\tau) \to h(\tau)$ for all $\tau \neq \tau^+$. By the dominated convergence theorem $h_k \to h$ in $L^1_{\mathrm{loc}}(\mathbb{R})$. Hence, $\mu_k \rightharpoonup \mathrm{DH}(\psi_\bullet)$. □

Corollary 7.10 *For any $\psi_\bullet \in \mathcal{TC}^1_{\mathcal{I}}(X,\theta)$. The Duistermaat–Heckman measure* $\mathrm{DH}(\psi_\bullet)$ *is independent of the choice of the valuation ν.*

7.4 Applications to non-Archimedean geometry

Assume that L is ample and θ is a Kähler form. We write $\omega = \theta$ instead.

Finite-energy geodesic rays Let $\mathcal{E}^1(X,\omega)$ denote the space of ω-psh functions with finite energy:

$$\mathcal{E}^1(X,\omega) := \left\{ \varphi \in \mathrm{PSH}(X,\omega) \;\middle|\; \int_X \omega_\varphi^n = \int_X \omega^n, \int_X |\varphi|\, \omega_\varphi^n < \infty \right\}.$$

See [24] for a detailed introduction. Recall that $\mathcal{E}^1(X,\omega)$ admits a natural metric d_1: for $\varphi, \psi \in \mathcal{E}^1(X,\omega)$, given by

$$d_1(\varphi,\psi) := \mathrm{E}(\varphi) + \mathrm{E}(\psi) - 2\mathrm{E}(\varphi \wedge \psi).$$

Here

$$\varphi \wedge \psi := \sup\{\eta \in \mathrm{PSH}(X,\omega) \mid \eta \le \varphi, \eta \le \psi\}.$$

In [27, Theorem 2.10], Darvas, Di Nezza and Lu proved that $\varphi \wedge \psi \in \mathcal{E}^1(X,\omega)$. They proved in [25, Section 3] that d_1 is indeed a metric. The Monge–Ampère energy functional $\mathrm{E}\colon \mathcal{E}^1(X,\omega) \to \mathbb{R}$ is defined as

$$\mathrm{E}(\varphi) = \frac{1}{n+1} \sum_{i=0}^{n} \int_X \varphi\, \omega_\varphi^i \wedge \omega^{n-i}.$$

In this case, let $\mathcal{R}^1(X,\omega)$ denote the set of geodesic rays in $\mathcal{E}^1(X,\omega)$ emanating from 0. For a detailed study of $\mathcal{R}^1(X,\omega)$, we refer to [30]. Here we only recall the definition of the metric on $\mathcal{R}^1(X,\omega)$. Given $\ell, \ell' \in \mathcal{R}^1(X,\omega)$, we define

$$d_1(\ell,\ell') := \lim_{t\to\infty} \frac{1}{t} d_1(\ell_t, \ell'_t).$$

By [20, Corollary 5.5], $t \mapsto d_1(\ell_t,\ell'_t)$ is convex, guaranteeing the existence of the limit. It is shown in [30] that $(\mathcal{R}^1(X,\omega), d_1)$ is a complete metric space.

The following notion is introduced in [54]:

Definition 7.11 A *rooftop metric space* is a triple $(E,d,\wedge)$: (E,d) is a metric space and $\wedge\colon E\times E \to E$ is an associative, commutative binary operator on E satisfying

$$d(a\wedge c, b\wedge c) \le d(a,b) \quad \text{for any } a,b,c \in E.$$

For $\ell, \ell' \in \mathcal{R}^1(X,\omega)$, define $\ell \wedge \ell'$ as the greatest geodesic in $\mathcal{R}^1(X,\omega)$ that lies below both ℓ and ℓ'. It is shown in [54, Theorem 7.6] that $\wedge$ is well-defined and $(\mathcal{R}^1(X,\omega), d_1, \wedge)$ is a complete rooftop metric space.

The energy functional $\mathbf{E}\colon \mathcal{R}^1(X,\omega) \to \mathbb{R}$ is defined as

$$\mathbf{E}(\ell) := \mathrm{E}(\ell_1).$$

Recall that we have the following two maps: Given any $\ell \in \mathcal{R}^1(X,\omega)$, its *inverse Legendre transform* is defined as

$$\hat{\ell}_\tau := \inf_{t \geq 0} (\ell_t - t\tau).$$

Conversely, given any $\psi_\bullet \in \mathcal{TC}^1(X,\omega)$, we define its *Legendre transform* by

$$\check{\psi}_t := \sup_{\tau \in \mathbb{R}} (\psi_\tau + t\tau).$$

They are inverse to each other, as proved in [32, Theorem 3.7].

Non-Archimedean pluripotential theory Let X^{an} be the Berkovich analytification of X with respect to the trivial valuation on X and L^{an} be the analytification of L. See Section 2.7 for a brief introduction. In the same section, we also recalled the definition of the space $\mathcal{E}^1(L^{\mathrm{an}})$ of non-Archimedean psh metrics on L^{an} with finite energy and the energy functional $\mathrm{E}\colon \mathcal{E}^1(L^{\mathrm{an}}) \to \mathbb{R}$.

Next we briefly explain the relation between the non-Archimedean pluripotential theory and the complex pluripotential theory. Firstly, given a geodesic ray $\ell \in \mathcal{R}^1(X,\omega)$, one can associate a non-Archimedean potential $\ell^{\mathrm{an}} \in \mathcal{E}^1(L^{\mathrm{an}})$ as in [5, Definition 4.2, Theorem 6.2]. The construction of ℓ^{an} requires the notion of Gauss extension of valuations, as explained in [5, Section 3.1]. The map

$$\mathcal{R}^1(X,\omega) \to \mathcal{E}^1(L^{\mathrm{an}})$$

is surjective but not injective. It admits a canonical section

$$\iota\colon \mathcal{E}^1(L^{\mathrm{an}}) \hookrightarrow \mathcal{R}^1(X,\omega)$$

sending $\phi \in \mathcal{E}^1(L^{\mathrm{an}})$ to the maximal element $\ell \in \mathcal{E}^1(L^{\mathrm{an}})$ with $\ell^{\mathrm{an}} = \phi$. See [5, Theorem 6.6].

The geodesics lying in the image of ι are known as *maximal geodesic rays* or *approximable geodesic rays*. Moreover,

$$\mathbf{E}(\iota(\alpha)) = \mathrm{E}(\alpha) \tag{7-10}$$

for any $\alpha \in \mathcal{E}^1(L^{\mathrm{an}})$; see [5, Corollary 6.7].

Maximal geodesic rays are closely related to test curves:

Theorem 7.12 *The Legendre transform is a bijection from $\mathcal{TC}^1_{\mathcal{I}}(X,\omega)$ (resp. $\mathcal{TC}^1(X,\omega)$) to $\iota(\mathcal{E}^1(L^{\mathrm{an}}))$ (resp. $\mathcal{R}^1(X,\omega)$); the inverse is given by the inverse Legendre transform. Further, for any $\psi_\bullet \in \mathcal{TC}^1(X,\omega)$,*

$$\mathbf{E}(\psi_\bullet) = \mathbf{E}(\check{\psi}). \tag{7-11}$$

This is one of the main theorems of [32, Theorems 3.7 and 3.17]. It is based on the previous work [49; 25].

Duistermaat–Heckman measures The space $\mathcal{E}^1(L^{\mathrm{an}})$ is closely related to the theory of test configurations. For the latter, we refer to [16, Section 2] for a brief introduction. Recall that two test configurations $(\mathcal{X},\mathcal{L})$ and $(\mathcal{X}',\mathcal{L}')$ of (X,L) are said to be equivalent if they can be dominated by a common test configuration; see [16, Definition 6.1]. There is a natural injection from the set of equivalence classes of test configurations to $\mathcal{E}^1(L^{\mathrm{an}})$. Moreover, this injection has dense image and $\mathcal{E}^1(L^{\mathrm{an}})$ is the d_1-completion of the space of test configurations (modulo the equivalence relation). These results are explained in detail in [32, Section 3.2].

Given a test configuration $(\mathcal{X},\mathcal{L})$, Witt Nyström [51] constructed a naturally defined Radon measure $\mathrm{DH}(\mathcal{X},\mathcal{L})$ on $\mathbb{R}$, called the *Duistermaat–Heckman measure*. See [16, Section 3.2] for more details. It is not hard to see from the definition that $\mathrm{DH}(\mathcal{X},\mathcal{L})$ depends only on the equivalence class of $(\mathcal{X},\mathcal{L})$.

In the sequel, we will define the Duistermaat–Heckman measure of an element in $\mathcal{E}^1(L^{\mathrm{an}})$. As the space $\mathcal{E}^1(L^{\mathrm{an}})$ is the completion of the space of test configurations (modulo the equivalence relation), our definition can be seen as an extension of Witt Nyström's results [51].

Definition 7.13 For any $\alpha \in \mathcal{E}^1(L^{\mathrm{an}})$, define the *Duistermaat–Heckman measure* of α as

$$\mathrm{DH}(\alpha) := \mathrm{DH}(\widehat{\iota(\alpha)}).$$

We get a map $\mathrm{DH}\colon \mathcal{E}^1(L^{\mathrm{an}}) \to \mathcal{M}(\mathbb{R})$. Here $\mathcal{M}(\mathbb{R})$ denotes the space of Radon measures on $\mathbb{R}$.

For the proof of the next theorem, we need to recall several basic constructions of test curves.

The space $\mathcal{TC}^1(X,\omega)$ is a rooftop metric space. Its rooftop structures $(d_1,\wedge)$ are induced from the corresponding structures on $\mathcal{R}^1(X,\omega)$.

Corollary 7.14 *Let $\psi_\bullet,\varphi_\bullet,\eta_\bullet \in \mathcal{TC}^1(X,\omega)$.*

(1) *The rooftop operator on $\mathcal{TC}^1(X,\omega)$ is given by*

$$(\psi\wedge\varphi)_\tau = \psi_\tau \wedge \varphi_\tau. \tag{7-12}$$

It is the maximal element in $\mathcal{TC}^1(X,\omega)$ that lies below both $\psi_\bullet$ and $\varphi_\bullet$. In particular,

$$d_1((\psi\wedge\eta)_\bullet,(\varphi\wedge\eta)_\bullet) \le d_1(\psi_\bullet,\varphi_\bullet). \tag{7-13}$$

(2) *The metric on $\mathcal{TC}^1(X,\omega)$ is given by*

$$d_1(\psi_\bullet,\varphi_\bullet) := \mathbf{E}(\psi_\bullet)+\mathbf{E}(\varphi_\bullet)-2\mathbf{E}((\psi\wedge\varphi)_\bullet). \tag{7-14}$$

Proof (1) Note that (7-13) is part of our definition of a rooftop structure.

Observe that the bijection $\mathcal{TC}^1(X,\omega)\to\mathcal{R}^1(X,\omega)$ is order-preserving. In order to prove our claim, it suffices to show that $(\varphi\wedge\psi)_\bullet$ defined by (7-12) is indeed in $\mathcal{TC}^1(X,\omega)$, which is obvious.

(2) This follows simply from (1) and (7-11). □

If $\varphi_\bullet, \psi_\bullet \in \mathcal{TC}^1_{\mathcal{I}}(X,\omega)$, then $(\psi \wedge \varphi)_\bullet \in \mathcal{TC}^1_{\mathcal{I}}(X,\omega)$ as well. This follows from the simple observation that the rooftop of two $\mathcal{I}$-model potentials is still $\mathcal{I}$-model. Now the d_1 metric on $\mathcal{TC}^1(X,\omega)$ restricts to a metric d_1 on $\mathcal{TC}^1_{\mathcal{I}}(X,\omega)$. The rooftop structure also restricts to a rooftop structure on $\mathcal{TC}^1_{\mathcal{I}}(X,\omega)$.

We need the following constructions on test curves:

(1) **Increasing limit** Let $\psi^\alpha_\bullet \in \mathcal{TC}^1(X,\omega)$ be an increasing net. Assume that $\tau^+_{\psi^\alpha}$ is bounded from above. Define

$$\widetilde{\psi}_\tau := C\bigl[\sup_\alpha{}^* \psi^\alpha_\tau\bigr].$$

Let $\tau^+ = \inf\{\tau \mid \widetilde{\psi}_\tau = -\infty\}$. We define

$$\psi_\tau = \begin{cases} \widetilde{\psi}_\tau & \text{if } \tau \neq \tau^+, \\ \lim_{\sigma\to\tau^+-} \widetilde{\psi}_\sigma & \text{if } \tau \neq \tau^+. \end{cases}$$

It is easy to verify that $\psi_\bullet \in \mathcal{TC}^1(X,\omega)$.

(2) **Decreasing limit** Let $\psi^\alpha_\bullet \in \mathcal{TC}^1(X,\omega)$ be an increasing net and $\eta_\bullet \in \mathcal{TC}^1(X,\omega)$. Assume that $\psi^\alpha_\bullet \geq \eta_\bullet$ for all α. Define

$$(\inf \psi)_\tau := \inf_\alpha \psi^\alpha_\tau.$$

Then if $(\inf_\psi)_\bullet$ is not identically $-\infty$, then $(\inf\psi)_\bullet \in \mathcal{TC}^1(X,\omega)$.

(3) **Max** Let $\varphi_\bullet, \psi_\bullet \in \mathcal{TC}^1(X,\omega)$. There is the smallest test curve $(\varphi \vee \psi)_\bullet \in \mathcal{TC}^1(X,\omega)$ such that $(\varphi\vee\psi)_\bullet \geq \varphi_\bullet$ and $(\varphi\vee\psi)_\bullet \geq \psi_\bullet$. In fact, we could simply define

$$(\varphi\vee\psi)_\tau := \inf\{\eta_\tau \mid \eta_\bullet \in \mathcal{TC}^1(X,\omega), \eta_\bullet \geq \varphi_\bullet, \eta_\bullet \geq \psi_\bullet\}.$$

In terms of the Legendre transform, $(\varphi\vee\psi)\check{}$ is the minimal geodesic ray lying above both $\check{\varphi}$ and $\check{\psi}$. We observe that

$$d_1(\varphi_\bullet, \psi_\bullet) \leq d_1(\varphi_\bullet, (\varphi\vee\psi)_\bullet) + d_1(\psi_\bullet, (\varphi\vee\psi)_\bullet) \leq C_0 d_1(\varphi_\bullet, \psi_\bullet) \tag{7-15}$$

for some $C_0(n) > 0$. See [29, Proposition 2.15] for the proof of the latter inequality. Moreover, if $\eta_\bullet \in \mathcal{TC}^1(X,\omega)$ and if $\varphi_\bullet \leq \psi_\bullet$, then

$$d_1((\varphi\vee\eta)_\bullet, (\psi\vee\eta)_\bullet) \leq d_1(\varphi_\bullet, \psi_\bullet). \tag{7-16}$$

This follows from the corresponding inequality of geodesic rays, which in turn follows from Proposition 4.12 of [54] (Proposition 6.8 in the arXiv version).

We observe that the operator $\vee$ is associative and commutative; hence, we could also define $\psi^1_\bullet \vee \cdots \vee \psi^k_\bullet$ in the obvious way.

Lemma 7.15 *Let $\psi^j_\bullet, \psi_\bullet \in \mathcal{TC}^1(X,\omega)$. Assume that one of the following conditions holds*:

(1) *$\psi^j_\bullet$ is increasing and $\psi_\bullet$ is the increasing limit of $\psi^j_\bullet$.*

(2) *$\psi^j_\bullet$ is decreasing and $\psi_\bullet = (\inf\psi)_\bullet$.*

Then $\psi^j_\bullet \xrightarrow{d_1} \psi_\bullet$.

Proof We assume that condition (2) holds; the other case is similar. First observe that $\tau^+_{\psi^j} \to \tau^+_{\psi}$. It suffices to observe that

$$d_1(\psi^j_\bullet, \psi_\bullet) = (\tau^+_{\psi^j} - \tau^+_{\psi}) \int_X \omega^n + \int_{-\infty}^{\infty} \left(\int_X \omega^n_{\psi^j_\tau} - \int_X \omega^n_{\psi_\tau} \right) \mathrm{d}\tau.$$

The assertion is a simple consequence of dominated convergence theorem. □

Theorem 7.16 *The map* $\mathrm{DH}\colon \mathcal{E}^1(L^{\mathrm{an}}) \to \mathcal{M}(\mathbb{R})$ *is continuous.*

For any $\alpha \in \mathcal{E}^1(L^{\mathrm{an}})$,

$$\int_{\mathbb{R}} x \, \mathrm{dDH}(\alpha)(x) = \mathrm{E}(\alpha), \tag{7-17}$$

$$\int_{\mathbb{R}} \mathrm{DH}(\alpha) = \frac{1}{n!}(L^n). \tag{7-18}$$

Proof We first prove the continuity of DH.

By the dominated convergence theorem, it suffices to show that $G[\psi_\bullet](x)$ depends continuously on $\psi_\bullet$ for almost all $x \in \operatorname{Int}\Delta(L)$. To be more precise, let $\psi^j_\bullet \in \mathcal{TC}^1_{\mathcal{I}}(X,\omega)$ be a sequence converging to $\psi_\bullet$. We want to show that

$$G[\psi^j_\bullet](x) \to G[\psi_\bullet](x)$$

for almost all $x \in \operatorname{Int}\Delta(L)$. We will reduce to the case where $\psi^j_\bullet$ is either increasing or decreasing. In these cases, it suffices to show that $G[\psi^j_\bullet] \to G[\psi_\bullet]$ in L^1. By (7-4) and (7-9), this amounts to showing that $\mathbf{E}(\psi^j_\bullet) \to \mathbf{E}(\psi_\bullet)$. The latter follows from Lemma 7.15.

In order to make the reduction, we will prove that after passing to a subsequence, there exists an increasing sequence $\varphi^j_\bullet \in \mathcal{TC}^1_{\mathcal{I}}(X,\omega)$ and a decreasing sequence $\eta^j_\bullet \in \mathcal{TC}^1_{\mathcal{I}}(X,\omega)$ such that $\varphi^j_\bullet \le \psi^j_\bullet \le \eta^j_\bullet$ and $\varphi^j_\bullet \xrightarrow{d_1} \psi_\bullet$, $\eta^j_\bullet \xrightarrow{d_1} \psi_\bullet$. In fact, we can relax the requirement to $\varphi^j_\bullet, \eta^j_\bullet \in \mathcal{TC}^1(X,\omega)$, not necessarily $\mathcal{I}$-model. Then it suffices to replace both test curves by their pointwise $\mathcal{I}$-projections, which satisfy the same conditions by [32, Theorem 3.18].

Up to subtracting a subsequence, we may assume that for all j,

$$d_1(\psi^j_\bullet, \psi_\bullet) \le 2^{-j}.$$

For $k \ge j \ge 0$, we set

$$\eta^{j,k}_\bullet := \psi^j_\bullet \vee \cdots \vee \psi^k_\bullet \in \mathcal{TC}^1(X,\omega).$$

Let $\eta^j_\bullet \in \mathcal{TC}^1(X,\omega)$ be the increasing limit of $\eta^{j,k}_\bullet$ as $k \to \infty$. We then have

$$\begin{aligned} d_1(\eta^{j,k}_\bullet, \psi_\bullet) \le{}& d_1(\psi_\bullet, (\psi \vee \psi^j)_\bullet) + d_1((\psi \vee \psi^j)_\bullet, (\psi \vee \psi^j \vee \psi^{j+1})_\bullet) + \cdots \\ &+ d_1((\psi \vee \psi^j \vee \cdots \vee \psi^{k-1})_\bullet, (\psi \vee \psi^j \vee \cdots \vee \psi^k)_\bullet) \\ \le{}& d_1(\psi_\bullet, (\psi \vee \psi^j)_\bullet) + \cdots + d_1(\psi_\bullet, (\psi \vee \psi^k)_\bullet) \\ \le{}& C_0 \sum_{i=j}^{k} d_1(\psi_\bullet, \psi^i_\bullet) \le C_0 2^{1-j}. \end{aligned}$$

Here the second inequality follows from (7-16), the third inequality follows from (7-15). Then by Lemma 7.15, we find that $d_1(\eta^j_\bullet, \psi_\bullet) \le C2^{1-j}$. Thus, $\eta^j_\bullet \xrightarrow{d_1} \psi_\bullet$.

Similarly, for $k \ge j \ge 0$, let

$$\varphi^{j,k}_\bullet := \psi^j_\bullet \wedge \cdots \wedge \psi^k_\bullet \in \mathcal{TC}^1(X,\omega).$$

The same argument as above shows that for $k \ge j \ge 0$, $d_1(\varphi^{j,k}_\bullet, \psi_\bullet) \le 2^{1-j}$. Let

$$\psi^j_\tau := \inf_{k \ge j} \varphi^{j,k}_\tau.$$

By the monotone convergence theorem, $\psi^j \in \mathcal{TC}^1(X,\omega)$. Thus, by Lemma 7.15, $d_1(\varphi^j_\bullet, \psi_\bullet) \le 2^{1-j}$.

Next we prove (7-17). Let $\alpha \in \mathcal{E}^1(L^{\mathrm{an}})$. Let $\psi_\bullet$ be the test curve corresponding to α. We need to compute

$$\int_{\mathbb{R}} x\,\mathrm{DH}(\alpha)(x) = \int_{\Delta(L)} G[\psi_\bullet]\,\mathrm{d}\lambda.$$

By (7-3), (7-4) and (7-9), the right-hand side is just $\mathbf{E}(\psi_\bullet)$, which is equal to $\mathrm{E}(\alpha)$ by (7-10) and (7-11).

Finally, (7-18) follows from (7-8). □

Remark 7.17 On the subspace $\mathcal{H}^{\mathrm{NA}}$, the Duistermaat–Heckman measure is the same as the one defined in [16, Section 3.2]. This follows from Theorem 5.28 and [11, Theorem A]. On the other hand, in [40, Definition 3.56], Inoue defined the Duistermaat–Heckman measure for a general non-Archimedean metric on L^{an}. As explained in [40, Remark 1.4], his definition agrees with ours for metrics in $\mathcal{E}^1(L^{\mathrm{an}})$.

8 Toric setting

This section is devoted to a toric interpretation of the partial Okounkov body construction.

8.1 Technical lemmata

Lemma 8.1 *Let $\alpha, \beta_1, \dots, \beta_m \in \mathbb{Z}^n$. Let Δ be the convex polytope generated by $\beta_1, \dots, \beta_m$. Then the following are equivalent*:

(1) *The function*

(8-1) $$|z^\alpha|^2 \left(\sum_{i=1}^m |z^{\beta_i}|^2 \right)^{-1}$$

is bounded on $\mathbb{C}^{*n}$.

(2) $\alpha \in \Delta$.

Proof (2) $\Longrightarrow$ (1) Write $\alpha = \sum_i t_i \beta_i$, where $t_i \in [0,1]$ and $\sum_i t_i = 1$. Then

$$|z^\alpha|^2 \left(\sum_{i=1}^m |z^{\beta_i}|^2 \right)^{-1} = \prod_i |z^{\beta_i}|^{2t_i} \left(\sum_{i=1}^m |z^{\beta_i}|^2 \right)^{-1} \le \prod_i \sum_j |z^{\beta_j}|^{2t_i} \left(\sum_{i=1}^m |z^{\beta_i}|^2 \right)^{-1} \le 1.$$

(1) $\Longrightarrow$ (2) Assume that $\alpha \notin \Delta$. Let H be a hyperplane that separates α and Δ. Say H is defined by $a_1x_1 + \cdots + a_nx_n = C$. Set

$$z(t) := (t^{a_1}, \dots, t^{a_n}).$$

Then clearly (8-1) evaluated at $z(t)$ is not bounded. □

Lemma 8.2 *Let $\beta_1, \dots, \beta_m \in \mathbb{N}^n$ and $\beta \in \mathbb{R}^n$. Then the following are equivalent*:

(1) $\log \sum_{i=1}^m \mathrm{e}^{x\cdot\beta_i} - (x, \beta)$ *is bounded from below.*

(2) β *is in the convex hull of the* β_i.

Proof The proof follows the same pattern as Lemma 8.1. □

8.2 Toric Okounkov bodies

Let X be an n-dimensional smooth projective toric variety, corresponding to a smooth complete fan Σ in $N_{\mathbb{R}} \cong \mathbb{R}^n$. Let N be the lattice in $N_{\mathbb{R}}$, whose dual is the character lattice M. Let $T := N \otimes_{\mathbb{Z}} \mathbb{C}^*$ be the corresponding torus. Define $M_{\mathbb{R}} = N_{\mathbb{R}}^{\vee}$. Given any T-invariant divisor D on X, let $P_D \subseteq M_{\mathbb{R}}$ be the polyhedron associated with D.

Let $D_1, \dots, D_s$ be the class of prime T-invariant divisors on X, each corresponding to a ray ρ_i in Σ. Let v_i be the primitive generator of ρ_i. Any T-invariant admissible flag $Y_\bullet$ has the following form after renumbering the D_i:

$$Y_i = D_1 \cap \cdots \cap D_i.$$

Now the v_i induce an isomorphism $\Phi\colon M \to \mathbb{Z}^n$, $u \mapsto ((u, v_i))_i$. Let $\Phi_{\mathbb{R}}\colon M_{\mathbb{R}} \to \mathbb{R}^n$ be the extension of ϕ to $M_{\mathbb{R}}$ and σ be the cone generated by the v_i. Let U_σ be the corresponding orbit of T. Given any T-invariant line bundle, there is a unique T-invariant divisor D with $D|_{U_\sigma} = 0$ such that $\mathcal{O}_X(D) = L$.

It is shown in [45, Proposition 6.1] that

$$\Gamma_k(L) = \Phi_{\mathbb{R}}\left((kP_D) \cap M\right) \tag{8-2}$$

for sufficiently divisible k. We will omit $\Phi_{\mathbb{R}}$ from out notations from now on.

Let T_c be the compact torus in T. Next consider a T_c-invariant metric ϕ on L. An unpublished result of Yi Yao says that in the toric setting, two invariant potentials ϕ' and ϕ'' are $\mathcal{I}$-equivalent if and only if $\overline{\nabla\phi'_{\mathbb{R}}(\mathbb{R}^n)} = \overline{\nabla\phi''_{\mathbb{R}}(\mathbb{R}^n)}$. In other words, in the toric setting, for the invariant potentials, the $P[\bullet]$-envelope is the same as the $P[\bullet]_{\mathcal{I}}$-envelope. In particular,

$$\operatorname{vol}(L, \phi) = \frac{1}{n!}\int_X (\mathrm{dd^c}\phi)^n$$

always holds, without having to take the $P[\bullet]_{\mathcal{I}}$-envelope. For the proof of a more general result, we refer to [7, Theorem 3.13, Proposition 3.11].

Let U_0 be the maximal orbit of T. The basis (v_i) allows us to identify $U_0 = \mathbb{C}^{*n}$. We denote the coordinates on $\mathbb{C}^{*n}$ by $(z_1, \dots, z_n)$, $z_i = x_i + \mathrm{i} y_i$. Fix a T-invariant section s_0 of L on U_0 corresponding to D. Then we can identify ϕ with a T_c-invariant function on U_0. Given the identification $U_0 = \mathbb{C}^{*n}$, ϕ can be identified with a convex function $\phi_{\mathbb{R}} : \mathbb{R}^n \to \mathbb{R}$ such that $\nabla \phi_{\mathbb{R}} \subseteq P_D$. We let $P_{D,\phi}$ be the closure of the image of $\nabla\phi$. By [3, Lemma 2.5], $P_{D,\phi}$ corresponds to the closure of

$$Q_{D,\phi} := \{ y \in M_{\mathbb{R}} \mid \phi(x) - (x, y) \text{ is bounded from below} \}.$$

We will be more explicit at this point. Assume that

$$\phi = \log \sum_{i=1}^{a} |s_i|^2 + \mathcal{O}(1),$$

where $s_i \in \mathrm{H}^0(X, L)$. Let β_i be the lattice points in P_D corresponding to s_i. In this case, $Q_{D,\phi}$ is just the convex polytope generated by the β_i by Lemma 8.2.

Consider $\alpha \in M \cap P_D$. It corresponds to a Laurent polynomial z^α on $\mathbb{C}^{*n}$. Observe that $\alpha \in Q_{D,\phi}$ if and only if $|z^\alpha|^2 \mathrm{e}^{-\phi}$ is bounded from above. This is just a reformulation of Lemma 8.1.

Thus, we find

$$\Gamma_k(W^0(L, \phi)) = (k Q_{D,\phi}) \cap M \tag{8-3}$$

when k is sufficiently divisible. Hence, $\Delta(L, \phi) \supseteq P_{D,\phi}$. Comparing the volumes, we find that equality holds.

Next we deal with T_c-invariant ϕ such that $\mathrm{dd^c}\phi$ is a Kähler current. Let ϕ^j be an equivariant quasi-equisingular approximation of ϕ constructed as in [34, Corollary 13.23]. Then by definition,

$$\Delta(L, \phi) = \bigcap_j \Delta(L, \phi^j).$$

On the other hand,

$$P_{D,\phi} \subseteq \bigcap_j P_{D,\phi^j}.$$

Hence, $P_{D,\phi} \subseteq \Delta(L, \phi)$. On the other hand, the volume of both sides agree, so they are indeed equal thanks to the assumption that ϕ has analytic singularities.

In general, if ϕ is T_c-invariant and has positive volume. Let $\psi \le \phi$ be a potential with $\mathrm{dd^c}\psi$ being a Kähler current. We may guarantee that ψ is T_c-invariant. Then by definition, if we set $\phi_\epsilon = (1-\epsilon)\phi + \epsilon\psi$, then

$$\Delta(L, \phi) = \overline{\bigcup_{\epsilon \in (0,1)} \Delta(L, \phi_\epsilon)},$$

while

$$P_{D,\phi} \supseteq \overline{\bigcap_{\epsilon} P_{D, P[\phi_\epsilon]_{\mathcal{I}}}}.$$

Thus, $\Delta(L, \phi) \supseteq P_{D,\phi}$. Comparing the volumes, we find that these convex bodies are equal.

Theorem 8.3 *Let ϕ be a T_c-invariant psh metric on L with positive volume. Then*

$$\Delta(L,\phi) = P_{D,\phi}$$

under the identification $\Phi_{\mathbb{R}}$ as above.

8.3 Mixed volumes of line bundles

Let X and T be as in Section 8.2.

Lemma 8.4 *Let $L_1,\dots,L_n$ be big and nef T-invariant line bundles on X. Assume that the flag is T-invariant. Then*

$$\frac{1}{n!}(L_1,\dots,L_n) = \operatorname{vol}(\Delta(L_1),\dots,\Delta(L_n)). \tag{8-4}$$

Here vol denotes the mixed volume functional. We refer to [50, Section 5.1] for the precise definition.

As pointed out by Rémi Reboulet, this result is already proved in [18, Proposition 3.4.3].

Proof **Step 1** We first assume that all the L_i are ample.

In this case, we know that for any $t_i \in \mathbb{N}$ for $i = 1,\dots,n$,

$$\Delta\Big(\sum_{i=1}^n t_i L_i\Big) = \sum_{i=1}^n t_i \Delta(L_i)$$

by [42, Theorem 3.1]. Hence,

$$\operatorname{vol}\Delta\Big(\sum_{i=1}^n t_i L_i\Big) = \sum_{\alpha\in\mathbb{N}^n,|\alpha|=n} \binom{n}{\alpha} t^\alpha \operatorname{vol}\big(\Delta(L_1)^{\alpha_1},\dots,\Delta(L_n)^{\alpha_n}\big).$$

On the other hand, by (1-3),

$$\operatorname{vol}\Delta\Big(\sum_{i=1}^n t_i L_i\Big) = \frac{1}{n!}\sum_{\alpha\in\mathbb{N}^n,|\alpha|=n} \binom{n}{\alpha} t^\alpha (L_1^{\alpha_1},\dots,L_n^{\alpha_n}).$$

Comparing the coefficients, we find (8-4).

Step 2 General case.

The results of Step 1 generalize immediately to ample $\mathbb{Q}$-divisors. Hence, the nef case follows from a simple perturbation argument. □

The following example is due to Chen Jiang.

Example 8.5 If the flag is not toric invariant, Lemma 8.4 fails. For example, consider $X = \mathbb{P}^1 \times \mathbb{P}^1$, $L_1 = \mathcal{O}(1,2)$ and $L_2 = \mathcal{O}(2,1)$. Take a flag $X = Y_0 \supseteq Y_1 \supseteq Y_2$ with Y_1 being the diagonal. In this case, (8-4) fails.

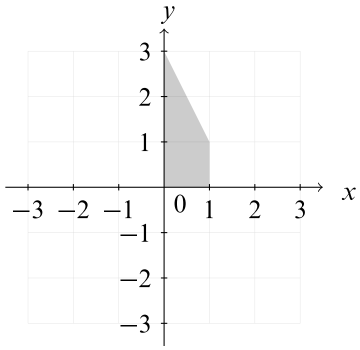


Figure 1: Okounkov body.

In this case, $(L_1, L_2) = 5$. By a simple computation using [45, Theorem 6.4], we find $\Delta(L_1) = \Delta(L_2)$ is the trapezoid shown in Figure 1. In particular,

$$\operatorname{vol}(\Delta(L_1), \Delta(L_2)) = 2 < \frac{5}{2!}.$$

For simplicity, we call (L, ϕ) a *T-invariant Hermitian big line bundle* on X if (T, ϕ) is a Hermitian big line bundle on X, L is T-invariant and ϕ is T_c-invariant.

Corollary 8.6 *Let (L_i, ϕ_i) for $i = 1, \dots, n$ be T-invariant Hermitian big line bundles on X with positive volumes. If the T-invariant flag satisfies that Y_n is not contained in any of the polar loci of the ϕ_i, then*

$$\frac{1}{n!}\int_X \mathrm{dd^c}\phi_1 \wedge \dots \wedge \mathrm{dd^c}\phi_n = \operatorname{vol}\bigl(\Delta(L_1, \phi_1), \dots, \Delta(L_n, \phi_n)\bigr). \tag{8-5}$$

Proof According to Proposition 5.16, by perturbing L_i, we may assume that each $\mathrm{dd^c}\phi_i$ is a Kähler current.

Observe that both sides of (8-5) are continuous under d_S-approximations of ϕ_i: the left-hand side follows from Theorem 4.2 and the right-hand side follows from Theorem 5.9.

Hence, by [33, Lemma 3.7], we may assume that each ϕ_i has analytic singularities. Taking a birational resolution, we may assume that ϕ_i has analytic singularities along normal crossing $\mathbb{Q}$-divisor E_i. By Remark 5.1, we reduce to the situation of Lemma 8.4. □

We have finished the proof of Theorem D.

Corollary 8.7 *Let $L_1, \dots, L_n$ be big T-invariant line bundles on X. Assume that the flag $(Y_\bullet)$ is T-invariant and Y_n is not contained in the non-Kähler locus of any $c_1(L_i)$. Then*

$$\frac{1}{n!}\langle L_1, \dots, L_n\rangle = \operatorname{vol}\bigl(\Delta(L_1), \dots, \Delta(L_n)\bigr). \tag{8-6}$$

Here $\langle \bullet \rangle$ denotes the movable intersection in the sense of [12; 15].

Proof It suffices to apply Corollary 8.6 to the case where ϕ_i has minimal singularities. □

Finally, we propose the following conjecture concerning the mixed volume of partial Okounkov bodies in the nontoric setting:

Conjecture 8.8 *Let (L_i, ϕ_i) for $i = 1, \dots, n$ be Hermitian big line bundles on X (not necessarily a toric variety) with positive volumes. Then*

$$\frac{1}{n!}\int_X \mathrm{dd^c}\phi_1 \wedge \cdots \wedge \mathrm{dd^c}\phi_n = \sup_\nu \operatorname{vol}\big(\Delta_\nu(L_1, \phi_1), \dots, \Delta_\nu(L_n, \phi_n)\big), \tag{8-7}$$

where ν runs over all rank n valuations $\mathbb{C}(X)^\times \to \mathbb{Z}^n$.

To the best of the author's knowledge, this conjecture is open even when the ϕ_i have minimal singularities.

References

[1] **O Amini**, **H Iriarte**, *Geometry of higher rank valuations*, preprint (2022) arXiv 2208.06237

[2] **V G Berkovich**, *Spectral theory and analytic geometry over non-Archimedean fields*, Math. Surv. Monogr. 33, Amer. Math. Soc., Providence, RI (1990) MR Zbl

[3] **R J Berman**, **B Berndtsson**, *Real Monge–Ampère equations and Kähler–Ricci solitons on toric log Fano varieties*, Ann. Fac. Sci. Toulouse Math. 22 (2013) 649–711 MR Zbl

[4] **R Berman**, **S Boucksom**, *Growth of balls of holomorphic sections and energy at equilibrium*, Invent. Math. 181 (2010) 337–394 MR Zbl

[5] **R J Berman**, **S Boucksom**, **M Jonsson**, *A variational approach to the Yau–Tian–Donaldson conjecture*, J. Amer. Math. Soc. 34 (2021) 605–652 MR Zbl

[6] **R Berman**, **S Boucksom**, **D Witt Nyström**, *Fekete points and convergence towards equilibrium measures on complex manifolds*, Acta Math. 207 (2011) 1–27 MR Zbl

[7] **A M Botero**, **J I Burgos Gil**, **D Holmes**, **R de Jong**, *Chern–Weil and Hilbert–Samuel formulae for singular Hermitian line bundles*, Doc. Math. 27 (2022) 2563–2624 MR Zbl

[8] **S Boucksom**, *Cônes positifs des variétés complexes compactes*, PhD thesis, Université Joseph-Fourier-Grenoble I (2002)

[9] **S Boucksom**, *Higher-dimensional Zariski decompositions*, preprint (2002) arXiv math/0204336

[10] **S Boucksom**, *Singularities of plurisubharmonic functions and multiplier ideals*, preprint (2017) Available at `http://sebastien.boucksom.perso.math.cnrs.fr/notes/L2.pdf`

[11] **S Boucksom**, **H Chen**, *Okounkov bodies of filtered linear series*, Compos. Math. 147 (2011) 1205–1229 MR Zbl

[12] **S Boucksom**, **J-P Demailly**, **M Păun**, **T Peternell**, *The pseudo-effective cone of a compact Kähler manifold and varieties of negative Kodaira dimension*, J. Algebraic Geom. 22 (2013) 201–248 MR Zbl

[13] **S Boucksom**, **P Eyssidieux**, **V Guedj**, **A Zeriahi**, *Monge–Ampère equations in big cohomology classes*, Acta Math. 205 (2010) 199–262 MR Zbl

[14] **S Boucksom**, **C Favre**, **M Jonsson**, *Valuations and plurisubharmonic singularities*, Publ. Res. Inst. Math. Sci. 44 (2008) 449–494 MR Zbl

[15] **S Boucksom**, **C Favre**, **M Jonsson**, *Differentiability of volumes of divisors and a problem of Teissier*, J. Algebraic Geom. 18 (2009) 279–308 MR Zbl

[16] **S Boucksom**, **T Hisamoto**, **M Jonsson**, *Uniform K-stability, Duistermaat–Heckman measures and singularities of pairs*, Ann. Inst. Fourier (Grenoble) 67 (2017) 743–841 MR Zbl

[17] **S Boucksom**, **M Jonsson**, *Global pluripotential theory over a trivially valued field*, Ann. Fac. Sci. Toulouse Math. 31 (2022) 647–836 MR Zbl

[18] **J I Burgos Gil**, **P Philippon**, **M Sombra**, *Arithmetic geometry of toric varieties: metrics, measures and heights*, Astérisque 360, Soc. Math. France, Paris (2014) MR Zbl

[19] **J Cao**, *Numerical dimension and a Kawamata–Viehweg–Nadel-type vanishing theorem on compact Kähler manifolds*, Compos. Math. 150 (2014) 1869–1902 MR Zbl

[20] **X Chen**, **J Cheng**, *On the constant scalar curvature Kähler metrics*, *III*: *General automorphism group*, preprint (2018) arXiv 1801.05907

[21] **S R Choi**, **Y Hyun**, **J Park**, **J Won**, *Asymptotic base loci via Okounkov bodies*, Adv. Math. 323 (2018) 784–810 MR Zbl

[22] **S R Choi**, **Y Hyun**, **J Park**, **J Won**, *Okounkov bodies associated to pseudoeffective divisors*, J. Lond. Math. Soc. 97 (2018) 170–195 MR Zbl

[23] **C Ciliberto**, **M Farnik**, **A Küronya**, **V Lozovanu**, **J Roé**, **C Shramov**, *Newton–Okounkov bodies sprouting on the valuative tree*, Rend. Circ. Mat. Palermo 66 (2017) 161–194 MR Zbl

[24] **T Darvas**, *Geometric pluripotential theory on Kähler manifolds*, from "Advances in complex geometry" (Y A Rubinstein, B Shiffman, editors), Contemp. Math. 735, Amer. Math. Soc., Providence, RI (2019) 1–104 MR Zbl

[25] **T Darvas**, **E Di Nezza**, **C H Lu**, *L^1 metric geometry of big cohomology classes*, Ann. Inst. Fourier (Grenoble) 68 (2018) 3053–3086 MR Zbl

[26] **T Darvas**, **E Di Nezza**, **C H Lu**, *Monotonicity of nonpluripolar products and complex Monge–Ampère equations with prescribed singularity*, Anal. PDE 11 (2018) 2049–2087 MR Zbl

[27] **T Darvas**, **E Di Nezza**, **C H Lu**, *On the singularity type of full mass currents in big cohomology classes*, Compos. Math. 154 (2018) 380–409 MR Zbl

[28] **T Darvas**, **E Di Nezza**, **C H Lu**, *Log-concavity of volume and complex Monge–Ampère equations with prescribed singularity*, Math. Ann. 379 (2021) 95–132 MR Zbl

[29] **T Darvas**, **E Di Nezza**, **H-C Lu**, *The metric geometry of singularity types*, J. Reine Angew. Math. 771 (2021) 137–170 MR Zbl

[30] **T Darvas**, **C H Lu**, *Geodesic stability, the space of rays and uniform convexity in Mabuchi geometry*, Geom. Topol. 24 (2020) 1907–1967 MR Zbl

[31] **T Darvas**, **R Reboulet**, **D Witt Nyström**, **M Xia**, **K Zhang**, *Transcendental Okounkov bodies*, preprint (2023) arXiv 2309.07584

[32] **T Darvas**, **M Xia**, *The closures of test configurations and algebraic singularity types*, Adv. Math. 397 (2022) art. id. 108198 MR Zbl

[33] **T Darvas**, **M Xia**, *The volume of pseudoeffective line bundles and partial equilibrium*, Geom. Topol. 28 (2024) 1957–1993 MR Zbl

[34] **J-P Demailly**, *Analytic methods in algebraic geometry*, Surv. Modern Math. 1, International, Somerville, MA (2012) MR Zbl

[35] **J-P Demailly**, *On the cohomology of pseudoeffective line bundles*, from "Complex geometry and dynamics" (J E Fornæss, M Irgens, E F Wold, editors), Abel Symp. 10, Springer (2015) 51–99 MR Zbl

[36] **J-P Demailly**, **T Peternell**, **M Schneider**, *Pseudo-effective line bundles on compact Kähler manifolds*, Int. J. Math. 12 (2001) 689–741 MR Zbl

[37] **Y Deng**, *Transcendental Morse inequality and generalized Okounkov bodies*, Algebr. Geom. 4 (2017) 177–202 MR Zbl

[38] **K Fujiwara**, **F Kato**, *Foundations of rigid geometry, I*, Eur. Math. Soc., Zürich (2018) MR Zbl

[39] **V Guedj**, **A Zeriahi**, *Degenerate complex Monge–Ampère equations*, EMS Tracts Math. 26, Eur. Math. Soc., Zürich (2017) Zbl

[40] **E Inoue**, *Entropies in μ-framework of canonical metrics and K-stability, II: Non-Archimedean aspect: non-Archimedean μ-entropy and μK-semistability*, preprint (2022) arXiv 2202.12168

[41] **S-Y Jow**, *Okounkov bodies and restricted volumes along very general curves*, Adv. Math. 223 (2010) 1356–1371 MR Zbl

[42] **K Kaveh**, *Note on cohomology rings of spherical varieties and volume polynomial*, J. Lie Theory 21 (2011) 263–283 MR Zbl

[43] **K Kaveh**, **A G Khovanskii**, *Newton–Okounkov bodies, semigroups of integral points, graded algebras and intersection theory*, Ann. of Math. 176 (2012) 925–978 MR Zbl

[44] **A G Khovanskii**, *Newton polyhedron, Hilbert polynomial and sums of finite sets*, Funktsional. Anal. i Prilozhen. 26 (1992) 57–63 MR Zbl In Russian; translated in Funct. Anal. Appl. 26 (1992) 276–281

[45] **R Lazarsfeld**, **M Mustaţă**, *Convex bodies associated to linear series*, Ann. Sci. École Norm. Sup. 42 (2009) 783–835 MR Zbl

[46] **X Ma**, **G Marinescu**, *Holomorphic Morse inequalities and Bergman kernels*, Progr. Math. 254, Birkhäuser, Basel (2007) MR Zbl

[47] **A Okounkov**, *Brunn–Minkowski inequality for multiplicities*, Invent. Math. 125 (1996) 405–411 MR Zbl

[48] **A Okounkov**, *Why would multiplicities be log-concave?*, from "The orbit method in geometry and physics" (C Duval, L Guieu, V Ovsienko, editors), Progr. Math. 213, Birkhäuser, Boston, MA (2003) 329–347 MR Zbl

[49] **J Ross**, **D Witt Nyström**, *Analytic test configurations and geodesic rays*, J. Symplectic Geom. 12 (2014) 125–169 MR Zbl

[50] **R Schneider**, *Convex bodies: the Brunn–Minkowski theory*, 2nd edition, Encycl. Math. Appl. 151, Cambridge Univ. Press (2014) MR Zbl

[51] **D Witt Nyström**, *Test configurations and Okounkov bodies*, Compos. Math. 148 (2012) 1736–1756 MR Zbl

[52] **D Witt Nyström**, *Transforming metrics on a line bundle to the Okounkov body*, Ann. Sci. École Norm. Sup. 47 (2014) 1111–1161 MR Zbl

[53] **D Witt Nyström**, *Monotonicity of non-pluripolar Monge–Ampère masses*, Indiana Univ. Math. J. 68 (2019) 579–591 MR Zbl

[54] **M Xia**, *Mabuchi geometry of big cohomology classes*, J. Reine Angew. Math. 798 (2023) 261–292 MR Zbl

[55] **M Xia**, *Pluripotential-theoretic stability thresholds*, Int. Math. Res. Not. 2023 (2023) 12324–12382 MR Zbl

[56] **M Xia**, *Nonpluripolar products on vector bundles and Chern–Weil formulae*, Math. Ann. 390 (2024) 3239–3316 MR Zbl

[57] **V P Zaharjuta**, *Transfinite diameter*, *Chebyshev constants*, *and capacity for compacta in* $\mathbb{C}^n$, Mat. Sb. 96(138) (1975) 374–389 MR Zbl In Russian; translated in Math. USSR-Sb. 25 (1975) 350–364

Department of Mathematical Sciences, Chalmers Tekniska Högskola
Gothenburg, Sweden

Institute of Geometry and Physics, University of Science and Technology of China (USTC)
Shanghai, China

xiamingchen2008@gmail.com

http://mingchenxia.github.io/home/

Proposed: Gang Tian Received: 21 March 2022
Seconded: Mark Gross, Leonid Polterovich Revised: 5 March 2024

Geometry & Topology **29**:3 (2025) 1345–1440
DOI: 10.2140/gt.2025.29.1345
Published: 31 May 2025

Global homotopy theory via partially lax limits

SIL LINSKENS
DENIS NARDIN
LUCA POL

We provide new ∞-categorical models for unstable and stable global homotopy theory. We use the notion of partially lax limits to formalize the idea that a global object is a collection of G-objects, one for each compact Lie group G, which are compatible with the restriction–inflation functors. More precisely, we show that the ∞-category of global spaces is equivalent to a partially lax limit of the functor sending a compact Lie group G to the ∞-category of G-spaces. We also prove the stable version of this result, showing that the ∞-category of global spectra is equivalent to the partially lax limit of a diagram of G-spectra. Finally, the techniques employed in the previous cases allow us to describe the ∞-category of proper G-spectra for a Lie group G, as a limit of a diagram of H-spectra for H running over all compact subgroups of G.

55N91, 55P91; 18N70

1 Introduction

It has been noted since the beginning of equivariant homotopy theory that there are equivariant objects which exist uniformly and compatibly for all compact Lie groups in a certain family, and which exhibit extra functoriality. For example, given compact Lie groups Π and G, there exists a construction for the classifying space of G-equivariant Π-principal bundles which is uniform on the group G and which is functorial on all continuous group homomorphisms [Schwede 2018, Remark 1.1.29]. Similarly, there are uniform constructions for many equivariant cohomology theories, such as K-theory, cobordism and stable cohomotopy, just to mention a few. The objects exhibiting such a "global" behavior are the subject of study of *global homotopy theory*.

In this paper we provide a new ∞-categorical model for global homotopy theory by formalizing the idea that a global stable/unstable object is a collection of G-objects, one for each compact Lie group G, which are compatible with the restriction–inflation functors. The key categorical construction that we will use to make this slogan precise is that of a partially lax limit, which we recall below. The main result of our paper is that this construction agrees with the models of global homotopy theory considered in the literature. Specifically we will compare it to the models of [Gepner and Henriques 2007] and [Schwede 2018] in the unstable and stable case, respectively. We first present our result in the simpler context of unstable global homotopy theory, and then consider the stable analogue of our main result. Finally we discuss an application of the techniques developed in this paper to proper equivariant homotopy theory.

Unstable global homotopy theory

Global spaces were first proposed in [Gepner and Henriques 2007] as a powerful framework for studying the homotopy theory of topological stacks and topological groupoids, which in turn generalize orbifolds and complexes of groups. This homotopy theory records the isotropy data of such objects as a particular diagram of fixed-point spaces. To make this precise, [Gepner and Henriques 2007] defined the ∞-category of *global spaces* as the presheaf ∞-category

$$\mathcal{S}_{\mathrm{gl}} = \mathrm{Fun}(\mathrm{Glo}^{\mathrm{op}}, \mathcal{S}).$$

Here Glo is the ∞-category whose objects are all compact Lie groups G, and whose morphism spaces are given by $\hom(H, G)_{hG}$; the homotopy orbits of the conjugation G-action on the space of continuous group homomorphisms. In particular, a global space X consists of the data of a fixed-point space X^G for every compact Lie group G, which are functorial in all continuous group homomorphisms. Furthermore, the conjugation actions have been trivialized, reflecting the fact that spaces of isotropy are insensitive to inner automorphisms.

This definition is motivated by Elmendorf's theorem in equivariant homotopy theory, which states that the ∞-category of *G-spaces* $\mathcal{S}_G$ is equivalent to the presheaf ∞-category on the G-orbit category $\boldsymbol{O}_G$. Here $\mathcal{S}_G$ is defined as the ∞-categorical localization of G-CW-complexes at the homotopy equivalences, and

$\boldsymbol{O}_G$ is the full subcategory of G-spaces spanned by the transitive G-spaces G/H for a closed subgroup $H \subseteq G$.

There is in fact a strong connection between equivariant and global homotopy theory. Let Orb denote the wide subcategory of Glo spanned by the injective group homomorphisms. Gepner and Henriques [2007] observed that the slice ∞-category $\mathrm{Orb}_{/G}$ is equivalent to the G-orbit category $\boldsymbol{O}_G$. In particular, this allows us to define a restriction functor

$$\mathrm{res}_G : \mathcal{S}_{\mathrm{gl}} \to \mathrm{Fun}(\boldsymbol{O}_G^{\mathrm{op}}, \mathcal{S}) \simeq \mathcal{S}_G$$

by precomposing with forgetful functor $\boldsymbol{O}_G \simeq \mathrm{Orb}_{/G} \to \mathrm{Glo}$. Thus a global space has an associated underlying G-space for all compact Lie groups G. Furthermore, that all these G-spaces come from the same global object imposes strong compatibility conditions among them.

We would like to understand how to recover a global space X from its restrictions $\mathrm{res}_G X$ to all compact Lie groups G, together with the previously mentioned compatibility conditions. The precise sense in which this is possible requires the notion of a (partially) lax limit, which we now recall, following [Gepner et al. 2017] and [Berman 2024].

Partially lax limits

Let $\mathcal{I}$ be an ∞-category and consider a functor $F : \mathcal{I} \to \mathrm{Cat}_\infty$. Intuitively, the *lax limit of* F is the ∞-category $\mathrm{laxlim}\, F$ whose objects consist of the following data:

- an object $X_i \in F(i)$ for each $i \in \mathcal{I}$, and
- compatible morphisms $f_\alpha : F(\alpha)(X_i) \to X_j$ for every arrow $\alpha : i \to j$ in $\mathcal{I}$.

A morphism $\{X_i, f_\alpha\} \to \{X_i', f_\alpha'\}$ is a suitably natural collection of maps $\{g_i : X_i \to X_i'\}$. More precisely, $\mathrm{laxlim}\, F$ is the ∞-category of sections of the cocartesian fibration associated to F. For our description we will require that for certain arrows α in $\mathcal{I}$, the map f_α is an equivalence. We therefore fix a collection of edges $\mathcal{W} \subset \mathcal{I}$ which contains all equivalences and which is stable under homotopy and composition, and denote by $\mathcal{I}^\dagger$ the resulting marked ∞-category. The *partially lax limit* of F is then the subcategory of $\mathrm{laxlim}\, F$ spanned by those objects $(\{X_i\}, \{f_\alpha\})$ for which the canonical map f_α is an equivalence for all edges $\alpha \in \mathcal{W}$. Note that if $\mathcal{W}$ contains only equivalences, then we recover the lax limit of F. On the other hand, if $\mathcal{W}$ contains all edges, we recover the usual notion of the limit of F. In particular we obtain canonical functors

$$\lim F \to \mathrm{laxlim}^\dagger F \to \mathrm{laxlim}\, F,$$

which indicates that a partially lax limit interpolates between the limit and the lax limit of a diagram. For exposition's sake, we have only defined the partially lax limit of a functor with values in Cat_∞, but there are similar definitions if we replace Cat_∞ with $\mathrm{Cat}_\infty^\otimes$, the ∞-category of symmetric monoidal ∞-categories. We refer the reader to Section 4 for more details on this construction.

As mentioned, in this paper we show that a global space can be thought of as a compatible collection of G-spaces. We can formalize what "compatible" means using the language of partially lax limits. To this end, let $(\mathrm{Glo}^{\mathrm{op}})^{\dagger}$ denote the ∞-category $\mathrm{Glo}^{\mathrm{op}}$ where we marked all the edges in $\mathrm{Orb}^{\mathrm{op}} \subseteq \mathrm{Glo}^{\mathrm{op}}$, ie all the injective edges. We prove the following theorem, which summarizes the main result of Section 6.

Theorem 6.17 *There exists a functor $\mathcal{S}_{\bullet}\colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_{\infty}^{\otimes}$ which sends a compact Lie group G to the ∞-category of G-spaces $\mathcal{S}_G$ endowed with the cartesian symmetric monoidal structure, and a continuous group homomorphism $\alpha\colon H \to G$ to the restriction–inflation functors. Furthermore, there is a symmetric monoidal equivalence*

$$\mathcal{S}_{\mathrm{gl}} \simeq \operatorname*{laxlim}^{\dagger}_{G \in (\mathrm{Glo}^{\mathrm{op}})^{\dagger}} \mathcal{S}_G$$

between the ∞-category of global spaces with the cartesian monoidal structure and the partially lax limit over $(\mathrm{Glo}^{\mathrm{op}})^{\dagger}$ of the diagram $\mathcal{S}_{\bullet}$.

By the above theorem a global space X consists of the following data and conditions:

- A G-space $\mathrm{res}_G X$ for each compact Lie group G.
- An H-equivariant map $f_{\alpha}\colon \alpha^{*}\mathrm{res}_G X \to \mathrm{res}_H X$ for each continuous group homomorphism $\alpha\colon H \to G$.
- The maps f_{α} are functorial, so that $f_{\beta\circ\alpha} \simeq f_{\beta} \circ \beta^{*}(f_{\alpha})$ for all composable maps α and β, and $f_{\mathrm{id}} = \mathrm{id}$.
- The map f_{α} is an equivalence for every continuous *injective* homomorphism α.
- A homotopy between the map f_{c_g} induced by the conjugation isomorphism and the map given by left multiplication by g, denoted by $l_g\colon c_g^{*}\mathrm{res}_G X \to \mathrm{res}_G X$.
- Higher coherences for the homotopies.

This is a precise formulation of the compatibility conditions encoded in a global space.

Global stable homotopy theory

Our discussion so far has been limited to the homotopy theory of global spaces, but there are also numerous examples of equivariant cohomology theories exhibiting a global behavior. These cohomology theories are represented by global spectra, and their study is called *global stable homotopy theory*.

The consideration of "global spectra" grew out of the literature on equivariant stable homotopy theory, and was considered in works such as [Greenlees and May 1997]. Morally, a global spectrum models a compatible family of equivariant spectra for all compact Lie groups at once. Our main result makes this moral precise, and provides the same description as in the unstable case.

There are multiple models for the homotopy theory of global spectra. In this paper we will use the framework developed by Schwede [2018]. His approach has the advantage of being very concrete; the category of global spectra is modeled by the usual category of orthogonal spectra but with a finer notion of equivalence, the global equivalences. The category of orthogonal spectra with the global stable model structure of [Schwede 2018, Theorem 4.3.17] underlies a symmetric monoidal ∞-category $\mathrm{Sp}_{\mathrm{gl}}$. As any orthogonal spectrum is a global spectrum, this approach comes with a good range of examples. For instance, there are global analogues of the sphere spectrum, cobordism, topological and algebraic K-theory spectra, Borel cohomology, symmetric product spectra and many others. Global spectra have also been shown to give cohomology theories on orbifolds and topological stacks in [Juran 2020], thereby establishing them as a natural home for (genuine) cohomology theories on topological stacks. As part of the framework developed by Schwede, the ∞-category of global spectra comes with symmetric monoidal restriction functors

$$\mathrm{res}_G \colon \mathrm{Sp}_{\mathrm{gl}} \to \mathrm{Sp}_G$$

into the ∞-category of G-spectra, for all compact Lie groups G. As a first indication that a global spectrum should consist of just this data, together with various comparison maps, note that the functors res_G are jointly conservative by the very definition of global equivalences.

However, not all equivariant spectra admit global refinements. In fact being a "global" object forces strong compatibility conditions between the underlying G-spectra for different G. For example, $\mathrm{res}_G X$ is always a split G-spectrum by [Schwede 2018, Remark 4.1.2] and its G-homotopy groups for all G together admit the structure of a global functor, see [Schwede 2018, Example 4.2.3]. We can again formalize how a global spectrum is determined by its restrictions for all compact Lie groups using the language of partially lax limits. Recall that $(\mathrm{Glo}^{\mathrm{op}})^{\dagger}$ denotes the ∞-category $\mathrm{Glo}^{\mathrm{op}}$, marked by all the edges in $\mathrm{Orb}^{\mathrm{op}}$, ie the injective group homomorphisms.

Theorem 11.10 *There exists a functor $\mathrm{Sp}_{\bullet} \colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_{\infty}^{\otimes}$ which sends a compact Lie group G to the symmetric monoidal ∞-category of G-spectra $\mathrm{Sp}_G^{\otimes}$, and a continuous group homomorphism $\alpha \colon H \to G$ to the restriction–inflation functor. Furthermore, there is a symmetric monoidal equivalence*

$$\mathrm{Sp}_{\mathrm{gl}} \simeq \operatorname*{laxlim^{\dagger}}_{G \in (\mathrm{Glo}^{\mathrm{op}})^{\dagger}} \mathrm{Sp}_G$$

between Schwede's ∞-category of global spectra, and the partially lax limit over $(\mathrm{Glo}^{\mathrm{op}})^{\dagger}$ of the diagram $\mathrm{Sp}_{\bullet}$.

Proper equivariant stable homotopy theory

The techniques employed in the proof of Theorem 11.10 can also be used in other settings. Given a (not necessarily compact) Lie group G, we can consider the ∞-category of proper G-spectra $\mathrm{Sp}_{G,\mathrm{pr}}$. This is the ∞-category underlying the category of orthogonal G-spectra with the proper stable model structure of [Degrijse et al. 2023], in which a map $f: X \to Y$ is a weak equivalence if and only if for all compact subgroups $H \leq G$, the map induced on homotopy groups $\pi_*^H(f): \pi_*^H(X) \to \pi_*^H(Y)$ is an isomorphism. Write $\boldsymbol{O}_{G,\mathrm{pr}}$ for the proper G-orbit category, which is defined to be the subcategory of $\boldsymbol{O}_G$ spanned by the cosets G/H, where H is a compact subgroup of G. Our techniques allow us to prove:

Theorem 12.11 *Let G be a Lie group. There is a symmetric monoidal equivalence*

$$\mathrm{Sp}_{G,\mathrm{pr}} \simeq \lim_{H \in \boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}}} \mathrm{Sp}_H$$

between the ∞-category of proper G-spectra and the limit of the functor $\mathrm{Sp}_\bullet$ restricted along the canonical functor $\iota_G: \boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}} \to \mathrm{Glo}^{\mathrm{op}}$ sending G/H to H.

Having introduced the main theorems of this article. We continue the introduction by discussing the proof strategy for each in some detail.

The proof strategy for Theorem 6.17

We begin with a discussion of the proof of the unstable result. Implicit in [Rezk 2014] is the following crucial observation (see also Proposition 6.13): the space of factorizations of any map $\alpha: H \to G$ in Glo into a surjective followed by an injective group homomorphism is contractible. In fewer words, the surjective and injective maps form an orthogonal factorization system on Glo. This is the main ingredient in the proof of Theorem 6.17, and moreover, we would like to argue that it is at the core of the relationship between global and G-equivariant homotopy theory.

This claim is justified by the following two facts. The first is that the functoriality under the restriction–inflation functors of the different ∞-categories of equivariant spaces is equivalent to the previous observation. The second is that the observation formally implies that one can recover a global space X from the $\mathrm{Glo}^{\mathrm{op}}$-indexed diagram of G-spaces $\mathrm{res}_G X$.

Let us first explain how the ∞-categories of equivariant spaces are functorial in the category $\mathrm{Glo}^{\mathrm{op}}$. Due to the existence of a nontrivial topology on the morphism spaces, this is not immediate. For example, note that exhibiting this functoriality also entails giving a homotopy coherent trivialization of the conjugation action on $\mathcal{S}_G$. The key is that the existence of the orthogonal factorization system allows one to define functors

$$\alpha_!: \mathrm{Orb}_{/H} \to \mathrm{Orb}_{/G}, \quad (K \hookrightarrow H) \mapsto (\alpha(K) \hookrightarrow G).$$

On objects, $\alpha_!$ factorizes the composite $K \hookrightarrow H \to G$ into a surjection followed by an injection, and then only remembers the injective part. The fact that such factorizations are unique is equivalent to the fact that this functor is well-defined. Precomposing with $\alpha_!^{\mathrm{op}}$ gives the standard restriction functor $\alpha^*: \mathcal{S}_G \to \mathcal{S}_H$. Furthermore, given this description of the individual restriction functors, it is clear that they are functorial in $\mathrm{Glo}^{\mathrm{op}}$.

Next we explain how the observation implies that one can recover a global space from its restrictions. When one takes an object $(\{\mathrm{res}_G X\}, \{f_\alpha\})$ of the partially lax limit over $\mathrm{Glo}^\dagger$ of the diagram $\mathscr{S}_\bullet$, the functoriality of the associated global space in injections is recorded by restricting to each $\mathrm{res}_G X$, and the functoriality in surjections is given by the morphisms f_α. One recovers the functoriality in all morphisms in Glo by factorizing an arbitrary morphism into an injection followed by a surjection. The ability to split the functoriality in this way again reduces to the observation that the surjective and injective maps form an orthogonal factorization system. We make precise all of the ideas sketched here in Section 6.

The proof strategy for Theorem 11.10

The proof of Theorem 11.10 is considerably more involved than its unstable analogue, and takes up the majority of the second half of the paper. Therefore we now give an overview of the proof as a roadmap for the reader.

Firstly, we discuss the existence of the functor $\mathrm{Sp}_\bullet$. Recall that a G-spectrum can be thought of as a pointed G-space together with a compatible collection of deloopings for all representation spheres. With modern tools we can give this construction a universal property: as a symmetric monoidal ∞-category, Sp_G is obtained from the ∞-category of pointed G-spaces by freely inverting the representation spheres S^V for every G-representation V; see [Gepner and Meier 2023, Appendix C]. This universal property, combined with the unstable functor $\mathscr{S}_\bullet$ of Theorem 6.17, immediately gives the functoriality of G-spectra in $\mathrm{Glo}^{\mathrm{op}}$ as in our theorem.

Unfortunately, constructing the functor $\mathrm{Sp}_\bullet$ via the universal property of equivariant spectra is unhelpful for our purposes, as it is too inexplicit for calculating the partially lax limit. For example, note that for a surjective group homomorphism $\alpha\colon H \to G$ and G-spectrum E, to obtain the H-spectrum $\alpha^* E$ one has to freely add deloopings with respect to representation spheres not in the image of $\alpha^*\colon \mathrm{Rep}(G) \to \mathrm{Rep}(H)$. This is a process which one cannot easily control.

Therefore, pivotal to our proof is an explicit construction of the functor $\mathrm{Sp}_\bullet$. The calculation of the partially lax limit of $\mathrm{Sp}_\bullet$ will then follow from this by a long series of nontrivial formal arguments. The crucial idea is to construct and calculate with a functoriality on prespectrum objects rather than at the level of spectrum objects. In this setting, we are able to build the functoriality of equivariant prespectra explicitly using the functoriality of the ∞-categories $\boldsymbol{O}_G$ and $\mathrm{Rep}(G)$, the category of representations and linear isometries.

To make this precise, let us first specify our model of G-prespectra. We define an ∞-category $\mathbf{OR}_G$, naturally fibered over $\boldsymbol{O}_G^{\mathrm{op}}$, whose objects are pairs (H, V), where H is a closed subgroup of G and V is an H-representation; see Definition 8.5. This is canonically symmetric promonoidal and so the ∞-category of functors $\mathrm{Fun}(\mathbf{OR}_G, \mathscr{S}_*)$ is symmetric monoidal via Day convolution. There is a functor $S_G\colon \mathbf{OR}_G \to \mathscr{S}_*$ which sends the object (H, V) to the pointed space $(S^V)^H$. This is a commutative algebra object in $\mathrm{Fun}(\mathbf{OR}_G, \mathcal{S}_*)$ via the universal property of Day convolution. The first ingredient of the proof is the following:

Step 1 *The ∞-category Sp_G is equivalent to an explicit Bousfield localization of the ∞-category*

$$\mathrm{PSp}_G := \mathrm{Mod}_{S_G} \mathrm{Fun}(\mathbf{OR}_G, \mathcal{S}_*).$$

We obtain this description by reinterpreting the construction of G-spectra as a Bousfield localization of the level model structure on orthogonal G-spectra internally to ∞-categories. This identification is the culmination of Sections 7 and 8, and the reader can find a precise statement as Proposition 7.30 and Corollary 8.14.

Having obtained this identification, we can build the functoriality of equivariant prespectra by exhibiting the pairs $(\mathbf{OR}_G, S_G)$ as functorial in $\mathrm{Glo}^{\mathrm{op}}$. In fact the categories $\mathbf{OR}_G$ will only be (pro)functorial in $\mathrm{Glo}^{\mathrm{op}}$, but this is a subtlety which we choose to gloss over in this introduction. To exhibit this functoriality, we build a global version of the category $\mathbf{OR}_G$ and the algebra object S_G, which we denote by $\mathbf{OR}_{\mathrm{gl}}$ and S_{gl}; see Definition 9.2. The ∞-category $\mathbf{OR}_{\mathrm{gl}}$ is naturally fibered over $\mathrm{Glo}^{\mathrm{op}}$ and has objects (G, V), where G is a compact Lie group and V is a G-representation, and $S_{\mathrm{gl}} \colon \mathbf{OR}_{\mathrm{gl}} \to \mathcal{S}_*$ sends (G, V) to the pointed space $(S^V)^G$.

There is a precise sense in which the pair $(\mathbf{OR}_{\mathrm{gl}}, S_{\mathrm{gl}})$ contain all of the functoriality of the pairs $(\mathbf{OR}_G, S_G)$ in Glo. For the group direction this stems from the fact that the surjections and injections form an orthogonal factorization system on Glo, while for the representation direction this follows from the observation that $\mathbf{OR}_{\mathrm{gl}}$ is a cocartesian fibration over $\mathrm{Glo}^{\mathrm{op}}$ classifying the functor $\mathrm{Rep}(-) \colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_\infty$ which sends a compact Lie group G to its category of G-representations, with functoriality given by restriction. These observations allow us to prove the following result; see Proposition 9.16.

Step 2 *There exists a functor*

$$\mathrm{PSp}_\bullet \colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_\infty^{\otimes}, \quad G \mapsto \mathrm{PSp}_G.$$

Furthermore the partially lax limit of $\mathrm{PSp}_\bullet$ over $(\mathrm{Glo}^{\mathrm{op}})^\dagger$ is given by $\mathrm{Mod}_{S_{\mathrm{gl}}} \mathrm{Fun}(\mathbf{OR}_{\mathrm{gl}}, \mathcal{S}_*)$.

We have shown in Step 1 that Sp_G is a Bousfield localization of PSp_G. We call a map in PSp_G a stable equivalence if it is inverted by the functor $\mathrm{PSp}_G \to \mathrm{Sp}_G$.

Step 3 *The diagram $\mathrm{PSp}_\bullet$ preserves stable equivalences, and therefore induces a diagram $\mathrm{Sp}_\bullet$. Furthermore, as indicated by the notation, this diagram is equivalent to the functoriality of equivariant spectra built at the beginning of this section using the universal property of Sp_G.*

In particular, on morphisms this diagram gives the standard restriction–inflation functors on equivariant spectra; see Corollary 10.6. The following result follows formally from this.

Step 4 *The partially lax limit of* $\mathrm{Sp}_\bullet$ *is given by an explicit Bousfield localization of the* ∞*-category*

$$\mathrm{Mod}_{S_{\mathrm{gl}}} \mathrm{Fun}(\mathbf{OR}_{\mathrm{gl}}, \mathcal{S}_*).$$

Finally, we compare this ∞-category to Schwede's model of global spectra, $\mathrm{Sp}_{\mathrm{gl}}$. Once again we do this by first translating his construction into one internal to ∞-categories. We define an ∞-category $\mathbf{OR}_{\mathrm{fgl}}$ as the subcategory of $\mathbf{OR}_{\mathrm{gl}}$ spanned by the objects (G, V), where V is a *faithful* G-representations. Restricting S_{gl} we obtain a commutative algebra object S_{fgl} in $\mathrm{Fun}(\mathbf{OR}_{\mathrm{fgl}}, \mathcal{S}_*)$. We then show:

Step 5 $\mathrm{Sp}_{\mathrm{gl}}$ *is an explicit Bousfield localization of the category* $\mathrm{Mod}_{S_{\mathrm{fgl}}}(\mathrm{Fun}(\mathbf{OR}_{\mathrm{fgl}}, \mathcal{S}_*))$.

The precise statement is obtained by combining Proposition 7.27 and Corollary 8.23. Finally we show in Section 11 that the canonical inclusion $j : \mathbf{OR}_{\mathrm{fgl}} \to \mathbf{OR}_{\mathrm{gl}}$ induces an adjunction

$$j_! : \mathrm{Mod}_{S_{\mathrm{fgl}}}(\mathrm{Fun}(\mathbf{OR}_{\mathrm{fgl}}, \mathcal{S}_*)) \leftrightarrows \mathrm{Mod}_{S_{\mathrm{gl}}}(\mathrm{Fun}(\mathbf{OR}_{\mathrm{gl}}, \mathcal{S}_*)) : j^*$$

on prespectrum objects. Then we show that this adjunction descends to an adjunction on the corresponding Bousfield localizations of Steps 4 and 5. Finally we prove that the fibrancy conditions imposed by these localizations cancel out the difference between all and faithful representations, so that we obtain an equivalence

$$\mathrm{Sp}_{\mathrm{gl}} \simeq \mathrm{laxlim}^{\dagger} \mathrm{Sp}_\bullet,$$

concluding the proof of Theorem 11.10.

Finally let us note that to fill in all of the details of this argument requires a long list of technical results about the relationship between various constructions applied to model categories and ∞-categories, Day convolution monoidal structures induced by promonoidal categories, and partially lax limits of symmetric monoidal categories. We have included these in Part I to make the paper self-contained, and because we failed to find a convenient reference for many of these facts.

Related work

There are many models of global unstable homotopy theory. The first was given in [Gepner and Henriques 2007], and since then others have been obtained in [Schwede 2018; 2020]. The second of these papers, together with [Körschgen 2018], proves that all these models induce the same ∞-category. Finally, we would like to mention the unpublished manuscript [Rezk 2014], which contains many of the ideas we exploit in Section 6.

There has been a lot of work towards finding a good framework for the study of global stable homotopy theory; see [Bohmann 2014; Greenlees and May 1997] and [Lewis et al. 1986, Chapter II]. Schwede's model [2018] has so far being the most successful one, in part because of its numerous applications to equivariant stable homotopy theory; see for example [Schwede 2017] and [Hausmann 2022]. Hausmann [2019] gave a model for global homotopy theory for the family of finite groups by endowing the category of symmetric spectra with a global model structure. There is also a model for G-global homotopy theory [Lenz 2025] which is a synthesis between classical equivariant homotopy theory and Schwede's global homotopy theory. This specializes to global homotopy theory by setting G to be the trivial group. Recently, Lenz [2022] gave an ∞-categorical model for global stable homotopy theory for the family of finite groups using spectral Mackey functors. However to the best of our knowledge, our model is the first ∞-categorical model for global stable homotopy theory for the family of all compact Lie groups and not just the finite ones.

Future directions

In this paper we focused only on global and proper equivariant homotopy theory, but it is quite natural to wonder if we can recover our two results as a special case of a more general one. For any Lie group G, we can in fact consider G-global homotopy theory which is a generalization of global and G-equivariant homotopy theory. We conjecture that G-global stable homotopy theory is equivalent to the partially lax limit of the functor $\mathrm{Sp}_\bullet$ restricted along the canonical functor $\mathrm{Glo}_{/G}^{\mathrm{op}} \to \mathrm{Glo}^{\mathrm{op}}$.

Organization of the paper

The paper is divided into three main parts.

In the first part we first discuss the relationship between model and ∞-categories. Then we recall the concept of a promonoidal ∞-category and use this to define the Day convolution product on functor categories. We then introduce the notions of partially lax (co)limits and collect various useful results that we will need throughout the paper. We finish Part I by describing the lax limits of symmetric monoidal ∞-categories in terms of the operadic norm functor.

The second part of the paper contains the proofs of our main results. In Section 6 we introduce the ∞-category of global spaces and prove Theorem 6.17. This is an unstable version of Theorem 11.10, and serves as a warm-up for the considerably more involved proof of the stable case. We therefore recommend that the reader read this section before moving forward. In Section 7 we recall various model structures on the categories of orthogonal G-spectra for a Lie group G, and hence define the underlying ∞-categories of proper G-spectra and of global spectra. In Section 8 we apply a variant of Elmendorf's theorem and use this to provide specific models for the ∞-categories of proper G-prespectra and global prespectra. In Section 9 we construct the functor $\mathrm{Sp}_\bullet$ from the introduction, and in Section 11 we identify the partially lax limits with the ∞-category of global spectra. Finally in Section 12, we apply the same techniques to describe the ∞-category of proper G-spectra as a limit, proving Theorem 12.11.

The third part of the paper contains an appendix on the tensor product of modules in an ∞-category.

Acknowledgements The authors thank Stefan Schwede for first suggesting that a description of global spectra as a partially lax limit should be possible. We would also like to thank Stefan Schwede, Lennart Meier, Markus Hausmann, Branko Juran and Bastiaan Cnossen for numerous helpful conversations on this topic. We also thank the referees for helpful suggestions, which improved the paper.

During the preparation of this text Linskens was a member of the Hausdorff Center for Mathematics at the University of Bonn funded by the German Research Foundation (DFG). This material is based upon work supported by the Swedish Research Council under grant 2016-06596 while Pol was in residence at the Institut Mittag-Leffler in Djursholm, Sweden, during the semester *Higher algebraic structures in algebra, topology, and geometry*. Linskens was supported by the DFG Schwerpunktprogramm 1786 *Homotopy theory and algebraic geometry* (project ID SCHW 860/1-1). Pol was supported by the SFB 1085 *Higher invariants* in Regensburg.

Part I Partially lax limits, promonoidal ∞-categories and Day convolution

In this part of the paper we introduce the necessary machinery to state and prove our main results. In the first section we give references for the passage from topological/model categories to ∞-categories. We then discuss the Day convolution product for functor ∞-categories, where the source is only assumed to be a promonoidal ∞-category. Finally we recall the notion of partially lax limits of ∞-categories and symmetric monoidal ∞-categories, and proof some useful properties about them.

2 From topological/model categories to ∞-categories

In this paper we will often need to pass from topological categories (or operads) and (symmetric monoidal) model categories to ∞-categories. In this section we recall how this is done, and provide relevant references. After this section we will largely leave these identifications implicit for the rest of the paper.

2.1 Topological categories and operads

We can promote a topological category $\mathscr{C}$ to an ∞-category by first applying the singular functor to the mapping spaces (see [Lurie 2009, Section 1.1.4]) and then applying the coherent (also called simplicial) nerve functor [Lurie 2009, Corollary 1.1.5.12]. This defines a functor

$$\mathrm{TopCat} \to \mathrm{Cat}_\infty$$

from topological categories to ∞-categories. Importantly, applying this functor to a topologically enriched category $\mathscr{C}$ preserves the set of objects and the weak homotopy type of the mapping space between any two objects; see [Lurie 2009, Theorem 1.1.5.13]. Throughout this paper we will not distinguish between the topological category and its ∞-categorical counterpart.

There is a similar functorial construction between topological operads and ∞-operads, which we now recall. Given a topological colored operad $\mathbb{O}$, we let $\mathbb{O}^{\otimes}$ denote the topological category whose objects are pairs $(I_+, (C_i)_{i\in I})$, where $I_+ \in \mathrm{Fin}_*$ and C_i are colors in $\mathbb{O}$. Given a pair of objects $C = (I_+, \{C_i\}_{i\in I})$ and $D = (J_+, \{D_j\}_{j\in J})$ in $\mathbb{O}^{\otimes}$, the morphism space $\mathbb{O}^{\otimes}(C, D)$ is given by

$$\coprod_{\alpha\colon I_+\to J_+} \prod_{j\in J} \mathbb{O}(\{C_i\}_{\alpha(i)=j}, D_j).$$

Composition is defined in the obvious way. This is the topological analogue of [Lurie 2017, Notation 2.1.1.22]. Note that $\mathbb{O}^{\otimes}$ admits a functor to Fin_*. By the process before, this induces a functor of ∞-categories $\mathbb{O}^{\otimes} \to \mathrm{Fin}_*$.

Lemma 2.1 *Let $\mathbb{O}$ be a topological colored operad. Then the forgetful functor $p\colon \mathbb{O}^{\otimes} \to \mathrm{Fin}_*$ defines an ∞-operad. Moreover, this construction is functorial in the sense that it sends maps of topological colored operads to maps of ∞-operads.*

Proof Recall that a topological category is seen as an ∞-category by applying the singular functor on mapping spaces and then by applying the coherent nerve functor to the resulting simplicial category. Since the singular functor preserves products and sends every object to a fibrant one, it sends the topological colored operad $\mathbb{O}$ to a fibrant[1] simplicial operad $\mathbb{O}_s$. Moreover by direct inspection, the singular functor sends the topological category $\mathbb{O}^{\otimes}$ defined above to $\mathbb{O}_s^{\otimes}$ as defined in [Lurie 2017, Notation 2.1.1.22]. Applying the coherent nerve to $\mathbb{O}_s^{\otimes} \to \mathrm{Fin}_*$ we obtain an ∞-operad by [Lurie 2017, Proposition 2.1.1.27], proving the first claim. A simple check shows that the formation of the topological category $\mathbb{O}^{\otimes}$ is functorial in maps of topological operads. Applying the singular functor and the coherent nerve then gives a functor of ∞-categories over Fin_*. Furthermore the cocartesian edges over inert edges are explicitly constructed in the proof of [Lurie 2017, Proposition 2.1.1.27], and the functor constructed clearly preserves these edges. □

2.2 Model categories and ∞-categories

We will very often pass from model categories to ∞-categories. Therefore we explain and give references for this passage.

Let $\mathcal{M}$ be a model category with class of weak equivalences denoted by W. We always assume that $\mathcal{M}$ has functorial factorizations. The model category $\mathcal{M}$ presents an ∞-category which we denote by $\mathcal{M}[W^{-1}]$. We may define $\mathcal{M}[W^{-1}]$ as the Dwyer–Kan localization of $N(\mathcal{M})$ at the weak equivalences of $\mathcal{M}$, ie as the initial ∞-category with a functor from $\mathcal{M}$ which inverts the morphisms in W. Write $\mathcal{M}^f$, $\mathcal{M}^c$ and $\mathcal{M}^\circ$ for the full subcategories of $\mathcal{M}$ spanned by the fibrant, cofibrant and bifibrant objects, respectively. The composite

$$N(\mathcal{M}^f) \to N(\mathcal{M}) \to \mathcal{M}[W^{-1}]$$

[1] Recall that a simplicial operad is fibrant if each multispace is a fibrant simplicial set; see [Lurie 2017, Definition 2.1.1.26].

is a Dwyer–Kan localization at the restriction of W to $\mathcal{M}^f$, and similarly for the case of cofibrant and bifibrant objects. See for example the discussion in [Lurie 2017, Remark 1.3.4.16].

If $\mathcal{M}$ is a topological model category, then the enriched structure gives another construction of $\mathcal{M}[W^{-1}]$. In this case, $\mathcal{M}[W^{-1}]$ is equivalent to the ∞-category associated to the topologically enriched category $\mathcal{M}^\circ$ as in the previous section; see [Lurie 2017, Theorem 1.3.4.20]. Throughout our paper it will be necessary to use all these different constructions of $\mathcal{M}[W^{-1}]$.

We note that if the model category $\mathcal{M}$ is cofibrantly generated and the underlying category is locally presentable, then $\mathcal{M}[W^{-1}]$ is a presentable ∞-category; see [Lurie 2017, Proposition 1.3.4.22]. Also we note that any Quillen adjunction of model categories $F\colon \mathcal{M}_0 \rightleftarrows \mathcal{M}_1 :G$ induces an adjunction of underlying ∞-categories $F\colon \mathcal{M}_0[W_0^{-1}] \rightleftarrows \mathcal{M}_1[W_1^{-1}] :G$ by [Hinich 2016, Proposition 1.5.1].

Next we may consider symmetric monoidal model categories. By [Lurie 2017, Proposition 4.1.7.6], if $\mathcal{M}$ is a symmetric monoidal model category then the ∞-category $\mathcal{M}[W^{-1}]$ admits a symmetric monoidal structure such that the localization functor $\mathcal{M}^c \to \mathcal{M}[W^{-1}]$ is strong monoidal, and if F is a symmetric monoidal left Quillen functor then F is again symmetric monoidal.

Once again we obtain a different construction of the symmetric monoidal ∞-category $\mathcal{M}[W^{-1}]$ when $\mathcal{M}$ is topological. Namely, one can first restrict to bifibrant objects and then form the topological colored operad $N^{\otimes}(\mathcal{M})$ with colors $X \in \mathcal{M}^\circ$ and multimorphism spaces

$$\mathrm{Mul}_{N^{\otimes}(\mathcal{M}^\circ)}(\{X_1,\dots,X_n\},Y) = \mathrm{Map}_{\mathcal{M}^\circ}(X_1 \otimes \cdots \otimes X_n, Y).$$

This then gives an ∞-operad by Lemma 2.1. By [Lurie 2017, Proposition 4.1.7.10] this is in fact a symmetric monoidal ∞-category whose underlying ∞-category is equivalent to $\mathcal{M}[W^{-1}]$. Furthermore, by [Lurie 2017, Corollary 4.1.7.16], these two methods of obtaining a symmetric monoidal structure on $\mathcal{M}[W^{-1}]$ are equivalent.

2.3 Pointed categories

Many of the typical constructions one applies to model categories admit an analogue internally to ∞-categories. Furthermore, in many cases these constructions are not only analogous but in fact equivalent.

For example we may consider the formation of pointed objects. Given a model category $\mathcal{M}$ with final object $*$, we can equip the slice category $\mathcal{M}_* = \mathcal{M}_{*/}$ with a model structure in which fibrations, cofibrations and weak equivalences are detected by the forgetful functor $\mathcal{M}_* \to \mathcal{M}$; see [Hovey 1999, Proposition 1.1.8]. If $\mathcal{M}$ is cofibrantly generated with set of generating cofibrations I and set of generating acyclic cofibrations J, then $\mathcal{M}_*$ is also cofibrantly generated by the sets I_+ and J_+; see [Hovey 1999, Lemma 2.1.21]. If $\mathcal{M}$ is symmetric monoidal with cofibrant unit given by $*$, then the slice category $\mathcal{M}_*$ with the smash product is again a symmetric monoidal model category with cofibrant unit; see [Hovey 1999, Proposition 4.2.9].

Let us now discuss the same construction for ∞-categories. Given a presentable symmetric monoidal ∞-category $(\mathscr{C}, \otimes)$, we can endow the slice $\mathscr{C}_* = \mathscr{C}_{*/}$ with a symmetric monoidal structure $\wedge_\otimes$ given as follows: for all $(* \to C), (* \to D) \in \mathscr{C}_*$, we define $C \wedge_\otimes D$ by the following pushout in $\mathscr{C}$:

$$\begin{tikzcd} C \otimes * \sqcup * \otimes D \arrow[r] \arrow[d] & C \otimes D \arrow[d] \\ * \otimes * \arrow[r] \arrow[ur, phantom, "\ulcorner"] & C \wedge_\otimes D \end{tikzcd}$$

The existence of such symmetric monoidal structure on $\mathscr{C}_*$ is a formal consequence of [Lurie 2017, Proposition 4.8.2.11] as we now explain. Indeed the cited reference shows that the functor $(-)_* : \mathrm{Pr}^{\mathrm{L}} \to \mathrm{Pr}^{\mathrm{L}}_*$ from presentable ∞-categories to pointed presentable ∞-categories is a smashing localization, so it induces a functor on commutative algebras $\mathrm{CAlg}(\mathrm{Pr}^{\mathrm{L}}) \to \mathrm{CAlg}(\mathrm{Pr}^{\mathrm{L}}_*)$ showing that a symmetric monoidal structure on $\mathscr{C}_*$ exists. Furthermore, [Lurie 2017, Proposition 4.8.2.11] implies that this symmetric monoidal structure is uniquely determined by the condition that the tensor product on $\mathscr{C}_*$ commutes with colimits on each variable and makes the functor $(-)_+ : \mathscr{C} \to \mathscr{C}_*$ into a symmetric monoidal functor. From this one obtains the concrete description of $\wedge_\otimes$ as given above.

Example 2.2 Applying this construction to $\mathscr{S}$ with the cartesian product returns $\mathscr{S}_*$, the category of pointed spaces with the smash product. We write $\mathscr{S}^\times$ for the ∞-operad giving the former, and $\mathscr{S}^\wedge_*$ for the latter.

We now give a result that connects these two constructions.

Proposition 2.3 *Let $\mathscr{M}$ be a symmetric monoidal model category with cofibrant final object, which is also the monoidal unit. Suppose that the underlying ∞-category $\mathscr{M}[W^{-1}]$ is presentable. Then the functor $(-)_+ : \mathscr{M} \to \mathscr{M}_*$ induces a symmetric monoidal equivalence*

$$(\mathscr{M}[W^{-1}])_* \simeq \mathscr{M}_*[W^{-1}].$$

Proof First note that the underlying ∞-category $\mathscr{M}_*[W^{-1}]$ models the ∞-categorical slice $(\mathscr{M}[W^{-1}])_*$; see for example [Cisinski 2019, Corollary 7.6.13]. Note also that $(-)_+ : \mathscr{M} \to \mathscr{M}_*$ is left Quillen and strong monoidal, and therefore we obtain a strong monoidal colimit-preserving functor

$$(-)_+ : \mathscr{M}[W^{-1}] \to \mathscr{M}_*[W^{-1}],$$

which is equivalent to the standard left adjoint $(-)_+$ under the equivalence $\mathscr{M}_*[W^{-1}] \simeq \mathscr{M}[W^{-1}]_*$ by inspection. Also, $\mathscr{M}_*[W^{-1}]$ is automatically presentable and closed monoidal. Now we can conclude the result, because there is a unique closed symmetric monoidal structure on $\mathscr{M}[W^{-1}]_*$ such that $(-)_+$ is strong monoidal. $\square$

Next we consider the formation of module categories. Recall that given a presentable symmetric monoidal ∞-category $\mathscr{C}$ and a commutative algebra object $S \in \mathrm{CAlg}(\mathscr{C})$, the category of S-modules in $\mathscr{C}$, $\mathrm{Mod}_S(\mathscr{C})$ is a symmetric monoidal ∞-category via the relative tensor product, constructed in Section 4.5.2 of [Lurie 2017]. We will always consider $\mathrm{Mod}_S(\mathscr{C})$ as symmetric monoidal in this way.

Proposition 2.4 *Let $\mathcal{M}$ be a symmetric monoidal and cofibrantly generated model category with weak equivalences W, generating cofibrations I and generating acyclic cofibrations J, and let A be a commutative algebra object in $\mathcal{M}$ whose underlying object is cofibrant. Suppose that $\mathrm{Mod}_A(\mathcal{M})$ admits a symmetric monoidal and cofibrantly generated model structure where fibrations and weak equivalences are tested on underlying objects, and the sets $A \otimes I$ and $A \otimes J$ form a set of generating cofibrations and generating acyclic cofibrations, respectively. Write W_m for the class of weak equivalences in $\mathrm{Mod}_A(\mathcal{M})$. Then applying Mod_A to the functor $\mathcal{M}^c \to \mathcal{M}[W^{-1}]$ induces a symmetric monoidal equivalence*

$$\mathrm{Mod}_A(\mathcal{M})[W_m^{-1}] \simeq \mathrm{Mod}_A(\mathcal{M}[W^{-1}]).$$

Proof This is essentially [Lurie 2017, 4.3.3.17]. However since the statement there does not literally apply, let us spell out the argument. We need to show that there exists a symmetric monoidal equivalence

$$\theta\colon N(\mathrm{Mod}_A(\mathcal{M})^c)[W_m^{-1}] \xrightarrow{\simeq} \mathrm{Mod}_A(N(\mathcal{M}^c)[W^{-1}]).$$

We start by noting that the forgetful functor $U\colon \mathrm{Mod}_A(\mathcal{M}) \to \mathcal{M}$ is left Quillen. One can verify this by observing that U sends the generating (acyclic) cofibrations to (acyclic) cofibrations, using that A is cofibrant and that $\mathcal{M}$ satisfies the pushout-product axiom. Since a cofibrant A-module is then also cofibrant in $\mathcal{M}$, there exists a symmetric monoidal functor

$$N(\mathrm{Mod}_A(\mathcal{M})^c) \to N(\mathrm{Mod}_A(\mathcal{M}^c)) \simeq \mathrm{Mod}_A(N(\mathcal{M}^c)).$$

Postcomposing with the symmetric monoidal functor $N(\mathcal{M}^c) \to N(\mathcal{M}^c)[W^{-1}]$ and using the universal property of symmetric monoidal localization we obtain a symmetric monoidal functor θ as claimed. To show that θ is an equivalence, we apply [Lurie 2017, 4.7.3.16] to the diagram

$$\begin{tikzcd} N(\mathrm{Mod}_A(\mathcal{M})^c)[W_m^{-1}] \arrow[rr, "\theta"] \arrow[dr, "U"'] & & \mathrm{Mod}_A(N(\mathcal{M}^c)[W^{-1}]) \arrow[dl, "U'"] \\ & N(\mathcal{M}^c)[W^{-1}] & \end{tikzcd}$$

We need to check:

(a) The ∞-categories $N(\mathrm{Mod}_A(\mathcal{M})^c)[W_m^{-1}]$ and $\mathrm{Mod}_A(N(\mathcal{M}^c)[W^{-1}])$ admit geometric realization of simplicial objects. In fact, both categories admit all colimits. For $N(\mathrm{Mod}_A(\mathcal{M})^c)[W_m^{-1}]$ this is [Barnea et al. 2017, Theorem 2.5.9]. For $\mathrm{Mod}_A(N(\mathcal{M}^c)[W^{-1}])$, we note that $N(\mathcal{M}^c)[W^{-1}]$ admits all colimits by the previous reference and that these can be calculated as homotopy colimits in the model category by [Barnea et al. 2017, Remark 2.5.7]. Since A is cofibrant, the functor $A \otimes -\colon \mathcal{M} \to \mathcal{M}$ is left Quillen and so it induces a colimit-preserving functor $N(\mathcal{M}^c)[W^{-1}] \to N(\mathcal{M}^c)[W^{-1}]$ by [Hinich 2016, Proposition 1.5.1]. Finally, we can invoke [Lurie 2017, Proposition 4.3.3.9] to deduce the existence of all colimits in $\mathrm{Mod}_A(N(\mathcal{M}^c)[W^{-1}])$.

(b) The functors U and U' admits left adjoints F and F'. The existence of a left adjoint to U follows from the fact that U is determined by a right Quillen functor. The existence of a left adjoint to U' follows from [Lurie 2017, Corollary 4.3.3.14].

(c) The functor U' is conservative and preserves geometric realizations of simplicial objects. This follows from [Lurie 2017, Corollary 4.3.3.2, Proposition 4.3.3.9].

(d) The functor U is conservative and preserves geometric realizations of simplicial objects. The first assertion is immediate from the definition of the weak equivalences in $\mathrm{Mod}_A(\mathcal{M})$, and the second follows from the fact that U is also a left Quillen functor.

(e) The natural map $U' \circ F' \to U \circ F$ is an equivalence. Unwinding the definitions, we are reduced to proving that if N is a cofibrant object of $\mathcal{M}$, then the natural map $N \to A \otimes N$ induces an equivalence $F'(N) \simeq A \otimes N$. This follows from the explicit description of F' given in [Lurie 2017, Corollary 4.3.3.13]. □

Remark 2.5 Suppose $\mathcal{M}$ is a symmetric monoidal cofibrantly generated model category. If $\mathcal{M}$ is locally presentable, then the existence of the model structure on $\mathrm{Mod}_A(\mathcal{M})$ as in Proposition 2.4 holds by [Schwede and Shipley 2000, Remark 4.2].

3 Promonoidal ∞-categories and Day convolution

We start this section by recalling the notion of a promonoidal ∞-category. We recall the definition of the operadic norm functor and use this to define the Day convolution product on a functor category. We then collect various important results about the Day convolution product which will be important later. We finish the section by giving a symmetric monoidal recognition criteria for presheaf categories, inspired by Elmendorf's theorem.

We start off by recalling the following useful notion from [Ayala and Francis 2020, Definition 0.7].

Definition 3.1 A functor $p\colon \mathcal{C} \to \mathcal{B}$ between ∞-categories is an *exponentiable fibration* if the pullback functor $p^*\colon \mathrm{Cat}_{\infty/\mathcal{B}} \to \mathrm{Cat}_{\infty/\mathcal{C}}$ admits a right adjoint p_*, which we call the pushforward.

Example 3.2 Both cocartesian and cartesian fibrations are exponentiable; see [Ayala and Francis 2020, Lemma 2.15].

Example 3.3 Exponentiable fibrations are stable under pullbacks; see [Ayala and Francis 2020, Corollary 1.17]

For any ∞-operad $\mathcal{O}^\otimes$, we let $\mathcal{O}^\otimes_{\mathrm{act}} := \mathcal{O}^\otimes \times_{\mathrm{Fin}_*} \mathrm{Fin} \subseteq \mathcal{O}^\otimes$ denote the subcategory of active arrows. We recall the following definition from [Shah 2021, Definition 10.2].

Definition 3.4 Let $\mathcal{O}^\otimes$ be an ∞-operad. A map of ∞-operads $p\colon \mathcal{C}^\otimes \to \mathcal{O}^\otimes$ defines a $\mathcal{O}^\otimes$*-promonoidal* ∞-category if the restricted functor $p_{\mathrm{act}}\colon \mathcal{C}^\otimes_{\mathrm{act}} \to \mathcal{O}^\otimes_{\mathrm{act}}$ is exponentiable. A functor of $\mathcal{O}^\otimes$-promonoidal ∞-categories is simply a map of $\mathcal{O}^\otimes$-operads.

Example 3.5 Any $\mathcal{O}^\otimes$-symmetric monoidal ∞-category is $\mathcal{O}^\otimes$-promonoidal by Example 3.2.

Example 3.6 Let $\mathscr{C}$ be an ∞-category. Then the ∞-operad $\mathscr{C}^{\amalg} \to \mathrm{Fin}_*$ of [Lurie 2017, Construction 2.4.3.1] is a symmetric promonoidal ∞-category. In fact

$$\mathscr{C}^{\amalg} \times_{\mathrm{Fin}_*} \mathrm{Fin} \to \mathrm{Fin}$$

is the cartesian fibration which classifies the functor sending I to $\mathrm{Fun}(I, \mathscr{C})$.

Example 3.7 Consider a cartesian fibration $p\colon \mathscr{C} \to \mathscr{I}$. Similarly to Example 3.6, one can show that the induced map $p^{\amalg}\colon \mathscr{C}^{\amalg} \to \mathscr{I}^{\amalg}$ exhibits $\mathscr{C}^{\amalg}$ as a $\mathscr{I}^{\amalg}$-promonoidal ∞-category.

The key property of promonoidal ∞-categories is that they induce operadic norm functors.

Definition 3.8 Let $p\colon \mathscr{C}^{\otimes} \to \mathscr{O}^{\otimes}$ be a $\mathscr{O}^{\otimes}$-promonoidal ∞-category. Then the functor

$$p^*\colon (\mathrm{Op}_\infty)_{/\mathscr{O}^{\otimes}} \to (\mathrm{Op}_\infty)_{/\mathscr{C}^{\otimes}}$$

has a right adjoint by [Shah 2021, Theorem/Construction 10.6], which we denote by N_p and call the *norm* along p. Note that p^* also has a left adjoint $p_!$ which is given by postcomposition with p.

The norm interacts well with pullbacks along maps of ∞-operads.

Lemma 3.9 *Let $p\colon \mathscr{C}^{\otimes} \to \mathscr{O}^{\otimes}$ be a $\mathscr{O}^{\otimes}$-promonoidal ∞-category and let $f\colon \mathscr{P}^{\otimes} \to \mathscr{O}^{\otimes}$ be a map of ∞-operads. Write $p'\colon \mathscr{C}^{\otimes} \times_{\mathscr{O}^{\otimes}} \mathscr{P}^{\otimes} \to \mathscr{P}^{\otimes}$ and $f'\colon \mathscr{C}^{\otimes} \times_{\mathscr{O}^{\otimes}} \mathscr{P}^{\otimes} \to \mathscr{O}^{\otimes}$ for the functors obtained via basechange. Then there is a natural equivalence of functors*

$$f^* N_p \simeq N_{p'} (f')^*\colon (\mathrm{Op}_\infty)_{/\mathscr{C}^{\otimes}} \to (\mathrm{Op}_\infty)_{/\mathscr{P}^{\otimes}}.$$

In other words, for every $\mathscr{D}^{\otimes} \in (\mathrm{Op}_\infty)_{/\mathscr{C}^{\otimes}}$ there is an equivalence of ∞-operads over $\mathscr{P}^{\otimes}$,

$$N_p(\mathscr{D}^{\otimes}) \times_{\mathscr{O}^{\otimes}} \mathscr{P}^{\otimes} \simeq N_{p'}(\mathscr{D}^{\otimes} \times_{\mathscr{O}^{\otimes}} \mathscr{P}^{\otimes}).$$

Proof To check that two right adjoint functors are equivalent it is enough to check that the left adjoints are equivalent. But the left adjoint of f^* is just postcomposition with f, so the thesis is equivalent to the fact that for every $\mathscr{E}^{\otimes} \in (\mathrm{Op}_\infty)_{/\mathscr{P}^{\otimes}}$, there is a natural equivalence

$$\mathscr{E}^{\otimes} \times_{\mathscr{O}^{\otimes}} \mathscr{C}^{\otimes} \simeq \mathscr{E}^{\otimes} \times_{\mathscr{P}^{\otimes}} (\mathscr{P}^{\otimes} \times_{\mathscr{O}^{\otimes}} \mathscr{C}^{\otimes})$$

and this is clear. $\square$

Remark 3.10 In a similar vein we observe that because $q^* p^* \simeq (pq)^*$, also $N_{pq} \simeq N_p N_q$.

Remark 3.11 Recall that passing to underlying ∞-categories gives a functor $U\colon \mathrm{Op}_\infty \to \mathrm{Cat}_\infty$ which admits a left adjoint F with essential image precisely those ∞-operads $q\colon \mathscr{P}^{\otimes} \to \mathrm{Fin}_*$ such that the functor q factors through $\mathrm{Triv} \subseteq \mathrm{Fin}_*$; see [Lurie 2017, Proposition 2.1.4.11]. In particular for any ∞-operad $\mathscr{O}^{\otimes}$, we obtain an adjunction on overcategories

$$F\colon (\mathrm{Cat}_\infty)_{/\mathscr{O}} \rightleftarrows (\mathrm{Op}_\infty)_{/\mathscr{O}^{\otimes}} : U.$$

See [Lurie 2009, Proposition 5.2.5.1]. Let $p\colon \mathscr{C}^{\otimes} \to \mathbb{O}^{\otimes}$ be a $\mathbb{O}^{\otimes}$-promonoidal ∞-category; we will now describe the effect of N_p on underlying ∞-categories. Observe that the underlying map $U(p)$ on ∞-categories is exponentiable, as it can be described as the pullback of p along $\mathbb{O} \subseteq \mathbb{O}^{\otimes}$, compare with Example 3.3. One can compute that the diagram of left adjoints

$$\begin{tikzcd} (\mathrm{Op}_\infty)_{/\mathscr{C}^{\otimes}} & (\mathrm{Op}_\infty)_{/\mathbb{O}^{\otimes}} \arrow[l, "p^*"'] \\ (\mathrm{Cat}_\infty)_{/\mathscr{C}} \arrow[u, "F"] & (\mathrm{Cat}_\infty)_{/\mathbb{O}} \arrow[l, "U(p)^*"'] \arrow[u, "F"'] \end{tikzcd}$$

commutes. Therefore the associated diagram of right adjoints

$$\begin{tikzcd} (\mathrm{Op}_\infty)_{/\mathscr{C}^{\otimes}} \arrow[r, "N_p"] \arrow[d, "U"'] & (\mathrm{Op}_\infty)_{/\mathbb{O}^{\otimes}} \arrow[d, "U"] \\ (\mathrm{Cat}_\infty)_{/\mathscr{C}} \arrow[r, "U(p)_*"] & (\mathrm{Cat}_\infty)_{/\mathbb{O}} \end{tikzcd}$$

also commutes, and we conclude that on underlying categories N_p is given by the pushforward $U(p)_*$.

We can now define the Day convolution functor.

Definition 3.12 Let $p\colon \mathscr{C}^{\otimes} \to \mathbb{O}^{\otimes}$ be a $\mathbb{O}^{\otimes}$-promonoidal ∞-category. The *Day convolution functor*

$$\mathrm{Fun}_{\mathbb{O}}(\mathscr{C}, -)^{\mathrm{Day}}\colon (\mathrm{Op}_\infty)_{/\mathbb{O}^{\otimes}} \to (\mathrm{Op}_\infty)_{/\mathbb{O}^{\otimes}}$$

is the right adjoint of the functor

$$p_! p^* = - \times_{\mathbb{O}^{\otimes}} \mathscr{C}^{\otimes}\colon (\mathrm{Op}_\infty)_{/\mathbb{O}^{\otimes}} \to (\mathrm{Op}_\infty)_{/\mathbb{O}^{\otimes}}.$$

This is a composite of right adjoints, and so we conclude that $\mathrm{Fun}_{\mathbb{O}}(\mathscr{C}, -)^{\mathrm{Day}} \simeq N_p p^*(-)$. This also shows the existence of $\mathrm{Fun}_{\mathbb{O}}(\mathscr{C}, -)^{\mathrm{Day}}$. When $\mathbb{O} = \mathrm{Fin}_*$, we will omit it from the notation.

Remark 3.13 Recall that $\mathrm{Alg}_{\mathscr{C}^{\otimes}}(\mathscr{D}^{\otimes})$ is defined to be the full subcategory of $\mathrm{Fun}_{/\mathrm{Fin}_*}(\mathscr{C}^{\otimes}, \mathscr{D}^{\otimes})$ spanned by the maps of operads, and that taking the maximal ∞-subgroupoid of this category gives the mapping spaces $\mathrm{Op}_\infty(\mathscr{C}^{\otimes}, \mathscr{D}^{\otimes})$. Therefore we may view $\mathrm{Alg}_{(-)}(-)$ as constituting an enrichment of Op_∞ in Cat_∞. A standard argument shows that the adjunction equivalence

$$\mathrm{Op}_\infty(\mathscr{P}^{\otimes}, \mathrm{Fun}(\mathscr{J}^{\otimes}, \mathscr{C}^{\otimes})^{\mathrm{Day}}) \simeq \mathrm{Op}_\infty(\mathscr{P}^{\otimes} \times_{\mathrm{Fin}_*} \mathscr{J}^{\otimes}, \mathscr{C}^{\otimes})$$

improves to an equivalence

$$\mathrm{Alg}_{\mathscr{P}^{\otimes}}(\mathrm{Fun}(\mathscr{J}^{\otimes}, \mathscr{C}^{\otimes})^{\mathrm{Day}}) \simeq \mathrm{Alg}_{\mathscr{P}^{\otimes} \times_{\mathrm{Fin}_*} \mathscr{J}^{\otimes}}(\mathscr{C}^{\otimes}).$$

Example 3.14 Recall from Example 3.6 that for any ∞-category $\mathscr{C}$, the ∞-operad $\mathscr{C}^{\amalg} \to \mathrm{Fin}_*$ is promonoidal. For every ∞-operad $\mathscr{D}^{\otimes}$, the Day convolution ∞-operad $\mathrm{Fun}(\mathscr{C}^{\amalg}, \mathscr{D}^{\otimes})^{\mathrm{Day}}$ is equivalent to the pointwise operad structure on $\mathrm{Fun}(\mathscr{C}, \mathscr{D})$. Indeed they corepresent the same functor by [Lurie 2017, Theorem 2.4.3.18].

The description of Day convolution combined with Remark 3.11 implies that on underlying categories $\mathrm{Fun}_{\mathbb{O}}(\mathscr{C}^{\otimes}, -)^{\mathrm{Day}}$ is given by $U(p)_*U(p)^*$. We can describe the fibers of this category explicitly.

Construction 3.15 Let $p\colon \mathscr{C} \to \mathscr{B}$ be an exponentiable fibration of ∞-categories and $q\colon \mathscr{D} \to \mathscr{B}$ any functor. Fix an arrow $f : b_0 \to b_1$ in $\mathscr{B}$ and let us write $\mathscr{C}_{b_i}$ and $\mathscr{D}_{b_i}$ for the fibers of p and q over b_i. The unit of the adjunction (p^*, p_*) gives a canonical functor $p_* p^*\mathscr{D} \to \mathscr{B}$ whose fiber over b_i can be identified with

$$(3.15.1)\qquad (p_* p^*\mathscr{D})_{b_i} \simeq \mathrm{Fun}_{\mathscr{B}}(\{b_i\}, p_* p^*\mathscr{D}) \simeq \mathrm{Fun}_{\mathscr{C}}(\mathscr{C} \times_{\mathscr{B}} \{b_i\}, \mathscr{C} \times_{\mathscr{B}} \mathscr{D}) \simeq \mathrm{Fun}(\mathscr{C}_i, \mathscr{D}_i).$$

Remark 3.16 One should be careful to note that, if the underlying ∞-category $\mathbb{O}$ of $\mathbb{O}^{\otimes}$ is not contractible, then the underlying ∞-category of $\mathrm{Fun}_{\mathbb{O}}(\mathscr{C}^{\otimes}, \mathscr{D}^{\otimes})^{\mathrm{Day}}$ is not the same as the ∞-category of functors over $\mathbb{O}$. Rather, it is a fibration over $\mathbb{O}$ whose global sections are $\mathrm{Fun}_{/\mathbb{O}}(\mathscr{C}, \mathscr{D})$. Compare also with the previous construction.

We would like to have a formula for the multimapping spaces for the Day convolution. We will achieve this in Lemma 3.25 below. In preparation for this result, we compute the mapping spaces in a pushforward. To state the result we recall the definition of twisted arrow ∞-categories, and the notion of coends.

Definition 3.17 Let $\epsilon\colon \Delta \to \Delta$ be the functor $[n] \mapsto [n] \star [n]^{\mathrm{op}} \simeq [2n+1]$. Let $\mathscr{I}$ be an ∞-category. The twisted arrow ∞-category $\mathrm{Tw}(\mathscr{I})$ is the associated ∞-category of the simplicial set $\epsilon^* N\mathscr{I}$. By definition, we have

$$\mathrm{Tw}(\mathscr{I})_n = \mathrm{Map}(\Delta^n \star (\Delta^n)^{\mathrm{op}}, \mathscr{I}).$$

The natural transformations $\Delta^{\bullet}$ and $(\Delta^{\bullet})^{\mathrm{op}} \to \Delta^{\bullet} \star (\Delta^{\bullet})^{\mathrm{op}}$ induce a functor $(s,t)\colon \mathrm{Tw}(\mathscr{I}) \to \mathscr{I} \times \mathscr{I}^{\mathrm{op}}$.

Remark 3.18 There are two possible conventions for defining $\mathrm{Tw}(-)$. In this paper we follow that of Lurie [2017, Section 5.2.1]. This is the opposite of the convention used in [Barwick 2017].

Example 3.19 The objects of $\mathrm{Tw}(\mathscr{I})$ are given by edges of $\mathscr{I}$. An edge from $f : x \to y$ to $f' : x' \to y'$ in $\mathrm{Tw}(\mathscr{I})$ is represented by a diagram

$$\begin{tikzcd} x \arrow[r] \arrow[d, "f"'] & x' \arrow[d, "f'"] \\ y & y' \arrow[l] \end{tikzcd}$$

Remark 3.20 The twisted arrow category is insensitive to taking opposites, meaning that $\mathrm{Tw}(\mathscr{I}^{\mathrm{op}}) \simeq \mathrm{Tw}(\mathscr{I})$. However under this equivalence (s,t) is sent to (t,s).

Definition 3.21 Given a functor $F\colon \mathscr{C} \times \mathscr{C}^{\mathrm{op}} \to \mathscr{S}$, we define the coend $\int^{x \in \mathscr{C}} F(x,x)$ to equal the colimit of the functor

$$\mathrm{Tw}(\mathscr{C}) \xrightarrow{(s,t)} \mathscr{C} \times \mathscr{C}^{\mathrm{op}} \xrightarrow{F} \mathscr{S}.$$

Dually for a functor $F\colon \mathscr{C}^{\mathrm{op}} \times \mathscr{C} \to \mathscr{C}$, we define the end $\int_{x \in \mathscr{C}} F(x,x)$ to be the limit of the functor

$$\mathrm{Tw}(\mathscr{C})^{\mathrm{op}} \xrightarrow{(s,t)^{\mathrm{op}}} \mathscr{C}^{\mathrm{op}} \times \mathscr{C} \xrightarrow{F} \mathscr{S}.$$

We are now ready to state the formula for multimapping spaces in the Day convolution.

Lemma 3.22 *Suppose we are in the setting of Construction 3.15. Let $F_i\colon \mathscr{C}_i \to \mathscr{D}_i$ be two objects of $(p_*p^*\mathscr{D})_{b_i}$, viewed as such via the equivalence (3.15.1). Then there is an equivalence*

$$\mathrm{Map}^f_{p_*p^*\mathscr{D}}(F_0,F_1) \simeq \int_{(x_0,x_1)\in\mathscr{C}_0^{\mathrm{op}}\times\mathscr{C}_1} \mathrm{Map}(\mathrm{Map}^f_{\mathscr{C}}(x_0,x_1),\mathrm{Map}^f_{\mathscr{D}}(F_0x_0,F_1x_1)), \tag{3.22.1}$$

where the left-hand side denotes the fiber over f of the canonical map

$$\mathrm{Map}_{p_*p^*\mathscr{D}}(F_0,F_1) \to \mathrm{Map}_{\mathscr{B}}(b_0,b_1).$$

Proof Let us write f as a map $\Delta^1 \to \mathscr{B}$. Then, by the definition of p_*, there is an equivalence

$$\mathrm{Map}_{\mathscr{B}}(\Delta^1, p_*p^*\mathscr{D}) \simeq \mathrm{Map}_{\mathscr{C}}(\Delta^1\times_{\mathscr{B}}\mathscr{C}, \mathscr{C}\times_{\mathscr{B}}\mathscr{D}) \simeq \mathrm{Map}_{\Delta^1}(\Delta^1\times_{\mathscr{B}}\mathscr{C}, \Delta^1\times_{\mathscr{B}}\mathscr{D}).$$

Therefore we have an equivalence

$$\begin{aligned}\mathrm{Map}^f_{p_*p^*\mathscr{D}}(F_0,F_1) &\simeq \{(F_0,F_1)\}\times_{\mathrm{Map}_{\mathscr{B}}(\partial\Delta^1,p_*p^*\mathscr{D})}\mathrm{Map}_{\mathscr{B}}(\Delta^1,p_*p^*\mathscr{D})\\ &\simeq \{(F_0,F_1)\}\times_{\mathrm{Map}(\mathscr{C}_0,\mathscr{D}_0)\times\mathrm{Map}(\mathscr{C}_1,\mathscr{D}_1)}\mathrm{Map}_{\Delta^1}(\Delta^1\times_{\mathscr{B}}\mathscr{C},\Delta^1\times_{\mathscr{B}}\mathscr{D}).\end{aligned}$$

Now from the proof of [Ayala and Francis 2020, Lemma 4.2] it follows that the map

$$\mathrm{Cat}_{\infty/\Delta^1} \to \mathrm{Cat}_\infty\times\mathrm{Cat}_\infty \qquad [\mathscr{C}\to\Delta^1]\mapsto(\mathscr{C}\times_{\Delta^1}\{0\},\mathscr{C}\times_{\Delta^1}\{1\}),$$

is a right fibration classified by the functor $\mathrm{Cat}_\infty\times\mathrm{Cat}_\infty\to\mathscr{S}$ sending $(\mathscr{C}_0,\mathscr{C}_1)$ to $\mathrm{Map}(\mathscr{C}_0^{op}\times\mathscr{C}_1,\mathscr{S})$. Therefore

$$\begin{aligned}\{(F_0,F_1)\}&\times_{\mathrm{Map}(\mathscr{C}_0,\mathscr{D}_0)\times\mathrm{Map}(\mathscr{C}_1,\mathscr{D}_1)}\mathrm{Map}_{\Delta^1}(\Delta^1\times_{\mathscr{B}}\mathscr{C},\Delta^1\times_{\mathscr{B}}\mathscr{D})\\ &\simeq \mathrm{Map}^{(F_0,F_1)}_{\mathrm{Cat}_{\infty/\Delta^1}}(\Delta^1\times_{\mathscr{B}}\mathscr{C},\Delta^1\times_{\mathscr{B}}\mathscr{D})\\ &\simeq \mathrm{Map}_{(\mathrm{Cat}_{\infty/\Delta^1})_{(\mathscr{C}_0,\mathscr{C}_1)}}(\Delta^1\times_{\mathscr{B}}\mathscr{C},(F_0,F_1)^*(\Delta^1\times_{\mathscr{B}}\mathscr{D}))\\ &\simeq \mathrm{Map}_{\mathrm{Fun}(\mathscr{C}_0^{\mathrm{op}}\times\mathscr{C}_1,\mathscr{S})}(\mathrm{Map}^f_{\mathscr{C}}(-,-),\mathrm{Map}^f_{\mathscr{D}}(F_0-,F_1-)).\end{aligned}$$

But this is exactly the thesis, thanks to [Gepner et al. 2017, Proposition 5.1]. □

Remark 3.23 In the setting of Lemma 3.22, suppose that q is equal to the projection $\mathscr{D}\times\mathscr{B}\to\mathscr{B}$ and that $\mathscr{D}$ is cocomplete. Then we can interpret formula (3.22.1) as saying that $p_*p^*\mathscr{D}$ is a cocartesian fibration and that given $f\colon i\to j$, the induced functor

$$f_!\colon \mathrm{Fun}(\mathscr{C}_i,\mathscr{D})\to\mathrm{Fun}(\mathscr{C}_j,\mathscr{D})$$

evaluated on a functor $F\colon\mathscr{C}_i\to\mathscr{D}$ gives the functor

$$\mathscr{C}_j\to\mathscr{D},\quad x_j\mapsto\int^{x_i\in\mathscr{C}_i}\mathrm{Map}_{\mathscr{C}_{ij}}(x_i,x_j)\times F(x_i),$$

where $\mathscr{C}_{ij} := \mathscr{C}\times_{\mathscr{B},f}[1]$. That is, $f_!F$ is computed by left Kan extending F along the inclusion $\mathscr{C}_i\subseteq\mathscr{C}_{ij}$ and then restricting to $\mathscr{C}_j\subseteq\mathscr{C}_{ij}$. In particular, if $\mathscr{C}_{ij}\to[1]$ is a cartesian fibration we have $f_!F\simeq F\circ f^*$, where $f^*\colon\mathscr{C}_1\to\mathscr{C}_0$ is the pullback.

Recall the following notion of multimapping spaces.

Definition 3.24 Let $\mathscr{C}^\otimes \to \mathbb{O}^\otimes$ be a map of ∞-operads and let $\phi\colon \{x_i\} \to y$ be an active morphism of $\mathbb{O}^\otimes$ with target in $\mathbb{O} := (\mathbb{O}^\otimes)_{1_+}$. For every $\{c_i\} \in (\mathscr{C}^\otimes)_{\{x_i\}} \simeq \prod_i \mathscr{C}_{x_i}$ and $d \in \mathscr{C}_y$, objects of $\mathscr{C}^\otimes$ over the source and target of ϕ, we define the ϕ-multimapping space in $\mathscr{C}^\otimes$ as the space of morphisms $\{c_i\} \to d$ above ϕ:

$$\mathrm{Mul}^{\phi}_{\mathscr{C}^\otimes}(\{c_i\}, d) := \mathrm{Map}_{\mathscr{C}^\otimes}(\{c_i\}, d) \times_{\mathrm{Map}_{\mathbb{O}^\otimes}(\{x_i\}, y)} \{\phi\}.$$

We say that $\mathscr{C}^\otimes$ is *representable* if for every active morphism ϕ and objects $\{c_i\}$, the functor

$$\mathrm{Mul}^{\phi}_{\mathscr{C}^\otimes}(\{c_i\}, -)\colon \mathscr{C} \to \mathscr{S}$$

is corepresentable. In this case we write $\bigotimes_\phi \{c_i\}$ for the corepresenting object and we call it the ϕ-tensor product of $\{c_i\}$. This is equivalent to the functor $\mathscr{C}^\otimes \to \mathbb{O}^\otimes$ being a locally cocartesian fibration.

We are ready to prove the formula for the multimapping spaces in the Day convolution.

Lemma 3.25 *Let $\mathbb{O}^\otimes$ be an ∞-operad, $\mathscr{C}^\otimes$ be an $\mathbb{O}^\otimes$-promonoidal ∞-category and $\mathscr{D}^\otimes$ be an ∞-operad over $\mathbb{O}^\otimes$. Then the multimapping spaces in $\mathrm{Fun}_{\mathbb{O}}(\mathscr{C}^\otimes, \mathscr{D}^\otimes)^{\mathrm{Day}}$ are given by the natural equivalence*

$$\mathrm{Mul}^{\phi}_{\mathrm{Fun}_{\mathbb{O}}(\mathscr{C}^\otimes, \mathscr{D}^\otimes)^{\mathrm{Day}}}(\{F_i\}, G) \simeq \int_{c' \in \mathscr{C}_y} \int_{\{c_i\} \in (\prod_i \mathscr{C}_{x_i})^{\mathrm{op}}} \mathrm{Map}\big(\mathrm{Mul}^{\phi}_{\mathscr{C}}(\{c_i\}, c'), \mathrm{Mul}^{\phi}_{\mathscr{D}}(\{F_i c_i\}, Gc')\big)$$

for all active morphisms $\phi\colon \{x_i\} \to y$, and objects $\{F_i\} \in \prod_i \mathrm{Fun}(\mathscr{C}_{x_i}, \mathscr{D}_{x_i})$, $G \in \mathrm{Fun}(\mathscr{C}_y, \mathscr{D}_y)$.

Proof We will use [Lurie 2017, Proposition 2.2.6.6]. However the cited result has the hypothesis that $\mathscr{C}^\otimes$ is a $\mathbb{O}^\otimes$-monoidal ∞-category. We note that this is only used to ensure the existence of the norm (after replacing the appeal to [Lurie 2009, Proposition 3.3.1.3] with [Lurie 2017, Proposition B.3.14]). Therefore, in view of [Shah 2021, Theorem/Construction 10.6], we can safely apply this result when $\mathscr{C}^\otimes$ is only $\mathbb{O}^\otimes$-promonoidal.

Then, arguing as in the proof of [Lurie 2017, Proposition 2.2.6.11], we obtain an equivalence

$$\mathrm{Mul}^{\phi}_{\mathrm{Fun}_{\mathbb{O}}(\mathscr{C}^\otimes, \mathscr{D}^\otimes)^{\mathrm{Day}}}(\{F_i\}, G) \simeq \{(F, G)\} \times_{\mathrm{Fun}_{/\mathbb{O}^\otimes}(\partial\Delta^1 \times_{\mathbb{O}^\otimes} \mathscr{C}^\otimes, \mathscr{D}^\otimes)} \mathrm{Fun}_{/\mathbb{O}^\otimes}(\Delta^1 \times_{\mathbb{O}^\otimes} \mathscr{C}^\otimes, \mathscr{D}^\otimes),$$

where $\Delta^1 \to \mathbb{O}^\otimes$ picks out the active arrow ϕ and $F\colon \mathscr{C}^\otimes_{\{x_i\}} \to \mathscr{D}^\otimes_{\{x_i\}}$ is the functor sending $\{c_i\}$ to $\{F_i c_i\}$. Let $\mathscr{C}^{\mathrm{act}} := \mathscr{C}^\otimes \times_{\mathbb{O}^\otimes} \mathbb{O}^{\mathrm{act}}$ and $\mathscr{D}^{\mathrm{act}} := \mathscr{D}^\otimes \times_{\mathbb{O}^\otimes} \mathbb{O}^{\mathrm{act}}$ be the subcategories of active arrows. Since $\Delta^1 \to \mathbb{O}^\otimes$ factors through $\mathbb{O}^{\mathrm{act}}$, we have an equivalence

$$\begin{aligned}\mathrm{Mul}_{\mathrm{Fun}(\mathscr{C}^\otimes, \mathscr{D}^\otimes)^{\mathrm{Day}}}(\{F_i\}_{i \in I}, G) &\simeq \{(F, G)\} \times_{\mathrm{Fun}_{/\mathbb{O}^{\mathrm{act}}}(\partial\Delta^1 \times_{\mathbb{O}^{\mathrm{act}}} \mathscr{C}^{\mathrm{act}}, \mathscr{D}^{\mathrm{act}})} \mathrm{Fun}_{/\mathbb{O}^{\mathrm{act}}}(\Delta^1 \times_{\mathrm{Fin}} \mathscr{C}^{\mathrm{act}}, \mathscr{D}^{\mathrm{act}})\\ &\simeq \mathrm{Map}_{(p^{\mathrm{act}})_*(p^{\mathrm{act}})^*\mathscr{D}^{\mathrm{act}}}(F, G),\end{aligned}$$

where the last equality makes sense since p^{act} is an exponentiable fibration. Therefore the thesis follows from Lemma 3.22. ◻

Definition 3.26 We say that an $\mathbb{O}^{\otimes}$-monoidal ∞-category $\mathcal{D}^{\otimes} \to \mathbb{O}^{\otimes}$ is *compatible with colimits* if for every object $x \in \mathbb{O}$, the fiber $\mathcal{D}_x$ has all small colimits, and if for every active arrow ϕ, the ϕ-tensor product commutes with all small colimits separately in each variable; see [Lurie 2017, Definition 3.1.1.18] for a more precise formulation. If moreover each fiber is presentable, then we say $\mathcal{D}^{\otimes}$ is a presentably $\mathbb{O}^{\otimes}$-monoidal ∞-category.

Example 3.27 The underlying ∞-category of a symmetric monoidal model category is compatible with colimits as the tensor product is a left Quillen bifunctor by the pushout-product axiom.

Remark 3.28 Recall that every cocomplete ∞-category $\mathscr{C}$ is canonically tensored over $\mathcal{S}$. Namely, for every $X \in \mathcal{S}$ and $C \in \mathscr{C}$, we define $X \times C$ to equal $\operatorname{colim}(\mathrm{const}_C \colon X \to \mathscr{C})$, the colimit over X of the constant functor at C.

Corollary 3.29 *Fix an ∞-operad $\mathbb{O}^{\otimes}$. Let $\mathscr{C}^{\otimes}$ be a small $\mathbb{O}^{\otimes}$-promonoidal ∞-category and let $\mathcal{D}^{\otimes}$ be a $\mathbb{O}^{\otimes}$-monoidal ∞-category which is compatible with colimits. Then*:

(a) $\mathrm{Fun}_{\mathbb{O}}(\mathscr{C}^{\otimes}, \mathcal{D}^{\otimes})^{\mathrm{Day}}$ *is an $\mathbb{O}^{\otimes}$-monoidal ∞-category which is again compatible with colimits.*

Suppose furthermore that $\mathbb{O}^{\otimes} \simeq \mathrm{Fin}_*$ is the commutative ∞-operad.

(b) *The unit of $\mathrm{Fun}(\mathscr{C}^{\otimes}, \mathcal{D}^{\otimes})^{\mathrm{Day}}$ is given by $1_{\mathrm{Day}} := \mathrm{Mul}_{\mathcal{D}}(\varnothing, -) \times 1_{\mathcal{D}}$, and the tensor product is given by*
$$(F \otimes^{\mathrm{Day}} G)(-) \simeq \int^{(c_1, c_2) \in \mathscr{C}^2} \mathrm{Mul}_{\mathscr{C}}(\{c_1, c_2\}, -) \times (F(c_1) \otimes G(c_2)).$$
In particular, when $\mathcal{D}$ is the ∞-category of spaces with the cartesian symmetric monoidal structure, we have
$$\mathrm{Map}_{\mathscr{C}}(x, -) \otimes^{\mathrm{Day}} \mathrm{Map}_{\mathscr{C}}(y, -) \simeq \mathrm{Mul}_{\mathscr{C}}(\{x, y\}, -)$$
for every $x, y \in \mathscr{C}$.

Proof If $\mathcal{D}^{\otimes}$ is $\mathbb{O}^{\otimes}$-monoidal, it follows from the formula of Lemma 3.25 that $\mathrm{Fun}(\mathscr{C}^{\otimes}, \mathcal{D}^{\otimes})^{\mathrm{Day}}$ is representable and that the ϕ-tensor product is given by
$$\bigotimes_{\phi} \{F_i\}_{i \in I} \simeq \int^{\{c_i\} \in \prod_{i \in I} \mathscr{C}_{o_i}} \mathrm{Mul}^{\phi}_{\mathscr{C}^{\otimes}}(\{c_i\}_{i \in I}, -) \times \bigotimes_{\phi} \{F_i(c_i)\}_{i \in I}.$$
This shows the existence of locally cartesian edges in $\mathrm{Fun}(\mathscr{C}^{\otimes}, \mathcal{D}^{\otimes})^{\mathrm{Day}}$. Because the tensor product functors in $\mathcal{D}^{\otimes}$ commutes with colimits in each variable, one can calculate that the composite of locally cartesian edges is locally cartesian, and therefore $\mathrm{Fun}_{\mathbb{O}}(\mathscr{C}^{\otimes}, \mathcal{D}^{\otimes})^{\mathrm{Day}}$ is a $\mathbb{O}^{\otimes}$-monoidal ∞-category. The fibers are clearly cocomplete, and from the formula for the tensor product it follows that the tensor in $\mathrm{Fun}(\mathscr{C}, \mathcal{D})^{\otimes}$ commutes with colimits in each variable.

Finally the statement for the tensor product of corepresentable functors follows from the formula above and the Yoneda lemma. □

Notation 3.30 Suppose we are in the situation of the previous corollary, and suppose that $\mathcal{O}^{\otimes} \simeq \mathrm{Fin}_*$. In the case that both $\mathcal{C}^{\otimes}$ and $\mathcal{D}^{\otimes}$ are canonically (pro)monoidal, then we write $\mathcal{C}\!-\!\mathcal{D}$ for the symmetric monoidal category given by the ∞-operad $\mathrm{Fun}(\mathcal{C}^{\otimes}, \mathcal{D}^{\otimes})^{\mathrm{Day}}$. The two examples which will arise constantly are $\mathcal{C}$-$\mathcal{S}$ and $\mathcal{C}$-$\mathcal{S}_*$, where $\mathcal{S}$ is symmetric monoidal via the cartesian product, and $\mathcal{S}_*$ via the smash product. Nevertheless, when we refer to the ∞-operad inducing the symmetric monoidal structure on $\mathcal{C}\!-\!\mathcal{D}$, we will continue to write $\mathrm{Fun}(\mathcal{C}^{\otimes}, \mathcal{D}^{\otimes})^{\mathrm{Day}}$. While this distinction is mathematically meaningless, we find it notationally convenient.

We next turn to the functoriality of Day convolution.

Construction 3.31 Let $\mathcal{O}^{\otimes}$ be an ∞-operad and suppose $f : \mathcal{I}^{\otimes} \to \mathcal{J}^{\otimes}$ is a map of $\mathcal{O}^{\otimes}$-promonoidal ∞-categories. Then for every two ∞-operads $\mathcal{C}^{\otimes}$ and $\mathcal{P}^{\otimes}$ over $\mathcal{O}^{\otimes}$ we have a natural transformation

$$\begin{aligned}\mathrm{Alg}_{\mathcal{P}^{\otimes}/\mathcal{O}^{\otimes}}(\mathrm{Fun}_{\mathcal{O}^{\otimes}}(\mathcal{J}^{\otimes}, \mathcal{C}^{\otimes})^{\mathrm{Day}}) \simeq \mathrm{Alg}_{\mathcal{P}^{\otimes} \times_{\mathcal{O}^{\otimes}} \mathcal{J}^{\otimes}}(\mathcal{C}^{\otimes}) &\to \mathrm{Alg}_{\mathcal{P}^{\otimes} \times_{\mathcal{O}^{\otimes}} \mathcal{I}^{\otimes}}(\mathcal{C}^{\otimes}) \\ &\simeq \mathrm{Alg}_{\mathcal{P}^{\otimes}/\mathcal{O}^{\otimes}}(\mathrm{Fun}_{\mathcal{O}^{\otimes}}(\mathcal{I}^{\otimes}, \mathcal{C}^{\otimes})^{\mathrm{Day}})\end{aligned}$$

given by precomposition along $\mathcal{P}^{\otimes} \times_{\mathcal{O}^{\otimes}} \mathcal{I}^{\otimes} \to \mathcal{P}^{\otimes} \times_{\mathcal{O}^{\otimes}} \mathcal{J}^{\otimes}$. Since this is natural in $\mathcal{P}^{\otimes}$, it induces a map in $(\mathrm{Op}_{\infty})_{/\mathcal{O}^{\otimes}}$

$$f^* : \mathrm{Fun}_{\mathcal{O}^{\otimes}}(\mathcal{J}^{\otimes}, \mathcal{C}^{\otimes})^{\mathrm{Day}} \to \mathrm{Fun}_{\mathcal{O}^{\otimes}}(\mathcal{I}^{\otimes}, \mathcal{C}^{\otimes})^{\mathrm{Day}}.$$

Definition 3.32 Consider $\mathcal{C}^{\otimes}, \mathcal{D}^{\otimes} \in (\mathrm{Op}_{\infty})_{/\mathcal{O}^{\otimes}}$. An *operadic adjunction* between $\mathcal{C}^{\otimes}$ and $\mathcal{D}^{\otimes}$ is a relative adjunction over $\mathcal{O}^{\otimes}$ in the sense of [Lurie 2017, Definition 7.3.2.2] such that both functors are maps of ∞-operads. This notion is equivalent to an adjunction in the $(\infty, 2)$-category of ∞-operads; see [Riehl and Verity 2016, Observation 4.3.2].

Remark 3.33 If $\mathcal{C}^{\otimes}$ and $\mathcal{D}^{\otimes}$ are both $\mathcal{O}^{\otimes}$-monoidal then an operadic left adjoint $f : \mathcal{C}^{\otimes} \to \mathcal{D}^{\otimes}$ is automatically $\mathcal{O}^{\otimes}$-monoidal by [Lurie 2017, Proposition 7.3.2.6].

Proposition 3.34 *Let $\mathcal{O}^{\otimes}$ be an ∞-operad and let $f : \mathcal{I}^{\otimes} \to \mathcal{J}^{\otimes}$ a map of $\mathcal{O}^{\otimes}$-promonoidal ∞-categories. Suppose $\mathcal{C}^{\otimes}$ is a presentably $\mathcal{O}^{\otimes}$-monoidal ∞-category. Let us consider the lax $\mathcal{O}^{\otimes}$-monoidal functor*

$$f^* : \mathrm{Fun}_{\mathcal{O}^{\otimes}}(\mathcal{J}^{\otimes}, \mathcal{C}^{\otimes})^{\mathrm{Day}} \to \mathrm{Fun}_{\mathcal{O}^{\otimes}}(\mathcal{I}^{\otimes}, \mathcal{C}^{\otimes})^{\mathrm{Day}}.$$

(a) *Suppose that for every active arrow $\phi : \{t_i\}_i \to t$ in $\mathcal{O}^{\otimes}$ the natural map*

$$(f_t)_! \,\mathrm{Mul}^{\phi}_{\mathcal{I}}(\{x_i\}_i, -) \to \mathrm{Mul}^{\phi}_{\mathcal{J}}(\{f_{t_i} x_i\}_i, -)$$

adjoint to

$$\mathrm{Mul}^{\phi}_{\mathcal{I}}(\{x_i\}_i, -) \to \mathrm{Mul}^{\phi}_{\mathcal{J}}(\{f_{t_i} x_i\}_i, f_t(-))$$

is an equivalence for every family of objects $\{x_i\}_i$. Then f^* has a left operadic adjoint $f_!$ that is $\mathcal{O}^{\otimes}$-monoidal.

(b) *Suppose f has an operadic right adjoint $g : \mathcal{J}^{\otimes} \to \mathcal{I}^{\otimes}$. Then there is a natural equivalence of maps of ∞-operads $f_! \simeq g^*$, and moreover this functor is $\mathcal{O}^{\otimes}$-monoidal.*

Proof We will use [Lurie 2017, Proposition 7.3.2.11] applied to the functor f^* over $\mathcal{O}^{\otimes}$. Since on the fiber over $t_i \in \mathcal{O}$ this is just given by precomposition by f_{t_i}, the functor on the fiber over $\{t_i\}_i$

$$\prod_i \mathrm{Fun}(\mathcal{J}_{t_i}, \mathcal{C}_{t_i}) \to \prod_i \mathrm{Fun}(\mathcal{I}_{t_i}, \mathcal{C}_{t_i})$$

has a left adjoint, given by the left Kan extension $(f_{t_i})_!$ on every component. In particular, this collection of left adjoints commutes with the pushforwards along inert maps. So it suffices to show that this collection of left adjoints commute with the pushforwards along active maps. Let $\phi\colon (t_i)_i \to t$ be an active map. Then we need to show that the map

$$(f_t)_!\left(\bigotimes_i^{\phi} F_i\right) \to \bigotimes_i^{\phi} (f_{t_i})_! F_i$$

is an equivalence. But then this follows from our hypothesis together with the description of Corollary 3.29.

Suppose now that f has an operadic right adjoint g. Since g^* is an operadic left adjoint to f^*, it follows immediately that $f_! = g^*$. So it remains only to check the two final conditions. But we have

$$(f_t)_! \,\mathrm{Mul}^{\phi}_{\mathcal{J}}(\{x_i\}_i, -) \simeq \mathrm{Mul}^{\phi}_{\mathcal{J}}(\{x_i\}_i, g_t -) \simeq \mathrm{Mul}^{\phi}_{\mathcal{J}}(\{f_{t_i} x_i\}_i, -),$$

since g is an operadic right adjoint of f. $\square$

Remark 3.35 If $\mathcal{O}^{\otimes} = \mathrm{Fin}_*$ and $\mathcal{I}^{\otimes}$ and $\mathcal{J}^{\otimes}$ are both symmetric monoidal, then the conditions ensuring the symmetric monoidality of $f_!$ are equivalent to f being a symmetric monoidal functor (since $f_!$ restricts to f on representables). Thus the above proposition gives an alternative proof of [Ben-Moshe and Schlank 2024, Proposition 3.6].

3.1 Symmetric monoidal structures on copresheaf categories

We finish this section by classifying all possible closed symmetric monoidal structures on the copresheaf ∞-category $\mathrm{Fun}(\mathcal{I}, \mathcal{S})$ in terms of symmetric promonoidal structures on $\mathcal{I}$; see Theorem 3.37.

Lemma 3.36 *Let $\mathcal{I}$ be a small ∞-category and let us suppose that the presheaf category* $\mathrm{Fun}(\mathcal{I}, \mathcal{S})$ *is equipped with a symmetric monoidal structure* $\mathrm{Fun}(\mathcal{I}, \mathcal{S})^{\otimes}$ *which is compatible with colimits. Equip $\mathcal{I}$ with the full suboperad structure $\mathcal{I}^{\otimes}$ induced by the Yoneda embedding $\mathcal{I} \subseteq \mathrm{Fun}(\mathcal{I}, \mathcal{S})^{\mathrm{op}}$. Then $\mathcal{I}^{\otimes}$ is symmetric promonoidal.*

Proof For brevity let us write $\mathcal{D}^{\otimes} = \mathrm{Fun}(\mathcal{I}, \mathcal{S})^{\otimes}$. Recall from Definition 3.4 that $\mathcal{I}^{\otimes}$ is promonoidal if the functor $\mathcal{I}^{\otimes} \to \mathrm{Fin}_*$ is exponentiable over $\mathrm{Fin} \simeq (\mathrm{Fin}_*)^{\mathrm{act}}$. By the characterization of exponentiability in [Ayala and Francis 2020, Lemma 1.10(c)], we need to show that for every map $f: I \to J$ in Fin, every $x \in \mathcal{I}^I$ and every $z \in \mathcal{I}$ the map

$$\int^{y \in \mathcal{I}^J} \mathrm{Mul}_{\mathcal{I}}(\{y_j\}_{j \in J}, z) \times \prod_{j \in J} \mathrm{Mul}_{\mathcal{I}}(\{x_i\}_{i \in f^{-1} j}, y_j) \to \mathrm{Mul}_{\mathcal{I}}(\{x_i\}_{i \in I}, z)$$

is an equivalence. Using that $\mathcal{I} \subseteq \mathcal{D}^{\mathrm{op}}$ is a full suboperad, this is equivalent to asking that the map

$$\int^{y \in \mathcal{I}^J} \prod_{j \in J} \mathrm{Map}_{\mathcal{D}}\Bigl(y_j, \bigotimes_{i \in f^{-1}j} x_i\Bigr) \times \mathrm{Map}_{\mathcal{D}}\Bigl(z, \bigotimes_{j \in J} y_j\Bigr) \to \mathrm{Map}_{\mathcal{D}}\Bigl(z, \bigotimes_{i \in I} x_i\Bigr)$$

is an equivalence of spaces. But since $\mathrm{Map}_{\mathcal{D}}(z, -)$ commutes with all colimits (as $z \in \mathcal{I}$ is tiny) it is enough to show that the map

$$\int^{y \in \mathcal{I}^J} \Bigl(\prod_{j \in J} \mathrm{Map}_{\mathcal{D}}\Bigl(y_j, \bigotimes_{i \in f^{-1}j} x_i\Bigr)\Bigr) \otimes \bigotimes_{j \in J} y_j \to \bigotimes_{i \in I} x_i$$

is an equivalence. Since the tensor product in $\mathcal{D}$ commutes with colimits in each variable, we can bring all the colimits inside (using that $\mathrm{Tw}(\mathcal{I}^J) \simeq \mathrm{Tw}(\mathcal{I})^J$). We are reduced to proving that the map

$$\bigotimes_{j \in J} \int^{y_j \in \mathcal{C}} \mathrm{Map}_{\mathcal{D}}\Bigl(y_j, \bigotimes_{i \in f^{-1}j} x_i\Bigr) \otimes y_j \to \bigotimes_{i \in I} x_i$$

is an equivalence. But this follows from the fact that for any $j \in J$ and $w \in \mathcal{D}$, the map

$$\int^{y_j \in \mathcal{C}} \mathrm{Map}(y_j, w) \times y_j \simeq \operatorname*{colim}_{y_j \in \mathcal{C}_{/w}} y_j \to w$$

is an equivalence, which is just another form of the Yoneda lemma. □

We are ready to prove our classification result.

Theorem 3.37 *Let $\mathcal{I}$ be a small ∞-category and suppose $\mathrm{Fun}(\mathcal{I}, \mathcal{S})$ is equipped with a symmetric monoidal structure $\mathrm{Fun}(\mathcal{I}, \mathcal{S})^{\otimes}$ which is compatible with colimits. Equip $\mathcal{I}^{\otimes}$ with the ∞-operad structure induced by the Yoneda embedding $\mathcal{I} \subseteq \mathrm{Fun}(\mathcal{I}, \mathcal{S})^{\mathrm{op}}$. Then $\mathcal{I}^{\otimes}$ is symmetric promonoidal and the symmetric monoidal structure on $\mathrm{Fun}(\mathcal{I}, \mathcal{S})$ is equivalent to the one induced by Day convolution with the symmetric promonoidal structure on $\mathcal{I}^{\otimes}$.*

Proof It follows from Lemma 3.36 that $\mathcal{I}^{\otimes}$ is symmetric promonoidal. Consider the composite

$$\mathcal{I}^{\otimes} \times_{\mathrm{Fin}_*} \mathrm{Fun}(\mathcal{I}, \mathcal{S})^{\otimes} \to (\mathrm{Fun}(\mathcal{I}, \mathcal{S})^{\mathrm{op}})^{\otimes} \times_{\mathrm{Fin}_*} \mathrm{Fun}(\mathcal{I}, \mathcal{S})^{\otimes} \to \mathcal{S}^{\times}$$

of lax symmetric monoidal functors, where the first functor is induced by the Yoneda embedding and the second is the lax symmetric monoidal enhancement of the mapping space functor constructed in [Glasman 2016, Section 3]. By the universal property of the Day convolution, we obtain a map of ∞-operads

$$\mathrm{Fun}(\mathcal{I}, \mathcal{S})^{\otimes} \to \mathrm{Fun}(\mathcal{I}^{\otimes}, \mathcal{S}^{\times})^{\mathrm{Day}},$$

which is the identity on underlying ∞-categories. Therefore to prove our thesis it will suffice to show that this functor is symmetric monoidal. Since $\mathrm{Fun}(\mathcal{I}, \mathcal{S})$ is generated under colimits by the corepresentable functors and both tensor products commute with colimits in each variable, it is enough to check that the maps

$$\mathrm{Mul}_{\mathcal{I}}(\varnothing,-) \simeq 1 \to 1^{\mathrm{Day}},$$
$$\mathrm{Mul}_{\mathcal{I}}(\{x,y\},-) \simeq \mathrm{Map}_{\mathcal{I}}(x,-)\otimes \mathrm{Map}_{\mathcal{I}}(y,-) \to \mathrm{Map}_{\mathcal{I}}(x,-)\otimes^{\mathrm{Day}} \mathrm{Map}_{\mathcal{I}}(y,-)$$

are equivalences for all $x, y \in \mathcal{I}$. But this follows from Corollary 3.29. □

Recall that the ∞-category of pointed objects in a presentably symmetric monoidal ∞-category is canonically symmetric monoidal. For later use we also record how taking pointed objects in a category of diagram spaces interacts with the Day convolution symmetric monoidal structure.

Proposition 3.38 *Consider a small promonoidal ∞-category $\mathcal{I}$, and a presentably symmetric monoidal ∞-category $\mathcal{C}$. There exists a symmetric monoidal equivalence*

$$(\mathcal{I}\text{–}\mathcal{C})_* \simeq \mathcal{I}\text{–}\mathcal{C}_*.$$

Proof Consider the lax monoidal functor $\mathcal{I}\text{–}\mathcal{C} \to \mathcal{I}\text{–}\mathcal{C}_*$ induced by the universal property of Day convolution by the composite

$$\mathrm{Fun}(\mathcal{I}^{\otimes}, \mathcal{C}^{\otimes}) \times_{\mathrm{Fin}_*} \mathcal{I}^{\otimes} \to \mathcal{C}^{\otimes} \xrightarrow{(-)_+} (\mathcal{C}_*)^{\wedge\otimes}.$$

Because $(-)_+$ is strong monoidal and colimit-preserving, one calculates that this functor is in fact strong monoidal. Therefore by [Lurie 2017, Proposition 4.8.2.11] we obtain an induced strong monoidal functor $(\mathcal{I}\text{–}\mathcal{C})_* \to \mathcal{I}\text{–}\mathcal{C}_*$, which is easily seen to be the identity on underlying categories. □

3.2 A symmetric monoidal Elmendorf's theorem

In this subsection we give a general ∞-categorical version of Elmendorf's theorem. We then enhance this to a symmetric monoidal equivalence.

Theorem 3.39 (Elmendorf) *Let $\mathcal{C}$ be a cocomplete ∞-category and let $i : \mathcal{C}_0 \to \mathcal{C}$ be the inclusion of a small full subcategory satisfying the following conditions*:

(a) *The objects of $\mathcal{C}_0$ are tiny*: *for all $c \in \mathcal{C}_0$, the functor $\mathrm{Map}_{\mathcal{C}}(c,-)$ preserves small colimits.*

(b) *The collection of objects $\{c_0 \in \mathcal{C}_0\}$ is jointly conservative*: *an arrow f in $\mathcal{C}$ is an equivalence if and only if $\mathrm{Map}_{\mathcal{C}}(c_0, f)$ is so for all $c_0 \in \mathcal{C}_0$.*

Then the restricted Yoneda functor induces an equivalence of ∞-categories $j : \mathcal{C} \simeq \mathcal{P}(\mathcal{C}_0)$.

Proof By the universal property of the category of presheaves [Lurie 2009, Theorem 5.1.5.6], there exists a colimit-preserving functor $L : \mathcal{P}(\mathcal{C}_0) \to \mathcal{C}$ such that $Lj_0 \simeq i$, where $j_0 : \mathcal{C}_0 \to \mathcal{P}(\mathcal{C}_0)$ denotes the Yoneda embedding. By the adjoint functor theorem [Nguyen et al. 2020, Corollary 4.1.4], the functor L admits a right adjoint $R : \mathcal{C} \to \mathcal{P}(\mathcal{C}_0)$, which is defined via the formula

$$R_c(c_0) = \mathrm{Map}_{\mathcal{C}}(Lj_0(c_0), c) \simeq \mathrm{Map}_{\mathcal{C}}(i(c_0), c)$$

for all $c \in \mathscr{C}$ and $c_0 \in \mathscr{C}_0$. Therefore R can be identified with the restricted Yoneda functor $j : \mathscr{C} \to \mathscr{P}(\mathscr{C}_0)$. We note that the functor j preserves all small colimits since for all $c_0 \in \mathscr{C}_0$, the functor

$$\mathrm{Map}_{\mathscr{C}}(c_0, -) : \mathscr{C} \xrightarrow{j} \mathscr{P}(\mathscr{C}_0) \xrightarrow{\mathrm{ev}_{c_0}} \mathscr{C}_0$$

does so by condition (a). As equivalences in $\mathscr{P}(\mathscr{C}_0)$ are detected pointwise, the same argument as above using condition (b) then shows that j is conservative. Note that the unit map $\eta : 1 \to jL$ is an equivalence on all objects in the image of j_0 as by construction $jLj_0 \simeq ji = j_0$. It follows that the unit map is an equivalence on all objects as $\mathscr{P}(\mathscr{C}_0)$ is generated under colimits by the representable functors and all the functors involved preserve colimits. Using the triangle identities of the adjunction we then find that $j(\epsilon)$ is an equivalence and so the counit map $\epsilon : Lj \to 1$ is an equivalence by conservativity of j. Thus j and L are inverse equivalences. $\square$

Example 3.40 Let G be a topological group and let $G\mathscr{T}$ be a convenient category of G-spaces. There is a model structure on $G\mathscr{T}$ where a map $f : X \to Y$ of G-spaces is a weak equivalence (resp. fibration) if $f^H : X^H \to Y^H$ is a weak homotopy equivalence (resp. Serre fibration) for all closed subgroups $H \leq G$; see [Schwede 2018, Proposition B.7]. Let $\mathscr{S}_G$ denote the underlying ∞-category of this model category, which is cocomplete by [Barnea et al. 2017, Theorem 2.5.9]. Moreover, colimits in $\mathscr{S}_G$ of projective cofibrant diagrams can be calculated as homotopy colimits in $G\mathscr{T}$ by [Barnea et al. 2017, Remark 2.5.7]. Let $\boldsymbol{O}_G \leq \mathscr{S}_G$ be the full subcategory of G-spaces spanned by the cosets G/H where H runs over all closed subgroups of G. Note that $G/H \in \mathscr{S}_G$ corepresents the H-fixed-point functors so the collection of cosets $\{G/H \mid H \leq G\}$ is jointly conservative by definition of weak equivalences in $G\mathscr{T}$. The fact that $G/H \in \mathscr{S}_G$ is tiny then follows from the fact that the H-fixed-point functor commutes with all small homotopy colimits [Schwede 2018, Proposition B.1(i) and (ii)]. Then the theorem above gives an equivalence $\boldsymbol{O}_G^{\mathrm{op}}\text{-}\mathscr{S} \simeq \mathscr{S}_G$. Therefore the previous theorem is a generalization of the classical theorem of Elmendorf [1983].

Under suitable assumptions we now enhance this to a symmetric monoidal equivalence, where we endow the presheaf category with Day convolution for a promonoidal structure on subcategory of tiny objects.

Corollary 3.41 *Suppose we are in the setting of Theorem 3.39 and that, furthermore, the following hold:*

(a) *$\mathscr{C}$ admits a symmetric monoidal structure $\mathscr{C}^{\otimes}$ which is compatible with colimits.*

(b) *$\mathscr{C}_0$ admits an ∞-operad structure $\mathscr{C}_0^{\otimes}$.*

(c) *The inclusion $i : \mathscr{C}_0 \to \mathscr{C}$ lifts to a fully faithful functor of ∞-operads $i^{\otimes} : \mathscr{C}_0^{\otimes} \to \mathscr{C}^{\otimes}$.*

Then $\mathscr{C}_0^{\otimes}$ is a symmetric promonoidal ∞-category and the restricted Yoneda embedding induces a symmetric monoidal equivalence $\mathscr{P}(\mathscr{C}_0)^{\mathrm{Day}} \simeq \mathscr{C}^{\otimes}$.

Proof By Theorem 3.39 there is a commutative diagram

$$\begin{tikzcd} \mathscr{C}_0 \arrow[r, "i"] \arrow[d, "j_0"'] & \mathscr{C} \arrow[dl, "j", "\sim"'] \\ \mathscr{P}(\mathscr{C}_0) & \end{tikzcd}$$

We can equip $\mathcal{P}(\mathcal{C}_0)$ with a symmetric monoidal structure $\mathcal{P}(\mathcal{C}_0)^{\otimes}$ induced by $\mathcal{C}^{\otimes}$ via j, and hence obtain a symmetric monoidal equivalence $j^{\otimes}\colon \mathcal{C}^{\otimes} \to \mathcal{P}(\mathcal{C}_0)^{\otimes}$. Combining this with condition (c) we obtain another commutative diagram

$$\begin{tikzcd} \mathcal{C}_0^{\otimes} \arrow[r, "i^{\otimes}"] \arrow[d, "j_0^{\otimes}"'] & \mathcal{C}^{\otimes} \arrow[dl, "j^{\otimes}", "\sim"'] \\ \mathcal{P}(\mathcal{C}_0)^{\otimes} & \end{tikzcd}$$

of ∞-operads. It is only left to note that by Theorem 3.37, the ∞-category $\mathcal{C}_0^{\otimes}$ is symmetric promonoidal and that the symmetric monoidal structure on $\mathcal{P}(\mathcal{C}_0)^{\otimes}$ coincides with the Day convolution product. $\square$

4 Partially lax limits

In this section we recall the necessary background on (partially) lax (co)limits and collect some important properties that we will use throughout the paper. The main references for this material are [Gepner et al. 2017; Berman 2024].

The notion of a partially lax limit over an ∞-category $\mathcal{I}$ is defined with reference to a collection of edges of $\mathcal{I}$. To make this precise we make the following definition.

Definition 4.1 A marked ∞-category is an ∞-category $\mathcal{C}$ along with a collection of edges $\mathcal{W} \subseteq \mathrm{Map}(\Delta^1, \mathcal{C})$ which contains all equivalences and which is stable under homotopy and composition. Given two marked ∞-categories $\mathcal{C}$ and $\mathcal{D}$, we write $\mathrm{Fun}^{\dagger}(\mathcal{C}, \mathcal{D})$ for the subcategory spanned by marked functors; those functors that preserve marked edges. We write $\mathrm{Cat}_{\infty}^{\dagger}$ for the ∞-category of marked ∞-categories. For the existence see [Lurie 2017, Construction 4.1.7.1].

Example 4.2 Let $\mathcal{C}$ be an ∞-category.

(a) There is a maximal marking $\mathcal{C}^{\sharp}$ where all morphisms are marked.

(b) There is a minimal marking $\mathcal{C}^{\flat}$ where only the equivalences are marked.

(c) Given a (co)cartesian fibration $p\colon \mathcal{C} \to \mathcal{I}^{\dagger}$ over a marked ∞-category, there is a marking $\mathcal{C}^{p}$ in which the (co)cartesian morphisms living over marked edges are marked.

Partially lax limits in an ∞-category $\mathcal{C}$ are also defined with reference to a cotensoring of $\mathcal{C}$ by Cat_{∞}. For the purposes of this paper, this is nothing but a functor $[-,-]\colon \mathrm{Cat}_{\infty}^{\mathrm{op}} \times \mathcal{C} \to \mathcal{C}$. The following examples are all naturally cotensored over Cat_{∞}.

Example 4.3 In the following $\mathcal{I}$ is an ∞-category.

(a) Clearly Cat_{∞} is cotensored over itself with cotensor given by $[\mathcal{I}, \mathcal{C}] = \mathrm{Fun}(\mathcal{I}, \mathcal{C})$.

(b) The ∞-category $\mathrm{Cat}_{\infty}^{\dagger}$ is cotensored over Cat_{∞} by considering $[\mathcal{I}, \mathcal{C}^{\dagger}] = \mathrm{Fun}(\mathcal{I}, \mathcal{C}^{\dagger})$, where we mark all those natural transformations whose components are all marked in $\mathcal{C}^{\dagger}$.

(c) The ∞-category of symmetric monoidal categories $\mathrm{Cat}_\infty^\otimes$ is cotensored over Cat_∞ by endowing the ∞-category $\mathrm{Fun}(\mathcal{I},\mathcal{C})$ with the pointwise symmetric monoidal structure $q\colon \mathrm{Fun}(\mathcal{I},\mathcal{C})^\otimes \to \mathrm{Fin}_*$ which is defined as follows. If $p\colon \mathcal{C}^\otimes \to \mathrm{Fin}_*$ is the cocartesian fibration witnessing the symmetric monoidal structure of $\mathcal{C}$, then we construct the pullback

$$\begin{tikzcd} \mathrm{Fun}(\mathcal{I},\mathcal{C})^\otimes \arrow[r, "q"] \arrow[d] & \mathrm{Fin}_* \arrow[d, "\mathrm{const}"] \\ \mathrm{Fun}(\mathcal{I},\mathcal{C}^\otimes) \arrow[r, "p_*"] & \mathrm{Fun}(\mathcal{I},\mathrm{Fin}_*) \end{tikzcd}$$

Note that by construction we have $\mathrm{Fun}(\mathcal{I},\mathcal{C})^\otimes_{\langle n\rangle} \simeq \mathrm{Fun}(\mathcal{I},\mathcal{C}^\otimes_{\langle n\rangle})$ for all $\langle n\rangle \in \mathrm{Fin}_*$. From this we immediately see that q satisfies the Segal conditions. The map p_* is a cocartesian fibration by the dual of [Lurie 2009, Proposition 3.1.2.1], and so by base-change [Lurie 2009, Proposition 2.4.2.3], q is too. Therefore q gives a symmetric monoidal structure on $\mathrm{Fun}(\mathcal{I},\mathcal{C})$.

(d) We can generalize the previous example as follows. Let $\mathcal{O}^\otimes \to \mathrm{Fin}_*$ be an ∞-operad. The ∞-category of ∞-operads Op_∞ is cotensored over Cat_∞ by endowing the ∞-category $\mathrm{Fun}(\mathcal{I},\mathcal{O})$ with the pointwise operadic structure induced by the map $\mathrm{Fun}(\mathcal{I},\mathcal{O}^\otimes) \times_{\mathrm{Fun}(\mathcal{I},\mathrm{Fin}_*)} \mathrm{Fin}_* \to \mathrm{Fin}_*$.

Similarly, partially lax colimits in $\mathcal{C}$ are defined with reference to a tensoring of $\mathcal{C}$ by Cat_∞. Once again, while more structured tensorings are typically useful, for our purposes it suffices for this to be a functor $(-)\otimes(-)\colon \mathrm{Cat}_\infty \times \mathcal{C} \to \mathcal{C}$. The most important example will be Cat_∞, for which the cartesian product gives a tensoring.

We now move on to the definition of partially lax (co)limits. For this we need to recall some categorical constructions. Recall the following result.

Lemma 4.4 [Lurie 2017, Proposition 4.1.7.2] *The minimal functor $(-)^\flat\colon \mathrm{Cat}_\infty \to \mathrm{Cat}_\infty^\dagger$ admits a left adjoint denoted by $|-|$.*

The ∞-category $|\mathcal{C}^\dagger|$ is obtained from $\mathcal{C}$ by adjoining formal inverses to all the marked morphisms, and so we call $|-|$ the localization functor.

Example 4.5 Given a model category $\mathcal{M}$, we may view it as a marked ∞-category by marking the weak equivalences in $\mathcal{M}$. Then $|\mathcal{M}| \simeq \mathcal{M}[W^{-1}]$.

Next we define marked slice categories.

Construction 4.6 Let $\mathcal{C}$ be an ∞-category. There is a functor $\mathcal{C}_{/-}\colon \mathcal{C} \to \mathrm{Cat}_\infty$ sending $x \in \mathcal{C}$ to the slice category $\mathcal{C}_{/x}$. This is obtained by straightening the cocartesian fibration given by the target map $t\colon \mathrm{Ar}(\mathcal{C}) := \mathrm{Fun}(\Delta^1,\mathcal{C}) \to \mathcal{C}$. One checks that a diagram

$$\begin{tikzcd} f_0 \arrow[r] \arrow[d, "f"'] & g_0 \arrow[d, "g"] \\ f_1 \arrow[r] & g_1 \end{tikzcd}$$

is a t-cocartesian edge if the top horizontal arrow is an equivalence. If $\mathscr{C}^\dagger$ is marked, then $\mathscr{C}^\dagger_{/x}$ has an induced marking where a morphism is marked if its image under the forgetful functor $\mathscr{C}^\dagger_{/x} \to \mathscr{C}^\dagger$ is a marked morphism. It is easy to see that this construction is functorial on x, and so we obtain a functor $\mathscr{C}^\dagger_{/-} \colon \mathscr{C} \to \mathrm{Cat}^\dagger_\infty$.

We are finally ready to introduce the notion of partially lax (co)limit. Recall the definition of the twisted arrow ∞-category from Definition 3.17.

Definition 4.7 Consider a functor $F \colon \mathscr{I} \to \mathscr{C}$ and choose a marking $\mathscr{I}^\dagger$.

(a) If $\mathscr{C}$ is cotensored over Cat_∞, then the partially lax limit of F is the limit of the composite

$$\mathrm{Tw}(\mathscr{I})^{\mathrm{op}} \xrightarrow{(s,t)^{\mathrm{op}}} \mathscr{I}^{\mathrm{op}} \times \mathscr{I} \xrightarrow{|\mathscr{I}^\dagger_{/-}| \times F} \mathrm{Cat}_\infty^{\mathrm{op}} \times \mathscr{C} \xrightarrow{[-,-]} \mathscr{C}.$$

We abbreviate this by $\mathrm{laxlim}^\dagger F$.

(b) If $\mathscr{C}$ is tensored over Cat_∞, then the partially lax colimit of F is the colimit of the composite

$$\mathrm{Tw}(\mathscr{I}) \xrightarrow{(s,t)} \mathscr{I} \times \mathscr{I}^{\mathrm{op}} \xrightarrow{F \times |(\mathscr{I}^{\mathrm{op}})^\dagger_{/-}|} \mathscr{C} \times \mathrm{Cat}_\infty \xrightarrow{-\otimes-} \mathscr{C}.$$

We abbreviate this by $\mathrm{laxcolim}^\dagger F$.

Remark 4.8 If we choose the minimal marking $\mathscr{I}^\flat$, then we recover the notion of lax (co)limit of [Gepner et al. 2017]. If we choose the maximal marking $\mathscr{I}^\sharp$, then we recover the usual notion of (co)limit; see [Berman 2024, Proposition 3.6].

In some cases we have a concrete description of the partially lax (co)limit.

Theorem 4.9 [Berman 2024, Theorem 4.4] *Let $\mathscr{I}^\dagger$ be a small marked ∞-category and let $F \colon \mathscr{I} \to \mathrm{Cat}_\infty$ be a functor. Consider the source of the (co)cartesian fibrations $\mathrm{Un}^{\mathrm{ct}}(F) \to \mathscr{I}^{\mathrm{op}}$ and $\mathrm{Un}^{\mathrm{co}}(F) \to \mathscr{I}$ as marked via Example 4.2(c).*

(a) *The partially lax limit of F is the ∞-category of marked sections of $p \colon \mathrm{Un}^{\mathrm{co}}(F) \to \mathscr{I}^\dagger$. In other words, we have*

$$\mathrm{laxlim}^\dagger F \simeq \mathrm{Fun}^\dagger_{/I^\dagger}(\mathscr{I}^\dagger, \mathrm{Un}^{\mathrm{co}}(F)).$$

(b) *The partially lax colimit of F is given by the localization of $\mathrm{Un}^{\mathrm{ct}}(F)$ at the marked edges. In other words, we have*

$$\mathrm{laxcolim}^\dagger F = |\mathrm{Un}^{\mathrm{ct}}(F)|.$$

Remark 4.10 The previous result gives a more explicit description of the partially lax limit of F. Recall that informally the Grothendieck construction $\mathrm{Un}^{\mathrm{co}}(F)$ is the ∞-category whose objects are pairs (X, i) where $i \in \mathscr{I}$ and $X \in F(i)$. A morphism from (X, i) to (Y, j) is a pair (φ, f) where $f \colon i \to j$ is a morphism in $\mathscr{I}$ and $\varphi \colon F(f)(X) \to Y$ is a morphism in $F(j)$. Then the previous result informally implies that $\mathrm{laxlim}^\dagger F$ is equivalent to the ∞-category whose objects are coherent collections of objects

$(X_i \in F(i))_{i\in\mathcal{I}}$ together with maps $\varphi_f\colon F(f)(X_i)\to X_j$ for every arrow $f\colon i\to j$ in $\mathcal{I}$, such that the map φ_f is an equivalence whenever f is marked.

We record some useful properties of partially lax (co)limits.

Proposition 4.11 *Let $\mathcal{I}^\dagger$ be a marked ∞-category and let $F\colon\mathcal{I}\to\mathrm{Cat}_\infty$ be a functor. Given any other ∞-category $\mathscr{C}$, we have an equivalence*

$$\mathrm{Fun}(\operatorname*{laxcolim}^\dagger_{\mathcal{I}} F,\mathscr{C})\simeq \operatorname*{laxlim}^\dagger_{\mathcal{I}^{\mathrm{op}}}\mathrm{Fun}(F(-),\mathscr{C}).$$

Proof The partially lax colimit of $F\colon\mathcal{I}\to\mathrm{Cat}_\infty$ is by definition calculated via the formula

$$\operatorname{laxcolim}^\dagger F=\operatorname*{colim}_{\mathrm{Tw}(\mathcal{I})} F\times|(\mathcal{I}^{\mathrm{op}})^\dagger_{/-}|.$$

Postcomposing by the limit-preserving functor $\mathrm{Fun}(-,\mathscr{C})\colon\mathrm{Cat}_\infty^{\mathrm{op}}\to\mathrm{Cat}_\infty$, we deduce that the ∞-category $\mathrm{Fun}(\operatorname{laxcolim}^\dagger F,\mathscr{C})$ is the limit of the diagram

$$\mathrm{Tw}(\mathcal{I})^{\mathrm{op}}\xrightarrow{(s,t)^{\mathrm{op}}}\mathcal{I}^{\mathrm{op}}\times\mathcal{I}\xrightarrow{(F,|(\mathcal{I}^{\mathrm{op}})^\dagger_{/-}|)^{\mathrm{op}}}\mathrm{Cat}_\infty^{\mathrm{op}}\times\mathrm{Cat}_\infty^{\mathrm{op}}\xrightarrow{-\times-}\mathrm{Cat}_\infty^{\mathrm{op}}\xrightarrow{\mathrm{Fun}(-,\mathscr{C})}\mathrm{Cat}_\infty. \tag{4.11.1}$$

By adjunction, we find that the composite of the final three functors is equivalent to

$$\mathrm{Fun}(-,-)\circ(|(\mathcal{I}^{\mathrm{op}})^\dagger_{/-}|,\mathrm{Fun}(F(-),\mathscr{C}))\circ\sigma\colon\mathcal{I}^{\mathrm{op}}\times\mathcal{I}\to\mathrm{Cat}_\infty,$$

where σ is the symmetry isomorphism of the product. As indicated in Remark 3.20, the following triangle commutes:

$$\begin{tikzcd}
\mathrm{Tw}(\mathcal{I})^{\mathrm{op}} \arrow[rr, "\sim"] \arrow[dr, "(s,t)^{\mathrm{op}}"'] & & \mathrm{Tw}(\mathcal{I}^{\mathrm{op}})^{\mathrm{op}} \arrow[dl, "(t,s)^{\mathrm{op}}"] \\
& \mathcal{I}^{\mathrm{op}}\times\mathcal{I} &
\end{tikzcd}$$

These two observations allow us to rewrite equation (4.11.1) and conclude that $\mathrm{Fun}(\operatorname{laxcolim}^\dagger F,\mathscr{C})$ is the limit of the functor

$$\mathrm{Tw}(\mathcal{I}^{\mathrm{op}})^{\mathrm{op}}\xrightarrow{(s,t)^{\mathrm{op}}}\mathcal{I}\times\mathcal{I}^{\mathrm{op}}\xrightarrow{(|(\mathcal{I}^{\mathrm{op}})^\dagger_{/-}|,\mathrm{Fun}(F(-),\mathscr{C}))}\mathrm{Cat}_\infty^{\mathrm{op}}\times\mathrm{Cat}_\infty\xrightarrow{\mathrm{Fun}(-,-)}\mathrm{Cat}_\infty,$$

which is exactly the definition of the partially lax limit of $\mathrm{Fun}(F(-),\mathscr{C})\colon\mathcal{I}^{\mathrm{op}}\to\mathrm{Cat}_\infty$. □

We finish this section by discussing how (partially) lax limits interact with localizations. Later on we will use these results to pass from (partially) lax limits of prespectra to that of spectra.

Lemma 4.12 *Let $\mathcal{I}$ be an ∞-category and let $F\colon\mathcal{I}\to\mathrm{Cat}_\infty$ be a functor. Suppose that for every $i\in\mathcal{I}$ we are given a reflexive subcategory $Gi\subseteq Fi$ with left adjoint $L_i\colon Fi\to Gi$. If for every arrow $f\colon i\to j$ of $\mathcal{I}$, the pushforward functor $f_*\colon Fi\to Fj$ sends L_i-equivalences to L_j-equivalences, then there is a functor $G\colon\mathcal{I}\to\mathrm{Cat}_\infty$ and a natural transformation $L\colon F\Rightarrow G$ whose i^{th} component is given by $L_i\colon Fi\to Gi$. Furthermore, the functor*

$$\operatorname*{laxlim}_{\mathcal{I}} L\colon\operatorname*{laxlim}_{\mathcal{I}} F\to\operatorname*{laxlim}_{\mathcal{I}} G$$

has a fully faithful right adjoint.

Proof Let us consider the Grothendieck construction $\mathrm{Un}^{\mathrm{co}}(F) \to \mathcal{I}$ of F. This is the cocartesian fibration classified by F under the straightening-unstraightening equivalence, so in particular the fiber over $i \in \mathcal{I}$ can be canonically identified with Fi. Let $\mathscr{E} \subseteq \mathrm{Un}^{\mathrm{co}}(F)$ be the full subcategory spanned by the objects of $Gi \subseteq \mathrm{Un}^{\mathrm{co}}(F)$ for all $i \in \mathcal{I}$.

We claim that $\mathscr{E} \to \mathcal{I}$ is a cocartesian fibration whose cocartesian edges are those that can be factored in $\mathrm{Un}^{\mathrm{co}}(F)$ as a cocartesian edge of $\mathrm{Un}^{\mathrm{co}}(F)$ followed by a L_i-equivalence in the fiber over i. More precisely, if $f : i \to j$ is an arrow of $\mathcal{I}$ and $x \in Gi$, then the cocartesian lift of f starting from x is the composition $x \to f_* x \to L_j(f_* x)$ where the first arrow is the cocartesian lift of f in $\mathrm{Un}^{\mathrm{co}}(F)$.

Indeed, for every $z \in Gj$, we have

$$\mathrm{Map}^f_{\mathscr{E}}(x, z) \simeq \mathrm{Map}_{Fj}(f_* x, z) \simeq \mathrm{Map}_{Gj}(L_j f_* x, z),$$

and so those edges are locally cocartesian. Furthermore, it is easy to see they are stable under composition (using the fact that L-equivalences are stable under pushforward), therefore they are cocartesian arrows by [Lurie 2009, Lemma 2.4.2.7].

The inclusion $\iota : \mathscr{E} \subseteq \mathrm{Un}^{\mathrm{co}}(F)$ has a relative left adjoint, which is a map of cocartesian fibrations by [Lurie 2017, Proposition 7.3.2.11]. Therefore there is a functor $G : \mathcal{I} \to \mathrm{Cat}_\infty$ and a natural transformation $L : F \Rightarrow G$ such that $\mathscr{E}$ can be identified with $\mathrm{Un}^{\mathrm{co}}(G)$ in such a way that the induced map $L : \mathrm{Un}^{\mathrm{co}}(F) \to \mathrm{Un}^{\mathrm{co}}(G)$ agrees with $L_i : Fi \to Gi$ on each fiber.

Finally, by Theorem 4.9 the lax limit of F and G are computed by the ∞-categories of sections of the respective cocartesian fibrations, and $\mathrm{laxlim}_{\mathcal{I}} L$ is given by postcomposition with L. Therefore postcomposition with ι gives a fully faithful right adjoint to $\mathrm{laxlim}_{\mathcal{I}} L$. □

Lemma 4.13 *Suppose we are in the situation of Lemma 4.12, and suppose $\mathcal{I}$ is equipped with a marking $\mathcal{I}^\dagger$ such that for every marked edge $f : i \to j$ the pushforward functor $f_* : Fi \to Fj$ sends Gi into Gj. Then the functor*

$$\operatorname*{laxlim}_{\mathcal{I}^\dagger} L : \operatorname*{laxlim}_{\mathcal{I}^\dagger} F \to \operatorname*{laxlim}_{\mathcal{I}^\dagger} G$$

has a fully faithful right adjoint. In particular, $\mathrm{laxlim}_{\mathcal{I}^\dagger} L$ *is a localization functor.*

Proof It suffices to show that the right adjoint of Lemma 4.12 sends $\mathrm{laxlim}_{\mathcal{I}^\dagger} G$ into $\mathrm{laxlim}_{\mathcal{I}^\dagger} F$. Recall that the partially lax limit can be calculated as the subcategory of sections spanned by those sending marked edges to cocartesian arrows. Thus, we ought to show that the right adjoint preserves cocartesian arrows lying over marked edges. But the right adjoint is given by postcomposing a section with the inclusion $\mathrm{Un}^{\mathrm{co}}(G) \to \mathrm{Un}^{\mathrm{co}}(F)$, and so by the description of cocartesian edges given in Lemma 4.12 and by our hypothesis, it sends cocartesian arrows over marked edges to cocartesian arrows (here we are implicitly using that an L_i-equivalence between objects of Gi is automatically an equivalence in Fi and so in particular a cocartesian arrow). □

For later reference we record the following immediate corollary of Lemma 4.12.

Corollary 4.14 *Let $\mathcal{I}$ be an ∞-category and let $F\colon \mathcal{I} \to \mathrm{Cat}_\infty^\otimes$ be a functor. Suppose that for every $i \in \mathcal{I}$, we are given a reflexive subcategory $Gi \subseteq Fi$ with left adjoint $L_i\colon Fi \to Gi$ which is compatible with the symmetric monoidal structure in the sense of* [Lurie 2017, Definition 2.2.1.6]. *Suppose furthermore that for every arrow $f\colon i \to j$ in I, the pushforward functor $f_*\colon Fi \to Fj$ sends L_i-equivalences to L_j-equivalences. Then there exists a functor $G\colon \mathcal{I} \to \mathrm{Cat}_\infty^\otimes$ and a symmetric monoidal natural transformation $L\colon F \Rightarrow G$ whose i^{th} component is given by $L_i\colon Fi \to Gi$.*

Proof $\mathrm{Cat}_\infty^\otimes$ embeds as a subcategory of $\mathrm{Fun}(\mathrm{Fin}_*, \mathrm{Cat}_\infty)$, so consider the functor $\widetilde{F}\colon \mathrm{Fin}_* \times \mathcal{I} \to \mathrm{Cat}_\infty$ induced by F, so that $\widetilde{F}(A_+, i) \simeq (Fi)^A$ (the fiber over A of $Fi \to \mathrm{Fin}_*$). If we let $\widetilde{G}(A_+, i) = (Gi)^A \subseteq \widetilde{F}(A_+, i)$, we can apply Lemma 4.12 to $\widetilde{F}$. To see that the pushforwards respect local equivalences, it suffices to prove this separately for maps of the form (σ, id) and (id, f) in $\mathrm{Fin}_* \times \mathcal{I}$. However, both of these cases are ensured by our assumptions. Therefore there exists a functor

$$\widetilde{G}\colon \mathrm{Fin}_* \times \mathcal{I} \to \mathrm{Cat}_\infty$$

and a natural transformation $\widetilde{L}\colon \widetilde{F} \Rightarrow \widetilde{G}$ as desired. By construction $\widetilde{G}$ satisfies the Segal conditions, and so it induces a functor $G\colon \mathcal{I} \to \mathrm{Cat}_\infty^\otimes$ with a symmetric monoidal natural transformation $L\colon F \Rightarrow G$ as desired. $\square$

5 Partially lax limits of symmetric monoidal ∞-categories

Recall that Op_∞ is canonically cotensored over Cat_∞ by Example 4.3. Therefore we immediately obtain a definition of partially lax limits of diagrams in Op_∞. In this section we will collect some important properties of partially lax limits of symmetric monoidal ∞-categories and ∞-operads. In particular the calculations of Proposition 5.8 and Theorem 5.10 are used repeatedly in part two. The first is analogous to the calculation of the (partially) lax limit of a diagram of ∞-categories, and as such it is stated in terms of an unstraightening equivalence for symmetric monoidal categories, which we recall in Proposition 5.5.

Remark 5.1 If $\mathcal{P}^\otimes$ is another ∞-operad, it follows from the definition and [Lurie 2017, Remark 2.1.3.4] that there is a natural equivalence

$$\mathrm{Alg}_{\mathcal{P}^\otimes}\bigl(\operatorname*{laxlim}_{i\in I} \mathcal{O}_i^\otimes\bigr) \simeq \operatorname*{laxlim}_{i\in I} \mathrm{Alg}_{\mathcal{P}^\otimes}(\mathcal{O}^\otimes).$$

Such a natural equivalence then also uniquely determines the lax limit. Since $\mathrm{Cat}_\infty^\otimes \subseteq \mathrm{Op}_\infty$ is a subcategory closed under limits and cotensoring, it is also closed under partially lax limits. In particular we conclude that for every family of symmetric monoidal ∞-categories $\mathcal{C}_\bullet$ and every symmetric monoidal ∞-category $\mathcal{D}$, there is a natural equivalence

$$\mathrm{Fun}^\otimes\bigl(\mathcal{D}, \operatorname*{laxlim}_{i\in I} \mathcal{C}_i\bigr) \simeq \operatorname*{laxlim}_{i\in I} \mathrm{Fun}^\otimes(\mathcal{D}, \mathcal{C}_i).$$

We note that the underlying ∞-category functor $U\colon \mathrm{Op}_\infty \to \mathrm{Cat}_\infty$ preserves limits and commutes with cotensoring, and therefore preserves partially lax limits. Therefore the previous construction equips the partially lax limit of a family of symmetric monoidal ∞-categories with a canonical symmetric monoidal structure, which satisfies the expected universal property.

Remark 5.2 There is always a canonical map $\operatorname{laxlim}^\dagger \mathcal{O}_i^\otimes \to \operatorname{laxlim} \mathcal{O}_i^\otimes$. This functor is induced on limits by a natural transformation which is pointwise given by the inclusion of a fully faithful suboperad. Thus we conclude that the partially lax limit is always a fully faithful suboperad of the lax limit. In practice this means that we can determine which suboperad by considering the induced map on underlying categories.

In the second part of the paper we will build diagrams of symmetric monoidal ∞-categories indexed on $\mathrm{Glo}^{\mathrm{op}}$. Central to our constructions of these diagrams is an operadic variant of straightening/unstraightening, which we will recall now.

Notation 5.3 Recall from [Lurie 2017, 2.4.3.5] that for every ∞-category $\mathcal{I}$ there is a functor of ∞-operads $c\colon \mathcal{I} \times \mathrm{Fin}_* \to \mathcal{I}^{\amalg}$ sending (x, A_+) to the constant family $\{x\}_{a\in A} \in \mathcal{I}^{\amalg}_{A_+}$.

Construction 5.4 Let $\mathcal{I}$ be an ∞-category and let $\mathscr{C}^\otimes$ be an $\mathcal{I}^{\amalg}$-monoidal ∞-category. Then the commutative diagram of cocartesian fibrations

$$\begin{tikzcd} \mathscr{C}^\otimes \times_{\mathcal{I}^{\amalg}} (\mathcal{I} \times \mathrm{Fin}_*) \arrow[rr, "\mathrm{pr}_2"] \arrow[dr, "\mathrm{pr}_{\mathcal{I}}"'] & & \mathcal{I} \times \mathrm{Fin}_* \arrow[dl, "\mathrm{pr}_1"] \\ & \mathcal{I} & \end{tikzcd}$$

is classified by a functor $\mathscr{C}_\bullet\colon \mathcal{I} \to (\mathrm{Cat}_\infty)_{/\mathrm{Fin}_*}$, which lands in $\mathrm{Cat}_\infty^\otimes$. We refer to $\mathscr{C}_\bullet$ as the family of symmetric monoidal ∞-categories classifying $\mathscr{C}^\otimes$.

Proposition 5.5 *The previous construction furnishes an equivalence between the ∞-category of $\mathcal{I}^{\amalg}$-monoidal categories and* $\mathrm{Fun}(\mathcal{I}, \mathrm{Cat}_\infty^\otimes)$.

Proof This is [Drew and Gallauer 2022, Corollary A.12]. □

Definition 5.6 Consider a map of ∞-operads $p\colon \mathcal{O}^\otimes \to \mathcal{I}^{\amalg}$. Any object $i \in \mathcal{I}$ induces a functor

$$\{i\} \times \mathrm{Fin}_* \hookrightarrow \mathcal{I} \times \mathrm{Fin}_* \xrightarrow{c} \mathcal{I}^{\amalg};$$

see Notation 5.3. Equivalently, the map above can be obtained by applying $(-)^{\amalg}$ to the map $\Delta^0 \to \mathcal{I}$ defined by $i \in \mathcal{I}$. Inspired by the equivalence of Proposition 5.5 we will refer to the pullback $\mathcal{O}^\otimes \times_{\mathcal{I}^{\amalg}} \mathrm{Fin}_*$ as the *operadic fiber* of p at $i \in \mathcal{I}$. If p is an $\mathcal{I}^{\amalg}$-monoidal ∞-category, then its operadic fiber at i is a symmetric monoidal ∞-category, and corresponds to the value of the functor $\mathscr{C}_\bullet$ at i.

The following example will be crucial for later applications.

Example 5.7 Let $p\colon \mathscr{C}^\otimes \to \mathcal{I}^{\amalg}$ be a $\mathcal{I}^{\amalg}$-promonoidal ∞-category and let $\mathscr{D}^\otimes \to \mathcal{I}^{\amalg}$ be a map of ∞-operads which is compatible with colimits. Then the operadic fiber of the Day convolution

$\mathrm{Fun}_{\mathcal{I}}(\mathcal{C}^{\otimes},\mathcal{D}^{\otimes})^{\mathrm{Day}}$ over $i\in\mathcal{I}$ is given by the symmetric monoidal ∞-category $\mathcal{C}_i\!-\!\mathcal{D}_i$, where $\mathcal{C}_i$ and $\mathcal{D}_i$ are the operadic fibers over i of $\mathcal{C}^{\otimes}$ and $\mathcal{D}^{\otimes}$, respectively. To see this, first recall that $\mathcal{C}_i\!-\!\mathcal{D}_i$ is defined to be $\mathrm{Fun}(\mathcal{C}_i,\mathcal{D}_i)$ with the Day convolution symmetric monoidal structure. Then the claim follows from the following computation using Lemma 3.9:

$$(N_p\,p^*\mathcal{D}^{\otimes})\times_{\mathcal{I}^{\amalg}}\mathrm{Fin}_*\simeq N_{p_i}(p^*\mathcal{D}^{\otimes}\times_{\mathcal{C}^{\otimes}}\mathcal{C}_i^{\otimes})\simeq N_{p_i}\,p_i^*\mathcal{D}_i^{\otimes}=\mathcal{C}_i\!-\!\mathcal{D}_i.$$

Recall that the lax limit of a diagram of ∞-categories was calculated by taking sections of the associated cocartesian fibration. Similarly, we can describe the lax limit of $\mathcal{C}_\bullet$ in terms of (suitable) sections of the ∞-operad $\mathcal{C}^{\otimes}$.

Proposition 5.8 *Let $\mathcal{C}^{\otimes}\to\mathcal{I}^{\amalg}$ be a $\mathcal{I}^{\amalg}$-monoidal ∞-category, and write $\mathcal{C}_\bullet\colon\mathcal{I}\to\mathrm{Cat}_\infty^{\otimes}$ for the associated diagram of symmetric monoidal ∞-categories. Then there is a natural equivalence of symmetric monoidal ∞-categories*

$$\mathrm{laxlim}\,\mathcal{C}_\bullet\simeq N_{\mathcal{I}^{\amalg}}\mathcal{C}^{\otimes},$$

where the right-hand side is the norm along $\mathcal{I}^{\amalg}\to\mathrm{Fin}_*$, which is well-defined by Example 3.6.

Proof We will show that the right-hand side has the universal property of the lax limit. By the universal property of the norm, for any ∞-operad $\mathcal{P}^{\otimes}$ we have an equivalence

$$\mathrm{Alg}_{\mathcal{P}^{\otimes}}(N_{\mathcal{I}^{\amalg}}\mathcal{C}^{\otimes})\simeq\mathrm{Alg}_{\mathcal{P}^{\otimes}\times_{\mathrm{Fin}_*}\mathcal{I}^{\amalg}/\mathcal{I}^{\amalg}}(\mathcal{C}^{\otimes}).$$

By [Lurie 2017, Theorem 2.4.3.18], we can write

$$\begin{aligned}\mathrm{Alg}_{\mathcal{P}^{\otimes}\times_{\mathrm{Fin}_*}\mathcal{I}^{\amalg}/\mathcal{I}^{\amalg}}(\mathcal{C}^{\otimes})&\simeq\mathrm{Alg}_{\mathcal{P}^{\otimes}\times_{\mathrm{Fin}_*}\mathcal{I}^{\amalg}}(\mathcal{C}^{\otimes})\times_{\mathrm{Alg}_{\mathcal{P}^{\otimes}\times_{\mathrm{Fin}_*}\mathcal{I}^{\amalg}}(\mathcal{I}^{\amalg})}\{\mathrm{pr}_2\}\\&\simeq\mathrm{Fun}(\mathcal{I},\mathrm{Alg}_{\mathcal{P}^{\otimes}}(\mathcal{C}^{\otimes}))\times_{\mathrm{Fun}(\mathcal{I},\mathrm{Alg}_{\mathcal{P}^{\otimes}}(\mathcal{I}^{\amalg}))}\{\mathrm{pr}_2\},\end{aligned}$$

where $\mathrm{pr}_2\colon\mathcal{P}^{\otimes}\times_{\mathrm{Fin}_*}\mathcal{I}^{\amalg}\to\mathcal{I}^{\amalg}$ is the projection. In other words, we have shown that $\mathrm{Alg}_{\mathcal{P}^{\otimes}}(N_{\mathcal{I}^{\amalg}}\mathcal{C}^{\otimes})$ is the ∞-category of sections of the functor

$$\mathrm{Alg}_{\mathcal{P}^{\otimes}}(\mathcal{C}^{\otimes})\times_{\mathrm{Alg}_{\mathcal{P}^{\otimes}}(\mathcal{I}^{\amalg})}\mathcal{I}\to\mathcal{I},$$

which is exactly the cocartesian fibration classified by $i\mapsto\mathrm{Alg}_{\mathcal{P}^{\otimes}}(\mathcal{C}_i^{\otimes})$. Our thesis then follows from Theorem 4.9. □

Remark 5.9 Let $p\colon\mathcal{C}^{\otimes}\to\mathcal{I}^{\amalg}$ be an $\mathcal{I}^{\amalg}$-monoidal ∞-category, and write $\mathcal{C}_\bullet\colon\mathcal{I}\to\mathrm{Cat}_\infty^{\otimes}$ for the associated diagram of symmetric monoidal ∞-categories. Then by the discussion in Remark 3.11, the underlying category of $N_{\mathcal{I}^{\amalg}}\mathcal{C}^{\otimes}$ is given by $\mathrm{Fun}_{/\mathcal{I}}(\mathcal{I},\mathcal{C})$. Therefore the proposition above is an operadic analogue of Theorem 4.9(b). Since we know that the partially lax limit of a diagram of ∞-operads is a fully faithful suboperad of the lax limit, the previous result also allows us to calculate the partially lax limit of $\mathcal{C}_\bullet$. Namely it is the fully faithful symmetric monoidal subcategory of $N_{\mathcal{I}^{\amalg}}\mathcal{C}^{\otimes}$ determined by the fully faithful subcategory $\mathrm{laxlim}^{\dagger}\,\mathcal{C}_\bullet\subset\mathrm{laxlim}\,\mathcal{C}_\bullet$.

We finish this section by proving that the formation of (partially) lax limits of symmetric monoidal categories commutes with taking modules, in a precise sense. This will be a key observation for the second part of the paper, and crucially uses the equivalence $N_{\mathcal{I}^{\amalg}}\mathcal{C}^{\otimes} \simeq \operatorname{laxlim}\mathcal{C}_\bullet$.

Theorem 5.10 *Let $\mathcal{C}^{\otimes} \to \mathcal{I}^{\amalg}$ be a $\mathcal{I}^{\amalg}$-monoidal ∞-category which is compatible with colimits, and write $\mathcal{C}_\bullet\colon \mathcal{I} \to \mathrm{Cat}_\infty^{\otimes}$ for the associated diagram of symmetric monoidal ∞-categories. Let $S \in \mathrm{CAlg}(\operatorname{laxlim}\mathcal{C}_\bullet)$ be a commutative algebra in the lax limit, which corresponds to a (partially lax) family of commutative algebras $S_i \in \mathrm{CAlg}(\mathcal{C}_i)$. Then there is a functor*

$$\mathrm{Mod}_{S_\bullet}(\mathcal{C}_\bullet)\colon \mathcal{I} \to \mathrm{Cat}_\infty^{\otimes}, \quad i \mapsto \mathrm{Mod}_{S_i}(\mathcal{C}_i),$$

and an equivalence of symmetric monoidal ∞-categories

$$\operatorname{laxlim}\mathrm{Mod}_{S_\bullet}(\mathcal{C}_\bullet) \simeq \mathrm{Mod}_S(\operatorname{laxlim}\mathcal{C}_\bullet).$$

Moreover, there is a natural transformation $\mathcal{C}_\bullet \to \mathrm{Mod}_{S_\bullet}(\mathcal{C}_\bullet)$ sending $x \in \mathcal{C}_i$ to the free S_i-module $S_i \otimes x$, which induces the functor $S \otimes -$ on the lax limit.

The proof of the previous result will require some preparation and some results from the appendix. For this reason we recommend the reader to skip this part on a first reading.

We start our journey by studying how the lax limit interacts with the tensor product of algebras.

Construction 5.11 By [Lurie 2017, Proposition 3.2.4.6] there is an equivalence of ∞-operads

$$\mathcal{I}^{\amalg} \otimes_{BV} \mathrm{Fin}_* \simeq \mathcal{I}^{\amalg},$$

where $\otimes_{BV}$ is the Boardman–Vogt tensor product, and so there exists a unique bifunctor of ∞-operads $\mathcal{I}^{\amalg} \times \mathrm{Fin}_* \to \mathcal{I}^{\amalg}$. For any ∞-operad $\mathcal{O}^{\otimes}$ we obtain a bifunctor of ∞-operads $m_{\mathcal{O}}$, which is given by the composition

$$\mathcal{I}^{\amalg} \times \mathcal{O}^{\otimes} \to \mathcal{I}^{\amalg} \times \mathrm{Fin}_* \to \mathcal{I}^{\amalg}.$$

Thus, for every map of ∞-operad $\mathcal{C}^{\otimes} \to \mathcal{I}^{\amalg}$ [Lurie 2017, Construction 3.2.4.1] produces a map of ∞-operads

$$\mathrm{Alg}_{\mathcal{O}^{\otimes}/\mathcal{I}^{\amalg}}(\mathcal{C})^{\otimes} \to \mathcal{I}^{\amalg},$$

whose operadic fiber over $i \in I$ is given by $\mathrm{Alg}_{\mathcal{O}^{\otimes}}(\mathcal{C}_i)^{\otimes}$. Suppose that $\mathcal{C}^{\otimes}$ is a $\mathcal{I}^{\amalg}$-monoidal category. Then by [Lurie 2017, Proposition 3.2.4.3.(3)] $\mathrm{Alg}_{\mathcal{O}^{\otimes}}(\mathcal{C}_i)^{\otimes}$ is also a $\mathcal{I}^{\amalg}$-monoidal ∞-category. In this case, Proposition 5.5 gives a functor $\mathcal{I} \to \mathrm{Cat}_\infty^{\otimes}$ sending $i \in \mathcal{I}$ to $\mathrm{Alg}_{\mathcal{O}^{\otimes}}(\mathcal{C}_i)^{\otimes}$. We will now compute the lax limit of this functor.

Lemma 5.12 *Let $\mathcal{I}$ be an ∞-category, $\mathcal{C}^{\otimes} \to \mathcal{I}^{\amalg}$ a map of ∞-operads and $\mathcal{O}^{\otimes}$ an ∞-operad. Then there is a natural equivalence of ∞-operads*

$$\mathrm{Alg}_{\mathcal{O}^{\otimes}}(N_{\mathcal{I}^{\amalg}}\mathcal{C}^{\otimes})^{\otimes} \simeq N_{\mathcal{I}^{\amalg}}\mathrm{Alg}_{\mathcal{O}^{\otimes}/\mathcal{I}^{\amalg}}(\mathcal{C})^{\otimes}.$$

In particular if $\mathcal{C}^{\otimes}$ is $\mathcal{I}^{\amalg}$-symmetric monoidal we have a natural equivalence of ∞-operads

$$\mathrm{Alg}_{\mathcal{O}^{\otimes}}\big(\operatorname*{laxlim}_{i \in \mathcal{I}}\mathcal{C}_i\big)^{\otimes} \simeq \operatorname*{laxlim}_{i \in \mathcal{I}}\mathrm{Alg}_{\mathcal{O}}(\mathcal{C}_i)^{\otimes}.$$

Proof We will prove that both sides represent the same functor in the ∞-category of ∞-operads. Let $\mathcal{P}^\otimes$ be an ∞-operad. Then

$$\begin{aligned}\mathrm{Alg}_{\mathcal{P}^\otimes} N_{\mathcal{I}^\amalg} \mathrm{Alg}_{\mathcal{O}^\otimes/\mathcal{I}^\amalg} \mathcal{C}^\otimes &\simeq \mathrm{Alg}_{\mathcal{P}^\otimes \times_{\mathrm{Fin}_*} \mathcal{I}^\amalg / I^\amalg}(\mathrm{Alg}_{\mathcal{O}^\otimes}(\mathcal{C})^\otimes \times_{\mathrm{Alg}_{\mathcal{O}^\otimes}(\mathcal{I})^\otimes} \mathcal{I}^\amalg) \\ &\simeq \mathrm{Alg}_{\mathcal{P}^\otimes \times_{\mathrm{Fin}_*} \mathcal{I}^\amalg / \mathrm{Alg}_{\mathcal{O}^\otimes}(\mathcal{I})^\otimes}(\mathrm{Alg}_{\mathcal{O}^\otimes}(\mathcal{C})^\otimes) \\ &\simeq \mathrm{Alg}_{\mathcal{P}^\otimes \times_{\mathrm{Fin}_*} \mathcal{I}^\amalg}(\mathrm{Alg}_{\mathcal{O}^\otimes}(\mathcal{C})^\otimes) \times_{\mathrm{Alg}_{\mathcal{P}^\otimes \times_{\mathrm{Fin}_*} \mathcal{I}^\amalg}(\mathrm{Alg}_{\mathcal{O}^\otimes}(\mathcal{I})^\otimes)} \{\mathrm{pr}_2\} \\ &\simeq \mathrm{Alg}_{(\mathcal{P}^\otimes \otimes_{BV} \mathcal{O}^\otimes) \times_{\mathrm{Fin}_*} \mathcal{I}^\amalg}(\mathcal{C}^\otimes) \times_{\mathrm{Alg}_{(\mathcal{P}^\otimes \otimes_{BV} \mathcal{O}^\otimes) \times_{\mathrm{Fin}_*} \mathcal{I}^\amalg}(\mathcal{I}^\amalg)} \{\mathrm{pr}_2\} \\ &\simeq \mathrm{Alg}_{(\mathcal{P}^\otimes \otimes_{BV} \mathcal{O}^\otimes) \times_{\mathrm{Fin}_*} \mathcal{I}^\amalg / \mathcal{I}^\amalg}(\mathcal{C})^\otimes \\ &\simeq \mathrm{Alg}_{\mathcal{P}^\otimes}(\mathrm{Alg}_{\mathcal{O}^\otimes}(N_{\mathcal{I}^\amalg}\mathcal{C})^\otimes).\end{aligned}$$

Here $\otimes_{BV}$ is the Boardman–Vogt tensor product of ∞-operads of [Lurie 2017, Section 2.2.5]. □

We are ready to prove the main result of this section.

Proof of Theorem 5.10 By the definition of the norm we have an equivalence

$$\mathrm{CAlg}(N_{\mathcal{I}^\amalg}\mathcal{C}^\otimes) \simeq \mathrm{Alg}_{\mathcal{I}^\amalg/\mathcal{I}^\amalg}(\mathcal{C}^\otimes) \simeq \mathrm{Alg}_{\mathcal{I}^\amalg}(\mathrm{Alg}_{\mathrm{Fin}_*/\mathcal{I}^\amalg}(\mathcal{C})^\otimes),$$

therefore we can also consider S as a section of $\mathrm{Alg}_{\mathrm{Fin}_*/\mathcal{I}^\amalg}(\mathcal{C})^\otimes \to \mathcal{I}^\amalg$ in Op_∞.

By Theorem 12.21 and Lemma 5.12 there is an equivalence

$$\begin{aligned}\mathrm{Mod}_S(N_{\mathcal{I}^\amalg}\mathcal{C}^\otimes)^\otimes &\simeq \mathrm{Alg}_{\mathcal{CM}}(N_{\mathcal{I}^\amalg}\mathcal{C})^\otimes \times_{\mathrm{CAlg}(N_{\mathcal{I}^\amalg}\mathcal{C})^\otimes} \mathrm{Fin}_* \\ &\simeq N_{\mathcal{I}^\amalg}(\mathrm{Alg}_{\mathcal{CM}/\mathcal{I}^\amalg}(\mathcal{C})^\otimes \times_{\mathrm{Alg}_{\mathrm{Fin}_*/\mathcal{I}^\amalg}(\mathcal{C})^\otimes} \mathcal{I}^\amalg),\end{aligned}$$

where $\mathcal{I}^\amalg \to \mathrm{Alg}_{\mathrm{Fin}_*/\mathcal{I}^\amalg}(\mathcal{C})^\otimes$ is the section corresponding to S. Moreover, by Lemma 12.20,

$$\mathrm{Alg}_{\mathcal{CM}/\mathcal{I}^\amalg}(\mathcal{C})^\otimes \times_{\mathrm{Alg}_{\mathrm{Fin}_*/\mathcal{I}^\amalg}(\mathcal{C})^\otimes} \mathcal{I}^\amalg \to \mathcal{I}^\amalg$$

is an $\mathcal{I}^\amalg$-monoidal ∞-category. Then Theorem 12.21 shows that the corresponding family of symmetric monoidal ∞-categories is exactly

$$i \mapsto \mathrm{Mod}_{S_i}(\mathcal{C}_i),$$

and so our thesis follows from Proposition 5.8.

Finally, let us construct the symmetric monoidal functor $\mathcal{C}_i^\otimes \to \mathrm{Mod}_{S_i}(\mathcal{C}_i)^\otimes$. There is a map of $\mathcal{I}^\amalg$-monoidal ∞-categories

$$\mathrm{Alg}_{\mathcal{CM}^\otimes/\mathcal{I}^\amalg}(\mathcal{C})^\otimes \to \mathrm{Alg}_{\mathrm{Fin}_*/\mathcal{I}^\amalg}(\mathcal{C})^\otimes \times_{\mathcal{I}^\amalg} \mathcal{C}^\otimes$$

induced by the map of ∞-operads $\mathrm{Fin}_* \boxplus \mathrm{Triv}^\otimes \to \mathcal{CM}^\otimes$ picking the algebra and the underlying object of the module. By [Lurie 2017, Corollary 4.2.4.4] this has a left adjoint on every fiber, which is compatible with the pushforwards by [Lurie 2017, Corollary 4.2.4.8], and so by [Lurie 2017, Corollary 7.3.2.12] it has a relative left adjoint F which is an $\mathcal{I}^\amalg$-monoidal functor. Then the functor we want is the composite

$$\mathcal{C}^\otimes \xrightarrow{(S,\mathrm{id})} \mathrm{Alg}_{\mathrm{Fin}_*/\mathcal{I}^\amalg}(\mathcal{C})^\otimes \times_{\mathcal{I}^\amalg} \mathcal{C}^\otimes \xrightarrow{F} \mathrm{Alg}_{\mathcal{CM}^\otimes/\mathcal{I}^\amalg}(\mathcal{C})^\otimes.$$

This induces the desired functor on the lax limit, since applying $N_{\mathcal{I}^\amalg}$ preserve operadic adjunctions. □

Part II ∞-categories of global objects as partially lax limits

In this second part of the paper we prove that various ∞-categories of global objects admit a description using (partially lax) limits. In Theorem 6.17, we show that the ∞-category of global spaces is equivalent to the partially lax limit of the functor sending a compact Lie group G to the ∞-category of G-spaces. Our main result is Theorem 11.10, which describes the ∞-category of global spectra as a partially lax limit of G-spectra where G runs over all compact Lie groups G. Finally, the techniques employed in the previous cases allow us to prove that for any Lie group G, the ∞-category of proper G-spectra is a limit of H-spectra for H running over all compact subgroups of G. The precise statement can be found in Theorem 12.11.

Remark To not burden the notation even more, we have decided to state Theorems 6.17 and 11.10 for the family of all compact Lie groups. However, the proofs hold verbatim for any family of compact Lie groups which is closed under isomorphisms, finite products, passage to subgroups and passage to quotients (ie any multiplicative global family in the language of [Schwede 2018]). If the family is not closed under finite products, then the equivalences of the two theorems still hold without symmetric monoidal structures. This is due to the fact that the model structure constructed in [Schwede 2018] is only shown to be symmetric monoidal for a multiplicative global family. We note that our result in fact allows us to define a symmetric monoidal structure on global spectra with respect to any global family, as a partially lax limit of symmetric monoidal categories is automatically symmetric monoidal.

6 Global spaces as a partially lax limit

In this section we show that the ∞-category of global spaces is equivalent to a certain partially lax limit of the functor which sends a group G to the ∞-category of G-spaces $\mathcal{S}_G$. This is an unstable version of our main result, and serves as a warm up for the considerable more details involved in that proof. We start off by recalling a few relevant definitions.

Definition 6.1 The *global category* Glo is the ∞-category associated to the topological category whose objects are compact Lie groups and whose mapping spaces are given by

$$\mathrm{Map}_{\mathrm{Glo}}(H,G) := |\mathrm{Hom}(H,G)/\!\!/G|,$$

the geometric realization of the action groupoid of G acting on the space of continuous group homomorphisms $\mathrm{Hom}(H,G)$ by conjugation. Composition is induced by the composition of group homomorphisms.

We define Orb and $\mathrm{Glo}^{\mathrm{sur}}$ to be the wide subcategory of Glo whose hom-spaces are given by those path-components of $\mathrm{Map}_{\mathrm{Glo}}(H,G)$ spanned by the injective and surjective group homomorphisms respectively. For later applications it will be convenient to mark all the edges in the full subcategory $\mathrm{Orb} \subseteq \mathrm{Glo}$; we denote this marking by $\mathrm{Glo}^{\dagger}$. Finally, we let Rep denote the homotopy category of Glo, that is, the category whose objects are compact Lie groups and whose morphisms are given by conjugacy classes of continuous group homomorphisms.

Remark 6.2 The definition of Glo agrees with the definition given in [Gepner and Henriques 2007, Section 4] restricted to compact Lie groups, up to one difference. We apply thin geometric realization to the action groupoids to obtain a topologically enriched category, while the original definition uses fat geometric realization. Up to a technical condition, the two conventions define Dwyer–Kan equivalent topological categories. See [Körschgen 2018, Remark 3.10] for a more detailed discussion. Note as well that [Gepner and Henriques 2007] uses the name Orb for both Glo and what we call Orb.

Key to the main properties of Glo is the following description of the mapping spaces.

Proposition 6.3 *Let G and H be two compact Lie groups. Then*

$$\mathrm{Hom}(H,G) \simeq \coprod_{[\alpha]\in \mathrm{Rep}(H,G)} \alpha G \quad \textit{and} \quad \mathrm{Glo}(H,G) \simeq \coprod_{[\alpha]\in \mathrm{Rep}(H,G)} BC(\alpha),$$

where αG denotes the orbit of α under the G-conjugation action, and $C(\alpha)$ denotes the centralizer of the image of α.

Proof See [Körschgen 2018, Propositions 2.4 and 2.5] for a proof of the first and second statement, respectively. □

Proposition 6.4 *Let $f : H \to G$ be a map in* Glo*. The induced map $f_* \colon \mathrm{Glo}(K,H) \to \mathrm{Glo}(K,G)$ on mapping spaces corresponds under the equivalences of Proposition 6.3 to the composite of the map*

$$\coprod_{[\alpha]\in \mathrm{Rep}(H,G)} Bf \colon \coprod_{[\alpha]\in \mathrm{Rep}(K,H)} BC(\alpha) \to \coprod_{[\alpha]\in \mathrm{Rep}(K,H)} BC(f\alpha)$$

with the map

$$\coprod_{[\alpha]\in \mathrm{Rep}(K,H)} BC(f\alpha) \to \coprod_{[\beta]\in \mathrm{Rep}(K,G)} BC(\beta)$$

which is the identity on individual path-components and acts on π_0 by $f_* \colon \mathrm{Rep}(K,H) \to \mathrm{Rep}(K,G)$.

Proof The statement on π_0 follows from the fact that Rep is the homotopy category of Glo. Therefore, it suffices to restrict to one path component, and analyze the effect of f. The relationship $f_*(c_h \alpha) = c_{f(h)} f\alpha$ implies that f_* acts as f when restricted to a map $\alpha H \to f\alpha G$. This implies that the induced map $BC(\alpha) \to BC(f\alpha)$ equals Bf. □

Definition 6.5 The ∞-category of *global spaces* $\mathcal{S}_{\mathrm{gl}}$ is the category of functors from $\mathrm{Glo}^{\mathrm{op}}$ to $\mathcal{S}$. This admits a symmetric monoidal structure by pointwise product. This is equivalent to the symmetric monoidal category $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}\text{-}\mathcal{S}$.

Remark 6.6 Schwede [2020] proves that the underlying ∞-category of orthogonal spaces equipped with the positive global model structure of [Schwede 2018, Proposition 1.2.23] is equivalent to presheafs on a topologically enriched category $\boldsymbol{O}_{\mathrm{gl}}$. Furthermore, in [Körschgen 2018] it is shown that $\boldsymbol{O}_{\mathrm{gl}}$ is Dwyer–Kan equivalent to Glo. Therefore the two models of global spaces define the same ∞-category. In fact, the two ∞-categories are symmetric monoidal equivalent since they are both endowed with the cartesian monoidal structure; see [Schwede 2018, Theorem 1.3.2].

Before stating and proving the main result of this section, we need some preparation. In the following we fix an ∞-category $\mathscr{C}$ with an orthogonal factorization system $(\mathscr{C}^L, \mathscr{C}^R)$. For a detailed discussion and a definition of orthogonal factorization systems on ∞-categories, the reader may consult [Lurie 2009, Section 5.2.8]. We write $\mathscr{C}^L$ for the left class of maps and $\mathscr{C}^R$ for the right class. We will denote edges in $\mathscr{C}^L$ by $\twoheadrightarrow$ and edges in $\mathscr{C}^R$ by $\rightarrowtail$.

Proposition 6.7 *Let $\mathscr{C}$ be an ∞-category equipped with an orthogonal factorization $(\mathscr{C}^L, \mathscr{C}^R)$. Write $\mathrm{Ar}_R(\mathscr{C})$ for the full subcategory of the arrow category of $\mathscr{C}$ spanned by the edges in $\mathscr{C}^R$. Then the target projection $t\colon \mathrm{Ar}_R(\mathscr{C}) \to \mathscr{C}$ is a cocartesian fibration. Furthermore an edge in $\mathrm{Ar}_R(\mathscr{C})$ is t-cocartesian if and only if it is of the form*

(6.7.1)
$$\begin{tikzcd} X \arrow[r, two heads] \arrow[d, tail] & Y \arrow[d, tail] \\ X' \arrow[r] & Y' \end{tikzcd}$$

Proof Consider an edge in $\mathrm{Ar}_R(\mathscr{C})$:
$$\begin{tikzcd} X \arrow[r] \arrow[d, tail] & Y \arrow[d, tail] \\ X' \arrow[r] & Y' \end{tikzcd}$$

This is cocartesian if and only if, given a 2-simplex in $\mathscr{C}$ and a (2,0)-horn in $\mathrm{Ar}_R(\mathscr{C})$, there is a contractible choice of extensions. This corresponds to showing that given a diagram in $\mathscr{C}$
$$\begin{tikzcd} X \arrow[r] \arrow[d, tail] \arrow[rr, bend left] & Y \arrow[d, tail] \arrow[dr] & Z \arrow[d, tail] \\ X' \arrow[r] & Y' \arrow[r] & Z' \end{tikzcd}$$

its extensions to a 2-simplex in $\mathrm{Ar}_R(\mathscr{C})$ form a contractible space. However, completing this diagram is equivalent to supplying an edge $Y \to Z$ which makes the diagram below commute:
$$\begin{tikzcd} X \arrow[r] \arrow[d] & Z \arrow[d, tail] \\ Y \arrow[r] \arrow[ur, dotted] & Z' \end{tikzcd}$$

There is a contractible choice of such factorizations if and only if $X \to Y$ is in $\mathscr{C}^L$. This shows that an edge is t-cocartesian if and only if it is of the form of equation (6.7.1). Next, fix an edge in $\mathscr{C}$ and a lift of its source in $\mathrm{Ar}_R(\mathscr{C})$. This corresponds to a diagram

$$\begin{tikzcd} X \arrow[d, tail] \arrow[dr] & \\ X' \arrow[r] & Y' \end{tikzcd}$$

Factorizing the composite $X \to Y'$ extends this to an edge

$$\begin{tikzcd} X \arrow[r, two heads] \arrow[d, tail] \arrow[dr] & Y \arrow[d, tail] \\ X' \arrow[r] & Y' \end{tikzcd}$$

in $\mathrm{Ar}_R(\mathscr{C})$, which is t-cocartesian. $\square$

We record the following fact for later reference.

Lemma 6.8 *The constant functor $s_0\colon \mathscr{C} \to \mathrm{Ar}_R(\mathscr{C})$ is a fully faithful left adjoint to the source functor $s\colon \mathrm{Ar}_R(\mathscr{C}) \to \mathscr{C}$.*

Construction 6.9 Suppose we are in the setting of Proposition 6.7. Straightening the cocartesian fibration $t\colon \mathrm{Ar}_R(\mathscr{C}) \to \mathscr{C}$ gives a functor

$$\mathscr{C}^R_{/-}\colon \mathscr{C} \to \mathrm{Cat}_\infty.$$

To justify our notation let us unravel the effect of this functor. By definition, the evaluation of $\mathscr{C}^R_{/-}$ at an object $X \in \mathscr{C}$ is given by $\mathrm{Ar}_R(\mathscr{C})_X$; the fiber of t at X. By construction this is the full subcategory of $\mathscr{C}_{/X}$ on the objects $C \rightarrowtail X$ in $\mathscr{C}^R$. A priori an edge in this full subcategory is given by a diagram

$$\begin{tikzcd} X \arrow[r] \arrow[dr, tail] & X' \arrow[d, tail] \\ & Y \end{tikzcd}$$

However the edge $X \to X'$ is necessarily also in $\mathscr{C}^R$ by [Lurie 2009, Proposition 5.2.8.6(3)], and therefore $\mathrm{Ar}_R(\mathscr{C})_X$ is in fact equivalent to $\mathscr{C}^R_{/X}$. Next consider an edge $f\colon Y \to Y'$. Then the induced map $f_*\colon \mathscr{C}^R_{/Y} \to \mathscr{C}^R_{/Y'}$ sends an object $X \rightarrowtail Y$ to an object $X' \rightarrowtail Y'$ such that the following diagram commutes:

$$\begin{tikzcd} X \arrow[r, two heads] \arrow[d, tail] & X' \arrow[d, tail] \\ Y \arrow[r, "f"] & Y' \end{tikzcd}$$

In particular, if $f \in \mathscr{C}^R$ this is nothing but the standard functoriality of the slices $\mathscr{C}^R_{/-}$. Therefore the functor $\mathscr{C}^R_{/-}\colon \mathscr{C} \to \mathrm{Cat}_\infty$ extends the functoriality of the slices of $\mathscr{C}^R$ to all of $\mathscr{C}$.

Proposition 6.10 *Let $\mathscr{C}$ be an ∞-category equipped with a factorization system $(\mathscr{C}^L, \mathscr{C}^R)$. The partially lax colimit of $(-)^{\mathrm{op}} \circ \mathscr{C}^R_{/-} \colon \mathscr{C} \to \mathrm{Cat}_\infty$ with respect to the marking $\mathscr{C}^R \subset \mathscr{C}$ is equivalent to $\mathscr{C}^{\mathrm{op}}$.*

Proof Recall that the partially lax colimit of a functor $F \colon \mathscr{C} \to \mathrm{Cat}_\infty$ is the localization of $\mathrm{Un}^{\mathrm{ct}}(F)$ at the cartesian edges which live above marked edges; see Theorem 4.9(b). In the case $F = (-)^{\mathrm{op}} \circ \mathscr{C}^R_{/-}$, we observe that $\mathrm{Un}^{\mathrm{ct}}(F) \simeq \mathrm{Un}^{\mathrm{co}}(\mathscr{C}^R_{/-})^{\mathrm{op}}$ and so we conclude that the partially lax colimit of F is equal to the opposite of $\mathrm{Ar}_R(\mathscr{C})$ localized at the edges of the form

$$\begin{tikzcd} X \arrow[r, two heads] \arrow[d, tail] & X' \arrow[d, tail] \\ Y \arrow[r, tail] & Y' \end{tikzcd}$$

However, note that because edges in $\mathscr{C}^R$ are left cancellable, $X \to X'$ is not only in $\mathscr{C}^L$ but also in $\mathscr{C}^R$. Therefore $X \to X'$ is in fact an equivalence. We will write M for this collection of edges. We claim that localizing at the edges of M is equivalent to localizing at the larger class of edges M' of the form

$$\begin{tikzcd} X \arrow[r, "\sim"] \arrow[d, tail] & X' \arrow[d, tail] \\ Y \arrow[r] & Y' \end{tikzcd}$$

where we do not impose any conditions on the edge $Y \to Y'$. To see this note that such an edge in M' fits into the following diagram:

$$\begin{tikzcd} X \arrow[r, "\sim"] \arrow[d, "\sim"'] & X' \arrow[r, equal] \arrow[d, tail] & X' \arrow[d, tail] \\ X' \arrow[r, tail] \arrow[rr, tail, bend right] & Y \arrow[r] & Y' \end{tikzcd}$$

Both the first edge and the composite are in M, and so therefore M' is contained in the two-out-of-three closure of M. So it is enough to calculate the localization of $\mathrm{Ar}_R(\mathscr{C})$ at M'. Note that the source functor $s \colon \mathrm{Ar}_R(\mathscr{C}) \to \mathscr{C}$ sends an edge to an equivalence if and only if it is in M'. Then Lemma 6.8 implies that $\mathscr{C}$ is a Bousfield colocalization of $\mathrm{Ar}_R(\mathscr{C})$ at M'. So we conclude that the partially lax colimit of ${}^{\mathrm{op}} \circ \mathscr{C}^R_{/-}$ is equivalent to $\mathscr{C}^{\mathrm{op}}$, finishing the proof. □

Example 6.11 There are two extreme cases of the previous result. If $\mathscr{C}^R = \mathscr{C}$ and $\mathscr{C}^L = \iota\mathscr{C}$, then

$$\mathrm{colim}((\mathscr{C}_{/-})^{\mathrm{op}} \colon \mathscr{C} \to \mathrm{Cat}_\infty) \cong \mathscr{C}^{\mathrm{op}}.$$

If $\mathscr{C}^R = \iota\mathscr{C}$ and $\mathscr{C}^L = \mathscr{C}$, then

$$\mathrm{laxcolim}(\iota\mathscr{C}_{/-} \colon \mathscr{C} \to \mathrm{Cat}_\infty) \cong \mathscr{C}^{\mathrm{op}}.$$

Now that we have introduced the main tools we need, we can build our functor and compute its partially lax limit. This relies on two important observations. The first key insight is the following, which was first stated in [Gepner and Henriques 2007] and originally proven as [Rezk 2014, Example 3.5.1].

Lemma 6.12 *For all compact Lie groups G, the assignment $G/K \mapsto (K \hookrightarrow G)$ defines an equivalence $\boldsymbol{O}_G \simeq \mathrm{Orb}_{/G}$.*

Proof Observe that the spaces $\boldsymbol{O}_G(G/H, G/K)$ are homeomorphic to the space $\{g \in G \mid c_g(H) \subseteq K\}/K$. The latter space is equivalent to the homotopy orbits $\{g \in G \mid c_g(H) \subseteq K\}_{\mathrm{h}K}$ as the K-space is free; see for example [Körschgen 2018, Theorem A.7]. Therefore we can define a functor $F' \colon \boldsymbol{O}_G \to \mathrm{Glo}$, which sends G/H to H, and on mapping spaces acts as homotopy orbits of the K-equivariant inclusion

$$\{g \in G \mid c_g(H) \subseteq K\} \to \mathrm{hom}(H, K), \quad g \mapsto [c_g \colon H \to K].$$

Note that the ∞-category $\boldsymbol{O}_G$ has a final object G/G, and therefore F' induces a functor $\boldsymbol{O}_G \to \mathrm{Glo}_{/G}$, which in fact factors through $\mathrm{Orb}_{/G}$. We claim that the induced functor $F \colon \boldsymbol{O}_G \to \mathrm{Orb}_{/G}$ is an equivalence of ∞-categories. First note that F is clearly essentially surjective. To deduce that the functor is fully faithful pick two objects G/H and G/K, which we identify with inclusions $i \colon H \hookrightarrow G$ and $j \colon K \hookrightarrow G$. Recall that the mapping space between G/H and G/K is empty if and only if H is not subconjugate to K. In this case the mapping space in Orb/G between i and j is also empty. Now suppose that this is not the case. Consider the square

$$\begin{tikzcd}
\{g \in G \mid c_g(H) \subseteq K\}_{\mathrm{h}K} \arrow[r] \arrow[d] & \mathrm{Hom}(H, K)_{\mathrm{h}K} \arrow[d] \\
* \arrow[r] & \mathrm{Hom}(H, G)_{\mathrm{h}G}
\end{tikzcd}$$

To prove F is fully faithful it suffices to prove that this square is homotopy cartesian. For every K-space X, $(G \times_K X)_{\mathrm{h}G} \simeq X_{\mathrm{h}K}$, so that the above square is equivalent to

$$\begin{tikzcd}
\big(G \times_K \{g \in G \mid c_g(H) \subseteq K\}\big)_{\mathrm{h}G} \arrow[r] \arrow[d] & (G \times_K \mathrm{Hom}(H, K))_{\mathrm{h}G} \arrow[d] \\
G_{\mathrm{h}G} \arrow[r] & \mathrm{Hom}(H, G)_{\mathrm{h}G}
\end{tikzcd}$$

Because taking homotopy orbits preserves homotopy pullback diagrams, it suffices to show that the square

$$\begin{tikzcd}
G \times_K \{g \in G \mid c_g(H) \subseteq K\} \arrow[r] \arrow[d] & G \times_K \mathrm{Hom}(H, K) \arrow[d] \\
G \arrow[r] & \mathrm{Hom}(H, G)
\end{tikzcd}$$

is homotopy cartesian. In fact it is easily shown to be a pullback square of topological spaces, and the bottom horizontal arrow is a Serre fibration. To see this we note that the map $G \to \mathrm{Hom}(H, G)$ factors through one component of the decomposition of Proposition 6.3, and therefore is equivalent to the quotient map $G \to G/C(H)$, which is a fibration by [Körschgen 2018, Theorem A.9]. □

The second insight is the following, which was also observed in [Rezk 2014].

Proposition 6.13 *The subcategories* $\mathrm{Glo}^{\mathrm{sur}}$ *and* Orb *are the left and right classes, respectively, of an orthogonal factorization system on* Glo.

Proof We will apply [Lurie 2009, Proposition 5.2.8.17] to the subcategories $\mathrm{Glo}^{\mathrm{sur}}$ and Orb. Clearly these subcategories contain all the equivalences and are closed under equivalences in Ar(Glo). Therefore it suffices to prove that given a diagram

$$\begin{tikzcd} H \arrow[r] \arrow[d, "f"'] & J \arrow[d, hook, "g"] \\ G \arrow[r] \arrow[ur, dotted] & K \end{tikzcd}$$

the space of dotted diagonal fillers is contractible. As noted in [Lurie 2009, Remark 5.2.8.3], this is equivalent to the map

$$\mathrm{Map}_{\mathrm{Glo}_{H/}}(H \xrightarrow{f} G, H \to J) \xrightarrow{g} \mathrm{Map}_{\mathrm{Glo}_{H/}}(H \xrightarrow{f} G, H \to K)$$

being a weak homotopy equivalence for every lift of g to a map in $\mathrm{Glo}_{H/}$ from $H \to J$ to $H \to K$. Proposition 6.4 shows that when f is surjective the map

$$\mathrm{Map}_{\mathrm{Glo}}(G, J) \xrightarrow{f^*} \mathrm{Map}_{\mathrm{Glo}}(H, J)$$

is an inclusion of path-components for every J.

Therefore the space $\mathrm{Map}_{\mathrm{Glo}_{H/}}(H \xrightarrow{f} G, H \to J)$, being the homotopy fiber of this map, is either empty or contractible. Translating back this reduces our task to simply proving the existence of a lift in the square above. This is a simple exercise in group theory. □

Remark 6.14 When we restrict to finite groups, Glo is equivalent to the full subcategory of $\mathcal{S}$ given by the connected 1-truncated spaces. In this case the orthogonal factorization system constructed above is a restriction of the standard mono/epi factorization system of any ∞-topos. However, in the generality of compact Lie groups, no such description applies.

We are finally ready to construct the functor.

Construction 6.15 Applying Construction 6.9 to the orthogonal factorization system $(\mathrm{Glo}^{\mathrm{sur}}, \mathrm{Orb})$ yields a functor $\mathrm{Orb}_{/-}\colon \mathrm{Glo} \to \mathrm{Cat}_\infty$. Postcomposing the opposite of this functor with $\mathrm{Fun}((-)^{\mathrm{op}}, \mathcal{S})\colon \mathrm{Cat}_\infty^{\mathrm{op}} \to \mathrm{Cat}_\infty$ gives the desired functor

$$\mathcal{S}_\bullet\colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_\infty.$$

Also note that $\mathcal{S}_\bullet$ clearly factors through product-preserving functors, and so enhances to a functor

$$\mathcal{S}_\bullet\colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_\infty^{\otimes},$$

where each category $(\mathrm{Orb}_{/G})^{\mathrm{op}}$-$\mathcal{S}$ is given the cartesian monoidal structure.

Lemma 6.12 and Elmendorf's theorem for G-spaces, see Example 3.40, imply that the value of $\mathcal{S}_\bullet$ at the object G is equivalent to the ∞-category of G-spaces $\mathcal{S}_G$. However, we owe the reader the following consistency check, which implies that the functor $\mathcal{S}_\bullet$ also has the expected functoriality.

Proposition 6.16 *Let $\alpha\colon H \to G$ be a continuous group homomorphism. Then the diagram*

$$\begin{tikzcd} \mathrm{Fun}((\mathrm{Orb}_{/G})^{\mathrm{op}}, \mathcal{S}) \arrow[r, "\simeq"] \arrow[d, "\mathcal{S}_\alpha"'] & \mathcal{S}_G \arrow[d, "\alpha^*"] \\ \mathrm{Fun}((\mathrm{Orb}_{/H})^{\mathrm{op}}, \mathcal{S}) \arrow[r, "\simeq"] & \mathcal{S}_H \end{tikzcd}$$

commutes. Here the horizontal equivalences are obtained by applying Lemma 6.12 and Example 3.40.

Proof It is enough to check that the analogous diagram, where the vertical maps are replaced with left adjoints, commutes. For this, let us denote by L_α and $\alpha_!$ the left adjoints of $\mathcal{S}_\alpha$ and α^*, respectively. Note that the inclusion $\iota_H\colon \mathrm{Orb}_{/H} \hookrightarrow \mathrm{Glo}_{/H}$ has a left adjoint L_H, which on objects sends $K \xrightarrow{\beta} H$ to $\beta(K) \hookrightarrow H$. By the universal property of the presheaf categories there exists a unique cocontinuous functor (the left Kan extension along ι_H)

$$(\iota_H)_!\colon \mathrm{Fun}((\mathrm{Orb}_{/H})^{\mathrm{op}}, \mathcal{S}) \to \mathrm{Fun}((\mathrm{Glo}_{/H})^{\mathrm{op}}, \mathcal{S}),$$

which agrees with ι_H on representables. In a similar fashion, we define functors $(L_G)_!$ and $(\alpha_*)_!$, where $\alpha_*\colon \mathrm{Glo}_{/H} \to \mathrm{Glo}_{/G}$ is postcomposition by α. We claim that the following diagram commutes:

$$\begin{tikzcd} \mathrm{Fun}((\mathrm{Orb}_{/H})^{\mathrm{op}}, \mathcal{S}) \arrow[r, hook, "(\iota_H)_!"] \arrow[d, "L_\alpha"'] & \mathrm{Fun}((\mathrm{Glo}_{/H})^{\mathrm{op}}, \mathcal{S}) \arrow[d, "(\alpha_*)_!"] \\ \mathrm{Fun}((\mathrm{Orb}_{/G})^{\mathrm{op}}, \mathcal{S}) & \mathrm{Fun}((\mathrm{Glo}_{/G})^{\mathrm{op}}, \mathcal{S}) \arrow[l, "(L_G)_!"'] \end{tikzcd}$$

This is easily seen by comparing the result on generators, and using that all the functors in the diagram commute with all colimits. Using this diagram we can reduce to a statement on the level of model categories. Namely, all three functors which make up the long way around in the diagram above can be modeled by left Quillen functors between enriched functor categories with the projective model structure. Indeed, the right adjoint of $(\iota_H)_!$ is given by restriction along ι_H, which is clearly a right Quillen functor. A similar argument also works for $(L_G)_!$ and $(\alpha_*)_!$. After precomposing and postcomposing with the equivalences

$$\mathcal{T}_H \simeq \mathrm{Fun}^{\mathrm{top}}((\mathrm{Orb}_{/H})^{\mathrm{op}}, \mathcal{T}) \quad \text{and} \quad \mathrm{Fun}^{\mathrm{top}}((\mathrm{Orb}_{/G})^{\mathrm{op}}, \mathcal{T}) \simeq \mathcal{T}_G$$

constructed in [Rezk 2014, Proposition 3.5.1], which agree with the equivalences constructed by [Gepner and Meier 2023] by inspection, we can apply the explicit description for $(L_G)_!$ and $(\iota_H)_!$ given in [Rezk 2014, Section 5.3], where $(L_G)_!$ is denoted by Π_G and $(\iota_H)_!$ by Δ_H, to deduce that the functor $L_\alpha\colon \mathcal{T}_H \to \mathcal{T}_G$ is equivalent to induction of H-spaces. $\square$

We have now constructed our functor. Therefore we are left to prove that the partially lax limit is given by the ∞-category of global spaces.

Theorem 6.17 *Let* $\mathrm{Glo}^\dagger$ *denote the marked* ∞*-category from Definition 6.1. Then the partially lax limit over* $(\mathrm{Glo}^\dagger)^{\mathrm{op}}$ *of the diagram from Construction 6.15*

$$\mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_\infty^\otimes, \quad G \mapsto \mathcal{S}_G,$$

is equivalent to the ∞*-category of global spaces, equipped with the cartesian monoidal structure.*

Proof Recall that $\mathcal{S}_G = \mathrm{Fun}(\boldsymbol{O}_G^{\mathrm{op}}, \mathcal{S})$ and that $\boldsymbol{O}_G \simeq \mathrm{Orb}_{/G}$. First we prove the result on underlying categories. Proposition 4.11 implies that it suffices to prove an equivalence between the partially lax colimit of $(\mathrm{Orb}_{/-})^{\mathrm{op}}$ and $\mathrm{Glo}^{\mathrm{op}}$. However, this follows from Proposition 6.10 applied to the factorization system $(\mathrm{Glo}^{\mathrm{sur}}, \mathrm{Orb})$ on Glo. Now we deduce the symmetric monoidal statement. First observe that the equivalence constructed before trivially lifts to a symmetric monoidal equivalence, where both sides are given the cartesian symmetric monoidal structure. Then note that the subcategory of Op_∞ spanned by the cartesian operads is closed under partially lax limits. This implies that $\mathcal{S}_{\mathrm{gl}}$ is equivalent to the partially lax limit of the diagram $\mathcal{S}^\bullet\colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_\infty^\otimes$, but now taken in symmetric monoidal ∞-categories. $\square$

7 ∞-categories of equivariant prespectra

In this section we define the ∞-categories of G-(pre)spectra for a Lie group G, and we introduce the ∞-category of global (pre)spectra. We will do this by first defining the relevant level model structures, which present the ∞-categories of prespectra objects, and then defining the stable model category as a Bousfield localization. This will then present the ∞-categories of spectra objects. The material in this section is classical, and largely well-known. Nevertheless we include the details of the model structures, mainly to emphasize that the level model structure on Sp_G^O is induced formally from the level model structure on $\mathcal{I}$-$G\mathcal{T}$. While not a deep statement, it is crucial to our proof strategy. In particular, this observation will allow us to interpret the construction of the level model structure ∞-categorically, as will be explained in this section.

Definition 7.1 Let $\mathcal{I}$ denote the topological category whose objects are finite-dimensional inner product spaces V, and morphism space $\mathcal{I}(V, W)$ is given by the space of linear isometric isomorphisms from V to W.

Definition 7.2 Let G be a Lie group (not necessarily compact). We write $\mathcal{I}\text{-}G\mathcal{T}$ for the enriched category of continuous functors from $\mathcal{I}$ into G-spaces, and call this the category of $\mathcal{I}$-G-spaces. When G is the trivial group, we simply write $\mathcal{I}\text{-}\mathcal{T}$ and refer to it as the category of $\mathcal{I}$-spaces.

Remark 7.3 As discussed in [Bohmann 2014, Section 5], the category of $\mathcal{I}$-G-spaces (as defined above) is equivalent as a topological category to the category of $\mathcal{I}_G$-spaces as defined by Mandell and May in [2002, Chapter II, Definition 2.3].

Remark 7.4 The category $\mathcal{I}\text{-}G\mathcal{T}$ has a symmetric monoidal structures given by enriched Day convolution; see [Mandell and May 2002, Chapter II, Proposition 3.7]. Given $X, Y \in \mathcal{I}\text{-}G\mathcal{T}$ we have the formula

$$(X \otimes Y)(V) := \int^{(W,W')\in\mathcal{I}\times\mathcal{I}} \mathcal{I}(W \oplus W', V) \times X(W) \times Y(W').$$

Remark 7.5 Given any $\mathcal{I}$-G-space X and an inner product space V, the value $X(V)$ admits a $G\times O(V)$-action. If V is given the structure of an H-representation $\rho\colon H \to O(V)$, then we can equip $X(V)$ with an H-action by restricting along

$$H \xrightarrow{\Delta} H \times H \xrightarrow{i\times\rho} G \times O(V).$$

We will always consider the value $X(V)$ with this H-action in the following.

Construction 7.6 (free $\mathcal{I}$-G-space) For every H-representation V, there is an evaluation functor

$$\mathrm{ev}_V\colon \mathcal{I}\text{-}G\mathcal{T} \to H\mathcal{T}, \quad X \mapsto X(V).$$

This functor admits a left adjoint $G \times_H \mathcal{I}_V$, given by the formula

$$G \times_H \mathcal{I}_V A = G \times_H (\mathcal{I}(V,-) \times A).$$

When $A = *$, we simply write $G \times_H \mathcal{I}_V$ and when $G = H$, we write $\mathcal{I}_V(-)$. By construction, the $\mathcal{I}$-G-space $G \times_H \mathcal{I}_V$ corepresents the functor $X \mapsto X(V)^H$.

For all compact subgroups H and K of G, all H-representations V and all K-representations W, there is an isomorphism of $\mathcal{I}$-G-spaces

$$(G \times_H \mathcal{I}_V) \otimes (G \times_K \mathcal{I}_W) \cong \Delta^*(G \times G \times_{H\times K} \mathcal{I}_{V\oplus W}), \tag{7.6.1}$$

where $\Delta\colon G \to G \times G$ is the diagonal embedding. This can be checked directly by applying the formula of the Day convolution product from Remark 7.4 and using that induction commutes with colimits.

We will now proceed to equip the category of $\mathcal{I}$-G-spaces with the level model structure. The following will be the weak equivalences, fibrations and cofibrations of this model structure.

Definition 7.7 Let G be a Lie group and let $f: X \to Y$ be a morphism in $\mathcal{I}$-$G\mathcal{T}$.

(a) We say f is a *level equivalence* if for any compact subgroup $H \leq G$ and any H-representation V, the map $f(V)^H: X(V)^H \to Y(V)^H$ is a weak homotopy equivalence of spaces.

(b) We say f is a *level fibration* if for any compact subgroup $H \leq G$ and any H-representation V, the map $f(V)^H: X(V)^H \to Y(V)^H$ is a Serre fibration.

(c) We say f is a *level cofibration* if for every $m \geq 0$, the map $f(\mathbb{R}^m): X(\mathbb{R}^m) \to Y(\mathbb{R}^m)$ is a Com-cofibration of $G \times O(m)$-spaces, see [Degrijse et al. 2023, Definition 1.1.2], and moreover the $O(m)$-action is free away from the image of $f(\mathbb{R}^m)$.

For all $m \geq 0$, we let $\mathscr{C}_G(m)$ denote the family of compact subgroups Γ of $G \times O(m)$ such that $\Gamma \cap (1 \times O(m))$ consists only of the neutral element. These are precisely the graph subgroups of a continuous homomorphism to $O(m)$ defined on some compact subgroup of G. The category of $G \times O(m)$-spaces admits a $\mathscr{C}_G(m)$-projective model structure by [Schwede 2018, Proposition B.7]. We have the following useful characterization of the level equivalences, cofibrations and fibrations.

Lemma 7.8 *Let G be a Lie group and let $f: X \to Y$ be a morphism in $\mathcal{I}$-$G\mathcal{T}$. The following are equivalent*:

(a) *The map $f: X \to Y$ is a level equivalence* (*resp. level fibration*).

(b) *The map $f(\mathbb{R}^m): X(\mathbb{R}^m) \to Y(\mathbb{R}^m)$ is a weak equivalence* (*resp. fibration*) *in the $\mathscr{C}_G(m)$-projective model structure for all $m \geq 0$.*

Furthermore, the following are equivalent:

(c) *The map $f: X \to Y$ is a level cofibration.*

(d) *The map $f(\mathbb{R}^m): X(\mathbb{R}^m) \to Y(\mathbb{R}^m)$ is a cofibration in the $\mathscr{C}_G(m)$-projective model structure for all $m \geq 0$.*

Proof Let $H \leq G$ be a compact subgroup and let V be an H-representation. Choose a linear isometric isomorphism $\varphi: V \cong \mathbb{R}^m$ and define a group homomorphism

$$\rho: G \to O(m), \quad g \mapsto \varphi \circ (g \cdot -) \circ \varphi^{-1}.$$

The homeomorphism $X(\varphi): X(V) \simeq X(\mathbb{R}^m)$ restricts to a homeomorphism

$$X(V)^H \simeq X(\mathbb{R}^m)^{\Gamma(\rho)},$$

where $\Gamma(\rho) = \{(h, \rho(h)) \in H \times O(m)\}$ by the definition of the H-action given in Remark 7.5. From this description, it is clear that (b) implies (a). Conversely, given $\Gamma \in \mathscr{C}_G(m)$, we can always find a continuous group homomorphism $\alpha: H \to O(m)$ for $H \leq G$ compact such that $\Gamma = \Gamma(\alpha)$. By definition of the H-action, we have $X(\mathbb{R}^m)^H = X(\mathbb{R}^m)^\Gamma$, showing that (a) implies (b). Finally, that (c) and (d) are equivalent follows from (the topological version of) [Stephan 2016, Proposition 2.16]. $\square$

Theorem 7.9 *Let G be a Lie group. The category $\mathscr{I}$-$G\mathscr{T}$ admits a cofibrantly generated and topological model structure in which the weak equivalences are the level equivalences, the fibrations are the level fibrations and the cofibrations are the level cofibrations. The set of generating cofibrations I_G and acyclic cofibrations J_G are given by*

$$\begin{aligned} I_G &= \{G \times_H \mathscr{I}_V \partial D^n \to G \times_H \mathscr{I}_V D^n \mid H \leq G, n \geq 0\}, \\ J_G &= \{G \times_H \mathscr{I}_V (D^n \times \{0\}) \to G \times_H \mathscr{I}_V (D^n \times [0,1]) \mid H \leq G, n \geq 0\}, \end{aligned}$$

where H runs over all compact subgroups of G and V runs over all H-representations. We call this the ***(proper) level model structure****.*

Proof We observe that the category $\mathscr{I}$-$G\mathscr{T}$ is equivalent to $\prod_{m \geq 0}(G \times O(m))\mathscr{T}$. We can endow this latter category with the product of the $\mathscr{C}_G(m)$-projective model structures on $G \times O(m)$-spaces. By Lemma 7.8, the induced model structure on $\mathscr{I}$-$G\mathscr{T}$ has weak equivalences, fibrations and cofibrations as in the theorem. Also we note that the right lifting property against the sets I_G and J_G detect the level fibrations and level acyclic fibrations respectively, by the adjunction isomorphism

$$\mathrm{Hom}_{\mathscr{I}\text{-}G\mathscr{T}}(G \times_H \mathscr{I}_V A, X) \simeq \mathrm{Hom}_{\mathscr{T}}(A, X(V)^H)$$

for A a nonequivariant space. Finally, we observe that resulting model structure is again topological by [Schwede 2018, Proposition B.5]. □

As discussed in [Degrijse et al. 2023, Proposition 1.1.6], a continuous homomorphism $\alpha\colon K \to G$ between Lie groups gives rise to adjoint functors between the associated category of equivariant spaces

$$\begin{tikzcd} G\mathscr{T} \arrow[r, "\alpha^*"] & K\mathscr{T} \arrow[l, bend right, "G\times_\alpha -"'] \arrow[l, bend left, "\mathrm{Map}^\alpha(G,-)"] \end{tikzcd}$$

which by levelwise application gives rise to an adjoint triple

$$\begin{tikzcd} \mathscr{I}\text{-}G\mathscr{T} \arrow[r, "\alpha^*"] & \mathscr{I}\text{-}K\mathscr{T}. \arrow[l, bend right, "G\times_\alpha -"'] \arrow[l, bend left, "\mathrm{Map}^\alpha(G,-)"] \end{tikzcd}$$

Proposition 7.10 *Let $\alpha\colon K \to G$ be a continuous group homomorphism between Lie groups.*

(a) *Then α^* preserves level fibrations and level equivalences. Thus the adjoint pair $(G \times_\alpha -, \alpha^*)$ is Quillen.*

(b) *If α has closed image and compact kernel, then the adjoint pair $(\alpha^*, \mathrm{Map}^\alpha(G, -))$ is also Quillen with respect to the level model structure.*

Proof Part (a) follows from [Degrijse et al. 2023, Proposition 1.1.6(ii)]. Suppose that α has closed image and compact kernel and note that by (a), it suffices to check that α^* preserves level cofibrations. We start by noting that the image of $\alpha \times O(m)$ is closed in $G \times O(m)$ since the image of α is closed in G. Moreover, the kernel of $\alpha \times O(m)$ is $\ker(\alpha) \times 1$, which is compact by hypothesis. So restriction along $\alpha \times O(m)$ takes Com-cofibrations of $G\times O(m)$-spaces to Com-cofibrations of $K\times O(m)$-spaces by [Degrijse et al. 2023, Proposition 1.1.6(iii)]. Now let $i\colon A \to B$ be a level cofibration of $\mathscr{I}$-G-spaces so that $i(\mathbb{R}^m)$ is a Com-cofibration of $G\times O(m)$-spaces. By the previous discussion, $\alpha^*(i(\mathbb{R}^m))$ is a Com-cofibration of $K\times O(m)$-spaces. Moreover, the $O(m)$-action is unchanged, so it still acts freely off the image of $\alpha^* i$. This shows that α^* preserves cofibrations as required. □

Proposition 7.11 *The level model structures on $\mathscr{I}$-$G\mathscr{T}$ is symmetric monoidal with cofibrant unit object.*

Proof Let us show that the pushout-product axiom holds. By a standard reduction [Hovey 1999, 4.2.5], it suffices to check that the pushout product $f \,\square\, g$ is

(i) a cofibration if f and g belong to the set of generating cofibrations,

(ii) an acyclic cofibration if furthermore f or g is a generating acyclic cofibration.

In this case we may assume $f = G \times_H \mathscr{I}_V f'$ and $g = G \times_K \mathscr{I}_W g'$ and so

$$f \,\square\, g = \Delta^*(G \times G \times_{H\times K} \mathscr{I}_{V\oplus W} f' \,\square\, g')$$

by equation (7.6.1). Since $\mathscr{T}$ is a symmetric monoidal model category, the pushout product $f' \square g'$ satisfies conditions (i) and (ii) above. By Proposition 7.10 we see that the functors

$$\Delta^*\colon \mathscr{I}\text{-}(G \times G)\mathscr{T} \to \mathscr{I}\text{-}G\mathscr{T}$$

are left Quillen. Moreover, it is clear from the definition of the model structures that the functor $\mathrm{ev}_{V\oplus W}\colon \mathscr{I}\text{-}(G\times G)\mathscr{T} \to (H\times K)\mathscr{T}$ is right Quillen, and therefore $(G\times G)\times_{H\times K} \mathscr{I}_{V\oplus W}$ is left Quillen. From these observations it follows that the pushout-product axiom holds for $\mathscr{I}$-$G\mathscr{T}$ too. Finally, the unit axiom holds since the unit object $* = G \times_G \mathscr{I}_0$ is cofibrant. □

In Section 2.3 we discussed how to induce a model structure on pointed objects. We will apply these results to the category $\mathscr{I}$-$G\mathscr{T}$ with the level model structure. Note first that the category of pointed objects in $\mathscr{I}$-$G\mathscr{T}$ is equivalent to $\mathscr{I}$-$G\mathscr{T}_*$, the category of continuous functors from $\mathscr{I}$ to $G\mathscr{T}_*$, the category of based G-spaces.

Proposition 7.12 *Let G be a Lie group. The category $\mathscr{I}$-$G\mathscr{T}_*$ admits a **proper level model structure** in which the weak equivalences, fibrations and cofibrations are detected by the forgetful functor $\mathscr{I}$-$G\mathscr{T}_* \to \mathscr{I}$-$G\mathscr{T}$. This model structure is topological, cofibrantly generated by the sets $(I_G)_+$ and $(J_G)_+$, symmetric monoidal, and the unit object is cofibrant. Moreover, there exists a symmetric monoidal equivalence of ∞-categories*

$$\mathscr{I}\text{-}G\mathscr{T}_*[W_{\mathrm{lvl}}^{-1}] \simeq (\mathscr{I}\text{-}G\mathscr{T}[W_{\mathrm{lvl}}^{-1}])_*.$$

Proof The first part follows from the discussion in Section 2.3 and [Schwede 2018, Proposition B.5]. For the final claim apply Proposition 2.3 together with the fact that $\mathcal{I}\text{-}G\mathcal{T}[W_{\mathrm{lvl}}^{-1}]$ is presentable by Theorem 8.9. □

We now change gears and consider the global analogue of the previous discussion. Recall that for any G-representation V and $\mathcal{I}$-space X, the value $X(V)$ admits a natural G-action by restricting along the canonical morphism $G \to O(V)$; see Remark 7.5.

Definition 7.13 Let $f : X \to Y$ be a morphism in $\mathcal{I}\text{-}\mathcal{T}$.

(a) We say f is a *faithful level equivalence* if for every compact Lie group G and every faithful G-representation V, the map $f(V) \colon X(V) \to Y(V)$ is a G-weak equivalence: for all closed subgroups $H \le G$, the induced map $f(V)^H : X(V)^H \to Y(V)^H$ is a weak homotopy equivalence of spaces.

(b) We say f is a *faithful level fibration* if for every compact Lie group G and every faithful G-representation V, the map $f(V) \colon X(V) \to Y(V)$ is a fibration in the projective model structure of G-spaces.

The following result is a reformulation of [Schwede 2018, Lemmas 1.2.7, 1.2.8] in our context.

Lemma 7.14 *Let $f : X \to Y$ be a morphism in $\mathcal{I}\text{-}\mathcal{T}$. Then the following are equivalent*:

(a) *The map $f(V) \colon X(V)^G \to Y(V)^G$ is a weak homotopy equivalence* (*resp. Serre fibration*) *for every compact Lie group G and every G-representation V.*

(b) *The map $f : X \to Y$ is a faithful level equivalence* (*resp. faithful level fibration*).

(c) *The map $f(\mathbb{R}^m) \colon X(\mathbb{R}^m) \to Y(\mathbb{R}^m)$ is an $O(m)$-weak equivalence* (*resp. $O(m)$-fibration*) *for every $m \ge 0$.*

Proof It is clear that (a) implies (b), which implies (c). Suppose that (c) holds and let V be a G-representation. As in the proof of Lemma 7.8 we can choose a linear isometric isomorphism $\varphi \colon V \simeq \mathbb{R}^m$ and define a group homomorphism $\rho \colon G \to O(m)$ such that

$$X(V)^G \simeq X(\mathbb{R}^m)^{\rho(G)},$$

showing that (c) implies (a). □

Construction 7.15 (semifree $\mathcal{I}$-space) For every G-representation V, there is an evaluation functor

$$\mathrm{ev}_{G,V} \colon \mathcal{I}\text{-}\mathcal{T} \to G\mathcal{T}, \quad X \mapsto X(V),$$

which admits a left adjoint $\mathcal{I}_{G,V}$ given by the formula $\mathcal{I}_{G,V}(A) = \mathcal{I}(V,-) \times_G A$. When $A = *$, we simply write $\mathcal{I}_{G,V}$. For all H-representations V and K-representations W, there is an isomorphism of $\mathcal{I}$-G-spaces

(7.15.1) $$\mathcal{I}_{H,V} \otimes \mathcal{I}_{K,W} \cong \mathcal{I}_{H\times K,V\oplus W}.$$

One can check this using the formula in Remark 7.4 or by mimicking the proof of [Schwede 2018, Example 1.3.3].

The next result is an analogue of [Schwede 2018, Proposition 1.2.10], adapted to our context.

Theorem 7.16 *The category $\mathcal{I}$-$\mathcal{T}$ admits a topological, cofibrantly generated model structure in which the weak equivalences are the faithful level equivalences $W_{\text{f-lvl}}$ and the fibrations are the faithful level fibrations. The set of generating cofibrations I and acyclic cofibrations J are given by*

$$I = \{\mathcal{I}_{G,V}(\partial D^n) \to \mathcal{I}_{G,V}(D^n)\} \quad \text{and} \quad J = \{\mathcal{I}_{G,V}(D^n \times \{0\}) \to \mathcal{I}_{G,V}(D^n \times [0,1])\},$$

where G runs over all compact Lie groups, V over all faithful G-representations, and $n \geq 0$. This is a symmetric monoidal model category with cofibrant unit object. We call this the ***faithful level model structure****.*

Proof We can identify $\mathcal{I}$-$\mathcal{T}$ with the category $\prod_{m\geq 0} O(m)\mathcal{T}$ and endow the latter category with the product of the standard model structures on $O(m)$-spaces. The induced model structure on $\mathcal{I}$-$\mathcal{T}$ has weak equivalences and fibrations as in the theorem by Lemma 7.14. We note that the right lifting property against the sets I and J detect the level fibrations and level acyclic fibrations respectively, by the adjunction isomorphism

$$\mathrm{Hom}_{\mathcal{I}\text{-}\mathcal{T}}(\mathcal{I}_{H,V} A, X) \simeq \mathrm{Hom}_{\mathcal{T}}(A, X(V)^H)$$

for A a nonequivariant space. Let us next show that the pushout-product axiom holds. As explained in the proof of Proposition 7.11, it suffices to check that the pushout product $f \,\square\, g$ is an (acyclic) cofibration if f and g belong to the set of generating (acyclic) cofibrations. In any case we have $f = \mathcal{I}_{G,V} f'$ and $g = \mathcal{I}_{H,W} g'$. But then $f \,\square\, g = \mathcal{I}_{G\times H, V\oplus W} f' \,\square\, g'$ by equation (7.15.1). Since $G\mathcal{T}$ is a symmetric monoidal model category, it suffices to check that the functor $\mathcal{I}_{G\times H,V\oplus W}$ is left Quillen. This is clear since $\mathrm{ev}_{G\times H,V\oplus W}$ is right Quillen by definition of the faithful level model structure. The pushout-product axiom then follows. Finally, the unit axiom holds since the unit object $* = \mathcal{I}_{e,0}$ is cofibrant and the model structure is topological by [Schwede 2018, Proposition B.5]. □

As before we obtain an induced model structured on pointed objects.

Proposition 7.17 *The category $\mathcal{I}$-$\mathcal{T}_*$ admits a faithful level model structure in which the weak equivalences, fibrations and cofibrations are detected by the forgetful functor $\mathcal{I}$-$\mathcal{T}_* \to \mathcal{I}$-$\mathcal{T}$. This model structure is topological, cofibrantly generated by the set I_+ and J_+, symmetric monoidal and the unit object is cofibrant. Finally, there exists a symmetric monoidal equivalence of ∞-categories*

$$\mathcal{I}\text{-}\mathcal{T}_*[W_{\text{f-lvl}}^{-1}] \simeq (\mathcal{I}\text{-}\mathcal{T}[W_{\text{f-lvl}}^{-1}])_*.$$

Proof The first two claims follows from the discussion in Section 2.3 and [Schwede 2018, Proposition B.5]. For the final claim apply Proposition 2.3, using the fact that $\mathcal{I}$-$\mathcal{T}[W_{\text{f-lvl}}^{-1}]$ is presentable. We will show this in Theorem 8.19. □

We now pass from pointed objects to prespectrum objects. Observe that the category of pointed $\mathcal{I}$-G-spaces has a commutative algebra object S_G given by the functor sending V to its one-point compactification S^V equipped with the trivial G-action. If we are thinking of the category of $\mathcal{I}$-spaces with the faithful level model structure, we will write S_{fgl} for S_e, to emphasize that the sphere should be thought of as evaluated on all faithful representations of all groups (fgl stands for faithful global).

Definition 7.18 Let G be a Lie group. Following [Mandell and May 2002, Chapter II Proposition 3.8], we define the topological category Sp_G^O of orthogonal G-spectra to be the category of S_G-modules in $\mathcal{I}$-$G\mathcal{T}_*$. These categories inherit induced model structures:

(a) The category of orthogonal G-spectra admits a (*proper*) *level model structure*, whose weak equivalences and fibrations are created by the forgetful functor $\mathrm{Sp}_G^O \to \mathcal{I}$-$G\mathcal{T}_*$, where the target is endowed with the level model structure. This is a cofibrantly generated, proper, topological model category; see the proof of [Degrijse et al. 2023, Theorem 1.2.22]. We also obtain that a set of generating cofibrations and acyclic cofibrations are given by the maps $S_G \otimes I_G$ and $S_G \otimes J_G$, where $S_G \otimes -$ denotes the left adjoint to the forgetful functor $\mathrm{Sp}_G^O \to \mathcal{I}$-$G\mathcal{T}_*$.

(b) The category of orthogonal spectra admits a *faithful level model structure*, whose weak equivalences and fibrations are created by the forgetful functor $\mathrm{Sp}^O \to \mathcal{I}$-$\mathcal{T}_*$, where the target is endowed with the faithful level model structure; see [Schwede 2018, Propositions 4.3.5]. From this result we obtain that the faithful level model structure is cofibrantly generated and topological, with a set of generating cofibrations and acyclic cofibrations given by $S_{\mathrm{fgl}} \otimes I$ and $S_{\mathrm{fgl}} \otimes J$, where $S_{\mathrm{fgl}} \otimes -$ denotes the left adjoint to the forgetful functor $\mathrm{Sp}^O \to \mathcal{I}$-$\mathcal{T}_*$.

Remark 7.19 By combining straightforward generalizations of [Mandell and May 2002, Theorem 4.3] and [Schwede 2018, Remark 3.1.8] to Lie groups, we conclude that Sp_G^O is equivalent to the category of orthogonal spectra defined in [Degrijse et al. 2023, Definition 1.1.9].

As discussed in [Mandell and May 2002, Chapter II Section 3], the category of orthogonal G-spectra admits a closed symmetric monoidal structure.

Proposition 7.20 *Let G be a Lie group.*

(a) *The level model structure on Sp_G^O is symmetric monoidal.*

(b) *The faithful level model structure on Sp^O is symmetric monoidal.*

Proof The proof that the pushout product axiom holds for Sp_G^O is similar to that given in Proposition 7.11 for $\mathcal{I}$-G-spaces. The explicit argument for cofibrations can be found in [Degrijse et al. 2023, Proposition 1.2.28(i)] and we note that a slight modification of that argument then also gives the statement for acyclic cofibrations. The argument that the faithful level model structure satisfies the pushout-product axiom is similar to that given in Theorem 7.16. The argument for cofibrations can also be found in [Schwede 2018, Proposition 4.3.23] and a slight modification of that argument also gives the statement for acyclic cofibrations. □

Definition 7.21 We define the ∞-category PSp_G of G*-prespectra* to be the symmetric monoidal ∞-category associated to the symmetric monoidal model category Sp_G^O with the level model structure. Similarly, we define the ∞-category $\mathrm{PSp}_{\mathrm{fgl}}$ of *faithful global prespectra* to be the symmetric monoidal ∞-category associated to the symmetric monoidal model category Sp^O with the faithful level model structure.

We have emphasized how the level model structures on Sp_G^O and Sp^O are induced by the level model structure on $\mathcal{I}\text{-}G\mathcal{T}_*$ and $\mathcal{I}\text{-}\mathcal{T}_*$, respectively, by taking modules. This allows us to reinterpret the passage to modules internally to ∞-categories.

Proposition 7.22 *There are symmetric monoidal equivalences*

$$\mathrm{PSp}_G \simeq \mathrm{Mod}_{S_G}(\mathcal{I}\text{-}G\mathcal{T}[W_{\mathrm{lvl}}^{-1}]_*) \quad \text{and} \quad \mathrm{PSp}_{\mathrm{fgl}} \simeq \mathrm{Mod}_{S_{\mathrm{fgl}}}(\mathcal{I}\text{-}\mathcal{T}[W_{\text{f-lvl}}^{-1}]_*).$$

Proof Apply Proposition 2.4. ◻

Finally, we pass from the level model structure to the stable model structure, which will present the categories of global and genuine G-spectra. Fix a complete G-universe $\mathcal{U}_G$ and write $s(\mathcal{U}_G)$ for the poset, under inclusion, of finite-dimensional G-subrepresentations of $\mathcal{U}_G$. The G-equivariant homotopy groups of an orthogonal G-spectrum X are given by

$$\pi_k^G(X) = \begin{cases} \mathrm{colim}_{V\in s(\mathcal{U}_G)}[S^{k+V}, X(V)]_*^G & \text{for } k \geq 0, \\ \mathrm{colim}_{V\in s(\mathcal{U}_G)}[S^{V}, X(\mathbb{R}^{-k}\oplus V)]_*^G & \text{for } k \leq 0, \end{cases}$$

where the connecting maps in the colimit system are induced by the structure maps, and $[-,-]_*^G$ means G-equivariant homotopy classes of based G-maps. Note that the same definition works even if X is an orthogonal spectrum, since the value $X(V)$ admits a G-action as discussed before Definition 7.13. Moreover, everything is functorial with respect to morphisms of orthogonal (G-)spectra. We finally note that the definition above a priori depends on a choice of complete G-universe. However, the functors associated to different complete G-universes are naturally isomorphic, and so the choice is immaterial.

Definition 7.23 Let G be a Lie group.

- A morphism $f: X \to Y$ of orthogonal G-spectra is a $\underline{\pi}_*$*-isomorphism* if $\pi_*^H(f): \pi_*^H(X) \to \pi_*^H(Y)$ is an isomorphism for all compact subgroups $H \leq G$. The $\underline{\pi}_*$-isomorphisms are part of a cofibrantly generated, topological, stable and symmetric monoidal model structure on the category of orthogonal G-spectra [Degrijse et al. 2023, Theorem 1.2.22], called the G*-stable model structure*.
- A morphism $f: X \to Y$ of orthogonal spectra is a *global equivalence* if $\pi_*^H(f): \pi_*^H(X) \to \pi_*^H(Y)$ is an isomorphism for all compact Lie groups H. The global equivalences are part of a cofibrantly generated, topological, proper, stable and symmetric monoidal model structure on the category of orthogonal spectra [Schwede 2018, Theorem 4.3.17, Proposition 4.3.24], called the *global model structure*.

Definition 7.24 We define the symmetric monoidal ∞-category Sp_G of G*-spectra* to be the underlying ∞-category of orthogonal G-spectra with the G-stable model structure. Similarly, we define the symmetric monoidal ∞-category $\mathrm{Sp}_{\mathrm{gl}}$ of *global spectra* to be the underlying ∞-category of orthogonal spectra with the global model structure.

We now make precise the observation that Sp_G and $\mathrm{Sp}_{\mathrm{gl}}$ are Bousfield localizations of PSp_G and $\mathrm{PSp}_{\mathrm{fgl}}$, respectively, at an explicit collection of weak equivalences. We begin with global spectra.

Construction 7.25 Given a compact Lie group G and a G-representation V, consider the adjoint pairs

$$\mathrm{Sp}^O \underset{S_{\mathrm{gl}}\otimes -}{\overset{\text{forget}}{\rightleftarrows}} \mathscr{I}\text{-}\mathscr{T}_* \underset{\mathscr{I}_{G,V}}{\overset{\mathrm{ev}_{G,V}}{\rightleftarrows}} G\mathscr{T}_*.$$

Following [Schwede 2018, Construction 4.1.23], we denote the composite $S_{\mathrm{gl}} \otimes \mathscr{I}_{G,V}$ by $F_{G,V}$. Note that the adjoint pairs above are Quillen with respect to the global level structure and so they yield corresponding adjoint pairs of underlying ∞-categories. As discussed before [Schwede 2018, Theorem 4.1.29], there are maps in Sp^O

$$\lambda_{G,V,W}\colon F_{G,V\oplus W} S^V \to F_{G,W} S^0$$

for all compact Lie groups G and G-representations V and W. We can view these maps in $\mathrm{PSp}_{\mathrm{gl}}$ since the domain and codomain of $\lambda_{G,V,W}$ are bifibrant. Consider the diagram

$$\begin{tikzcd}
G\mathscr{T}_*(S^0, X(W)) \arrow[r, "\tilde{\sigma}_{G,V,W}"] & G\mathscr{T}_*(S^V, X(V\oplus W)) \arrow[d, "\sim"] \\
\mathrm{Sp}^O(F_{G,W}S^0, X) \arrow[u, "\sim"] \arrow[r] & \mathrm{Sp}^O(F_{G,V\oplus W}S^V, X)
\end{tikzcd}$$

where the vertical maps are the adjunction isomorphisms and the top map is the adjoint structure map of X. The bottom map is equal to precomposition by $\lambda_{G,V,W}$. In particular, taking $X = F_{G,W}S^0$, we may define $\lambda_{G,V,W}$ as the image of the identity of $F_{G,W}S^0$ under the bottom map. Note also that $\lambda_{G,V,W}$ is equivalent to $F_{G,W}S^0 \otimes \lambda_{G,V,0}$, and that $\lambda_{G,V,0}$ is adjoint to the identity.

Remark 7.26 Both characterizations of $\lambda_{G,V,W}$ given above also uniquely specify the map on the level of ∞-categories.

Proposition 7.27 $\mathrm{Sp}_{\mathrm{gl}}$ *is a Bousfield localization of* $\mathrm{PSp}_{\mathrm{fgl}}$. *Furthermore, an object in* $\mathrm{PSp}_{\mathrm{fgl}}$ *lies in* $\mathrm{Sp}_{\mathrm{gl}}$ *if and only if it is local with respect to the morphisms* $\{\lambda_{G,V,W}\}$ *for all compact Lie groups G and G-representations V and W with W faithful.*

Proof Let Λ denote the set of maps $\lambda_{G,V,W}$ for G, V and W as in the proposition. We write $\mathrm{Sp}^O_{\mathrm{lvl}}$ and $\mathrm{Sp}^O_{\mathrm{gl}}$ for the category of orthogonal spectra endowed with the faithful level model structure and the global stable model structure, respectively. We will show that $\mathrm{Sp}^O_{\mathrm{gl}}$ is a left Bousfield localization (in the model categorical sense) of $\mathrm{Sp}^O_{\mathrm{lvl}}$ at the set Λ, that is, $L_\Lambda \mathrm{Sp}^O_{\mathrm{lvl}} = \mathrm{Sp}^O_{\mathrm{gl}}$. Because both can be checked on underlying homotopy categories, Bousfield localizations of model categories present Bousfield localizations of ∞-categories. Therefore the claim in the proposition will follow by passing to underlying ∞-categories. By definition $X \in \mathrm{Sp}^O_{\mathrm{lvl}}$ is Λ-local (and so fibrant in the Bousfield localization) if and only if X is fibrant in $\mathrm{Sp}^O_{\mathrm{lvl}}$ (which always holds in this case), and the canonical map of homotopy function complexes

$$\lambda^*_{G,V,W} \colon \mathrm{Map}(F_{G,W} S^0, X) \to \mathrm{Map}(F_{G,V\oplus W} S^V, X)$$

is an equivalence for all $\lambda_{G,V,W} \in \Lambda$. By adjunction this is equivalent to asking that $X(W)^G \to \Omega^V(X(V \oplus W))^G$ be an equivalence for all G, V and W as in the proposition. In other words, X is a global Ω-spectrum; see [Schwede 2018, Definition 4.3.8]. By [Schwede 2018, Theorem 4.3.17] these are precisely the fibrant objects $\mathrm{Sp}^O_{\mathrm{gl}}$. Since $L_\Lambda \mathrm{Sp}^O_{\mathrm{lvl}}$ and $\mathrm{Sp}^O_{\mathrm{gl}}$ have the same cofibrations and fibrant objects, the two model structures coincide by [Joyal 2008, Proposition E.1.10]. □

We repeat this analysis for Sp_G and PSp_G.

Construction 7.28 Let H be a compact subgroup of a Lie group G, and let V be an H-representation. We have a sequence of adjoint pairs

$$\mathrm{Sp}^O_G \underset{S_G\otimes -}{\overset{\text{forget}}{\rightleftarrows}} \mathcal{I}\text{-}G\mathcal{T}_* \underset{G_+\wedge_H \mathcal{I}_V}{\overset{\mathrm{ev}_V}{\rightleftarrows}} H\mathcal{T}_*,$$

which are Quillen with respect to the proper level model structure, and so they define adjoint pairs at the level of underlying ∞-categories. The composite $S_G \otimes (G_+ \wedge_H \mathcal{I}_V)$ will also be denoted by $G \ltimes_H F_V$ following [Degrijse et al. 2023, Example 1.1.15]. This notation is justified by the fact that $G \ltimes_H F_V$ is also equivalent to the induction of the H-prespectrum F_V as one can easily verify. For all pairs of H-representations V and W, there are maps in Sp^O_G

$$G \ltimes_H \lambda_{V,W} \colon G \ltimes_H F_{V\oplus W} S^V \to G \ltimes_H F_W,$$

see [Degrijse et al. 2023, equation 1.2.19]. We can view these maps in PSp_G as the domains and codomains are bifibrant. Similarly to before, $G \ltimes_H \lambda_{V,W}$ is determined by the property that the map

$$\mathrm{Sp}^O_G(G \ltimes_H F_W, X) \to \mathrm{Sp}^O_G(G \ltimes_H F_{V\oplus W} S^V, X),$$

defined so that the diagram

$$\begin{tikzcd}
H\mathcal{T}_*(S^0, X(W)) \arrow[r, "\mathrm{res}^G_H(\tilde{\sigma}_{V,W})"] & H\mathcal{T}_*(S^V, X(V \oplus W)) \arrow[d, "\sim"] \\
\mathrm{Sp}^O_G(G \ltimes_H F_W, X) \arrow[u, "\sim"] \arrow[r] & \mathrm{Sp}^O_G(G \ltimes_H F_{V\oplus W} S^V, X)
\end{tikzcd}$$

commutes, is equal to precomposition by $G \ltimes_H \lambda_{H,V,W}$. Also, $G \ltimes \lambda_{V,W}$ is equal to $G \ltimes_H F_W S^0 \otimes \lambda_{V,0}$ and $\lambda_{V,0}$ is adjoint to the identity on S^V.

Remark 7.29 Once again, the characterizations of $G \ltimes_H \lambda_{V,W}$ given above also uniquely specify the map on the level of ∞-categories.

Proposition 7.30 *Let G be a Lie group. Then Sp_G is a Bousfield localization of PSp_G. Furthermore, an object in PSp_G lies in Sp_G if and only if it is local with respect to the morphisms $\{G \ltimes_H \lambda_{V,W}\}$ for all compact subgroups $H \le G$ and H-representations V and W. Equivalently, $X \in \mathrm{PSp}_G$ lies in Sp_G if and only if for all compact subgroups $H \le G$, the object $\mathrm{res}^G_H X \in \mathrm{PSp}_H$ is local with respect to morphisms $\{\lambda_{V,W}\}$ for all H-representations V and W.*

Proof The proof is similar to that of Proposition 7.27 but now we use the characterization of fibrant objects in the proper stable model structure given in [Degrijse et al. 2023, Theorem 1.2.22(v)]. The second claim follows from the first one by adjunction. □

8 Models for ∞-categories of equivariant prespectra

In the previous section we introduced the ∞-categories of equivariant and global (pre)spectra, and exhibited the spectrum objects as local objects in the relevant category of prespectra with respect to an explicit class of weak equivalences. Furthermore, we observed that the construction of PSp_G admitted a reinterpretation internal to ∞-categories, by first passing to pointed objects in $\mathcal{I}\text{-}G\mathcal{T}[W_{\mathrm{lvl}}^{-1}]$ and then taking modules over S_G. Similarly, we observed that

$$\mathrm{PSp}_{\mathrm{fgl}} \simeq \mathrm{Mod}_{S_{\mathrm{fgl}}}(\mathcal{I}\text{-}\mathcal{T}[W_{\text{f-lvl}}^{-1}]_*).$$

Furthermore, these equivalences were symmetric monoidal.

However this is only part of the story, because the ∞-categories $\mathcal{I}\text{-}G\mathcal{T}[W_{\mathrm{lvl}}^{-1}]$ and $\mathcal{I}\text{-}\mathcal{T}[W_{\text{f-lvl}}^{-1}]$ are still too inexplicit for our arguments. Luckily we can give explicit models of these ∞-categories. Consider the case of $\mathcal{I}\text{-}G\mathcal{T}[W_{\mathrm{lvl}}^{-1}]$. By construction this ∞-category records the fixed-point spaces $X(V)^H$ for every (compact) subgroup H of G and every H-representation V of an $\mathcal{I}$-G-space X. By functoriality, these different fixed-point spaces are related by subconjugacy relationships in H and equivariant linear isometries in V. We will prove that the ∞-category $\mathcal{I}\text{-}G\mathcal{T}$ is in fact freely generated under these properties. More precisely, we will exhibit an equivalence

$$\mathcal{I}\text{-}G\mathcal{T}[W_{\mathrm{lvl}}^{-1}] \simeq \mathbf{OR}_G\text{-}\mathcal{S},$$

where the ∞-category $\mathbf{OR}_G$ indexes pairs (H, V), each one of which records one of the fixed-point spaces $X(V)^H$ of an $\mathcal{I}$-G-space X. Similarly, we will prove that

$$\mathcal{I}\text{-}\mathcal{T}[W_{\text{f-lvl}}^{-1}] \simeq \mathbf{OR}_{\mathrm{fgl}}\text{-}\mathcal{S},$$

where the ∞-category $\mathbf{OR}_{\mathrm{fgl}}$ indexes pairs (G, V), where G is a compact Lie group and V is a faithful G-representation.

In total we will obtain equivalences

$$\mathrm{PSp}_G \simeq \mathrm{Mod}_{S_G}(\mathbf{OR}_G\text{-}\mathscr{S}_*) \quad \text{and} \quad \mathrm{PSp}_{\mathrm{fgl}} \simeq \mathrm{Mod}_{S_{\mathrm{fgl}}}(\mathbf{OR}_{\mathrm{fgl}}\text{-}\mathscr{S}_*).$$

It will be in this guise that we will think of the ∞-category of G-prespectra and global prespectra for the remainder of the paper.

Finally, to make future constructions symmetric monoidal it will be important to understand how the symmetric monoidal structures transfer under the equivalences

$$\mathscr{I}\text{-}G\mathscr{T}[W_{\mathrm{lvl}}^{-1}] \simeq \mathbf{OR}_G\text{-}\mathscr{S} \quad \text{and} \quad \mathscr{I}\text{-}\mathscr{T}[W_{\text{f-lvl}}^{-1}] \simeq \mathbf{OR}_{\mathrm{fgl}}\text{-}\mathscr{S}.$$

We may immediately apply Theorem 3.37 to conclude that the monoidal structure on $\mathscr{I}\text{-}G\mathscr{T}[W_{\mathrm{lvl}}^{-1}]$ and $\mathscr{I}\text{-}\mathscr{T}[W_{\text{f-lvl}}^{-1}]$ are induced by Day convolution from the restricted promonoidal structure on $\mathbf{OR}_G$. We will make these promonoidal structures explicit.

To show that $\mathscr{I}\text{-}G\mathscr{T}[W_{\mathrm{lvl}}^{-1}]$ and $\mathscr{I}\text{-}\mathscr{T}[W_{\mathrm{lvl}}^{-1}]$ are equivalent to categories of copresheafs on an explicit set of generators, we will apply a version of Elmendorf's theorem; see Corollary 3.41. The application of this theorem to $\mathscr{I}\text{-}G\mathscr{T}[W_{\mathrm{lvl}}^{-1}]$ and $\mathscr{I}\text{-}\mathscr{T}[W_{\text{f-lvl}}^{-1}]$ has a similar flavor, but are logically distinct. Therefore we treat each case separately.

8.1 $\mathscr{I}$-G-spaces and $\mathbf{OR}_G$-spaces

We begin with $\mathscr{I}\text{-}G\mathscr{T}[W_{\mathrm{lvl}}^{-1}]$.

Remark 8.1 Let G be a Lie group and consider a map $\varphi\colon G/K \to G/H$ in the orbit category $\boldsymbol{O}_G$. Giving φ is equivalent to giving $gH \in (G/H)^K$, that is an element $gH \in G/H$ such that $c_g(K) = g^{-1}Kg \subseteq H$. When we need to emphasize this correspondence between gH and φ we will use subscripts φ_g and g_φ. Since $g_{\psi\circ\varphi}H = g_\varphi g_\psi H$, composition of maps corresponds to multiplication with reverse order.

Definition 8.2 For a Lie group G, the *proper G-orbit category* $\boldsymbol{O}_{G,\mathrm{pr}}$ is the full subcategory of $\boldsymbol{O}_G$ spanned by those cosets G/H with $H \leq G$ compact.

Let G be a Lie group and $H, K \leq G$ be compact subgroups. Given an H-representation V and a K-representation W, we can consider the space $G \times_H \mathscr{I}(V, W)$, where H acts on G by right translation, and on $\mathscr{I}(V, W)$ via $h.\varphi = \varphi h^{-1}$. Note that K acts diagonally on $G \times_H \mathscr{I}(V, W)$ via G and W. We have the following helpful criterion.

Lemma 8.3 *An element $[g, \varphi] \in G \times_H \mathscr{I}(V, W)$ is K-fixed if and only if $c_g(K) \subseteq H$ and $k.\varphi(v) = \varphi(c_g(k)v)$ for all $k \in K$ and $v \in V$.*

Proof An element $[g, \varphi] \in G \times_H \mathscr{I}(V, W)$ is K-fixed if and only if $[kg, k.\varphi] = [g, \varphi]$ for all $k \in K$. This means that there exists $h \in H$ such that $kg = gh$ and $k.\varphi = \varphi h$ for all $k \in K$. In other words g is such that $c_g(K) \subseteq H$ and φ is K-equivariant in the sense that $k.\varphi = \varphi c_g(k)$ for all $k \in K$. □

Lemma 8.4 *Let G be a Lie group and $H, K, L \leq G$ be compact subgroups. Let V be an H-representation, W a K-representation and U an L-representation. Then the map*

$$\circ\colon (G \times_K \mathscr{I}(W,U))^L \times (G \times_H \mathscr{I}(V,W))^K \to (G \times_H \mathscr{I}(V,U))^L$$

given by $([g',\psi],[g,\varphi]) \mapsto [g'g,\psi\varphi]$ is well-defined and continuous. Furthermore, upon varying the objects, the collection of maps so obtained is associative and unital.

Proof Let us first show that the map does not depend on the chosen representatives. For $h \in H$ and $k \in K$ we have $[g,\varphi] = [gh,\varphi h]$ and $[g',\psi] = [g'k,\psi k]$ so we ought to check that $[g'g,\psi\varphi] = [g'kgh,\psi k\varphi h]$. Using that $c_g(K) \subseteq H$ and φ is K-equivariant with respect to the c_g-twisted action, we can write

$$[g'kgh,\psi k\varphi h] = [g'g\,\underbrace{c_g(k)h}_{\in H},\psi k\varphi h] = [g'g,\psi k\varphi h(c_g(k)h)^{-1}] = [g'g,\psi k\varphi c_g(k^{-1})] = [g'g,\psi\varphi],$$

as required. We verify that $[g'g,\psi\varphi]$ is K-fixed using the criterion from Lemma 8.3. Using that $c_{g'}(L) \subseteq K$ and $c_g(K) \subseteq H$ we immediately see that $c_{g'g}(L) \subseteq H$. Using the twisted equivariance of ψ and φ we see that

$$l.\psi\varphi = \psi\overbrace{c_{g'}(l)}^{\in K}\varphi = \psi\varphi c_g(c_{g'}(l)) = \psi\varphi c_{g'g}(l) \quad \text{for all } l \in L.$$

Therefore $\psi\varphi$ is twisted equivariant and $[g'g,\psi\varphi]$ is indeed K-fixed. Finally, the map is associative, unital and continuous, since multiplication and composition maps are so. □

We now formally define the ∞-category $\mathbf{OR}_G$.

Definition 8.5 Let G be a Lie group. We define a topological category $\mathbf{OR}_G$ whose objects are pairs (H,V) of a compact subgroup $H \leq G$ and an H-representation V. The morphism spaces are given by

$$\mathbf{OR}_G((H,V),(K,W)) = (G \times_H \mathscr{I}(V,W))^K.$$

Composition is given by the maps

$$\circ\colon \mathbf{OR}_G((K,W),(L,U)) \times \mathbf{OR}_G((H,V),(K,W)) \to \mathbf{OR}_G((H,V),(L,U))$$

defined in Lemma 8.4. Note that there is a projection map

$$\mathbf{OR}_G((H,V),(K,W)) \to (G/H)^K = \boldsymbol{O}_{G,\mathrm{pr}}(G/K,G/H), \quad [g,\varphi] \mapsto [gH],$$

which extends to a functor $\pi_G\colon \mathbf{OR}_G \to \boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}}$.

Example 8.6 Let $G = e$ be the trivial group. Then the topological category $\mathbf{OR}_G$ is equivalent to $\mathscr{I}$.

Example 8.7 By definition, $\mathbf{OR}_G((H,V),(e,W)) = G \times_H \mathscr{I}(V,W)$, which is a space with an action of

$$\mathbf{OR}_G((e,W),(e,W)) = G \times O(W).$$

One can identify the functor $\mathbf{OR}_G((H,V),(e,-))\colon \mathscr{I} \to G\mathscr{T}$ with the free $\mathscr{I}$-G-space $G \times_H \mathscr{I}_V$.

Definition 8.8 We let $\mathbf{OR}_G$-$\mathcal{S}$ denote the ∞-category of $\mathbf{OR}_G$-spaces, given by the ∞-category of functors $\mathbf{OR}_G \to \mathcal{S}$.

We are finally ready to prove the main result of this subsection.

Theorem 8.9 *Let G be a Lie group. Then there is an equivalence of ∞-categories*

$$\mathcal{I}\text{-}G\mathcal{T}[W_{\mathrm{lvl}}^{-1}] \simeq \mathbf{OR}_G\text{-}\mathcal{S}.$$

Proof The discussion in Example 8.7 shows that there exists a functor of topological categories (and so of ∞-categories)

$$\mathbf{OR}_G^{\mathrm{op}} \to \mathcal{I}\text{-}G\mathcal{T}, \quad (H,V) \mapsto \mathbf{OR}_G((H,V),(e,-)) = G \times_H \mathcal{I}_V.$$

This is fully faithful by definition of $\mathbf{OR}_G$. Since the $\mathcal{I}$-G-spaces $G \times_H \mathcal{I}_V$ are bifibrant in the level model structure, the composite

$$L\colon \mathbf{OR}_G^{\mathrm{op}} \to \mathcal{I}\text{-}G\mathcal{T} \to \mathcal{I}\text{-}G\mathcal{T}[W_{\mathrm{lvl}}^{-1}], \quad (H,V) \mapsto G \times_H \mathcal{I}_V,$$

is also fully faithful. We apply Theorem 3.39 to the functor L. We note that the $\mathcal{I}$-G-space $G \times_H \mathcal{I}_V$ corepresents the functor $X \mapsto X(V)^H$. This functor commutes with small homotopy colimits since:

- The H-fixed-point functor preserves small homotopy colimits as discussed in Example 3.40.
- The evaluation functor $X \mapsto X(V)$ preserves small homotopy colimits. Indeed, this functor preserves all colimits (as they are calculated pointwise), level equivalences by definition, and (acyclic) cofibrations (as one can verify by checking on the generating (acyclic) cofibrations).

Finally, the collection of objects $\{G \times_H \mathcal{I}_V \mid (H,V) \in \mathbf{OR}_G\}$ is jointly conservative by definition of the level equivalences. Thus the required equivalence follows from Theorem 3.39. $\square$

Next we explain how to upgrade the equivalence above to an equivalence of symmetric monoidal ∞-categories.

Construction 8.10 We enhance the topological category $\mathbf{OR}_G$ to a topological colored operad as follows. The colors are simply the objects of $\mathbf{OR}_G$, and the space of multimorphisms from $\{(H_i,V_i)\}_{i\in I}$ to (K,W) is given by

$$\mathbf{OR}_G(\{(H_i,V_i)\}_{i\in I},(K,W)) = \left(\left(\prod_{i\in I} G\right) \times_{(\prod_{i\in I} H_i)} \mathcal{I}\left(\bigoplus_{i\in I} V_i, W\right)\right)^K.$$

By Lemma 8.3, a point of this space is equivalent to the following data:

- For all $i \in I$, an element $g_i H_i \in G/H_i$ such that $c_{g_i}(K) \subseteq H_i$.
- A linear isometry $\varphi = \sum_i \varphi_i : \bigoplus_i V_i \to W$ such that $k.\varphi_i(v) = \varphi_i(c_{g_i}(k)v)$ for all $v \in V_i$, $k \in K$ and $i \in I$.

For every map $I \to J$ of finite sets with fibers $\{I_j\}_{j\in J}$, every finite collections of objects $\{(H_i, V_i)\})_{i\in I}$ and $\{(K_j, W_j)\}_{j\in J}$, and every $(L, U) \in \mathbf{OR}_G$ we have a composition map

$$\prod_{j\in J} \mathbf{OR}_G(\{(H_i, V_i)\}_{i\in I_j}, (K_j, W_j)) \times \mathbf{OR}_G(\{(K_j, W_j)\}_{j\in J}, (L, U)) \to \mathbf{OR}_G(\{(H_i, V_i)\}_{i\in I}, (L, U)),$$

which is defined by the formulas

$$\Bigl(\bigoplus_{i\in I_j} V_i \to W_j, \bigoplus_{j\in J} W_j \to U\Bigr) \mapsto \Bigl(\bigoplus_{i\in I} V_i = \bigoplus_{j\in J} \bigoplus_{i\in I_j} V_i \to \bigoplus_{j\in J} W_j \to U\Bigr)$$

and

$$\bigl((g_i H_i)_{i\in I_j}, (g_j K_j)_{j\in J}\bigr) \mapsto (g_j g_i H_i)_{j\in J, i\in I_j}.$$

Note that for any color $(H, V) \in \mathbf{OR}_G$, there is an identity element $[eH, 1_V] \in \mathbf{OR}_G((H, V), (H, V))$. Using Lemma 8.3 one can check that this composition is continuous, associative and unital and so that $\mathbf{OR}_G$ is indeed a topological colored operad. We leave the details to the interested reader.

Remark 8.11 We can endow the topological category $\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}}$ with a topological colored operad structure whose colors are the objects of $\boldsymbol{O}_{G,\mathrm{pr}}$, and whose multimorphism spaces are given by

$$\boldsymbol{O}_{G,\mathrm{pr}}(\{G/H_i\}_{i\in I}, G/K) = \boldsymbol{O}_{G,\mathrm{pr}}\Bigl(G/K, \prod_{i\in I} G/H_i\Bigr) = \Bigl(\prod_{i\in I} G/H_i\Bigr)^K$$

with composition defined in the obvious way. The associated ∞-operad models the cocartesian monoidal structure. There is a canonical projection functor of topological colored operads

$$\pi_G : \mathbf{OR}_G \to \boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}}.$$

By Lemma 2.1, we can lift π_G to a map of ∞-operads $\mathbf{OR}_G^{\otimes} \to (\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}})^{\amalg}$, which by abuse of notation we still denote by π_G.

Recall that because $\mathcal{I}\text{-}G\mathcal{T}$ is a symmetric monoidal topological model category, we can construct a topological colored operad whose colors are given by the bifibrant objects of $\mathcal{I}\text{-}G\mathcal{T}$ and the multimorphism spaces are given by

$$\mathrm{Mul}_{N^{\otimes}((\mathcal{I}\text{-}G\mathcal{T}^{\circ})^{\mathrm{op}})}(\{X_i\}, Y) = \mathcal{I}\text{-}G\mathcal{T}\Bigl(Y, \bigotimes_{i\in I} X_i\Bigr).$$

Furthermore the associated ∞-operad models the symmetric monoidal structure on $(\mathcal{I}\text{-}G\mathcal{T}[W_{\mathrm{lvl}}^{-1}])^{\mathrm{op}}$.

Lemma 8.12 *The functor L of Theorem 8.9 lifts to a fully faithful functor of topological colored operads.*

Proof We define a functor between colored operads by

$$\mathbf{OR}_G \to (\mathcal{I}\text{-}G\mathcal{T}^{\circ})^{\mathrm{op}}, \quad \{(H_i, V_i)\} \mapsto \mathbf{OR}_G\Bigl(\bigotimes (H_i, V_i), (e, -)\Bigr).$$

Using equation (7.6.1), we can rewrite this functor in more familiar terms as

$$\mathbf{OR}_G(\{(H_i,V_i)\},(e,-)) = \left(\prod_i G\right) \times_{(\prod_i H_i)} \mathscr{I}\left(\bigoplus_i V_i,-\right) \simeq \bigotimes_i (G \times_{H_i} \mathscr{I}_{V_i}).$$

By construction, this functor defines a colored operad map which lifts L. Using this description of the functor and the fact that $G \times_H \mathscr{I}_W$ corepresents the functor $X \mapsto X(W)^K$, we also see that the map induced on multimorphism spaces

$$\mathbf{OR}_G(\{(H_i,V_i)\}_{i\in I},(K,W)) \to \mathscr{I}\text{-}G\mathscr{T}\left(G\times_K \mathscr{I}_W, \bigotimes_{i\in I} G\times_{H_i}\mathscr{I}_{V_i}\right)$$

is a homeomorphism. Therefore the functor of colored operads is fully faithful. □

The map L of topological operads constructed above induces a map $L\colon \mathbf{OR}_G^{\otimes} \to (\mathscr{I}\text{-}G\mathscr{T}[W_{\mathrm{lvl}}^{-1}]^{\otimes})^{\mathrm{op}}$ of ∞-operads. Furthermore this functor is again fully faithful.

Corollary 8.13 *The functor $L\colon \mathbf{OR}_G^{\otimes} \to (\mathscr{I}\text{-}G\mathscr{T}_*[W_{\mathrm{lvl}}^{-1}]^{\otimes})^{\mathrm{op}}$ induces a symmetric monoidal equivalence*

$$\mathscr{I}\text{-}G\mathscr{T}[W_{\mathrm{lvl}}^{-1}] \simeq \mathbf{OR}_G\text{-}\mathscr{S},$$

where the right-hand side is equipped with the Day convolution product.

Proof This follows from Corollary 3.41, where we argue as in Theorem 8.9 and use Lemma 8.12. □

As a convenient reference, let us summarize the final description of G-prespectrum objects, which combines all of the identifications obtained.

Corollary 8.14 *Let G be a Lie group. Then there is a symmetric monoidal equivalence*

$$\mathrm{PSp}_G \simeq \mathrm{Mod}_{S_G}(\mathbf{OR}_G\text{-}\mathscr{S}_*).$$

Proof Combine Corollary 8.13 and Propositions 7.22 and 3.38. □

Remark 8.15 We will often implicitly identify PSp_G with $\mathrm{Mod}_{S_G}(\mathbf{OR}_G\text{-}\mathscr{S}_*)$ for the remainder of the paper.

8.2 $\mathscr{I}$-spaces and $\mathbf{OR}_{\mathrm{fgl}}$-spaces.

We now undertake a similar analysis for the ∞-category of $\mathscr{I}$-spaces localized at the faithful level equivalences. Many of the details are similar, so we will be briefer in this section than in the previous one.

Definition 8.16 We define a topological category $\mathbf{OR}_{\mathrm{fgl}}$ whose objects are pairs (G,V), where G is a compact Lie group and V is a faithful G-representation. The morphism spaces are given by

$$\mathbf{OR}_{\mathrm{fgl}}((G,V),(H,W)) = (\mathscr{I}(V,W)/G)^H.$$

There is a composition map

$$\circ\colon \mathbf{OR}_{\mathrm{fgl}}((H,W),(L,U)) \times \mathbf{OR}_{\mathrm{fgl}}((G,V),(H,W)) \to \mathbf{OR}_{\mathrm{fgl}}((G,V),(L,U))$$

given by $([\psi],[\varphi]) \mapsto [\psi\circ\varphi]$. Similarly to Lemma 8.4, one may verify that this composition is well-defined, associative, unital and continuous.

Example 8.17 By definition $\mathbf{OR}_{\mathrm{fgl}}((G,V),(e,W)) = \mathscr{I}(V,W)/G$. Thus we can identify the functor

$$\mathbf{OR}_{\mathrm{fgl}}((G,V),(e,-))\colon \mathscr{I} \to \mathscr{T}$$

with the semifree $\mathscr{I}$-space $\mathscr{I}_{G,V}$ from Construction 7.15. Recall this $\mathscr{I}$-space corepresents the functor $X \mapsto X(V)^G$.

Definition 8.18 We let $\mathbf{OR}_{\mathrm{fgl}}$-$\mathscr{S}$ denote the ∞-category of $\mathbf{OR}_{\mathrm{fgl}}$-spaces which is the ∞-category of functors $\mathbf{OR}_{\mathrm{fgl}} \to \mathscr{S}$. We also write $\mathbf{OR}_{\mathrm{fgl}}$-$\mathscr{S}_*$ for the ∞-category of functors $\mathbf{OR}_{\mathrm{fgl}} \to \mathscr{S}_*$.

We now prove the main result of this subsection.

Theorem 8.19 *There is an equivalences of ∞-categories*

$$\mathscr{I}\text{-}\mathscr{T}[W_{\text{f-lvl}}^{-1}] \simeq \mathbf{OR}_{\mathrm{fgl}}\text{-}\mathscr{S}.$$

Proof The discussion in Example 8.17 shows that there exists a functor of topological categories (and so of ∞-categories)

$$(\mathbf{OR}_{\mathrm{fgl}})^{\mathrm{op}} \to \mathscr{I}\text{-}\mathscr{T}, \quad (G,V) \mapsto \mathbf{OR}_{\mathrm{fgl}}((G,V),(e,-)) = \mathscr{I}_{G,V}.$$

This is fully faithful by definition of $\mathbf{OR}_{\mathrm{fgl}}$. Since the $\mathscr{I}$-spaces $\mathscr{I}_{G,V}$ are bifibrant in the faithful level model structure, the composite

$$(\mathbf{OR}_{\mathrm{fgl}})^{\mathrm{op}} \to \mathscr{I}\text{-}\mathscr{T} \to \mathscr{I}\text{-}\mathscr{T}[W_{\text{f-lvl}}^{-1}]$$

is also fully faithful. We note that the semifree $\mathscr{I}$-space $\mathscr{I}_{G,V}$ corepresents the functor $X \mapsto X(V)^G$, which commutes with small homotopy colimits. Indeed the G-fixed-point functor commutes with small homotopy colimits by the discussion in Example 3.40, and so does the evaluation functor $X \mapsto X(V)$ since it preserves all colimits (as they are calculated pointwise), faithful level equivalences by definitions and cofibrations (as one can verify by checking on the set of generating cofibrations). Finally, the collection of objects $\{\mathscr{I}_{G,V} \mid (G,V) \in \mathbf{OR}_{\mathrm{fgl}}\}$ is jointly conservative by definition of the faithful level equivalences. Thus the claimed equivalence follows by applying Theorem 3.39. $\square$

We now discuss how the symmetric monoidal structure on $\mathscr{I}\text{-}\mathscr{T}^c[W_{\text{f-lvl}}^{-1}]$ translates to $\mathbf{OR}_{\mathrm{fgl}}$-$\mathscr{S}_*$.

Lemma 8.20 *The topological category $\mathbf{OR}_{\mathrm{fgl}}$ is symmetric monoidal with unit object $(e,0)$ and tensor product given by $(G,V)\otimes(H,W)=(G\times H,V\oplus W)$. In particular, the ∞-category of $\mathbf{OR}_{\mathrm{fgl}}$-spaces admits a symmetric monoidal structure given by Day convolution.*

Proof The first claim is a straightforward verification. The second claim follows from Corollary 3.29. □

Write $\mathbf{OR}_{\mathrm{fgl}}^{\otimes}$ for the ∞-operad associated to symmetric monoidal topological category $\mathbf{OR}_{\mathrm{fgl}}$.

Lemma 8.21 *The functor $L_{\mathrm{gl}}\colon \mathbf{OR}_{\mathrm{fgl}}\to(\mathscr{I}\text{-}\mathscr{T}[W_{\text{f-lvl}}^{-1}])^{\mathrm{op}}$ given by $(G,V)\mapsto\mathscr{I}_{G,V}$ lifts to a fully faithful symmetric monoidal functor of ∞-categories,*

$$L_{\mathrm{gl}}\colon \mathbf{OR}_{\mathrm{fgl}}\to(\mathscr{I}\text{-}\mathscr{T}[W_{\text{f-lvl}}^{-1}])^{\mathrm{op}}.$$

Proof It suffices to observe that (7.15.1) implies that $L_{\mathrm{gl}}\colon\mathbf{OR}_{\mathrm{fgl}}\to\mathscr{I}\text{-}\mathscr{T}$ is a strong monoidal functor. □

Corollary 8.22 *There is a symmetric monoidal equivalence*

$$\mathscr{I}\text{-}\mathscr{T}[W_{\mathrm{lvl}}^{-1}]\simeq\mathbf{OR}_{\mathrm{fgl}}\text{-}\mathscr{S},$$

where the right-hand side is symmetric monoidal via Day convolution.

Proof This follows from Corollary 3.41, where we argue as in Theorem 8.19 and use Lemma 8.21. □

Summarizing all of the identifications made, we have the following description of the symmetric monoidal ∞-category of faithful global prespectra.

Corollary 8.23 *There is a symmetric monoidal equivalence*

$$\mathrm{PSp}_{\mathrm{fgl}}\simeq\mathrm{Mod}_{S_{\mathrm{fgl}}}(\mathbf{OR}_{\mathrm{fgl}}\text{-}\mathscr{S}_*).$$

Proof Combine Proposition 7.22, Corollary 8.22 and Proposition 3.38. □

Remark 8.24 We will often implicitly identify $\mathrm{PSp}_{\mathrm{fgl}}$ with $\mathrm{Mod}_{S_{\mathrm{fgl}}}(\mathbf{OR}_{\mathrm{fgl}}\text{-}\mathscr{S}_*)$.

9 Functoriality of equivariant prespectra

The goal of this section is to construct a functor $\mathrm{PSp}_\bullet\colon\mathrm{Glo}^{\mathrm{op}}\to\mathrm{Cat}_\infty^{\otimes}$ sending a compact Lie group G to the symmetric monoidal ∞-category of G-prespectra of Definition 7.18, and to compute its (partially) lax limit. By Corollary 8.14, the ∞-category of G-prespectra can be identified with the category of modules over a certain object S_G in $\mathbf{OR}_G\text{-}\mathscr{S}_*$. Therefore our first step is to construct a functor sending a compact Lie group G to the ∞-category $\mathbf{OR}_G\text{-}\mathscr{S}_*$.

In the unstable case we observed that the relevant functoriality was induced by the functoriality of the partial slices $\mathrm{Orb}_{/G}$ in Glo. Formally, the functoriality of the categories $\mathbf{OR}_G\text{-}\mathcal{S}_*$ is induced by a (pro)functoriality of the categories $\mathbf{OR}_G$, and we will see that this is once again given by "passing to the slices" of a global analogue $\mathbf{OR}_{\mathrm{gl}}$ of the individual equivariant categories $\mathbf{OR}_G$. The category $\mathbf{OR}_{\mathrm{gl}}$ will be fibered over Glo and its objects will consist of pairs (G, V), where G is a compact Lie group and V is an arbitrary G-representation. Furthermore we will see that restricting to faithful representations, we recover $\mathbf{OR}_{\mathrm{fgl}}$.

Construction 9.1 Let G, H be compact Lie groups and let V, W be orthogonal G and H-representations respectively. We equip the topological space

$$\mathrm{Hom}(H, G) \times \mathcal{I}(V, W)$$

with the right G-action and the left H-action given by

$$(\alpha, \varphi) \cdot g = (c_g \alpha, \varphi g^{-1}) \quad \text{and} \quad h \cdot (\alpha, \varphi) = (\alpha, h\varphi\alpha(h)^{-1}).$$

There is a residual G-action on the fixed points $(\mathrm{Hom}(H, G) \times \mathcal{I}(V, W))^H$ since the G and H-actions commute. By definition, the fixed-point space can be characterized as the space of pairs (α, φ), where $\alpha\colon H \to G$ is a Lie group homomorphism and $\varphi\colon V \to W$ is an H-equivariant isometry (where H acts on V via α). If K is another compact Lie group and U is an orthogonal K-representation, we define a composition map

$$\begin{aligned}(\mathrm{Hom}(H, G) \times \mathcal{I}(V, W))^H \times (\mathrm{Hom}(K, H) \times \mathcal{I}(W, U))^K &\to (\mathrm{Hom}(K, G) \times \mathcal{I}(V, U))^K,\\ (\alpha, \varphi) \cdot (\beta, \psi) &= (\alpha\beta, \varphi\psi),\end{aligned}$$

that is compatible with the various actions, so that it induces an associative and unital composition map on the respective action groupoids:

$$(\mathrm{Hom}(H, G) \times \mathcal{I}(V, W))^H /\!\!/ G \times (\mathrm{Hom}(K, H) \times \mathcal{I}(W, U))^K /\!\!/ H \to (\mathrm{Hom}(K, G) \times \mathcal{I}(V, U))^K /\!\!/ G.$$

Definition 9.2 Let $\mathbf{OR}_{\mathrm{gl}}$ be the topological category whose objects are pairs (G, V), where G is a compact Lie group and V is an orthogonal G-representation. Its morphism spaces are defined to be

$$\mathbf{OR}_{\mathrm{gl}}((G, V), (H, W)) = |(\mathrm{Hom}(H, G) \times \mathcal{I}(V, W))^H /\!\!/ G|,$$

where $|- /\!\!/ G|$ is the geometric realization of the action groupoid of G on $\mathcal{I}(V, W)$ (as in Definition 6.1). As in Lemma 8.20, one sees that $\mathbf{OR}_{\mathrm{gl}}$ admits a symmetric monoidal structure given by $(G, V) \otimes (H, W) \simeq (G \times H, V \oplus W)$. We write $\mathbf{OR}_{\mathrm{gl}}^{\otimes}$ for the associated ∞-operad.

The next result tells us that the ∞-category $\mathbf{OR}_{\mathrm{fgl}}$ from Definition 8.16 is equivalent to the subcategory of $\mathbf{OR}_{\mathrm{gl}}$ spanned by the faithful representations.

Lemma 9.3 *Let $\mathscr{C}$ be the symmetric monoidal subcategory of $\mathbf{OR}_{\mathrm{gl}}$ spanned by (G,V), where V is a faithful G-representation. Then there is a symmetric monoidal functor of topological categories $\mathscr{C}\to\mathbf{OR}_{\mathrm{fgl}}$ sending (G,V) to (G,V), which induces a homotopy equivalence on mapping spaces (and so it is an equivalence of the underlying ∞-categories).*

Proof The functor is the identity on objects, so it suffices to define it on mapping spaces. For any $(G,V),(H,W)\in\mathscr{C}$, let us consider the map

$$p\colon(\mathrm{Hom}(H,G)\times\mathscr{I}(V,W))^H\to(\mathscr{I}(V,W)/G)^H$$

sending (α,φ) to $[\varphi]$. We claim that this map exhibits the target as the quotient of the source by G. Firstly, note that the map is G-equivariant. Let us show that its fibers are exactly the G-orbits. Suppose we have a point $[\varphi]$ in the target and let us choose a representative $\varphi\colon V\to W$. Then we know that $h\cdot[\varphi]=[h\varphi]=[\varphi]$ for every $h\in H$. Then necessarily there exists $\alpha(h)\in G$ such that $h\varphi=\varphi\alpha(h)^{-1}$. Note that the element $\alpha(h)$ is unique since V is a faithful G-representation. Then the map $h\mapsto\alpha(h)$ is a Lie group homomorphism and its graph is closed in $H\times G$ (since it is a fiber of the continuous map $H\times G\to\mathscr{I}(V,W)$ sending (h,g) to $h\varphi g^{-1}$), so it is continuous. Then it is clear that (α,φ) is a preimage of $[\varphi]$, and so p is surjective.

On the other hand, if (α,φ) and (α',φ') have the same image under p, then there is some $g\in G$ such that $\varphi'=\varphi g$. A simple computation as before shows that this forces $\alpha'=c_g\alpha$ (since the G-action on $\mathscr{I}(V,W)$ is faithful, α' is determined by φ'). Moreover, the action of G on $(\mathrm{Hom}(H,G)\times\mathscr{I}(V,W))^H$ is free and proper, and so p is a principal G-bundle. In particular it induces a natural equivalence of topological groupoids

$$(\mathrm{Hom}(H,G)\times\mathscr{I}(V,W))^H/\!\!/G\simeq(\mathscr{I}(V,W)/G)^H,$$

and so a homotopy equivalence

$$|(\mathrm{Hom}(H,G)\times\mathscr{I}(V,W))^H/\!\!/G|\simeq(\mathscr{I}(V,W)/G)^H.$$

Finally, it is easy to check that p is compatible with composition and sends the identity to the identity. Therefore it induces an equivalence of ∞-categories $\mathscr{C}\to\mathbf{OR}_{\mathrm{fgl}}$. We leave to the reader to check that the above can be given the structure of a symmetric monoidal equivalence. □

Remark 9.4 There is a pair of functors of topological categories

$$s_0\colon\mathrm{Glo}^{\mathrm{op}}\to\mathbf{OR}_{\mathrm{gl}},\ \ \pi_{\mathrm{gl}}\colon\mathbf{OR}_{\mathrm{gl}}\to\mathrm{Glo}^{\mathrm{op}}$$

given by $s_0(G)=(G,0)$ and $\pi_{\mathrm{gl}}(G,V)=G$ on objects. Note that s_0 and π_{gl} are both symmetric monoidal, where Glo is symmetric monoidal under the cartesian product (and therefore $\mathrm{Glo}^{\mathrm{op}}$ is equipped with the cocartesian symmetric monoidal structure). This implies that the functors π_{gl} and s_0 lift to maps of ∞-operads $\pi_{\mathrm{gl}}\colon\mathbf{OR}_{\mathrm{gl}}^{\otimes}\to(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$ and $s_0\colon(\mathrm{Glo}^{\mathrm{op}})^{\amalg}\to\mathbf{OR}_{\mathrm{gl}}^{\otimes}$, respectively.

Lemma 9.5 *Let $\{(G_i, V_i)\}, (H, W)$ be objects of $\mathbf{OR}_{\mathrm{gl}}^{\otimes}$, and consider the map*

$$\pi_{\mathrm{gl}}\colon \mathrm{Mul}_{\mathbf{OR}_{\mathrm{gl}}}(\{(G_i, V_i)\}, (H, W)) \to \mathrm{Mul}_{\mathrm{Glo}^{\mathrm{op}}}(\{G_i\}, H).$$

The homotopy fiber of this map over a group homomorphism $\alpha\colon H \to \prod_i G_i \in (\mathrm{Glo}^{\mathrm{op}})^{\amalg}$ is equivalent to the space of H-equivariant isometries $\bigoplus_i V_i \to W$, where H acts on $\bigoplus_i V_i$ via α.

Proof Put $V = \bigoplus_i V_i$ and $G = \prod_i G_i$ so that $\alpha\colon H \to G$, and we can rewrite the map induced by π_{gl} as

$$\mathrm{Map}_{\mathbf{OR}_{\mathrm{gl}}}((G, V), (H, W)) \to \mathrm{Map}_{\mathrm{Glo}^{\mathrm{op}}}(G, H) = \mathrm{Map}_{\mathrm{Glo}}(H, G).$$

We recall from Proposition 6.3 that the G-space $\mathrm{Hom}(H, G)$ decomposes as a disjoint union of orbits

$$\mathrm{Hom}(H, G) \simeq \coprod_{(\alpha)} G/C(\alpha),$$

where α is a conjugacy class of homomorphisms and $C(\alpha)$ is the centralizer of the image of α. Therefore we have a decomposition

$$\mathrm{Map}_{\mathbf{OR}_{\mathrm{gl}}}((G, V), (H, W)) \simeq \big((\mathrm{Hom}(H, G) \times \mathcal{I}(V, W))^{H}\big)_{hG} \simeq \coprod_{(\alpha)} \mathcal{I}(V, W)^{H}_{hC(\alpha)},$$

depending on the choice of an α in each conjugacy class. This lies above the decomposition

$$\mathrm{Map}_{\mathrm{Glo}}(H, G) \simeq \coprod_{(\alpha)} BC(\alpha)$$

from Proposition 6.3 via the canonical maps $\mathcal{I}(V, W)^{H} \to *$. Therefore the homotopy fiber over α is precisely $\mathcal{I}(V, W)^{H}$. □

Lemma 9.6 *The functor $\pi_{\mathrm{gl}}\colon \mathbf{OR}_{\mathrm{gl}}^{\otimes} \to (\mathrm{Glo}^{\mathrm{op}})^{\amalg}$ is a cocartesian fibration, and therefore exhibits $\mathbf{OR}_{\mathrm{gl}}^{\otimes}$ as a $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$-monoidal ∞-category.*

Proof Consider $\{(G_i, V_i)\}_{i \in I} \in \mathbf{OR}_{\mathrm{gl}}^{\otimes}$, and let us set $V = \bigoplus_i V_i$ and $G = \prod_i G_i$ so that V is naturally a G-representation. Since π_{gl} is a map of ∞-operads, it is enough to find cocartesian lifts over active morphisms whose target is in $\mathrm{Glo}^{\mathrm{op}}$. A multimorphism from $\{G_i\}$ to H in $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$ is the datum of a continuous group homomorphism $\alpha\colon H \to G$. Consider the multimorphism $f \in \mathbf{OR}_{\mathrm{gl}}^{\otimes}(\{(G_i, V_i)\}, (H, \alpha^* V))$ lying over the map α which is represented by the element

$$[\alpha, 1_V] \in |(\mathrm{Hom}(H, G) \times \mathcal{I}(V, \alpha^* V))^{H} /\!/ G|.$$

We claim that this is a cocartesian edge. This follows from the fact that for all $(L, W) \in \mathbf{OR}_{\mathrm{gl}}^{\otimes}$, the square

$$\begin{tikzcd}
\mathrm{Mul}_{\mathbf{OR}_{\mathrm{gl}}}((H, \alpha^* V), (L, W)) \arrow[r, "f^*"] \arrow[d, "\pi_{\mathrm{gl}}"] & \mathrm{Mul}_{\mathbf{OR}_{\mathrm{gl}}}(\{(G_i, V_i)\}, (L, W)) \arrow[d, "\pi_{\mathrm{gl}}"] \\
\mathrm{Mul}_{\mathrm{Glo}^{\mathrm{op}}}(H, L) \arrow[r, "\alpha^*"] & \mathrm{Mul}_{\mathrm{Glo}^{\mathrm{op}}}(\{G_i\}, L)
\end{tikzcd}$$

is a homotopy pullback of spaces. We can verify this by checking that the vertical fibers are equivalent. This is now a consequence of Lemma 9.5. □

Definition 9.7 We define $\mathrm{Rep}\colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_\infty^\otimes$ to be the functor corresponding to $\mathbf{OR}_{\mathrm{gl}}^\otimes$ under the equivalence of Proposition 5.5.

Remark 9.8 $\mathrm{Rep}(G)$ is the ∞-category corresponding to the topologically enriched category with objects V a G-representation, and morphism spaces $\mathrm{Rep}(V,W) = \mathcal{I}(V,W)^G$, the space of G-equivariant linear isometries from V to W. This is a symmetric monoidal category via direct sum. The functoriality in Glo is given by restriction of representations along group homomorphisms.

Recall from Remark 8.11 that there is a map of ∞-operads $\pi_G\colon \mathbf{OR}_G^\otimes \to (\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}})^{\amalg}$. Also note that there is a canonical functor $\boldsymbol{O}_{G,\mathrm{pr}} \to \mathrm{Glo}$ which sends an object G/H to H and acts as

$$\boldsymbol{O}_{G,\mathrm{pr}}(G/H, G/K) \simeq \{g \in G \mid c_g(H) \subseteq K\}_{\mathrm{h}K} \to \hom(H,K)_{\mathrm{h}K}, \quad g \mapsto [c_g\colon H \to K].$$

This is an immediate generalization of the functor used in Lemma 6.12 to (not necessarily compact) Lie groups. We denote the opposite of this functor by ι_G. It induces a map of cocartesian ∞-operads, which we denote by $\iota_G^{\amalg}$. We are now ready to state the next result.

Lemma 9.9 *Let G be a Lie group. Then there is a canonical map of ∞-operads $\nu_G\colon \mathbf{OR}_G^\otimes \to \mathbf{OR}_{\mathrm{gl}}^\otimes$ and a cartesian square of ∞-operads*

$$\begin{tikzcd}
\mathbf{OR}_G^\otimes \arrow[r, "\nu_G"] \arrow[d, "\pi_G"'] & \mathbf{OR}_{\mathrm{gl}}^\otimes \arrow[d, "\pi_{\mathrm{gl}}"] \\
(\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}})^{\amalg} \arrow[r, "\iota_G^{\amalg}"] & (\mathrm{Glo}^{\mathrm{op}})^{\amalg}
\end{tikzcd}$$

Proof It will suffice to construct the map ν_G at the level of topological colored operads and then apply Lemma 2.1. Recall from Definition 8.5 that

$$\mathbf{OR}_G((H,V),(K,W)) = (G \times_H \mathcal{I}(V,W))^K,$$

where $G \times_H \mathcal{I}(V,W)$ is the quotient of $G \times \mathcal{I}(V,W)$ by the right H-action $(g,\varphi)\cdot h = (gh, \varphi h)$. Since the H-action is free, we can identify the quotient with the homotopy quotient (see [Körschgen 2018, Theorem A.7] for example) and so there is a canonical identification

$$\mathbf{OR}_G((H,V),(K,W)) = |(G \times \mathcal{I}(V,W))^K /\!\!/ H|$$

that respects composition. Moreover, under this identification, the multilinear spaces of the colored operad structure are given by

$$\mathbf{OR}_G(\{(H_i,V_i)\}_i,(K,W)) = \left| \left(\prod_i G \times \mathcal{I}\left(\bigoplus_i V_i, W\right)\right)^K /\!\!/ \prod H_i \right|.$$

Therefore we may define a functor of topological colored operad $\mathbf{OR}_G \to \mathbf{OR}_{\mathrm{gl}}$ by sending (H,V) to (H,V) and on the multimorphism spaces we take the map which is induced by the map of topological groupoids

$$\Big(\prod_i G \times \mathcal{I}\Big(\bigoplus_i V_i, W\Big)\Big)^K /\!\!/ \prod_i H_i \to \Big(\mathrm{Hom}\Big(K, \prod_i H_i\Big) \times \mathcal{I}\Big(\bigoplus_i V_i, W\Big)\Big)^K /\!\!/ \prod_i H_i,$$
$$(\{g_i\}, \varphi) \mapsto ((c_{g_i}|_K)_i, \varphi).$$

A tedious but simple calculation shows that these maps respect composition. This defines a map $\nu_G \colon \mathbf{OR}_G^{\otimes} \to \mathbf{OR}_{\mathrm{gl}}^{\otimes}$ as required.

Another tedious calculation shows that the square in the lemma commutes (already as a square of topological operads) and that it is a pullback on 0-vertices. Therefore it is enough to show that every induced square

$$\begin{tikzcd}
\mathrm{Mul}_{\mathbf{OR}_G}(\{(H_i,V_i)\},(K,W)) \arrow[r, "\nu_G"] \arrow[d, "\pi_G"] & \mathrm{Mul}_{\mathbf{OR}_{\mathrm{gl}}}(\{(H_i,V_i)\},(K,W)) \arrow[d, "\pi_{\mathrm{gl}}"] \\
\mathrm{Mul}_{\boldsymbol{O}^{\mathrm{op}}_{G,\mathrm{pr}}}(\{G/H_i\}, G/K) \arrow[r, "\iota_G"] & \mathrm{Mul}_{\mathrm{Glo}^{\mathrm{op}}}(\{H_i\}, K)
\end{tikzcd}$$

of multimorphism spaces is a homotopy pullback. It suffices to check that the vertical homotopy fibers are equivalent. A morphism $\varphi \colon G/K \to \prod G/H_i$ in $\boldsymbol{O}_{G,\mathrm{pr}}$ amounts to giving elements $g_i \in G$ such that $c_{g_i}(K) \subseteq H_i$. The homotopy fiber of π_G over φ is given by the space of K-equivariant isometries $\bigoplus_i V_i \to W$, where K acts on each V_i via c_{g_i}. The map ι_G sends φ to $(c_{g_i} \colon K \to H_i)$ and the homotopy fiber over this is again the space of K-equivariant isometries as above by Lemma 9.5. As the vertical homotopy fibers are equivalent, the square is a pullback of ∞-operads. □

We write $\mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo})$ for the *full* subcategory of $\mathrm{Ar}(\mathrm{Glo})$ spanned by the injective group homomorphisms.

Definition 9.10 We define $\mathbf{OR}^{\otimes}$ via the pullback of ∞-operads

$$\begin{tikzcd}
\mathbf{OR}^{\otimes} \arrow[r] \arrow[d, "\pi_{\mathrm{inj}}"'] & \mathbf{OR}_{\mathrm{gl}}^{\otimes} \arrow[d, "\pi_{\mathrm{gl}}"] \\
(\mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo})^{\mathrm{op}})^{\amalg} \arrow[r, "s^{\mathrm{op}}"] & (\mathrm{Glo}^{\mathrm{op}})^{\amalg}
\end{tikzcd}$$

Thus an object of $\mathbf{OR}$, the underlying ∞-category of $\mathbf{OR}^{\otimes}$, is a pair $(\alpha \colon H \to G, V)$, where α is injective and V is a H-representation.

Lemma 9.11 *The composition*

$$\pi \colon \mathbf{OR}^{\otimes} \xrightarrow{\pi_{\mathrm{inj}}} (\mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo})^{\mathrm{op}})^{\amalg} \xrightarrow{t^{\mathrm{op}}} (\mathrm{Glo}^{\mathrm{op}})^{\amalg}$$

gives $\mathbf{OR}^{\otimes}$ *the structure of a* $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$*-promonoidal* ∞*-category, whose operadic fiber over* G *is exactly* $\mathbf{OR}_G^{\otimes}$.

Proof We will show that each of the two maps in the defining composite is promonoidal in turn. Note that both are maps of ∞-operads. The map π_{inj} is a pullback of a cocartesian fibration, and therefore again cocartesian. The second map is then promonoidal by Example 3.7.

Finally we note that the operadic fiber of t^{op} over G is $(\boldsymbol{O}_G^{\mathrm{op}})^{\amalg}$ by Lemma 6.12 and the observation that $(-)^{\amalg}$ preserves pullbacks. Therefore, the calculation of the operadic fiber follows from Lemma 9.9 and the observation that the composite

$$(\boldsymbol{O}_G^{\mathrm{op}})^{\amalg} \to (\mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo})^{\mathrm{op}})^{\amalg} \xrightarrow{t^{\mathrm{op}}} (\mathrm{Glo}^{\mathrm{op}})^{\amalg}$$

is equivalent to $\iota_G^{\amalg}$. $\square$

Because π is a promonoidal category over $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$ with operadic fiber $\mathbf{OR}_G^{\otimes}$, morally it represents a profunctor of promonoidal ∞-categories. Therefore we can extract an honest symmetric monoidal functor by taking copresheafs. This will be the functor $\mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_\infty^{\otimes}$ sending G to $\mathbf{OR}_G\text{-}\mathscr{S}_*$.

Definition 9.12 The Day convolution $\mathrm{Fun}_{\mathrm{Glo}^{\mathrm{op}}}(\mathbf{OR}^{\otimes}, \mathscr{S}_*^{\wedge} \times (\mathrm{Glo}^{\mathrm{op}})^{\amalg})^{\mathrm{Day}}$ is a $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$-monoidal ∞-category, whose operadic fiber over $G \in \mathrm{Glo}$ equals

$$\begin{aligned}\mathrm{Fun}_{\mathrm{Glo}^{\mathrm{op}}}(\mathbf{OR}^{\otimes}, \mathscr{S}_*^{\wedge} \times (\mathrm{Glo}^{\mathrm{op}})^{\amalg})^{\mathrm{Day}} \times_{(\mathrm{Glo}^{\mathrm{op}})^{\amalg}} \mathrm{Fin}_* &\simeq \mathrm{Fun}(\mathbf{OR}^{\otimes} \times_{(\mathrm{Glo}^{\mathrm{op}})^{\amalg}} \mathrm{Fin}_*, \mathscr{S}_*^{\wedge})^{\mathrm{Day}} \\ &\simeq \mathbf{OR}_G\text{-}\mathscr{S}_*\end{aligned}$$

by Example 5.7 and Lemma 9.11. We define $\mathbf{OR}_\bullet\text{-}\mathscr{S}_* \colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_\infty^{\otimes}$ to be the functor associated to it under the equivalence of Proposition 5.5.

Lemma 9.13 *Let $\mathbf{OR}$ be the underlying category of the ∞-operad $\mathbf{OR}^{\otimes}$. Then the projection map*

$$\pi \colon \mathbf{OR} \to \mathrm{Glo}^{\mathrm{op}}$$

is cartesian over $\mathrm{Orb}^{\mathrm{op}}$, and an edge $(\sigma, \phi) \in \mathbf{OR}$ is π-cartesian if and only if $s^{\mathrm{op}}(\sigma)$ and ϕ are equivalences.

Proof Suppose we have an injection $\alpha \colon H \to G$, and an object $(\beta \colon K \to H, V) \in \mathbf{OR}$. As noted before, the map $t^{\mathrm{op}} \colon \mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo})^{\mathrm{op}} \to \mathrm{Glo}^{\mathrm{op}}$ is a cartesian fibration. Furthermore, over an injection $\alpha \colon H \to G$, cartesian lifts with target $\beta \colon K \to H$ are given by squares σ:

$$\begin{tikzcd} K \arrow[d, "\alpha\beta"'] & K \arrow[l, "\sim"'] \arrow[d, "\beta"] \\ G & H \arrow[l, "\alpha"] \end{tikzcd}$$

In particular, we note that cartesian lifts of injections are sent to equivalences by the source functor $s^{\mathrm{op}} \colon \mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo})^{\mathrm{op}} \to \mathrm{Glo}^{\mathrm{op}}$. Lifting $s^{\mathrm{op}}(\sigma)$ to an equivalence $\phi \in \mathbf{OR}_{\mathrm{gl}}$ with target (K, V), we obtain an edge (σ, ϕ) which lies over α and ends at (β, V). Because both components of the edge (σ, ϕ) in $\mathbf{OR}$ are π-cartesian, the edge (σ, ϕ) is itself π-cartesian. This shows that there are enough cartesian edges in $\mathbf{OR}$ over injections, and that they are exactly of the form claimed. $\square$

Lemma 9.14 *The projection map*

$$\mathbf{OR}^{\otimes} \to \mathbf{OR}_{\mathrm{gl}}^{\otimes}$$

induces a fully faithful symmetric monoidal functor

$$\mathbf{OR}_{\mathrm{gl}}\text{-}\mathcal{S}_* \to \mathbf{OR}\text{-}\mathcal{S}_*$$

via restriction, with essential image those functors $F\colon \mathbf{OR} \to \mathcal{S}_*$ that send cartesian arrows over $\mathrm{Orb}^{\mathrm{op}}$ *to equivalences.*

Proof Recall from Lemma 6.8 that the source projection $\mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo}) \to \mathrm{Glo}$ has a fully faithful left adjoint $\mathrm{Glo} \to \mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo})$ given by the diagonal embedding. Therefore, by the functoriality of the cocartesian operad [Lurie 2017, Proposition 2.4.3.16], it follows that the source projection

$$(\mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo})^{\mathrm{op}})^{\amalg} \to (\mathrm{Glo}^{\mathrm{op}})^{\amalg}$$

has a fully faithful operadic right adjoint. Since Bousfield localizations are stable under basechange, it follows that the projection

$$\mathbf{OR}^{\otimes} \to \mathbf{OR}_{\mathrm{gl}}^{\otimes}$$

again has a fully faithful operadic right adjoint. Therefore $\mathbf{OR} \to \mathbf{OR}_{\mathrm{gl}}$ is a Bousfield localization on underlying ∞-categories and moreover the fully faithful functor

$$\mathbf{OR}_{\mathrm{gl}}\text{-}\mathcal{S} \to \mathbf{OR}\text{-}\mathcal{S}_*$$

is symmetric monoidal by Proposition 3.34(b). Finally, because $\mathbf{OR} \to \mathbf{OR}_{\mathrm{gl}}$ is a Bousfield localization, the essential image of the functor $\mathrm{Fun}(\mathbf{OR}_{\mathrm{gl}}, \mathcal{S}_*) \to \mathrm{Fun}(\mathbf{OR}, \mathcal{S}_*)$ is given by those functors which send the edges inverted by the map $\mathbf{OR} \to \mathbf{OR}_{\mathrm{gl}}$ to equivalences. But these are exactly the cartesian arrows over the injections by Lemma 9.13. □

Lemma 9.15 *There are symmetric monoidal equivalences*

$$\operatorname*{laxlim}_{G \in \mathrm{Glo}^{\mathrm{op}}} \mathbf{OR}_G\text{-}\mathcal{S}_* \simeq \mathbf{OR}\text{-}\mathcal{S} \quad \textit{and} \quad \operatorname*{laxlim^{\dagger}}_{G \in \mathrm{Glo}^{\mathrm{op}}} \mathbf{OR}_G\text{-}\mathcal{S}_* \simeq \mathbf{OR}_{\mathrm{gl}}\text{-}\mathcal{S}_*,$$

where the lax limit is marked over the subcategory $\mathrm{Orb} \subseteq \mathrm{Glo}$ *of all objects and injective maps.*

Proof By Proposition 5.8 there is a symmetric monoidal equivalence

$$\operatorname*{laxlim}_{G \in \mathrm{Glo}^{\mathrm{op}}} \mathbf{OR}_G\text{-}\mathcal{S}_* \simeq N_p \,\mathrm{Fun}_{\mathrm{Glo}^{\mathrm{op}}}(\mathbf{OR}^{\otimes}, \mathcal{S}_*^{\wedge} \times (\mathrm{Glo}^{\mathrm{op}})^{\amalg})^{\mathrm{Day}},$$

where $p\colon (\mathrm{Glo}^{\mathrm{op}})^{\amalg} \to \mathrm{Fin}_*$ is the structure morphism of $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$. Applying the formula of Day convolution twice (see Definition 3.12), and the transitivity of norms of operads, we obtain

$$\operatorname{laxlim} \mathbf{OR}_\bullet\text{-}\mathcal{S}_* \simeq N_p N_\pi \pi^*(\mathcal{S}_*^{\wedge} \times (\mathrm{Glo}^{\mathrm{op}})^{\amalg}) \simeq N_{\pi p}(\pi^* p^* \mathcal{S}_*^{\wedge}) \simeq \mathrm{Fun}(\mathbf{OR}^{\otimes}, \mathcal{S}_*^{\wedge})^{\mathrm{Day}} = \mathbf{OR}\text{-}\mathcal{S}_*.$$

To compute the partially lax limit we appeal to Remark 5.2 to reduce to a statement on underlying categories. Combining Remarks 3.11 and 5.9, we conclude that the underlying ∞-category of the ∞-operad

$N_p \operatorname{Fun}_{\mathrm{Glo}^{\mathrm{op}}}(\mathbf{OR}^{\otimes}, \mathcal{S}_*^{\wedge} \times (\mathrm{Glo}^{\mathrm{op}})^{\amalg})^{\mathrm{Day}}$ is given by sections of the cocartesian fibration $\pi_*\pi^*(\mathcal{S}_* \times \mathrm{Glo}^{\mathrm{op}})$, where by slight abuse of notation we write $\pi = U(\pi)$. Therefore we may calculate

$$\operatorname{Fun}_{/\mathrm{Glo}^{\mathrm{op}}}(\mathrm{Glo}^{\mathrm{op}}, \pi_*\pi^*(\mathcal{S}_* \times \mathrm{Glo}^{\mathrm{op}})) \simeq \operatorname{Fun}_{/\mathrm{Glo}^{\mathrm{op}}}(\mathbf{OR}, \mathcal{S}_* \times \mathrm{Glo}^{\mathrm{op}}) \simeq \operatorname{Fun}(\mathbf{OR}, \mathcal{S}_*),$$

using the definition of the left adjoints π^* and $\pi_!$. Now by Theorem 4.9 the partially lax limit of the diagram in question is given by the full subcategory of the left-most category spanned by those sections which map edges in $\mathrm{Orb}^{\mathrm{op}}$ to cocartesian arrows. We now apply [Lurie 2009, Corollary 3.2.2.13] (with $p\colon \mathbf{OR} \times_{\mathrm{Glo}^{\mathrm{op}}} \mathrm{Orb}^{\mathrm{op}} \to \mathrm{Orb}^{\mathrm{op}}$, $q\colon \mathcal{S}_* \times \mathrm{Orb}^{\mathrm{op}} \to \mathrm{Orb}^{\mathrm{op}}$ and $T = \pi_*\pi^*(\mathcal{S}_* \times \mathrm{Glo}^{\mathrm{op}}) \times_{\mathrm{Glo}^{\mathrm{op}}} \mathrm{Orb}^{\mathrm{op}})$ together with Lemma 9.13, to see that these sections corresponds to those functors in $\operatorname{Fun}_{/\mathrm{Glo}^{\mathrm{op}}}(\mathbf{OR}, \mathcal{S}_* \times \mathrm{Glo}^{\mathrm{op}})$ which send cartesian edges over $\mathrm{Orb}^{\mathrm{op}}$ to cocartesian edges of $\mathcal{S}_* \times \mathrm{Orb}^{\mathrm{op}} \to \mathrm{Orb}^{\mathrm{op}}$. These are exactly those maps which are equivalences in the first component, and therefore such sections correspond to functors $F\colon \mathbf{OR} \to \mathcal{S}_*$ which map cartesian edges over Orb to equivalences. Therefore we conclude by applying Lemma 9.14. □

Proposition 9.16 *There exists a functor* $\mathrm{PSp}_\bullet\colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_\infty^{\otimes}$ *sending* G *to* PSp_G. *Moreover, there is a symmetric monoidal equivalence*

$$\operatorname*{laxlim}^{\dagger}_{G \in \mathrm{Glo}^{\mathrm{op}}} \mathrm{PSp}_G \simeq \mathrm{Mod}_{S_{\mathrm{gl}}}(\mathbf{OR}_{\mathrm{gl}}\text{-}\mathcal{S}_*).$$

Proof There is a lax symmetric monoidal topologically enriched functor $S_{\mathrm{gl}}\colon \mathbf{OR}_{\mathrm{gl}} \to \mathcal{S}_*$ sending (G, V) to $(S^V)^G$. This induces a lax symmetric monoidal functor of ∞-operads, which uniquely specifies a commutative algebra in $\mathbf{OR}\text{-}\mathcal{S}_*$ by [Lurie 2017, Example 2.2.6.9], where we view $\mathbf{OR}_{\mathrm{gl}}\text{-}\mathcal{S}_*$ as a symmetric monoidal subcategory of $\mathbf{OR}\text{-}\mathcal{S}_*$ using Lemma 9.14. Applying Theorem 5.10 to the lax limit of Lemma 9.15 shows that there is a functor sending G to $\mathrm{Mod}_{S_G}(\mathbf{OR}_G\text{-}\mathcal{S}_*) \simeq \mathrm{PSp}_G$ (see Corollary 8.14), whose lax limit is $\mathrm{Mod}_{S_{\mathrm{gl}}}(\mathbf{OR}\text{-}\mathcal{S}_*)$.

Finally, we have to calculate the subcategory corresponding to the partially lax limit. Because the natural transformation $\mathrm{PSp}_G \to \mathbf{OR}_G\text{-}\mathcal{S}_*$ is pointwise conservative, we can check that an object lies in the partially lax limit of PSp_G by checking that its image lies in the partially lax limit of $\mathbf{OR}_G\text{-}\mathcal{S}_*$. In other words, we have a pullback square of symmetric monoidal ∞-categories

$$\begin{tikzcd}
\operatorname{laxlim}^{\dagger}_G \mathrm{PSp}_G \arrow[r] \arrow[d] & \operatorname{laxlim}_G \mathrm{PSp}_G \arrow[d] \\
\operatorname{laxlim}^{\dagger}_G \mathbf{OR}_G\text{-}\mathcal{S}_* \arrow[r] & \operatorname{laxlim}_G \mathbf{OR}_G\text{-}\mathcal{S}_*
\end{tikzcd}$$

Therefore, by Lemma 9.15 and the previous paragraph we have a symmetric monoidal equivalence

$$\operatorname*{laxlim}^{\dagger}_{G \in \mathrm{Glo}^{\mathrm{op}}} \mathrm{PSp}_G \simeq \mathrm{Mod}_{S_{\mathrm{gl}}}(\operatorname{Fun}(\mathbf{OR}, \mathcal{S}_*)) \times_{\operatorname{Fun}(\mathbf{OR}, \mathcal{S}_*)} \operatorname{Fun}(\mathbf{OR}_{\mathrm{gl}}, \mathcal{S}_*).$$

Finally, since $S_{\mathrm{gl}} \in \operatorname{Fun}(\mathbf{OR}_{\mathrm{gl}}, \mathcal{S}_*)$, this implies that

$$\operatorname*{laxlim}^{\dagger}_{G \in \mathrm{Glo}^{\mathrm{op}}} \mathrm{PSp}_G \simeq \mathrm{Mod}_{S_{\mathrm{gl}}}(\mathbf{OR}_{\mathrm{gl}}\text{-}\mathcal{S}_*). \qquad □$$

Notation 9.17 We write $\mathrm{PSp}^{\dagger}_{\mathrm{gl}}$ for the ∞-category $\mathrm{Mod}_{S_{\mathrm{gl}}}(\mathbf{OR}_{\mathrm{gl}}\text{-}\mathscr{S}_*)$, and identify it with $\mathrm{laxlim}^{\dagger}\,\mathrm{PSp}_\bullet$.

Recall the definition of the diagram $\mathscr{S}_\bullet\colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}^{\otimes}_{\infty}$ from Construction 6.15, which sends a group G to the ∞-category of G-spaces. We would like to construct a natural transformation $\Sigma^\infty\colon \mathscr{S}_\bullet \to \mathrm{PSp}_\bullet$, whose component at G is given by an analogue of the suspension prespectrum functor. Morally, this sends a G-space X to the S_G-module $(H,V) \mapsto (X \wedge S^V)^H$. We make this precise in the next construction. Let us first fix some notation: we write $\mathscr{S}_{\bullet,*}$ for the composite $(-)_* \circ \mathscr{S}_\bullet$ of $\mathscr{S}_\bullet$ with the functor which sends a presentably symmetric monoidal category to the ∞-category of pointed objects.

Construction 9.18 We will construct natural transformations of functors $\mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}^{\otimes}_{\infty}$

$$\mathscr{S}_\bullet \to \mathscr{S}_{\bullet,*} \to \mathrm{PSp}_\bullet.$$

The first natural transformation is simply given by postcomposing $\mathscr{S}_\bullet$ with the natural transformation $(-)_+\colon \mathrm{id} \to (-)_*$ of functors $(\mathrm{Pr}^{\mathrm{L}})^{\otimes} \to (\mathrm{Pr}^{\mathrm{L}})^{\otimes}$.

For the second natural transformation, we will construct it as a composite

$$\mathscr{S}_{\bullet,*} \to \mathbf{OR}_\bullet\text{-}\mathscr{S}_* \to \mathrm{PSp}_\bullet.$$

For the latter transformation $\mathbf{OR}_\bullet\text{-}\mathscr{S}_* \to \mathrm{PSp}_\bullet$, we simply note that the free module functors

$$S_G \otimes -\colon \mathbf{OR}_G\text{-}\mathscr{S}_* \to \mathrm{Mod}_{S_G}(\mathbf{OR}_G\text{-}\mathscr{S}_*) \simeq \mathrm{PSp}_G$$

are symmetric monoidal and fit into a natural transformation by the second half of Theorem 5.10.

For the first, it will be technically convenient to construct the natural transformation $\mathscr{S}^{\wedge}_{\bullet,*} \to \mathbf{OR}_\bullet\text{-}\mathscr{S}_*$ as a map of $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$-monoidal ∞-categories and then to use Proposition 5.5.

For this, we need to pin down the $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$-monoidal ∞-category which corresponds to $\mathscr{S}_{\bullet,*}$ under Proposition 5.5. Note that the map $t^{\mathrm{op}}\colon (\mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo})^{\mathrm{op}})^{\amalg} \to (\mathrm{Glo}^{\mathrm{op}})^{\amalg}$ exhibits $\mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo})^{\mathrm{op}}$ as a $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$-monoidal category; see Example 3.7. We claim that $\mathscr{S}_{\bullet,*}$ corresponds to the Day convolution

$$\mathrm{Fun}_{\mathrm{Glo}^{\mathrm{op}}}((\mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo})^{\mathrm{op}})^{\amalg}, \mathscr{S}^{\wedge}_{*} \times (\mathrm{Glo}^{\mathrm{op}})^{\amalg})^{\mathrm{Day}}.$$

To see this, we first note that

$$\mathrm{Fun}_{\mathrm{Glo}^{\mathrm{op}}}((\mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo})^{\mathrm{op}})^{\amalg}, \mathscr{S}^{\times} \times (\mathrm{Glo}^{\mathrm{op}})^{\amalg})^{\mathrm{Day}}$$

classifies $\mathscr{S}^{\times}_{\bullet}$, because it does so on underlying categories (combine Remark 3.11 and [Gepner et al. 2017, Proposition 7.3]) and the forgetful functor $\mathrm{Cat}^{\otimes}_{\infty} \to \mathrm{Cat}_\infty$ is faithful when restricted to cartesian monoidal ∞-categories. Now we observe that the $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$-monoidal functor

$$((-)_+)_*\colon \mathrm{Fun}_{\mathrm{Glo}^{\mathrm{op}}}(\mathrm{Ar}_{\mathrm{inj}}((\mathrm{Glo}^{\mathrm{op}})^{\amalg}, \mathscr{S}^{\times} \times (\mathrm{Glo}^{\mathrm{op}})^{\amalg})^{\mathrm{Day}} \to \mathrm{Fun}_{\mathrm{Glo}^{\mathrm{op}}}(\mathrm{Ar}_{\mathrm{inj}}((\mathrm{Glo}^{\mathrm{op}})^{\amalg}, \mathscr{S}^{\wedge}_{*} \times (\mathrm{Glo}^{\mathrm{op}})^{\amalg})^{\mathrm{Day}}$$

agrees pointwise with $(-)_+$, and therefore by the universal property of taking pointed objects (see [Lurie 2017, Proposition 4.8.2.11]), $\mathrm{Fun}_{\mathrm{Glo}^{\mathrm{op}}}(\mathrm{Ar}_{\mathrm{inj}}((\mathrm{Glo}^{\mathrm{op}})^{\amalg}, \mathscr{S}^{\wedge}_{*} \times (\mathrm{Glo}^{\mathrm{op}})^{\amalg})^{\mathrm{Day}}$ must classify $\mathscr{S}_{\bullet,*}$.

Now we can construct the $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$-monoidal functor which will induce $\mathscr{S}_{\bullet,*} \to \mathbf{OR}_{\bullet}\text{-}\mathscr{S}_*$. Pulling back the functor s_0 of Remark 9.4 along t^{op} we obtain a commutative diagram

$$\begin{tikzcd} \mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo})^{\mathrm{op}})^{\amalg} \arrow[rr, "s_{0,\mathrm{inj}}"] \arrow[dr, "t^{\mathrm{op}}"'] & & \mathbf{OR}^{\otimes} \arrow[dl, "\pi"] \\ & (\mathrm{Glo}^{\mathrm{op}})^{\amalg} & \end{tikzcd}$$

where t^{op} and π exhibit the sources as $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$-promonoidal ∞-categories by Lemma 9.11, so that $s_{0,\mathrm{inj}}$ is a map of $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$-promonoidal ∞-categories. One can then verify that $s_{0,\mathrm{inj}}$ satisfies the hypotheses of Proposition 3.34(a), and there exists a $(\mathrm{Glo}^{\mathrm{op}})^{\amalg}$-monoidal functor

$$(s_{0,\mathrm{inj}})_!\colon \mathrm{Fun}_{\mathrm{Glo}^{\mathrm{op}}}((\mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo})^{\mathrm{op}})^{\amalg}, \mathscr{S}_*^{\wedge} \times (\mathrm{Glo}^{\mathrm{op}})^{\amalg})^{\mathrm{Day}} \to \mathrm{Fun}_{\mathrm{Glo}^{\mathrm{op}}}(\mathbf{OR}^{\otimes}, \mathscr{S}_*^{\wedge} \times (\mathrm{Glo}^{\mathrm{op}})^{\amalg})^{\mathrm{Day}},$$

which then induces the required natural transformation. This description shows as well that the component at G coincides with $\mathscr{I}_0$, and so the composite functor $\mathscr{S}_{G,*} \to \mathrm{PSp}_G$ is analogous to the usual suspension prespectrum functor $F_0(-)$. We will formulate a precise statement to this effect as Proposition 10.5.

10 Functoriality of equivariant spectra

In the previous section we have constructed the functor

$$\mathrm{PSp}_{\bullet}\colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_{\infty}^{\otimes},$$

and calculated its partially lax limit. In this section we will show that this functor descends to a diagram $\mathrm{Sp}_{\bullet}$, where on every level we restrict to the subcategory of spectrum objects. Furthermore, we will prove that the functoriality obtained in this way agrees with the standard functoriality of equivariant spectra under the restriction–inflation functors. Finally, we will compute the partially lax limit of $\mathrm{Sp}_{\bullet}$ as a Bousfield localization of $\mathrm{PSp}_{\mathrm{gl}}^{\dagger} = \mathrm{laxlim}^{\dagger}\,\mathrm{PSp}_{\bullet}$.

Given a continuous group homomorphism $\alpha\colon H \to G$ between compact Lie groups, we write

$$\alpha^*\colon \mathbf{OR}_G\text{-}\mathscr{S}_* \to \mathbf{OR}_H\text{-}\mathscr{S}_*$$

for the symmetric monoidal functor induced by α. Our goals require a better understanding of α^*. We start by studying the interaction between α^* and the Quillen adjunction of Construction 7.6

$$\mathscr{I}_V\colon G\mathscr{T} \leftrightarrows \mathscr{I}\text{-}G\mathscr{T} :\mathrm{ev}_V$$

for a given G-representation V. However, before we do this, we first need to understand how these adjunctions manifest themselves under the equivalences

$$\mathscr{S}_G \simeq \boldsymbol{O}_G\text{-}\mathscr{S} \quad \text{and} \quad \mathbf{OR}_G\text{-}\mathscr{S} \simeq \mathscr{I}\text{-}G\mathscr{T}$$

of Example 3.40 and Theorem 8.9.

Remark 10.1 Consider $X \in \mathscr{I}\text{-}G\mathscr{T}$ and a G-representation V. Then the G-space $X(V)$ corresponds to the presheaf

$$G/H \mapsto X(V)^H \simeq \mathrm{Map}_{\mathscr{I}\text{-}G\mathscr{T}}(G \times_H \mathscr{I}_{V|_H}, X).$$

Note that $G \times_H \mathscr{I}_{V|_H}$ is the image of $(H, V|_H)$ under the embedding L of Theorem 8.9. Therefore, if we let $s_V : \boldsymbol{O}_G^{\mathrm{op}} \to \mathbf{OR}_G$ be the cocartesian section of π_G sending G/G to (G, V), we have $s(G/H) \simeq (H, V|_H)$, so we can identify ev_V with

$$s_V^* : \mathbf{OR}_G\text{-}\mathscr{S} \to \mathscr{S}_G, \quad X \mapsto X \circ s_V,$$

and similarly for the pointed version. It follows that the derived functor associated to $\mathscr{I}_V$ is given by the left Kan extension functor $(s_V)_!$. Finally, we can compute that this is given by

$$(\mathscr{I}_V X)(H, W) \simeq \mathscr{I}(V, W)^H \times X^H,$$

by the following lemma.

Lemma 10.2 *Let $\pi : \mathscr{E} \to \mathscr{B}$ be a cocartesian fibration of ∞-categories and $s : \mathscr{B} \to \mathscr{E}$ be a cocartesian section. For every functor $F : \mathscr{B} \to \mathscr{C}$ where $\mathscr{C}$ is a cocomplete ∞-category, we can compute the left Kan extension along s by*

$$(s_! F)(e) \simeq \mathrm{Map}_{\pi^{-1}(\pi e)}(s\pi(e), e) \times F(\pi(e)) \quad \textit{for all } e \in E.$$

Proof By the usual formula for left Kan extensions we have that

$$(s_! F)(e) \simeq \operatorname*{colim}_{b \in \mathscr{B} \times_{\mathscr{E}} \mathscr{E}_{/e}} F(b).$$

We claim that the projection $\mathscr{B} \times_{\mathscr{E}} \mathscr{E}_{/e} \to \mathscr{B}_{/\pi e}$ is a left fibration with fiber over $f : b \to \pi e$ given by $\mathrm{Map}^f_{\mathscr{E}}(s(b), e)$. In particular, since F is constant along the fibers of this fibration and $\mathscr{B}_{/\pi e}$ has a final object, we have

$$\operatorname*{colim}_{b \in \mathscr{B} \times_{\mathscr{E}} \mathscr{E}_{/e}} F(b) \simeq \operatorname*{colim}_{[f : b \to \pi e] \in \mathscr{B}_{/\pi e}} \mathrm{Map}^f_{\mathscr{E}}(s(b), e) \times F(b) \simeq \mathrm{Map}_{\pi^{-1}(\pi e)}(s\pi(e), e) \times F(\pi(e)).$$

It only remains to prove that the functor $\mathscr{B} \times_{\mathscr{E}} \mathscr{E}_{/e} \to \mathscr{B}_{/\pi e}$ is a left fibration. That is, we need to show that for every diagram

$$\begin{tikzcd} \Lambda^n_i \arrow[r] \arrow[d] & \mathscr{B} \times_{\mathscr{E}} \mathscr{E}_{/e} \arrow[d] \\ \Delta^n \arrow[r] \arrow[ur, dashed] & \mathscr{B}_{/\pi e} \end{tikzcd}$$

with $0 \le i < n$, there exists a dotted arrow completing the diagram. Using the definition of slice ∞-categories, this is equivalent to finding a dotted arrow completing the dotted diagram

$$\begin{tikzcd} \Lambda^n_i \star \Delta^0 \arrow[r, "F"] \arrow[d] & \mathscr{E} \arrow[d, "\pi"] \\ \Delta^n \star \Delta^0 \simeq \Delta^{n+1} \arrow[r, "G"] \arrow[ur, dashed] & \mathscr{B} \end{tikzcd}$$

where F restricted to $\Lambda^n_i \subseteq \Lambda^{n+1}_i$ is given by the restriction of sG. This diagram is a diagram of marked simplicial sets when we give $\mathscr{B}$ the total marking, $\mathscr{E}$ the cocartesian marking and on the left column the marking $(\Lambda^n_i)^\sharp \star \Delta^0 \to (\Delta^n)^\sharp \star \Delta^0$. Since the left vertical arrow is left marked anodyne by [Shah 2023, Lemma 4.10], the lift exists. $\square$

Having understood the adjunction $\mathscr{I}_V \dashv \mathrm{ev}_V$, we now discuss how this interacts with the functor α^*.

Proposition 10.3 *Let us fix an arrow $\alpha\colon H \to G$ in* Glo*.*

(1) *Given a pointed G-space X, there is a natural equivalence*

$$\alpha^*\mathscr{I}_V X \simeq \mathscr{I}_{\alpha^*V}(\alpha^*X).$$

(2) *Given a pointed $\mathbf{OR}_G$-space Y, there is a natural equivalence*

$$\alpha^*\,\mathrm{ev}_V\, Y \simeq \mathrm{ev}_{\alpha^*V}\,\alpha^*Y.$$

(3) *Under the two previous identifications, the counit natural transformation*

$$\mathscr{I}_V\,\mathrm{ev}_V\, X \to X$$

is sent by α^* to

$$\mathscr{I}_{\alpha^*V}\,\mathrm{ev}_{\alpha^*}\, V(\alpha^*X) \to \alpha^*X,$$

the counit natural transformation for α^*V applied to α^*X.

Proof Write $\boldsymbol{O}_\alpha \simeq \mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo}) \times_{\mathrm{Glo}} [1]$ (using the target map $t\colon \mathrm{Ar}_{\mathrm{inj}}(\mathrm{Glo}) \to \mathrm{Glo}$) and let $i_0\colon \boldsymbol{O}_H \to \boldsymbol{O}_\alpha$ and $i_1\colon \boldsymbol{O}_G \to \boldsymbol{O}_\alpha$ be the inclusions of the fibers over 0 and 1, respectively. Similarly, write $\mathbf{OR}_\alpha := \mathbf{OR}_{\mathrm{gl}} \times_{\mathrm{Glo}^{\mathrm{op}}} \boldsymbol{O}_\alpha^{\mathrm{op}}$ and $j_0, j_1\colon \mathbf{OR}_H, \mathbf{OR}_G \to \mathbf{OR}_\alpha$ for the inclusion of the fiber of $\mathbf{OR}_\alpha \to [1]^{\mathrm{op}}$ over 0 and 1, respectively. Therefore by Remark 3.23 we can identify

$$\alpha^* \simeq i_0^*(i_1)_!\colon \mathscr{S}_{G,*} \to \mathscr{S}_{H,*} \quad \text{and} \quad \alpha^* \simeq j_0^*(j_1)_!\colon \mathbf{OR}_G\text{-}\mathscr{S}_* \to \mathbf{OR}_H\text{-}\mathscr{S}_*.$$

Let $s_V\colon \boldsymbol{O}_G^{\mathrm{op}} \to \mathbf{OR}_G$ be the cocartesian section of $\pi_G\colon \mathbf{OR}_G \to \boldsymbol{O}_G^{\mathrm{op}}$ which sends G/G to (G,V). Similarly let $s\colon \boldsymbol{O}_\alpha^{\mathrm{op}} \to \mathbf{OR}_\alpha$ be the cocartesian section sending the initial object $i_1(G/G)$ of $\boldsymbol{O}_\alpha^{\mathrm{op}}$ to $j_1(G,V)$. Then s restricts to s_V on $\boldsymbol{O}_G^{\mathrm{op}}$ and to s_{α^*V} on $\boldsymbol{O}_H^{\mathrm{op}}$, since a cocartesian section is determined by where it sends the initial object. Therefore by Remark 10.1 we obtain

$$\alpha^*\mathscr{I}_V X \simeq \alpha^*(s_V)_!X \simeq j_0^*(j_1 s_V)_!X \simeq j_0^* s_!(i_1)_!X$$

for every pointed G-space X. Using the formula for $s_!$ described in Lemma 10.2 we see that the above can be identified with $(s_{\alpha^*V})_! i_0^*(i_1)_!X$, thus proving the first statement.

Now let Y be an $\mathbf{OR}_G$-space. Then we claim that $s^*(j_1)_!Y$ is left Kan extended from $\boldsymbol{O}_G^{\mathrm{op}}$. In fact this happens if and only if $s^*(j_1)_!Y$ sends the arrows $(G,\alpha L) \to (H,L)$ in $\boldsymbol{O}_\alpha^{\mathrm{op}}$ to equivalences. But the arrow

$$[s(G,\alpha L) \to s(H,L)] \simeq [(G,\alpha L, V) \to (H, L, \alpha^*V)]$$

is a terminal object of $\mathbf{OR}_G \times_{\mathbf{OR}_\alpha} (\mathbf{OR}_\alpha)_{/(H,L,\alpha^*V)}$ and so it is sent to an equivalence by $(j_1)_!Y$. This implies that

$$\mathrm{ev}_{\alpha^*V}\,\alpha^*Y \simeq s_{\alpha^*V}^*(j_1)_!Y \simeq j_0^*s^*(j_1)_!Y \simeq j_0^*(j_1)_!(s_V)^*Y \simeq \alpha^*\,\mathrm{ev}_V\,Y,$$

proving the second statement.

Finally we consider for every $\mathbf{OR}_G$-space Y, the natural transformation

$$s_!s^*(j_1)_!Y \to (j_1)_!Y,$$

and note that this is a natural transformation of functors left Kan extended from $\mathbf{OR}_G$, which restricts to

$$(s_V)_!s_V^*Y \to Y \text{ and } (s_{\alpha^*V})_!s_{\alpha^*V}^*\alpha^*Y \to \alpha^*Y$$

on the fibers over 0 and 1, respectively. Thus α^* sends the former to the latter, showing the third statement. $\square$

With this result we can show that $\mathrm{PSp}_\bullet$ restricts to a functor on spectrum objects.

Proposition 10.4 *There exists a functor* $\mathrm{Sp}_\bullet\colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_\infty^\otimes$ *and a natural transformation of functors*

$$L_\bullet\colon \mathrm{PSp}_\bullet \to \mathrm{Sp}_\bullet,$$

whose component for a fixed G *is the spectrification functor* $L_G\colon \mathrm{PSp}_G \to \mathrm{Sp}_G$.

Proof Consider a group homomorphism $\alpha\colon H \to G$. We claim that the functor $\mathrm{PSp}_\alpha\colon \mathrm{PSp}_G \to \mathrm{PSp}_H$ preserves stable equivalences. It suffices to show that it preserves the generating equivalences $G \rtimes_K \lambda_{V,W}$ of Proposition 7.30. Moreover, since G is compact, we can restrict to the cofinal set W of K-representations that are extended from G.

First note that $\lambda_{V,W} \simeq (G \rtimes_K F_V(S^0)) \otimes \lambda_{0,W}$. Since PSp_α is symmetric monoidal by construction and stable equivalences are stable under tensor product, it suffices to show that $\mathrm{PSp}_\alpha(\lambda_{0,W})$ is a stable equivalence. We claim it is equivalent to λ_{0,α^*W}. In fact, $\lambda_{0,W}$ is exactly the counit of the adjunction $F_W \dashv \mathrm{ev}_W$ of Construction 7.28 applied to S_G. Therefore we can factor it as

$$(F_W\,\mathrm{ev}_W)S_G \simeq (S_G \otimes -)\mathcal{I}_W\,\mathrm{ev}_W\,US_G \to (S_G \otimes -)US_G \to S_G,$$

where $(S_G \otimes -) \dashv U$ is the free-forgetful adjunction between PSp_G and $\mathbf{OR}_G$-$\mathcal{S}_*$, and the arrows are the counits of the respective adjunctions. Then our claim follows from Theorem 5.10 and Proposition 10.3.

Knowing that PSp_α preserves stable equivalences, we can combine Construction 9.18 and Corollary 4.14 to obtain $\mathrm{Sp}_\bullet$ and the natural transformation $L_\bullet\colon \mathrm{PSp}_\bullet \to \mathrm{Sp}_\bullet$. $\square$

Recall that we constructed a natural transformation $\Sigma_\bullet^\infty\colon \mathcal{S}_{\bullet,*} \to \mathrm{PSp}_\bullet$ in Construction 9.18, which pointwise was our analogue of the suspension prespectrum functor. We may compose this with the natural transformation $L_\bullet$ to obtain a new natural transformation, which we again denote by $\Sigma_\bullet^\infty$.

Proposition 10.5 *The component of $\Sigma^\infty_\bullet\colon \mathscr{S}_{\bullet,*} \to \mathrm{Sp}_\bullet$ at the group G is equivalent to the standard suspension spectrum functor.*

Proof Considering the component at G, we observe that the functor Σ^∞_G is defined as the composition

$$\mathscr{S}_{G,*} \to \mathbf{OR}_G\text{-}\mathscr{S}_* \to \mathrm{Mod}_{S_G}(\mathbf{OR}_G\text{-}\mathscr{S}_*) \simeq \mathrm{PSp}_G \to \mathrm{Sp}_G,$$

where the first functor is $\mathscr{I}_0$ (ie precomposition along $\mathbf{OR}_G \to \boldsymbol{O}_G^{\mathrm{op}}$), the second functor is the free S_G-module functor ($S_G \otimes -$) and the third functor is the localization functor. These functors are all modeled by left Quillen functors

$$G\mathscr{T}_* \to \mathscr{I}\text{-}G\mathscr{T}_* \to \mathrm{Sp}^O_G \to \mathrm{Sp}^O_G$$

given by the constant $\mathscr{I}$-G-space, the free S_G-module and the identity, respectively. Therefore Σ^∞_G is modeled by their composition, which is exactly the suspension spectrum functor constructed in [Mandell and May 2002]. □

This suffices for us to conclude that the functoriality of $\mathrm{Sp}_\bullet$ agrees morphismwise with the functoriality of equivariant spectra in restriction, by the universal property of G-spectra.

Corollary 10.6 *The functor $\mathrm{Sp}_\bullet\colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}^\otimes_\infty$ sends a compact Lie group G to Sp_G, and a continuous group homomorphism $\alpha\colon H \to G$ to the restriction functor $\alpha^*\colon \mathrm{Sp}_G \to \mathrm{Sp}_H$.*

Proof Consider the commutative diagram

$$\begin{tikzcd}
\mathrm{Sp}_G \arrow[r, "\mathrm{Sp}_\alpha"] & \mathrm{Sp}_H \\
\mathscr{S}_{G,*} \arrow[u, "\Sigma^\infty_G"] \arrow[r, "\mathscr{S}^\alpha_*"] & \mathscr{S}_{H,*} \arrow[u, "\Sigma^\infty_H"'] \\
\mathscr{S}_G \arrow[u, "(-)_+"] \arrow[r, "\mathscr{S}^\alpha"] & \mathscr{S}_H \arrow[u, "(-)_+"']
\end{tikzcd}$$

of symmetric monoidal functors. By the universal property of G-spectra [Gepner and Meier 2023, Corollary C.7], the functor Sp_α is uniquely determined by $\mathscr{S}_{\alpha,*}$, and this is completely determined by $\mathscr{S}_\alpha$ by [Lurie 2017, Proposition 4.8.2.11]. Finally, Proposition 6.16 identifies the functor $\mathscr{S}^\alpha$ with α^*. □

Remark 10.7 The argument of Corollary 10.6 in fact shows that the natural transformation $\Sigma_{\bullet,*}\colon \mathscr{S}_{\bullet,*} \to \mathrm{Sp}_\bullet$ admits a universal property. This forces $\mathrm{Sp}_\bullet$ to coincide with the construction of [Bachmann and Hoyois 2021, Section 9] on the subcategory of Glo spanned by finite groups. This suggests a possible comparison between ultracommutative Fin-global ring spectra in the sense of [Schwede 2018] and normed spectra in the sense of [Bachmann and Hoyois 2021].

We have now constructed $\mathrm{Sp}_\bullet$ and shown that it agrees with the standard functoriality of equivariant spectra. We will write $\mathrm{Sp}^\dagger_{\mathrm{gl}}$ for the partially lax limit $\mathrm{laxlim}^\dagger \mathrm{Sp}_\bullet$. We would like to describe $\mathrm{Sp}^\dagger_{\mathrm{gl}}$ as a Bousfield localization of $\mathrm{PSp}^\dagger_{\mathrm{gl}}$ by applying Lemma 4.13. To do this requires the following two lemmas.

Proposition 10.8 *Let $\alpha\colon H \to G$ be an injective group homomorphism. Then the functor $\alpha^*\colon \mathbf{OR}_G\text{-}\mathscr{S} \to \mathbf{OR}_H\text{-}\mathscr{S}$ has a left adjoint $\alpha_!$. Moreover, under the identification of Theorem 8.9, the adjunction $\alpha_! \dashv \alpha^*$ corresponds to the Quillen adjunction $G \times_H - \dashv \alpha^*$ of Proposition 7.10.*

In particular, for $X \in \mathbf{OR}_H\text{-}\mathscr{S}$ and $Y \in \mathbf{OR}_G\text{-}\mathscr{S}$, the comparison map

$$\alpha_!(X \otimes \alpha^* Y) \to \alpha_! X \otimes Y$$

adjoint to $X \otimes \alpha^* Y \to \alpha^*\alpha_! X \otimes \alpha^* Y$ is an equivalence.

Proof By the description of Remark 3.23 and Lemma 9.13 it follows that $\alpha^*\colon \mathbf{OR}_H\text{-}\mathscr{S} \to \mathbf{OR}_G\text{-}\mathscr{S}$ is given by precomposition along the functor $p_\alpha\colon \mathbf{OR}_H \to \mathbf{OR}_G$ obtained by basechange from $\boldsymbol{O}_H^{\mathrm{op}} \to \boldsymbol{O}_G^{\mathrm{op}}$. In particular, it has a left adjoint $\alpha_!$ given by left Kan extension along p_α.

In the proof of Theorem 8.9 we have constructed a functor $L_H\colon \mathbf{OR}_H^{\mathrm{op}} \to \mathscr{I}\text{-}H\mathscr{T}[W_{\mathrm{lvl}}^{-1}]$ sending (K, W) to $H \times_K \mathscr{I}_W$. We claim that there is a commutative diagram

$$\begin{tikzcd}
\mathbf{OR}_H^{\mathrm{op}} \arrow[r, "\text{Yoneda}"] \arrow[d, "p_\alpha"'] \arrow[rr, bend left, "L_H"] & \mathbf{OR}_H\text{-}\mathscr{S} \arrow[r, "\sim"] \arrow[d, "\alpha_!"] & \mathscr{I}\text{-}H\mathscr{T}[W_{\mathrm{lvl}}^{-1}] \arrow[d, "G\times_H -"] \\
\mathbf{OR}_G^{\mathrm{op}} \arrow[r, "\text{Yoneda}"'] \arrow[rr, bend right, "L_G"'] & \mathbf{OR}_G\text{-}\mathscr{S} \arrow[r, "\sim"] & \mathscr{I}\text{-}G\mathscr{T}[W_{\mathrm{lvl}}^{-1}]
\end{tikzcd}$$

where the horizontal equivalences are given by Theorem 8.9. The diagram on the left commutes by the universal property of presheaf categories and the outer square commutes by direct verification using the formulas of L_G and L_H. Therefore a generation argument, using that all the functors preserve colimits, shows that the rightmost diagram commutes too. The rightmost vertical functor can be modeled by a left Quillen functor by Proposition 7.10, so the first claim follows.

Finally, since the map

$$G \times_H (X \otimes Y) \to (G \times_H X) \otimes Y$$

is an isomorphism in $\mathscr{I}\text{-}G\mathscr{T}$, it follows that the derived formula holds as well. $\square$

Lemma 10.9 *Let $\alpha\colon H \to G$ be an injective homomorphism of compact Lie groups. Then $\mathrm{PSp}_\alpha\colon \mathrm{PSp}_G \to \mathrm{PSp}_H$ sends Sp_G into Sp_H.*

Proof PSp_α sends X to $S_H \otimes_{\alpha^* S_G} \alpha^* X \simeq \alpha^* X$, since α is injective. Therefore PSp_α preserves all small limits and colimits, since α^* does, and so it has a left adjoint L_α. Moreover, by Proposition 10.8 there is an equivalence

$$L_\alpha(X \otimes \mathrm{PSp}_\alpha Y) \simeq L_\alpha(X) \otimes Y.$$

To prove that $\alpha^*(\mathrm{Sp}_G) \subseteq \mathrm{Sp}_H$ it suffices to show that L_α preserves stable equivalences. By cofinality the stable equivalences in Sp_H are generated by those of the form $H \times_M \lambda_{V,W|_M}$, where $M < H$ is a closed subgroup, V is an M-representation and W is a G-representation. But then

$$L_\alpha(H \times_M \lambda_{V,W|_M}) \simeq L_\alpha((H \times_M F_V S^0) \otimes \alpha^* \lambda_{0,W|_H}) \simeq L_\alpha(H \times_M F_V S^0) \otimes \lambda_{0,W}.$$

Since stable equivalences are stable under tensoring and $\lambda_{0,W}$ is a stable equivalence, this proves the thesis. □

Given a compact Lie group $G \in \mathrm{Glo}$, we denote by $U_G^{\mathrm{gl}} \colon \mathrm{PSp}_{\mathrm{gl}}^{\dagger} \to \mathrm{PSp}_G$ the canonical functors associated to the universal cone.

Proposition 10.10 *The ∞-category $\mathrm{Sp}_{\mathrm{gl}}^{\dagger}$ is a Bousfield localization of $\mathrm{PSp}_{\mathrm{gl}}^{\dagger}$. We denote the associated left adjoint by $L_{\mathrm{gl}} \colon \mathrm{PSp}_{\mathrm{gl}}^{\dagger} \to \mathrm{Sp}_{\mathrm{gl}}^{\dagger}$. Furthermore, the following conditions are equivalent for an object $X \in \mathrm{PSp}_{\mathrm{gl}}^{\dagger}$:*

(a) *X is in $\mathrm{Sp}_{\mathrm{gl}}^{\dagger}$.*
(b) *For every compact Lie group G, the G-prespectrum $U_G^{\mathrm{gl}}(X)$ is in Sp_G.*
(c) *For every compact Lie group G, the G-prespectrum $U_G^{\mathrm{gl}} X$ is local with respect to the maps $\lambda_{V,W}$ defined in Construction 7.28 for any G-representations V and W.*

Proof Recall that $\mathrm{Sp}_\bullet$ was constructed in Proposition 10.4 by localizing the functor $\mathrm{PSp}_\bullet$ using Lemma 4.13. By the same lemma together with Lemma 10.9, we conclude that $\mathrm{Sp}_{\mathrm{gl}}^{\dagger}$ is a Bousfield localization and that conditions (a) and (b) are equivalent. By Proposition 7.30, condition (b) is equivalent to the condition that for every compact Lie group G and closed subgroup $H \leq G$, the H-prespectrum $\mathrm{res}_H^G U_G^{\mathrm{gl}} X$ is local with respect to the maps $\{\lambda_{V,W}\}$, where V and W vary over all H-representations. By construction we have $U_H^{\mathrm{gl}} = \mathrm{res}_H^G \circ U_G^{\mathrm{gl}}$, so (b) and (c) are equivalent. □

11 Global spectra as a partially lax limit

Recall the functors $\mathrm{PSp}_\bullet, \mathrm{Sp}_\bullet \colon \mathrm{Glo}^{\mathrm{op}} \to \mathrm{Cat}_\infty^{\otimes}$ constructed in Propositions 9.16 and 10.4. We also defined

$$\mathrm{PSp}_{\mathrm{gl}}^{\dagger} := \operatorname*{laxlim}^{\dagger}_{\mathrm{Glo}^{\mathrm{op}}} \mathrm{PSp}_G \quad \text{and} \quad \mathrm{Sp}_{\mathrm{gl}}^{\dagger} := \operatorname*{laxlim}^{\dagger}_{\mathrm{Glo}^{\mathrm{op}}} \mathrm{Sp}_G.$$

The goal of this section is to show that $\mathrm{Sp}_{\mathrm{gl}}^{\dagger}$ is symmetric monoidally equivalent to Schwede's ∞-category of global spectra $\mathrm{Sp}_{\mathrm{gl}}$, whose definition is recalled in Definition 7.23. Our proof will go roughly as follows:

- We will first construct a symmetric monoidal adjunction
$$j_!\colon \mathrm{PSp}_{\mathrm{fgl}} \simeq \mathrm{Mod}_{S_{\mathrm{fgl}}}(\mathbf{OR}_{\mathrm{fgl}}\text{-}\mathcal{S}_*) \leftrightarrows \mathrm{Mod}_{S_{\mathrm{gl}}}(\mathbf{OR}_{\mathrm{gl}}\text{-}\mathcal{S}_*) \simeq \mathrm{PSp}_{\mathrm{gl}}^{\dagger} : j^*$$
between prespectra objects, where the equivalences are given by Proposition 9.16 and Corollary 8.23.
- We note that there are Bousfield localizations $\mathrm{Sp}_{\mathrm{gl}} \subset \mathrm{PSp}_{\mathrm{fgl}}$ and $\mathrm{Sp}_{\mathrm{gl}}^{\dagger} \subset \mathrm{PSp}_{\mathrm{gl}}^{\dagger}$. We denote by $L_{\mathrm{gl}}\colon \mathrm{PSp}_{\mathrm{gl}}^{\dagger} \to \mathrm{Sp}_{\mathrm{gl}}^{\dagger}$ the localization functor.
- We will then check that j^* preserves spectrum objects, and therefore obtain an induced adjunction
$$L_{\mathrm{gl}} \circ j_!\colon \mathrm{Sp}_{\mathrm{gl}} \leftrightarrows \mathrm{Sp}_{\mathrm{gl}}^{\dagger} : j^*$$
between the respective localizations.
- We will show that this adjunction is in fact an equivalence, by showing that j^* is conservative on spectrum objects, and that the unit of the adjunction $(L_{\mathrm{gl}} \circ j_!, j^*)$ is an equivalence.

We start by constructing an adjunction between prespectrum objects. By Lemma 9.3 we can identify $\mathbf{OR}_{\mathrm{fgl}}$ with the full subcategory of $\mathbf{OR}_{\mathrm{gl}}$ spanned by (G, V), where V is a faithful G-representation. Then the canonical inclusion $j : \mathbf{OR}_{\mathrm{fgl}} \hookrightarrow \mathbf{OR}_{\mathrm{gl}}$ induces an adjunction
$$j_!\colon \mathbf{OR}_{\mathrm{fgl}}\text{-}\mathcal{S}_* \leftrightarrows \mathbf{OR}_{\mathrm{gl}}\text{-}\mathcal{S}_* : j^*.$$
Note that $j_!$ is fully faithful as it is given as a left Kan extension along a fully faithful functor. Moreover the functor $j_!$ is strong monoidal by Proposition 3.34.

Proposition 11.1 *The inclusion $j : \mathbf{OR}_{\mathrm{fgl}} \hookrightarrow \mathbf{OR}_{\mathrm{gl}}$ admits a right adjoint q, which is given on objects by*
$$(G, V) \mapsto (G/\ker(V), V),$$
where $\ker(V) < G$ is the subgroup of $g \in G$ acting trivially on V. In particular, the left Kan extension $j_!$ is equivalent to the functor q^* given by precomposition by q.

Proof The $G/\ker(V)$-representation V is clearly faithful, so to prove the thesis it is enough to show that for every $(H, W) \in \mathbf{OR}_{\mathrm{fgl}}$, the map $(G/\ker(V), V) \to (G, V)$ induces an equivalence on mapping spaces
$$\mathrm{Map}_{\mathbf{OR}_{\mathrm{gl}}}((H, W), (G/\ker V)) \xrightarrow{\sim} \mathrm{Map}_{\mathbf{OR}_{\mathrm{gl}}}((H, W), (G, V)).$$
By Definition 9.2, this means we need to show that the map
$$(\mathrm{Hom}(G/\ker V, H) \times \mathcal{I}(W, V))_{hH}^{G/\ker V} \to (\mathrm{Hom}(G, H) \times \mathcal{I}(W, V))_{hH}^{G}$$
given by precomposition with $G \to G/\ker V$ on the first coordinate, is a homotopy equivalence. In fact we will show that
$$(\mathrm{Hom}(G/\ker V, H) \times \mathcal{I}(W, V))^{G/\ker V} \to (\mathrm{Hom}(G, H) \times \mathcal{I}(W, V))^{G}$$

is a homeomorphism. Since it is a continuous map of compact Hausdorff topological spaces, it suffices to show that it is bijective. As $\mathrm{Hom}(G/\ker V, H) \to \mathrm{Hom}(G, H)$ is injective, so is the above map. Therefore to conclude we need to show it is surjective.

Concretely this means that if we have a map $\alpha\colon G \to H$ and an isometry $\varphi\colon W \to V$ that is G-equivariant, we need to show that α is trivial when restricted to $\ker V$. But if $g \in \ker V$, then g acts as the identity on V, and therefore $\alpha(g)$ acts as the identity on W (since φ is G-equivariant). Since W is a faithful H-representation this implies that $\alpha(g) = 1$, as required. $\square$

Note that it is clear from the definitions that $j^*S_{\mathrm{gl}} \simeq S_{\mathrm{fgl}}$ as commutative algebra objects. As an application of the previous proposition we find:

Corollary 11.2 *The counit map $\epsilon\colon j_!S_{\mathrm{fgl}} \to S_{\mathrm{gl}}$ is an equivalence of commutative algebra objects. In particular, the functors $j_! \dashv j^*$ induce an adjunction*

$$j_!\colon \mathrm{PSp}_{\mathrm{fgl}} \simeq \mathrm{Mod}_{S_{\mathrm{fgl}}}(\mathbf{OR}_{\mathrm{fgl}}\text{-}\mathcal{S}_*) \leftrightarrows \mathrm{Mod}_{S_{\mathrm{gl}}}(\mathbf{OR}_{\mathrm{gl}}\text{-}\mathcal{S}_*) \simeq \mathrm{PSp}^{\dagger}_{\mathrm{gl}} : j^*.$$

Proof Because j is strong monoidal, the counit is canonically a map of commutative algebra objects. Therefore for all $(G, V) \in \mathbf{OR}_{\mathrm{gl}}$ we compute

$$j_!(S_{\mathrm{fgl}})(G, V) \simeq S_{\mathrm{fgl}}(q(G, V)) = (S^V)^{G/\ker(V)} \simeq (S^V)^G = S_{\mathrm{gl}}(G, V).$$

Because $j_!$ and j^* are strong and lax monoidal, respectively, and they swap the two algebra objects, they induce functors as in the statement, which are evidently adjoint. $\square$

We will now use the adjunction

$$j_!\colon \mathrm{PSp}_{\mathrm{fgl}} \leftrightarrows \mathrm{PSp}^{\dagger}_{\mathrm{gl}} : j^*$$

to induce an adjunction at the level of spectrum objects. To do this we need to see how the adjunction $(j_!, j^*)$ interacts with the full subcategories of spectrum objects. To this end we briefly rephrase the discussion of local objects in $\mathrm{PSp}_{\mathrm{fgl}}$ given at the end of Section 7.

Remark 11.3 Recall from Proposition 7.27 that $\mathrm{Sp}_{\mathrm{gl}}$ is a Bousfield localization of $\mathrm{PSp}_{\mathrm{fgl}}$ at the morphisms $\{\lambda_{G,V,W}\}$ where G is a compact Lie group and V and W are G-representations with W faithful. Because $j_!\colon \mathrm{PSp}_{\mathrm{fgl}} \to \mathrm{PSp}^{\dagger}_{\mathrm{gl}}$ is fully faithful, we can equivalently require that $j_!X$ is local with respect to the maps $j_!(\lambda_{G,V,W})$, where W is a faithful representation. These maps again corepresent the G-fixed points of the adjoint structure map $\tilde{\sigma}_{G,V,W}$, and therefore we will denote them by $\lambda^{\dagger}_{G,V,W}$, and similarly we will write $F^{\dagger}_{G,V}$ for $j_!F_{G,V}$.

We have seen in Construction 7.25 that for any compact Lie group G and G-representation V, there is a functor $\mathrm{ev}_{G,V}\colon \mathrm{PSp}_{\mathrm{fgl}} \to \mathcal{S}_{G,*}$ that sends a faithful global prespectrum X to the G-space $X(V)$. Under the equivalence

$$\mathrm{PSp}_{\mathrm{fgl}} \simeq \mathrm{Mod}_{S_{\mathrm{fgl}}}(\mathbf{OR}_{\mathrm{fgl}}\text{-}\mathcal{S}_*),$$

this functor can be modeled as follows. Consider the cocartesian section $s_V\colon \boldsymbol{O}_G^{\mathrm{op}} \to \mathbf{OR}_G$ which is determined by the object $(G,V)\in\mathbf{OR}_G$, and write k_V for the composite $\boldsymbol{O}_G^{\mathrm{op}} \xrightarrow{s_V} \mathbf{OR}_G \xrightarrow{\nu_G} \mathbf{OR}_{\mathrm{gl}}$. If V is faithful then k_V lands in $\mathbf{OR}_{\mathrm{fgl}}$ and so we can define $\mathrm{ev}_{G,V}$ as the composite of right adjoints

$$\mathrm{Mod}_{S_{\mathrm{fgl}}}(\mathbf{OR}_{\mathrm{fgl}}\text{-}\mathcal{S}_*) \xrightarrow{\mathrm{fgt}} \mathbf{OR}_{\mathrm{fgl}}\text{-}\mathcal{S}_* \xrightarrow{k_V^*} \mathcal{S}_{G,*}.$$

Similarly, as discussed in Construction 7.28, there is a functor $\mathrm{ev}_V\colon \mathrm{PSp}_G \to \mathcal{S}_{G,*}$ sending a G-prespectrum X to the G-space $X(V)$. Under the equivalence

$$\mathrm{PSp}_G \simeq \mathrm{Mod}_{S_G}(\mathbf{OR}_G\text{-}\mathcal{S}_*)$$

this functor is modeled by the composite

$$\mathrm{Mod}_{S_G}(\mathbf{OR}_G\text{-}\mathcal{S}_*) \xrightarrow{\mathrm{fgt}} \mathbf{OR}_G\text{-}\mathcal{S}_* \xrightarrow{s_V^*} \mathcal{S}_{G,*}.$$

See also Remark 10.1.

Remark 11.4 From the previous discussion we conclude that there is a commutative diagram of right adjoints

$$\begin{tikzcd}
\mathrm{PSp}_{\mathrm{fgl}} \arrow[d, "\sim"'] & \mathrm{PSp}^{\dagger}_{\mathrm{gl}} \arrow[l, "j^*"'] \arrow[r, "U_G^{\mathrm{gl}}"] \arrow[d, "\sim"] & \mathrm{PSp}_G \arrow[d, "\sim"] \\
\mathrm{Mod}_{S_{\mathrm{fgl}}}(\mathbf{OR}_{\mathrm{fgl}}\text{-}\mathcal{S}_*) \arrow[d, "\mathrm{fgt}"'] & \mathrm{Mod}_{S^{\dagger}_{\mathrm{gl}}}(\mathbf{OR}_{\mathrm{gl}}\text{-}\mathcal{S}_*) \arrow[l, "j^*"'] \arrow[r, "\nu_G^*"] \arrow[d, "\mathrm{fgt}"] & \mathrm{Mod}_{S_G}(\mathbf{OR}_G\text{-}\mathcal{S}_*) \arrow[d, "\mathrm{fgt}"] \\
\mathbf{OR}_{\mathrm{fgl}}\text{-}\mathcal{S}_* \arrow[dr, "k_W^*"'] & \mathbf{OR}_{\mathrm{gl}}\text{-}\mathcal{S}_* \arrow[l, "j^*"'] \arrow[r, "\nu_G^*"] \arrow[d, "k_W^*"] & \mathbf{OR}_G - \mathcal{S}_* \arrow[dl, "s_W^*"] \\
& \mathcal{S}_{G,*} &
\end{tikzcd}$$

Using that the corresponding diagram of left adjoints commute, we see that for all $X\in\mathrm{PSp}^{\dagger}_{\mathrm{gl}}$ and G-representations V and W with W faithful, the diagram

$$\begin{tikzcd}
\mathcal{S}_{G,*}(S^0, X(W)) \arrow[r, "\tilde{\sigma}_{V,W}"] & \mathcal{S}_{G,*}(S^V, X(V\oplus W)) \arrow[d, "\sim"] \\
\mathrm{PSp}_G(F_W S^0, U_G^{\mathrm{gl}}(X)) \arrow[u, "\sim"] \arrow[r, "\lambda^*_{V,W}"] & \mathrm{PSp}_G(F_{V\oplus W} S^V, U_G^{\mathrm{gl}}(X)) \arrow[d, "\sim"] \\
\mathrm{PSp}_{\mathrm{fgl}}(F_{G,W} S^0, j^*X) \arrow[u, "\sim"] \arrow[r, "\lambda^*_{G,V,W}"] & \mathrm{PSp}_{\mathrm{fgl}}(F_{G,V\oplus W} S^V, j^*X) \arrow[d, "\sim"'] \\
\mathrm{PSp}^{\dagger}_{\mathrm{gl}}(F^{\dagger}_{G,W} S^0, X) \arrow[u, "\sim"] \arrow[r, "(\lambda^{\dagger}_{G,V,W})^*"] & \mathrm{PSp}^{\dagger}_{\mathrm{gl}}(F^{\dagger}_{G,V\oplus W} S^V, X)
\end{tikzcd} \tag{11.4.1}$$

commutes, so all the various λ-maps correspond to each other under the various adjunctions.

Given any compact Lie group G and any faithful G-representation W, we define a functor

$$U^{\mathrm{fgl}}_{G,W} : \mathrm{PSp}_{\mathrm{fgl}} \to \mathrm{PSp}_G$$

as the composite

$$\mathrm{PSp}_{\mathrm{fgl}} \xrightarrow{j_!} \mathrm{PSp}^{\dagger}_{\mathrm{gl}} \xrightarrow{U^{\mathrm{gl}}_G} \mathrm{PSp}_G \xrightarrow{\mathrm{sh}_W} \mathrm{PSp}_G,$$

where sh_W denotes the shift W-functor, given by cotensoring by $F_W S^0$.

Theorem 11.5 *An object $X \in \mathrm{PSp}_{\mathrm{fgl}}$ is in $\mathrm{Sp}_{\mathrm{gl}}$ if and only if, for every compact Lie group G and faithful G-representation W, the object $U^{\mathrm{fgl}}_{G,W}(X)$ is in Sp_G. Moreover, the functors $\{U^{\mathrm{fgl}}_{G,W}\}_{(G,W)}$ are also jointly conservative.*

Proof By Remark 11.3, we know that $X \in \mathrm{PSp}_{\mathrm{fgl}}$ is in $\mathrm{Sp}_{\mathrm{gl}}$ if and only if $j_! X \in \mathrm{PSp}^{\dagger}_{\mathrm{gl}}$ is local with respect to the set of maps $\{\lambda^{\dagger}_{G,V,W}\}$, where G runs over all compact Lie groups and V and W are G-representations with W faithful. The commutative diagram (11.4.1), together with the fact that $j^* j_! X \simeq X$, shows that this is equivalent to asking that for all compact Lie groups G, the object $U^{\mathrm{gl}}_G(j_! X)$ is local with respect to $\{\lambda_{G,V,W}\}$, where V and W are as above.

We next note that by definition, given an arbitrary G prespectrum Y, the map

$$\lambda^*_{U,V} : \mathrm{PSp}_G(F_V S^0, \mathrm{sh}_W Y) \to \mathrm{PSp}_G(F_{U \oplus V} S^U, \mathrm{sh}_W Y)$$

is equivalent to $\lambda^*_{U,V\oplus W}$. Also recall that given a faithful G-representation W, $W \oplus U$ is also faithful for any G-representation U.

These two observations combine to imply that $U^{\mathrm{gl}}_G(j_! X)$ is local with respect to $\{\lambda_{V,W}\}$ for G, V and W as above if and only if for all compact Lie groups G and faithful G-representations W, the object $\mathrm{sh}_W U^{\mathrm{gl}}_G j_!(X) = U^{\mathrm{fgl}}_{G,W} X$ is local with respect to $\{\lambda_{V,U}\}$ for arbitrary G-representations V and U.

On the other hand by Proposition 7.30, $U^{\mathrm{fgl}}_{G,W} X$ is in Sp_G if and only if for all closed subgroups $H \le G$, the H-prespectrum $\mathrm{res}^G_H U^{\mathrm{fgl}}_{G,W} X = U^{\mathrm{fgl}}_{H,\mathrm{res}^G_H W} X$ is local with respect to $\{\lambda_{V,U}\}$ for arbitrary H-representations V and U, and W a faithful G-representation. Varying these statements over all compact Lie groups, we find that $U^{\mathrm{fgl}}_{G,W} X$ is in Sp_G for all compact Lie groups G and all faithful G-representations W if and only if for all G and all faithful G-representations W, the G-prespectrum $U^{\mathrm{fgl}}_{G,W} X$ is $\{\lambda_{V,U}\}$-local for arbitrary G-representation V and U. This is identical to the condition of the previous paragraph, and so we obtain the first claim in the theorem. For the second statement, note that after forgetting module structures, the functor $U^{\mathrm{fgl}}_{G,W}$ is given by restriction along the functor

$$\mathrm{sh}_W : \mathbf{OR}_G \to \mathbf{OR}_{\mathrm{fgl}}, (G/H, U) \mapsto (H, U \oplus \mathrm{res}^G_H(W)).$$

The claim then follows from the fact that the functors $\{\mathrm{sh}_W\}_{(G,W)}$, where G runs over all compact Lie groups and W all faithful G-representations, are jointly essentially surjective. $\square$

The following is the key fact about the right adjoint j^*.

Proposition 11.6 *Let G be a compact Lie group and let W be a faithful G-representation. Then the following square commutes*:

$$\begin{tikzcd} \mathrm{PSp}_G \arrow[d, "\mathrm{sh}_W"'] & \mathrm{PSp}_{\mathrm{gl}}^{\dagger} \arrow[l, "U_G^{\mathrm{gl}}"'] \arrow[d, "j^*"] \\ \mathrm{PSp}_G & \mathrm{PSp}_{\mathrm{fgl}} \arrow[l, "U_{G,W}^{\mathrm{fgl}}"'] \end{tikzcd}$$

Proof The unit of the adjunction $j_! \dashv j^*$ provides a natural transformation

$$U_{G,W}^{\mathrm{fgl}} j^* = \mathrm{sh}_W U_G^{\mathrm{gl}} j_! j^* \to \mathrm{sh}_W U_G^{\mathrm{gl}},$$

which we claim is a natural equivalence. This follows from the fact that on underlying objects $\mathrm{sh}_W U_G^{\mathrm{gl}}$ is given by restriction along the functor

$$\mathbf{OR}_G \to \mathbf{OR}_{\mathrm{gl}}, \quad (H, V) \mapsto (H, \mathrm{res}_H^G(W) \oplus V).$$

This only sees levels in the image of $\mathbf{OR}_{\mathrm{fgl}}$, where the unit is an equivalence. □

Corollary 11.7 *Suppose $X \in \mathrm{Sp}_{\mathrm{gl}}^{\dagger}$. Then $j^*(X) \in \mathrm{Sp}_{\mathrm{gl}}$. In particular we obtain a functor*

$$j^* \colon \mathrm{Sp}_{\mathrm{gl}}^{\dagger} \to \mathrm{Sp}_{\mathrm{gl}},$$

which admits a left adjoint given by $L_{\mathrm{gl}} \circ j_!$.

Proof Because X is in $\mathrm{Sp}_{\mathrm{gl}}^{\dagger}$, we obtain that $U_G^{\mathrm{gl}}(X)$ is a G-spectrum by Proposition 10.10. Note that the functor sh_W preserves G-spectra for every G-representation W. We deduce using Proposition 11.6 that $U_{G,W}^{\mathrm{fgl}} j^*(X)$ is a G-spectrum for every G and W faithful. Therefore by Theorem 11.5 $j^*(X)$ is contained in $\mathrm{Sp}_{\mathrm{gl}}$. □

Proposition 11.8 *The map $j^* \colon \mathrm{Sp}_{\mathrm{gl}}^{\dagger} \to \mathrm{Sp}_{\mathrm{gl}}$ is conservative.*

Proof Let $f \colon X \to Y$ be a map in $\mathrm{PSp}_{\mathrm{gl}}^{\dagger}$ such that $j^*(f)$ is an equivalence. This implies that $f_{(G,W)}$ is an equivalence of spaces for every faithful G-representation W. We finish the argument by proving that if f is in fact a map between objects in $\mathrm{Sp}_{\mathrm{gl}}^{\dagger}$, then $f_{(G,V)}$ is an equivalence for *every* G-representation V if and only if it is an equivalence for *faithful* G-representations. The forward direction is trivial. For the converse, note that because $\mathrm{PSp}_{\mathrm{gl}}^{\dagger}$ is a partially lax limit, the collection of functors $\{U_G^{\mathrm{gl}}\}_G$ is jointly conservative. Now our assumptions tell us that $U_G^{\mathrm{gl}}(f)_{(G,W)}$ is an equivalence for every faithful G-representation W. But because f is in fact in $\mathrm{Sp}_{\mathrm{gl}}^{\dagger}$, both the source and target of $U_G^{\mathrm{gl}}(f)$ are G-spectra. Therefore our claim reduces to the fact that a map between G-spectra, which is an equivalence on faithful levels, is already an equivalence. The collection of faithful representations is cofinal in all representations, and so this is clear. □

Theorem 11.9 *The unit of the adjunction*

$$L_{\mathrm{gl}} \circ j_! \colon \mathrm{Sp}_{\mathrm{gl}} \leftrightarrows \mathrm{Sp}_{\mathrm{gl}}^{\dagger} : j^*$$

is an equivalence.

Proof Consider $X \in \mathrm{Sp}_{\mathrm{gl}}$. Let $\eta_X : X \to j^* L_{\mathrm{gl}} j_! X$ be the unit of the adjunction $L_{\mathrm{gl}} \circ j_! \leftrightarrows j^*$ evaluated at X. This adjunction is given as a composite of two adjunctions and so the unit is given by the composite

$$X \xrightarrow{\eta'} j^* j_! X \xrightarrow{j^*(\gamma)} j^* L_{\mathrm{gl}} j_! X,$$

where η' is the unit of the adjunction $j_! \dashv j^*$ and γ exhibits $L_{\mathrm{gl}} j_! X$ as the localization of $j_! X$ in $\mathrm{PSp}_{\mathrm{gl}}^{\dagger}$. However, recall that $j_!$ is fully faithful and therefore the first of the two maps is an equivalence. So it suffices to prove that the second map is also an equivalence.

The functors $U_{G,W}^{\mathrm{fgl}}$ are jointly conservative, and so we will prove that $U_{G,W}^{\mathrm{fgl}}(j^*(\gamma))$ is an equivalence for every (G, W) where W is faithful. Applying Proposition 11.6 we conclude that $U_{G,W}^{\mathrm{fgl}}(j^*(\gamma))$ is equivalent to

$$\mathrm{sh}^W U_G^{\mathrm{gl}}(\gamma) \colon \mathrm{sh}^W U_G^{\mathrm{gl}} j_! X \to \mathrm{sh}^W U_G^{\mathrm{gl}} L_{\mathrm{gl}} j_! X.$$

By Proposition 10.10, $U_G^{\mathrm{gl}}(\gamma)$ is equivalent to

$$\gamma_G : U_G^{\mathrm{gl}} j_! X \to L_G U_G^{\mathrm{gl}} j_! X,$$

where γ_G exhibits $L_G U_G^{\mathrm{gl}} j_! X$ as the localization of $U_G^{\mathrm{gl}} j_! X$ in PSp_G. Spectrification of G-prespectra commutes with sh^W, and therefore $\mathrm{sh}^W(\gamma_G)$ gives the localization of $U_{G,W}^{\mathrm{fgl}}(X) = \mathrm{sh}^W U_G^{\mathrm{gl}} j_! X$ in PSp_G. Recall that $X \in \mathrm{Sp}_{\mathrm{gl}}$, and so $U_{G,W}^{\mathrm{fgl}}(X)$ is a G-Ω-spectrum by Theorem 11.5. Therefore $\mathrm{sh}^W(\gamma_G)$ is an equivalence, concluding the proof. □

Theorem 11.10 *There is a symmetric monoidal equivalence* $j^* \colon \mathrm{Sp}_{\mathrm{gl}}^{\dagger} := \mathrm{laxlim}^{\dagger}{}_G\, \mathrm{Sp}_G \to \mathrm{Sp}_{\mathrm{gl}}$.

Proof We have proven that $j_! \dashv j^*$ is an adjunction in which the right adjoint is conservative, and the unit is a natural equivalence. Therefore the functors are an adjoint equivalence. Moreover $j_!$ is strong monoidal, which implies that j^*, as its inverse, is also strong monoidal. □

12 Proper equivariant spectra as a limit

The goal of this section is to exhibit the ∞-category of genuine proper G spectra Sp_G as a limit over the proper orbit category $\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}}$ of a diagram

$$\mathrm{Sp}_{(-)} \colon \boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}} \to \mathrm{Cat}_\infty, \quad G/H \to \mathrm{Sp}_H.$$

In contrast to the case of global spectra, once the diagram has been constructed, the identification of the limit will be almost immediate. In fact even the general strategy for constructing the diagram is essentially identical. For this reason we will be brief and refer to Section 9 for the relevant details.

Recall from Lemma 9.9 that the ∞-operad $\mathbf{OR}_G^{\otimes}$ fits into a pullback

$$\begin{tikzcd} \mathbf{OR}_G^{\otimes} \arrow[r, "\nu_G"] \arrow[d, "\pi_G"'] & \mathbf{OR}_{\mathrm{gl}}^{\otimes} \arrow[d, "\pi_{\mathrm{gl}}"] \\ (\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}})^{\amalg} \arrow[r, "\iota_G^{\amalg}"] & (\mathrm{Glo}^{\mathrm{op}})^{\amalg} \end{tikzcd}$$

Because $\mathbf{OR}_{\mathrm{gl}}^{\otimes} \to (\mathrm{Glo}^{\mathrm{op}})^{\amalg}$ is a cocartesian fibration which by definition classifies the functor $\mathrm{Rep}(-)$, we immediately obtain:

Proposition 12.1 *For every Lie group G, the forgetful functor $\pi_G : \mathbf{OR}_G^{\otimes} \to (\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}})^{\amalg}$ is a cocartesian fibration which classifies the functor*

$$\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}} \to \mathrm{Cat}_{\infty}^{\otimes}, \quad G/H \mapsto \mathrm{Rep}(H).$$

Definition 12.2 We define $\widetilde{\mathbf{OR}}_G^{\otimes}$ via the following pullback of operads:

$$\begin{tikzcd} \widetilde{\mathbf{OR}}_G^{\otimes} \arrow[r] \arrow[d, "\pi_{\mathrm{Ar}}"] & \mathbf{OR}_G^{\otimes} \arrow[d] \\ (\mathrm{Ar}(\boldsymbol{O}_{G,\mathrm{pr}})^{\mathrm{op}})^{\amalg} \arrow[r, "s^{\mathrm{op}}"] & (\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}})^{\amalg} \end{tikzcd}$$

We consider $\widetilde{\mathbf{OR}}_G^{\otimes}$ as living over $\boldsymbol{O}_{G,\mathrm{pr}}$ via the composite

$$\pi : \widetilde{\mathbf{OR}}_G^{\otimes} \xrightarrow{\pi_{\mathrm{Ar}}} (\mathrm{Ar}(\boldsymbol{O}_{G,\mathrm{pr}})^{\mathrm{op}})^{\amalg} \xrightarrow{t^{\mathrm{op}}} (\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}})^{\amalg}.$$

Just as in Lemma 9.11, we can show that $\widetilde{\mathbf{OR}}_G^{\otimes}$ is a pro-$(\boldsymbol{O}_G)^{\amalg}$-monoidal category.

Proposition 12.3 *The functor $\pi : \widetilde{\mathbf{OR}}_G \to (\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}})$, given by restricting π to underlying categories, is a cartesian fibration. Furthermore, an edge $(f,g) \in \widetilde{\mathbf{OR}}_G$ is cartesian if and only if $s^{\mathrm{op}}(f)$ and g are equivalences.*

Proof The proof is analogous to Lemma 9.13. □

Proposition 12.4
$$\widetilde{\mathbf{OR}}_G^{\otimes} \times_{(\boldsymbol{O}_G^{\mathrm{op}})^{\amalg}} \{G/H\} \simeq \mathbf{OR}_H^{\otimes}.$$

Proof The pullback $P = \widetilde{\mathbf{OR}}_G^{\otimes} \times_{(\boldsymbol{O}_G^{\mathrm{op}})^{\amalg}} \{G/H\}$ fits into the diagram

$$\begin{tikzcd} P \arrow[r] \arrow[d] & \widetilde{\mathbf{OR}}_G \arrow[r] \arrow[d] & \mathbf{OR}_G^{\otimes} \arrow[r] \arrow[d] & \mathbf{OR}_{\mathrm{gl}}^{\otimes} \arrow[d] \\ (\boldsymbol{O}_H^{\mathrm{op}})^{\amalg} \arrow[r] \arrow[d] & (\mathrm{Ar}(\boldsymbol{O}_{G,\mathrm{pr}})^{\mathrm{op}})^{\amalg} \arrow[r] \arrow[d] & (\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}})^{\amalg} \arrow[r] & (\mathrm{Glo}^{\mathrm{op}})^{\amalg} \\ \{G/H\} \arrow[r] & (\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}})^{\amalg} & & \end{tikzcd}$$

in which every square is a pullback. One can show by direct computation that the middle composite $(\boldsymbol{O}_H^{\mathrm{op}})^{\amalg} \to (\mathrm{Glo}^{\mathrm{op}})^{\amalg}$ is equivalent to $\iota_H^{\amalg}$. Therefore the result follows from Lemma 9.9. □

Definition 12.5 Consider the Day convolution operad

$$\mathrm{Fun}_{\boldsymbol{O}_{G,\mathrm{pr}}}(\widetilde{\mathbf{OR}}_G^{\otimes}, \mathcal{S}_*^{\wedge} \times (\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}})^{\amalg})^{\mathrm{Day}}.$$

Just as in Section 9, this is an $(\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}})^{\amalg}$-monoidal category. We define

$$\mathbf{OR}_\bullet\text{-}\mathcal{S}_* \colon \boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}} \to \mathrm{Cat}_\infty^{\otimes}$$

to be the functor associated to it by the equivalence of Proposition 5.5. By Proposition 12.4, the value of $\mathbf{OR}_\bullet\text{-}\mathcal{S}_*$ at G/H is equivalent to $\mathbf{OR}_H\text{-}\mathcal{S}_*$.

Proposition 12.6 *The projection map $\widetilde{\mathbf{OR}}_G^{\otimes} \to \mathbf{OR}_G^{\otimes}$ induces a fully faithful symmetric monoidal functor $\mathbf{OR}_G\text{-}\mathcal{S}_* \to \widetilde{\mathbf{OR}}_G\text{-}\mathcal{S}_*$, given by restriction. A functor $F \colon \widetilde{\mathbf{OR}}_G \to \mathcal{S}_*$ is in its essential image if and only if F sends π-cartesian edges to equivalences.*

Proof The argument is identical to that of Lemma 9.14. □

Lemma 12.7 *There is a symmetric monoidal equivalence*

$$\lim_{\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}}} \mathbf{OR}_\bullet\text{-}\mathcal{S}_* \simeq \mathbf{OR}_G\text{-}\mathcal{S}_*.$$

Proof The calculation at the beginning of the proof of Lemma 9.15 shows that the lax limit of the diagram $\mathbf{OR}_\bullet\text{-}\mathcal{S}_*$ is equivalent to the symmetric monoidal category $\widetilde{\mathbf{OR}}_G\text{-}\mathcal{S}_*$ To compute the actual limit, we can once again argue on underlying categories by appealing to Remark 5.2. Note that by Remark 3.11, the underlying category of $\widetilde{\mathbf{OR}}_G\text{-}\mathcal{S}_*$ is equivalent to $\mathrm{Fun}(\widetilde{\mathbf{OR}}_G, \mathcal{S}_*)$. The analysis of the second half of the proof of Lemma 9.15 implies that the limit is equivalent to the full subcategory spanned by the functors which send π-cartesian edges to equivalences. By Proposition 12.6 this subcategory is equivalent to $\mathrm{Fun}(\mathbf{OR}_G, \mathcal{S}_*)$. □

Recall from Definition 7.18 that $\mathbf{OR}_G$-spaces admit an algebra object S_G, whose restriction to $\mathbf{OR}_H$-spaces for H a compact subgroup of G is equivalent to S_H.

Corollary 12.8 *There exists a functor $\mathrm{PSp}_\bullet \colon \boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}} \to \mathrm{Cat}_\infty^{\otimes}$, and one calculates*

$$\lim_{\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}}} \mathrm{PSp}_\bullet \simeq \mathrm{Mod}_{S_G}(\mathbf{OR}_G\text{-}\mathcal{S}).$$

Proof Once again, $\mathrm{PSp}_\bullet$ is defined as $\mathrm{Mod}_{S^\bullet}(\mathbf{OR}_\bullet\text{-}\mathcal{S}_*)$, using Theorem 5.10. An argument as in Proposition 9.16 allows us to calculate the limit. □

So far we have constructed and computed the limit of the diagram $\mathrm{PSp}_\bullet \colon \boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}} \to \mathrm{Cat}_\infty^{\otimes}$. Given a map $\alpha \colon H \hookrightarrow K \subset G$ in $\boldsymbol{O}_{G,\mathrm{pr}}$, the induced map $\mathrm{PSp}_K \to \mathrm{PSp}_H$ is by construction equivalent to the global functoriality constructed in Section 9 evaluated at α. Therefore the results there imply that PSp_α preserves spectrum objects, and so we obtain a diagram $\mathrm{Sp}_\bullet\colon \boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}} \to \mathrm{Cat}_\infty^\otimes$. Furthermore, Corollary 10.6 implies that $\mathrm{Sp}_\alpha\colon \mathrm{Sp}_K \to \mathrm{Sp}_H$ agrees with the standard restriction functor between equivariant spectra. To calculate the limit of $\mathrm{Sp}_\bullet$, we apply Lemma 4.13 to conclude:

Corollary 12.9 *The limit $\lim_{\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}}} \mathrm{Sp}_\bullet$ is a Bousfield localization of $\mathrm{Mod}_{S_G}(\mathbf{OR}_G\text{-}\mathcal{S}_*)$ at the objects X whose restriction to $\mathrm{Mod}_{S_H}(\mathbf{OR}_H\text{-}\mathcal{S}_*)$ is an H-spectrum for every compact subgroup H of G.*

Recall from Section 8 that the category of genuine proper G-spectra is also a Bousfield localization of $\mathrm{Mod}_{S_G}(\mathbf{OR}_G\text{-}\mathcal{S}_*)$. Therefore it remains to show that the two subcategories agree.

Proposition 12.10 *An object $X \in \mathrm{Mod}_{S_G}(\mathbf{OR}_G\text{-}\mathcal{S}_*)$ is a G-spectrum if and only if for every compact subgroup $H \le G$, the restriction of X to $\mathrm{Mod}_{S_H}(\mathbf{OR}_H\text{-}\mathcal{S}_*)$ is a H-spectrum.*

Proof Recall from Proposition 7.30 that an object $X \in \mathrm{PSp}_G$ is a G-spectrum if and only if for all compact subgroups $H \le G$, the object $\mathrm{res}^G_H X$ is local with respect to $\lambda_{H,V,W}$. Now by definition, $\mathrm{res}^G_H X$ is a G-spectrum if and only if $\mathrm{res}^H_K \mathrm{res}^G_H X$ is local with respect to $\lambda_{K,V,W}$. However, because $\mathrm{res}^H_K \mathrm{res}^G_H = \mathrm{res}^G_K$, we conclude that the two conditions of the theorem agree. $\square$

Thus we can conclude the main theorem of this section:

Theorem 12.11 *The category of proper G-spectra is equivalent to the limit of the diagram*

$$\mathrm{Sp}_\bullet\colon \boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}} \to \mathrm{Cat}_\infty^\otimes.$$

In symbols,

$$\mathrm{Sp}_G \simeq \lim_{\boldsymbol{O}_{G,\mathrm{pr}}^{\mathrm{op}}} \mathrm{Sp}_\bullet.$$

Appendix Tensor product of modules in an ∞-category

The goal of this section is to provide a proof of Theorem 12.21 below, which will be useful when studying lax limits of ∞-categories of modules. This section uses some technical results about the theory of ∞-operads as developed in [Lurie 2017] and so it should be skipped on a first reading.

Definition 12.12 We define $\mathscr{CM}^\otimes$ to be the ∞-operad corresponding to the symmetric multicategory with two objects a and m with

$$\mathrm{Mul}(\{x_i\}, a) = \begin{cases} * & \text{if for all } i,\ x_i = a, \\ \varnothing & \text{otherwise,} \end{cases} \qquad \mathrm{Mul}(\{x_i\}, m) = \begin{cases} * & \text{if } |\{i \mid x_i = m\}| = 1, \\ \varnothing & \text{otherwise.} \end{cases}$$

We know by [Glasman 2014, Proposition 7] or [Hinich 2015, Lemma B.1.1] that for every ∞-operad $\mathscr{C}^\otimes$ there is a natural equivalence of ∞-categories

$$\mathrm{Mod}^{\mathrm{Fin}_*}(\mathscr{C}) \simeq \mathrm{Alg}_{\mathscr{CM}^\otimes}(\mathscr{C}).$$

Our goal is to give a similar description of the tensor product of modules over a commutative algebra, that is, of the family of ∞-operads $\mathrm{Mod}(\mathscr{C})^{\otimes}$. In order to do so we will introduce a variant of $\mathscr{CM}^{\otimes}$ which parametrizes finite sets of modules.

Construction 12.13 Let $\widetilde{\mathscr{CM}}^{\otimes}$ be the category whose objects are triples $(\langle n\rangle, \langle m\rangle, \{S_i\}_{i=1,\dots,n})$, where $\langle n\rangle, \langle m\rangle \in \mathrm{Fin}_*$ and $\{S_i\}$ is a family of pairwise disjoint subsets of $\langle m\rangle$. A map

$$(\langle n\rangle, \langle m\rangle, \{S_i\}) \to (\langle n'\rangle, \langle m'\rangle, \{S'_i\})$$

is a pair of maps $f:\langle n\rangle \to \langle n'\rangle$ and $g:\langle m\rangle \to \langle m'\rangle$ in Fin_* such that

- for every $i \in \langle n\rangle^\circ$, we have $g(S_i) \subseteq S'_{f(i)} \cup \{*\}$, where $S'_* = \varnothing$,
- for every $i \in f^{-1}\langle n'\rangle^\circ$ and every $s' \in S'_{f(i)}$, there is exactly one $s \in S_i$ such that $g(s) = s'$.

Lemma 12.14 *The projection $\widetilde{\mathscr{CM}}^{\otimes} \to \mathrm{Fin}_* \times \mathrm{Fin}_*$ that forgets the subsets $\{S_i\}$ is a Fin_*-family of ∞-operads in the sense of* [Lurie 2017, Definition 2.3.2.10], *with inert arrows exactly those arrows that are sent to an equivalence by the first projection and to an inert arrow by the second projection.*

Proof The inert arrows are the arrows

$$(\mathrm{id}_{\langle n\rangle}, f): (\langle n\rangle, \langle m\rangle, \{S_i\}) \to (\langle n\rangle, \langle m'\rangle, \{f(S_i) \cap \langle m'\rangle^\circ\}),$$

where $f:\langle m\rangle \to \langle m'\rangle$ is an inert arrow in Fin_*. It is easy to check that they satisfy all necessary properties. $\square$

Notation 12.15 For every ∞-category $X \to \mathrm{Fin}_*$ with a functor to Fin_*, we will write $\widetilde{\mathscr{CM}}^{\otimes}_X$ for the X-family of ∞-operads $X \times_{\mathrm{Fin}_*} \widetilde{\mathscr{CM}}^{\otimes}$, where we pull back along the composite

$$\widetilde{\mathscr{CM}}^{\otimes} \to \mathrm{Fin}_* \times \mathrm{Fin}_* \xrightarrow{\mathrm{pr}_1} \mathrm{Fin}_*.$$

Note that $\widetilde{\mathscr{CM}}^{\otimes}_{\langle 1\rangle}$ is equivalent to $\mathscr{CM}^{\otimes}$. Intuitively, the fiber $\widetilde{\mathscr{CM}}^{\otimes}_{\langle n\rangle}$ is the ∞-operad controlling pairs $(A, \{M_i\})$, where A is a commutative algebra and $\{M_i\}$ is an n-tuple of A-modules.

We will write a_n for the object $(\langle n\rangle, \langle 1\rangle, \{\varnothing\})$ and $m_{j,n}$ for the object $(\langle n\rangle, \langle 1\rangle, \{S_i\})$, where $S_i = \varnothing$ for $i \neq j$ and $S_j = \{1\}$. It's easy to see these are all the objects of the underlying category of the generalized operad

$$\widetilde{\mathscr{CM}}^{\otimes} \to \mathrm{Fin}_* \times \mathrm{Fin}_* \xrightarrow{\mathrm{pr}_2} \mathrm{Fin}_*.$$

First we will prove a generalization of [Glasman 2014, Proposition 7] that shows how $\widetilde{\mathscr{CM}}^{\otimes}$ controls the tensor product of modules over commutative algebras.

Proposition 12.16 *Let $X \in (\mathrm{Cat}_\infty)_{/\mathrm{Fin}_*}$ be an ∞-category over Fin_*, and let $\mathscr{C}^{\otimes} \in \mathrm{Op}_\infty$ be an ∞-operad. Then there is a natural equivalence*

$$\mathrm{Alg}_{\widetilde{\mathscr{CM}}_X}(\mathscr{C}^{\otimes}) \simeq \mathrm{Fun}_{/\mathrm{Fin}_*}(X, \mathrm{Mod}^{\mathrm{Fin}_*}(\mathscr{C})^{\otimes}).$$

Proof Let $\mathcal{K} \subseteq \mathrm{Ar}(\mathrm{Fin}_*)$ be the full subcategory of semi-inert arrows [Lurie 2017, Notation 3.3.2.1]. Consider the pullback

$$\begin{tikzcd} X \times_{\mathrm{Fin}_*} \mathcal{K} \arrow[r] \arrow[d] & X \arrow[d] \\ \mathcal{K} \arrow[r, "s"] \arrow[d, "t"] & \mathrm{Fin}_* \\ \mathrm{Fin}_* & \end{tikzcd}$$

We will say that an arrow (f, g) in $X \times_{\mathrm{Fin}_*} \mathcal{K}$ is inert if f is an equivalence and $t(g)$ is an inert edge of Fin_*(this is different from the convention in [Lurie 2017], but it is more suited to the current proof). Then recall that by [Lurie 2017, Construction 3.3.3.1] the ∞-category $\mathrm{Mod}(\mathcal{C})^{\otimes}$ is defined so that there is a natural fully faithful inclusion

$$\mathrm{Fun}_{/\mathrm{Fin}_*}(X, \mathrm{Mod}^{\mathrm{Fin}_*}(\mathcal{C})^{\otimes}) \to \mathrm{Fun}_{/\mathrm{Fin}_*}(X \times_{\mathrm{Fin}_*} \mathcal{K}, \mathcal{C}^{\otimes}),$$

where $X \times_{\mathrm{Fin}_*} \mathcal{K}$ lives over Fin_* by the vertical composite in the diagram above, with essential image those functors sending inert arrows of $X \times_{\mathrm{Fin}_*} \mathcal{K}$ to inert arrows.

There is a functor $\mathcal{K} \to \widetilde{\mathcal{CM}}$ sending a semi-inert arrow $[s\colon \langle n\rangle \to \langle m\rangle]$ to $(\langle n\rangle, \langle m\rangle, \{\{s(i)\} \cap \langle m\rangle^\circ\}_i)$. It identifies $\mathcal{K}$ with the full subcategory of $\widetilde{\mathcal{CM}}$ spanned by those triples $(\langle n\rangle, \langle m\rangle, \{S_i\})$ where $|S_i| \le 1$ for every $i \in \langle n\rangle^\circ$. Moreover, an arrow in $X \times_{\mathrm{Fin}_*} \mathcal{K}$ is inert if and only if its image in $\widetilde{\mathcal{CM}}_X$ is inert. Therefore restricting along this inclusion induces a natural transformation

$$\mathrm{Alg}_{\widetilde{\mathcal{CM}}_X}(\mathcal{C}^{\otimes}) \to \mathrm{Fun}_{/\mathrm{Fin}_*}(X, \mathrm{Mod}^{\mathrm{Fin}_*}(\mathcal{C})^{\otimes}).$$

Our goal now is to prove that this is an equivalence of ∞-categories. This follows from [Lurie 2009, Proposition 4.3.2.15] together with the following two statements, where we write $p\colon \mathcal{C}^{\otimes} \to \mathrm{Fin}_*$ for the structure map of $\mathcal{C}^{\otimes}$:

(1) Every map $F\colon X \times_{\mathrm{Fin}_*} \mathcal{K} \to \mathcal{C}^{\otimes}$ over Fin_* that sends inert arrows to inert arrows admits a right p-Kan extension to $\widetilde{\mathcal{CM}}_X$ that sends inert arrows to inert arrows.

(2) A functor $F\colon \widetilde{\mathcal{CM}}_X \to \mathcal{C}^{\otimes}$ which sends inert arrows to inert arrows is the right p-Kan extension of its restriction to $X \times_{\mathrm{Fin}_*} \mathcal{K}$.

Let $(x, \langle m\rangle, \{S_i\})$ be an object of $\widetilde{\mathcal{CM}}_X$ and write $S = \coprod_i S_i \subseteq \langle m\rangle^\circ$. Let us consider the functor

$$\mathcal{P}(S)^{\mathrm{op}} \to \widetilde{\mathcal{CM}}_X$$

sending a subset $A \subseteq S$ to $(x, \langle m\rangle/(S \setminus A), \{A \cap S_i\})$ and all arrows to inert arrows. This induces a functor

$$\mathcal{P}(S)^{\mathrm{op}} \to (\widetilde{\mathcal{CM}}_X)_{(x,\langle m\rangle,\{S_i\})/},$$

which sends A to the inert morphism collapsing all elements of S not in A to the basepoint. If we let $\mathcal{Q}(S) \subseteq \mathcal{P}(S)$ be the subposet of those elements A such that $|A \cap S_i| \le 1$ for every i, we obtain a functor

$$\mathcal{Q}(S)^{\mathrm{op}} \to (X \times_{\mathrm{Fin}_*} \mathcal{K})_{(x,\langle m\rangle,\{S_i\})/}$$

to the comma category, which has a right adjoint given by

$$[(f,g)\colon (x,\langle m\rangle,\{S_i\}) \to (x',\langle m'\rangle,\{S_i'\})] \mapsto g^{-1}\biggl(\coprod_i S_i'\biggr)\cap S,$$

and therefore is coinitial. Thus, by [Lurie 2009, Proposition 4.3.1.7 and Lemma 4.3.2.13] it suffices to show the following two conditions:

(1) Let $F\colon X\times_{\mathrm{Fin}_*}\mathcal{K}\to\mathscr{C}^\otimes$ send inert arrows to inert arrows. Then the composition
$$\mathfrak{Q}(S)^{\mathrm{op}}\to X\times_{\mathrm{Fin}_*}\mathcal{K}\to\mathscr{C}^\otimes$$
has a p-limit diagram sending all edges to inert edges.

(2) Let $F\colon \widetilde{\mathscr{C}\mathscr{M}}_X\to\mathscr{C}^\otimes$ send inert arrows to inert arrows. Then the composition
$$(\mathfrak{Q}(S)^{\mathrm{op}})^{\triangleleft}\to\mathscr{P}(S)^{\mathrm{op}}\to\widetilde{\mathscr{C}\mathscr{M}}_X\to\mathscr{C}^\otimes$$
is a p-limit diagram, where the first functor sends the cone point to $S\subseteq S$.

Both of them are now an immediate consequence of the characterization of p-limit diagrams in terms of mapping spaces [Lurie 2009, Remark 4.3.1.2] and the definition of ∞-operads. □

Now we will obtain a description of inert and cocartesian arrows of $\mathrm{Mod}^{\mathrm{Fin}_*}(\mathscr{C})^\otimes$ in terms of the model of Proposition 12.16.

Construction 12.17 (bar construction) There is a functor
$$B\colon (\Delta^{\mathrm{op}})^{\triangleright}\to\widetilde{\mathscr{C}\mathscr{M}}^\otimes$$
sending $[n]$ to $(\langle 2\rangle,\mathrm{Hom}_\Delta([n],[1])_+,\{\{r_0\},\{r_1\}\})$, where r_i is the constant arrow at i, and sending the point at ∞ to $m_{1,1}=(\langle 1\rangle,\langle 1\rangle,\{1\})$. Concretely this sends $[n]$ to the object $(m_{2,1},a,\dots,a,m_{2,2})$ in the fiber over $\langle n+2\rangle$ of the ∞-operad $\widetilde{\mathscr{C}\mathscr{M}}^\otimes_{\langle 2\rangle}$ (and so it encodes the bar construction in $\widetilde{\mathscr{C}\mathscr{M}}^\otimes_{\langle 2\rangle}$).

Lemma 12.18 *Let $e\colon\Delta^1\to\mathrm{Mod}^{\mathrm{Fin}_*}(\mathscr{C})^\otimes$ be an arrow, and let $e_0\colon\langle n\rangle\to\langle n'\rangle$ be the image of e in Fin_*. Write*
$$F_e\colon\widetilde{\mathscr{C}\mathscr{M}}^\otimes_{\Delta^1}\to\mathscr{C}^\otimes$$
for the functor corresponding to e via the isomorphism of Proposition 12.16.

(1) *The arrow e is inert if and only if e_0 is inert and F_e sends the arrows $a_n\to a_{n'}$ and $m_{i,n}\to m_{e_0 i,n'}$ to cocartesian arrows.*

(2) *Suppose that $\mathscr{C}^\otimes$ is a symmetric monoidal ∞-category compatible with geometric realizations, and that e_0 is the unique active arrow from $\langle 2\rangle$ to $\langle 1\rangle$. Then e is cocartesian if and only if F_e sends the arrow $a_2\to a_1$ to a cocartesian arrow and the composition*
$$(\Delta^{\mathrm{op}})^{\triangleright}\xrightarrow{B}\widetilde{\mathscr{C}\mathscr{M}}^\otimes_{\Delta^1}\xrightarrow{F_e}\mathscr{C}^\otimes$$
is an operadic colimit diagram.

Proof This is immediate from the proofs of [Lurie 2017, Proposition 3.3.3.10 and Theorem 4.5.2.1] and the identification of Proposition 12.16. □

Construction 12.19 There is a square of ∞-categories

$$\begin{tikzcd}
\mathrm{Fin}_* \times \mathrm{Fin}_* \arrow[r, "{(1,\wedge)}"] \arrow[d, "j_1"] & \mathrm{Fin}_* \times \mathrm{Fin}_* \arrow[d, "j_2"] \\
\mathrm{Fin}_* \times \mathscr{C}\mathscr{M}^{\otimes} \arrow[r, "\phi"] & \widetilde{\mathscr{C}\mathscr{M}}^{\otimes}
\end{tikzcd}$$

where

- the top horizontal arrow sends $(\langle n\rangle, \langle m\rangle)$ to $(\langle n\rangle, \langle n\rangle \wedge \langle m\rangle)$,
- the arrow j_1 sends $(\langle n\rangle, \langle m\rangle)$ to $(\langle n\rangle, (\langle m\rangle, \varnothing)) \in \mathrm{Fin}_* \times \mathscr{C}\mathscr{M}^{\otimes}$,
- the arrow j_2 sends $(\langle n\rangle, \langle m\rangle)$ to $(\langle n\rangle, \langle m\rangle, \{\varnothing\}) \in \widetilde{\mathscr{C}\mathscr{M}}^{\otimes}$,
- the arrow ϕ sends $(\langle n\rangle, (\langle m\rangle, S)) \in \mathrm{Fin}_* \times \mathscr{C}\mathscr{M}^{\otimes}$ to $(\langle n\rangle, \langle n\rangle \wedge \langle m\rangle, \{\{i\} \times S\})$.

Since each of these functors sends inert arrows to inert arrows, it induces for every $X \in (\mathrm{Cat}_\infty)_{/\mathrm{Fin}_*}$ a natural square

$$\begin{tikzcd}
\mathrm{Fun}_{/\mathrm{Fin}_*}(X, \mathrm{Mod}^{\mathrm{Fin}_*}(\mathscr{C})^{\otimes}) \simeq \mathrm{Alg}_{\widetilde{\mathscr{C}\mathscr{M}}^{\otimes}_X}(\mathscr{C}^{\otimes}) \arrow[r] \arrow[d] & \mathrm{Fun}_{/\mathrm{Fin}_*}(X, \mathrm{Alg}_{\mathscr{C}\mathscr{M}^{\otimes}}(\mathscr{C})^{\otimes}) \arrow[d] \\
\mathrm{Fun}_{/\mathrm{Fin}_*}(X, \mathrm{Fin}_* \times \mathrm{CAlg}(\mathscr{C})) \simeq \mathrm{Alg}_{X \times \mathrm{Fin}_*}(\mathscr{C}^{\otimes}) \arrow[r] & \mathrm{Fun}_{/\mathrm{Fin}_*}(X, \mathrm{CAlg}(\mathscr{C})^{\otimes})
\end{tikzcd}$$

and therefore a natural square of ∞-categories over Fin_*

(12.19.1)
$$\begin{tikzcd}
\mathrm{Mod}^{\mathrm{Fin}_*}(\mathscr{C})^{\otimes} \arrow[r] \arrow[d] & \mathrm{Alg}_{\mathscr{C}\mathscr{M}^{\otimes}}(\mathscr{C})^{\otimes} \arrow[d] \\
\mathrm{Fin}_* \times \mathrm{CAlg}(\mathscr{C}) \arrow[r] & \mathrm{CAlg}(\mathscr{C})^{\otimes}
\end{tikzcd}$$

Our goal now is to show that the square (12.19.1) is cartesian. To do so we will show that the right vertical arrow is a cocartesian fibration in favorable situations.

Lemma 12.20 *Let $\mathscr{I}$ be an ∞-category and $\mathscr{C}^{\otimes} \to \mathscr{I}^{\amalg}$ be an $\mathscr{I}^{\amalg}$-monoidal ∞-category compatible with geometric realizations. Then the map of ∞-operads*

$$p_{\mathscr{I}} : \mathrm{Alg}_{\mathscr{C}\mathscr{M}^{\otimes}/\mathscr{I}^{\amalg}}(\mathscr{C})^{\otimes} \to \mathrm{Alg}_{\mathrm{Fin}_*/\mathscr{I}^{\amalg}}(\mathscr{C})^{\otimes}$$

is a cocartesian fibration.

Proof By [Lurie 2017, Proposition 3.2.4.3.(3)] this is a map of cocartesian fibrations over $\mathscr{I}^{\amalg}$. Moreover, the fiber over $\{x_j\}_{j\in J} \in \mathscr{I}^{\amalg}$ is given by

$$\prod_{j\in J} \mathrm{Mod}(\mathscr{C}_{x_j}) \to \prod_{j \in J} \mathrm{CAlg}(\mathscr{C}_{x_j}),$$

and therefore it is a cocartesian fibration by [Lurie 2017, Theorem 4.5.3.1]. Therefore by [Lurie 2009, Proposition 2.4.2.11] $p_{\mathscr{I}}$ is a locally cocartesian fibration with locally cocartesian arrows those given by the composition of a fiberwise cocartesian arrow and a cocartesian arrow over $\mathcal{I}^{\amalg}$. In order to prove it is a cocartesian fibration it suffices to show then that the composition of two locally cocartesian arrow is locally cartesian, that is, that fiberwise cocartesian arrows are stable under pushforward along arrows in $\mathcal{I}^{\amalg}$. Unwrapping the various cases it suffices to show that for every $x, y \in \mathcal{I}$ and arrow $f : x \to y$, the squares

$$\begin{tikzcd} \mathrm{Mod}(\mathscr{C}_x) \times \mathrm{Mod}(\mathscr{C}_x) \arrow[r, "\otimes"] \arrow[d] & \mathrm{Mod}(\mathscr{C}_x) \arrow[d] \\ \mathrm{CAlg}(\mathscr{C}_x) \times \mathrm{CAlg}(\mathscr{C}_x) \arrow[r, "\otimes"] & \mathrm{CAlg}(\mathscr{C}_x) \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} \mathrm{Mod}(\mathscr{C}_x) \arrow[r, "f_*"] \arrow[d] & \mathrm{Mod}(\mathscr{C}_y) \arrow[d] \\ \mathrm{CAlg}(\mathscr{C}_x) \arrow[r, "f_*"] & \mathrm{CAlg}(\mathscr{C}_y) \end{tikzcd}$$

are maps of cocartesian fibrations. That is, that for every two maps of commutative algebras $A \to A'$, $B \to B'$, A-module M and B-module N, the canonical maps

$$(M \otimes N) \otimes_{A \otimes B} (A' \otimes B') \simeq (M \otimes_A A') \otimes (N \otimes_B B') \quad \text{and} \quad f_*(M \otimes_A B) \simeq f_* M \otimes_{f_* A} f_* B$$

are equivalences. This is easily seen to be true since f_* is symmetric monoidal and commutes with geometric realization, and the tensor product commutes with geometric realization in each variable. $\square$

Finally we arrive at the main result of this section.

Theorem 12.21 *The square (12.19.1) is cartesian for every ∞-operad $\mathscr{C}^{\otimes}$.*

Proof Let us do first the case where $\mathscr{C}^{\otimes}$ is a symmetric monoidal ∞-category compatible with geometric realizations. Then both vertical arrows are cocartesian fibrations by [Lurie 2017, Theorem 4.5.3.1] and Lemma 12.20. Moreover, the description of cocartesian arrows in Lemma 12.18 and [Lurie 2017, Proposition 3.2.4.3.(4)] shows that

$$\mathrm{Mod}^{\mathrm{Fin}_*}(\mathscr{C})^{\otimes} \to (\mathrm{Fin}_* \times \mathrm{CAlg}(\mathscr{C})) \times_{\mathrm{CAlg}(\mathscr{C})^{\otimes}} \mathrm{Alg}_{\mathscr{C}\mathscr{M}^{\otimes}}(\mathscr{C})^{\otimes}$$

is a map of cocartesian fibrations over Fin_*. So it suffices to show that it induces an equivalence on fibers. Since it is a map of generalized operads, it suffices to show it induces an equivalence on the fibers over $\langle 0 \rangle$ and $\langle 1 \rangle$. But this is immediate by Proposition 12.16.

Now let us show the result for small ∞-operads $\mathscr{C}$. Indeed, it is clear by inspection that if the square (12.19.1) is cartesian for an ∞-operad, then it is cartesian for any full suboperad. But every small ∞-operad embeds as a full suboperad of a symmetric monoidal ∞-category compatible with small colimits. Indeed, this is just the composition $\mathscr{C}^{\otimes} \to \mathrm{Env}\,\mathscr{C}^{\otimes} \to \mathscr{P}(\mathrm{Env}\,\mathscr{C})^{\otimes}$, where $\mathrm{Env}\,\mathscr{C}^{\otimes}$ is the symmetric monoidal envelope of $\mathscr{C}^{\otimes}$, and the second arrow is the Yoneda embedding.

Finally, since every ∞-operad is a sufficiently filtered union of small suboperads, the thesis is true for any ∞-operad. $\square$

References

[Ayala and Francis 2020] **D Ayala**, **J Francis**, *Fibrations of ∞-categories*, High. Struct. 4 (2020) 168–265 MR Zbl

[Bachmann and Hoyois 2021] **T Bachmann**, **M Hoyois**, *Norms in motivic homotopy theory*, Astérisque 425, Soc. Math. France, Paris (2021) MR Zbl

[Barnea et al. 2017] **I Barnea**, **Y Harpaz**, **G Horel**, *Pro-categories in homotopy theory*, Algebr. Geom. Topol. 17 (2017) 567–643 MR Zbl

[Barwick 2017] **C Barwick**, *Spectral Mackey functors and equivariant algebraic K-theory, I*, Adv. Math. 304 (2017) 646–727 MR Zbl

[Ben-Moshe and Schlank 2024] **S Ben-Moshe**, **T M Schlank**, *Higher semiadditive algebraic K-theory and redshift*, Compos. Math. 160 (2024) 237–287 MR Zbl

[Berman 2024] **J D Berman**, *On lax limits in ∞-categories*, Proc. Amer. Math. Soc. 152 (2024) 5055–5066 MR Zbl

[Bohmann 2014] **A M Bohmann**, *Global orthogonal spectra*, Homology Homotopy Appl. 16 (2014) 313–332 MR Zbl

[Cisinski 2019] **D-C Cisinski**, *Higher categories and homotopical algebra*, Cambridge Stud. Adv. Math. 180, Cambridge Univ. Press (2019) MR Zbl

[Degrijse et al. 2023] **D Degrijse**, **M Hausmann**, **W Lück**, **I Patchkoria**, **S Schwede**, *Proper equivariant stable homotopy theory*, Mem. Amer. Math. Soc. 1432, Amer. Math. Soc., Providence, RI (2023) MR Zbl

[Drew and Gallauer 2022] **B Drew**, **M Gallauer**, *The universal six-functor formalism*, Ann. K-Theory 7 (2022) 599–649 MR Zbl

[Elmendorf 1983] **A D Elmendorf**, *Systems of fixed point sets*, Trans. Amer. Math. Soc. 277 (1983) 275–284 MR Zbl

[Gepner and Henriques 2007] **D Gepner**, **A Henriques**, *Homotopy theory of orbispaces*, preprint (2007) arXiv math/0701916

[Gepner and Meier 2023] **D Gepner**, **L Meier**, *On equivariant topological modular forms*, Compos. Math. 159 (2023) 2638–2693 MR Zbl

[Gepner et al. 2017] **D Gepner**, **R Haugseng**, **T Nikolaus**, *Lax colimits and free fibrations in ∞-categories*, Doc. Math. 22 (2017) 1225–1266 MR Zbl

[Glasman 2014] **S Glasman**, *A note on commutative algebras and their modules in quasicategories*, preprint (2014) arXiv 1409.3506

[Glasman 2016] **S Glasman**, *Day convolution for ∞-categories*, Math. Res. Lett. 23 (2016) 1369–1385 MR Zbl

[Greenlees and May 1997] **J P C Greenlees**, **J P May**, *Localization and completion theorems for MU-module spectra*, Ann. of Math. 146 (1997) 509–544 MR Zbl

[Hausmann 2019] **M Hausmann**, *Symmetric spectra model global homotopy theory of finite groups*, Algebr. Geom. Topol. 19 (2019) 1413–1452 MR Zbl

[Hausmann 2022] **M Hausmann**, *Global group laws and equivariant bordism rings*, Ann. of Math. 195 (2022) 841–910 MR Zbl

[Hinich 2015] **V Hinich**, *Rectification of algebras and modules*, Doc. Math. 20 (2015) 879–926 MR Zbl

[Hinich 2016] **V Hinich**, *Dwyer–Kan localization revisited*, Homology Homotopy Appl. 18 (2016) 27–48 MR Zbl

[Hovey 1999] **M Hovey**, *Model categories*, Math. Surv. Monogr. 63, Amer. Math. Soc., Providence, RI (1999) MR Zbl

[Joyal 2008] **A Joyal**, *The theory of quasi-categories and its applications*, preprint (2008) Available at https://ncatlab.org/nlab/files/JoyalTheoryOfQuasiCategories.pdf

[Juran 2020] **B Juran**, *Orbifolds, orbispaces and global homotopy theory*, preprint (2020) arXiv 2006.12374

[Körschgen 2018] **A Körschgen**, *A comparison of two models of orbispaces*, Homology Homotopy Appl. 20 (2018) 329–358 MR Zbl

[Lenz 2022] **T Lenz**, *Global algebraic K-theory is Swan K-theory*, preprint (2022) arXiv 2202.07272

[Lenz 2025] **T Lenz**, *G-global homotopy theory and algebraic K-theory*, Mem. Amer. Math. Soc. 1545, Amer. Math. Soc., Providence, RI (2025) MR Zbl

[Lewis et al. 1986] **L G Lewis, Jr**, **J P May**, **M Steinberger**, **J E McClure**, *Equivariant stable homotopy theory*, Lecture Notes in Math. 1213, Springer (1986) MR Zbl

[Lurie 2009] **J Lurie**, *Higher topos theory*, Ann. of Math. Stud. 170, Princeton Univ. Press (2009) MR Zbl

[Lurie 2017] **J Lurie**, *Higher algebra*, preprint (2017) Available at https://url.msp.org/Lurie-HA

[Mandell and May 2002] **M A Mandell**, **J P May**, *Equivariant orthogonal spectra and S-modules*, Mem. Amer. Math. Soc. 755, Amer. Math. Soc., Providence, RI (2002) MR Zbl

[Nguyen et al. 2020] **H K Nguyen**, **G Raptis**, **C Schrade**, *Adjoint functor theorems for ∞-categories*, J. Lond. Math. Soc. 101 (2020) 659–681 MR Zbl

[Rezk 2014] **C Rezk**, *Global homotopy theory and cohesion*, preprint (2014) Available at https://ncatlab.org/nlab/files/Rezk14.pdf

[Riehl and Verity 2016] **E Riehl**, **D Verity**, *Homotopy coherent adjunctions and the formal theory of monads*, Adv. Math. 286 (2016) 802–888 MR Zbl

[Schwede 2017] **S Schwede**, *Equivariant properties of symmetric products*, J. Amer. Math. Soc. 30 (2017) 673–711 MR Zbl

[Schwede 2018] **S Schwede**, *Global homotopy theory*, New Math. Monogr. 34, Cambridge Univ. Press (2018) MR Zbl

[Schwede 2020] **S Schwede**, *Orbispaces, orthogonal spaces, and the universal compact Lie group*, Math. Z. 294 (2020) 71–107 MR Zbl

[Schwede and Shipley 2000] **S Schwede**, **B E Shipley**, *Algebras and modules in monoidal model categories*, Proc. Lond. Math. Soc. 80 (2000) 491–511 MR Zbl

[Shah 2021] **J Shah**, *Parametrized higher category theory, II: Universal constructions*, preprint (2021) arXiv 2109.11954

[Shah 2023] **J Shah**, *Parametrized higher category theory*, Algebr. Geom. Topol. 23 (2023) 509–644 MR Zbl

[Stephan 2016] **M Stephan**, *On equivariant homotopy theory for model categories*, Homology Homotopy Appl. 18 (2016) 183–208 MR Zbl

Fakultät für Mathematik, Universität Regensburg
Regensburg, Germany

sil.linskens@mathematik.uni-regensburg.de, den.nardin@gmail.com, luca.pol@ur.de

Proposed: Jesper Grodal
Seconded: Mark Behrens, Haynes R Miller

Received: 22 June 2022
Revised: 5 March 2024

Geometry & Topology **29**:3 (2025) 1441–1488
DOI: 10.2140/gt.2025.29.1441
Published: 31 May 2025

An h-principle for complements of discriminants

ALEXIS AUMONIER

We compare spaces of nonsingular algebraic sections of ample vector bundles to spaces of continuous sections of jet bundles. Under some conditions, we provide an isomorphism in homology in a range of degrees growing with the jet ampleness. As an application, when $\mathcal{L}$ is a very ample line bundle on a smooth projective complex variety, we prove that the rational cohomology of the space of nonsingular algebraic sections of $\mathcal{L}^{\otimes d}$ stabilises as $d \to \infty$ and compute the stable cohomology. We also prove that the integral homology does not stabilise, using tools from stable homotopy theory.

55R80; 14J10, 14J70

1 Introduction

Our purpose here is to study spaces of nonsingular holomorphic sections of vector bundles by comparing them to spaces of continuous sections of appropriate jet bundles. The latter are particularly amenable to computations using tools from homotopy theory.

Given a holomorphic line bundle $\mathcal{L}$ on a smooth projective complex variety X, one may consider the vector space of all holomorphic global sections $\Gamma_{\mathrm{hol}}(X;\mathcal{L})$. To each section $s \in \Gamma_{\mathrm{hol}}(X;\mathcal{L})$ is associated a geometric object, its vanishing set

$$V(s) := \{x \in X \mid s(x) = 0\} \subset X,$$

and s is called nonsingular whenever its derivative $ds \in \Gamma_{\mathrm{hol}}(\Omega^1_X \otimes \mathcal{L})$ does not vanish on $V(s)$. This implies in particular that $V(s)$ is a smooth subvariety of X. It has been known for a century now that when $\mathcal{L}$ is a very ample line bundle, the Bertini theorem implies that a generic section is nonsingular. There is thus a Zariski open subset

$$\Gamma_{\mathrm{hol,ns}}(X;\mathcal{L}) \subset \Gamma_{\mathrm{hol}}(X;\mathcal{L})$$

consisting of those nonsingular sections, which geometrically can be interpreted as a moduli space of equations of certain smooth hypersurfaces in X. A prime example is the space $\Gamma_{\mathrm{hol,ns}}(\mathbb{CP}^n;\mathcal{O}(d))$ (sometimes modded out by $\mathbb{C}^*$ or $\mathrm{GL}_{n+1}(\mathbb{C})$) of smooth hypersurfaces of degree d in the complex projective space $\mathbb{CP}^n$.

The cohomology ring of $\Gamma_{\mathrm{hol,ns}}(X;\mathcal{L})$, sometimes known as the ring of characteristic classes, is therefore an important object in the study of hypersurface bundles. We give a way of computing it in a range.

Before revealing our main theorem, we will extend the classical situation above in two directions. To begin, instead of limiting ourselves to line bundles, we will look at sections of bundles of possibly higher rank. Furthermore, we observe that being nonsingular imposes conditions on the value and derivative of a global section. We will generalise this situation by looking at a broader class of conditions on higher-order derivatives, thus encompassing various other flavours of moduli spaces: hypersurfaces with simple nodes, smooth complete intersections, etc.

Having said this, let X be a smooth projective complex variety and $\mathcal{E}$ be a holomorphic vector bundle on X. One can construct a new holomorphic vector bundle $J^r\mathcal{E}$, called the r^{th} jet bundle of $\mathcal{E}$, together with a map on global sections $j^r\colon \Gamma_{\text{hol}}(\mathcal{E}) \to \Gamma_{\text{hol}}(J^r\mathcal{E})$. Intuitively, for a section s of $\mathcal{E}$, the associated section $j^r(s)$ of the jet bundle records all derivatives of s up to order r. For $\mathfrak{T} \subset J^r\mathcal{E}$ a subset which we think of as "forbidden derivatives", we say that a section s of $\mathcal{E}$ is nonsingular if $j^r(s)(x) \notin \mathfrak{T}$ for all $x \in X$. For instance, when $\mathcal{E}$ is a line bundle and $\mathfrak{T} \subset J^1\mathcal{E}$ is the zero section, we recover the classical notion of nonsingular sections discussed at the beginning of this article.

Theorem 1.1 (see Theorem 2.13 for full generality) *Let X be a smooth complex projective variety and $\mathcal{E}$ be a holomorphic vector bundle on it. Let $r \geq 0$ be an integer and $\mathfrak{T} \subset J^r\mathcal{E}$ be a closed subvariety of the r^{th} jet bundle of $\mathcal{E}$ of codimension at least $\dim_{\mathbb{C}} X + 1$. We write*

$$\Gamma_{\text{hol,ns}}(\mathcal{E}) := \{s \in \Gamma_{\text{hol}}(\mathcal{E}) \mid \forall x \in X,\ j^r(s)(x) \notin \mathfrak{T}\}$$

for the space of nonsingular holomorphic sections of $\mathcal{E}$. If $\mathcal{E}$ is d-jet ample, the composition

$$\Gamma_{\text{hol,ns}}(\mathcal{E}) \xrightarrow{j^r} \Gamma_{\text{hol}}(J^r\mathcal{E} - \mathfrak{T}) \hookrightarrow \Gamma_{\mathcal{C}^0}(J^r\mathcal{E} - \mathfrak{T})$$

induces an isomorphism in integral homology in the range of degrees $* < (d-r)/(r+1)$.

The theorem above can be strengthened, and in Section 2 we introduce a more general class of allowed subsets $\mathfrak{T} \subset J^r\mathcal{E}$ of the jet bundle as well as give a sharper range of degrees. We also take advantage of that section to give the definition of jet ampleness and jet bundles in algebraic geometry.

Remark 1.2 By the Whitney approximation theorem, the spaces of continuous ($\mathcal{C}^0$) sections and smooth ($\mathcal{C}^\infty$) sections of a fibre bundle are homotopy equivalent. Thus, in our main theorem, instead of first taking the jet and then including inside the continuous sections, we could have tried to argue in the reverse order:

$$\begin{tikzcd} \Gamma_{\text{hol,ns}}(\mathcal{E}) \arrow[r] \arrow[d, hook] & \Gamma_{\text{hol}}(J^r\mathcal{E} - \mathfrak{T}) \arrow[d, hook] \\ \Gamma_{\mathcal{C}^\infty,\text{ns}}(\mathcal{E}) \arrow[r] & \Gamma_{\mathcal{C}^\infty}(J^r\mathcal{E} - \mathfrak{T}) \end{tikzcd}$$

We caution the reader about one subtle point: $J^r\mathcal{E}$ is the holomorphic jet bundle of $\mathcal{E}$, which does not agree with the smooth jet bundle of the underlying real vector bundle of $\mathcal{E}$. In particular $\Gamma_{\mathcal{C}^\infty,\text{ns}}(\mathcal{E})$ is defined analogously to its holomorphic counterpart by imposing conditions on the *complex* derivatives of smooth sections. It seems likely that the map $\Gamma_{\mathcal{C}^\infty,\text{ns}}(\mathcal{E}) \to \Gamma_{\mathcal{C}^\infty}(J^r\mathcal{E} - \mathfrak{T})$ can be studied using the same arguments as given by Vassiliev in the real case [29], but we shall not comment further on that matter.

1.1 Motivations and applications

Our main theorem can be seen as a holomorphic analogue of the work of Vassiliev on spaces of maps with mild singularities [29, Chapter III]. In another current of thought, we should also mention the seminal work of Segal [25] on spaces of rational maps, where the idea was born that holomorphic maps should approximate continuous ones, and that this approximation becomes better with the growth of ampleness.

We were also very much influenced by the work of Vakil and Wood on stability results in the Grothendieck ring of varieties [28]. There they conjectured that for a very ample line bundle $\mathcal{L}$ on a smooth projective complex variety, the space of nonsingular sections of $\mathcal{L}^{\otimes d}$ should exhibit cohomological stability. In the special case of the projective space, Tommasi obtained the following result:

Theorem 1.3 (Tommasi [27]) *Let $d, n \geq 1$ be integers. Let $U_{d,n}$ be the space of nonsingular holomorphic sections of $\mathcal{O}(d)$ on $\mathbb{CP}^n$. The rational cohomology of $U_{d,n}$ is isomorphic to the rational cohomology of the space $\mathrm{GL}_{n+1}(\mathbb{C})$ in degrees $* < \frac{1}{2}(d+1)$.*

She furthermore has investigated an extension of this result to arbitrary smooth projective varieties (personal communication, 2021). Using different techniques, O Banerjee also confirmed the conjecture of Vakil and Wood in the case of smooth projective curves [2].

The present work was strongly motivated by the result of Tommasi and the wish to understand the stable cohomology from a more homotopy-theoretic point of view. At the time of writing, let us in particular mention the following result:

Theorem 1.4 (Tommasi, personal communication, 2021) *Let X be a smooth projective complex variety of dimension n and $\mathcal{L}$ be a very ample line bundle on X. Let $d \geq 1$ be an integer and U_d be the space of nonsingular holomorphic sections of $\mathcal{L}^{\otimes d}$. There is a Vassiliev spectral sequence converging to the homology of U_d. Working with rational coefficients, this spectral sequence degenerates on the E_2-page in the stable range if and only if the stable cohomology is a free commutative graded algebra on the cohomology of X shifted by one degree.*

Assuming this degeneration, the rational cohomology of U_d in degrees $* < \lfloor \frac{1}{2}(d+1) \rfloor$ is given by the free commutative graded algebra $\Lambda(H^{*-1}(X;\mathbb{Q}))$ on the cohomology of X shifted by one degree.

In the last section (Section 8) we apply our main theorem to spaces of smooth hypersurfaces to prove a homological stability result with rational coefficients.

Theorem 1.5 (see Theorem 8.2) *Let X be a smooth projective complex variety and $\mathcal{L}$ be a very ample line bundle on X. The rational cohomology ring of the space $\Gamma_{\mathrm{hol,ns}}(\mathcal{L}^d)$ of nonsingular sections (in the classical sense) of the d^{th} tensor power of $\mathcal{L}$ is isomorphic to $\Lambda(H^{*-1}(X;\mathbb{Q}))$ in degrees $* < \frac{1}{2}(d-1)$.*

Firstly, this agrees with the work in progress of Tommasi. In fact, one can use our main theorem to show the degeneration of the Vassiliev spectral sequence she constructed. Secondly, in contrast to many other instances of homological stability, note that there are no natural stabilisation maps from spaces of nonsingular sections of $\mathcal{L}^d$ to those of $\mathcal{L}^{d+1}$. Thus we only mean that the cohomology rings abstractly stabilise, and the answer only depends on X and not on $\mathcal{L}$. After the apparition of the first version of the present article, and using different tools, Das and Howe proved a version of the above theorem for hypersurfaces in algebraic varieties over any algebraically closed field [11].

On the other hand, we find it quite interesting that there is in general no integral homological stability. In fact, we prove the following result about the moduli space of smooth hypersurfaces of degree d in $\mathbb{CP}^2$:

Proposition 8.10 *Let $d \geq 6$ be an integer. We have*

$$H_2(\Gamma_{\mathrm{hol,ns}}(\mathbb{CP}^2, \mathcal{O}(d)); \mathbb{Z}/2) \cong \begin{cases} \mathbb{Z}/2 & d \equiv 0 \bmod 2, \\ 0 & d \equiv 1 \bmod 2. \end{cases}$$

Besides the phenomenon this result illustrates, its proof showcases the potential of homotopical methods allowed by our main theorem. Indeed, the computation comes down to simple manipulations of Steenrod squares where the parity of d is reflected in the Chern class of $\mathcal{O}(d)$. In contrast, a more classical approach following the work of Vassiliev [30] would require knowledge of nontrivial differentials in spectral sequences that quickly grow out of hand when d increases.

For good measure, we also study the p-torsion in the homology of $\Gamma_{\mathrm{hol,ns}}(\mathcal{L}^d)$ and show that it stabilises when $p \geq \dim_{\mathbb{C}} X + 2$ and $d \to \infty$; see Proposition 8.15.

Our results are also inspired by analogies with theorems in arithmetic statistics, such as Poonen's Bertini theorem over finite fields [24], and in motivic statistics in the Grothendieck ring of varieties, as in [28] or [6]. The recent results of Bilu and Howe particularly influenced the current formulation of our main theorem and we would like to recommend the introduction of their paper [6] to the reader interested in an overview of these analogies. Finally, we also wish to mention that I Banerjee recently announced a result relating nonsingular sections of a line bundle on an algebraic curve and smooth sections of the same line bundle [1].

Acknowledgements

I would like to thank my PhD advisor Søren Galatius for suggesting to compare algebraic sections to continuous sections of jet bundles. It is a pleasure to thank him for his encouragement and many helpful discussions. I would also like to thank Orsola Tommasi for discussing and sharing her work with me, as well as Ronno Das for helpful discussions related to this project. I was supported by the Danish National Research Foundation through the Copenhagen Centre for Geometry and Topology (DNRF151) as well as the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 682922).

2 Statement of the main theorem

We begin with a few preliminary definitions before stating precisely our main theorem. Throughout this article, X is a smooth projective complex variety and $\mathcal{E}$ is a holomorphic vector bundle on X. We denote by Γ the space of sections of a vector bundle, and decorate it with subscripts "hol" or "$\mathcal{C}^0$" to indicate holomorphic or continuous sections, respectively. We will make extensive use of cohomology with compact support, which we denote by H_c^*, and refer to [8] for its definition and standard properties. All homology and cohomology groups will be taken with integral coefficients, unless otherwise specified. We recall the following definition of jet ampleness:

Definition 2.1 (compare [4]) Let $k \geq 0$ be an integer. Let $x_1, \dots, x_t$ be t distinct points in X and $(k_1, \dots, k_t)$ be a t-tuple of positive integers such that $\sum_i k_i = k+1$. Denote by $\mathcal{O}$ the structure sheaf of X and by $\mathfrak{m}_i$ the maximal ideal sheaf corresponding to x_i. We regard the tensor product $\bigotimes_{i=1}^t \mathfrak{m}_i^{k_i}$ as a subsheaf of $\mathcal{O}$ under the multiplication map $\bigotimes_{i=1}^t \mathfrak{m}_i^{k_i} \to \mathcal{O}$. We say that $\mathcal{E}$ is *k-jet ample* if the evaluation map

$$\Gamma_{\mathrm{hol}}(\mathcal{E}) \to \Gamma_{\mathrm{hol}}\Bigl(\mathcal{E} \otimes \Bigl(\mathcal{O}/\bigotimes_{i=1}^t \mathfrak{m}_i^{k_i}\Bigr)\Bigr) \cong \bigoplus_{i=1}^t \Gamma_{\mathrm{hol}}(\mathcal{E} \otimes (\mathcal{O}/\mathfrak{m}_i^{k_i}))$$

is surjective for any $x_1, \dots, x_t$ and $k_1, \dots, k_t$ as above.

Example 2.2 A vector bundle $\mathcal{E}$ is 0-jet ample if and only if it is spanned by its global sections. In the case of a line bundle, 1-jet ampleness corresponds to the usual notion of very ampleness. On a curve, a line bundle is k-jet ample whenever it is k-very ample. However, on higher-dimensional varieties, a k-jet ample line bundle is also k-very ample but the converse is not true in general. Finally, and most importantly for us, if $\mathcal{A}$ and $\mathcal{B}$ are holomorphic vector bundles which are a- and b-jet ample, respectively, then their tensor product $\mathcal{A} \otimes \mathcal{B}$ is $(a+b)$-jet ample; see [4, Proposition 2.3].

To ease the readability of various statements we will use the following notation:

Definition 2.3 For a holomorphic vector bundle $\mathcal{E}$ on X and an integer $r \in \mathbb{N}$, we define $N(\mathcal{E}, r) \geq 0$ to be the largest integer N such that $\mathcal{E}$ is $((N+1)(r+1)-1)$-jet ample. If no such integer exists, we set $N(\mathcal{E}, r) = -1$, although we do not consider such cases here.

Let us also recall the construction of the jet bundle from [14, IV.16.7], where it is called the sheaf of principal parts. The diagonal morphism $\Delta\colon X \to X \times X$ gives a surjection of sheaves $\Delta^\sharp\colon \Delta^*\mathcal{O}_{X\times X} \to \mathcal{O}_X$. Denoting by $\mathcal{I}$ the kernel, $\mathcal{O}_X \cong \Delta^*\mathcal{O}_{X\times X}/\mathcal{I}$. For an integer $r \geq 0$, we define the r^{th} jet bundle of $\mathcal{O}_X$ to be

$$J^r\mathcal{O}_X := \Delta^*\mathcal{O}_{X\times X}/\mathcal{I}^{r+1}.$$

The projections $p_i\colon X \times X \to X$ give two $\mathcal{O}_X$-algebra structures on $J^r\mathcal{O}_X$ and, unless otherwise specified, we use the one given by the first projection p_1. The other morphism induced by p_2 is denoted by

$$d_X^r\colon \mathcal{O}_X \to J^r\mathcal{O}_X.$$

For a holomorphic vector bundle $\mathcal{E}$ on X, we define its r^{th} *jet bundle* to be

$$J^r\mathcal{E} := J^r\mathcal{O}_X \otimes_{\mathcal{O}_X} \mathcal{E}, \tag{1}$$

where $J^r\mathcal{O}_X$ is seen as an $\mathcal{O}_X$-module via the morphism d^r_X for the tensor product, and the result is regarded as an $\mathcal{O}_X$-module again via p_1. It comes with the morphism

$$d^r_{X,\mathcal{E}} := d^r_X \otimes \mathcal{E} \colon \mathcal{E} \to J^r\mathcal{O}_X \otimes_{\mathcal{O}_X} \mathcal{E} = J^r\mathcal{E}.$$

Taking global sections, we obtain the *jet map*:

$$j^r = \Gamma(d^r_{X,\mathcal{E}}) \colon \Gamma_{\mathrm{hol}}(\mathcal{E}) \to \Gamma_{\mathrm{hol}}(J^r\mathcal{E}). \tag{2}$$

The most important observation for us is that if $x \in X$ is a point with maximal ideal sheaf $\mathfrak{m}$, the fibre $(J^r\mathcal{E})|_x$ is naturally identified with the complex vector space $\mathcal{E}_x/\mathfrak{m}_x^{r+1}\mathcal{E}_x$. Furthermore, the composition

$$\mathcal{E}_x \xrightarrow{(d^r_{X,\mathcal{E}})_x} (J^r\mathcal{E})_x \to (J^r\mathcal{E})|_x = \mathcal{E}_x/\mathfrak{m}_x^{r+1}\mathcal{E}_x$$

is the natural quotient map. (Here, and everywhere else, we write $\mathcal{E}_x$ for the stalk of the sheaf $\mathcal{E}$ and $\mathcal{E}|_x = \mathcal{E}_x/\mathfrak{m}_x\mathcal{E}_x$ for the fibre of the bundle $\mathcal{E}$.) Intuitively, for a holomorphic section s of $\mathcal{E}$, one should think of the value of $j^r(s)$ at a point $x \in X$ as the tuple of all derivatives of s at x up to order r. In particular, the following lemma is a direct consequence of the definitions:

Lemma 2.4 *Let $\mathcal{E}$ be a holomorphic vector bundle on X and let $N(\mathcal{E}, r)$ be as in Definition 2.3. Let $(x_0, \dots, x_p)$ be a tuple of $p+1$ distinct points in X. If $p \le N(\mathcal{E}, r)$, the simultaneous evaluation of the jet map (2) at these points,*

$$j^r_{(x_0,\dots,x_p)} \colon \Gamma_{\mathrm{hol}}(\mathcal{E}) \to (J^r\mathcal{E})|_{x_0} \times \cdots \times (J^r\mathcal{E})|_{x_p}, \quad s \mapsto (j^r(s)(x_0), \dots, j^r(s)(x_p)),$$

is surjective. □

We shall now explain what we precisely mean by restricting the behaviour of sections of $\mathcal{E}$. In particular, we will require certain subsets of the jet bundle to be "semialgebraic". This is a technical condition chosen for two reasons: to make the proofs of Section 4 precise, and to prove our main theorem in a good degree of generality. Our arguments rely on multiple properties of these sets: they admit cell decompositions, have a well-defined dimension and they behave well under projections and closure. (See Section 4.2 for their single but crucial use.)[1]

There is a well-studied concept of real semialgebraic subsets of a Euclidean space. They are subsets defined by polynomial equations and inequalities.

Definition 2.5 (compare [7]) A semialgebraic subset of $\mathbb{R}^n$ is a union of finitely many subsets of the form

$$\{x \in \mathbb{R}^n \mid P(x) = 0,\ Q_1(x) > 0, \dots, Q_l(x) > 0\},$$

where $l \in \mathbb{N}$ and $P, Q_1, \dots, Q_l \in \mathbb{R}[X_1, \dots, X_n]$.

[1]As the referee pertinently pointed out, the semialgebraicity conditions could potentially be rephrased in the language of o-minimal structures. We have refrained to do so to keep the technicalities to a minimum.

We adapt the definition to families, ie to subsets of vector bundles, by demanding the standard definition be satisfied locally in charts. This is well defined because an algebraic variety X has an atlas whose transition functions are algebraic, and hence respect the semialgebraicity.

Let us be more precise. First, we briefly recall the notion of an algebraic atlas on X. To lighten the notation, we let n be the complex dimension of X and m be the complex rank of $J^r\mathcal{E}$. We denote by $V(-)$ the vanishing set of the tuple of polynomials.

The variety X can be covered by Zariski open subsets, each of the form

$$U \cong V(f_1, \dots, f_{d-n}) \subset \mathbb{C}^d$$

for some integer $d \geq 1$ and polynomials $f_1, \dots, f_{d-n}$. Furthermore, if U and W are Zariski open subsets of X with $\alpha\colon U \cong V(f_1, \dots, f_{d-n}) \subset \mathbb{C}^d$ and $\beta\colon W \cong V(g_1, \dots, g_{d'-n}) \subset \mathbb{C}^{d'}$, the homeomorphism on the intersection

$$\alpha(W \cap U) \cap V(f_1, \dots, f_{d-n}) \xrightarrow{\cong} W \cap U \xrightarrow{\cong} \beta(U \cap W) \cap V(g_1, \dots, g_{d'-n})$$

is given by a rational function whose domain is a subset of $\mathbb{C}^d$ and codomain is a subset of $\mathbb{C}^{d'}$. Recall also that the algebraic vector bundle $J^r\mathcal{E}$ is equivalently given by the data of trivialising Zariski open subsets $U_i \subset X$ (over which $J^r\mathcal{E}|_{U_i} \cong U_i \times \mathbb{C}^m$) and transition functions on overlaps $U_i \cap U_j \to \mathrm{GL}_m(\mathbb{C})$. Most importantly for us, the transition functions are regular morphisms.

Definition 2.6 Let n be the complex dimension of X and m be the complex rank of $J^r\mathcal{E}$. A subset $\mathfrak{T} \subset J^r\mathcal{E}$ is *real semialgebraic* if there exists a cover $X = \bigcup U_i$ by Zariski open subsets such that the following conditions hold for each i:

(i) The jet bundle may be trivialised over U_i via a map $\varphi_i : J^r\mathcal{E}|_{U_i} \xrightarrow{\cong} U_i \times \mathbb{C}^m$.

(ii) There is a chart $\phi_i : U_i \xrightarrow{\cong} V(f_1^i, \dots, f_{d_i-n}^i) \subset \mathbb{C}^{d_i}$ for some polynomials $f_1^i, \dots, f_{d_i-n}^i$.

(iii) The image in $\mathbb{R}^{2(d_i+m)}$ of $\mathfrak{T}|_{U_i}$ via the map

$$J^r\mathcal{E}|_{U_i} \xrightarrow{\varphi_i} U_i \times \mathbb{C}^m \xrightarrow{\phi_i \times \mathrm{id}} V(f_1^i, \dots, f_{d_i-n}^i) \times \mathbb{C}^m \subset \mathbb{C}^{d_i+m} \cong \mathbb{R}^{2(d_i+m)}$$

is a real semialgebraic subset. (Here $\mathfrak{T}|_{U_i}$ is the restriction of $\mathfrak{T}$ above U_i.)

We will often drop the adjective "real" as we will never consider any complex analogue. In essence, a subset $\mathfrak{T} \subset J^r\mathcal{E}$ is semialgebraic in the sense of Definition 2.6 when it is semialgebraic in the usual way when "read in charts". As all the change-of-coordinates maps described above are rational functions, being semialgebraic is independent of the choice of the cover. Indeed, the image of a semialgebraic set by a rational function is still semialgebraic; see [7, Section 2.2].

A semialgebraic subset has a well-defined dimension (as in [7, Section 2.8]) which can be thought of as the maximal dimension in a decomposition into cells of the form $]0,1[^d$; see [7, Corollary 2.8.9]. We therefore get a well-defined dimension for a semialgebraic subset $\mathfrak{T} \subset J^r\mathcal{E}$ by looking at the dimensions when "reading in charts":

Definition 2.7 Let $\mathfrak{T} \subset J^r\mathcal{E}$ be a semialgebraic subset. Let $X = \bigcup U_i$ be a finite cover as in Definition 2.6 (the finiteness can always be arranged by compactness of X) and write $\mathfrak{T}_{U_i} \subset \mathbb{R}^{2(d_i+m)}$ for the semialgebraic sets obtained using (iii). Each of them has a well-defined dimension and we let the *dimension of* $\mathfrak{T}$ be their maximum.

In the following definition, we denote by $\mathrm{rk}_{\mathbb{C}}\, J^r\mathcal{E}$ the complex rank of $J^r\mathcal{E}$.

Definition 2.8 A subset $\mathfrak{T} \subset J^r\mathcal{E}$ is an *admissible Taylor condition* if it is closed, real semialgebraic and has dimension at most $2(\mathrm{rk}_{\mathbb{C}}\, J^r\mathcal{E} - 1)$. We will use the notation $\mathfrak{T}|_x := (J^r\mathcal{E})|_x \cap \mathfrak{T}$ for the fibre above a point $x \in X$.

Remark 2.9 Although our definition is quite technical and general, the typical admissible Taylor conditions arise as subvarieties of high-enough codimension. Indeed, any closed subvariety $\mathfrak{T} \subset J^r\mathcal{E}$ of the jet bundle of complex codimension at least $\dim_{\mathbb{C}} X + 1$ defines an admissible Taylor condition.

Motivated by the previous remark, and to help general bookkeeping, we will use the following notation:

Definition 2.10 The (real) *excess codimension* of an admissible Taylor condition $\mathfrak{T}$ is the number $e(\mathfrak{T}) = \mathrm{codim}_{\mathbb{R}}\, \mathfrak{T} - \dim_{\mathbb{R}} X \geq 2$, where $\mathrm{codim}_{\mathbb{R}}\, \mathfrak{T}$ is the real codimension of $\mathfrak{T}$ in the jet bundle $J^r\mathcal{E}$.

We are now ready to define what it means for a section to be singular with respect to an admissible Taylor condition $\mathfrak{T}$:

Definition 2.11 A holomorphic section s of the vector bundle $\mathcal{E}$ is said to be *singular* if there exists a point $x \in X$ such that $j^r(s)(x) \in \mathfrak{T}|_x$. Similarly, a (continuous) section s of the vector bundle $J^r\mathcal{E}$ is said to be *singular* if there exists a point $x \in X$ such that $s(x) \in \mathfrak{T}|_x$. A section that is not singular is said to be *nonsingular*.

Example 2.12 If $\mathcal{E}$ is a line bundle, we may take $\mathfrak{T}$ to be the zero section of $J^1\mathcal{E}$. It is an admissible Taylor condition, and a singular section is one that vanishes at a point on X where its derivative also vanishes. In particular, if s is a nonsingular section, its zero set $Z(s) := \{x \in X \mid s(x) = 0\} \subset X$ is a smooth submanifold.

When talking about spaces of sections Γ, we will use the subscript "ns" to denote the subspace of ***n**onsingular* sections. The following is our main result:

Theorem 2.13 *Let $r \geq 0$ and $N \geq 1$ be integers. Let $\mathcal{E}$ be an $((N+1)(r+1)-1)$-jet ample vector bundle on X and let $\mathfrak{T} \subset J^r\mathcal{E}$ be an admissible Taylor condition. The composition*

$$\Gamma_{\mathrm{hol,ns}}(\mathcal{E}) \xrightarrow{j^r} \Gamma_{\mathrm{hol,ns}}(J^r\mathcal{E}) \hookrightarrow \Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^r\mathcal{E})$$

induces an isomorphism in homology

$$H_*(\Gamma_{\mathrm{hol,ns}}(\mathcal{E}); \mathbb{Z}) \to H_*(\Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^r\mathcal{E}); \mathbb{Z})$$

in the range of degrees $* < N(e(\mathfrak{T}) - 1) + e(\mathfrak{T}) - 2$.

2.1 Outline

We present here a detailed summary of the arguments exposed in Sections 3–7. We will produce a sequence of vector spaces

$$\Gamma_{-1} \to \Gamma_0 \to \Gamma_1 \to \Gamma_2 \to \cdots \to \Gamma_\infty$$

where $\Gamma_{-1} = \Gamma_{\mathrm{hol}}(\mathcal{E})$, $\Gamma_\infty = \Gamma_{\mathcal{C}^0}(J^r\mathcal{E})$ and $\Gamma_j = \Gamma_{\mathrm{hol}}(J^r\mathcal{E} \otimes \mathcal{L}^j) \otimes \overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^j)}$ for $0 \le j < \infty$ (see Section 5). There is a discriminant $\Sigma_\infty \subset \Gamma_\infty$ inducing discriminants $\Sigma_j \subset \Gamma_j$ by preimage, and such that $\Gamma_{\mathrm{hol,ns}}(\mathcal{E}) = \Gamma_{-1} - \Sigma_{-1}$ and $\Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^r\mathcal{E}) = \Gamma_\infty - \Sigma_\infty$. For $n < m < \infty$ one gets a map of long exact sequences

$$\begin{tikzcd}
\cdots \arrow[r] & H^i_c(\Gamma_n - \Sigma_n) \arrow[r] \arrow[d] & H^i_c(\Gamma_n) \arrow[r] \arrow[d] & H^i_c(\Sigma_n) \arrow[r] \arrow[d] & H^{i+1}_c(\Gamma_n - \Sigma_n) \arrow[r] \arrow[d] & \cdots \\
\cdots \arrow[r] & H^{i+d}_c(\Gamma_m - \Sigma_m) \arrow[r] & H^{i+d}_c(\Gamma_m) \arrow[r] & H^{i+d}_c(\Sigma_m) \arrow[r] & H^{i+d+1}_c(\Gamma_m - \Sigma_m) \arrow[r] & \cdots
\end{tikzcd}$$

where $d = \dim_{\mathbb{R}} \Gamma_m - \dim_{\mathbb{R}} \Gamma_n$. This only relies on a few properties: each Γ_j is a finite-dimensional vector space, Σ_j is a closed subset satisfying mild point–set conditions and the inverse image of Γ_m is Γ_n. The five lemma shows that $H^i_c(\Sigma_n) \to H^{i+d}_c(\Sigma_m)$ needs to be proven to be an isomorphism in a range for the main theorem to follow. This will be done one step at a time by showing that one gets a map of spectral sequences associated to resolutions of these discriminant loci, which induces an isomorphism on the first page in a range. To finish the argument, we invoke the Stone–Weierstrass theorem to prove a weak homotopy equivalence $\mathrm{colim}_j(\Gamma_j - \Sigma_j) \simeq \Gamma_\infty - \Sigma_\infty$.

We construct the resolution of a discriminant locus and its associated spectral sequence in Section 3. We study in details its first page in Section 4. In Section 5, we construct the Γ_j interpolating between holomorphic and continuous sections. In Section 6, we construct a morphism of spectral sequences and use it to compare various spaces of sections. We prove the last homotopy equivalence and finish proving our main theorem in Section 7. Lastly, in Section 8, we apply our results to study spaces of nonsingular sections of a very ample line bundle on a projective variety.

3 Resolution of singularities

In this section, we choose an admissible Taylor condition $\mathfrak{T} \subset J^r\mathcal{E}$ inside the r^{th} jet bundle of a holomorphic vector bundle $\mathcal{E}$ on X, and write for brevity

$$\Gamma = \Gamma_{\mathrm{hol}}(\mathcal{E}) \quad \text{and} \quad \Sigma = \Gamma_{\mathrm{hol}}(\mathcal{E}) - \Gamma_{\mathrm{hol,ns}}(\mathcal{E}),$$

for the vector space Γ of all holomorphic sections of $\mathcal{E}$ and its subspace Σ of singular sections. We also define the *singular space* of a section $f \in \Gamma$

$$\mathrm{Sing}(f) := \{x \in X \mid j^r(f)(x) \in \mathfrak{T}\} \subset X \tag{3}$$

as the space of points where f is singular (as in Definition 2.11). Our final goal, Theorem 2.13, is to understand the homology of the space of nonsingular sections $\Gamma_{\mathrm{hol,ns}}(\mathcal{E}) = \Gamma - \Sigma$. By Alexander duality [8, Theorem V.9.3]

$$H_c^i(\Sigma) \cong \widetilde{H}_{2\dim_{\mathbb{C}}\Gamma - i - 1}(\Gamma - \Sigma),$$

it is equivalent to understand the compactly supported cohomology of its complement Σ. To achieve that, we want to construct a spectral sequence converging to $H_c^*(\Sigma)$. This spectral sequence arises from a resolution of the space Σ, which we define in this section.

Remark 3.1 By cohomology with compact support and homology, we shall always technically mean sheaf (co)homology and rely accordingly on the theory exposed in [8]. Of course, all the spaces of interest to us are homologically locally connected and sheaf and singular (co)homologies agree for them.

3.1 Construction of the resolution

We will construct a space $R\mathfrak{X} \twoheadrightarrow \Sigma$ recording points in the singular space $\mathrm{Sing}(f)$. Accordingly, the inverse image of a section $f \in \Sigma$ with $j+1$ singularities will be a j-simplex Δ^j with vertices indexed by the singular points. It will follow that $R\mathfrak{X} \to \Sigma$, or rather a slightly modified construction $R^N_{\mathrm{cone}}\mathfrak{X} \to \Sigma$, induces an isomorphism in cohomology with compact supports. This construction will be advantageously filtered by subspaces $R^j\mathfrak{X}$ related via pushout diagrams resembling the skeletal decomposition of a simplicial space. This filtration will then yield a spectral sequence computing the cohomology of Σ.

This is inspired by the so-called truncated resolution of Mostovoy [21], but written in a more functorial way as in [31].

In what follows, the space Γ is given its canonical topology coming from the fact that it is a finite-dimensional complex vector space. Let F be the category whose objects are the finite sets $[n] := \{0, \dots, n\}$ for $n \geq 0$ and whose morphisms are *all* maps of sets $[n] \to [m]$. Let Top be the category of topological spaces and continuous maps between them. We define the functor

$$\mathfrak{X}\colon \mathsf{F}^{\mathrm{op}} \to \mathsf{Top}, \quad [n] \mapsto \mathfrak{X}[n] := \{(f, s_0, \dots, s_n) \in \Gamma \times X^{n+1} \mid \forall i,\ s_i \in \mathrm{Sing}(f)\}, \tag{4}$$

where $\mathfrak{X}[n]$ is given the subspace topology from $\Gamma \times X^{n+1}$. On morphisms, for a map of sets $g\colon [n] \to [m]$, we define

$$\mathfrak{X}(g)\colon \mathfrak{X}[m] \to \mathfrak{X}[n], \quad (f, s_0, \dots, s_m) \mapsto (f, s_{g(0)}, \dots, s_{g(n)}).$$

For an integer $k \geq 0$, we denote by $\mathsf{F}_{\leq k}$ the full subcategory of F on objects $[n]$ for $n \leq k$. We also write

$$|\Delta^n| = \{(t_0, \dots, t_n) \mid \forall i,\ 0 \leq t_i \leq 1 \text{ and } t_0 + \dots + t_n = 1\} \subset \mathbb{R}^{n+1}$$

for the standard topological n-simplex, and denote by $\partial|\Delta^n|$ its boundary. In particular, the assignment $[n] \mapsto |\Delta^n|$ gives a functor $\mathsf{F} \to \mathsf{Top}$. For an integer $j \geq 0$, we define the j^{th} *geometric realisation of* $\mathfrak{X}$ by the coend

$$R^j\mathfrak{X} := \int^{[n] \in \mathsf{F}_{\leq j}} \mathfrak{X}[n] \times |\Delta^n| = \Big(\bigsqcup_{0 \leq n \leq j} \mathfrak{X}[n] \times |\Delta^n|\Big)\Big/\!\sim, \tag{5}$$

where the equivalence relation $\sim$ is generated by $(\mathfrak{X}(g)(z), t) \sim (z, g_*(t))$ for all maps $g\colon [n] \to [m]$ in F. (Here $g_*\colon |\Delta^n| \to |\Delta^m|$ denotes the usual map induced on the simplices by functoriality.) This is of course reminiscent of the classical geometric realisation of a simplicial space. Note however that here a cell $|\Delta^n|$ in the geometric realisation is indexed by an *unordered* set of singularities, even though the functor $\mathfrak{X}$ is defined using ordered tuples. Indeed, all the permutations $[n] \to [n]$ are valid morphisms in our category F.

Let $j \geq 1$ be an integer. We now describe how $R^j\mathfrak{X}$ may be obtained from $R^{j-1}\mathfrak{X}$ via a pushout diagram. Let L_j be the set

$$L_j := \{(f, s_0, \dots, s_j) \in \Gamma \times X^{j+1} \mid \exists l \neq k \text{ such that } s_l = s_k\} \subset \mathfrak{X}[j], \tag{6}$$

topologised as a subspace of $\mathfrak{X}[j]$. This should be thought of as the analogue of the "latching object" of a simplicial space. We denote by

$$L_j \times_{\mathfrak{S}_{j+1}} |\Delta^j|$$

the quotient space of $L_j \times |\Delta^j|$ by the symmetric group $\mathfrak{S}_{j+1}$ acting on L_j by permuting the singularities s_i, and on $|\Delta^j|$ by permuting the coordinates. Denote by $\hat{\ }$ the omission of an element in a tuple.

Lemma 3.2 *The formula*

$$\begin{aligned}&((f, s_0, \dots, s_j), (t_0, \dots, t_j))\\ &\qquad \mapsto ((f, s_0, \dots, \hat{s}_l, \dots, s_j), (t_0, \dots, t_k + t_l, \dots, \hat{t}_l, \dots, t_j)) \quad \textit{if there exists } k \neq l \textit{ such that } s_l = s_k\end{aligned}$$

gives a well-defined map $L_j \times_{\mathfrak{S}_{j+1}} |\Delta^j| \to R^{j-1}\mathfrak{X}$.

Proof The formula appears ill-defined as we are arbitrarily choosing two indices k and l. The identifications made by the coend formula (5) show that any choice will yield the same class in the quotient. □

Recall that a point $t = (t_0, \dots, t_j) \in |\Delta^j|$ is in the boundary $\partial|\Delta^j|$ if one of its coordinates vanishes. An argument similar to the proof of Lemma 3.2 gives the following:

Lemma 3.3 *The formula*

$$((f, s_0, \dots, s_j), (t_0, \dots, t_j)) \mapsto ((f, s_0, \dots, \hat{s}_l, \dots, s_j), (t_0, \dots, \hat{t}_l, \dots, t_j)) \quad \textit{if } t_l = 0$$

gives a well-defined map $\mathfrak{X}[j] \times_{\mathfrak{S}_{j+1}} \partial|\Delta^j| \to R^{j-1}\mathfrak{X}$. □

Consider the following pushout diagram of spaces:

$$\begin{tikzcd}
L_j \times_{\mathfrak{S}_{j+1}} \partial|\Delta^j| \arrow[r, hook] \arrow[d, hook] \arrow[dr, phantom, "\ulcorner"] & \mathfrak{X}[j] \times_{\mathfrak{S}_{j+1}} \partial|\Delta^j| \arrow[d] \\
L_j \times_{\mathfrak{S}_{j+1}} |\Delta^j| \arrow[r] & (L_j \times_{\mathfrak{S}_{j+1}} |\Delta^j|) \cup (\mathfrak{X}[j] \times_{\mathfrak{S}_{j+1}} \partial|\Delta^j|)
\end{tikzcd}$$

Equivalently, the pushout is the union of the top-right and bottom-left spaces inside $\mathfrak{X}[j] \times_{\mathfrak{S}_{j+1}} |\Delta^j|$. The maps defined in Lemmas 3.2 and 3.3 glue to a continuous map

$$\alpha_{j-1}\colon (L_j \times_{\mathfrak{S}_{j+1}} |\Delta^j|) \cup (\mathfrak{X}[j] \times_{\mathfrak{S}_{j+1}} \partial|\Delta^j|) \to R^{j-1}\mathfrak{X}.$$

The natural map $\mathfrak{X}[j] \times |\Delta^j| \to R^j\mathfrak{X}$ factors through the quotient by the symmetric group action and gives a map

$$\beta_j : \mathfrak{X}[j] \times_{\mathfrak{S}_{j+1}} |\Delta^j| \to R^j\mathfrak{X}.$$

From the coend formula (5) and the inclusion of the full subcategory $\mathsf{F}_{\leq j-1} \subset \mathsf{F}_{\leq j}$, we also get a natural map $R^{j-1}\mathfrak{X} \to R^j\mathfrak{X}$.

Proposition 3.4 *The following square is a pushout diagram of topological spaces*:

$$\begin{tikzcd}
(L_j \times_{\mathfrak{S}_{j+1}} |\Delta^j|) \cup (\mathfrak{X}[j] \times_{\mathfrak{S}_{j+1}} \partial|\Delta^j|) \arrow[r, "\alpha_{j-1}"] \arrow[d, hook] & R^{j-1}\mathfrak{X} \arrow[d] \\
\mathfrak{X}[j] \times_{\mathfrak{S}_{j+1}} |\Delta^j| \arrow[r, "\beta_j"'] \arrow[ur, phantom, "\ulcorner", very near end] & R^j\mathfrak{X}
\end{tikzcd} \tag{7}$$

Proof We may construct the pushout P as the quotient

$$P := (R^{j-1}\mathfrak{X} \sqcup \mathfrak{X}[j] \times_{\mathfrak{S}_{j+1}} |\Delta^j|)/\sim.$$

One may check that the map β_j together with the natural map $R^{j-1}\mathfrak{X} \to R^j\mathfrak{X}$ gives a map from the disjoint union above which factors through the quotient. Hence we get a well-defined map $P \to R^j\mathfrak{X}$. We now construct a continuous inverse. Recall that $R^j\mathfrak{X}$ is defined in (5) as a quotient of

$$\left(\coprod_{0 \leq n \leq j-1} \mathfrak{X}[n] \times |\Delta^n| \right) \sqcup (\mathfrak{X}[j] \times |\Delta^j|).$$

The natural map $\left(\coprod_{0 \leq n \leq j-1} \mathfrak{X}[n] \times |\Delta^n|\right) \to R^{j-1}\mathfrak{X} \to P$ together with the identity of $\mathfrak{X}[j] \times |\Delta^j|$ gives a map from the disjoint union that factors through the quotient and yields a well-defined map $R^j\mathfrak{X} \to P$. One may finally verify that it is the inverse of the map $P \to R^j\mathfrak{X}$ constructed above. $\square$

We now turn to proving some topological results about our constructions.

Lemma 3.5 *For any integer $n \geq 0$, the subspace $\mathfrak{X}[n] \subset \Gamma \times X^{n+1}$ defined in (4) is closed.*

Proof Let $\mathrm{ev}\colon \Gamma \times X^{n+1} \to (J^r\mathcal{E})^{n+1}$ be the simultaneous evaluation of the jet map j^r (defined in (2)) at $(n+1)$ points of X. We observe directly from the definitions that $\mathfrak{X}[n] = \mathrm{ev}^{-1}(\mathfrak{T}^{n+1})$, and hence is closed as the inverse image of a closed set. $\square$

Lemma 3.6 *For any $n \geq 0$, the map $\rho_n\colon \mathfrak{X}[n] \to \Gamma$ given by $(f, s_0, \dots, s_n) \mapsto f$ is a proper map.*

Proof The projection onto the first factor $\Gamma \times X^{n+1} \to \Gamma$ is proper as X^{n+1} is compact. Hence so is its restriction ρ_n to the closed subspace $\mathfrak{X}[n]$. $\square$

In particular, the map ρ_n is closed, so $\Sigma = \rho_1(\mathfrak{X}[1])$ is closed in Γ. We have natural projections maps $\mathfrak{X}[n] \times |\Delta^n| \to \mathfrak{X}[n] \xrightarrow{\rho_n} \Gamma$ for any $n \geq 0$. They give rise to a map

$$\tau_j : R^j\mathfrak{X} \to \Sigma \tag{8}$$

for every integer $j \geq 0$.

Lemma 3.7 *For any integer $j \geq 0$, the map $\tau_j : R^j\mathfrak{X} \to \Sigma$ is a proper map.*

Proof We have to show that the preimage of any compact set is compact. Equivalently, because Σ is locally compact and Hausdorff, we will show that τ_j is a closed map with compact fibres. From Lemma 3.6, for any n the map ρ_n is closed and hence so is the composition $\mathfrak{X}[n] \times |\Delta^n| \to \mathfrak{X}[n] \xrightarrow{\rho_n} \Gamma$. This implies that τ_j is closed. It remains to see that it has compact fibres. If $f \in \Sigma$, we observe that $\tau_j^{-1}(f) = \beta_j(\rho_j^{-1}(f))$, which is compact as $\rho_j^{-1}(f)$ is, by Lemma 3.6. □

A major advantage of the pushout square (7) is that it allows us to prove the following topological lemma:

Lemma 3.8 *For any integer $j \geq 0$, the space $R^j\mathfrak{X}$ is paracompact and Hausdorff. Furthermore, the natural map $R^{j-1}\mathfrak{X} \to R^j\mathfrak{X}$ is a closed embedding.*

Proof Firstly, from Lemma 3.5, $R^0\mathfrak{X} = \mathfrak{X}[0] \subset \Gamma \times X$ is a closed subset, and hence is itself paracompact Hausdorff. Then the lemma is proven inductively using the pushout diagram (7) together with the fact that

$$((L_j \times_{\mathfrak{S}_{j+1}} |\Delta^j|) \cup (\mathfrak{X}[j] \times_{\mathfrak{S}_{j+1}} \partial|\Delta^j|)) \hookrightarrow \mathfrak{X}[j] \times_{\mathfrak{S}_{j+1}} |\Delta^j|$$

is a closed embedding. □

In the sequel, using the closed embedding of Lemma 3.8 just above, we will simply write $R^{j-1}\mathfrak{X} \subset R^j\mathfrak{X}$. For an integer $j \geq 0$, we let

$$Y_j := \{(f, s_0, \dots, s_j) \in \mathfrak{X}[j] \mid s_l \neq s_k \text{ if } l \neq k\} = \mathfrak{X}[j] - L_j \subset \mathfrak{X}[j] \tag{9}$$

be the subspace of $\mathfrak{X}[j]$ where the singularities are pairwise distinct. For later use, we record the following homeomorphism, which is a direct consequence of the pushout square (7) and the fact that the vertical maps therein are closed embeddings:

$$R^j\mathfrak{X} - R^{j-1}\mathfrak{X} \cong Y_j \times_{\mathfrak{S}_{j+1}} \mathrm{Interior}(|\Delta^j|). \tag{10}$$

Let us now discuss why $\tau_j : R^j\mathfrak{X} \to \Sigma$ needs to be slightly modified to obtain a meaningful "resolution" of Σ. The fibre $\tau_j^{-1}(f)$ above a section $f \in \Sigma$ that has at most $j+1$ singularities is by construction a j-simplex. Hence it is contractible, and one might hope that τ_j induces an isomorphism in cohomology. This is unfortunately not the case. Indeed, $\tau_j^{-1}(f)$ is not contractible if f has at least $j+2$ singularities. To fix this problem, we will modify $R^j(\Sigma)$ by gluing a cone over each fibre $\tau_j^{-1}(f)$ which is not contractible. The precise construction is as follows.

Let $N \geq 0$ be an integer. We let

$$\Sigma_{\geq N+2} := \{f \in \Gamma \mid \#\mathrm{Sing}(f) \geq N+2\} \subset \Sigma \tag{11}$$

denote the subspace of those sections with at least $N+2$ singularities. We denote by $\overline{\Sigma_{\geq N+2}}$ its closure in Σ (or equivalently, in Γ). Observe that the surjectivity of the map τ_N implies the following equality:

$$\tau_N(\tau_N^{-1}(\overline{\Sigma_{\geq N+2}})) = \overline{\Sigma_{\geq N+2}}.$$

We glue fibrewise a cone over each $f \in \overline{\Sigma_{\geq N+2}}$ by defining the space $R^N_{\mathrm{cone}}(\Sigma)$ as the following homotopy pushout:

$$\begin{tikzcd} \tau_N^{-1}(\overline{\Sigma_{\geq N+2}}) \arrow[r, hook] \arrow[d, "\tau_N"'] \arrow[dr, phantom, "\mathrm{ho}\ulcorner"] & R^N\mathfrak{X} \arrow[d] \\ \overline{\Sigma_{\geq N+2}} \arrow[r] & R^N_{\mathrm{cone}}\mathfrak{X}. \end{tikzcd} \tag{12}$$

All three defining spaces in the corners of (12) map to Σ, and hence we obtain a surjective projection map

$$\pi\colon R^N_{\mathrm{cone}}\mathfrak{X} \to \Sigma. \tag{13}$$

We want to prove that π induces an isomorphism in cohomology with compact supports. We begin with a couple of lemmas.

Lemma 3.9 *The map $\pi\colon R^N_{\mathrm{cone}}\mathfrak{X} \to \Sigma$ is proper.*

Proof We will prove that is it closed with compact fibres, which implies the properness. By definition of the homotopy pushout, $R^N_{\mathrm{cone}}\mathfrak{X}$ is a quotient of the following disjoint union:

$$R^N\mathfrak{X} \sqcup \tau_N^{-1}(\overline{\Sigma_{\geq N+2}}) \times [0,1] \sqcup \overline{\Sigma_{\geq N+2}}.$$

The map π is induced by the following three maps: the projection $\tau_N\colon R^N\mathfrak{X} \to \Sigma$, the projection $\tau_N^{-1}(\overline{\Sigma_{\geq N+2}}) \times [0,1] \to \tau_N^{-1}(\overline{\Sigma_{\geq N+2}}) \to \Sigma$ and the inclusion $\overline{\Sigma_{\geq N+2}} \hookrightarrow \Sigma$. The first two are closed by Lemma 3.7 and the last one is the inclusion of a closed subset, and hence closed.

Finally, we prove that the fibres of π are compact. We saw in the proof of Lemma 3.7 that for any $f \in \Sigma$, the fibre $\tau_N^{-1}(f)$ was compact. Now, $\pi^{-1}(f)$ is either $\tau_N^{-1}(f)$ if $f \in \Sigma - \overline{\Sigma_{\geq N+2}}$ or a cone over it if $f \in \overline{\Sigma_{\geq N+2}}$. In any case it is compact. □

Lemma 3.10 *The space $R^N_{\mathrm{cone}}\mathfrak{X}$ is paracompact, locally compact and Hausdorff.*

Proof The paracompactness and Hausdorffness follow from the definition as a homotopy pushout and Lemma 3.8. It is locally compact as it maps properly to the locally compact space Σ. □

The most important corollary is the following:

Proposition 3.11 *The map $\pi\colon R^N_{\mathrm{cone}}\mathfrak{X} \to \Sigma$ induces an isomorphism in cohomology with compact supports.*

Proof The properness of π proved in Lemma 3.9 implies that it induces a well-defined map in cohomology with compact supports. We also observed in the proof of that lemma that a fibre of π is either a simplex or a cone, and hence contractible. The proposition then follows from the Vietoris–Begle theorem [8, V.6.1]. □

3.2 Construction of the spectral sequence

Let $N \geq 1$ be an integer. Recall from Lemma 3.8 that we have closed embeddings $R^{j-1}\mathfrak{X} \subset R^j\mathfrak{X}$. We define the following filtration on $R^N_{\mathrm{cone}}\mathfrak{X}$:

$$F_0 = R^0\mathfrak{X} \subset F_1 = R^1\mathfrak{X} \subset \cdots \subset F_N = R^N\mathfrak{X} \subset F_{N+1} = R^N_{\mathrm{cone}}\mathfrak{X}.$$

Following standard arguments, we obtain from the filtration a spectral sequence

$$E_1^{p,q} = H_c^{p+q}(F_p - F_{p-1}) \Rightarrow H_c^{p+q}(R^N_{\mathrm{cone}}\mathfrak{X}).$$

Using Proposition 3.11 and Alexander duality, we obtain

$$H_c^{p+q}(R^N_{\mathrm{cone}}\mathfrak{X}) \cong H_c^{p+q}(\Sigma) \cong \widetilde{H}_{2\dim_{\mathbb{C}}\Gamma-(p+q)-1}(\Gamma - \Sigma),$$

where $\widetilde{H}$ denotes reduced singular homology. Letting $s = -p-1$ and $t = 2\dim_{\mathbb{C}}\Gamma - q$, we regrade our spectral sequence and obtain the following:

Proposition 3.12 *There is a spectral sequence on the second quadrant $s \leq -1$ and $t \geq 0$:*

$$E^1_{s,t} = H_c^{2\dim_{\mathbb{C}}\Gamma-1-s-t}(F_{-s-1} - F_{-s-2}; \mathbb{Z}) \Rightarrow \widetilde{H}_{s+t}(\Gamma - \Sigma; \mathbb{Z}).$$

The differential d^r on the r^{th} page of the spectral sequence has bidegree $(-r, r-1)$, ie it is a morphism $d^r_{s,t}\colon E^r_{s,t} \to E^r_{s-r,t+r-1}$.

4 Cohomology groups on the E^1-page

As in the last section, we choose a holomorphic vector bundle $\mathcal{E}$ on X and an admissible Taylor condition $\mathfrak{T} \subset J^r\mathcal{E}$ inside the r^{th} jet bundle of $\mathcal{E}$. For the remainder of this section, we also let

$$N = N(\mathcal{E}, r)$$

be the largest integer $N \geq 0$ such that $\mathcal{E}$ is $((N+1)(r+1)-1)$-jet ample as in Definition 2.3. As discussed in the introduction, we assume that such an N exists. If not, the statements in this section are either trivially false, or trivially true as they describe elements of the empty set. For brevity, we still use the notation

$$\Gamma = \Gamma_{\mathrm{hol}}(\mathcal{E}) \quad \text{and} \quad \Sigma = \Gamma_{\mathrm{hol}}(\mathcal{E}) - \Gamma_{\mathrm{hol,ns}}(\mathcal{E}),$$

as well as $\mathfrak{X}$ for the associated functor $\mathsf{F}^{\mathrm{op}} \to \mathsf{Top}$ as in (4).

We will study the first page of the spectral sequence from Proposition 3.12 converging to the cohomology of $R^N_{\mathrm{cone}}\mathfrak{X}$. For convenience, we summarise the results of this section as follows:

Proposition 4.1 *Let $\mathcal{E}$ be a holomorphic vector bundle on X and $\mathfrak{T} \subset J^r\mathcal{E}$ be an admissible Taylor condition. Let $N = N(\mathcal{E}, r)$. The resolution and its filtration described in Section 3 give rise to a spectral sequence on the second quadrant $s \leq -1$ and $t \geq 0$ converging to the homology of the space of nonsingular sections $\Gamma_{\mathrm{hol,ns}}(\mathcal{E})$:*

$$E^1_{s,t} = H_c^{2\dim_{\mathbb{C}}\Gamma-1-s-t}(F_{-s-1} - F_{-s-2}; \mathbb{Z}) \Rightarrow \widetilde{H}_{s+t}(\Gamma_{\mathrm{hol,ns}}(\mathcal{E}); \mathbb{Z}).$$

The differentials on the r^{th} page have bidegree $(-r, r-1)$. Furthermore:

(i) ***Proposition 4.4*** *For $-N-1 \leq s \leq -1$, we have the isomorphisms*

$$E^1_{s,t} \cong H_c^{-t-2s\,\mathrm{rk}_{\mathbb{C}}\,J^r\mathcal{E}}(\mathfrak{T}^{(-s)}; \mathbb{Z}^{\mathrm{sign}})$$

for all $t \geq 0$, with $\mathfrak{T}^{(-s)}$ defined in (14).

(ii) ***Proposition 4.10*** *For $t < (N+1)e(\mathfrak{T})$,*

$$E^1_{-N-2,t} = 0.$$

As a visual aid, we have drawn the spectral sequence in Figure 1, where we have chosen to fix $e(\mathfrak{T}) = 2$ to lighten the notation. We briefly describe the various zones. Firstly, the only possibly nonvanishing groups lie in the coloured squares. All groups $E^r_{s,t}$ with $s \leq -N-3$ are zero as the filtration finishes after $N+1$ steps. According to Proposition 4.10, the groups below the horizontal solid line in the column $s = -N-2$ vanish. The differentials coming from the groups below the upper staircase never hit groups in the column where $s = -N-2$ and $t \geq 2N+2$. Finally, the lower staircase delimits the zone of total degree $* \leq N-1$. We have also drawn some differentials d^r to the group $E^r_{-N-2,2N+2}$ for $r = 1, 2, 3$ and $N+1$.

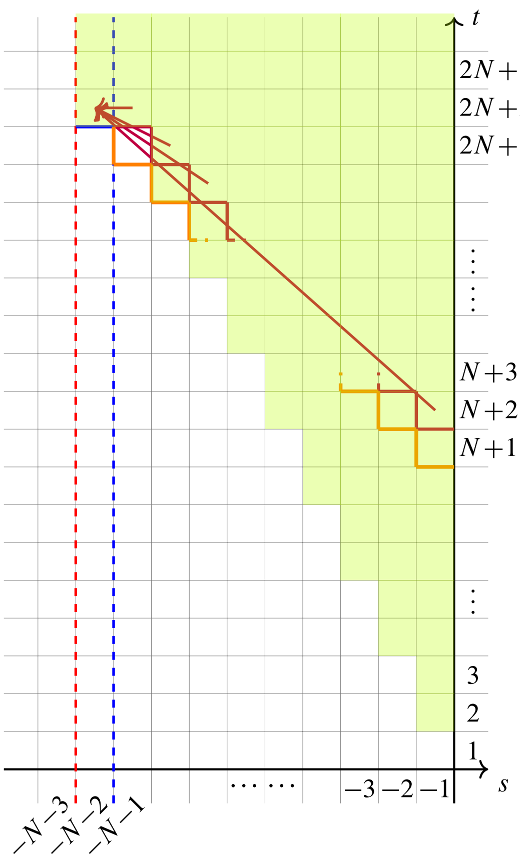


Figure 1: The first page of the spectral sequence when $e(\mathfrak{T}) = 2$.

4.1 The first steps of the filtration

For an integer $j \geq 0$, recall from (9) the space

$$Y_j = \{(f, s_0, \dots, s_j) \in \mathfrak{X}[j] \mid s_l \neq s_k \text{ if } l \neq k\} \subset \mathfrak{X}[j].$$

Lemma 4.2 *For $0 \leq j \leq N(\mathcal{E}, r)$, there is a fibre bundle*

$$\mathrm{Interior}(|\Delta^j|) \to F_j - F_{j-1} \to Y_j/\mathfrak{S}_{j+1}.$$

Proof Recall from the definition of the filtration on $R^N_{\mathrm{cone}}\mathfrak{X}$ that $F_j = R^j\mathfrak{X}$ for $0 \leq j \leq N$. As a consequence of the pushout square (7), we observed in (10) that we have the following homeomorphism:

$$R^j\mathfrak{X} - R^{j-1}\mathfrak{X} \cong Y_j \times_{\mathfrak{S}_{j+1}} \mathrm{Interior}(|\Delta^j|).$$

Projecting down to the first factor gives the required fibre bundle. □

By an *affine bundle* we mean a torsor for a vector bundle. In the sequel, they will arise naturally from fibrewise surjective linear maps between vector bundles. For any integer $j \geq 1$, the bundle $(J^r\mathcal{E})^j$ projects down to X^j and we may consider its restriction to the open subset $\mathrm{Conf}_j(X) \subset X^j$ of those tuples of points which are pairwise distinct. The symmetric group $\mathfrak{S}_j$ acts on these spaces by permuting the coordinates. In particular, it acts on the subspace $\mathfrak{T}^j \subset (J^r\mathcal{E})^j$ and we let

$$\mathfrak{T}^{(j)} := (\mathfrak{T}^j|_{\mathrm{Conf}_j(X)})/\mathfrak{S}_j \tag{14}$$

be the orbit space of the restriction of $\mathfrak{T}^j$ over the subspace $\mathrm{Conf}_j(X) \subset X^j$.

Lemma 4.3 *Let $0 \leq j \leq N(\mathcal{E}, r)$ be an integer and recall from (9) the space Y_j of those tuples $(f, s_0, \dots, s_j) \in \Gamma \times \mathrm{Conf}_{j+1}(X)$ where f is singular at the s_i. We may simultaneously evaluate the jet map at these points*:

$$Y_j \to \mathfrak{T}^{j+1}|_{\mathrm{Conf}_{j+1}(X)}, \quad (f, s_0, \dots, s_j) \mapsto (j^r(f)(s_0), \dots, j^r(f)(s_j)).$$

Taking $\mathfrak{S}_{j+1}$-orbits on the domain and codomain of this map yields an affine bundle

$$Y_j/\mathfrak{S}_{j+1} \to \mathfrak{T}^{(j+1)}$$

whose fibre has complex dimension $\dim_{\mathbb{C}} \Gamma - (j+1)\,\mathrm{rk}_{\mathbb{C}}\, J^r\mathcal{E}$. (Here $\mathrm{rk}_{\mathbb{C}}\, J^r\mathcal{E}$ denotes the complex rank of the vector bundle $J^r\mathcal{E}$.)

Proof The simultaneous evaluation of the jet map gives a map

$$\begin{tikzcd} \Gamma \times \mathrm{Conf}_{j+1}(X) \arrow[rr] \arrow[dr] & & (J^r\mathcal{E})^{j+1}|_{\mathrm{Conf}_{j+1}(X)} \arrow[dl] \\ & \mathrm{Conf}_{j+1}(X) & \end{tikzcd} \tag{15}$$

of vector bundles over the configuration space $\mathrm{Conf}_{j+1}(X)$. Under the assumption $0 \leq j \leq N(\mathcal{E}, r)$, Lemma 2.4 shows that this map of bundles is fibrewise surjective. Therefore the top map of (15) is an affine bundle. Subtracting the ranks, we obtain that its fibre has complex dimension $\dim_{\mathbb{C}} \Gamma - (j+1)\,\mathrm{rk}_{\mathbb{C}}\, J^r\mathcal{E}$.

Now, the pullback of the affine bundle (15) to the subspace $\mathfrak{T}^{j+1}|_{\mathrm{Conf}_{j+1}(X)}$ is an affine bundle with total space Y_j. Finally, taking $\mathfrak{S}_{j+1}$-orbits yields the affine bundle

$$Y_j/\mathfrak{S}_{j+1} \to (\mathfrak{T}^{j+1}|_{\mathrm{Conf}_{j+1}(X)})/\mathfrak{S}_{j+1} = \mathfrak{T}^{(j+1)},$$

which still has the rank that we have computed above. □

The quotient maps $Y_j \to Y_j/\mathfrak{S}_{j+1}$ and $\mathfrak{T}^{j+1}|_{\mathrm{Conf}_{j+1}(X)} \to \mathfrak{T}^{(j+1)}$ are principal $\mathfrak{S}_{j+1}$-bundles and hence are classified by (homotopy classes of) maps to the classifying space $B\mathfrak{S}_{j+1}$. Composing with the sign representation $B\mathfrak{S}_{j+1} \xrightarrow{B\,\mathrm{sign}} B\mathbb{Z}/2$, we obtain two well-defined homotopy classes of maps:

$$Y_j/\mathfrak{S}_{j+1} \to B\mathbb{Z}/2 \quad \text{and} \quad \mathfrak{T}^{(j+1)} \to B\mathbb{Z}/2.$$

We will write $\mathbb{Z}^{\mathrm{sign}}$ for the corresponding local coefficient systems.

Proposition 4.4 *Let $-N(\mathcal{E},r)-1 \le s \le -1$. Then we have the isomorphism*

$$E^1_{s,t} \cong H_c^{-t-2s\,\mathrm{rk}_{\mathbb{C}}\,J^r\mathcal{E}}(\mathfrak{T}^{(-s)};\mathbb{Z}^{\mathrm{sign}}),$$

where $\mathfrak{T}^{(-s)}$ is the space defined in (14) and $\mathbb{Z}^{\mathrm{sign}}$ is the local coefficient system described above.

Proof Recall from Proposition 3.12 that the first page of the spectral sequence is given by

$$E^1_{s,t} = H_c^{2\dim_{\mathbb{C}}\Gamma-1-s-t}(R^{-s-1}\mathfrak{X} - R^{-s-2}\mathfrak{X};\mathbb{Z}).$$

Via a homeomorphism $\mathrm{Interior}(|\Delta^j|) \cong \mathbb{R}^j$, we see that the fibre bundle of Lemma 4.2 is homeomorphic to a vector bundle. Applying the Thom isomorphism to the latter, we obtain

$$E^1_{s,t} \cong H_c^{2\dim_{\mathbb{C}}\Gamma-t}(Y_{-s-1}/\mathfrak{S}_{-s};\mathbb{Z}^{\mathrm{sign}}).$$

Another application of the Thom isomorphism using Lemma 4.3 yields

$$E^1_{s,t} \cong H_c^{-t-2s\,\mathrm{rk}_{\mathbb{C}}\,J^r\mathcal{E}}(\mathfrak{T}^{(-s)};\mathbb{Z}^{\mathrm{sign}}).$$ □

4.2 The last step of the filtration

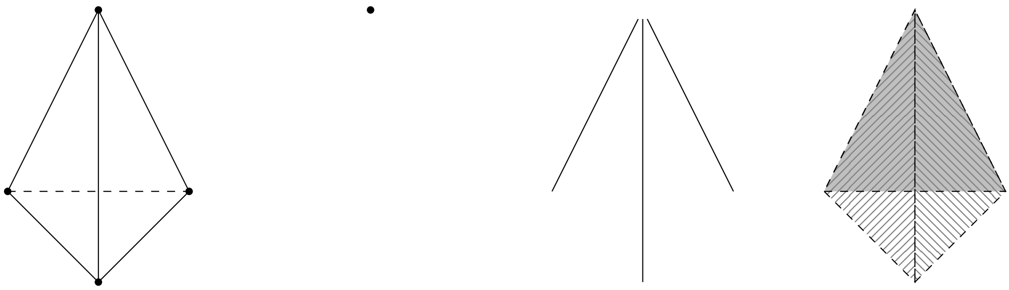
We study the last nontrivial part of the E^1-page, that is, the column $s = -N(\mathcal{E},r)-2$, where

$$E^1_{-N-2,t} = H_c^{2\dim_{\mathbb{C}}\Gamma+1+N-t}(R^N_{\mathrm{cone}}\mathfrak{X} - R^N\mathfrak{X};\mathbb{Z}).$$

The methods from the last section do not apply to the space $R^N_{\mathrm{cone}}\mathfrak{X} - R^N\mathfrak{X}$ and we will not be able to express the cohomology groups $E^1_{-N-2,t}$ in terms of other "known" groups. However, using the technical assumptions made in Definition 2.8 about the Taylor condition $\mathfrak{T}$, we will obtain a vanishing result for $E^1_{-N-2,t}$. This will be enough for the proof of our main theorem.

Recall the projection map $\tau_N : R^N\mathfrak{X} \to \Sigma$ from (8). From the homotopy pushout square (12), we obtain the homeomorphism

$$R^N_{\mathrm{cone}}\mathfrak{X} - R^N\mathfrak{X} \cong ((\tau_N^{-1}(\overline{\Sigma_{\ge N+2}}) \times]0,1]) \sqcup \overline{\Sigma_{\ge N+2}})/\sim,$$



$\pi^{-1}(f)$ $\pi^{-1}(f) \cap \mathrm{Str}_{-1}$ $\pi^{-1}(f) \cap \mathrm{Str}_0$ $\pi^{-1}(f) \cap \mathrm{Str}_1$ (the three sides only)

Figure 2: Decomposition of the open cone.

where $(z,1) \in \tau_N^{-1}(\overline{\Sigma_{\geq N+2}}) \times]0,1]$ is identified with $\tau_N(z) \in \overline{\Sigma_{\geq N+2}}$ in the quotient. Indeed, there is a natural continuous bijection from the right-hand side to the left-hand side. It is in fact a homeomorphism, as the top arrow in the homotopy pushout square (12) is the inclusion of a closed subset. In other words, this is the fibrewise (for the map τ_N) open cone over $\overline{\Sigma_{\geq N+2}}$. We stratify this space by the following locally closed subspaces (this is analogous to [27, Lemma 18]):

$$\begin{aligned} \mathrm{Str}_{-1} &:= \overline{\Sigma_{\geq N+2}}, \\ \mathrm{Str}_0 &:= (\tau_N^{-1}(\overline{\Sigma_{\geq N+2}}) \times]0,1[) \cap (R^0\mathfrak{X} \times]0,1[), \\ \mathrm{Str}_j &:= (\tau_N^{-1}(\overline{\Sigma_{\geq N+2}}) \times]0,1[) \cap ((R^j\mathfrak{X} - R^{j-1}\mathfrak{X}) \times]0,1[) \quad \text{for } 1 \leq j \leq N. \end{aligned}$$

For $0 \leq j \leq N$, let

$$Y_j^{\geq N+2} := \{(f, s_0, \dots, s_j) \in \Gamma \times \mathrm{Conf}_{j+1}(X) \mid f \in \overline{\Sigma_{\geq N+2}} \text{ and } s_i \in \mathrm{Sing}(f)\} \subset Y_j. \tag{16}$$

Using the homeomorphism (10) identifying the difference between two consecutive steps of the resolution, we have a homeomorphism

$$\mathrm{Str}_j \cong (Y_j^{\geq N+2} \times_{\mathfrak{S}_{j+1}} |\mathring{\Delta}^j|) \times]0,1[\tag{17}$$

for $0 \leq j \leq N$, where $|\mathring{\Delta}^j|$ denotes the interior of the simplex.

It is easier to think about this stratification by looking at one fibre $\pi^{-1}(f)$ at a time. Then we are just decomposing an open cone over a union of simplices into the following pieces: the apex (corresponding to $\mathrm{Str}_{-1} \cap \pi^{-1}(f)$), the open segments from the 0-simplices to the apex (corresponding to $\mathrm{Str}_0 \cap \pi^{-1}(f)$), the open (filled) triangles between the 1-simplices and the apex, etc. Figure 2 shows the strata in a single fibre $\pi^{-1}(f)$ when f has three singular points and $N=1$. In this case, $\tau_N^{-1}(f)$ consists of three 1-simplices glued together (ie a triangle), so $\pi^{-1}(f)$ is the cone over that triangle.

If we find an integer $D \geq 0$ such that $H_c^k(\mathrm{Str}_j) = 0$ for all $-1 \leq j \leq N$ and all $k > D$, then the same result will hold for the union, ie $H_c^k(R^N_{\mathrm{cone}}\mathfrak{X} - R^N\mathfrak{X}) = 0$ for $k > D$. In what follows, we set out to find such a D as small as we can. With that in mind, we make the following ad hoc definition of cohomological dimension:

Definition 4.5 A space Z has *cohomological dimension* D with respect to a local coefficient system $\mathcal{A}$ if D is the smallest integer such that $H_c^k(Z;\mathcal{A}) = 0$ for all $k > D$. We will denote it by $\operatorname{cohodim}(Z,\mathcal{A})$, or simply $\operatorname{cohodim}(Z)$ if $\mathcal{A} = \mathbb{Z}$.

The only nontrivial local coefficient system we will need is $\mathbb{Z}^{\mathrm{sign}}$, which is induced on the quotient $Y_j^{\geq N+2}/\mathfrak{S}_{j+1}$ by the sign representation $\mathfrak{S}_{j+1} \to \mathbb{Z}/2$.

Lemma 4.6 *For* $0 \leq j \leq N$, *we have*

$$\operatorname{cohodim}(\mathrm{Str}_j) = 1 + j + \operatorname{cohodim}(Y_j^{\geq N+2}/\mathfrak{S}_{j+1}, \mathbb{Z}^{\mathrm{sign}}).$$

Proof From the homeomorphism (17), we have a trivial fibre bundle

$$]0,1[\to \mathrm{Str}_j \to Y_j^{\geq N+2} \times_{\mathfrak{S}_{j+1}} |\mathring{\Delta}^j|.$$

This implies that $\operatorname{cohodim}(\mathrm{Str}_j) = 1 + \operatorname{cohodim}(Y_j^{\geq N+2} \times_{\mathfrak{S}_{j+1}} |\mathring{\Delta}^j|)$. Now, we have another fibre bundle:

$$|\mathring{\Delta}^j| \to Y_j^{\geq N+2} \times_{\mathfrak{S}_{j+1}} |\mathring{\Delta}^j| \to Y_j^{\geq N+2}/\mathfrak{S}_{j+1}.$$

Hence, by the Thom isomorphism, we obtain

$$\operatorname{cohodim}(Y_j^{\geq N+2} \times_{\mathfrak{S}_{j+1}} |\mathring{\Delta}^j|) = j + \operatorname{cohodim}(Y_j^{\geq N+2}/\mathfrak{S}_{j+1}, \mathbb{Z}^{\mathrm{sign}}). \qquad \square$$

We thus have reduced our problem to studying the cohomology of $Y_j^{\geq N+2}/\mathfrak{S}_{j+1}$ for $0 \leq j \leq N$, as well as that of $\overline{\Sigma_{\geq N+2}}$. We shall do so by comparing these spaces to a known one, namely the space

$$Y_N = \{(f, s_0, \dots, s_N) \in \Gamma \times \mathrm{Conf}_{N+1}(X) \mid s_i \in \mathrm{Sing}(f)\}.$$

We first introduce some notation. Using charts on X, we may cover Y_N by finitely many semialgebraic sets, whose intersections are also semialgebraic. Recall, eg from [7, Theorem 2.3.6], that every semialgebraic set is the disjoint union of cells, each homeomorphic to an open disc $]0,1[^d$ for some $d \geq 0$. The largest d in such a decomposition is called the dimension of the semialgebraic set. Let $\dim Y_N$ be the largest of the dimensions of the semialgebraic sets in a cover of Y_N. (It depends a priori on the chosen cover, but we suppress this from the notation.) The following is a crucial result for controlling our spectral sequence:

Lemma 4.7 *For* $0 \leq j \leq N$, *we have*

$$\dim Y_N \geq \operatorname{cohodim}(Y_j^{\geq N+2}/\mathfrak{S}_{j+1}, \mathbb{Z}^{\mathrm{sign}}).$$

Proof Forgetting the last singularity yields a map

$$Y_{N+1} \to Y_N, \quad (f, s_0, \dots, s_{N+1}) \mapsto (f, s_0, \dots, s_N),$$

and we will write $Y_N^{\geq N+2} \subset Y_N$ for its image. As the projection map is semialgebraic (when read in charts), its image is semialgebraic (in charts) and $\dim Y_N^{\geq N+2} \leq \dim Y_N$. Let $0 \leq j \leq N$. Only remembering the $(j+1)^{\mathrm{st}}$ singularities gives a map

$$Y_N^{\geq N+2} \to Y_j^{\geq N+2}, \quad (f, s_0, \dots, s_N) \mapsto (f, s_0, \dots, s_j). \tag{18}$$

Notice that this map is not surjective, as it may happen that a section $f \in \overline{\Sigma_{\geq N+2}}$ has fewer than $N+1$ singularities. We study the map (18) locally via charts. Let $U_0, \dots, U_N \subset X$ be charts on X as in Definition 2.6. Then the subsets

$$U := \{(f, s_0, \dots, s_j) \in \overline{\Sigma_{\geq N+2}} \times U_0 \times \cdots \times U_j \mid s_k \in \mathrm{Sing}(f),\, s_i \neq s_j \;\forall i \neq j\} \subset Y_j^{\geq N+2}$$

and

$$V := \{(f, s_0, \dots, s_N) \in \Gamma \times U_0 \times \cdots \times U_N \mid s_k \in \mathrm{Sing}(f),\, s_i \neq s_j \;\forall i \neq j\} \cap Y_N^{\geq N+2} \subset Y_N^{\geq N+2}$$

are semialgebraic. Indeed, they are the preimages of the semialgebraic sets $\mathfrak{T}^{j+1}$ and $\mathfrak{T}^{N+1}$, respectively, via the simultaneous evaluation of the jet map, which is algebraic and hence is semialgebraic; see [7, Proposition 2.2.7]. The restriction of the map (18) to U and V is an algebraic map, and hence a semialgebraic map, $\phi \colon V \to U$ between semialgebraic sets. Using [7, Theorem 2.8.8] we obtain the following inequality on the dimensions (as defined above using cell decompositions):

$$\dim(V) \geq \dim(\phi(V)).$$

Furthermore, the definition of $Y_j^{\geq N+2}$ implies that the semialgebraic map $\phi \colon V \to U$ has dense image, ie $\overline{\phi(V)} = U$. Using that the closure has the same dimension [7, Proposition 2.8.2] and the inequality above, we obtain

$$\dim(V) \geq \dim(U).$$

Varying the charts $U_0, \dots, U_N \subset X$, we may cover the domain and codomain of (18) by subsets defined like U and V. If U' and V' are two other such subsets, then $U \cap U'$ and $V \cap V'$ are also semialgebraic sets because they are intersections of semialgebraic sets. (This follows from Definition 2.5.) Hence the argument shows that the inequality on the dimensions also holds on intersections. Let $\dim Y_j^{\geq N+2}$ denote the maximum of the dimensions in a cover of $Y_j^{\geq N+2}$ by semialgebraic sets. Then an argument using the Mayer–Vietoris spectral sequence shows that the cohomological dimension of $Y_j^{\geq N+2}$ is less than its dimension $\dim Y_j^{\geq N+2}$. Therefore

$$\dim Y_N \geq \dim Y_N^{\geq N+2} \geq \dim Y_j^{\geq N+2} \geq \mathrm{cohodim}(Y_j^{\geq N+2}). \tag{19}$$

Finally, from the principal $\mathfrak{S}_{j+1}$-bundle $Y_j^{\geq N+2} \to Y_j^{\geq N+2}/\mathfrak{S}_{j+1}$, we see that the dimension of the orbit space is the same as that of $Y_j^{\geq N+2}$. Therefore the inequality (19) holds when replacing the rightmost term with $\mathrm{cohodim}(Y_j^{\geq N+2}/\mathfrak{S}_{j+1}, \mathbb{Z}^{\mathrm{sign}})$. □

Repeating the proof with the map $Y_N^{\geq N+2} \to \overline{\Sigma_{\geq N+2}}$, $(f, s_0, \dots, s_N) \mapsto f$ yields:

Lemma 4.8 *The following inequality holds*:

$$\dim Y_N \geq \mathrm{cohodim}(\overline{\Sigma_{\geq N+2}}, \mathbb{Z}).$$ □

The final computation to be made is the content of the following lemma. It uses the notation $e(\mathfrak{T})$ of excess codimension established in Definition 2.10.

Lemma 4.9 *The dimension of Y_N satisfies*

$$\dim Y_N \leq 2\dim_{\mathbb{C}} \Gamma - (N+1)e(\mathfrak{T}).$$

Proof The proof of Lemma 4.3 shows that the simultaneous evaluation of the jet map

$$Y_N \to \mathfrak{T}^{N+1}|_{\mathrm{Conf}_{N+1}(X)}, \quad (f, s_0, \dots, s_N) \mapsto (j^r(f)(s_0), \dots, j^r(f)(s_N))$$

is an affine bundle whose fibre has complex dimension $\dim_{\mathbb{C}} \Gamma - (N+1)\operatorname{rk}_{\mathbb{C}} J^r\mathcal{E}$. Therefore, on dimensions,

$$\dim Y_N \leq \dim(\mathfrak{T}^{N+1}|_{\mathrm{Conf}_{N+1}(X)}) + 2\dim_{\mathbb{C}} \Gamma - 2(N+1)\operatorname{rk}_{\mathbb{C}} J^r\mathcal{E}.$$

Now, because $\mathfrak{T}$ is a semialgebraic subset of $J^r\mathcal{E}$ of dimension less than $2\operatorname{rk}_{\mathbb{C}} J^r\mathcal{E} - e(\mathfrak{T})$,

$$\dim(\mathfrak{T}^{N+1}|_{\mathrm{Conf}_{N+1}(X)}) \leq (N+1)(2\operatorname{rk}_{\mathbb{C}} J^r\mathcal{E} - e(\mathfrak{T})).$$

The lemma is then proven by combining these two inequalities. □

Assembling all the estimations we have obtained so far, we can state and prove the following:

Proposition 4.10 *The cohomology groups in the column $s = -N(\mathcal{E}, r) - 2$ on the first page of the spectral sequence*

$$E^1_{-N-2,t} = H_c^{2\dim_{\mathbb{C}} \Gamma + 1 + N - t}(R^N_{\mathrm{cone}}\mathfrak{X} - R^N\mathfrak{X}; \mathbb{Z})$$

vanish for $t < (N+1)e(\mathfrak{T})$.

Proof A direct inspection of the spectral sequence associated to the stratification Str_j on $R^N_{\mathrm{cone}}\mathfrak{X} - R^N\mathfrak{X}$ shows that

$$\operatorname{cohodim}(R^N_{\mathrm{cone}}\mathfrak{X} - R^N\mathfrak{X}) \leq \max_j \operatorname{cohodim}(\mathrm{Str}_j).$$

For $0 \leq j \leq N$, combining Lemmas 4.6, 4.7 and 4.9, we get

$$\operatorname{cohodim}(\mathrm{Str}_j) \leq 1 + j + 2\dim_{\mathbb{C}} \Gamma - (N+1)e(\mathfrak{T}) \leq 2\dim_{\mathbb{C}} \Gamma - N(e(\mathfrak{T}) - 1) - (e(\mathfrak{T}) - 1).$$

Similarly, using Lemmas 4.8 and 4.9, we obtain

$$\operatorname{cohodim}(\mathrm{Str}_{-1}) \leq 2\dim_{\mathbb{C}} \Gamma - (N+1)e(\mathfrak{T}).$$

Therefore $\operatorname{cohodim}(R^N_{\mathrm{cone}}\mathfrak{X} - R^N\mathfrak{X}) \leq 2\dim_{\mathbb{C}} \Gamma - N(e(\mathfrak{T}) - 1) - (e(\mathfrak{T}) - 1)$ and the result follows. □

5 Interpolating holomorphic and continuous sections

In this section, we introduce and study section spaces that lie in between holomorphic and continuous sections of the jet bundle $J^r\mathcal{E}$. They will be written as combinations of holomorphic and "antiholomorphic" sections. We first explain how to take the complex conjugate of a holomorphic section. We then construct these spaces and finish by explaining how the resolution and the spectral sequence from the previous sections can be adapted to them.

5.1 Complex conjugation of sections

Using the fact that X is projective, we choose once and for all a very ample holomorphic line bundle $\mathcal{L}$ on it as well as a basis $z_0, \dots, z_M$ of the complex vector space of holomorphic global sections $\Gamma_{\mathrm{hol}}(\mathcal{L})$.

We denote by $\bar{\mathcal{L}}$ the complex conjugate line bundle of $\mathcal{L}$. It is obtained from the underlying real vector bundle of $\mathcal{L}$ by having the complex numbers act by multiplication by their complex conjugates. We regard it as a smooth complex line bundle. We now define a complex conjugation operation $\mathcal{L} \to \bar{\mathcal{L}}$. Recall that the line bundle $\mathcal{L}$ may be constructed as a quotient

$$\mathcal{L} := \Bigl(\bigsqcup_i U_i \times \mathbb{C}\Bigr) / (x, v_i) \sim (x, t_{ji}(v_i))$$

from the data $(\{U_i\}_i, (t_{ij})_{i,j})$ of trivialising open sets $U_i \subset X$ and transition functions

$$t_{ij} : U_i \cap U_j \to \mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^*$$

satisfying a cocycle condition. Similarly, $\bar{\mathcal{L}}$ may be constructed via such a quotient by replacing the transition functions by their complex conjugates $\bar{t}_{ij}$. The formula

$$\bigsqcup_i U_i \times \mathbb{C} \to \bigsqcup_i U_i \times \mathbb{C}, \quad (x, v) \mapsto (x, \bar{v}),$$

then gives a well-defined $\mathbb{R}$-linear isomorphism $\mathcal{L} \to \bar{\mathcal{L}}$. On continuous global sections, we thus obtain an $\mathbb{R}$-linear *complex conjugation operation*:

$$\bar{\cdot} : \Gamma_{\mathcal{C}^0}(\mathcal{L}) \to \Gamma_{\mathcal{C}^0}(\bar{\mathcal{L}}). \tag{20}$$

For a complex vector space V, we denote by $\bar{V}$ the $\mathbb{C}$-vector space whose underlying set is V with the $\mathbb{C}$-module structure given by multiplication by the complex conjugate. We get a $\mathbb{C}$-linear map

$$\overline{\Gamma_{\mathrm{hol}}(\mathcal{L})} \hookrightarrow \overline{\Gamma_{\mathcal{C}^0}(\mathcal{L})} \xrightarrow{(20)} \Gamma_{\mathcal{C}^0}(\bar{\mathcal{L}}). \tag{21}$$

We let

$$\eta := \sum_{j=0}^{M} z_j \otimes \bar{z}_j \in \Gamma_{\mathrm{hol}}(\mathcal{L}) \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L})}. \tag{22}$$

We note that although η depends on a choice of basis of $\Gamma_{\mathrm{hol}}(\mathcal{L})$, our results will be independent of this choice. Its image via the composition of the map (21) and the multiplication map $\Gamma_{\mathcal{C}^0}(\mathcal{L}) \otimes_{\mathbb{C}} \Gamma_{\mathcal{C}^0}(\bar{\mathcal{L}}) \to \Gamma_{\mathcal{C}^0}(\mathcal{L} \otimes \bar{\mathcal{L}})$ is a never vanishing section. It therefore gives an explicit trivialisation of the smooth complex line bundle $\mathcal{L} \otimes \bar{\mathcal{L}} \cong X \times \mathbb{C}$. In particular, we obtain an isomorphism on the level of continuous sections:

$$\Gamma_{\mathcal{C}^0}(\mathcal{L} \otimes \bar{\mathcal{L}}) \cong \Gamma_{\mathcal{C}^0}(X \times \mathbb{C}) = \mathcal{C}^0(X, \mathbb{C}). \tag{23}$$

5.2 Stabilisation

For every integer $k \geq 0$, we now construct the following commutative diagram:

$$\begin{tikzcd}
\Gamma_{\mathrm{hol}}((J^r\mathcal{E}) \otimes \mathcal{L}^k) \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)} \arrow[d, "\gamma_k"'] \arrow[dr, "\varphi_k"] & \\
& \Gamma_{\mathcal{C}^0}(J^r\mathcal{E}) \\
\Gamma_{\mathrm{hol}}((J^r\mathcal{E}) \otimes \mathcal{L}^{k+1}) \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^{k+1})} \arrow[ur, "\varphi_{k+1}"'] &
\end{tikzcd} \tag{24}$$

The horizontal maps are given by the composition

$$\begin{aligned}
\varphi_k \colon \Gamma_{\mathrm{hol}}((J^r\mathcal{E}) \otimes \mathcal{L}^k) \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)} &\to \Gamma_{\mathcal{C}^0}(J^r\mathcal{E} \otimes \mathcal{L}^k) \otimes_{\mathbb{C}} \Gamma_{\mathcal{C}^0}(\bar{\mathcal{L}}^k) \\
&\to \Gamma_{\mathcal{C}^0}(J^r\mathcal{E} \otimes \mathcal{L}^k \otimes \bar{\mathcal{L}}^k) \cong \Gamma_{\mathcal{C}^0}(J^r\mathcal{E}),
\end{aligned} \tag{25}$$

where the first arrow is induced by the map (21), the second arrow is the multiplication map and the last isomorphism is (23) applied to $(\mathcal{L} \otimes \bar{\mathcal{L}})^k \cong \mathcal{L}^k \otimes \bar{\mathcal{L}}^k$.

We construct the vertical map in the diagram (24) as the composition

$$\begin{aligned}
\gamma_k \colon \Gamma_{\mathrm{hol}}((J^r\mathcal{E}) \otimes \mathcal{L}^k) \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)} & \\
\to \Gamma_{\mathrm{hol}}((J^r\mathcal{E}) \otimes \mathcal{L}^k) \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)} &\otimes_{\mathbb{C}} (\Gamma_{\mathrm{hol}}(\mathcal{L}) \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L})}) \\
\cong (\Gamma_{\mathrm{hol}}((J^r\mathcal{E}) \otimes \mathcal{L}^k) \otimes_{\mathbb{C}} \Gamma_{\mathrm{hol}}(\mathcal{L})) &\otimes_{\mathbb{C}} (\overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)} \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L})}) \\
&\to \Gamma_{\mathrm{hol}}((J^r\mathcal{E}) \otimes \mathcal{L}^{k+1}) \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^{k+1})},
\end{aligned} \tag{26}$$

where the first arrow is given by tensoring with the element η defined in (22), the isomorphism is given by reordering the factors and the last arrow is given by the multiplication maps.

The commutativity of the diagram (24) follows directly from the fact that η is sent to the constant function equal to 1 via the isomorphism (23). Loosely speaking, the vertical map γ_k is a "multiplication by η", which amounts to multiplying a continuous section of $J^r\mathcal{E}$ by the constant function 1 after using the chosen identification (23).

Example 5.1 If $X = \mathbb{CP}^n$, $\mathcal{L} = \mathcal{O}(1)$ and $\mathcal{E} = \mathcal{O}(d+1)$, then $\Gamma_{\mathrm{hol}}(\mathcal{E})$ is the space of homogeneous polynomials of degree $d+1$ in $n+1$ variables. One may also prove an isomorphism $J^1(\mathcal{O}(d+1)) \cong \mathcal{O}(d)^{\oplus(n+1)}$ as holomorphic vector bundles; see [12, Proposition 2.2] for a proof.

We may then view $\Gamma_{\mathrm{hol}}((J^1\mathcal{E}) \otimes \mathcal{L}^k) \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)}$ as the space of $(n+1)$-tuples of homogeneous polynomials of bidegree $(d+k, k)$, that is, of degree $d+k$ in the variables z_i and of degree k in the complex conjugate variables $\bar{z}_i$. In this case, the image of η in $\Gamma_{\mathcal{C}^0}(\mathcal{L} \otimes \bar{\mathcal{L}})$ is $|z|^2 := z_0\bar{z}_0 + \cdots + z_n\bar{z}_n$. The isomorphism $\Gamma_{\mathcal{C}^0}(\mathcal{L} \otimes \bar{\mathcal{L}}) \cong \mathcal{C}^0(X, \mathbb{C})$ corresponding to (23) sends a section s to the map

$$z = [z_0 : \cdots : z_n] \in \mathbb{CP}^n \mapsto \frac{s(z)}{|z|^2} \in \mathbb{C}.$$

Under these identifications, the map γ_k is then

$$(f_0,\dots,f_n)\mapsto((z_0\bar z_0+\cdots+z_n\bar z_n)f_0,\dots,(z_0\bar z_0+\cdots+z_n\bar z_n)f_n),$$

which sends a tuple of polynomials of bidegree $(d+k,k)$ to one of bidegree $(d+k+1,k+1)$; compare [20] for a related situation.

We will need the following small result, analogous to Lemma 2.4. Let $(x_0,\dots,x_p)$ be a tuple of points in X. We may evaluate a continuous section of $J^r\mathcal{E}$ simultaneously at all these points:

$$\mathrm{ev}_{(x_0,\dots,x_p)}\colon \Gamma_{\mathcal{C}^0}(J^r\mathcal{E})\to(J^r\mathcal{E})|_{x_0}\times\cdots\times(J^r\mathcal{E})|_{x_p},\quad s\mapsto(s(x_0),\dots,s(x_p)). \tag{27}$$

Lemma 5.2 *Let $\mathcal{E}$ be a holomorphic vector bundle on X and $N(\mathcal{E},r)\in\mathbb{N}$ be as in Definition 2.3. Let $(x_0,\dots,x_p)$ be a tuple of $p+1$ distinct points in X. If $p\le N(\mathcal{E},r)$, the composition*

$$\Gamma_{\mathrm{hol}}((J^r\mathcal{E})\otimes\mathcal{L}^k)\otimes_{\mathbb{C}}\overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)}\xrightarrow{\varphi_k}\Gamma_{\mathcal{C}^0}(J^r\mathcal{E})\to(J^r\mathcal{E})|_{x_0}\times\cdots\times(J^r\mathcal{E})|_{x_p}$$

of the map φ_k of (25) and the simultaneous evaluation (27) is surjective.

Proof The case $k=0$ is a direct consequence of Lemma 2.4. The result for $k\ge1$ then follows from the commutativity of the diagram (24). □

5.3 Nonsingular sections

We define

$$\mathcal{N}(k)\subset\Gamma_{\mathrm{hol}}((J^r\mathcal{E})\otimes\mathcal{L}^k)\otimes_{\mathbb{C}}\overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)}$$

to be subspace of elements sent to nonsingular sections of $J^r\mathcal{E}$ (as in Definition 2.11) under the map φ_k defined in (25). We say that an $s\in\Gamma_{\mathrm{hol}}((J^r\mathcal{E})\otimes\mathcal{L}^k)\otimes_{\mathbb{C}}\overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)}$ is nonsingular if it is in the subspace $\mathcal{N}(k)$. We define the singular subset to be the complement

$$\mathcal{S}(k):=(\Gamma_{\mathrm{hol}}((J^r\mathcal{E})\otimes\mathcal{L}^k)\otimes_{\mathbb{C}}\overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)})-\mathcal{N}(k).$$

Remark 5.3 When $k=0$, $\mathcal{N}(0)\subset\Gamma_{\mathrm{hol}}(J^r\mathcal{E})$ is the usual subspace of nonsingular sections of $J^r\mathcal{E}$ as in Definition 2.11.

Example 5.4 In the case $X=\mathbb{CP}^n$, $\mathcal{L}=\mathcal{O}(1)$ and $\mathcal{E}=\mathcal{O}(d+1)$, recall from Example 5.1 that the space $\Gamma_{\mathrm{hol}}((J^1\mathcal{E})\otimes\mathcal{L}^k)\otimes_{\mathbb{C}}\overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)}$ corresponds to $(n+1)$-tuples of homogeneous polynomials of degree $d+k$ in the holomorphic variables z_i and of degree k in the complex conjugate variables $\bar z_i$. Under this identification, if the Taylor condition $\mathfrak{T}\subset J^1(\mathcal{O}(d+1))$ is the zero section, the space of nonsingular sections $\mathcal{N}(k)$ contains exactly those $(n+1)$-tuples of polynomials that never vanish simultaneously.

5.4 Resolution and spectral sequences

We now explain how the results from Section 3 can be adapted to the case

$$\Gamma=\Gamma_{\mathrm{hol}}((J^r\mathcal{E})\otimes\mathcal{L}^k)\otimes_{\mathbb{C}}\overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)}\quad\text{and}\quad\Sigma=\mathcal{S}(k)$$

to construct a resolution of $\mathcal{S}(k)$ and a spectral sequence converging to its cohomology, or equivalently to the homology of $\mathcal{N}(k)$ by Alexander duality. In this case, the definition of the singular space (3) of $f \in \Gamma$ has to be changed to

$$\mathrm{Sing}(f) := \{x \in X \mid \varphi_k(f)(x) \in \mathfrak{T}\} \subset X.$$

In particular, in the case $k = 0$, it agrees with Definition 2.11. The topological results about the resolution just follow from the fact that $\mathfrak{T} \subset J^r\mathcal{E}$ is closed. In particular, Lemma 3.5 still holds with its proof nearly unchanged: one has to replace the jet map j^r by φ_k. The construction of the spectral sequence is then unchanged.

The computations of cohomology groups on the E^1-page from Section 4 can also be adapted in this case. We first describe what to adapt for the first steps of the filtration. The analogue of Lemma 4.3 with the jet map j^r replaced by φ_k still holds as the key point is the surjectivity established in Lemma 5.2. The other result, Lemma 4.2, remains unchanged. Hence Proposition 4.4 is true in our new setting.

The adaptations are similar to examine the last step $R^N_{\mathrm{cone}}\mathfrak{X} - R^N\mathfrak{X}$. Indeed, the same stratification works, as well as the cohomological dimension estimates. In details, Lemma 4.6 is unchanged, and Lemma 4.9 is proved similarly by just replacing the jet map by φ_k. The other two results, Lemmas 4.7 and 4.8, also hold when rewriting the proof by changing the jet map j^r by φ_k. Indeed, the key ingredients were the semialgebraicity of the Taylor condition $\mathfrak{T}$ (which remains unchanged), and the fact that the jet map was complex algebraic, and hence real semialgebraic. The map φ_k is no longer complex algebraic, but is given by a ratio of algebraic maps and complex conjugates of algebraic maps. In particular, it is real semialgebraic. This is enough for the proof to go through.

To sum up, we have the following analogue of Proposition 4.1:

Proposition 5.5 *Let $\mathcal{E}$ be a holomorphic vector bundle on X and $\mathfrak{T} \subset J^r\mathcal{E}$ be an admissible Taylor condition. Let*

$$\Gamma = \Gamma_{\mathrm{hol}}((J^r\mathcal{E}) \otimes \mathcal{L}^k) \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)}$$

and $\mathcal{N}(k) \subset \Gamma$ be the subspace of nonsingular sections. Let $N = N(\mathcal{E}, r)$. The resolution and its filtration described in Section 3 give rise to a spectral sequence on the second quadrant $s \leq -1$ and $t \geq 0$ converging to the homology of the space of nonsingular sections:

$$E^1_{s,t} = H_c^{2\dim_{\mathbb{C}}\Gamma - 1 - s - t}(F_{-s-1} - F_{-s-2}; \mathbb{Z}) \Rightarrow \widetilde{H}_{s+t}(\mathcal{N}(k); \mathbb{Z}).$$

The differentials on the r^{th} page have bidegree $(-r, r-1)$. Furthermore, for $-N-1 \leq s \leq -1$, we have the following isomorphisms for all $t \geq 0$:

$$E^1_{s,t} \cong H_c^{-t-2s\,\mathrm{rk}_{\mathbb{C}} J^r\mathcal{E}}(\mathfrak{T}^{(-s)}; \mathbb{Z}^{\mathrm{sign}}).$$

Moreover, for $t < (N+1)e(\mathfrak{T})$,

$$E^1_{-N-2,t} = 0.$$

Lastly, let us mention that in the particular example where $X = \mathbb{CP}^n$, $\mathcal{L} = \mathcal{O}(1)$, $\mathcal{E} = \mathcal{O}(d+1)$ and $\mathfrak{T} \subset J^1\mathcal{E}$ is the zero section, the spectral sequence is completely analogous to that of [21].

6 Comparison of spectral sequences

From our definition of nonsingularity, it follows that the jet map j^r sends a nonsingular section f of $\mathcal{E}$ to a nonsingular section $j^r(f)$ of $J^r\mathcal{E}$. Likewise, the stabilisation map described in (26) sends elements in $\mathcal{N}(k)$ to elements in $\mathcal{N}(k+1)$. We shall see that these maps induce isomorphisms in homology in a range of degrees up to around $N = N(\mathcal{E}, r)$. We first explain the argument for the jet map j^r and then go through the required modifications for the stabilisation map.

6.1 The case of the jet map

Reading Propositions 4.1 and 5.5, we may observe that we have similar-looking spectral sequences, one converging to the homology of $\Gamma_{\mathrm{hol,ns}}(\mathcal{E})$ and the other one to that of $\Gamma_{\mathrm{hol,ns}}(J^r\mathcal{E}) = \mathcal{N}(0)$. In particular, in the range $-N-1 \le s \le -1$, the terms $E^1_{s,t}$ are given by the same cohomology groups

$$E^1_{s,t} \cong H_c^{-t-2s\,\mathrm{rk}_{\mathbb{C}}\, J^r\mathcal{E}}(\mathfrak{T}^{(-s)}; \mathbb{Z}^{\mathrm{sign}})$$

in both spectral sequences. If we had a morphism of spectral sequences that happened to be an isomorphism in this range, then, using the vanishing result $E^1_{-N-2,t} = 0$ for $t < (N+1)e(\mathfrak{T})$, the morphism induced on the E^∞-page would be an isomorphism in the range of degrees $* < N(e(\mathfrak{T})-1)+e(\mathfrak{T})-2$. (See Figure 1, where we have drawn some differentials.) We shall construct such a morphism of spectral sequences, whilst making sure that it is compatible with the morphism induced on homology by the jet map j^r:

$$\widetilde{H}_{s+t}(\Gamma_{\mathrm{hol,ns}}(\mathcal{E})) \to \widetilde{H}_{s+t}(\Gamma_{\mathrm{hol,ns}}(J^r\mathcal{E})).$$

For the sake of completeness, we recall when a morphism is compatible with a morphism of spectral sequences; see eg [32, Section 5.2]. If two spectral sequences $E^r_{p,q}$ and $E'^r_{p,q}$ converge to H_* and H'_*, respectively, we say that a map $h\colon H_* \to H'_*$ is *compatible* with a morphism $f : E \to E'$ if h maps F_pH_n to $F_pH'_n$ (here F_p denotes the filtration) and the associated maps $F_pH_n/F_{p-1}H_n \to F_pH'_n/F_{p-1}H'_n$ correspond to $f^\infty_{p,q}\colon E^\infty_{p,q} \to E'^\infty_{p,q}$ (where $q = n-p$) under the isomorphisms $E^\infty_{p,q} \cong F_pH_n/F_{p-1}H_n$ and $E'^\infty_{p,q} \cong F_pH'_n/F_{p-1}H'_n$. The main point being that if f is an isomorphism in a range, then h also is an isomorphism in a range; see [32, Comparison Theorem 5.2.12].

Let $d_1 := 2\dim_{\mathbb{C}} \Gamma_{\mathrm{hol}}(\mathcal{E})$ and $d_2 := 2\dim_{\mathbb{C}} \Gamma_{\mathrm{hol}}(J^r\mathcal{E})$ be the real dimensions of the complex vector spaces of sections. We define the shriek morphism $j^!$ as the unique morphism making the square

$$\begin{tikzcd}
\widetilde{H}_*(\Gamma_{\mathrm{hol,ns}}(\mathcal{E})) \arrow[r, "(j^r)_*"] \arrow[d, "\cong"'] & \widetilde{H}_*(\Gamma_{\mathrm{hol,ns}}(J^r\mathcal{E})) \arrow[d, "\cong"] \\
H_c^{d_1-1-*}(\Gamma_{\mathrm{hol}}(\mathcal{E}) - \Gamma_{\mathrm{hol,ns}}(\mathcal{E})) \arrow[r, dashed, "j^!"'] & H_c^{d_2-1-*}(\Gamma_{\mathrm{hol}}(J^r\mathcal{E}) - \Gamma_{\mathrm{hol,ns}}(J^r\mathcal{E}))
\end{tikzcd} \tag{28}$$

commutative, where the vertical isomorphisms are given by Alexander duality and the top map is induced by the jet map j^r in homology. As our spectral sequences actually converge to the Čech cohomology with compact support of the singular subspaces, we will construct our morphism of spectral sequences so that it is compatible with $j^!$.

The spectral sequences arose from filtrations, so we now recall some notation from Section 3. We let $\mathfrak{X}$ be the functor $\mathsf{F}^{\mathrm{op}} \to \mathrm{Top}$ constructed there using $\Gamma = \Gamma_{\mathrm{hol}}(\mathcal{E})$ and $\Sigma = \Gamma_{\mathrm{hol}}(\mathcal{E}) - \Gamma_{\mathrm{hol,ns}}(\mathcal{E})$. As we have explained in Section 5.4, the resolution also works for $\Gamma_{\mathrm{hol}}(J^r\mathcal{E})$ and its singular subspace, and we let $\mathfrak{Y}\colon \mathsf{F}^{\mathrm{op}} \to \mathrm{Top}$ be the associated functor in this case. We denote the filtration of $R^N_{\mathrm{cone}}\mathfrak{X}$ by

$$F^1_{-1} = \varnothing \subset F^1_0 = R^0\mathfrak{X} \subset \cdots \subset F^1_N = R^N\mathfrak{X} \subset F^1_{N+1} = R^N_{\mathrm{cone}}\mathfrak{X},$$

and the analogous one of $R^N_{\mathrm{cone}}\mathfrak{Y}$ by

$$F^2_{-1} = \varnothing \subset F^2_0 = R^0\mathfrak{Y} \subset \cdots \subset F^2_N = R^N\mathfrak{Y} \subset F^2_{N+1} = R^N_{\mathrm{cone}}\mathfrak{Y}. \tag{29}$$

We will slightly abuse notation and also write

$$j^!\colon H^*_c(R^N_{\mathrm{cone}}\mathfrak{X}) \to H^{*+d_2-d_1}_c(R^N_{\mathrm{cone}}\mathfrak{Y}) \tag{30}$$

for the bottom map defined by making the following square commutative:

$$\begin{tikzcd}
H^*_c(\Gamma_{\mathrm{hol}}(\mathcal{E}) - \Gamma_{\mathrm{hol,ns}}(\mathcal{E})) \arrow[r, "j^!"] \arrow[d, "\cong"'] & H^{*+d_2-d_1}_c(\Gamma_{\mathrm{hol}}(\mathcal{E}) - \Gamma_{\mathrm{hol,ns}}(J^r\mathcal{E})) \arrow[d, "\cong"] \\
H^*_c(R^N_{\mathrm{cone}}\mathfrak{X}) \arrow[r, dashed, "j^!"'] & H^{*+d_2-d_1}_c(R^N_{\mathrm{cone}}\mathfrak{Y})
\end{tikzcd}$$

Recall from the general theory that the spectral sequence associated to the filtration F^i_* for $i = 1, 2$, arises from an exact couple $(H^\bullet_c(F^i_*), H^\bullet_c(F^i_* - F^i_{*-1}))$. The map of spectral sequences that we want is then constructed via a map of exact couples as in the following lemma:

Lemma 6.1 *Let $\delta = d_2 - d_1 = 2(\dim_{\mathbb{C}} \Gamma_{\mathrm{hol}}(\mathcal{E}) - \dim_{\mathbb{C}} \Gamma_{\mathrm{hol}}(J^r\mathcal{E}))$. There exists a morphism of exact couples*

$$(j^!_p, j^!_{(p)})_{p\geq 0}\colon (H^*_c(F^1_p), H^*_c(F^1_p - F^1_{p-1})) \to (H^{*+\delta}_c(F^2_p), H^{*+\delta}_c(F^2_p - F^2_{p-1}))$$

satisfying the following two assertions:

(i) *For $0 \leq p \leq N$, the map $j^!_{(p)}$ in the diagram*

$$\begin{tikzcd}
H^*_c(F^1_p - F^1_{p-1}) \arrow[r, "\cong"] \arrow[d, "j^!_{(p)}"'] & H^\bullet_c(\mathfrak{T}^{(p+1)}; \mathbb{Z}^{\mathrm{sign}}) \\
H^{*+\delta}_c(F^2_p - F^2_{p-1}) \arrow[r, "\cong"'] & H^\bullet_c(\mathfrak{T}^{(p+1)}; \mathbb{Z}^{\mathrm{sign}})
\end{tikzcd} \tag{31}$$

is an isomorphism, where

$$\bullet = * - 2\dim_{\mathbb{C}} \Gamma_{\mathrm{hol}}(\mathcal{E}) - p + 2(p+1)\,\mathrm{rk}_{\mathbb{C}}\, J^r\mathcal{E},$$

and the horizontal isomorphisms are given by Thom isomorphisms as in Proposition 4.4.

(ii) *The map $j^!_{N+1}$ is equal to the shriek map (30).*

Unpacking the definition of a morphism of exact couples, we see that it amounts to providing morphisms $j_p^!$ and $j_{(p)}^!$ for $0 \le p \le N+1$ such that the diagram

$$\begin{tikzcd}
H_c^{*-1}(F_{p-1}^1) \arrow[r] \arrow[d, "j_{p-1}^!"] & H_c^*(F_p^1 - F_{p-1}^1) \arrow[r] \arrow[d, "j_{(p)}^!"] & H_c^*(F_p^1) \arrow[r] \arrow[d, "j_p^!"] & H_c^*(F_{p-1}^1) \arrow[d, "j_{p-1}^!"] \\
H_c^{*-1+\delta}(F_{p-1}^2) \arrow[r] & H_c^{*+\delta}(F_p^2 - F_{p-1}^2) \arrow[r] & H_c^{*+\delta}(F_p^2) \arrow[r] & H_c^{*+\delta}(F_{p-1}^2)
\end{tikzcd}$$

commutes, where the horizontal morphisms in the diagram are given by the long exact sequence of the pair (F_p^i, F_{p-1}^i) for $i = 1, 2$.

This result says exactly what we need: there a morphism of spectral sequences compatible with $j^!$ (by (ii)) and giving an isomorphism in the vertical strip $-N-1 \le s \le 1$ (by (i)). The lemma, as well as the strategy of proof, is adapted from [31, Proposition 4.7]. First, let us state the most important consequence:

Proposition 6.2 *For a holomorphic vector bundle $\mathcal{E}$ on X, the jet map*

$$j^r : \Gamma_{\mathrm{hol,ns}}(\mathcal{E}) \to \Gamma_{\mathrm{hol,ns}}(J^r\mathcal{E})$$

induces an isomorphism in homology in the range of degrees $* < N(\mathcal{E}, r)(e(\mathfrak{T}) - 1) + e(\mathfrak{T}) - 2$. □

To understand how to construct the degree-shifting morphisms of Lemma 6.1, it is helpful to give a description of the shriek map between cohomology groups arising from Alexander duality as in the diagram (28). We shall do so generally first (following [31, Appendix D]) and then specialise to our situation to prove the lemma at hand.

6.1.1 Alexander duality and shriek maps Let $p\colon E \to B$ be a vector bundle between oriented paracompact topological manifolds of dimensions n and m, respectively. Let $j : K \subset E$ be a closed subset, and let $i : B \hookrightarrow E$ be the zero section. We will see B as a submanifold of E via i. Using Alexander duality (the vertical isomorphisms in the diagram below), we may define the *shriek map*

$$i^! : H_c^*(B \cap K) \to H_c^{*+(n-m)}(K) \tag{32}$$

to be the unique morphism making the following diagram commute:

$$\begin{tikzcd}
H_*(B, B - B \cap K) \arrow[r, "i_*"] & H_*(E, E - K) \\
H_c^{m-*}(B \cap K) \arrow[u, "\cong"] \arrow[r, dashed, "i^!"'] & H_c^{n-*}(K) \arrow[u, "\cong"']
\end{tikzcd}$$

The goal of this section is to give a more intrinsic definition of $i^!$ that will allow us to define the required morphisms in Lemma 6.1.

Firstly, Vokřínek proves in [31, Proposition D.1] the following:

Lemma 6.3 *The diagram*

$$\begin{tikzcd} H_*(B, B-B\cap K) \arrow[rr, "i_*"] & & H_*(E, E-K) \\ H_c^{m-*}(B\cap K) \arrow[u, "\cong"] & H^{m-*}_{p^{-1}c}(K) \arrow[l, "k^*"] \arrow[r, "-\cup j^*\tau"'] & H_c^{n-*}(K) \arrow[u, "\cong"'] \end{tikzcd}$$

commutes, where the vertical isomorphisms are given by Alexander duality, $k\colon B\cap K \hookrightarrow K$ *is the inclusion,* $\tau \in H^\delta(D(E), S(E))$ *is the Thom class of* p *and* $p^{-1}c$ *is the family of supports defined as*

$$p^{-1}c = \{F \subset K \mid F \text{ closed and } \overline{p(F)} \subset B\cap K \text{ is compact}\},$$

so that $H^*_{p^{-1}c}$ *denotes cohomology with supports in* $p^{-1}c$. (See eg [8, Chapter II.2].)

Sketch of proof We repeat Vokřínek's proof here for convenience. First, we explain the morphisms in Alexander duality. Recall from eg [8, Corollary V.10.2] that we have fundamental classes $[B] \in H_m^{\mathrm{BM}}(B)$ and $[E] \in H_n^{\mathrm{BM}}(E)$, where H_*^{BM} denotes Borel–Moore homology (also known as homology with closed support). Using the proper inclusions $(E, \varnothing) \hookrightarrow (E, E-K)$ and $(B, \varnothing) \hookrightarrow (B, B-B\cap K)$, they give rise to classes $o_E \in H_n^{\mathrm{BM}}(E, E-K)$ and $o_B \in H_m^{\mathrm{BM}}(B, B-B\cap K)$. If $U \subset E$ is a closed neighbourhood of K, we get a morphism

$$H_c^{n-*}(U) \xrightarrow{-\cap o_E|U} H_*(U, U-K) \to H_*(E, E-K),$$

where $o_E|U$ is the image of o_E via the excision isomorphism $H_n^{\mathrm{BM}}(E, E-K) \cong H_n^{\mathrm{BM}}(U, U-K)$. (Note that it is important for U to be closed, so that the inclusion $U \hookrightarrow E$ is proper, and hence induces a morphism in Borel–Moore homology.) Likewise, we get a morphism

$$H_c^{m-*}(B\cap U) \xrightarrow{-\cap o_B|U} H_*(B\cap U, B\cap(U-K)) \to H_*(B, B-B\cap K).$$

Now, the isomorphisms in Alexander duality are given by taking the colimit over all closed neighbourhoods U of K of the two morphisms constructed above; this is explained in [8, V.9]. Hence, to prove the lemma, it suffices to check commutativity of the diagram

$$\begin{tikzcd} H_*(B\cap U, B\cap(U-K)) \arrow[rr, "g_*"] & & H_*(U, U-K) \\ H_c^{m-*}(B\cap U) \arrow[u, "-\cap o_B|U"] & H^{m-*}_{p^{-1}c}(U) \arrow[l, "g^*"] \arrow[r, "-\cup h^*\tau"'] \arrow[ur, "-\cap g_*(o_B|U)"] & H_c^{n-*}(U) \arrow[u, "-\cap o_E"'] \end{tikzcd}$$

where $g\colon B\cap U \hookrightarrow U$ and $h\colon U \hookrightarrow E$ are the inclusions. The left part commutes by naturality of the cap products. The right part commutes by observing that the fundamental classes can be chosen to correspond under the Thom isomorphism, which implies that $h^*\tau \cap o_E|U = g_*o_B|U$, and finishes the proof. □

In the statement of Lemma 6.3, if the morphism k^* were invertible, the shriek map (32) would be given by "$(k^*)^{-1}$" followed by taking the cup product with the "Thom class" $j^*\tau$. However, it is not invertible in general. There is nevertheless a way around that problem, which we explain below, using ε-small neighbourhoods of $B \cap K$ in K and the continuity property of cohomology.

We choose, once and for all, a bundle metric on $p\colon E \to B$. For a real number $\varepsilon > 0$, denote by D_ε (resp. S_ε, $\mathring{D}_\varepsilon$) the closed disc (resp. sphere, open disc) subbundle of $E \to B$ of radius ε (for the chosen metric). In [31, Lemma D.2], Vokřínek proves:

Lemma 6.4 *The diagram*

$$\begin{tikzcd}
H_*(B, B - B\cap K) \arrow[rr] & & H_*(E\cap \mathring{D}_\varepsilon, (E-K)\cap \mathring{D}_\varepsilon) \arrow[r] & H_*(E, E-K) \\
H_c^{m-*}(B\cap K) \arrow[u, "\cong"] & & H_c^{n-*}(K\cap \mathring{D}_\varepsilon) \arrow[u, "\cong"'] \arrow[r] \arrow[d, "\cong"'] & H_c^{n-*}(K) \arrow[u, "\cong"'] \\
& H_c^{m-*}(K\cap D_\varepsilon) \arrow[ul, "(l_\varepsilon)_*"] \arrow[r, "-\cup\tau_\varepsilon"'] & H_c^{n-*}(K\cap D_\varepsilon, K\cap S_\varepsilon) &
\end{tikzcd} \tag{33}$$

commutes, where the vertical isomorphisms on the first row are given by Alexander duality, the one on the second row follows from general results about cohomology with compact supports, $l_\varepsilon\colon B\cap K \hookrightarrow K\cap D_\varepsilon$ is the inclusion, τ_ε is the restriction of the Thom class of $E \to B$ and the rightmost horizontal arrows are induced by the inclusions. (Recall that cohomology with compact supports is covariant for open inclusions.)

Sketch of proof The left part of the diagram can be shown to commute by a proof analogous to that of Lemma 6.3. The right-hand square is seen to commute by a direct verification. □

Taking the limit $\varepsilon \to 0$, the morphisms $(l_\varepsilon)_*$ induce a morphism from the colimit

$$\operatorname*{colim}_{\varepsilon\to 0} H_c^{m-*}(K\cap D_\varepsilon) \to H_c^{m-*}(B\cap K)$$

which is an isomorphism by the continuity property of cohomology with compact supports; see eg [8, Theorem II.14.4]. We finally obtain another description of the shriek map $i^!$:

Proposition 6.5 (compare [31, Theorem D.3]) *The shriek map $i^!$ defined in (32) is equal to the composite obtained as one goes along the bottom path in the diagram (33) above:*

$$\begin{aligned} i^!\colon H_c^{m-*}(B\cap K) &\xleftarrow{\cong} \operatorname*{colim}_{\varepsilon\to 0} H_c^{m-*}(K\cap D_\varepsilon) \\ &\to \operatorname*{colim}_{\varepsilon\to 0} H_c^{n-*}(K\cap D_\varepsilon, K\cap S_\varepsilon) \cong \operatorname*{colim}_{\varepsilon\to 0} H_c^{n-*}(K\cap \mathring{D}_\varepsilon) \to H_c^{n-*}(K). \end{aligned}$$

Furthermore, in the case where both E and B are themselves vector bundles over the same base, $K = E$ and $i : B \hookrightarrow E$ is the inclusion of a subbundle, the shriek map $i^!$ is the Thom isomorphism of the bundle $E \to B$ given by choosing a splitting of i.

Proof The first part follows from Lemmas 6.3 and 6.4. The second part is shown by direct inspection of the construction. □

6.1.2 The proof of Lemma 6.1 We shall apply the general theory described in the last section to our case. To lighten the notation, we write

$$\Gamma_1 := \Gamma_{\mathrm{hol}}(\mathcal{E}), \quad \Sigma_1 := \Gamma_{\mathrm{hol}}(\mathcal{E}) - \Gamma_{\mathrm{hol,ns}}(\mathcal{E})$$

and

$$\Gamma_2 := \Gamma_{\mathrm{hol}}(J^r\mathcal{E}), \quad \Sigma_2 := \Gamma_{\mathrm{hol}}(J^r\mathcal{E}) - \Gamma_{\mathrm{hol,ns}}(J^r\mathcal{E}).$$

The jet map j^r gives a linear embedding of Γ_1 into Γ_2 such that the image of the singular subspace is precisely given by the intersection with the bigger singular subspace:

$$j^r(\Sigma_1) = j^r(\Gamma_1) \cap \Sigma_2.$$

Choosing a complementary linear subspace of $j^r(\Gamma_1)$ inside Γ_2, we obtain a projection giving a vector bundle

$$\Gamma_2 \to j^r(\Gamma_1) \cong \Gamma_1 \tag{34}$$

of real rank $\delta = d_2 - d_1$. Below, we apply Vokřínek's results to this situation.

We first set up the notation. Let $\varepsilon > 0$ be a positive real number and denote by D_ε (resp. S_ε, $\mathring{D}_\varepsilon$) the closed disc (resp. sphere, open disc) subbundle of radius ε of the vector bundle (34). Recall from (29) the functor $\mathfrak{Y}$ giving rise to the resolution of Σ_2. We also define $\mathfrak{Y}_{D_\varepsilon} \colon \mathsf{F}^{\mathrm{op}} \to \mathsf{Top}$ to be the subfunctor of $\mathfrak{Y}$ given by

$$\mathfrak{Y}_{D_\varepsilon}[n] := \{(f, s_0, \dots, s_n) \in \mathfrak{Y}[n] \mid f \in D_\varepsilon\},$$

and likewise for $\mathfrak{Y}_{S_\varepsilon} \subset \mathfrak{Y}$ and $\mathfrak{Y}_{\mathring{D}_\varepsilon} \subset \mathfrak{Y}$ using only sections $f \in S_\varepsilon$ or $\mathring{D}_\varepsilon$. Let $\tau_\varepsilon \in H^\delta(\Sigma_2 \cap D_\varepsilon, \Sigma_2 \cap S_\varepsilon)$ be the restriction of the Thom class of the vector bundle (34) to Σ_2. (Recall that the Thom class is an element of $H^\delta(D_\varepsilon, S_\varepsilon)$.) In all what follows, we see $\Gamma_1 \subset \Gamma_2$ via the embedding $j = j^r$. Let $l_\varepsilon \colon \Sigma_1 \hookrightarrow \Sigma_2 \cap D_\varepsilon$ be the inclusion (which is proper, and hence induces a morphism on compactly supported cohomology). We explained in Proposition 6.5 that the shriek map $j^!$ is obtained from the zigzag

$$H_c^*(\Sigma_1) \xleftarrow{(l_\varepsilon)_*} H_c^*(\Sigma_2 \cap D_\varepsilon) \xrightarrow{-\cup\tau_\varepsilon} H_c^{*+\delta}(\Sigma_2 \cap D_\varepsilon, \Sigma_2 \cap S_\varepsilon) \cong H_c^{*+\delta}(\Sigma_2 \cap \mathring{D}_\varepsilon) \to H_c^{*+\delta}(\Sigma_2)$$

by taking a colimit as $\varepsilon \to 0$.

We mimic that construction at the level of the resolutions. Let $0 \le p \le N+1$ be an integer. Recall from (29) that F_p^i denoted the p^{th} step of the filtration of the resolution of Σ_i. We denote by F^2_{p,D_ε}, F^2_{p,S_ε} and $F^2_{p,\mathring{D}_\varepsilon}$ the analogous filtrations on the resolutions obtained from the subfunctors $\mathfrak{Y}_{D_\varepsilon}$, $\mathfrak{Y}_{S_\varepsilon}$ and $\mathfrak{Y}_{\mathring{D}_\varepsilon}$, respectively. Because a singular point of a section $f \in \Gamma_1$ is also a singular point of $j^r(f) \in \Gamma_2$, the jet map gives a map on resolutions

$$\mathfrak{X}[p] \to \mathfrak{Y}[p], \quad (f, s_0, \dots, s_p) \mapsto (j^r(f), s_0, \dots, s_p),$$

which preserves the filtrations. Let $\tilde{l}_\varepsilon\colon F_p^1 \hookrightarrow F^2_{p,D_\varepsilon}$ be the induced inclusion. Let $\gamma_\varepsilon \in H^\delta(F^2_{p,D_\varepsilon}, F^2_{p,S_\varepsilon})$ be the pullback of τ_ε along $(F^2_{p,D_\varepsilon}, F^2_{p,S_\varepsilon}) \to (\Sigma_2 \cap D_\varepsilon, \Sigma_2 \cap S_\varepsilon)$. The diagram

$$\begin{tikzcd}
H_c^*(F_p^1) & H_c^*(F^2_{p,D_\varepsilon}) \arrow[l, "(\tilde{l}_\varepsilon)_*"'] \arrow[r, "-\cup\gamma_\varepsilon"] & H_c^{*+\delta}(F^2_{p,D_\varepsilon}, F^2_{p,S_\varepsilon}) \cong H_c^{*+\delta}(F^2_{p,\mathring{D}_\varepsilon}) \arrow[r] & H_c^{*+\delta}(F_p^2) \\
H_c^*(\Sigma_1) \arrow[u] & H_c^*(\Sigma_2 \cap D_\varepsilon) \arrow[u] \arrow[l, "(l_\varepsilon)_*"] \arrow[r, "-\cup\tau_\varepsilon"'] & H_c^{*+\delta}(\Sigma_2 \cap D_\varepsilon, \Sigma_2 \cap S_\varepsilon) \cong H_c^{*+\delta}(\Sigma_2 \cap \mathring{D}_\varepsilon) \arrow[u] \arrow[r] & H_c^{*+\delta}(\Sigma_2) \arrow[u]
\end{tikzcd}$$

then commutes by naturality of all the constructions involved, where all the vertical maps are induced by the proper projections $F_p^i \to \Sigma_i$. The morphism $j_p^!\colon H_c^*(F_p^1) \to H_c^{*+\delta}(F_p^2)$ is then defined as the colimit, when $\varepsilon \to 0$, of the top composition in the diagram above. (Recall that $(\tilde{l}_\varepsilon)_*$ is an isomorphism in the colimit, by continuity of cohomology.) In particular, when $p = N+1$, the vertical maps are isomorphisms (by 3.11), which proves Lemma 6.1(ii) by noticing that the bottom composition is the shriek map $j^!$.

The morphisms $j^!_{(p)}\colon H_c^*(F_p^1 - F_{p-1}^1) \to H_c^{*+\delta}(F_p^2 - F_{p-1}^2)$ are defined analogously, ie by the colimit as $\varepsilon \to 0$ of the zigzag

$$\begin{aligned} H_c^*(F_p^1 - F_{p-1}^1) \leftarrow H_c^*(F^2_{p,D_\varepsilon} - F^2_{p-1,D_\varepsilon}) \to H_c^{*+\delta}(F^2_{p,D_\varepsilon} - F^2_{p-1,D_\varepsilon}, F^2_{p,S_\varepsilon} - F^2_{p-1,S_\varepsilon}) \\ \cong H_c^{*+\delta}(F^2_{p,\mathring{D}_\varepsilon} - F^2_{p-1,\mathring{D}_\varepsilon}) \to H_c^{*+\delta}(F_p^2 - F_{p-1}^2), \end{aligned}$$

where, as before, the first morphism is induced by the inclusion, the second morphism is the cup product with the Thom class and the third is induced covariantly by the open inclusion.

One may check, using naturality of the various constructions involved, that the morphisms $j_p^!$ and $j^!_{(p)}$ give a morphism of exact couples. This amounts to staring at the following commutative diagram:

$$\begin{tikzcd}
H_c^{*-1}(F^1_{p-1}) \arrow[r] & H_c^*(F_p^1 - F^1_{p-1}) \arrow[r] & H_c^*(F_p^1) \arrow[r] & H_c^*(F^1_{p-1}) \\
H_c^{*-1}(F^2_{p-1,D_\varepsilon}) \arrow[r] \arrow[u] \arrow[d] & H_c^*(F^2_{p,D_\varepsilon} - F^2_{p-1,D_\varepsilon}) \arrow[r] \arrow[u] \arrow[d] & H_c^*(F^2_{p,D_\varepsilon}) \arrow[r] \arrow[u] \arrow[d] & H_c^*(F^2_{p-1,D_\varepsilon}) \arrow[u] \arrow[d] \\
H_c^{*-1+\delta}(F^2_{p-1,D_\varepsilon}, F^2_{p-1,S_\varepsilon}) \arrow[r] \arrow[d, "\cong"] & H_c^{*+\delta}(F^2_{p,D_\varepsilon} - F^2_{p-1,D_\varepsilon}, F^2_{p,S_\varepsilon} - F^2_{p-1,S_\varepsilon}) \arrow[r] \arrow[d, "\cong"] & H_c^{*+\delta}(F^2_{p,D_\varepsilon}, F^2_{p,S_\varepsilon}) \arrow[r] \arrow[d, "\cong"] & H_c^{*+\delta}(F^2_{p-1,D_\varepsilon}, F^2_{p-1,S_\varepsilon}) \arrow[d, "\cong"] \\
H_c^{*-1+\delta}(F^2_{p-1,\mathring{D}_\varepsilon}) \arrow[r] \arrow[d] & H_c^{*+\delta}(F^2_{p,\mathring{D}_\varepsilon} - F^2_{p-1,\mathring{D}_\varepsilon}) \arrow[r] \arrow[d] & H_c^{*+\delta}(F^2_{p,\mathring{D}_\varepsilon}) \arrow[r] \arrow[d] & H_c^{*+\delta}(F^2_{p-1,\mathring{D}_\varepsilon}) \arrow[d] \\
H_c^{*-1+\delta}(F^2_{p-1}) \arrow[r] & H_c^{*+\delta}(F_p^2 - F^2_{p-1}) \arrow[r] & H_c^{*+\delta}(F_p^2) \arrow[r] & H_c^{*+\delta}(F^2_{p-1})
\end{tikzcd}$$

To conclude the proof, we verify Lemma 6.1(i), ie that the morphism

$$j^!_{(p)}\colon H_c^*(F_p^1 - F_{p-1}^1) \to H_c^{*+\delta}(F_p^2 - F_{p-1}^2)$$

is an isomorphism. Recall from (10) that

$$F_p^2 - F_{p-1}^2 \cong Y_p(\mathfrak{Y}) \times_{\mathfrak{S}_{p+1}} |\mathring{\Delta}^p| \quad \text{and} \quad F_p^1 - F_{p-1}^1 \cong Y_p(\mathfrak{X}) \times_{\mathfrak{S}_{p+1}} |\mathring{\Delta}^p|,$$

where we defined as in (9) the subspace

$$Y_p(\mathfrak{Y}) := \{(f, s_0, \dots, s_p) \in \mathfrak{Y}[p] \mid s_l \neq s_k \text{ if } l \neq k\} \subset \mathfrak{Y}[p],$$

and likewise for $Y_p(\mathfrak{X}) \subset \mathfrak{X}[p]$. Recall also that these spaces were vector bundles over $\mathfrak{T}^{(p+1)}$; see Section 4. Hence we have an inclusion of vector bundles:

$$\begin{tikzcd} F_p^1 - F_{p-1}^1 \arrow[rr, hook] \arrow[dr] & & F_p^2 - F_{p-1}^2 \arrow[dl] \\ & \mathfrak{T}^{(p+1)} & \end{tikzcd}$$

Now the second part of Proposition 6.5 applies and finishes the proof. $\square$

6.2 The case of the stabilisation map

Choose some integer $k \geq 0$. We now describe how the argument of the previous section can be made with the stabilisation map

$$\gamma_k : \Gamma_{\mathrm{hol}}((J^r \mathcal{E}) \otimes \mathcal{L}^k) \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)} \to \Gamma_{\mathrm{hol}}((J^r \mathcal{E}) \otimes \mathcal{L}^{k+1}) \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^{k+1})}$$

from (26). It is a linear embedding; hence, by choosing a complementary subspace, we get a vector bundle

$$\Gamma_{\mathrm{hol}}((J^r \mathcal{E}) \otimes \mathcal{L}^{k+1}) \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^{k+1})} \to \gamma_k(\Gamma_{\mathrm{hol}}((J^r \mathcal{E}) \otimes \mathcal{L}^k) \otimes_{\mathbb{C}} \overline{\Gamma_{\mathrm{hol}}(\mathcal{L}^k)})$$

analogous to the one in (34). From the commutativity of the diagram (24), we see that a singularity $x \in X$ for $f \in \mathcal{S}(k)$ is also a singularity of $\gamma_k(f) \in \mathcal{S}(k+1)$. Therefore we also get a map induced on the respective resolutions of $\mathcal{S}(k)$ and $\mathcal{S}(k+1)$. Together with the fact that nonsingular sections are sent to nonsingular sections, this is enough for the argument to be repeated in that case.

Proposition 6.6 *The restriction of the stabilisation map γ_k to the nonsingular subspaces*

$$\gamma_k : \mathcal{N}(k) \to \mathcal{N}(k+1)$$

induces an isomorphism in homology in the range of degrees $* < N(\mathcal{E}, r)(e(\mathfrak{T}) - 1) + e(\mathfrak{T}) - 2$. $\square$

Combining Propositions 6.2 and 6.6, we obtain the following:

Proposition 6.7 *Each map in the composition*

$$\Gamma_{\mathrm{hol,ns}}(\mathcal{E}) \to \Gamma_{\mathrm{hol,ns}}(J^r \mathcal{E}) = \mathcal{N}(0) \to \operatorname*{colim}_{k \to \infty} \mathcal{N}(k)$$

induces an isomorphism in homology in the range of degrees $* < N(\mathcal{E}, r)(e(\mathfrak{T}) - 1) + e(\mathfrak{T}) - 2$. $\square$

7 Comparison of holomorphic and continuous sections

We shall relate $\operatorname{colim}_k \mathcal{N}(k)$ to the space $\Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^r \mathcal{E})$ of nonsingular continuous sections of the jet bundle. Recall from the stabilisation diagram (24) that every nonsingular space $\mathcal{N}(k)$ maps via φ_k to $\Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^r \mathcal{E})$. The aim of this section is to prove the following result about the map induced from the colimit:

Proposition 7.1 *The map*

$$\operatorname*{colim}_{k\to\infty} \mathcal{N}(k) \to \Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^r\mathcal{E}) \tag{35}$$

is a weak homotopy equivalence.

Combining this result with Proposition 6.7 readily implies Theorem 2.13. Proposition 7.1 is a direct consequence of the openness of the subspace of nonsingular sections, which follows from the fact that the admissible Taylor condition $\mathfrak{T} \subset J^r\mathcal{E}$ is closed (see the discussion after Lemma 3.6), and the following:

Lemma 7.2 *Let F be a finite CW-complex. The map*

$$\mathcal{C}^0\big(F, \operatorname*{colim}_{k\to\infty} \mathcal{N}(k)\big) \to \mathcal{C}^0(F, \Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^r\mathcal{E}))$$

induced by (35) has a dense image.

As in [20], we will need an adaptation of the classical Stone–Weierstrass theorem for real vector bundles.

Theorem 7.3 (Stone–Weierstrass) *Let $E \to B$ be a finite-rank real vector bundle over a compact Hausdorff space. Let $A \subset \mathcal{C}^0(B,\mathbb{R})$ be a subalgebra and $\{s_j\}_{j\in J}$ be a set of sections such that*

(i) *the subalgebra A separates the points of B: for any $x, y \in B$, there exists $h \in A$ such that $h(x) \neq h(y)$,*

(ii) *for any $x \in B$, there exists $h \in A$ such that $h(x) \neq 0$,*

(iii) *for any $x \in B$, the fibre E_x is spanned by the $s_j(x)$ as an $\mathbb{R}$-vector space.*

Then the A-module generated by the s_j is dense for the sup-norm (induced by the choice of any inner product on E) in the space of all continuous sections of E.

Proof of Lemma 7.2 Let F be a finite CW-complex. By adjunction, a continuous map $F \to \Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^r\mathcal{E})$ corresponds to a section of the underlying real vector bundle of $J^r\mathcal{E}\times F \to X \times F$. We shall apply Theorem 7.3 to that vector bundle.

Recall that we have chosen in Section 5 a very ample line bundle $\mathcal{L}$ on X and explained how to define the complex conjugate $\bar{s}$ of a section s of $\mathcal{L}$. For any integer $k \geq 0$, define the squared norm of a holomorphic section of $\mathcal{L}$ by

$$|\cdot|^2 \colon \Gamma_{\mathrm{hol}}(\mathcal{L}^k) \to \Gamma_{\mathcal{C}^0}(\mathcal{L}^k \otimes \bar{\mathcal{L}}^k) \cong \mathcal{C}^0(X,\mathbb{C}), \quad s \mapsto |s|^2 := s\bar{s},$$

where the isomorphism with continuous maps was obtained in (23). Notice that $|s|^2$ is in fact a real-valued function $X \to \mathbb{R} \subset \mathbb{C}$. We also let

$$A_k := \{|g(\cdot,\cdot)|^2 \colon X \times F \to \mathbb{R} \mid g \in \mathcal{C}^0(F, \Gamma_{\mathrm{hol}}(\mathcal{L}^k))\} \subset \mathcal{C}^0(X \times F, \mathbb{R}),$$

where if $g \in \mathcal{C}^0(F, \Gamma_{\mathrm{hol}}(\mathcal{L}^k))$, we see $g(\cdot,\cdot)$ as a map from $X \times F$ to $\mathcal{L}^k$ by adjunction. Keeping the notation from Theorem 7.3, we let A be the subalgebra of $\mathcal{C}^0(X \times F, \mathbb{R})$ generated by all the A_k for $k \geq 0$. For the set of sections as in Theorem 7.3, take

$$\{(x,u) \mapsto (s(x,u),u) \colon X \times F \to J^r\mathcal{E} \times F \mid s \in \mathcal{C}^0(F, \Gamma_{\mathrm{hol}}(J^r\mathcal{E}))\}, \tag{36}$$

where again, for $s \in \mathcal{C}^0(F, \Gamma_{\mathrm{hol}}(J^r\mathcal{E}))$, we see $s(\cdot,\cdot)$ as a map from $X \times F$ to $J^r\mathcal{E}$ by adjunction. We may now check the conditions of Theorem 7.3.

(i) Let $(x,u) \neq (x',u') \in X \times F$. Consider the first case, where $x \neq x'$. For $k \geq 2$, $\mathcal{L}^k$ is 2-very ample (see Example 2.2). Hence there exists a section $s \in \Gamma_{\mathrm{hol}}(\mathcal{L}^2)$ such that $s(x) \neq 0$ and $s(x') = 0$. Then the map $(x,u) \mapsto |s(x)|^2$ is in A_k and separates (x,u) and (x',u') as $|s(x)|^2 \neq 0$ and $|s(x')|^2 = 0$. In the other case, where $x = x'$, we have that $u \neq u'$. By the 1-very ampleness of $\mathcal{L}$ we may choose $s \in \Gamma_{\mathrm{hol}}(\mathcal{L})$ such that $s(x) = s(x') \neq 0$. Let $\rho\colon F \to \mathbb{R}_+$ be a bump function such that $\rho(u) = 0$ and $\rho(u') = 1$. Then the map $(x,u) \mapsto |\rho(u)s(x)|^2$ is in A_1 and separates the points. Indeed it is vanishing at (x,u) but nonvanishing at (x',u').

(ii) The second point is exactly what we have just proved in the first case of (i).

(iii) It suffices to prove that the fibre of $J^r\mathcal{E}$ above $x \in X$ is spanned by the sections $s(x)$ for $s \in \Gamma_{\mathrm{hol}}(J^r\mathcal{E})$. This is implied by the 0-jet ampleness of $\mathcal{E}$ (see Example 2.2).

By construction, any element in the image of the map

$$\mathcal{C}^0\big(F, \operatorname*{colim}_{k\to\infty} \mathcal{N}(k)\big) \to \mathcal{C}^0(F, \Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^r\mathcal{E}))$$

is, by adjunction, in the A-module generated by the set (36). □

8 Applications

8.1 Nonsingular sections of line bundles

Our first application concerns the case of nonsingular sections of line bundles, which was the starting motivation for this work. Here, a direct corollary of our main theorem reads as:

Corollary 8.1 *Let X be a smooth projective complex variety and $\mathcal{L}$ be a very ample line bundle on it. Let $d \geq 1$ be an integer. The jet map*

$$j^1\colon \Gamma_{\mathrm{hol,ns}}(\mathcal{L}^d) \to \Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^1\mathcal{L}^d)$$

from nonsingular holomorphic sections of $\mathcal{L}^d$ to continuous never-vanishing sections of $J^1\mathcal{L}^d$, induces an isomorphism in homology in the range of degrees $* < \frac{1}{2}(d-1)$.

Proof It is a straightforward application of Theorem 2.13 by taking the admissible Taylor condition $\mathfrak{T}$ to be the zero section of $J^1\mathcal{L}^d$ and recalling from Example 2.2 that if $\mathcal{L}$ is very ample, then the tensor power $\mathcal{L}^d$ is d-very ample. □

More interestingly, we can compute the stable rational cohomology. This agrees with a computation made by Tommasi (personal communication, 2021).

Theorem 8.2 *Let $n = \dim_{\mathbb{C}} X$ be the complex dimension of X. For $d \geq 1$, there is a rational homotopy equivalence*

$$\Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^1\mathcal{L}^d) \xrightarrow{\simeq_{\mathbb{Q}}} \prod_{i=1}^{2n+1} K(H_{i-1}(X;\mathbb{Q}), i).$$

In particular, the rational cohomology of $\Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^1\mathcal{L}^d)$ is given by the free commutative graded algebra

$$\Lambda(H^{*-1}(X;\mathbb{Q})),$$

on the cohomology of X shifted by one degree.

Remark 8.3 This result implies in particular that the rational (co)homology of $\Gamma_{\mathrm{hol,ns}}(\mathcal{L}^d)$ stabilises as $d \to \infty$. As we will see below, the integral cohomology does not stabilise in general.

Remark 8.4 The stable cohomology only depends on the topology of X. This is in accordance with the analogies between topology and arithmetic and motivic statistics mentioned in the introduction. In both the results of Poonen and Vakil–Wood, the limit is expressed by a zeta function which only depends on X.

Example 8.5 For $X = \mathbb{CP}^n$ and $\mathcal{L} = \mathcal{O}(1)$, we find that the stable rational cohomology is the exterior algebra

$$\Lambda_{\mathbb{Q}}(t_1, t_3, \dots, t_{2n+1})$$

where t_i is in degree i. This agrees with the result of Tommasi in [27].

Proof of Theorem 8.2 Recall that the nonsingular sections of $J^1\mathcal{L}^d$ are precisely the never-vanishing ones. We choose a Riemannian metric once and for all and denote by $\mathrm{Sph}(J^1\mathcal{L}^d) \to X$ the unit sphere bundle of the vector bundle $J^1\mathcal{L}^d$. We may scale a never-vanishing section to have norm equal to 1 (for the chosen metric) in each fibre. We thus obtain a homotopy equivalence

$$\Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^1\mathcal{L}^d) \xrightarrow{\simeq} \Gamma_{\mathcal{C}^0}(\mathrm{Sph}(J^1\mathcal{L}^d)).$$

We now rationalise the sphere bundle in the following sense. By [17, Theorem 3.2], there is a fibration $S^{2n+1}_{\mathbb{Q}} \to \mathrm{Sph}(J^1\mathcal{L}^d)_{\mathbb{Q}} \to X$ and a morphism of fibrations

$$\begin{tikzcd}
S^{2n+1} \arrow[r] \arrow[d] & S^{2n+1}_{\mathbb{Q}} \arrow[d] \\
\mathrm{Sph}(J^1\mathcal{L}^d) \arrow[r] \arrow[dr] & \mathrm{Sph}(J^1\mathcal{L}^d)_{\mathbb{Q}} \arrow[d] \\
 & X
\end{tikzcd}$$

such that the map induced on the fibres $S^{2n+1} \to S^{2n+1}_{\mathbb{Q}} \simeq K(\mathbb{Q}, 2n+1)$ is a rationalisation. As X is homotopy equivalent to a finite CW-complex and S^{2n+1} is nilpotent (it is indeed simply connected), we may use [19, Theorem 5.3] to show that the map $\mathrm{Sph}(J^1\mathcal{L}^d) \to \mathrm{Sph}(J^1\mathcal{L}^d)_{\mathbb{Q}}$ induces a map

$$\Gamma_{\mathcal{C}^0}(\mathrm{Sph}(J^1\mathcal{L}^d)) \xrightarrow{\simeq_{\mathbb{Q}}} \Gamma_{\mathcal{C}^0}(\mathrm{Sph}(J^1\mathcal{L}^d)_{\mathbb{Q}}),$$

which is a rationalisation. (In general, one has to restrict to some path component. However both spaces are connected in our situation.) Now, oriented rational odd sphere bundles are classified by their Euler class; see eg [13, II.15.b]. In our situation, the orientation is induced from the canonical one on the complex vector bundle $J^1\mathcal{L}^d$ and the Euler class vanishes for dimensional reasons. It follows directly that $\mathrm{Sph}(J^1\mathcal{L}^d)_{\mathbb{Q}} \to X$ is a trivial bundle. Therefore

$$\Gamma_{\mathcal{C}^0}(\mathrm{Sph}(J^1\mathcal{L}^d)_{\mathbb{Q}}) \cong \mathrm{map}(X, K(\mathbb{Q}, 2n+1)),$$

where $\mathrm{map}(-,-)$ denotes the space of continuous functions with its compact open topology. Finally, in [26] (see also [15] for an accessible reference), Thom proves that this mapping space is homotopy equivalent to a product of Eilenberg–MacLane spaces

$$\mathrm{map}(X, K(\mathbb{Q}, 2n+1)) \simeq \prod_{i=0}^{2n+1} K(H^{2n+1-i}(X;\mathbb{Q}), i) \simeq \prod_{i=0}^{2n+1} K(H_{i-1}(X;\mathbb{Q}), i),$$

where the last equivalence comes from Poincaré duality. More precisely, he proves that if

$$\mathrm{ev}\colon \mathrm{map}(X, K(\mathbb{Q}, 2n+1)) \times X \to K(\mathbb{Q}, 2n+1)$$

is the evaluation map, and $\chi \in H^{2n+1}(K(\mathbb{Q}, 2n+1);\mathbb{Q})$ is the fundamental class, we may write

$$\mathrm{ev}^*(\chi) = \sum_i \chi_i,$$

where $\chi_i \in H^i(\mathrm{map}(X, K(\mathbb{Q}, 2n+1)); H^{2n+1-i}(X;\mathbb{Q}))$. Then the projection

$$\mathrm{map}(X, K(\mathbb{Q}, 2n+1)) \to K(H^{2n+1-i}(X;\mathbb{Q}), i)$$

is determined by the cohomology class χ_i. □

8.1.1 Geometric description of the stable classes and mixed Hodge structures As a Zariski open subset of the affine space $\Gamma_{\mathrm{hol}}(\mathcal{L}^d)$, the subspace $\Gamma_{\mathrm{hol,ns}}(\mathcal{L}^d)$ inherits a structure of complex variety and its cohomology thus has a natural mixed Hodge structure. On the other hand, we may endow the stable cohomology computed in Theorem 8.2 with a mixed Hodge structure defined as follows. Recall that the cohomology $H^*(X;\mathbb{Q})$ can be equipped with a mixed Hodge structure using the structure of complex variety on X, and denote by $\mathbb{Q}(-1)$ the Tate–Hodge structure of pure weight 2. By first tensoring these structures and then applying the symmetric algebra functor, we obtain a mixed Hodge structure on the stable cohomology. In this section, we show the following:

Proposition 8.6 *The morphism of Theorem 8.2,*

$$\Lambda(H^{*-1}(X;\mathbb{Q})\otimes\mathbb{Q}(-1))\to H^*(\Gamma_{\mathrm{hol,ns}}(\mathcal{L}^d);\mathbb{Q}),$$

is compatible with the mixed Hodge structures.

Proof By the universal property of the (graded) symmetric algebra, it is enough to see that the morphism

$$H^{*-1}(X;\mathbb{Q})\otimes\mathbb{Q}(-1)\to H^*(\Gamma_{\mathrm{hol,ns}}(\mathcal{L}^d);\mathbb{Q})$$

respects the mixed Hodge structures. We do this by giving a more geometric description of this map. Let

$$\pi\colon \Gamma_{\mathrm{hol,ns}}(\mathcal{L}^d)\times X\to\Gamma_{\mathrm{hol,ns}}(\mathcal{L}^d)$$

be the trivial fibre bundle, and let

$$j\colon \Gamma_{\mathrm{hol,ns}}(\mathcal{L}^d)\times X\to J^1\mathcal{L}^d-\{0\}$$

be the jet evaluation. By integrating along the fibres of π, we obtain in cohomology a morphism of mixed Hodge structures:

$$\pi_!\circ j^*\colon H^*(J^1\mathcal{L}^d-\{0\})\otimes\mathbb{Q}(n)\to H^{*-2n}(\Gamma_{\mathrm{hol,ns}}(\mathcal{L}^d)).$$

The extra Tate twist $\mathbb{Q}(n)$ comes from the definition of the Gysin map $\pi_!$ via Poincaré duality; see [23, Corollary 6.25]. As the Euler class of the jet bundle vanishes for dimensional reasons, we compute that

$$H^*(J^1\mathcal{L}^d-\{0\};\mathbb{Q})\cong H^*(X;\mathbb{Q})\otimes H^*(\mathbb{C}^{n+1}-\{0\};\mathbb{Q}).$$

Now $H^{2n+1}(\mathbb{C}^{n+1}-\{0\};\mathbb{Q})\cong\mathbb{Q}(-n-1)$, so we have obtained a morphism of mixed Hodge structures:

$$\pi_!\circ j^*\colon H^*(X)\otimes\mathbb{Q}(-1)\to H^{*+1}(\Gamma_{\mathrm{hol,ns}}(\mathcal{L}^d)).$$

We claim that this coincides with the morphism given in Theorem 8.2. The proof is an exercise in algebraic topology and uses the description of the mapping space given at the end of the proof of Theorem 8.2. □

8.2 Integral homology and stability

In this section, we focus on the special case where $X=\mathbb{CP}^1$ and $\mathcal{L}=\mathcal{O}(1)$. That is, we study the space

$$U_d:=\Gamma_{\mathrm{hol,ns}}(\mathbb{CP}^1,\mathcal{O}(d))$$

of nonsingular homogeneous polynomials in two variables of degree d. From Corollary 8.1, we know that the jet map

$$j^1\colon U_d\to\Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^1\mathcal{O}(d))$$

induces an isomorphism in integral homology in the range of degrees $*<\frac{1}{2}(d-1)$. We prove that the section space on the right-hand side does not depend on $d\ge 1$, and hence that the integral homology of U_d stabilises.

Theorem 8.7 *For $d \geq 1$, we have a homotopy equivalence*

$$\Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^1\mathcal{O}_{\mathbb{CP}^1}(d)) \simeq \mathrm{map}(S^2, S^3).$$

In particular

$$H_*(U_d; \mathbb{Z}) \cong H_*(\mathrm{map}(S^2, S^3); \mathbb{Z})$$

in the range of degrees $* < \frac{1}{2}(d-1)$.

Remark 8.8 Choosing a basepoint $b \in S^2$ and using the Lie group structure on $S^3 \cong SU(2)$, we obtain a homeomorphism

$$\mathrm{map}(S^2, S^3) \xrightarrow{\cong} S^3 \times \mathrm{map}_*(S^2, S^3) = S^3 \times \Omega^2 S^3, \quad f \mapsto (f(b), f(b)^{-1} f),$$

which can be used to compute the integral homology. This can be done one prime at a time. Indeed, the p-primary elements have order exactly p by [22, Corollary 10.26.5]. This p-primary part can then be computed directly from the Bockstein spectral sequence and the knowledge of the $\mathbb{Z}/p$ homology, which is recalled in [22, Corollary 10.26.4]. We can also note that the homology of $\Omega^2 S^3 \simeq \Omega^2_0 S^2$ is the stable homology of braid groups studied in [10, Paper III, Appendix A].[2]

Remark 8.9 In the next section, we will show that one cannot expect integral homological stability in general. The case $X = \mathbb{CP}^1$ should be seen as a very particular phenomenon.

Proof Recall from the proof of Theorem 8.2 that we have to study continuous sections of the sphere bundle of the jet bundle:

$$S^3 \to \mathrm{Sph}(J^1\mathcal{O}_{\mathbb{CP}^1}(d)) \to \mathbb{CP}^1.$$

One sees that this bundle is classified by the second Stiefel–Whitney class of the jet bundle, ie the reduction modulo 2 of its first Chern class. Using that $d \geq 1$ and [12, Proposition 2.2], we obtain an isomorphism of vector bundles

$$J^1\mathcal{O}_{\mathbb{CP}^1}(d) \cong \mathcal{O}_{\mathbb{CP}^1}(d-1)^{\oplus 2}.$$

We compute the first Chern class to be

$$c_1(J^1\mathcal{O}_{\mathbb{CP}^1}(d)) = c_1(\mathcal{O}_{\mathbb{CP}^1}(d-1)^{\oplus 2}) = 2c_1(\mathcal{O}_{\mathbb{CP}^1}(d-1)),$$

so its reduction modulo 2 vanishes regardless of d. As the sphere bundle was classified by this class, this shows that it is trivial. Therefore

$$\Gamma_{\mathcal{C}^0,\mathrm{ns}}(J^1\mathcal{O}_{\mathbb{CP}^1}(d)) \simeq \Gamma_{\mathcal{C}^0}\big(\mathrm{Sph}(J^1\mathcal{O}_{\mathbb{CP}^1}(d))\big) \simeq \mathrm{map}(S^2, S^3). \qquad \square$$

[2] Many thanks to Antoine Touzé for explaining this computation to me, and to the referee for pointing out the connection to braid groups.

8.3 Integral homology and nonstability

As we indicated in Remark 8.3, the rational cohomology groups of the spaces $\Gamma_{\mathrm{hol,ns}}(\mathcal{L}^d)$ stabilise. That is, for a fixed $i \geq 0$, the i^{th} rational cohomology group is independent of d as long as $i \leq \frac{1}{2}(d-1)$. In this section, to contrast with the very special case of the previous one, we show that one cannot expect integral stability in general.

Let us fix some notation for the remainder of this section: $d \geq 1$ is an integer, $\mathcal{L}$ is a very ample line bundle on a smooth projective complex variety X and $n = \dim_{\mathbb{C}} X$ is the complex dimension of X. As we will only be considering spaces of continuous sections, we will use Γ as a shorthand for $\Gamma_{\mathcal{C}^0}$.

The main result of this section is Theorem 8.11. To show its computational potential, we will show the following:

Proposition 8.10 *Let $d \geq 6$ be an integer. We have*

$$H_2\big(\Gamma_{\mathrm{hol,ns}}(\mathbb{CP}^2, \mathcal{O}(d)); \mathbb{Z}/2\big) \cong \begin{cases} \mathbb{Z}/2 & d \equiv 0 \bmod 2, \\ 0 & d \equiv 1 \bmod 2. \end{cases}$$

8.3.1 A comparison map As stated in Corollary 8.1, we are reduced to studying the homotopy type of the space of continuous sections of the sphere bundle $\mathrm{Sph}(J^1\mathcal{L}^d)$. Even though this is a purely homotopy-theoretic problem, its resolution is quite hard. We will therefore "linearise it" in the homotopical sense using spectra. This is made precise in the following result:

Theorem 8.11 *Let TX be the tangent bundle of X, and let $X^{J^1\mathcal{L}^d - TX}$ denote the Thom spectrum of the virtual bundle $J^1\mathcal{L}^d - TX$ of rank 2. There is a $2n$-connected map*

$$\Gamma(\mathrm{Sph}(J^1\mathcal{L}^d)) \to \Omega^{\infty+1} X^{J^1\mathcal{L}^d - TX}.$$

Our proof uses very lightly the theory of parametrised pointed spaces/spectra and is written using ∞-categories. We feel that this choice helps in conveying the main ideas more clearly. The unfamiliar reader is encouraged to think of bundles of pointed spaces/spectra, whilst resting assured that there exists a theory which renders all statements made here literally true. An encyclopaedic reference is [18]. As we shall only use basic adjunctions and Costenoble–Waner duality, we suggest to simply look at [16, Appendix A] for a very readable introduction.

We denote by S_* and Sp the ∞-categories of pointed spaces and spectra, respectively. Likewise, we let $\mathsf{S}_{*/X} = \mathsf{Fun}(X, \mathsf{S}_*)$ and $\mathsf{Sp}_{/X} = \mathsf{Fun}(X, \mathsf{Sp})$ be the ∞-categories of parametrised pointed spaces/spectra over X. (In the definitions, X is seen as an ∞-groupoid.) We let $r\colon X \to *$ be the unique map to the point. We will use the following three standard functors:

the restriction functor:	$r^*\colon \mathsf{S}_* \to \mathsf{S}_{*/X}$,
its right adjoint:	$r_*\colon \mathsf{S}_{*/X} \to \mathsf{S}_*$,
its left adjoint:	$r_!\colon \mathsf{S}_{*/X} \to \mathsf{S}_*$.

The right and left adjoints are given by right and left Kan extensions, respectively. In other words, r_* takes the limit of a functor $F \in \mathsf{S}_{*/X} = \mathsf{Fun}(X, \mathsf{S}_*)$, whilst $r_!$ takes its colimit. We will also use the analogous functors in the case of parametrised spectra with the same notation. It will be clear from the context which one we are using. The crucial fact for us is that for any bundle $Y \to X$ equipped with a section s (this gives the data of a *pointed* space over X), $r_*(Y)$ is the path component of s in the section space.

As a last piece of notation, we will use $\Sigma^\infty_{/X} \dashv \Omega^\infty_{/X}$ to denote the infinite suspension/loop space adjunction between parametrised pointed spaces and spectra, and use $\Sigma^\infty \dashv \Omega^\infty$ to denote the usual adjunction in the unparametrised setting.

On our way to the proof of Theorem 8.11, we first make some formal observations. Loosely speaking, we would like to say that the section space of a fibrewise infinite loop space is the infinite loop space of the "section spectrum". This is made precise in the lemma below.

Lemma 8.12 *Let $Y \in \mathsf{S}_{*/X}$ be a parametrised space over X. We have a natural equivalence of pointed spaces*:

$$\Omega^\infty r_*(\Sigma^\infty_{/X} Y) \simeq r_*(\Omega^\infty_{/X} \Sigma^\infty_{/X} Y).$$

Proof We use the Yoneda lemma and the adjunction $r^* \dashv r_*$. Let $Z \in \mathsf{S}_*$ be a pointed space. We have

$$\begin{aligned} \mathrm{map}_{\mathsf{S}_*}(Z, \Omega^\infty r_*(\Sigma^\infty_{/X} Y)) &\simeq \mathrm{map}_{\mathsf{Sp}}(\Sigma^\infty, r_*(\Sigma^\infty_{/X} Y)) \simeq \mathrm{map}_{\mathsf{Sp}_{/X}}(r^*\Sigma^\infty Z, \Sigma^\infty_{/X} Y) \\ &\simeq \mathrm{map}_{\mathsf{Sp}_{/X}}(\Sigma^\infty_{/X} r^* Z, \Sigma^\infty_{/X} Y) \simeq \mathrm{map}_{\mathsf{S}_{*/X}}(r^* Z, \Omega^\infty_{/X} \Sigma^\infty_{/X} Y) \\ &\simeq \mathrm{map}_{\mathsf{S}_*}(Z, r_*(\Omega^\infty_{/X} \Sigma^\infty_{/X} Y)). \end{aligned}$$

Almost all manipulations follow from the standard adjunctions. The third equivalence uses the fact that $r^*\Sigma^\infty Z$ is the trivial parametrised spectrum with fibre $\Sigma^\infty Z$, and hence is equivalent to $\Sigma^\infty_{/X} r^* Z$. □

We will need two more facts before proving Theorem 8.11. The first one is the following simple observation. If $V \to X$ is a vector bundle such that its associated sphere bundle $\mathrm{Sph}(V) \to X$ has a section s, then we may take the fibrewise infinite suspension $\Sigma^\infty_{/X} \mathrm{Sph}(V) \in \mathsf{Sp}_{/X}$, using s to give a basepoint in each fibre. On the other hand, we could have taken the fibrewise one-point compactification and then suspended using the added point at infinity as a basepoint in each fibre. Up to a suspension, these are the same parametrised spectra.

Lemma 8.13 *Let $V \to X$ be a vector bundle with a nonvanishing section, and let $\mathrm{Sph}(V) \to X$ be its associated sphere bundle. Let $\mathbb{S}^V_X$ denote the fibrewise infinite suspension of the fibrewise one-point compactification of V (using the point at infinity as the basepoint in each fibre). Then*

$$\Sigma^\infty_{/X} \mathrm{Sph}(V) \simeq \Omega_X \mathbb{S}^V_X,$$

where Ω_X denotes the desuspension in the category $\mathsf{Sp}_{/X}$.

Proof Let us scale a nonvanishing section s of V so that it has image in the sphere bundle. We write $D(V) \subset V$ for the unit disc bundle of V, which we point using s, and V^+ for the fibrewise one-point compactification. We obtain the lemma by applying the fibrewise infinite suspension $\Sigma^\infty_{/X}$ to the cofibre sequence $\mathrm{Sph}(V) \to D(V) \to V^+$ of parametrised pointed spaces over X. □

Recall that the ∞-category $\mathsf{Sp}_{/X}$ is symmetric monoidal, with monoidal unit $\mathbb{S}_X := r^*(\mathbb{S})$. (Here and everywhere else $\mathbb{S}$ denotes the sphere spectrum.) The usefulness of the whole machinery set up so far is contained in the following result. A classical reference is [18, Chapter 18]. In the language of ∞-categories, one may read the second section of [16, Appendix A].

Lemma 8.14 (Costenoble–Waner duality) *The Costenoble–Waner dualising spectrum of X is $\mathbb{S}_X^{-TX}$, the spherical fibration associated to the stable normal bundle of X. That is, we have an equivalence of functors:*

$$r_*(-) \simeq r_!(- \otimes_{\mathbb{S}_X} \mathbb{S}_X^{-TX}).$$

Proof of Theorem 8.11 We start by choosing once and for all a section s of the sphere bundle $\mathrm{Sph}(J^1\mathcal{L}^d)$, which provides us with a basepoint in every fibre. We may therefore apply the free infinite loop space functor $Q = \Omega^\infty\Sigma^\infty \colon \mathsf{S}_* \to \mathsf{S}_*$ fibrewise and obtain the following diagram of fibrations:

$$\begin{tikzcd}
S^{2n+1} \arrow[r] \arrow[d] & \Omega^\infty\Sigma^\infty S^{2n+1} \arrow[d] \\
\mathrm{Sph}(J^1\mathcal{L}^d) \arrow[r] \arrow[d] & \Omega^\infty_{/X}\Sigma^\infty_{/X}\,\mathrm{Sph}(J^1\mathcal{L}^d) \arrow[d] \\
X \arrow[r, equal] & X
\end{tikzcd}$$

By the Freudenthal suspension theorem, the map $S^{2n+1} \to \Omega^\infty\Sigma^\infty S^{2n+1}$ is $(4n+1)$-connected. Using that X is homotopy equivalent to a $2n$-dimensional CW-complex, and that $\Gamma(-)$ sends homotopy pushouts to homotopy pullbacks, a direct induction on the skeletal filtration shows that the map on section spaces

$$\Gamma(\mathrm{Sph}(J^1\mathcal{L}^d)) \to \Gamma(\Omega^\infty_{/X}\Sigma^\infty_{/X}\,\mathrm{Sph}(J^1\mathcal{L}^d))$$

is $2n$-connected. (Notice that both spaces are connected, so the choice of s was immaterial.) Using Lemma 8.12, we obtain

$$\Gamma(\Omega^\infty_{/X}\Sigma^\infty_{/X}\,\mathrm{Sph}(J^1\mathcal{L}^d)) \simeq r_*(\Omega^\infty_{/X}\Sigma^\infty_{/X}\,\mathrm{Sph}(J^1\mathcal{L}^d)) \simeq \Omega^\infty r_*(\Sigma^\infty_{/X}\,\mathrm{Sph}(J^1\mathcal{L}^d)).$$

We now make the purely formal computation

$$\begin{aligned} r_*(\Sigma^\infty_{/X}\,\mathrm{Sph}(J^1\mathcal{L}^d)) &\simeq r_!(\Sigma^\infty_{/X}\,\mathrm{Sph}(J^1\mathcal{L}^d) \otimes_{\mathbb{S}_X} \mathbb{S}_X^{-TX}) \simeq r_!(\Omega_X \mathbb{S}_X^{J^1\mathcal{L}^d} \otimes_{\mathbb{S}_X} \mathbb{S}_X^{-TX}) \\ &\simeq r_!(\Omega_X \mathbb{S}_X^{J^1\mathcal{L}^d - TX}) \simeq \Omega r_!(\mathbb{S}_X^{J^1\mathcal{L}^d - TX}) \simeq \Omega X^{J^1\mathcal{L}^d - TX}, \end{aligned}$$

where we used Lemma 8.14 for the first equivalence, Lemma 8.13 for the second, and recognised that the value of $r_!$ on a spherical fibration is the associated Thom spectrum. □

8.3.2 An example when $X = \mathbb{CP}^2$ To show how Theorem 8.11 can be applied in practice, we use it to prove Proposition 8.10. We hope that this will convince the reader of the computational power of homotopy-theoretic methods to study spaces of algebraic sections.

Following Theorem 8.11, we should investigate $\Omega^{\infty+1} X^{J^1\mathcal{L}^d - TX}$ when $X = \mathbb{CP}^2$ and $\mathcal{L} = \mathcal{O}(1)$. Because $J^1\mathcal{L}^d - TX$ is of rank 2, the spectrum $\Omega X^{J^2\mathcal{L}^d - TX}$ is 1-connective and the bottom homotopy group is $\pi_1 \cong \mathbb{Z}$ by the Hurewicz theorem. We consider the fibration

$$F \to \Omega^{\infty+1} X^{J^1\mathcal{L}^d - TX} \to S^1,$$

where F is the homotopy fibre of the rightmost map, which is taken to induce an isomorphism on π_1. A generator of $\pi_1(\Omega^{\infty+1} X^{J^1\mathcal{L}^d - TX}) \cong \mathbb{Z}$ gives a section of that fibration, and we obtain

$$\Omega^{\infty+1} X^{J^1\mathcal{L}^d - TX} \simeq S^1 \times F.$$

In particular, F is simply connected with $\pi_2(F) \cong \pi_2(\Omega^{\infty+1} X^{J^1\mathcal{L}^d - TX})$. By the Hurewicz theorem and the universal coefficient theorem, $H_2(F; \mathbb{Z}/2) \cong H_2(F; \mathbb{Z}) \otimes \mathbb{Z}/2 \cong \pi_2(F) \otimes \mathbb{Z}/2$. We thus wish to compute $\pi_2(\Omega^{\infty+1} X^{J^1\mathcal{L}^d - TX})$, which we will do using the Adams spectral sequence at the prime 2:

$$E_2^{s,t} = \operatorname{Ext}_{\mathcal{A}}^{s,t}(H^*(X^{J^1\mathcal{L}^d - TX}; \mathbb{Z}/2), \mathbb{Z}/2) \Rightarrow \pi_*(X^{J^1\mathcal{L}^d - TX})^\wedge_2.$$

(Hence we will only compute the 2-completed group, but this will be enough for our purposes.) The E_2-page is computed by knowing the cohomology $H^*(X^{J^1\mathcal{L}^d - TX}; \mathbb{Z}/2)$ as an algebra over the mod 2 Steenrod algebra $\mathcal{A}$. (See [3, Section 3.3] for a very readable introduction.) If U denotes the Thom class of $J^1\mathcal{L}^d - TX$, the classes in the cohomology of the Thom spectrum $X^{J^1\mathcal{L}^d - TX}$ are given via the Thom isomorphism as yU where $y \in H^*(X; \mathbb{Z}/2)$. At the prime 2, the Steenrod squares can then be computed from the formula

$$\operatorname{Sq}^k(yU) = \sum_{i+j=k} \operatorname{Sq}^i(y)\operatorname{Sq}^j(U) = \sum_{i+j=k} \operatorname{Sq}^i(y) w_j U,$$

where w_j is the j^{th} Stiefel–Whitney class of $J^1\mathcal{L}^d - TX$. In our case, writing $\mathbb{Z}/2[x]/(x^3)$ for the cohomology ring of $X = \mathbb{CP}^2$, the total Stiefel–Whitney class is given by:

$$w(J^1\mathcal{L}^d - TX) = \begin{cases} 1 & d \equiv 0 \bmod 2, \\ 1 + x & d \equiv 1 \bmod 2. \end{cases}$$

We used the handy tool [9] to compute the E_2-page for us, and obtained Figure 3. From this, standard arguments about differentials (see eg [3, Section 4.8]) show that

$$\pi_3(X^{J^1\mathcal{L}^d - TX})^\wedge_2 \cong \begin{cases} \mathbb{Z}/2 & d \equiv 0 \bmod 2, \\ 0 & d \equiv 1 \bmod 2. \end{cases}$$

Therefore

$$H_2(F; \mathbb{Z}/2) \cong \pi_2(F) \otimes \mathbb{Z}/2 \cong \pi_3(X^{J^1\mathcal{L}^d - TX}) \otimes \mathbb{Z}/2 \cong \begin{cases} \mathbb{Z}/2 & d \equiv 0 \bmod 2, \\ 0 & d \equiv 1 \bmod 2. \end{cases}$$

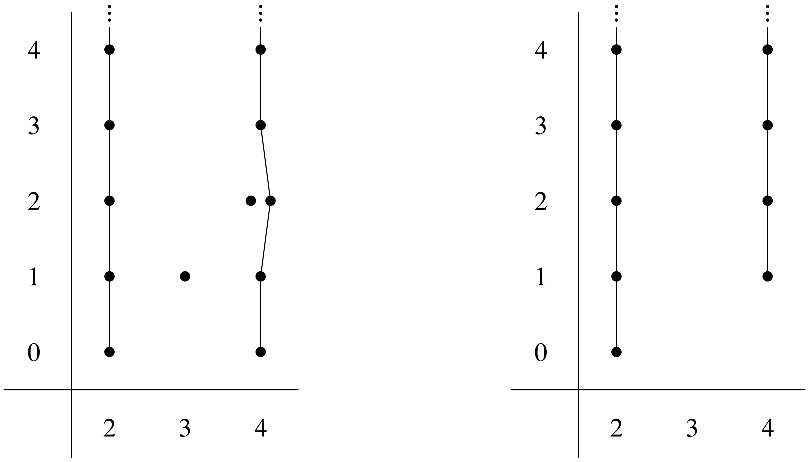


Figure 3: Left: $d \equiv 0 \mod 2$. Right: $d \equiv 1 \mod 2$. Following the established convention, we use the Adams grading: the horizontal axis is indexed by $t-s$, and the vertical one by s. Every dot represents a copy of $\mathbb{Z}/2$. The vertical lines represent multiplication by $h_0 \in \mathrm{Ext}^{1,1}_{\mathcal{A}}(\mathbb{Z}/2, \mathbb{Z}/2)$. We suggest to the unfamiliar reader to look at [3, Section 4.3] for more explanation.

Using the Künneth theorem, we obtain

$$H_2(\Omega^{\infty+1} X^{J^1\mathcal{L}^d - TX}; \mathbb{Z}/2) \cong H_2(S^1 \times F; \mathbb{Z}/2) \cong H_2(F; \mathbb{Z}/2),$$

which finishes the proof of Proposition 8.10.

8.4 Stability of p-torsion

In this final section, we study the p-torsion in the homology of the space $\Gamma_{\mathcal{C}^0}(\mathrm{Sph}(J^1\mathcal{L}^d))$. On the one hand, we have just seen in Proposition 8.10 that it depends on d in general. On the other hand, the result below shows that when the prime p is slightly bigger than the dimension of X, the p-torsion is independent of $\mathcal{L}$.

Proposition 8.15 *Let X be a smooth complex projective variety of complex dimension n and $\mathcal{L}$ be a holomorphic line bundle on it. Let p be a prime such that $p \geq n+2$. Then the fibrewise p-localisation of the sphere bundle $\mathrm{Sph}(J^1\mathcal{L}) \to X$ is trivial. In particular, we have an equivalence of p-local spaces*

$$\Gamma_{\mathcal{C}^0}(\mathrm{Sph}(J^1\mathcal{L}))_{(p)} \simeq \mathrm{map}(X, S^{2n+1}_{(p)}).$$

As an immediate consequence, combining the proposition above with Corollary 8.1 shows that the p-torsion in the homology of $\Gamma_{\mathrm{hol,ns}}(X; \mathcal{L}^d)$ stabilises when $p \geq \dim_{\mathbb{C}} X + 2$ and $d \to \infty$.

The proof uses the following result, which we learned from [5, Proposition 4.1]:

Lemma 8.16 *For $p \geq \frac{1}{2}k + \frac{3}{2}$, the space $\mathrm{map}_1(S^k_{(p)}, S^k_{(p)})$ of maps homotopic to the identity is $(k-1)$-connected.*

Proof The proof is given in [5], but we sketch it here for convenience. We shall assume that k is odd, as we will only use this case. Recall the evaluation fibration

$$\Omega^k_1 S^k_{(p)} \to \mathrm{map}_1(S^k_{(p)}, S^k_{(p)}) \to S^k_{(p)}.$$

Using the associated long exact sequence of homotopy groups, it suffices to show that $\pi_i(\Omega^k_1 S^k_{(p)})$ vanishes for $i \le k-1$. Using the assumption $p \ge \frac{1}{2}k + \frac{3}{2}$, this follows from Serre's calculations on p-torsion in the homotopy groups of spheres. □

Proof of Proposition 8.15 Let

$$S^{2n+1}_{(p)} \to \mathrm{Sph}(J^1\mathcal{L})_{(p)} \to X$$

be the fibrewise p-localisation of $\mathrm{Sph}(J^1\mathcal{L}) \to X$. By [19, Theorem 5.3], we have a homotopy equivalence

$$\Gamma_{\mathcal{C}^0}(\mathrm{Sph}(J^1\mathcal{L}))_{(p)} \simeq \Gamma_{\mathcal{C}^0}(\mathrm{Sph}(J^1\mathcal{L})_{(p)}).$$

As the sphere bundle is canonically oriented (using the complex orientation of $J^1\mathcal{L}$), the fibration $\mathrm{Sph}(J^1\mathcal{L})_{(p)} \to X$ is classified by a map

$$X \to B\,\mathrm{map}_1(S^{2n+1}_{(p)}, S^{2n+1}_{(p)}).$$

By Lemma 8.16, the codomain of that map is $(2n+1)$-connected. As the domain has real dimension $2n$, the classifying map must be nullhomotopic, thus showing that the fibration is trivial. □

References

[1] **I Banerjee**, *Stable cohomology of discriminant complements for an algebraic curve*, preprint (2020) arXiv 2010.14644 To appear in J. Topol. Anal.

[2] **O Banerjee**, *Filtration of cohomology via symmetric semisimplicial spaces*, Math. Z. 308 (2024) art. id. 16 MR Zbl

[3] **A Beaudry**, **J A Campbell**, *A guide for computing stable homotopy groups*, from "Topology and quantum theory in interaction" (D Ayala, D S Freed, R E Grady, editors), Contemp. Math. 718, Amer. Math. Soc., Providence, RI (2018) 89–136 MR Zbl

[4] **M C Beltrametti**, **S Di Rocco**, **A J Sommese**, *On generation of jets for vector bundles*, Rev. Mat. Complut. 12 (1999) 27–45 MR Zbl

[5] **M Bendersky**, **J Miller**, *Localization and homological stability of configuration spaces*, Q. J. Math. 65 (2014) 807–815 MR Zbl

[6] **M Bilu**, **S Howe**, *Motivic Euler products in motivic statistics*, Algebra Number Theory 15 (2021) 2195–2259 MR Zbl

[7] **J Bochnak**, **M Coste**, **M-F Roy**, *Real algebraic geometry*, Ergebnisse der Math. 36, Springer (1998) MR Zbl

[8] **G E Bredon**, *Sheaf theory*, 2nd edition, Graduate Texts in Math. 170, Springer (1997) MR Zbl

[9] **H Chatham**, **D Chua**, *The spectral sequences project*, GitHub repository (2021) Available at https://github.com/SpectralSequences/sseq

[10] **F R Cohen**, **T J Lada**, **J P May**, *The homology of iterated loop spaces*, Lecture Notes in Math. 533, Springer (1976) MR Zbl

[11] **R Das**, **S Howe**, *Cohomological and motivic inclusion-exclusion*, Compos. Math. 160 (2024) 2228–2283 MR Zbl

[12] **S Di Rocco**, **A J Sommese**, *Line bundles for which a projectivized jet bundle is a product*, Proc. Amer. Math. Soc. 129 (2001) 1659–1663 MR Zbl

[13] **Y Félix**, **S Halperin**, **J-C Thomas**, *Rational homotopy theory*, Graduate Texts in Math. 205, Springer (2001) MR Zbl

[14] **A Grothendieck**, *Éléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, IV*, Inst. Hautes Études Sci. Publ. Math. 32 (1967) 5–361 MR Zbl

[15] **A Haefliger**, *Rational homotopy of the space of sections of a nilpotent bundle*, Trans. Amer. Math. Soc. 273 (1982) 609–620 MR Zbl

[16] **M Land**, *Reducibility of low-dimensional Poincaré duality spaces*, Münster J. Math. 15 (2022) 47–81 MR Zbl

[17] **I Llerena**, *Localization of fibrations with nilpotent fibre*, Math. Z. 188 (1985) 397–410 MR Zbl

[18] **J P May**, **J Sigurdsson**, *Parametrized homotopy theory*, Math. Surv. Monogr. 132, Amer. Math. Soc., Providence, RI (2006) MR Zbl

[19] **J M Møller**, *Nilpotent spaces of sections*, Trans. Amer. Math. Soc. 303 (1987) 733–741 MR Zbl

[20] **J Mostovoy**, *Spaces of rational maps and the Stone–Weierstrass theorem*, Topology 45 (2006) 281–293 MR Zbl

[21] **J Mostovoy**, *Truncated simplicial resolutions and spaces of rational maps*, Q. J. Math. 63 (2012) 181–187 MR Zbl

[22] **J Neisendorfer**, *Algebraic methods in unstable homotopy theory*, New Math. Monogr. 12, Cambridge Univ. Press (2010) MR Zbl

[23] **C A M Peters**, **J H M Steenbrink**, *Mixed Hodge structures*, Ergebnisse der Math. (3) 52, Springer (2008) MR Zbl

[24] **B Poonen**, *Bertini theorems over finite fields*, Ann. of Math. 160 (2004) 1099–1127 MR Zbl

[25] **G Segal**, *The topology of spaces of rational functions*, Acta Math. 143 (1979) 39–72 MR Zbl

[26] **R Thom**, *L'homologie des espaces fonctionnels*, from "Colloque de topologie algébrique", Georges Thone, Liège (1957) 29–39 MR Zbl

[27] **O Tommasi**, *Stable cohomology of spaces of non-singular hypersurfaces*, Adv. Math. 265 (2014) 428–440 MR Zbl

[28] **R Vakil**, **M M Wood**, *Discriminants in the Grothendieck ring*, Duke Math. J. 164 (2015) 1139–1185 MR Zbl

[29] **V A Vassiliev**, *Complements of discriminants of smooth maps: topology and applications*, Transl. Math. Monogr. 98, Amer. Math. Soc., Providence, RI (1992) MR Zbl

[30] **V A Vassiliev**, *How to calculate the homology of spaces of nonsingular algebraic projective hypersurfaces*, Tr. Mat. Inst. Steklova 225 (1999) 132–152 MR Zbl In Russian; translated in Proc. Steklov Inst. Math. 225 (1999) 121–140

[31] **L Vokřínek**, *A generalization of Vassiliev's h-principle*, preprint (2007) arXiv 0705.0333

[32] **C A Weibel**, *An introduction to homological algebra*, Cambridge Stud. Adv. Math. 38, Cambridge Univ. Press (1994) MR Zbl

Department of Mathematical Sciences, University of Copenhagen
Copenhagen, Denmark

Current address: *Department of Mathematics, Stockholms universitet*
Stockholm, Sweden

aumonier.math@gmail.com

Proposed: Ulrike Tillmann Received: 11 July 2022
Seconded: Arend Bayer, Benson Farb Revised: 13 May 2024

Geometry & Topology **29**:3 (2025) 1489–1570
DOI: 10.2140/gt.2025.29.1489
Published: 31 May 2025

The motivic lambda algebra and motivic Hopf invariant one problem

WILLIAM BALDERRAMA
DOMINIC LEON CULVER
J D QUIGLEY

We investigate forms of the Hopf invariant one problem in motivic homotopy theory over arbitrary base fields of characteristic not equal to 2. Maps of Hopf invariant one classically arise from unital products on spheres, and one consequence of our work is a classification of motivic spheres represented by smooth schemes admitting a unital product.

The classical Hopf invariant one problem was resolved by Adams, following his introduction of the Adams spectral sequence. We introduce the motivic lambda algebra as a tool to carry out systematic computations in the motivic Adams spectral sequence. Using this, we compute the E_2-page of the $\mathbb{R}$-motivic Adams spectral sequence in filtrations $f \leq 3$. This universal case gives information over arbitrary base fields.

We then study the 1-line of the motivic Adams spectral sequence. We produce differentials $d_2(h_{a+1}) = (h_0 + \rho h_1)h_a^2$ over arbitrary base fields, which are motivic analogues of Adams' classical differentials. Unlike the classical case, the story does not end here, as the motivic 1-line is significantly richer than the classical 1-line. We determine all permanent cycles on the $\mathbb{R}$-motivic 1-line, and explicitly compute differentials in the universal cases of the prime fields $\mathbb{F}_q$ and $\mathbb{Q}$, as well as $\mathbb{Q}_p$ and $\mathbb{R}$.

55T15; 14F42, 55Q25, 55Q45, 55S10

1 Introduction

Motivic homotopy theory is a homotopy theory for algebraic varieties, developed by Morel and Voevodsky [1999]. Since its conception and subsequent use by Voevodsky [2003; 2011] to resolve the Milnor and Bloch–Kato conjectures, an immense amount of work has gone into the theory, with applications to algebraic geometry, algebraic number theory, and algebraic topology.

Motivic stable homotopy theory is the home of $\mathbb{A}^1$-invariants on algebraic varieties, such as algebraic K-theory, motivic cohomology, and algebraic cobordism. The universal such invariants are motivic stable homotopy groups, and as such the internal structure of the motivic stable homotopy groups of spheres reflects the broad-scale structure of the motivic stable homotopy category. These motivic stable stems encode deep geometric and number-theoretic information; for example, Morel [2004] showed that the Milnor–Witt K-theory of a field appears in its stable stems, and Röndigs, Spitzweck and Østvær [Röndigs et al. 2019; 2021] have identified motivic stable stems in low Milnor–Witt stem in terms of variants of Milnor K-theory, Hermitian K-theory, and motivic cohomology.

Motivic homotopy theory was originally developed to apply ideas and tools from homotopy theory to problems in algebraic geometry and algebraic K-theory. Information now flows the other way as well. After p-completion, $\mathbb{C}$-motivic stable stems capture information about classical stable stems that is not seen using classical techniques. This has led to the highly successful program of Gheorghe, Isaksen, Wang and Xu [Isaksen 2019; Isaksen et al. 2023; Gheorghe et al. 2021], yielding groundbreaking advances in computations of classical stable homotopy groups of spheres. A similar program using $\mathbb{R}$-motivic stable stems to capture information about C_2-equivariant stable stems has also developed [Burklund et al. 2020; Belmont and Isaksen 2022; Dugger and Isaksen 2017a; 2017b; Guillou and Isaksen 2020; Belmont et al. 2021]. More recently, Bachmann, Kong, Wang and Xu [Bachmann et al. 2022] related F-motivic stable homotopy theory over a general field F to classical complex cobordism.

All of this has motivated a swath of explicit computations of motivic stable stems over particular base fields F. We refer the reader to [Isaksen and Østvær 2020] for a general survey, but mention the following 2-primary computations:

- $F=\mathbb{C}$ Dugger and Isaksen [2010] computed the $\mathbb{C}$-motivic stable stems through the 36 stem, and these computations were pushed out to the 90 stem in [Isaksen 2019; Isaksen et al. 2023].
- $F=\mathbb{R}$ Dugger and Isaksen [2017a] computed the first four Milnor–Witt stems over $\mathbb{R}$, and Belmont and Isaksen [2022] expanded on this to compute the first 11 Milnor–Witt stems over $\mathbb{R}$.
- $F=\mathbb{F}_q$ Wilson [2016] and Wilson and Østvær [2017] computed the motivic stable homotopy groups of finite fields in motivic weight zero through topological dimension 18.

There are still many mysteries contained in the motivic stable stems. All of the above computations were enabled by the *motivic Adams spectral sequence*, originally introduced by Morel [1999] and further developed by Dugger and Isaksen [2010]. This is a motivic analogue of the classical Adams spectral sequence, which was developed by Adams [1958; 1960] to resolve the *Hopf invariant one problem*. Adams used this spectral sequence to prove that the only elements of Hopf invariant one in the classical stable stems π_*^{cl} are the classical Hopf maps $\eta_{\mathrm{cl}} \in \pi_1^{\mathrm{cl}}$, $\nu_{\mathrm{cl}} \in \pi_3^{\mathrm{cl}}$, and $\sigma_{\mathrm{cl}} \in \pi_7^{\mathrm{cl}}$. This theorem has a number of implications, including classifications of which spheres can be made into H-spaces, which spheres are parallelizable, which 2-dimensional modules over the Steenrod algebra can be realized by cell complexes, which dimensions a finite-dimensional real division algebra can have, and more.

This paper is concerned with topics surrounding motivic analogues of the classical Hopf invariant one problem. There is an element η in the motivic stable stems, represented by the canonical map $\eta\colon \mathbb{A}^2 \setminus \{0\} \to \mathbb{P}^1$, which refines the classical complex Hopf map η_{cl}. Hopkins and Morel — see [Morel 2004] — showed that η is one of the generators of the Milnor–Witt K-theory of the base field. This motivic η behaves quite differently from the classical Hopf map; most famously, η is not nilpotent, and is generally not 2-torsion. Because η is not nilpotent, one may consider the η-inverted stable stems $\pi_{*,*}[\eta^{-1}]$. These are closely related to Witt K-theory [Bachmann 2022; Bachmann and Hopkins 2020], and have been the subject of thorough investigation [Andrews and Miller 2017; Guillou and Isaksen 2015; 2016; Ormsby and Röndigs 2020; Wilson 2018].

Using the theory of Cayley–Dickson algebras, Dugger and Isaksen [2013] have shown that the classical quaternionic and octonionic Hopf maps ν_{cl} and σ_{cl} also admit geometric refinements to motivic classes ν and σ. All of these motivic Hopf maps η, ν, and σ are maps of Hopf invariant one, but, unlike classically, they are not the only such maps. For example, the classical stable stems include into the weight 0 portion of the motivic stable stems, and η_{cl}, ν_{cl}, and σ_{cl} give rise to distinct examples of maps of Hopf invariant one in the motivic setting. If we reformulate the condition of a map α having nontrivial Hopf invariant as asking that the homology of the 2-cell complex with attaching map α not split as a module over the motivic Steenrod algebra, then the situation becomes even richer: for example, σ_{cl}^2 admits an $\mathbb{R}$-motivic refinement to a map of nontrivial Hopf invariant in this sense, closely related to the nonexistent Hopf map coming next in the sequence η, ν, σ.

All of this motivates the present work, the purpose of which is three-fold:

(1) to analyze the motivic Hopf invariant one problem and deduce geometric consequences;

(2) to advance our understanding of motivic stable stems over general base fields;

(3) to introduce the *motivic lambda algebra*, a new tool for motivic computations.

As mentioned above, Adams resolved the Hopf invariant one problem by introducing and studying the Adams spectral sequence. Morel [1999] and Dugger and Isaksen [2010] have already introduced the *F-motivic Adams spectral sequence*, which takes the form

$$E_2^{*,*,*} = \mathrm{Ext}_{\mathcal{A}^F}^{*,*,*}(\mathbb{M}^F, \mathbb{M}^F) \Rightarrow \pi_{*,*}^F.$$

Here $\mathcal{A}^F$ is the F-motivic Steenrod algebra [Voevodsky 2003; Hoyois et al. 2017], which acts on $\mathbb{M}^F$, the mod 2 motivic cohomology of $\mathrm{Spec}(F)$. This spectral sequence converges to $\pi^F_{*,*}$, the homotopy groups of the $(2,\eta)$-completed F-motivic sphere [Hu et al. 2011a; Kylling and Wilson 2019]. Implicit is the assumption that 2 is invertible in F.

In this paper, we bring the motivic Adams spectral sequence back to its classical roots, using it to study the motivic Hopf invariant one problem. We do not follow Adams' original approach. Instead, at least in broad outline, we follow J S P Wang's approach [1967], which proceeded by first gaining a good understanding of the E_2-page of the Adams spectral sequence. Importing this approach to motivic homotopy theory requires analyzing the E_2-page of the motivic Adams spectral sequence over general base fields in ranges beyond what is known by previous techniques.

To carry out this analysis, we bring another tool from classical stable homotopy theory into the motivic context: the *lambda algebra*. The classical lambda algebra Λ^{cl} is a certain differential graded algebra, originally constructed by Bousfield, Curtis, Kan, Quillen, Rector and Schlesinger [Bousfield et al. 1966], whose homology recovers the E_2-page of the Adams spectral sequence. The classical lambda algebra is now a standard member of the homotopy theorist's toolbox, and we cannot hope to list all of its applications, but the following are a selection:

(1) Wang's computation [1967] of the E_2-page of the Adams spectral sequence through the 3-line, and subsequent simplified resolution of the Hopf invariant one problem;

(2) some of the first automated computations of the E_2-page of the Adams spectral sequence, including products and Massey products [Tangora 1985; 1993; 1994; Curtis et al. 1987];

(3) the construction of Brown–Gitler spectra [1973], which played an important role in analyzing the bo-resolution [Mahowald 1981; Shimamoto 1984], the proof of the immersion conjecture [Cohen 1985], and more [Mahowald 1977; Goerss 1999; Hunter and Kuhn 1999];

(4) the algebraic Atiyah–Hirzebruch spectral sequence for $\mathbb{R}P^\infty$ [Wang and Xu 2016], used as input to their proof of the nonexistence of exotic smooth structures on the 61-sphere [Wang and Xu 2017];

(5) the only complete computations of the 4- and 5-lines of the Adams E_2-term [Chen 2011; Lin 2008].

We expect that the motivic lambda algebra will likewise become a useful member of the motivic homotopy theorist's toolbox. We focus in particular on developing the lambda algebra and applying this to the motivic Hopf invariant one problem. We consider both the unstable problem, with applications to H-space structures on motivic spheres, and the stable problem, which is concerned with the 1-line of the motivic Adams spectral sequence. The motivic situation is substantially richer than the classical situation, and requires us to develop a number of new techniques for motivic computations across general base fields.

Adams' resolution of the classical Hopf invariant one problem asserted the existence of differentials $d_2(h_{a+1}) = h_0 h_a^2$ in the Adams spectral sequence. There are classes h_a in the F-motivic Adams spectral sequence for any field F, corresponding to the motivic Hopf maps discussed above for $a \le 3$. Using

Betti realization, it is possible to lift Adams' differentials to the $\mathbb{C}$-motivic Adams spectral sequence. It follows that, if F admits a complex embedding, then h_{a+1} must support a nontrivial differential for $a \geq 3$. However, this is insufficient to determine the precise target of the differential, as well as to determine what happens over other base fields, particularly fields of positive characteristic. The techniques we develop are geared towards resolving this sort of issue. We use these to obtain a number of new results; let us give the following here, as it is the most pleasant to state.

Theorem A (Theorem 7.3.1) *For an arbitrary base field F of characteristic not equal to* 2, *there are differentials of the form*

$$d_2(h_{a+1}) = (h_0 + \rho h_1)h_a^2$$

in the F-motivic Adams spectral sequence, *which are nonzero for $a \geq 3$.*

It is worth making a couple remarks to distinguish this from the classical result.

Remark 1.0.1 Classically, there is at most one possible nontrivial target for a d_2-differential on h_{a+1}. As suggested by the target in Theorem A, the motivic situation is more complicated. For example, when $F = \mathbb{R}$, we show that, if $a \geq 4$, then the group of potential values of $d_2(h_{a+1})$ is given by $\mathbb{F}_2\{h_0 h_a^2, \rho h_1 h_a^2\}$. The general picture is similar, except there may be additional interference coming from the mod 2 Milnor K-theory of F. This computation requires new techniques for computing the cohomology of the motivic Steenrod algebra, which is much richer than the analogous classical computation. ◁

Remark 1.0.2 Even once we have carried out the algebraic work of identifying potential values of $d_2(h_{a+1})$, the classical proof does not directly generalize to yield Theorem A. In spirit, our proof follows Wang's classical inductive proof [1967]. The base case of Wang's induction is the differential $d_2(h_4) = h_0 h_3^2$, which follows easily from graded commutativity of stable stems. By contrast, our base case must include the differential $d_2(h_5) = (h_0 + \rho h_1)h_4^2$. Over $\mathbb{R}$, this differential may be deduced by combining complex and real Betti realization, but a completely different argument is required to obtain the differential for other fields. To obtain this differential over other base fields, we use a certain motivic Hasse principle to reduce to considering fields with simple mod 2 Milnor K-theory, then analyze how the classical Kervaire class θ_4 appears in the motivic stable stems. ◁

Remark 1.0.3 There is another elegant proof of the classical Adams differential $d_2(h_{a+1}) = h_0^2 h_a$, due to Bruner [1986b, Corollary 1.5], which makes use of power operations in the Adams spectral sequence. Tilson [2017] has explored analogues of Bruner's results in the $\mathbb{R}$-motivic setting, but so far these methods have only succeeded in determining the $\mathbb{R}$-motivic differential $d_2(h_{a+1})$ for $a \leq 3$. ◁

1.1 Brief overview

Now let us give a very brief overview of what we do in this paper, before giving a more thorough summary in Section 1.2. This paper has three main parts. These parts are not independent, but none rely on the hardest aspects of the others.

The first part is purely algebraic, and is the most computationally intensive. In Section 2, we introduce the F*-motivic lambda algebra* (Theorem B), and in Section 4 we use the $\mathbb{R}$-motivic lambda algebra to compute $\mathrm{Ext}_{\mathbb{R}}$ in filtrations $f \leq 3$ (Theorem C). The result is quite complicated, with eight infinite families of multiplicative generators and numerous relations between these. As we explain in Section 7.1, this gives information about Ext_F for any base field F once the mod 2 Milnor K-theory of F is known.

The second part is shorter, and does not rely on the above computation. In Section 6, after some preliminaries in Section 5, we consider the motivic analogue of the Hopf invariant one problem in its classical *unstable* formulation, concerning unstable 2-cell complexes with specified cup product, as well as concerning geometric applications such as to H-space structures on motivic spheres. Our analysis proceeds by a novel reduction to the classical case and other known results, by first formulating a certain motivic Lefschetz principle (Proposition 5.2.1), then using this to build unstable "Betti realization" functors over arbitrary algebraically closed fields (Proposition 5.3.2). One consequence of this analysis is a complete classification of motivic spheres which are represented by smooth schemes admitting a unital product (Theorem D).

The third part is our main homotopical contribution. In Section 7, we give a detailed study of the 1-line of the F-motivic Adams spectral sequence. This work has a direct geometric interpretation: permanent cycles on the 1-line of the motivic Adams spectral sequence classify how the motivic Steenrod algebra can act on the cohomology of a motivic 2-cell complex. This section does not rely on the full strength of our computation of $\mathrm{Ext}_{\mathbb{R}}$, and should be understandable by the reader familiar with prior work on the $\mathbb{R}$-motivic Adams spectral sequence. The main theorems in this section are Theorem A above, together with much more detailed information about the 1-line of the F-motivic Adams spectral sequence for the particular fields $F = \mathbb{R}$, $F = \mathbb{F}_q$ with q an odd prime power, $F = \mathbb{Q}_p$ with p any prime, and $F = \mathbb{Q}$ (Theorem E). As this includes all the prime fields, these computations give information that applies to an arbitrary base field. When $F = \mathbb{R}$, we completely determine all permanent cycles on the 1-line by comparison with a computation in Borel C_2-equivariant homotopy theory (Theorem F); both the equivariant computation and the method of comparison are of independent interest.

1.2 Summary of results

We now summarize our work in more detail. We begin with our introduction of the motivic lambda algebra. The nature of the classical lambda algebra Λ^{cl} [Bousfield et al. 1966] was greatly clarified by Priddy [1970], who introduced the notion of a *Koszul algebra* and showed that Λ^{cl} is the Koszul complex of the classical Steenrod algebra. We carry out the motivic analogue of this, producing the following.

Theorem B (Section 2.4) *There is a differential graded algebra* Λ^F, *the* F*-motivic lambda algebra*, *with the following properties*:

(1) Λ^F *may be described explicitly in terms of generators*, *relations*, *and monomial basis.*

(2) *There is a surjective and multiplicative quasiisomorphism $C(\mathcal{A}^F) \to \Lambda^F$ from the cobar complex of the F-motivic Steenrod algebra to Λ^F. In particular, there is an isomorphism*

$$H_*\Lambda^F \cong \mathrm{Ext}^*_F$$

compatible with all products and Massey products. Moreover, the squaring operation $\mathrm{Sq}^0\colon \mathrm{Ext}^*_F \to \mathrm{Ext}^*_F$ lifts to a map $\theta\colon \Lambda^F \to \Lambda^F$ of differential graded algebras.

(3) *Λ^F generalizes the classical lambda algebra, in the sense that, if F is algebraically closed, then $\Lambda^F[\tau^{-1}] = \Lambda^{\mathrm{cl}}[\tau^{\pm 1}]$. In particular, it is considerably smaller than $C(\mathcal{A}^F)$.*

Here we have abbreviated $\mathrm{Ext}^{*,*,*}_{\mathcal{A}^F}(\mathbb{M}^F, \mathbb{M}^F)$ to Ext^*_F, where the single index refers to filtration, or homological degree, ie $\mathrm{Ext}^f_F = H^f(\mathcal{A}^F)$.

Remark 1.2.1 Several subtleties arise in the construction and identification of the motivic lambda algebra. We note two interesting points here:

(1) Priddy's notion [1970] of Koszul algebra is not general enough for our situation: $\mathcal{A}^F$ is generally not augmented as an $\mathbb{M}^F$-algebra, and $\mathbb{M}^F$ is generally not central in $\mathcal{A}^F$. This forces us to consider a more general notion of a Koszul algebra, as well as to find new arguments to prove that $\mathcal{A}^F$ is Koszul in this more general sense.

(2) As readers familiar with the motivic Adem relations might suspect, the elements τ and ρ of $\mathbb{M}^F$ appear in the relations defining the motivic lambda algebra, as well as in its differential and the endomorphism θ lifting Sq^0. Determining these formulas precisely is delicate and requires some careful arguments. $\triangleleft$

Remark 1.2.2 As indicated above, we construct the F-motivic lambda algebra as a certain Koszul complex for the F-motivic Steenrod algebra. The Koszul story produces other complexes as well: for any $\mathcal{A}^F$-modules M and N with M projective over $\mathbb{M}^F$, there are complexes $\Lambda^F(M, N)$ serving as small models of the cobar complex computing $\mathrm{Ext}_{\mathcal{A}^F}(M, N)$. An amusing special case of this produces a lambda algebra Λ^{C_2} for the C_2-equivariant Steenrod algebra (Remark 2.3.5). $\triangleleft$

We use the motivic lambda algebra to study Ext_F in low filtration. Before diving into our more extensive computations, we illustrate the structure of Λ^F with some simple examples in Section 3.1, showing how it may be used to give easy rederivations of some well-known low-dimensional relations in Ext_F. We then carry out our main algebraic computation in Section 4, where we prove the following. Note that $\mathrm{Ext}^0_{\mathbb{R}} = \mathbb{F}_2[\rho]$.

Theorem C *The structure of $\mathrm{Ext}_{\mathbb{R}}$ in filtrations $f \leq 3$ is as described in Section 4; in particular, the $\mathbb{F}_2[\rho]$-module structure is described in Theorem 4.2.12, including a description of multiplicative generators and the action of Sq^0, and the majority of the multiplicative structure is described in Theorem 4.3.7.*

Here we are justified in focusing on $\mathrm{Ext}_{\mathbb{R}}$ as it is, in a certain precise sense, the universal case (see Remark 2.2.8). We explain in Section 7.1 how to pass from information about $\mathrm{Ext}_{\mathbb{R}}$ to information about Ext_F for other base fields F.

Example 1.2.3 (Theorem 4.2.12(1)) The computation of $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$ is much more involved than the corresponding classical computation, and the result is much richer. We refer the reader to Section 4 for the full statements, but illustrate this here with the following sample. Classically, $\mathrm{Ext}_{\mathrm{cl}}^{\leq 3}$ is generated as an algebra by the classes h_a and c_a for $a \geq 0$. By contrast, a minimal multiplicative generating set of $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$ as an $\mathbb{F}_2[\rho]$-algebra is given by the classes in the following table:

multiplicative generator	ρ-torsion exponent
h_{a+1}	∞
c_{a+1}	∞
$\tau^{\lfloor 2^{a-1}(4n+1)\rfloor} h_a$	2^a
$\tau^{2^a(8n+1)} h_{a+2}^2$	$2^{a+1}\cdot 3$
$\tau^{\lfloor 2^{a-1}(2(16n+1)+1)\rfloor} h_{a+3}^2 h_a$	$2^a \cdot 13$
$\tau^{2^a(4(4n+1)+1)} h_{a+3}^3$	$2^a\cdot 7$
$\tau^{\lfloor 2^{a-1}(16n+1)\rfloor} c_a$	$2^a\cdot 7$
$\tau^{2^{a+1}(8n+1)} c_{a+1}$	$2^{a+2}\cdot 3$
$\tau^{\lfloor 2^{a-1}(2(4n+1)+1)\rfloor} c_a$	$2^a\cdot 3$

Here $a,n \geq 0$, and the ρ-torsion exponent of a class α is the minimal r for which $\rho^r\alpha = 0$; the classes h_{a+1} and c_{a+1} are ρ-torsion-free. Note that all of the classes listed are named for their image in $\mathrm{Ext}_{\mathbb{C}}$, and are not themselves products. $\triangleleft$

Example 1.2.4 Observe that the multiplicative generators h_a and c_a of $\mathrm{Ext}_{\mathrm{cl}}^{\leq 3}$ appear, with a shift, as ρ-torsion-free classes in $\mathrm{Ext}_{\mathbb{R}}$. This is a general phenomenon: Dugger and Isaksen [2017a, Theorem 4.1] produce an isomorphism $\mathrm{Ext}_{\mathbb{R}}[\rho^{-1}] \simeq \mathrm{Ext}_{\mathrm{dcl}}[\rho^{\pm 1}]$; here $\mathrm{Ext}_{\mathrm{dcl}} = \mathrm{Ext}_{\mathrm{cl}}$ only given a motivic grading such that $\mathrm{Ext}_{\mathrm{cl}}^{s,f} = \mathrm{Ext}_{\mathrm{dcl}}^{2s+f,f,s+f}$. As we discuss in Section 3.2, this in fact refines to a splitting $\mathrm{Ext}_{\mathbb{R}} \cong \mathrm{Ext}_{\mathrm{dcl}}[\rho] \oplus \mathrm{Ext}_{\mathbb{R}}^{\rho\text{-tors}}$, where $\mathrm{Ext}_{\mathbb{R}}^{\rho\text{-tors}} \subset \mathrm{Ext}_{\mathbb{R}}$ is the subgroup of ρ-torsion; moreover, this splitting is modeled by a multiplicatively split inclusion $\widetilde{\theta}\colon \Lambda^{\mathrm{dcl}} \to \Lambda^{\mathbb{R}}$. The general shape of $\mathrm{Ext}_{\mathbb{R}}$ forced by this may be illustrated by the description of the 1-line

$$\mathrm{Ext}_{\mathbb{R}}^1 = \mathbb{F}_2[\rho]\{h_a : a \geq 1\} \oplus \bigoplus_{a\geq 0} \mathbb{F}_2[\rho]/(\rho^{2^a})\{\tau^{\lfloor 2^{a-1}(4n+1)\rfloor} h_a : n \geq 0\}. \tag{1-1}$$ $\triangleleft$

As $\mathrm{Ext}_{\mathrm{cl}}^{\leq 3}$ is entirely understood by Wang's computation [1967], the hard work of Theorem C is in computing the ρ-torsion subgroup of $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$. This is the most computationally intensive part of the paper, and proceeds by a direct case analysis of monomials in $\Lambda^{\mathbb{R}}$ in low filtration. In the end, we find that $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$ carries the multiplicative generators listed in Example 1.2.3, and that there are many exotic relations between these generators. Our computation describes all of this.

With the algebraic computation of Theorem C in place, we turn to more homotopical topics, namely those surrounding the *Hopf invariant one problem*. There are (at least) *two* good motivic analogues of the Hopf invariant one problem: one which is unstable, concerning the construction of unstable 2-cell complexes with nontrivial cup product structure, and one which is stable, concerning the construction of stable 2-cell complexes with nontrivial $\mathcal{A}^F$-module structure. As we recall in Section 6.2, understanding the latter question is equivalent to understanding the 1-line of the F-motivic Adams spectral sequence; we get to this in Section 7, which we will discuss further below.

It is the former unstable formulation which has more direct geometric applications. For example, following [Dugger and Isaksen 2013] on the Hopf construction in motivic homotopy theory, it is directly tied up with the question of which unstable motivic spheres $S^{a,b}$ admit H-space structures (see Lemma 6.4.3). Here $S^{a,b}$ is the motivic sphere which is $\mathbb{A}^1$-homotopy equivalent to $\Sigma^{a-b}\mathbb{G}_m^{\wedge b}$. We discuss this unstable formulation in Section 6, which is independent of our other calculations. One pleasant consequence of this story is the following.

Theorem D (Theorem 6.4.5) *The only motivic spheres which are represented by smooth F-schemes admitting a unital product are $S^{0,0}$, $S^{1,1}$, $S^{3,2}$, and $S^{7,4}$.*

The statement of Theorem D is directly analogous to the classical result that the only spheres admitting unital products are S^0, S^1, S^3, and S^7. Classically, the nonexistence of H-space structures on any other spheres may be reduced to the Hopf invariant one problem, which was then established by Adams. This reduction makes use of the instability condition that $\mathrm{Sq}^a(x) = x^2$ whenever $x \in H^a(X)$ for some space X. There is an analogous instability condition for the motivic cohomology of a motivic space, but it holds only in a smaller range than we would need; as a consequence, some additional input is needed to analyze the unstable motivic Hopf invariant one problem (see Remark 6.3.2).

This additional input is interesting in itself. It follows from the formulation of the unstable motivic Hopf invariant one problem that, at least for nonexistence, one may reduce to the case where F is algebraically closed. In Section 5.2, we explain how work of Wilson and Østvær [2017] implies a certain *Lefschetz principle* for suitable 2-primary categories of cellular motivic spectra. When combined with Mandell's p-adic homotopy theory [2001], this gives a 2-primary unstable "Betti realization" functor for any algebraically closed field F, which is well behaved with respect to the mod 2 cohomology of motivic cell complexes; see Section 5.3. This gives a direct relation between motivic and classical homotopy theory, and we are then able to analyze the unstable motivic Hopf invariant one problem using a combination of classical results, work of Dugger and Isaksen [2013] on the motivic Hopf construction, and work of Asok, Doran and Fasel [Asok et al. 2017] on smooth models of motivic spheres.

Finally, in Section 7, we turn to a study of the 1-line of the F-motivic Adams spectral sequence. After a few preliminaries, we begin by proving Theorem A, producing the differentials

$$d_2(h_{a+1}) = (h_0 + \rho h_1)h_a^2$$

valid for any F (Theorem 7.3.1). As we mentioned above, the main content of this theorem is not the fact that the classes h_{a+1} for $a \geq 3$ support nonzero differentials, but the exact value of the target of these differentials. We mention two interesting aspects of this computation here:

First, in order to get a more explicit handle on possible targets of $d_2(h_{a+1})$, we reduce to considering the case where F is a prime field, ie $F = \mathbb{F}_p$ with p odd or $F = \mathbb{Q}$. The latter case is then handled with the aid of a *Hasse principle*. We explain how work of Ormsby and Østvær [2013] on the structure of $\mathbb{M}^{\mathbb{Q}}$ may be used to give a concrete description of $\mathrm{Ext}_{\mathbb{Q}}$ and of the Hasse map

$$\mathrm{Ext}_{\mathbb{Q}} \to \mathrm{Ext}_{\mathbb{R}} \times \prod_{p \text{ prime}} \mathrm{Ext}_{\mathbb{Q}_p}, \tag{1-2}$$

in particular proving this map is injective (Proposition 7.1.3). In this way we reduce to computing the differentials $d_2(h_n)$ over the fields $\mathbb{F}_p$ with p odd, $\mathbb{Q}_p$ with p prime, and $\mathbb{R}$.

Second, the classical argument, using the fact that $2\sigma^2 = 0$, may be used to compute $d_2(h_4)$, but a new argument is required to produce the differential $d_2(h_5) = (h_0 + \rho h_1)h_4^2$ (Proposition 7.3.3). Once this differential is resolved, the rest follow by an inductive argument analogous to Wang's classical argument [1967]. After a further reduction when $F = \mathbb{R}$, the differential $d_2(h_5)$ may be resolved uniformly in the above choices of base field. In short, to resolve this differential, we lift the Hurewicz map $\pi_*^{\mathrm{cl}} \to \pi_{*,0}^F$ to a map $\mathrm{Ext}_{\mathrm{cl}}^{*,*} \to \mathrm{Ext}_F^{*,*,0}$ of spectral sequences (Proposition 5.1.1) and, by considering the effect of this on the Kervaire class θ_4, deduce that $(h_0 + \rho h_1)h_4^2$ must be hit by h_5.

The story does not stop with the differentials $d_2(h_{a+1})$, as Ext_F^1 contains many more classes than these; recall for instance $\mathrm{Ext}_{\mathbb{R}}^1$ from (1-1). Having resolved these differentials, we move on to giving an explicit analysis of the 1-line of the F-motivic Adams spectral sequence for a number of base fields F. Our main results may be summarized in the following.

Theorem E *The following are carried out in Section 7:*

1. *In Theorem 7.4.9, we compute all d_2-differentials out of $\mathrm{Ext}_{\mathbb{R}}^1$, as well as all permanent cycles in $\mathrm{Ext}_{\mathbb{R}}^1$.*
2. *In Theorem 7.5.3, for q a prime power satisfying $q \equiv 1 \pmod 4$, we compute all Adams differentials out of $\mathrm{Ext}_{\mathbb{F}_q}^1$, in particular giving all permanent cycles in $\mathrm{Ext}_{\mathbb{F}_q}^1$.*
3. *In Theorem 7.5.6, for q a prime power satisfying $q \equiv 3 \pmod 4$, we compute all d_2-differentials out of $\mathrm{Ext}_{\mathbb{F}_q}^1$, as well as all higher differentials in stems $s \leq 7$, in particular giving all permanent cycles in $\mathrm{Ext}_{\mathbb{F}_q}^1$ in stems $s \leq 7$.*
4. *In Theorem 7.6.2, for p an odd prime, we give as much information about $\mathrm{Ext}_{\mathbb{Q}_p}^1$ as was given for $\mathrm{Ext}_{\mathbb{F}_p}^1$.*
5. *In Theorem 7.6.6, we compute all d_2-differentials out of $\mathrm{Ext}_{\mathbb{Q}_2}^1$, as well as all higher differentials in stems $s \leq 7$, in particular giving all permanent cycles in $\mathrm{Ext}_{\mathbb{Q}_2}^1$ in stems $s \leq 7$.*
6. *In Theorem 7.7.1, we give the same information for $\mathrm{Ext}_{\mathbb{Q}}^1$ as was given for $\mathrm{Ext}_{\mathbb{Q}_2}^1$.*

Cases (2)–(6) of Theorem E proceed by a direct analysis, combining the Hopf differentials we produced in Theorem A with arithmetic differentials that may be obtained by comparison with the F-motivic Adams spectral sequence for integral motivic cohomology. The latter has been computed by Kylling [2015] for $F = \mathbb{F}_q$ with q an odd prime power, by Ormsby [2011] for $F = \mathbb{Q}_p$ with p an odd prime, and by Ormsby and Østvær [2013] for $F = \mathbb{Q}_2$ and $F = \mathbb{Q}$. Case (6), where $F = \mathbb{Q}$, may be read off the cases $F = \mathbb{R}$ and $F = \mathbb{Q}_p$, using our good understanding of the Hasse map (1-2). As with Ormsby and Østvær's computations over $\mathbb{Q}$, the final description of the set of d_2-cycles in $\mathrm{Ext}^1_{\mathbb{Q}}$ is quite intricate, but we feel that our techniques show that understanding the $\mathbb{Q}$-motivic Adams spectral sequence for $\pi^{\mathbb{Q}}_{*,*}$ is an accessible problem ripe for future investigation.

The $\mathbb{R}$-motivic computation requires more work. Recall the structure of $\mathrm{Ext}^1_{\mathbb{R}}$ from (1-1). Theorem A describes what happens on the ρ-torsion-free summand of this, but says nothing about the large quantity of ρ-torsion classes. It is possible to use similar methods to compute all d_2-differentials supported on this ρ-torsion summand, and we do so in Proposition 7.4.8. However, this is insufficient to determine which classes in $\mathrm{Ext}^1_{\mathbb{R}}$ are permanent cycles, as higher differentials may, and indeed must, occur.

We resolve this by comparison with *Borel C_2-equivariant homotopy theory*. Behrens and Shah [2020] formulate and prove an equivalence

$$(\mathcal{S}\mathrm{p}^{\mathrm{cell}}_{\mathbb{R}})^{\wedge}_{(2,\rho)}[\tau^{-1}] \simeq \mathrm{Fun}(BC_2, \mathcal{S}\mathrm{p}^{\wedge}_2)$$

between the τ-periodic $(2,\rho)$-complete cellular $\mathbb{R}$-motivic category and the 2-complete Borel C_2-equivariant category. Define

$$\mathrm{Ext}^{s,f,w}_{BC_2} = \mathrm{Ext}^{s-w,f}_{\mathcal{A}^{\mathrm{cl}}}(\mathbb{F}_2, H^* P^{\infty}_w),$$

where P^{∞}_w is a stunted real projective space. These form the E_2-pages of the classical Adams spectral sequences for the stable cohomotopy groups of infinite stunted projective spaces. The equivalence of Behrens and Shah gives an effective method of computing these groups by "inverting τ" in $\mathrm{Ext}_{\mathbb{R}}$. The τ-periodic behavior of $\mathrm{Ext}_{\mathbb{R}}$ is plainly visible in our computation of $\mathrm{Ext}^{\leq 3}_{\mathbb{R}}$, allowing us to directly read off the structure of $\mathrm{Ext}^{\leq 3}_{BC_2}$ (Lemma 7.4.3). In particular,

$$\mathrm{Ext}^1_{BC_2} = \mathbb{F}_2[\rho]\{h_a : a \geq 1\} \oplus \bigoplus_{a \geq 0} \mathbb{F}_2[\rho]/(\rho^{2^a})\{\tau^{\lfloor 2^{a-1}(4n+1)\rfloor} h_a : n \in \mathbb{Z}\}$$

(compare (1-1)). We warn the reader that this naming of classes is incompatible with viewing Ext_{BC_2} as a collection of ordinary Adams spectral sequences; for example, h_0 does not detect 2, but instead the transfer $P^{\infty}_0 \to S^0$. We may use the relatively simple structure of these 1-lines to verify that $\mathrm{Ext}^1_{\mathbb{R}} \to \mathrm{Ext}^1_{BC_2}$ reflects permanent cycles (Lemma 7.4.4), and this reduces the identification of permanent cycles in $\mathrm{Ext}^1_{\mathbb{R}}$ to the identification of permanent cycles in $\mathrm{Ext}^1_{BC_2}$. The problem of ρ-torsion permanent cycles in $\mathrm{Ext}^1_{BC_2}$ turns out to be equivalent to the vector fields on spheres problem (Lemma 7.4.5), which was resolved by Adams [1962]. Together with known information regarding the ρ-torsion-free classes, this leads to the following classification of maps $\Sigma^c P^{\infty}_w \to S^0$ detected in Adams filtration 1.

Theorem F (Theorem 7.4.7) *For $a \geq 0$, write $a = c + 4d$ with $0 \leq c \leq 3$, and define $\psi(a) = 2^c + 8d$. Then the subgroup of permanent cycles in $\mathrm{Ext}^1_{BC_2}$ is given by*

$$\mathbb{F}_2[\rho]\{h_1, h_2, h_3, \rho h_4\} \oplus \bigoplus_{a \geq 0} \mathbb{F}_2[\rho]/(\rho^{\psi(a)})\{\rho^{2^a - \psi(a)} \tau^{\lfloor 2^{a-1}(4n+1) \rfloor} h_a : n \in \mathbb{Z}\}.$$

Moreover, one may characterize maps $\Sigma^c P^\infty_w \to S^0$ detected by each of these classes.

1.3 Future directions

The classical lambda algebra has been applied broadly in stable homotopy theory. This suggests several natural directions for future work, and we list a few here.

1.3.1 Homological computations The homology of the classical lambda algebra can be computed algorithmically via a method known as the *Curtis algorithm*. This procedure was refined and implemented by Tangora [1985] to compute the cohomology of the Steenrod algebra through internal degree 56, as well as to compute products and Massey products [Tangora 1993; 1994]; further computations of Curtis, Goerss, Mahowald and Milgram [Curtis et al. 1987] pushed this out to describe the cohomology of the Steenrod algebra through stem 51. More recently, the Curtis algorithm was used by Wang and Xu [2016] to compute the algebraic Atiyah–Hirzebruch spectral sequence for $\mathbb{R}P^\infty$, providing the data necessary for their proof of the uniqueness of the smooth structure on the 61-sphere [Wang and Xu 2017].

Our method for computing $\mathrm{Ext}^{\leq 3}_{\mathbb{R}}$ is closely related to the homology algorithm of [Tangora 1985], only modified to take into account the $\mathbb{F}_2[\rho]$-module structure of $\Lambda^{\mathbb{R}}$, as well to incorporate some additional flexibility in choosing representatives for the sake of a more digestible manual computation. By ignoring this additional flexibility and incorporating the ideas of [loc. cit., Section 3.4], one obtains a Curtis algorithm for computing the homology of the $\mathbb{R}$-motivic lambda algebra, as well as of other motivic lambda complexes. The effectiveness of these procedures in higher dimensions remains to be seen.

In addition to its use in computer-assisted computations, the classical lambda algebra has also been used in [Lin 2008; Chen 2011] to completely compute the cohomology of the classical Steenrod algebra through filtration 5. In principle, there should be no obstruction to continuing our computation of $\mathrm{Ext}^{\leq 3}_{\mathbb{R}}$ to higher filtrations, other than the rather more involved calculations and bookkeeping that this would necessarily take.

1.3.2 Motivic Brown–Gitler spectra Brown–Gitler spectra [1973] have many applications in classical algebraic topology, including Mahowald's analysis [1981; Shimamoto 1984] of the bo-resolution, Cohen's solution [1985] of the immersion conjecture, and more [Mahowald 1977; Hunter and Kuhn 1999; Goerss 1999]. The classical lambda algebra was essential for constructing and analyzing Brown–Gitler spectra [1973; Shimamoto 1984] as above, as well as [Goerss et al. 1986]. Culver and Quigley [2021] introduced a motivic analogue of the bo-resolution, the kq-resolution, and analyzed it over algebraically closed fields of characteristic zero. The analysis of the kq-resolution over more general base fields would be greatly simplified by the existence of motivic Brown–Gitler spectra.

1.3.3 Unstable motivic Adams spectral sequences The classical lambda algebra Λ^{cl} has certain subcomplexes $\Lambda^{\mathrm{cl}}(n)$ which form the E_1-page of an unstable Adams spectral sequence:

$$E_1 \cong \Lambda^{\mathrm{cl}}(n) \Rightarrow \pi_* S^n.$$

Moreover, James's 2-local fiber sequence [1957]

$$S^n \to \Omega S^{n+1} \to \Omega S^{2n+1},$$

which gives rise to the EHP sequence, is modeled by short exact sequences [Curtis 1971, Section 11]

$$0 \to \Lambda^{\mathrm{cl}}(n) \to \Lambda^{\mathrm{cl}}(n+1) \to \Sigma^n \Lambda^{\mathrm{cl}}(2n+1) \to 0,$$

which are useful for understanding both the unstable complexes $\Lambda^{\mathrm{cl}}(n)$ and the stable complex Λ^{cl}. It is natural to ask whether there are analogous subcomplexes of Λ^F related to a suitable motivic unstable Adams spectral sequence. The motivic situation seems to be much more delicate: it is not obvious how to define such subcomplexes of Λ^F, and the nature of the cohomology of motivic Eilenberg–Mac Lane spaces suggests that a motivic unstable Adams spectral sequence may not be as well behaved. A better understanding of these topics would shed light both on the nature of Λ^F and on unstable F-motivic homotopy theory.

1.4 Conventions

We maintain the following conventions throughout the paper:

(1) We work solely at the prime 2.

(2) We write F for a base field of characteristic not equal to 2.

(3) We write $\pi^F_{*,*}$ for the homotopy groups of the $(2,\eta)$-completed F-motivic sphere spectrum.

(4) Our homotopy and cohomology groups are bigraded by (s,w), where s is stem and w is weight.

(5) In particular, we write $S^{a,b}$ for the motivic sphere which is $\mathbb{A}^1$-homotopy equivalent to $\Sigma^{a-b}\mathbb{G}_m^{\wedge b}$.

(6) We write $H^{*,*}$ for reduced mod 2 F-motivic cohomology and H^* for reduced ordinary mod 2 cohomology.

(7) We write, for instance, $H^{*,*}(X_+)$ for the unreduced mod 2 motivic cohomology of X.

(8) We will use homological grading even for cohomology classes, in the sense that, if $x \in H^{a,b}(X)$, then we say $|x| = (-a,-b)$. This allows us to say, for instance, $|\tau| = (0,-1)$ and $|\rho| = (-1,-1)$, regardless or whether we are working with homology or cohomology.

(9) We write $\mathbb{M}^F = H^{*,*}(\mathrm{Spec}(F)_+)$ for the unreduced mod 2 motivic cohomology of a point.

(10) We write $\mathbb{M}^F_0$ for the portion of $\mathbb{M}^F$ concentrated on the line $s = w$, so that $\mathbb{M}^F = \mathbb{M}^F_0[\tau]$. (The ring $\mathbb{M}^F_0$ may be identified as the mod 2 Milnor K-theory of F, by work of Voevodsky; see [Isaksen and Østvær 2020, Section 2.1] for an overview of the structure of $\mathbb{M}^F$).

(11) We write Ext_F for the cohomology of the F-motivic Steenrod algebra, employing the grading conventions given in the following two points.

(12) We write Ext_F^f for the filtration f piece of Ext_F.

(13) We write $\mathrm{Ext}_F^{s,f,w} \subset \mathrm{Ext}_F^f$ for the subset of elements in filtration f with topological stem s and weight w.

(14) We use a subscript or superscript cl to denote classical objects; in particular, π_*^{cl} are the classical 2-completed stable stems, $\mathcal{A}^{\mathrm{cl}}$ is the classical mod 2 Steenrod algebra, and $\mathrm{Ext}_{\mathrm{cl}}$ is its cohomology.

(15) We take the binomial coefficient $\binom{a}{b}$ to be $\frac{a!}{b!(a-b)!}$ for $0 \le b \le a$, and to be zero otherwise.

Acknowledgements

The authors thank Uğur Yiğit for sharing a copy of his thesis, which proposes an alternative construction of a C_2-equivariant lambda algebra; Eva Belmont and Dan Isaksen for sharing data produced for their study of $\mathrm{Ext}_{\mathbb{R}}$; and the referees for helpful comments and suggestions. Balderrama thanks Nick Kuhn for some enlightening conversations about real projective spaces. Culver would like to thank the Max Planck Institute for providing a wonderful working environment and financial support while this paper was being written. Quigley thanks Eva Belmont, Dan Isaksen, and Inna Zakharevich for helpful discussions.

Part I The motivic lambda algebra

2 The motivic lambda algebra

In this section, we show that Priddy's construction [1970] of the lambda algebra as a certain Koszul complex can be extended to produce a motivic lambda algebra. As noted in Remark 1.2.1, a more refined notion of Koszulity is needed to handle the more exotic nature of the $\mathbb{M}^F$-algebra $\mathcal{A}^F$. The notion of a Koszul algebra has been generalized in various ways; see [Polishchuk and Positselski 2005] for an account of some developments in this area. We will use the formulation given in [Balderrama 2023, Section 3], as this gives a sufficiently general definition of Koszul algebra and explicit description of their associated Koszul complex. The reader familiar with Koszul algebras will find no surprises in this material.

In Section 2.2, we review the structure of the F-motivic Steenrod algebra $\mathcal{A}^F$. We show that $\mathcal{A}^F$ is in fact a Koszul algebra in Section 2.3, ultimately by reducing to Priddy's classical PBW criterion for Koszulity [Priddy 1970, Section 5]. The *F-motivic lambda algebra* Λ^F is then defined to be the Koszul complex of $\mathcal{A}^F$. We compute the structure of Λ^F explicitly, and introduce an endomorphism θ of Λ^F lifting the squaring operation Sq^0 on Ext_F. All of this structure is summarized in one place in Section 2.4.

2.1 Review of Koszul algebras

This section summarizes the definitions and facts from [Balderrama 2023, Section 3] regarding Koszul algebras which we will use to construct the motivic lambda algebra. We review this material in some detail, in order to specialize from the more abstract context considered there. Many of the results we need have appeared in varying levels of generality throughout the literature; in particular, the definition of Koszulity we use can be considered as a direct generalization of the homogeneous case considered by Rezk [2012, Section 4].

We fix throughout this subsection an associative algebra S to serve as our base ring, together with an associative algebra A which is an S-algebra in the sense of being equipped with an algebra map $S \to A$. Equivalently, A is a monoid in the category of S-bimodules. We abbreviate $\otimes = \otimes_S$.

We are most interested in the case where $S = \mathbb{M}^F$ and $A = \mathcal{A}^F$, and so, to avoid some subtle points regarding signs, we shall assume that S is of characteristic 2. In addition, we suppose throughout that A is projective as a left S-module.

Definition 2.1.1 Say that A is a *graded S-algebra* if we have chosen a decomposition $A = \bigoplus_{n \geq 0} A[n]$ of S-bimodules such that

(1) $S \cong A[0]$;

(2) the product on A restricts to $A[n] \otimes A[m] \to A[n+m]$.

Say that A is a *filtered S-algebra* if we have chosen a filtration $A \cong \operatorname{colim}_{n \to \infty} A_{\leq n}$ such that

(1) $S \cong A_{\leq 0}$;

(2) the product on A restricts to $A_{\leq n} \otimes A_{\leq m} \to A_{\leq n+m}$.

Finally, say that the filtration on a filtered S-algebra A is *projective* if (both A and) the associated graded algebra

$$\operatorname{gr} A := \bigoplus_{n \geq 0} A[n], \qquad A[n] := \operatorname{coker}(A_{\leq n-1} \to A_{\leq n})$$

are projective as left S-modules. ◁

Fix a left A-module M. Write $B^{\mathrm{un}}(A, A, M)$ and $B(A, A, M)$ for the unreduced and reduced bar resolutions of M relative to S; that is, for the unnormalized and normalized chain complexes associated to the standard monadic resolution of M with respect to the adjunction $\mathrm{L}\mathcal{M}\mathrm{od}_S \leftrightarrows \mathrm{L}\mathcal{M}\mathrm{od}_A$. These are projective left A-module resolutions provided that M is projective as a left S-module. If A is a filtered algebra, then $B^{\mathrm{un}}(A, A, M)$ is a filtered complex, with filtration defined by

$$B_n^{\mathrm{un}}(A, A, M)[\leq m] := \operatorname{Im}\Bigl(\bigoplus_{m_1 + \cdots + m_n = m} A \otimes A_{\leq m_1} \otimes \cdots \otimes A_{\leq m_n} \otimes M \to B_n^{\mathrm{un}}(A, A, M)\Bigr), \tag{2-1}$$

and this descends to a filtration of $B(A,A,M)$; compare for instance [Priddy 1970, Section 10; Rezk 2012, Section 4; Balderrama 2023, Section 3.5]. If A is augmented, then this augmentation makes S into an A-bimodule, allowing us to form the bar complex $B(A) := S \otimes_A B(A,A,S)$ and consider the homology $H_*(A) := H_*(B(A))$, and the filtration of (2-1) descends to a filtration on $B(A)$. If A is graded, then A is naturally filtered by $A_{\leq n} = \bigoplus_{i \leq n} A[i]$; this filtration is split in the sense that $A \cong \operatorname{gr} A$, and likewise the filtration on $B(A)$ is split with $\operatorname{gr} B(A) = \bigoplus_{m \geq 0} B(A)[m]$. This then passes to a splitting $H_*(A) \cong \bigoplus_{m \geq 0} H_*(A)[m]$.

Definition 2.1.2 [Rezk 2012, Definition 4.4; Balderrama 2023, Definition 3.5.3] We say that A is a *homogeneous Koszul S-algebra* provided that

(1) A has been given the structure of a graded S-algebra;

(2) $H_n(A)[m] = 0$ for $n \neq m$.

We say that A is a *Koszul S-algebra* if

(1) A has been equipped with a projective filtration;

(2) $\operatorname{gr} A$ is a homogeneous Koszul S-algebra. ◁

Suppose now that A is projectively filtered, and fix a left A-module M which is flat as a left S-module. The filtration of (2-1) on $B(A,A,M)$ induced by that on A satisfies $\operatorname{gr} B(A,A,M) \cong A \otimes B(\operatorname{gr} A) \otimes M$, and so the convergent spectral sequence associated to this filtration is of signature

$$E^1_{p,q} = A \otimes H_q(\operatorname{gr} A)[p] \otimes M \Rightarrow H_q B(A,A,M), \quad d^r_{p,q} \colon E^r_{p,q} \to E^r_{p-r,q-1}. \tag{2-2}$$

Definition 2.1.3 Let M be an A-module which is flat as a left S-module. The *Koszul resolution* of M is the augmented chain complex

$$M \leftarrow K(A,A,M)$$

defined by

$$K_p(A,A,M) = E^1_{p,p} = A \otimes H_p(\operatorname{gr} A)[p] \otimes M,$$

with differential given by the d^1-differential of the spectral sequence (2-2). When M is projective as a left S-module, we define the *Koszul complex* $K_A(M,M')$ as the cochain complex

$$K_A(M,M') := \operatorname{Hom}_A(K(A,A,M),M') \cong \operatorname{Hom}_S(H_*(\operatorname{gr} A) \otimes M, M'),$$

with differential inherited from that on $K(A,A,M)$. ◁

Observe that, by construction, $K(A,A,M)$ is a subcomplex of $B(A,A,M)$, and dually $K_A(M,M')$ is a quotient complex of the cobar complex $C_A(M,M') := \operatorname{Hom}_A(B(A,A,M),M')$. When A is Koszul, the spectral sequence of (2-2) collapses into the Koszul complex $K(A,A,M)$, proving the following.

Theorem 2.1.4 (see [Priddy 1970, Theorem 3.8; Rezk 2012, Proposition 4.8; Balderrama 2023, Theorem 3.5.5]) *Suppose that A is a Koszul S-algebra, and fix left A-modules M and M'.*

(1) *If M is flat over S, then there is an injective quasiisomorphism $K(A,A,M)\subset B(A,A,M)$.*

(2) *If M is projective over S, then there is a surjective quasiisomorphism $C_A(M,M')\to K_A(M,M')$.*

In particular, if M is projective over S, the homology of $K_A(M,M')$ is isomorphic to $\mathrm{Ext}_A(M,M')$. □

This allows us to define Koszul complexes in the generality we need. We now recall some facts from [Balderrama 2023, Sections 3.6–3.7] describing the structure of Koszul complexes; these are direct analogues of [Priddy 1970, Theorem 4.6]. We begin by fixing some conventions.

Definition 2.1.5 Fix a left S-module M. Then the dual $M^\vee=\mathrm{L}\mathcal{M}\mathrm{od}_S(M,S)$ carries the structure of a *right* S-module by

$$(f\cdot s)(m)=f(m)\cdot s.$$

If M is in fact an S-bimodule, then $M^\vee$ also carries an S-bimodule structure, with left S-module structure

$$(s\cdot f)(m)=(f(m\cdot s)).$$

Now, if M is a left S-module and M' is an S-bimodule, then there is a comparison map

$$c\colon M^\vee\otimes M'^\vee\to(M'\otimes M)^\vee,\quad c(f\otimes f')(m'\otimes m)=f'(m'f(m)).$$

If M is finitely presented and projective as a left S-module, then this map is an isomorphism. In general, if M'' is another left S-module, then we write

$$M^\vee\mathbin{\hat{\otimes}}M'':=\mathrm{L}\mathcal{M}\mathrm{od}_S(M,M''),$$

so that, in particular,

$$M^\vee\mathbin{\hat{\otimes}}M'^\vee\cong(M'\otimes M)^\vee;$$

in good cases, this may be realized as a topological tensor product, as the notation suggests. ◁

The theory of Koszul algebras is closely related to the theory of quadratic algebras; let us fix some notation.

Definition 2.1.6 Fix an S-bimodule B and subbimodule $R\subset B\otimes B$. The *quadratic algebra* generated by the pair (B,R) is the algebra

$$T(B,R):=\bigoplus_{n\geq 0}T_n(B,R),\quad T_n(B,R):=\mathrm{coker}\biggl(\sum_{i+j=n}B^{\otimes i-1}\otimes R\otimes B^{\otimes j-1}\to B^{\otimes n}\biggr),$$

with multiplication inherited from the tensor algebra $T(B)$. Similarly, given a subbimodule $R'\subset B^\vee\mathbin{\hat{\otimes}}B^\vee$ dual to a quotient of $B\otimes B$, we define the completed quadratic algebra

$$\hat{T}(B^\vee,R'):=\prod_{n\geq 0}\hat{T}_n(B^\vee,R'),\quad \hat{T}_n(B^\vee,R'):=\mathrm{coker}\biggl(\sum_{i+j=n}(B^\vee)^{\hat{\otimes}i-1}\mathbin{\hat{\otimes}}R'\mathbin{\hat{\otimes}}(B^\vee)^{\hat{\otimes}j-1}\to(B^\vee)^{\hat{\otimes}n}\biggr).$$

Say that (B,R) is a *quadratic datum* if $T(B,R)$ is projective. In this case, the *dual quadratic datum* to (B,R) is the pair $(B^\vee,R^\perp)$, where $R^\perp=(T_2(B,R))^\vee$. ◁

The cohomology of a homogeneous Koszul algebra may be explicitly described as follows.

Theorem 2.1.7 (see [Priddy 1970, Theorem 2.5; Rezk 2012, Proposition 4.12; Balderrama 2023, Theorem 3.6.4]) (1) *Let (B,R) be a quadratic datum. Then $H^1(T(B,R))[1]\cong B^\vee$, and the inclusion $B^\vee\subset H^*(T(B,R))$ extends to an isomorphism $\widehat{T}(B^\vee,R^\perp)\cong\prod_{n\geq 0}H^n(T(B,R))[n]$.*

(2) *Let $A=\bigoplus_{n\geq 0}A[n]$ be a homogeneous Koszul algebra, and let $R=\ker(A[1]\otimes A[1]\to A[2])$. Then $A\cong T(A[1],R)$ is quadratic, and $H^*(A)\cong\widehat{T}(A[1]^\vee,R^\perp)$.* □

Now fix a quadratic algebra $A=T(A[1],R)$ and left A-modules M and M', supposing that M is projective as a left S-module. We may use Theorem 2.1.7 to describe the Koszul complex $K_A(M,M')$. Recall that

$$K^n_A(M,M')=\mathrm{L}\mathcal{M}\mathrm{od}_A(A\otimes H_n(A)[n]\otimes M,M')\cong\mathrm{L}\mathcal{M}\mathrm{od}_S(H_n(A)[n]\otimes M,M').$$

If we suppose that $H_*(A)$ is projective as a left S-module, as holds if A is Koszul, then there is an isomorphism $(H_n(A))^\vee\cong H^n(A)$ of S-bimodules. In this case, we have

$$K^n_A(M,M')\cong\mathrm{L}\mathcal{M}\mathrm{od}_S(M,H^n(\mathrm{gr}\,A)[n]\mathbin{\widehat{\otimes}}M')\cong\mathrm{L}\mathcal{M}\mathrm{od}_S\big(M,\widehat{T}_n(A[1]^\vee,R^\perp)\mathbin{\widehat{\otimes}}M'\big);$$

Thus $K^*_A(M,M')$ is completely described as a graded object by Theorem 2.1.7.

It remains to describe the differential on $K_A(M,M')$. Observe first that, if M'' is an additional A-module, then there are pairings

$$\wr\colon K^n_A(M,M')\otimes_{\mathbb{Z}}K^{n'}_A(M',M'')\to K^{n+n'}_A(M,M'').$$

This is a pairing of chain complexes compatible with analogous pairings on cobar complexes and, when A is Koszul, it is a chain-level lift of the standard composition product in Ext_A. In addition, it may be described in terms of the product structure on $\widehat{T}(A[1]^\vee,R^\perp)$ as follows (see [Balderrama 2023, Sections 3.2 and 3.7]). Write μ for the multiplication on $\widehat{T}(A[1]^\vee,R^\perp)$. Then, given $f\colon M\to\widehat{T}_n(A[1]^\vee,R^\perp)\mathbin{\widehat{\otimes}}M'$ and $g\colon M'\to\widehat{T}_{n'}(A[1]^\vee,R^\perp)\mathbin{\widehat{\otimes}}M''$, we have

$$f\wr g=(\mu\otimes 1)\circ(1\otimes g)\circ f.$$

In the special case where $M=M'$, these pairings give $K_A(M,M)$ the structure of a differential graded algebra, and give $K_A(M,M')$ the structure of a differential graded $K_A(M,M)$-$K_A(M',M')$-bimodule. Note that $K^1_A(M,M)=\mathrm{L}\mathcal{M}\mathrm{od}_S(A[1]\otimes M,M)$. The A-module structure on M restricts to an element $Q^M\in K^1_A(M,M)$, and we have the following.

Theorem 2.1.8 [Balderrama 2023, Theorem 3.7.1] *The differential on $K_A(M,M')$ is given by*

$$\delta\colon K^n_A(M,M')\to K^{n+1}_A(M,M'),\quad \delta(f)=Q^M\wr f-f\wr Q^{M'}.$$

In particular, if $M=M'$, then $\delta(f)$ is the commutator $[Q^M,f]$. □

This theorem describes Koszul complexes for a homogeneous Koszul algebra. Suppose now that A is an arbitrary Koszul S-algebra, and continue to fix left A-modules M and M' with M projective as a left S-module. The additive and multiplicative structure of the Koszul complexes $K_A(M,M')$ depend only on the algebra $\operatorname{gr} A$ and left S-modules M and M', and so are still described by Theorem 2.1.7. In practice, the differential on $K_A(M, M')$ may be identified using the following.

Let $qR = \ker(A_{\le 1} \otimes A_{\le 1} \to A_{\le 2})$, and observe that $(A_{\le 1}, qR)$ is a quadratic datum. Let $A^{\mathrm{big}} = \bigoplus_{n\ge 0} A_{\le n}$. This is a graded algebra, and the inclusion $A_{\le 1} \subset A^{\mathrm{big}}$ extends to a map $T(A_{\le 1}, qR) \to A^{\mathrm{big}}$ of graded algebras.

Theorem 2.1.9 [Balderrama 2023, Theorem 3.7.3] (1) *$T(A_{\le 1}, qR) \to A^{\mathrm{big}}$ is an isomorphism of graded algebras.*

(2) *A^{big} is a homogeneous Koszul algebra.*

(3) *The surjection $A^{\mathrm{big}} \to A$ gives rise to short exact sequences*

$$0 \to K^n_A(M, M') \to K^n_{A^{\mathrm{big}}}(M, M') \to K^{n-1}_A(M, M') \to 0,$$

which are split if A is augmented.

In particular, $K_A(M, M') \subset K_{A^{\mathrm{big}}}(M, M')$ is a subcomplex with differential on the target described by Theorem 2.1.8. □

2.2 The motivic Steenrod algebra

We will construct the motivic lambda algebra by applying the theory recalled in Section 2.1 to the mod 2 motivic Steenrod algebra, whose structure we now recall. The conventions of Section 1.4 are in force throughout this section.

We note in particular that, following these conventions, we take the somewhat unconventional approach of consistently using *homological grading*. Thus, for example, $\tau \in H^{0,1}(\operatorname{Spec}(F)_+)$, but we shall write $|\tau| = (0, -1)$, as this is how it will appear in the lambda algebra.

We begin by recalling the general structure of the base ring $\mathbb{M}^F = H^{*,*}(\operatorname{Spec}(F)_+)$.

Example 2.2.1 For any F, we have $\mathbb{M}^F = \mathbb{M}^F_0[\tau]$, where

$$|\tau| = (0, -1)$$

and $\mathbb{M}^F_0 \subset \mathbb{M}^F$ is the subring concentrated on the line $s = w$, isomorphic to the Milnor K-theory of F taken mod 2. The following are some particular examples of the ring $\mathbb{M}^F_0$. We refer the reader to [Isaksen and Østvær 2020, Section 2.1] for further details.

- For $F = \overline{F}$ algebraically closed, such as $F = \mathbb{C}$, we have

$$\mathbb{M}^{\overline{F}}_0 \cong \mathbb{F}_2$$

- For $F = \mathbb{R}$ the real numbers, we have

$$\mathbb{M}^{\mathbb{R}}_0 \cong \mathbb{F}_2[\rho],$$

where $|\rho| = (-1, -1)$.

- For $F = \mathbb{F}_q$ a finite field of odd prime-power order q, we have
$$\mathbb{M}_0^{\mathbb{F}_q} \cong \begin{cases} \mathbb{F}_q[u]/u^2 & \text{if } q \equiv 1 \pmod 4, \\ \mathbb{F}_q[\rho]/\rho^2 & \text{if } q \equiv 3 \pmod 4, \end{cases}$$
where $|\rho| = |u| = (-1,-1)$.
- For $F = \mathbb{Q}_p$ the p-adic rationals with p an arbitrary prime, we have
$$\mathbb{M}_0^{\mathbb{Q}_p} \cong \begin{cases} \mathbb{F}_2[\pi,u]/(\pi^2,u^2) & \text{if } q \equiv 1 \pmod 4, \\ \mathbb{F}_2[\pi,\rho]/(\rho^2,\rho\pi+\pi^2) & \text{if } q \equiv 3 \pmod 4, \\ \mathbb{F}_2[\pi,\rho,u]/(\rho^3,u^2,\pi^2,\rho u,\rho\pi,\rho^2+u\pi) & \text{if } q = 2, \end{cases}$$
where $|\rho| = |u| = |\pi| = (-1,-1)$.

See also Section 7.1 for a discussion of $\mathbb{M}^{\mathbb{Q}}$. ◁

Voevodsky [2003] (with minor corrections by Riou [2012]) and Hoyois, Kelly and Østvær [Hoyois et al. 2017] have constructed Steenrod squares
$$\mathrm{Sq}^a\colon H^{m,n}(X) \to H^{m+a,n+\lfloor a/2 \rfloor}(X)$$
for $a \geq 0$ and shown that they generate the algebra $\mathcal{A}^F$ of natural operations in mod 2 motivic cohomology. It is convenient to take the convention that $\mathrm{Sq}^a = 0$ for $a < 0$. The relations between these squares are generated by $\mathrm{Sq}^0 = 1$ together with the *Adem relations*:

Theorem 2.2.2 [Voevodsky 2003, Theorem 10.2; Riou 2012, théorème 4.5.1; Hoyois et al. 2017, Theorem 5.1] *Fix positive integers a and b with $a < 2b$.*

If a is even and b is odd, then
$$\mathrm{Sq}^a\,\mathrm{Sq}^b = \sum_{0 \leq j \leq \lfloor a/2 \rfloor} \binom{b-1-j}{a-2j} \mathrm{Sq}^{a+b-j}\,\mathrm{Sq}^j + \sum_{\substack{1 \leq j \leq \lfloor a/2 \rfloor \\ j \text{ odd}}} \binom{b-1-j}{a-2j} \rho\,\mathrm{Sq}^{a+b-j-1}\,\mathrm{Sq}^j.$$

If a and b are odd, then
$$\mathrm{Sq}^a\,\mathrm{Sq}^b = \sum_{\substack{1 \leq j \leq \lfloor a/2 \rfloor \\ j \text{ odd}}} \binom{b-1-j}{a-2j} \mathrm{Sq}^{a+b-j}\,\mathrm{Sq}^j.$$

If a and b are even, then
$$\mathrm{Sq}^a\,\mathrm{Sq}^b = \sum_{0 \leq j \leq \lfloor a/2 \rfloor} \tau^{j \bmod 2} \binom{b-1-j}{a-2j} \mathrm{Sq}^{a+b-j}\,\mathrm{Sq}^j.$$

If a is odd and b is even, then
$$\mathrm{Sq}^a\,\mathrm{Sq}^b = \sum_{\substack{0 \leq j \leq \lfloor a/2 \rfloor \\ j \text{ even}}} \binom{b-1-j}{a-2j} \mathrm{Sq}^{a+b-j}\,\mathrm{Sq}^j + \sum_{\substack{1 \leq j \leq \lfloor a/2 \rfloor \\ j \text{ odd}}} \binom{b-1-j}{a-1-2j} \rho\,\mathrm{Sq}^{a+b-j-1}\,\mathrm{Sq}^j.$$

In all cases, the bounds on summation are not necessary, but give regions where the given binomial coefficients may be nonzero. □

As with the classical Steenrod algebra, $\mathcal{A}^F$ admits an admissible basis.

Definition 2.2.3 Given a sequence $I = (r_1, \dots, r_k)$ with $r_i > 0$ for all $1 \le i \le k$, we abbreviate $\mathrm{Sq}^I = \mathrm{Sq}^{r_1} \dots \mathrm{Sq}^{r_k}$. Say that Sq^I is *admissible* if $r_i \ge 2r_{i+1}$ for all $1 \le i \le k-1$. ◁

Proposition 2.2.4 [Voevodsky 2003, Section 11] *$\mathcal{A}^F$ is freely generated as a left $\mathbb{M}^F$-module by the admissible squares Sq^I.* □

The mod 2 motivic cohomology $H^{*,*}(X_+)$ of any smooth scheme X carries the structure of a left $\mathcal{A}$-module. These actions satisfy the following *Cartan formulas*.

Proposition 2.2.5 [Voevodsky 2003, Proposition 9.6; Riou 2012, Proposition 4.4.2] *Let $a \ge 0$ and $x, y \in H^{*,*}(X_+)$. Then*

$$\mathrm{Sq}^{2a}(xy) = \sum_{r=0}^{a} \mathrm{Sq}^{2r}(x)\,\mathrm{Sq}^{2a-2r}(y) + \tau \sum_{s=0}^{a-1} \mathrm{Sq}^{2s+1}(x)\,\mathrm{Sq}^{2a-2s-1}(y),$$

$$\mathrm{Sq}^{2a+1}(xy) = \sum_{r=0}^{a} (\mathrm{Sq}^{2r+1}(x)\,\mathrm{Sq}^{2ai-2r}(y) + \mathrm{Sq}^{2r}(x)\,\mathrm{Sq}^{2a-2r+1}(y)) + \rho \sum_{s=0}^{a-1} \mathrm{Sq}^{2s+1}(x)\,\mathrm{Sq}^{2a-2s-1}(y). \quad \square$$

The action of $\mathcal{A}^F$ on $\mathbb{M}^F$ is determined by these Cartan formulas and the following.

Proposition 2.2.6 [Voevodsky 2003; Röndigs and Østvær 2016, Appendix A] *The action of $\mathcal{A}^F$ on $\mathbb{M}^F$ satisfies*

$$\mathrm{Sq}^{\ge 1}(x) = 0 \quad \text{for } x \in \mathbb{M}_0^F, \qquad \mathrm{Sq}^1(\tau) = \rho, \qquad \mathrm{Sq}^{\ge 2}(\tau) = 0. \quad \square$$

As in the classical case, the Cartan formulas of Proposition 2.2.5 may be encoded in a coproduct on the algebra $\mathcal{A}^F$. The resulting structure is not quite a Hopf algebra, but is dual to a Hopf algebroid structure on the dual Steenrod algebra $(\mathcal{A}^F)^\vee$. This complication arises in part due to the following. The Steenrod algebra $\mathcal{A}^F$ is an $\mathbb{M}^F$-algebra, by way of the homomorphism $\mathbb{M}^F \to \mathcal{A}^F$ sending an element $x \in \mathbb{M}^F$ to the stable operation given by left multiplication by x. However, $\mathbb{M}^F$ does not land in the center of $\mathcal{A}^F$; equivalently, $\mathcal{A}^F$ has nontrivial $\mathbb{M}^F$-bimodule structure. We may describe this structure explicitly as follows.

Proposition 2.2.7 *The $\mathbb{M}^F$-bimodule structure of $\mathcal{A}^F$ is determined by*

$$\begin{aligned} \mathrm{Sq}^n\, x &= x\, \mathrm{Sq}^n \quad \text{for } x \in \mathbb{M}_0^F, \\ \mathrm{Sq}^{2n}\, \tau &= \tau\, \mathrm{Sq}^{2n} + \rho\tau\, \mathrm{Sq}^{2n-1}, \\ \mathrm{Sq}^{2n+1}\, \tau &= \tau\, \mathrm{Sq}^{2n+1} + \rho\, \mathrm{Sq}^{2n} + \rho^2\, \mathrm{Sq}^{2n-1}. \end{aligned}$$

Proof It suffices to show both sides of each equality coincide when evaluated on an arbitrary cohomology class. For example, for any X and $x \in H^{*,*}(X_+)$, we have

$$\begin{aligned}(\mathrm{Sq}^{2n}\tau)(x) = \mathrm{Sq}^{2n}(\tau x) &= \sum_{i+j=n}(\mathrm{Sq}^{2i}\tau)(\mathrm{Sq}^{2j}x) + \tau\sum_{i+j=n-1}(\mathrm{Sq}^{2i+1}\tau)(\mathrm{Sq}^{2j+1}x)\\ &= \tau\,\mathrm{Sq}^{2n}(x) + \rho\tau\,\mathrm{Sq}^{2n-1}(x)\end{aligned}$$

by Proposition 2.2.5. This proves the second equation, and the other cases are similar. □

Remark 2.2.8 Although we work in this section over an arbitrary base field F, there is a sense in which $F = \mathbb{R}$ represents the universal case: the class ρ may be defined over any field F, making $\mathbb{M}^F$ into an $\mathbb{M}^{\mathbb{R}}$-module, and in all cases we have

$$\mathcal{A}^F = \mathbb{M}^F \otimes_{\mathbb{M}^{\mathbb{R}}} \mathcal{A}^{\mathbb{R}}.$$

In fact, the formulas of Proposition 2.2.6 describe an action of $\mathcal{A}^{\mathbb{R}}$ on $\mathbb{M}^F$ for which

$$\mathrm{Ext}_F \cong \mathrm{Ext}_{\mathcal{A}^{\mathbb{R}}}(\mathbb{M}^{\mathbb{R}}, \mathbb{M}^F),$$

and at least additively this depends only on the $\mathbb{F}_2[\rho]$-module structure of $\mathbb{M}_0^F$.

It is worth putting this observation in a slightly more general context. The Cartan formulas of Proposition 2.2.5 give the category of left $\mathcal{A}^{\mathbb{R}}$-modules a symmetric monoidal structure. If R is a monoid in this category, then the tensor product $R \otimes_{\mathbb{M}^{\mathbb{R}}} \mathcal{A}^{\mathbb{R}}$ may be equipped with a product with the property that

$$\mathrm{L}\mathcal{M}\mathrm{od}_{R\otimes_{\mathbb{M}^{\mathbb{R}}}\mathcal{A}^{\mathbb{R}}} \simeq \mathrm{L}\mathcal{M}\mathrm{od}_R(\mathrm{L}\mathcal{M}\mathrm{od}_{\mathcal{A}^{\mathbb{R}}});$$

this is the *semitensor product* of [Massey and Peterson 1965]. Moreover, we have

$$\mathrm{Ext}_{R\otimes_{\mathbb{M}^{\mathbb{R}}}\mathcal{A}^{\mathbb{R}}}(R, R) \cong \mathrm{Ext}_{\mathcal{A}^{\mathbb{R}}}(\mathbb{M}^{\mathbb{R}}, R).$$

The algebras $\mathcal{A}^F$ are obtained in the case where $R = \mathbb{M}^F$. Another simple class of example is given by the algebras $\mathcal{A}^{\mathbb{R}}/(\rho^n, \tau^m)$, where n and m are such that τ^m is central in $\mathcal{A}^{\mathbb{R}}/(\rho^n)$. A more interesting example is the following: there is an isomorphism of algebras

$$\mathcal{A}^{C_2} \cong \mathbb{M}^{C_2} \otimes_{\mathbb{M}^{\mathbb{R}}} \mathcal{A}^{\mathbb{R}},$$

where $\mathcal{A}^{C_2}$ is the C_2-equivariant Steenrod algebra, $\mathbb{M}^{C_2}$ is the C_2-equivariant cohomology of a point, and $\mathcal{A}^{\mathbb{R}}$ acts on $\mathbb{M}^{C_2}$ as described, for instance, in [Guillou et al. 2020, Section 2] (building on [Hu and Kriz 2001]). ◁

2.3 The motivic lambda algebra

We now produce the motivic lambda algebra. For simplicity of notation, we consider the base field F as fixed, and abbreviate

$$\mathcal{A} = \mathcal{A}^F, \quad \mathbb{M} = \mathbb{M}^F$$

throughout this subsection.

2.3.1 Koszulity of $\mathcal{A}$ We begin by showing that $\mathcal{A}$ is Koszul. The algebra $\mathcal{A}$ is a projectively filtered $\mathbb{M}$-algebra under the length filtration: $\mathcal{A}_{\leq n} \subset \mathcal{A}$ is the submodule generated by squares Sq^I where I is a sequence of length at most n. In particular,

$$\mathcal{A}_{\leq 1} = \mathbb{M}\{\mathrm{Sq}^a : a \geq 0\}$$

as a left $\mathbb{M}$-module, with the understanding that $\mathrm{Sq}^0 = 1$ in $\mathcal{A}$. By Definition 2.1.3, to show that $\mathcal{A}$ is Koszul we must show that $\operatorname{gr} \mathcal{A}$ is homogeneous Koszul. To show that the classical Steenrod algebra is Koszul, Priddy [1970, Theorem 5.3] developed a *PBW criterion* for Koszulity. We cannot apply this criterion directly, in part due to the nontrivial $\mathbb{M}$-bimodule structure of $\operatorname{gr} \mathcal{A}$. Our strategy is to filter this issue away, thereby reducing to Priddy's criterion.

Theorem 2.3.1 *$\mathcal{A}$ is a Koszul $\mathbb{M}$-algebra.*

Proof As $\mathcal{A}$ is a projectively filtered algebra, we need only show that $\operatorname{gr} \mathcal{A}$ is a homogeneous Koszul algebra, ie that $H_n(\operatorname{gr} \mathcal{A})[m] = 0$ for $n \neq m$. To that end, we define a new filtration $\overline{F}_\bullet \operatorname{gr} \mathcal{A}$ on $\operatorname{gr} \mathcal{A}$ by declaring $\overline{F}_{\leq m} \operatorname{gr} \mathcal{A} \subset \operatorname{gr} \mathcal{A}$ to be generated by elements of the form Sq^I, where $I = (r_1, \ldots, r_k)$ is a sequence satisfying $r_1 + \cdots + r_k \leq m$. This filtration is multiplicative, and so we may form its associated graded algebra $\overline{\operatorname{gr}} \operatorname{gr} \mathcal{A}$.

The same construction employed in Section 2.1 shows that the filtration $\overline{F}_\bullet \operatorname{gr} \mathcal{A}$ induces a filtration on the bar complex $B(\operatorname{gr} \mathcal{A})$ with associated graded $B(\overline{\operatorname{gr}} \operatorname{gr} \mathcal{A})$. This filtration is compatible with the decomposition

$$B(\operatorname{gr} \mathcal{A}) \cong \bigoplus_{m \geq 0} B(\operatorname{gr} \mathcal{A})[m],$$

and so, for each m, there is a convergent spectral sequence

$$E_1^n = H_n B(\overline{\operatorname{gr}} \operatorname{gr} \mathcal{A})[m] \Rightarrow H_n(\operatorname{gr} \mathcal{A})[m].$$

It is thus sufficient to verify that $\overline{\operatorname{gr}} \operatorname{gr} \mathcal{A}$ is a homogeneous Koszul algebra with respect to the grading $\overline{\operatorname{gr}} \operatorname{gr} \mathcal{A} = \bigoplus_{m \geq 0} \overline{\operatorname{gr}} \operatorname{gr}^m \mathcal{A}$. By passing from $\operatorname{gr} \mathcal{A}$ to $\overline{\operatorname{gr}} \operatorname{gr} \mathcal{A}$, we have filtered away both the nontrivial $\mathbb{M}$-bimodule structure on $\operatorname{gr} \mathcal{A}$ described in Proposition 2.2.7 and the parts of the Adem relations involving ρ which appear in Theorem 2.2.2, and in the end we may identify

$$\overline{\operatorname{gr}} \operatorname{gr} \mathcal{A} \cong \mathbb{M}^F \otimes_{\mathbb{F}_2[\tau]} \operatorname{gr} \mathcal{A}^{\mathbb{C}}.$$

From here, it is easily seen that the admissible basis of $\overline{\operatorname{gr}} \operatorname{gr} \mathcal{A}$ satisfies Priddy's PBW criterion [1970, Section 5.1]. It now follows from [loc. cit., Theorem 5.3] that $\overline{\operatorname{gr}} \operatorname{gr} A$ is Koszul; the assumption in [loc. cit.] that the base is a field is not needed so long as everything in sight is free over the base. □

Remark 2.3.2 When $F = \mathbb{R}$, the filtration $\overline{F}_\bullet \operatorname{gr} A$ coincides with the ρ-adic filtration of $\operatorname{gr} A$. The use of $\overline{F}$ allows us to apply our argument uniformly to arbitrary base fields, but we could have also proved Theorem 2.3.1 in the $\mathbb{R}$-motivic case, and deduced the general case from this. Indeed, everything in Section 2.1 is compatible with base change (see [Balderrama 2023, Lemma 3.5.7]), so Koszulity of $\mathcal{A}^{\mathbb{R}}$

implies that any algebra obtained from the construction of Remark 2.2.8 is Koszul. As an example not explicitly covered by the statement of Theorem 2.3.1, $\mathcal{A}^{C_2}$ is Koszul over $\mathbb{M}^{C_2}$. ◁

Definition 2.3.3 The *F-motivic lambda algebra* Λ^F is the Koszul complex $K_{\mathcal{A}^F}(\mathbb{M}^F, \mathbb{M}^F)$ associated to the Koszul $\mathbb{M}^F$-algebra $\mathcal{A}^F$, as defined in Definition 2.1.3, where $\mathcal{A}^F$ acts on $\mathbb{M}^F$ as described in Proposition 2.2.6. ◁

We shall abbreviate $\Lambda = \Lambda^F$ throughout the rest of this subsection. Theorem 2.1.4 now implies the following.

Theorem 2.3.4 *Let $C(\mathcal{A}) = C_{\mathcal{A}}(\mathbb{M}, \mathbb{M})$ denote the cobar complex of $\mathcal{A}$. Then there is a surjective multiplicative quasiisomorphism*

$$C(\mathcal{A}) \to \Lambda.$$

In particular,

$$H_*\Lambda \cong \mathrm{Ext}^*_{\mathcal{A}}(\mathbb{M}, \mathbb{M}),$$

and this isomorphism is compatible with all products and Massey products. □

Remark 2.3.5 More generally, the theory recalled in Section 2.1 produces and describes Koszul complexes $K_{\mathcal{A}}(M, M')$ modeling the cobar complex $C_{\mathcal{A}}(M, M')$ for any left $\mathcal{A}$-modules M and M' with M projective over $\mathbb{M}$. Classically, the case where $M = H^*(\mathbb{R}P^\infty)$ and $M' = \mathbb{F}_2$ is of particular importance. Another amusing example is given over $F = \mathbb{R}$ with the observation that $K_{\mathcal{A}^{\mathbb{R}}}(\mathbb{M}^{\mathbb{R}}, \mathbb{M}^{C_2}) \cong K_{\mathcal{A}^{C_2}}(\mathbb{M}^{C_2}, \mathbb{M}^{C_2}) = \Lambda^{C_2}$ (see Remarks 2.2.8 and 2.3.2). ◁

2.3.2 The structure of the motivic lambda algebra We will now apply the theory recalled in Section 2.1 to describe Λ explicitly. First note that $\Lambda = \bigoplus_{m\geq 0} \Lambda[m]$ with $\Lambda[1] = (\mathcal{A}[1])^\vee$, where $\mathcal{A}[1] = \mathrm{coker}(\mathbb{M} \to \mathcal{A}_{\leq 1})$. As a left $\mathbb{M}$-module, we may identify

$$\mathcal{A}[1] = \mathbb{M}\{\mathrm{Sq}^r : r \geq 1\}.$$

Dualizing, we may identify

$$\Lambda[1] = \{\lambda_r : r \geq 0\}\mathbb{M}$$

as a right $\mathbb{M}$-module, where λ_r is dual to Sq^{r+1} in the given basis. Considering internal algebraic degrees yields $|\lambda_r| = \left(r+1, \left\lfloor \frac{1}{2}(r+1) \right\rfloor\right)$; following our conventions (Section 1.4), we subtract off the filtration from the algebraic stem to obtain the topological stem, and so instead write $|\lambda_r| = \left(r, \left\lceil \frac{1}{2} r \right\rceil\right)$.

We now begin by describing the multiplicative structure of Λ.

Proposition 2.3.6 *The left $\mathbb{M}$-module structure on $\Lambda[1]$ is determined by*

$$x\lambda_n = \lambda_n x \quad \text{for } x \in \mathbb{M}_0, \qquad \tau\lambda_{2n+1} = \lambda_{2n+1}\tau + \lambda_{2n+2}\rho, \qquad \tau\lambda_{2n} = \lambda_{2n}\tau + \lambda_{2n+1}\tau\rho + \lambda_{2n+2}\rho^2.$$

Proof This follows by dualizing Proposition 2.2.7. □

Proposition 2.3.7 *If a is odd or b is even, then*

$$\lambda_a\lambda_{2a+b+1} = \sum_{0\leq r<b/2} \lambda_{a+b-r}\lambda_{2a+1+r}\binom{b-r-1}{r},$$

and if a is even and b is odd, then

$$\lambda_a\lambda_{2a+b+1} = \sum_{0\leq r<b/2} \lambda_{a+b-r}\lambda_{2a+1+r}\binom{b-r-1}{r}\tau^{(r-1) \bmod 2} + \sum_{0\leq r\leq (b+1)/2} \lambda_{a+b+1-r}\lambda_{2a+1+r}\binom{\lfloor\frac{1}{2}b\rfloor-\lfloor\frac{1}{2}r\rfloor}{\lfloor\frac{1}{2}r\rfloor}\rho.$$

Proof By Theorem 2.1.7, the bimodule of relations defining Λ as a quadratic algebra with generating bimodule $\Lambda[1]$ may be identified as $\mathcal{A}[2]^\vee = \ker(\mathcal{A}[1]^\vee \otimes \mathcal{A}[1]^\vee \to R^\vee)$, where $R \subset \mathcal{A}[1]\otimes\mathcal{A}[1]$ is the projection of the subbimodule $qR \subset \mathcal{A}_{\leq 1}\otimes\mathcal{A}_{\leq 1}$ of Adem relations recalled in Theorem 2.2.2. It follows by direct computation that this kernel is generated by the indicated relations. □

Remark 2.3.8 Unless both a and b are even, the Adem relation expanding a product of the form $\lambda_a\lambda_b$ is exactly as in the classical lambda algebra. ◁

The additive structure of Λ may be understood just as in the classical case.

Definition 2.3.9 Given a sequence $I = (r_1,\dots,r_n)$, write $\lambda_I = \lambda_{r_1}\cdots\lambda_{r_n}$. Call the sequence I *coadmissible* if $2r_i \geq r_{i+1}$ for all $1\leq i\leq n-1$. ◁

Proposition 2.3.10 *Λ is freely generated as a right $\mathbb{M}$-module by classes of the form λ_I, where I is a coadmissible sequence.*

Proof The relations of Proposition 2.3.7 imply that the coadmissible classes λ_I generate Λ as a right $\mathbb{M}$-module, and we must only verify that they do so freely. Following Remarks 2.2.8 and 2.3.2, there is an isomorphism

$$\Lambda \cong \Lambda^{\mathbb{R}} \otimes_{\mathbb{M}^{\mathbb{R}}} \mathbb{M};$$

thus we may reduce to the case where $F=\mathbb{R}$. By construction, Λ is free as a right $\mathbb{M}$-module. Thus, to show that the coadmissible classes λ_I freely generate Λ over $\mathbb{M}$, it is sufficient to verify the same for $\Lambda/(\rho)[\tau^{-1}]$ over $\mathbb{M}/(\rho)[\tau^{-1}]$. There is an isomorphism $\Lambda/(\rho)[\tau^{-1}] \cong \Lambda^{\mathrm{cl}} \otimes_{\mathbb{F}_2} \mathbb{F}_2[\tau^{\pm 1}]$, so this follows from the classical case. □

Finally, we describe the differential on Λ by applying Theorem 2.1.8.

Proposition 2.3.11 *The differential on Λ is determined by the Leibniz rule, together with*

$$\delta(x) = 0 \quad \text{for } x\in\mathbb{M}_0 \qquad \delta(\tau) = \lambda_0\rho, \qquad \delta(\lambda_n) = \sum_{1\leq r\leq n/2} \lambda_{n-r}\lambda_{r-1}\binom{n-r}{r}.$$

Proof Recall the construction $\mathcal{A}^{\mathrm{big}} = \bigoplus_{m\geq 0} \mathcal{A}_{\leq m}$ used in the statement of Theorem 2.1.9. By inspection, we find that $\mathcal{A}^{\mathrm{big}}$ may be identified as the "big motivic Steenrod algebra", defined with generators and relations the same as $\mathcal{A}$ only without the stipulation that $\mathrm{Sq}^0 = 1$. Let $\Lambda^{\mathrm{big}} = K_{\mathcal{A}^{\mathrm{big}}}(\mathbb{M}, \mathbb{M})$, where $\mathcal{A}^{\mathrm{big}}$ acts on $\mathbb{M}$ through the quotient $\mathcal{A}^{\mathrm{big}} \to \mathcal{A}$, ie with Sq^0 acting by the identity.

Theorem 2.1.9 tells us that $\mathcal{A}^{\mathrm{big}}$ is a homogeneous Koszul algebra, and that there is an inclusion $\Lambda \subset \Lambda^{\mathrm{big}}$ of differential graded algebras. As $\mathcal{A}^{\mathrm{big}}$ is homogeneous Koszul, Theorem 2.1.7 applies to show that Λ^{big} is generated by classes λ_r for $r \geq -1$, subject to relations of the same form as described for Λ in Propositions 2.3.6 and 2.3.7. The inclusion $\Lambda \subset \Lambda^{\mathrm{big}}$ is the obvious one, identifying Λ as the subalgebra of Λ^{big} generated by the classes λ_r for $r \geq 0$.

Theorem 2.1.8 describes the differential on Λ^{big} as

$$\delta(f) = [Q, f] = Q \cdot f - f \cdot Q,$$

where $Q \in \Lambda^{\mathrm{big}}[1] \cong (\mathcal{A}^{\mathrm{big}}[1])^{\vee}$ is the map $\mathcal{A}^{\mathrm{big}}[1] \cong \mathcal{A}_{\leq 1} \otimes \mathbb{M} \to \mathbb{M}$ induced by the action of $\mathcal{A}^{\mathrm{big}}$ on $\mathbb{M}$. In the basis $\mathcal{A}^{\mathrm{big}}[1] = \mathcal{A}_{\leq 1} = \mathbb{M}\{\mathrm{Sq}^r : r \geq 0\}$, this map is the projection onto Sq^0, which by definition is the class $\lambda_{-1} \in \Lambda^{\mathrm{big}}$. So the differential on Λ^{big} is given by

$$\delta(f) = [\lambda_{-1}, f] = \lambda_{-1} f - f \lambda_{-1},$$

and $\Lambda \subset \Lambda^{\mathrm{big}}$ is closed under this. The proposition follows upon expanding out this commutator using the relations defining the algebra Λ^{big}. □

Remark 2.3.12 The description of the differential on Λ as the commutator $\delta(f) = [\lambda_{-1}, f]$ has appeared classically as well; see [Bruner 1988, page 83]. ◁

2.3.3 A closed formula for $\delta(\tau^n)$

Proposition 2.3.11 gives a recursive process for computing $\delta(\tau^n)$. It is possible to solve this recursion, and we do so here. Recall that the pair $(\mathbb{M}, \mathcal{A}^{\vee})$ carries the structure of a Hopf algebroid. In particular, $\mathcal{A}^{\vee}$ is a commutative ring, and $\mathcal{A}^{\vee}_{\leq 1}$ is a quotient of this ring. Now, the differential $\delta\colon \Lambda[0] \to \Lambda[1]$ may be described as the composite

$$\eta_R + \eta_L \colon \Lambda[0] = \mathbb{M} \to \mathcal{A}^{\vee} \to \mathcal{A}^{\vee}_{\leq 1} \to \mathrm{coker}(\mathbb{M} \to \mathcal{A}^{\vee}_{\leq 1}) = \Lambda[1],$$

where $\eta_L, \eta_R \colon \mathbb{M} \to \mathcal{A}^{\vee}$ are given by $\eta_R(m)(a) = \epsilon(ma)$ and $\eta_L(m)(a) = \epsilon(am)$, where $\epsilon\colon \mathcal{A} = \mathcal{A} \otimes_{\mathbb{M}} \mathbb{M} \to \mathbb{M}$ encodes the action of $\mathcal{A}$ on $\mathbb{M}$.

We may use this interpretation to compute $\delta(\tau^n)$. The full structure of the Hopf algebroid $(\mathbb{M}, \mathcal{A}^{\vee})$ was determined by Voevodsky [2003]; however, we only need a small piece of this, which is easily computed by hand from the structure of $\mathcal{A}$ recalled in Section 2.2. We record this piece in the following.

Lemma 2.3.13 *There is an isomorphism of rings*

$$\mathcal{A}^{\vee}_{\leq 1} = \mathbb{M}[\tau_0, \xi_1]/(\tau_0^2 + \xi_1 \tau_0 \rho + \xi_1 \tau),$$

where the quotient map

$$\mathcal{A}^{\vee}_{\leq 1} \to \Lambda[1]$$

acts by

$$\tau_0^\epsilon \xi_1^n \mapsto \lambda_{2n-1+\epsilon}$$

for $\epsilon \in \{0,1\}$ *and* $n \geq 0$, *with the interpretation that* $\lambda_{-1} = 0$. *Moreover, the maps* $\eta_L, \eta_R \colon \mathbb{M} \to \mathcal{A}_{\leq 1}^\vee$ *act by*

$$\eta_R(x) = x \quad \text{for } x \in \mathbb{M}, \qquad \eta_L(x) = x \quad \text{for } x \in \mathbb{M}_0, \qquad \eta_L(\tau) = \tau + \tau_0\rho.$$

Proof The structure of the ring $\mathcal{A}_{\leq 1}^\vee$ may be read off the coproduct of $\mathcal{A}$, as given in Proposition 2.2.5, and its relation with our basis of $\Lambda[1]$ then follows by construction. The behavior of the left and right units may be read off the $\mathbb{M}$-bimodule structure of $\mathcal{A}_{\leq 1}$ as given in Proposition 2.3.11, together with knowledge of the counit map $\epsilon \colon \mathcal{A}_{\leq 1} \to \mathbb{M}$ given in Proposition 2.2.6. □

The main input to our computation of $\delta(\tau^n)$ is the following elementary computation.

Lemma 2.3.14 *In the ring* $\mathcal{A}_{\leq 1}^\vee$, *we have*

$$\tau_0^n = \sum_{\substack{\epsilon \in \{0,1\} \\ (n-\epsilon)/2 \leq i \leq n-1}} \tau_0^\epsilon \xi_1^i \binom{i+\epsilon-1}{n-i-1} \tau^{n-i-\epsilon} \rho^{2i-n+\epsilon}.$$

These bounds on i are not necessary, but give a region where the binomial coefficients may be nonzero.

Proof We first compute τ^n in the quotient ring

$$\mathbb{F}_2[\tau_0, \xi_1]/(\tau_0^2 + \xi_1\tau_0 + \xi_1)$$

of $\mathcal{A}_{\leq 1}^\vee$, in which both τ and ρ are set to 1. Clearly,

$$\tau_0^n = \sum_{0 \leq i \leq n} (\xi_1^i c_{n,i} + \tau_0 \xi_1^i d_{n,i})$$

for some $c_{n,i}, d_{n,i} \in \mathbb{F}_2$. The relation

$$\tau_0^n = \xi_1(\tau_0^{n-1} + \tau_0^{n-2})$$

gives rise to recurrence relations

$$c_{n,i} = c_{n-1,i-1} + c_{n-2,i-1}, \quad d_{n,i} = d_{n-1,i-1} + d_{n-2,i-1}.$$

Set $c'_{i,n} = c_{n+i,i}$ and $d'_{i,n} = d_{n+i,i}$. Then these relations become

$$c'_{i,n} = c'_{i-1,n-1} + c'_{i-1,n}, \quad d'_{i,n} = d'_{i-1,n-1} + d'_{i-1,n},$$

exactly as seen in Pascal's triangle. Paired with the initial conditions

$$c'_{i,0} = c'_{0,1} = d'_{1,0} = 0, \quad c'_{1,1} = 1 = d'_{0,1},$$

we find that

$$c'_{i,n} = \binom{i-1}{n-1}, \quad d'_{n,i} = \binom{i}{n-1},$$

and thus

$$c_{n,i}=\binom{i-1}{n-i-1},\quad d_{n,i}=\binom{i}{n-i-1}.$$

Plugging this back in, we find

$$\tau_0^n=\sum_{0\le i\le n}\left(\xi_1^i\binom{i-1}{n-i-1}+\tau_0\xi_1^i\binom{i}{n-i-1}\right)=\sum_{\substack{\epsilon\in\{0,1\}\\0\le i\le n}}\tau_0^\epsilon\xi_1^i\binom{i+\epsilon-1}{n-i-1}.$$

To compute τ_0^n in $\mathcal{A}^\vee_{\le 1}$ itself, recall that $|\tau|=(0,-1)$, $|\rho|=(-1,-1)$, $|\tau_0|=(1,0)$, $|\xi_1|=(2,1)$. Solving

$$|\tau_0^n|=|\tau_0^\epsilon\xi_1^i\tau^a\rho^b|$$

yields

$$a=n-i-\epsilon,\quad b=2i-n+\epsilon.$$

It follows that

$$\tau_0^n=\sum_{\substack{\epsilon\in\{0,1\}\\0\le i\le n}}\tau_0^\epsilon\xi_1^i\binom{i+\epsilon-1}{n-i-1}\tau^{n-i-\epsilon}\rho^{2i-n+\epsilon}$$

in $\mathcal{A}^\vee_{\le 1}$. For this binomial coefficient to be nonzero, we require

$$0\le i+\epsilon-1,\quad 0\le n-i-1,\quad n-i-1\le i+\epsilon-1,$$

giving the stated bounds on summation. □

Proposition 2.3.15 *The differential δ satisfies*

$$\delta(\tau^n)=\sum_{r\ge 0}\lambda_r\binom{n+\lfloor\frac12 r\rfloor}{r+1}\tau^{n-\lfloor r/2\rfloor-1}\rho^{r+1}.$$

Proof Following Lemma 2.3.13, to compute $\delta(\tau^n)$ one may compute

$$\tau^n+(\tau+\tau_0\rho)^n$$

in terms of the standard basis of $\mathcal{A}^\vee_{\le 1}=\mathbb{M}[\tau_0,\xi_1]/(\tau_0^2+\xi_1\tau_0\rho+\xi_1\tau)$. Moreover, it is sufficient to work in the quotient of $\mathcal{A}^\vee_{\le 1}$ wherein τ and ρ are set to 1, as the necessary quantity of τ's and ρ's may be recovered by comparing degrees, just as in the proof of Lemma 2.3.14. Using Lemma 2.3.14, we find

$$1+(1+\tau_0)^n=\sum_{1\le k\le n}\binom{n}{k}\tau_0^k=\sum_{1\le k\le n}\binom{n}{k}\sum_{\substack{\epsilon\in\{0,1\}\\i\ge 0}}\binom{i+\epsilon-1}{k-i-1}\tau_0^\epsilon\xi_1^i;$$

here we are free to omit the bounds of summation on i, as they merely recorded when certain binomial coefficients were zero. The coefficient of $\tau_0^\epsilon\xi_1^i$ in this sum is

$$\sum_{1\le k\le n}\binom{n}{k}\binom{i+\epsilon-1}{k-i-1}=\sum_{1\le k\le n}\binom{n}{k}\binom{i+\epsilon-1}{2i+\epsilon-k}=\binom{n+i+\epsilon-1}{2i+\epsilon};$$

here the first equality uses the standard identity $\binom{a}{b} = \binom{a}{a-b}$, and the second uses Vandermonde's identity. Adding in a sufficient number of ρ's and τ's, and converting to $\Lambda[1]$, we learn

$$\delta(\tau^n) = \sum_{\substack{\epsilon\in\{0,1\},\, i\geq 0 \\ (i,\epsilon)\neq(0,0)}} \lambda_{2i+\epsilon-1}\binom{n+i+\epsilon-1}{2i+\epsilon}\tau^{n-i-\epsilon}\rho^{2i+\epsilon}.$$

Set $r = 2i+\epsilon-1$. Then $\lfloor \frac{1}{2}r \rfloor = i+\epsilon-1$, leading to the given description. □

2.3.4 Lift of Sq^0 The dual motivic Steenrod algebra $\mathcal{A}^\vee$ is a commutative Hopf algebroid, and thus its cohomology, which agrees by definition with $\mathrm{Ext}_{\mathcal{A}}(\mathbb{M},\mathbb{M})$, is equipped with algebraic Steenrod operations [Bruner 1986a]. The purpose of this section is to lift the operation Sq^0 to an endomorphism of Λ. Our approach essentially follows the proof of [Palmieri 2007, Proposition 1.4].

Let $C(\mathcal{A}) = C_{\mathcal{A}}(\mathbb{M},\mathbb{M})$ denote the cobar complex of the algebra $\mathcal{A}$; this is by definition the same as the cobar complex of the Hopf algebroid $\mathcal{A}^\vee$. As $\mathcal{A}^\vee$ is a commutative ring, $C(\mathcal{A})$ is the Moore complex of a cosimplicial commutative ring, and the levelwise Frobenius on this cosimplicial commutative ring induces a map

$$\sigma\colon C(\mathcal{A}) \to C(\mathcal{A}).$$

This is a map of differential graded algebras, and Sq^0 is the map induced by σ in homology.

Theorem 2.3.16 *The map $\sigma\colon C(\mathcal{A}) \to C(\mathcal{A})$ induced by the levelwise Frobenius descends to a map*

$$\theta\colon \Lambda \to \Lambda$$

of differential graded algebras. This map is given on generators by

$$\theta(x) = x^2 \quad \text{for } x\in\mathbb{M}, \qquad \theta(\lambda_{2n-1}) = \lambda_{4n-1}, \qquad \theta(\lambda_{2n}) = \lambda_{4n+1}\tau + \lambda_{4n+2}\rho.$$

Proof Recall $\mathcal{A}^{\mathrm{big}}$ and Λ^{big} from the proof of Proposition 2.3.11. Let $C(\mathcal{A}^{\mathrm{big}})$ be the cobar complex for $\mathcal{A}^{\mathrm{big}}$ with respect to augmentation of $\mathcal{A}^{\mathrm{big}}$, so that $H_*C(\mathcal{A}^{\mathrm{big}}) = \Lambda^{\mathrm{big}}$ as algebras. The levelwise Frobenius gives a map

$$\sigma\colon C(\mathcal{A}^{\mathrm{big}}) \to C(\mathcal{A}^{\mathrm{big}})$$

of differential graded algebras and, by taking homology, this induces a map

$$\theta'\colon \Lambda^{\mathrm{big}} \to \Lambda^{\mathrm{big}}$$

of algebras. We claim that to produce θ it suffices to show that θ' restricts to an endomorphism of $\Lambda \subset \Lambda^{\mathrm{big}}$ satisfying the given formulas. Indeed, there is an inclusion $C(\mathcal{A}) \subset C(\mathcal{A}^{\mathrm{big}})$ of algebras, which does not respect differentials but does commute with the levelwise Frobenius σ. It would thus follow that the restriction θ of θ' to Λ is induced by the levelwise Frobenius on $C(\mathcal{A})$. In particular, this would show that $\sigma\colon C(\mathcal{A}) \to C(\mathcal{A})$ indeed descends to an algebra map $\theta\colon \Lambda \to \Lambda$. That θ moreover respects the differential is inherited from σ.

To understand θ', it suffices to understand its effect on the generators of Λ^{big}, ie to understand the map

$$\theta'\colon \Lambda^{\mathrm{big}}[1] \to \Lambda^{\mathrm{big}}[1].$$

Recall that $\Lambda^{\mathrm{big}}[1] = (\mathcal{A}^{\mathrm{big}}[1])^\vee = \mathcal{A}^\vee_{\leq 1}$. This ring was described in Lemma 2.3.13, and θ' acts on it by the Frobenius. We find that θ' satisfies the same formulas as described for θ, only with the addition that $\theta'(\lambda_{-1}) = \lambda_{-1}$. In particular, θ' does restrict to Λ, and this restriction satisfies the stated formulas. $\square$

2.4 Summary

For ease of reference, let us summarize what we have learned in one place. As always, F is a base field of characteristic not equal to 2.

2.4.1 Generators There is a differential graded algebra Λ^F, the *F-motivic lambda algebra*, together with a multiplicative quasiisomorphism $C(\mathcal{A}^F) \to \Lambda^F$, where $C(\mathcal{A}^F)$ is the reduced cobar complex of $\mathcal{A}^F$. We write $\Lambda^F = \bigoplus_{m\geq 0} \Lambda^F[m]$, where the differential on Λ^F is of the form $\delta\colon \Lambda^F[m] \to \Lambda^F[m+1]$.

The F-motivic lambda algebra Λ^F is an $\mathbb{M}^F$-bimodule algebra, generated by classes $\lambda_r \in \Lambda^F[1]$ for $r \geq 0$. In the trigrading (stem, filtration, weight), we have

$$|\tau| = (0,0,-1), \quad |\rho| = (-1,0,-1), \quad |\lambda_a| = \left(a, 1, \left\lceil \tfrac{1}{2}a \right\rceil\right).$$

A right $\mathbb{M}^F$-module basis of Λ^F is given by those $\lambda_{r_1}\cdots\lambda_{r_n}$ with $2r_i \geq r_{i+1}$ for $1 \leq i \leq n-1$.

2.4.2 Relations The F-motivic lambda algebra is a quadratic $\mathbb{M}^F$-bimodule algebra. Recall that $\mathbb{M}^F = \mathbb{M}^F_0[\tau]$. The $\mathbb{M}^F$-bimodule structure of Λ^F is determined by

$$x\lambda_n = \lambda_n x \quad \text{for } x \in \mathbb{M}^F_0, \qquad \tau\lambda_{2n+1} = \lambda_{2n+1}\tau + \lambda_{2n+2}\rho, \qquad \tau\lambda_{2n} = \lambda_{2n}\tau + \lambda_{2n+1}\tau\rho + \lambda_{2n+2}\rho^2,$$

and the quadratic relations are given as follows. Fix $a, b \geq 0$. If a is odd or b is even, then

$$\lambda_a\lambda_{2a+b+1} = \sum_{0\leq r<b/2} \lambda_{a+b-r}\lambda_{2a+1+r}\binom{b-r-1}{r};$$

and if a is even and b is odd, then

$$\lambda_a\lambda_{2a+b+1} = \sum_{0\leq r<b/2} \lambda_{a+b-r}\lambda_{2a+1+r}\binom{b-r-1}{r}\tau^{(r-1) \bmod 2} + \sum_{0\leq r\leq \lceil b/2\rceil} \lambda_{a+b+1-r}\lambda_{2a+1+r}\binom{\lfloor \frac{1}{2}b\rfloor - \lfloor \frac{1}{2}r\rfloor}{\lfloor \frac{1}{2}r\rfloor}\rho.$$

2.4.3 Differentials The differential on Λ is determined by the Leibniz rule, together with

$$\delta(x) = 0 \quad \text{for } x \in \mathbb{M}^F_0, \quad \delta(\tau) = \lambda_0\rho, \quad \delta(\lambda_n) = \sum_{1\leq r\leq n/2} \lambda_{n-r}\lambda_{r-1}\binom{n-r}{r}.$$

Moreover, we have

$$\delta(\tau^n) = \sum_{r\geq 0} \lambda_r \binom{n + \lfloor \frac{1}{2}r\rfloor}{r+1} \tau^{n-\lfloor r/2\rfloor - 1}\rho^{r+1}.$$

2.4.4 The endomorphism θ The squaring operation $\mathrm{Sq}^0\colon \mathrm{Ext}_F^{s,f,w} \to \mathrm{Ext}_F^{2s+f,f,w+f}$ lifts to an endomorphism $\theta\colon \Lambda^F \to \Lambda^F$ of differential graded algebras, determined by

$$\theta(x) = x^2 \quad \text{for } x \in \mathbb{M}^F, \qquad \theta(\lambda_{2n-1}) = \lambda_{4n-1}, \qquad \theta(\lambda_{2n}) = \lambda_{4n+1}\tau + \lambda_{4n+2}\rho.$$

3 Some first examples, and the doubling map

3.1 First examples

Before continuing on, we give some basic examples illustrating the form of the motivic lambda algebra. In particular, we use Λ^F to define some classes in Ext_F, and reprove some well-known low-dimensional relations. This material is meant only to familiarize the reader with Λ^F; we give a more thorough and entirely self-contained investigation in Section 4.

Given a cycle $z \in \Lambda^F$, in this section we write $[z] \in \mathrm{Ext}_F$ for the corresponding cohomology class.

Lemma 3.1.1 *We have $\delta(\lambda_{2^a-1}) = 0$ for all $a \geq 0$.*

Proof The proof is identical to the proof of [Wang 1967, Proposition 2.2]. □

This allows us to define the following Hopf elements.

Definition 3.1.2 Let $h_a := [\lambda_{2^a-1}]$. ◁

Lemma 3.1.3 *If $\rho = 0$ in $\mathbb{M}^F$, such as when F is algebraically closed, then $\delta(\tau^n) = 0$ for all $n \geq 0$.*

Proof This is immediate from the differential $\delta(\tau) = \lambda_0\rho$. □

In general, if ρ is nilpotent in $\mathbb{M}^F$, then various powers of τ will be cycles in Λ^F. We shall write τ^n in place of $[\tau^n]$ in this case. We begin by considering some examples in the case where F is algebraically closed.

Proposition 3.1.4 *For F algebraically closed, there is a relation*

$$\tau \cdot h_1^3 = h_2 h_0^2.$$

Proof By definition, $\tau \cdot h_1^3 = [\lambda_1^3\tau]$ and $h_2h_0^2 = h_0^2h_2 = [\lambda_0^2\lambda_3]$. We have

$$\lambda_0^2\lambda_3 = \lambda_1^3\tau,$$

so these classes coincide in Ext_F. □

Proposition 3.1.5 *For F algebraically closed, there is a relation*

$$\tau \cdot h_1^4 = 0.$$

However, $h_1^n \neq 0$ for any n.

Proof Observe that $\lambda_0\lambda_1 = 0$, and thus $h_1 h_0 = 0$. Combined with Proposition 3.1.4, we find

$$\tau \cdot h_1^4 = \tau h_1^3 \cdot h_1 = h_2 h_0^2 \cdot h_1 = 0.$$

Alternatively, $\tau h_1^4 = [\lambda_1^4 \tau]$, and there is a differential

$$\delta(\lambda_2^2 \lambda_1) = \lambda_1^4 \tau.$$

On the other hand, for h_1^n to vanish, the class λ_1^n must be nullhomotopic, ie $\delta(x) = \lambda_1^n$ for some $x \in \Lambda$. The class x must live in stem $n+1$, weight n, and filtration $n-1$, and in this degree Λ is generated by the cycle $\lambda_3 \lambda_1^{n-2}$. So no such x exists. □

Next we consider some examples relevant to base fields F over which ρ does not vanish. We begin by defining some classes. Note that the differential

$$\delta(\tau) = \lambda_0 \rho$$

implies that $\delta(\tau^{2^n}) \equiv 0 \pmod{\rho^{2^n}}$. This allows for the following definition.

Definition 3.1.6 If $F = \mathbb{R}$, then

$$\tau^{2^{a-1}} h_a := \left[\frac{1}{\rho^{2^a}} \delta(\tau^{2^a}) \right]$$

for $a \geq 1$. In general, $\tau^{2^{a-1}} h_a \in \mathrm{Ext}_F$ is defined by pushing these classes forward along the map $\Lambda^{\mathbb{R}} \to \Lambda^F$ induced by $\mathbb{M}^{\mathbb{R}} \to \mathbb{M}^F$ (see Remark 2.2.8). ◁

Remark 3.1.7 Following our convention that Λ^F is considered primarily as a right $\mathbb{M}^F$-module, it would be more natural to write $h_a \tau^{2^{a-1}}$ for the classes introduced above. We have chosen instead to work with naming conventions compatible with those in [Belmont and Isaksen 2022], as no confusion should arise. ◁

Remark 3.1.8 If $\tau^{2^{a-1}}$ is a cycle in Ext_F, then $\tau^{2^{a-1}} h_a = \tau^{2^{a-1}} \cdot h_a$. ◁

Example 3.1.9 We have

$$\tau h_1 = [\lambda_1 \tau + \lambda_2 \rho], \quad \tau^2 h_2 = [\lambda_3 \tau^2 + \lambda_5 \tau \rho^2 + \lambda_6 \rho^3], \quad \tau^4 h_3 = [\lambda_7 \tau^4 + \lambda_{11} \tau^2 \rho^4 + \lambda_{13} \tau \rho^6 + \lambda_{14} \rho^7].$$

In fact, we may identify $\tau^{\lfloor 2^{a-1} \rfloor} h_a = [\tau^{2^a} \lambda_{2^a - 1}]$ for all $a \geq 1$. ◁

The following relation was proved over $\mathbb{R}$ by Dugger and Isaksen [2017a, Proof of Lemma 6.2] using Massey products and May's convergence theorem. We may use the lambda algebra to provide an explicit direct proof.

Proposition 3.1.10 *There is a relation*

$$(h_0 + \rho h_1) \cdot \tau h_1 = 0.$$

Proof By definition,

$$h_0 \cdot \tau h_1 = [\lambda_0(\lambda_1 \tau + \lambda_2 \rho)], \quad \rho h_1 \cdot \tau h_1 = [\rho \lambda_1(\lambda_1 \tau + \lambda_2 \rho)].$$

Expanding, we have

$$\lambda_0(\lambda_1\tau+\lambda_2\rho)=\lambda_1^2\tau\rho+\lambda_1\lambda_2\rho^2+\lambda_2\lambda_1\rho^2,\quad \rho h_1(\lambda_1\tau+\lambda_2\rho)=\lambda_1^2\tau\rho+\lambda_1\lambda_2\rho^2.$$

But

$$\delta(\lambda_3\tau\rho+\lambda_4\rho^2)=\lambda_2\lambda_1\rho^2,$$

so $h_0\cdot\tau h_1=\rho h_1\cdot\tau h_1$. The proposition follows. □

The fact that $\delta(\tau^n)\equiv 0 \pmod{\rho}$ allows for the following definition.

Definition 3.1.11 If $F=\mathbb{R}$, then

$$\tau^{2n}h_0:=\Big[\frac{1}{\rho}\delta(\tau^{2n+1})\Big].$$

In general, $\tau^{2n}h_0\in\mathrm{Ext}_F$ is defined by pushing these classes forward along the map $\Lambda^{\mathbb{R}}\to\Lambda^F$ induced by $\mathbb{M}^{\mathbb{R}}\to\mathbb{M}^F$ (see Remark 2.2.8). ◁

Example 3.1.12 We have

$$\begin{aligned}h_0&=[\lambda_0],\\ \tau^2h_0&=[\lambda_0\tau^2+\lambda_1\tau^2\rho+\lambda_3\tau\rho^3+\lambda_4\rho^4],\\ \tau^4h_0&=[\lambda_0\tau^4+\lambda_3\tau^3\rho^3+\lambda_4\tau^2\rho^4+\lambda_5\tau^2\rho^5+\lambda_7\tau\rho^7+\lambda_8\rho^8].\end{aligned}$$ ◁

The following proposition was originally proved over $\mathbb{R}$ by Dugger and Isaksen [2017a, Proof of Lemma 6.2] using Massey products, May's convergence theorem, and analysis of the ρ-Bockstein spectral sequence. Using the lambda algebra, the proof amounts to checking that the products of cycle representatives are equal.

Proposition 3.1.13 *There is a relation*

$$\tau^2h_0\cdot h_1=\rho(\tau h_1)^2.$$

Proof We may directly compute

$$\begin{aligned}\tau^2h_0\cdot h_1&=[(\lambda_0\tau^2+\lambda_1\tau^2\rho+\lambda_3\tau\rho^3+\lambda_4\rho^4)\lambda_1]\\ &=[\lambda_1^2\tau^2\rho+\lambda_2\lambda_1\tau\rho^2+\lambda_2^2\rho^3+\lambda_2\lambda_3\rho^4]\\ &=[\rho(\lambda_1\tau+\lambda_2\rho)^2]=\rho(\tau h_1)^2.\end{aligned}$$ □

3.2 The doubling map

Dugger and Isaksen [2017a, Theorem 4.1] produce an isomorphism

$$\mathrm{Ext}_{\mathrm{cl}}[\rho^{\pm1}]\cong\mathrm{Ext}_{\mathbb{R}}[\rho^{-1}],$$

which doubles internal degrees. We can lift this isomorphism to a quasiisomorphism of lambda algebras.

Proposition 3.2.1 *Let Λ^{dcl} denote the classic lambda algebra, only given a motivic grading where $|\lambda_n|$ has stem $2n+1$ and weight $n+1$. For any F, there is a retraction*

$$\Lambda^{\mathrm{dcl}} \xrightarrow{\widetilde{\theta}} \Lambda^F \to \Lambda^{\overline{F}} \xrightarrow{q} \Lambda^{\mathrm{dcl}}$$

with the following properties:

1. *All maps shown are maps of differential graded algebras respecting θ.*
2. *$\widetilde{\theta}$ is given on generators by $\widetilde{\theta}(\lambda_n) = \lambda_{2n+1}$.*
3. *q is given on generators by $q(\tau) = 0$, $q(\lambda_{2n}) = 0$, and $q(\lambda_{2n+1}) = \lambda_n$.*

Now say $F = \mathbb{R}$, and write $\mathrm{Ext}_{\mathbb{R}}^{\rho\text{-tors}} \subset \mathrm{Ext}_{\mathbb{R}}$ for the ρ-torsion subgroup of $\mathrm{Ext}_{\mathbb{R}}$.

4. *The map $\mathrm{Ext}_{\mathrm{dcl}}[\rho] \oplus \mathrm{Ext}_{\mathbb{R}}^{\rho\text{-tors}} \to \mathrm{Ext}_{\mathbb{R}}$ induced by $\widetilde{\theta}$ and the inclusion of ρ-torsion is an isomorphism.*
5. *In particular, $\widetilde{\theta}$ extends to a quasiisomorphism $\Lambda^{\mathrm{dcl}} \otimes_{\mathbb{F}_2} \mathbb{F}_2[\rho^{\pm 1}] \to \Lambda^{\mathbb{R}}[\rho^{-1}]$.*

Proof The assignments given in (2) and (3) are easily seen to extend to maps of differential graded algebras, proving (1), and that the resulting sequence is a retraction is clear. Evidently (4) implies (5), so we are left with proving (4).

It is equivalent to verify that the composite $\mathrm{Ext}_{\mathrm{dcl}}[\rho] \to \mathrm{Ext}_{\mathbb{R}} \to \mathrm{Ext}_{\mathbb{R}} / \mathrm{Ext}_{\mathbb{R}}^{\rho\text{-tors}}$ is an isomorphism. This is a split inclusion of free $\mathbb{F}_2[\rho]$-modules, so for it to be an isomorphism it is sufficient to verify that it is an isomorphism after inverting ρ, and for this it is sufficient for the injection $\mathrm{Ext}_{\mathrm{dcl}}[\rho^{\pm 1}] \to \mathrm{Ext}_{\mathbb{R}}[\rho^{-1}]$ to be an isomorphism. By Dugger and Isaksen's isomorphism [2017a, Theorem 4.1] $\mathrm{Ext}_{\mathbb{R}}[\rho^{-1}] \cong \mathrm{Ext}_{\mathrm{dcl}}[\rho^{\pm 1}]$, we find that our map $\mathrm{Ext}_{\mathrm{dcl}}[\rho^{\pm 1}] \to \mathrm{Ext}_{\mathbb{R}}[\rho^{-1}]$ is an injection between vector spaces of equal finite rank in each degree, and is thus an isomorphism. □

Remark 3.2.2 Proposition 3.2.1 has the following amusing corollary: there is a multiplicative injection

$$Q : \ker(\mathrm{Sq}^0 : \mathrm{Ext}_{\mathrm{cl}} \to \mathrm{Ext}_{\mathrm{cl}}) \to \mathrm{Ext}_{\mathbb{C}}^{\tau\text{-tors}},$$

acting in degrees as Sq^0 would. For example, as $\widetilde{\theta}\lambda_0^n = \lambda_1^n$, we find that $Q(h_0^n) = h_1^n$. This provides another explanation of the fact that h_1 is not nilpotent in $\mathrm{Ext}_{\mathbb{C}}$. It is natural to ask whether Q accounts for all indecomposable τ-torsion classes in $\mathrm{Ext}_{\mathbb{C}}$, but a counterexample is given by the class B_6 in stem 55 and filtration 7, as $\mathrm{Ext}_{\mathrm{cl}}^{24,7} = 0$. ◁

4 $\mathrm{Ext}_{\mathbb{R}}$ in filtrations $f \le 3$

In this section, we use the $\mathbb{R}$-motivic lambda algebra to compute $\mathrm{Ext}_{\mathbb{R}}^f$ for $f \le 3$. Throughout this section, we shall abbreviate

$$\Lambda = \Lambda^{\mathbb{R}}.$$

4.1 Preliminaries

We begin by describing our strategy for computing $\mathrm{Ext}_{\mathbb{R}}$. We rely on the following device, which uses ideas from Tangora's work [1985] on the classic lambda algebra to produce something like a chain-level lift of the ρ-Bockstein spectral sequence [Hill 2011]. While the algorithm is essentially standard, we include a detailed description since we were unable to find a reference with the algorithm in precisely the form we need in the sequel. We begin with some preliminary definitions.

Definition 4.1.1 Let $V = \mathbb{F}_2\{x_s : s \in S\}$ be a (locally) finite $\mathbb{F}_2$-vector space with ordered basis.

(1) The *leading term* of a class $x \in V$ is the largest term appearing when x is written as a sum of basis elements.

(2) We write $x < x'$ when the leading term of x is less than that of x'.

(3) Given another vector space $U = \mathbb{F}_2\{x_s : s \in T\}$ with ordered basis, map $\phi\colon V \to U$, and $s \in S$ and $t \in T$, we write
$$\phi(x_s + <) = y_t + <$$
for the *relation* that there exist some classes $u < x_s$ and $v < y_t$ for which $\phi(x_s + u) = y_t + v$. $\triangleleft$

The main technical lemma we need is the following. The reader is invited to skip this lemma on first reading; the details are not necessary to understand our computation, and we rephrase what we need in the context of Λ in Theorem 4.1.4.

Lemma 4.1.2 *Let (C, d) be a chain complex of locally finite and free $\mathbb{F}_2[\rho]$-modules, and suppose (for simplicity) that $H_*C[\rho^{-1}] = 0$. Choose an ordered basis $\mathbb{F}_2\{x_s : s \in T\}$ for $C/(\rho)$, and extend this to a basis $\mathbb{F}_2\{\rho^n x_s : (s,n) \in T \times \mathbb{N}\}$ for C, itself ordered by $\rho^n x_s < \rho^m x_t$ whenever $n > m$, or else $n = m$ and $s < t$. Let $\{\alpha_s : s \in B\}$ be a basis for $H_*(C/(\rho))$, indexed by a subset $B \subset T$ with the property that, for each α_s, there is some $z_s \in C$ with leading term x_s which projects to a cycle representative of α_s. Let $B_1 \subset B$ be the subset of those s for which x_s is the leading term of some cycle in C, and let $B_0 = B \setminus B_1$.*

There is then a unique injection $t\colon B_0 \to B$ such that
$$d(x_s + <) = \rho^{r(s)} x_{t(s)} + <$$
for all $s \in B_0$. Here $r(s) \geq 1$ is an integer uniquely determined by comparing the degrees of x_s and $x_{t(s)}$. Moreover, t restricts to a bijection $t\colon B_0 \cong B_1$, and there is an isomorphism
$$H_*C = \bigoplus_{s \in B_0} \mathbb{F}_2[\rho]/(\rho^{r(s)}),$$
where we may take the summand indexed by s to be generated by any class of the form $\rho^{-r(s)} \cdot d(x_s + <)$ with leading term $x_{t(s)}$.

Proof We begin by defining a function $t^{-1}\colon B_1 \to B$. Fix $b \in B_1$; we claim that there exists some $s \in B$ such that $d(x_s + <) = x_b + <$. The function t^{-1} will then be defined by declaring $t^{-1}(b)$ to be the minimal s for which $d(x_s + <) = x_b + <$.

Indeed, let z_b be a cycle with leading term x_b which projects to a cycle representative for α_b. As $H_*C[\rho^{-1}] = 0$, necessarily $\rho^r z_b$ is nullhomologous for some minimal $r \geq 1$. That is, there is some $y \in C$ not divisible by ρ such that $d(y) = \rho^r z_b$. If $y = x_s + <$ with $s \in B$, then we are done. Otherwise, as y is a cycle in $C/(\rho)$, necessarily y is homologous to some $x_s + u$ with $u < x_s$ and $s \in B$, in which case there exists some v with $d(v) = x_s + u + y$. We find that

$$d(x_s + <) = d(x_s + u) = d(x_s + u + d(v)) = d(y) = \rho^r z_b = \rho^r x_b + <,$$

as claimed. Thus we have produced the function t^{-1}.

Next we claim that t^{-1} restricts to a function $t^{-1}\colon B_1 \to B_0$. Indeed, suppose towards contradiction that there are some $b \in B_1$ such that $x_{t^{-1}(b)}$ is the leading term of some cycle. That is to say, suppose given $u, v < x_{t^{-1}(b)}$ such that

$$d(x_{t^{-1}(b)} + u) = x_b + <, \quad d(x_{t^{-1}(b)} + v) = 0.$$

Adding these together, we find

$$d(u + v) = x_b + <.$$

As $u + v < x_{t^{-1}(b)}$, this contradicts minimality of $t^{-1}(b)$. Thus we have a function $t^{-1}\colon B_1 \to B_0$.

Next we claim that t^{-1} is a bijection. It is a function between locally finite sets, and the assumption that $H_*C[\rho^{-1}] = 0$ implies that these sets have the same cardinality in each degree. So it is sufficient to verify that t^{-1} is an injection. Indeed, suppose towards contradiction that there were some $b < c$ in B_1 for which $t^{-1}(b) = s = t^{-1}(c)$. Thus there are $u, v < x_s$ such that

$$d(x_s + u) = x_b + <, \quad d(x_s + v) = x_c + <.$$

Adding these together, we find

$$d(u + v) = x_c + <.$$

As $u + v < x_s$, this contradicts minimality of $t^{-1}(c)$.

By taking the inverse of $t^{-1}\colon B_1 \to B_0$, we have thus proved the existence of a bijection $t\colon B_0 \to B_1$ with the property that $d(x_s + <) = x_{t(s)} + <$ for all $s \in B_0$. With this t, the given description of H_*C is clear; in effect, we have described how to choose a basis for C for which d is upper triangular, where, if a diagonal entry is divisible by ρ^r, so too are all entries above it. Compare the notion of a tag from [Tangora 1985].

It remains to verify uniqueness. Suppose towards contradiction that we have found some other injection $t'\colon B_0 \to B$ such that $d(x_s + <) = x_{t'(s)} + <$ for all $s \in B_0$. The condition that $t' \neq t$ means that there exists some $s \in B_0$ for which $d(x_s + <) = x_{t'(s)} + <$, but s is not minimal among possible $a \in B_0$ with $d(x_a + <) = x_{t'(s)} + <$. Choose such s with $t'(s)$ maximal, and let $a = t^{-1}(t'(s))$ be the minimal $a \in B_0$ with $d(x_a + <) = x_{t'(s)} + <$. So there are $u, v < x_a$ for which

$$d(x_a + u) = x_{t'(s)} + <, \quad d(x_a + v) = x_{t'(a)} + <.$$

Adding these together, we find that

$$d(u + v) = x_{t'(s)} + x_{t'(a)} + <,$$

where $u + v < x_a$. If $t'(a) < t'(s)$, then this reduces to

$$d(u + v) = x_{t'(s)} + <,$$

contradicting minimality of a. If $t'(s) < t'(a)$, then this reduces to

$$d(u + v) = x_{t'(a)} + <,$$

contradicting maximality of $t'(s)$. So there is no such t', proving that t is the unique injection satisfying the required property. □

We now specialize to the computation of $\mathrm{Ext}_{\mathbb{R}}$. Observe that by Proposition 3.2.1, we may reduce to considering only the ρ-torsion subgroup of $\mathrm{Ext}_{\mathbb{R}}$. In terms of Λ, this amounts to ignoring monomials of the form λ_I where I is a sequence of odd numbers. We will apply Lemma 4.1.2 to compute this ρ-torsion subgroup as follows.

We take as basis of $\Lambda/(\rho)$ the standard basis $\lambda_I \tau^n$ where I is coadmissible (Definition 2.3.9) and $n \geq 0$. We also need to order this basis. In the region where we will compute, our choice of order makes no difference, in the sense that all "error terms" appearing in "$+<$" will be divisible by ρ. But for concreteness let us say that $\lambda_I \tau^n < \lambda_J \tau^m$ if $n > m$, or else $n = m$ and $I < J$ lexicographically, ie if $I = (i_1, \dots, i_f)$ and $J = (j_1, \dots, j_f)$, then $i_1 < j_1$, or else $i_1 = j_1$ and $i_2 < j_2$, and so forth.

We must fix some further notation. Let $\{\alpha'_s : s \in S_0\}$ be a basis for $\mathrm{Ext}_{\mathrm{cl}}$, and write $\alpha_s \in \mathrm{Ext}_{\mathbb{C}}$ for the image of α'_s under the map induced by $\widetilde{\theta}\colon \Lambda^{\mathrm{dcl}} \to \Lambda^{\mathbb{C}}$ (see Proposition 3.2.1). Extend this to a minimal generating set $\{\alpha_s : s \in S\}$ for $\mathrm{Ext}_{\mathbb{C}}$ as an $\mathbb{F}_2[\tau]$-module. For $s \in S$, let n_s denote the τ-torsion exponent of α_s, so that $\{\alpha_s \tau^n : s \in S, n < n_s\}$ is an $\mathbb{F}_2$-basis for $\mathrm{Ext}_{\mathbb{C}}$. For each $s \in S$, choose a distinct coadmissible monomial $\lambda_{I(s)}$ which is the leading term of a cycle representative for α_s in $\Lambda_{\mathbb{C}}$, making this choice so that, if $s \in S_0$, then $\lambda_{I(s)}$ is in the image of $\widetilde{\theta}$. See the discussion following Proposition 4.2.1 for the particular choices we will take in our computation.

Let $B' = \{(s, n) : s \in S, n < n_s\}$. Given $b = (s, n) \in B'$, write $x_b = \lambda_{I(s)} \tau^n \in \Lambda^{\mathbb{R}}$. Let $B \subset B'$ be the subset of pairs not of the form $(s, 0)$ with $s \in S_0$. Let $B_1 \subset B$ be the subset of those b such that x_b is the leading term of some cycle, and let $B_0 = B \setminus B_1$. Let $B[f] \subset B$ be the subset of those b for which x_b is in filtration f, and extend this notation to all the indexing sets under consideration.

For our computation, we will produce, for every $b \in B_0[f]$ with $f \le 2$, some $t(b) \in B$ such that

$$\delta(x_b + <) = \rho^{r(b)} x_{t(b)} + <,$$

making this choice so that $t\colon B_0 \to B$ is injective. Here $r(b) \ge 1$ is some integer which may be determined by comparing the stems of x_b and $x_{t(b)}$.

Definition 4.1.3 In the above situation, we shall write $x_b \to x_{t(b)}\rho^{r(b)}$. ◁

Theorem 4.1.4 *Fix notation as above. Then*:

1. *t is uniquely determined* (*given our choice of ordered basis*).
2. *t restricts to bijections $t\colon B_0[f] \cong B_1[f+1]$.*
3. *The ρ-torsion subgroup of $\mathrm{Ext}_{\mathbb{R}}^{f+1}$ is isomorphic to*
$$\bigoplus_{b \in B_0[f]} \mathbb{F}_2[\rho]/(\rho^{r(b)}),$$
where the summand corresponding to $b \in B_0[f]$ is generated by any class of the form
$$\frac{\delta(x_b + <)}{\rho^{r(b)}}$$
with leading term $x_{t(b)}$.

Proof This follows by specializing Lemma 4.1.2 to the complementary summand of $\widetilde{\theta}\colon \Lambda^{\mathrm{cl}} \subset \Lambda$. □

Most notably, the ρ-torsion in $\mathrm{Ext}_{\mathbb{R}}^{f+1}$ is obtained by understanding differentials out of $\Lambda[f]$; this is significantly easier than finding cycles in $\Lambda[f+1]$ directly.

We end with two remarks, which could have been made in the more general context of Lemma 4.1.2.

Remark 4.1.5 More generally, $H^*(\mathcal{A}^{\mathbb{R}}/(\rho^m)) = H_*(\Lambda/(\rho^m))$ (denoted by $\mathrm{Ext}_{(m)}$ in Section 7) may be read off our computation as follows. For each $b \in B_0$, choose $u_b \in \Lambda$ such that $u_b < x_b$ and $\delta(x_b + u_b) = \rho^{r(b)} x_{t(b)} + <$, and let $z_b = \rho^{-r(b)} \cdot \delta(x_b + u_b)$. Then $H_*(\Lambda/(\rho^m))$ is given as follows:

1. For each $s \in S_0$, there is a summand of the form $\mathbb{F}_2[\rho]/(\rho^m)$, generated by the image of α_s.
2. For each $x_b \to \rho^{r(b)} x_{t(b)}$, there is a summand of the form $\mathbb{F}_2[\rho]/(\rho^{\min(m,r(b))})$, generated by the class with cycle representative z_s.
3. For each $x_b \to \rho^{r(b)} x_{t(b)}$, there is a summand of the form $\mathbb{F}_2[\rho]/(\rho^{m-\max(0,m-r(b))})$, generated by the class with cycle representative $\rho^{\max(0,m-r(b))}(x_b + u_b)$. ◁

Remark 4.1.6 Our approach to computing $\mathrm{Ext}_{\mathbb{R}}$ via Λ is closely related to the computation of $\mathrm{Ext}_{\mathbb{R}}$ via the ρ-Bockstein spectral sequence $\mathrm{Ext}_{\mathbb{C}}[\rho] \Rightarrow \mathrm{Ext}_{\mathbb{R}}$ [Hill 2011]. The precise relation is as follows. For $b = (s,n) \in B$, let $\alpha_b = \alpha_s \tau^n$, so that $\{\alpha_b : b \in B\}$ is a basis of $\mathrm{Ext}_{\mathbb{C}}$. Our ordering on Λ and choice of classes x_b gives B an order, thus making this into an ordered basis of $\mathrm{Ext}_{\mathbb{C}}$. Now, $x_b \to \rho^{r(b)} x_{t(b)}$ if and only if $d_{r(b)}(\alpha_b + <) = \rho^{r(b)} \alpha_{t(b)} + <$ in the ρ-Bockstein spectral sequence. ◁

The above discussion describes how we will compute $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$ as an $\mathbb{F}_2[\rho]$-module. The computation gives more, as it produces explicit cocycle representatives for our generators of $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$. We will use this in Section 4.3 to compute products in $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$.

4.2 $\mathrm{Ext}_{\mathbb{R}}^{f}$ for $f \leq 3$

We now proceed to the computation. We begin by understanding $\Lambda^{\mathbb{R}}/(\rho) \cong \Lambda^{\mathbb{C}}$.

Proposition 4.2.1 $\mathrm{Ext}_{\mathbb{C}}^{\leq 3}$ *is generated as a commutative* $\mathbb{F}_2[\tau]$*-algebra by classes* h_a *for* $a \geq 0$, *represented in* $\Lambda^{\mathbb{C}}$ *by* λ_{2^a-1}, *and* c_a *for* $a \geq 0$, *represented in* $\Lambda^{\mathbb{C}}$ *by* $\lambda_{2^a 3-1}\lambda^2_{2^{a+2}-1}$. *A full set of relations is given by*

$$h_{a+1}h_a = 0, \quad h_{a+2}^2 h_a = 0, \quad h_2 h_0^2 = \tau h_1^3, \quad h_{a+3}h_{a+1}^2 = h_{a+2}^3$$

for all $a \geq 0$. *This is free over* $\mathbb{F}_2[\tau]$, *with basis given by the classes in the following table*:

class	*constraints*
1	
h_a	$a \geq 0$
$h_a \cdot h_b$	$a \geq b \geq 0$ and $a \neq b+1$
$h_a \cdot h_b \cdot h_c$	$a \geq b \geq c \geq 0$ with $a \neq b+1$, $b \neq c+1$ and, if $b = c$ or $a = b$, then $a \neq c+2$
c_a	$a \geq 0$

The only such classes not in the image of $\widetilde{\theta}\colon \mathrm{Ext}_{\mathrm{dcl}} \to \mathrm{Ext}_{\mathbb{C}}$ *are those in which either* h_0 *or* c_0 *appears.*

Proof This is essentially well known, owing to work of Isaksen [2019] on the cohomology of the $\mathbb{C}$-motivic Steenrod algebra. Alternatively, one may compute $H_{\leq 3}(\Lambda^{\mathbb{C}}/(\tau))$ following Wang's approach [1967], and run the τ-Bockstein spectral sequence to recover $\mathrm{Ext}_{\mathbb{C}}^{\leq 3}$. One finds that $H_{\leq 3}(\Lambda^{\mathbb{C}}/(\tau))$ agrees with $\mathrm{Ext}_{\mathrm{cl}}^{\leq 3}$, with two exceptions:

(1) Instead of $h_0^2 \cdot h_2 = h_1^3$, one has $h_0^2 \cdot h_2 = 0$.

(2) There is a new cycle α represented by $\lambda_2^2\lambda_1$.

There is a τ-Bockstein differential $d_1(\alpha) = \tau h_1^4$, after which we recover the claimed $\mathbb{F}_2[\tau]$-module basis of $\mathrm{Ext}_{\mathbb{C}}^{\leq 3}$. The hidden extension $h_0^2 \cdot h_2 = \tau h_1^3$ was shown in Proposition 3.1.4; alternatively, it is the only relation compatible with $\mathrm{Sq}^0(h_0^2 \cdot h_2) = \tau^2 h_1^2 h_3 = \tau^2 h_2^3 = \mathrm{Sq}^0(\tau h_1^3)$. □

Proposition 4.2.1 describes a basis for $\mathrm{Ext}_{\mathbb{C}}^{\leq 3}$, thus giving our set $S[\leq 3]$. We must also choose lambda algebra representatives of these classes. We shall choose c_n to be represented by $\lambda_{2^n 3-1}\lambda^2_{2^{n+2}-1}$ and a product $h_{n_1}\cdots h_{n_k}$ with $n_1 \geq \cdots \geq n_k$ to be represented by $\lambda_{2^{n_1}-1}\cdots\lambda_{2^{n_k}-1}$. We warn that these representatives are not minimal; for example, we have chosen $\lambda_3\lambda_0$ as our representative for h_2h_0, rather than the minimal representative $\lambda_2\lambda_1$. However, they are easily defined and convenient enough for our computation.

The following identity will be used frequently in consolidating various cases in our computation. It is an immediate consequence of the description of θ given in Theorem 2.3.16.

Lemma 4.2.2 *We have*

$$\theta^a(\lambda_0\tau^n) = \lambda_{2^a-1}\tau^{\lfloor 2^{a-1}(2n+1)\rfloor} + O(\rho^{2^{a-1}})$$

for all $n \geq 0$, the error term being omitted when $a = 0$. □

Remark 4.2.3 Explicitly,

$$\lfloor 2^{a-1}(2n+1)\rfloor = \begin{cases} 2^{a-1}(2n+1) & \text{if } a \geq 1, \\ n & \text{if } a = 0. \end{cases}$$

This sort of pattern appears frequently throughout our computation, as a consequence of Lemma 4.2.2. ◁

We now produce the relation "→" described in Definition 4.1.3, proceeding filtration by filtration. To start, observe that $B_0[0] = \{\tau^n : n \geq 1\}$.

Proposition 4.2.4 *We have*

$$\delta(\tau^{2^a(2m+1)}) = \lambda_{2^a-1}\tau^{\lfloor 2^{a-1}(4m+1)\rfloor}\rho^{2^a} + O(\rho^{\lceil 2^a+2^{a-1}\rceil})$$

for all $a, m \geq 0$. In particular,

$$\tau^{2^a(2m+1)} \to \lambda_{2^a-1}\tau^{\lfloor 2^{a-1}(4m+1)\rfloor}\rho^{2^a}.$$

Proof When $a = 0$, as τ^2 is a cycle mod ρ^2, we may compute

$$\delta(\tau^{2m+1}) = \delta(\tau)\tau^{2m} + O(\rho^2) = \lambda_0\tau^{2m}\rho + O(\rho^2),$$

as claimed. By Lemma 4.2.2, applying θ^a for $a \geq 1$ to this yields

$$\begin{aligned}\delta(\tau^{2^a(2m+1)}) &= (\lambda_{2^a-1}\tau^{2^{a-1}(4m+1)} + O(\rho^{2^{a-1}}))\rho^{2^a} + O(\rho^{2^{a+1}}) \\ &= \lambda_{2^a-1}\tau^{\lfloor 2^{a-1}(4m+1)\rfloor}\rho^{2^a} + O(\rho^{2^a+2^{a-1}}).\end{aligned}$$

Combining the cases $a = 0$ and $a \geq 1$ yields the proposition. □

Corollary 4.2.5 *The set $B_0[1]$ consists of those $\lambda_{2^a-1}\tau^n$ such that n is not of the form $2^{a-1}(4m+1)$ for any m.* □

We have located the following indecomposable classes.

Definition 4.2.6 For $a, n \geq 0$, we declare

$$\tau^{\lfloor 2^{a-1}(4n+1)\rfloor}h_a$$

to be the class represented by

$$\rho^{-2^a}\cdot\delta(\tau^{2^a(2n+1)}).$$ ◁

We now compute out of $B_0[1]$.

Proposition 4.2.7 *For the combinations of a and b below, we have $\lambda_{2^b-1}\tau^{2^a(2m+1)} \to$ the following monomial:*

row	case	target
(1)	$a < b-1$ or $a = b$	$\lambda_{2^b-1}\lambda_{2^a-1}\tau^{\lfloor 2^{a-1}(4m+1)\rfloor}\rho^{2^a}$
(2)	$a > b+1$ and $b \neq 0$	$\lambda_{2^a-1}\lambda_{2^b-1}\tau^{\lfloor 2^{a-1}(4m+1)\rfloor}\rho^{2^a}$
(3)	$a = b-1$ and $m = 2n+1$	$\lambda_{2^b-1}^2\tau^{2^b(4n+1)}\rho^{2^b}$
(4)	$a = b+1$ and $b \neq 0$	$\lambda_{2^{b+1}-1}^2\tau^{\lfloor 2^{b-1}(8m+1)\rfloor}\rho^{2^b 3}$

Moreover, these cases are mutually exclusive and altogether exhaust $B_0[1]$.

Proof That these cases are mutually exclusive and altogether exhaust $B_0[1]$ is seen by direct inspection. As the monomials arising as targets are ρ-multiples of distinct elements of $B[2]$, it suffices to verify that for each claim of $x \to y$ we have $\delta(x + <) = y + <$.

(1) We have

$$\delta(\lambda_{2^b-1}\tau^{2^a(2m+1)}) = \lambda_{2^b-1}\lambda_{2^a-1}\tau^{\lfloor 2^{a-1}(4m+1)\rfloor}\rho^{2^a} + O(\rho^{2^a+2^{a-1}}).$$

(2) Note that

$$\tau^{2^a(2m+1)}\lambda_{2^b-1} = \lambda_{2^b-1}\tau^{2^a(2m+1)} + <,$$

as τ is central mod ρ. Now we have

$$\begin{aligned}\delta(\tau^{2^a(2m+1)}\lambda_{2^b-1}) &= (\lambda_{2^a-1}\tau^{\lfloor 2^{a-1}(4m+1)\rfloor}\rho^{2^a} + O(\rho^{2^a+2^{a-1}}))\lambda_{2^b-1}\\ &= \lambda_{2^a-1}\lambda_{2^b-1}\tau^{\lfloor 2^{a-1}(4m+1)\rfloor}\rho^{2^a} + O(\rho^{2^a+1}).\end{aligned}$$

(3) Note that

$$\theta^b(\lambda_0\tau^{2n+1}) = \lambda_{2^b-1}\tau^{\lfloor 2^{b-1}(4n+3)\rfloor} + O(\rho).$$

Now we have

$$\delta(\theta^b(\lambda_0\tau^{2n+1})) = \theta^b(\delta(\lambda_0\tau^{2n+1})) = \theta^b(\lambda_0^2\tau^{2n}\rho + O(\rho^2)) = \lambda_{2^b-1}^2\tau^{2^b(4n+1)}\rho^{2^b} + O(\rho^{2^b+1}).$$

(4) We have

$$\begin{aligned}\delta(\lambda_{2^b-1}\tau^{2^{b+1}(2m+1)}) &= \delta(\theta^{b-1}(\lambda_1\tau^{8m+4})) = \theta^{b-1}(\lambda_1\delta(\tau^4)\tau^{8m} + O(\rho^8))\\ &= \theta^{b-1}(\lambda_3^2\tau^{8m+1}\rho^6 + O(\rho^7)) = \lambda_{2^{b+1}-1}^2\tau^{2^{b-1}(8m+1)}\rho^{2^b 3} + O(\rho^{2^b 3+2^{b-1}}).\end{aligned}$$

Here the third equality uses the Adem relations $\lambda_1\lambda_3 = 0$ and $\lambda_1\lambda_5 = \lambda_3\lambda_3$ to determine the leading term of $\lambda_1\delta(\tau^4)$. □

Corollary 4.2.8 *The set $B_0[2]$ consists of those $\lambda_{2^b-1}\lambda_{2^c-1}\tau^n$ where $b = c$ or $b \geq c+2$, and where moreover:*

1. *$n \neq \lfloor 2^{b-1}(4m+1)\rfloor$ and $n \neq \lfloor 2^{c-1}(4m+1)\rfloor$ for any m.*
2. *If $b = c = 0$, then n is odd.*

(3) *If* $b = c \geq 1$, *then* $n \neq 2^b(4m+1)$ *for any* m.

(4) *If* $b = c \geq 2$, *then* $n \neq 2^{b-2}(8m+1)$ *for any* m. □

We have located the following indecomposable classes.

Definition 4.2.9 For $a, n \geq 0$, we declare

$$\tau^{2^a(8n+1)} h_{a+2}^2$$

to be the class represented by

$$\rho^{-2^{a+1}3} \cdot \delta(\lambda_{2^{a+1}-1} \tau^{2^{a+2}(2n+1)}) \qquad \triangleleft$$

We now compute out of $B_0[2]$.

Proposition 4.2.10 *For* $b = c$ *or* $b \geq c+2$, *we have* $\lambda_{2^b-1}\lambda_{2^c-1}\tau^{2^a(2m+1)} \to$ *the following monomial*:

#	case	target
(1)	$b = c = 0, a = -1, m = 2n+1$	$\lambda_0^3 \tau^{2n} \rho$
(2)	$b = c \geq 1, a = b-1, m = 2n+1$	$\lambda_{2^b-1}^3 \tau^{2^b(2n+1)} \rho^{2^b}$
(3)	$b = c \geq 0, a = c, m = 2n+1$	$\lambda_{2^b-1}^3 \tau^{\lfloor 2^{b-1}(4(2n+1)+1) \rfloor}$
(4)	$b = c \geq 1, a = b+1$	$\lambda_{2^{b-1}3-1} \lambda_{2^{b+1}-1}^2 \tau^{\lfloor 2^{b-2}(16m+1) \rfloor} \rho^{2^{b-1}7}$
(5)	$b = c \geq 1, a = b+2$	$\lambda_{2^{b+2}-1}^2 \lambda_{2^{b-1}-1} \tau^{\lfloor 2^{b-2}(2(16m+1)+1) \rfloor} \rho^{2^{b-1}13}$
(6)	$b = c \geq 1, a \geq b+3$	$\lambda_{2^a-1} \lambda_{2^b-1}^2 \tau^{\lfloor 2^{a-1}(4m+1) \rfloor} \rho^{2^a}$
(7)	$b = c \geq 2, a = b-2, m = 4n+2$	$\lambda_{2^b-1}^3 \tau^{\lfloor 2^{b-2}(2(4n+1)+1) \rfloor} \rho^{2^b}$
(8)	$b = c \geq 2, a = b-2, m = 2n+1$	$\lambda_{2^{b-2}3-1} \lambda_{2^b-1}^2 \tau^{\lfloor 2^{b-3}(2(4n+1)+1) \rfloor} \rho^{2^{b-2}3}$
(9)	$b = c \geq 3, a \leq b-3$	$\lambda_{2^b-1}^2 \lambda_{2^a-1} \tau^{\lfloor 2^{a-1}(4m+1) \rfloor} \rho^{2^a}$
(10)	$b-2 \geq c = 0, a = 0, m = 2n+1$	$\lambda_{2^b-1} \lambda_0^2 \tau^{2n} \rho$
(11)	$b-2 = c \geq 1, a = b$	$\tau^{2^c(8n+1)} \lambda_{2^c3-1} \lambda_{2^{c+2}-1}^2 \rho^{2^{c+1}3}$
(12)	$b-2 = c \geq 1, a \geq b+2$	$\lambda_{2^a-1} \lambda_{2^b-1} \lambda_{2^c-1} \tau^{\lfloor 2^{a-1}(4m+1) \rfloor} \rho^{2^a}$
(13)	$b-3 \geq c \geq 1, a \geq b, a \neq b+1$	$\lambda_{2^a-1} \lambda_{2^b-1} \lambda_{2^c-1} \tau^{\lfloor 2^{a-1}(4m+1) \rfloor} \rho^{2^a}$
(14)	$b-2 \geq c \geq 1, c \leq a < b, a \notin \{c+1, b-1\}$	$\lambda_{2^b-1} \lambda_{2^a-1} \lambda_{2^c-1} \tau^{\lfloor 2^{a-1}(4m+1) \rfloor} \rho^{2^a}$
(15)	$b-2 = c \geq 1, a = c-1, m = 2n+1$	$\lambda_{2^{c+1}-1}^3 \tau^{2^c(2n+1)} \rho^{2^c}$
(16)	$b-3 \geq c \geq 1, a = c-1, m = 2n+1$	$\lambda_{2^b-1} \lambda_{2^c-1}^2 \tau^{2^c(2n+1)} \rho^{2^c}$
(17)	$b-2 = c \geq 1, a = c+1, m = 2n+1$	$\lambda_{2^b-1}^3 \tau^{\lfloor 2^{b-3}(4(4n+1)+1) \rfloor} \rho^{2^{b-2}7}$
(18)	$b-3 = c \geq 1, a = c+1$	$\lambda_{2^{c+2}-1}^3 \tau^{\lfloor 2^{c-1}(8m+1) \rfloor} \rho^{2^c3}$
(19)	$b-4 \geq c \geq 1, a = c+1$	$\lambda_{2^b-1} \lambda_{2^{c+1}-1}^2 \tau^{\lfloor 2^{c-1}(8m+1) \rfloor} \rho^{2^c3}$
(20)	$b-3 \geq c \geq 1, a = b-1, m = 2n+1$	$\lambda_{2^b-1}^2 \lambda_{2^c-1} \tau^{2^b(2n+1)} \rho^{2^b}$
(21)	$b-2 \geq c \geq 1, a = b+1$	$\lambda_{2^{b+1}-1}^2 \lambda_{2^c-1} \tau^{\lfloor 2^{b-1}(8m+1) \rfloor} \rho^{2^b3}$
(22)	$b-2 \geq c \geq 2, a \leq c-2$	$\lambda_{2^b-1} \lambda_{2^c-1} \lambda_{2^a-1} \tau^{\lfloor 2^{a-1}(4m+1) \rfloor} \rho^{2^a}$

Moreover, these cases are mutually exclusive and altogether exhaust $B_0[2]$.

Proof That these cases are mutually exclusive and altogether exhaust $B_0[2]$ is seen by direct inspection. As the monomials arising as targets are ρ-multiples of distinct elements of $B[3]$, it suffices to verify that for each claim of $x \to y$ we have $\delta(x + <) = y + <$.

Each case represents a collection of families of monomials whose leading terms are connected by θ. Thus we may always reduce to the smallest possible c, except in cases (9) and (22), where doing so would place extra constraints on a. In addition, by working modulo the smallest power of ρ in which the proposed target does not vanish, we may always reduce to the smallest possible m.

We may further divide the list of cases provided into three types: those which require no calculations beyond those carried out in Proposition 4.2.7; cases (15) and (18); and the more interesting cases which do require additional calculation, producing new indecomposable classes in $\mathrm{Ext}^3_{\mathbb{R}}$. Here cases (15) and (18) are not really exceptional; they could be consolidated into cases (16) and (19), only this would require slightly modifying the setup of Section 4.1, and it is easier to just separate them out. The more interesting cases are (4), (5), (8), (11), and (17). The remaining less interesting cases may all be handled exactly the same way as the first two cases of Proposition 4.2.7 were handled. Thus we shall not handle them individually, and instead only illustrate this point with a verification of (21). With these reductions in place, the proposition is proved by the following calculations:

(4) Here we are claiming $\delta(\lambda_1^2\tau^4 + <) = \lambda_2\lambda_3^2\rho^7 + <$. In fact, $\delta(\lambda_1^2\tau^4) = \lambda_2\lambda_3^2\rho^7$ on the nose.

(5) Here we are claiming $\delta(\lambda_1^2\tau^8 + <) = \lambda_7^2\lambda_0\tau\rho^{13} + <$. Observe that $\delta(\lambda_1^2\tau^8) = \lambda_3^3\tau^4\rho^8 + O(\rho^{12})$, but $\lambda_3\tau^4\rho^4$ is already seen as a target in case (1). Thus some additional correction term must be added to $\lambda_1^2\tau^8$ to get down to $\lambda_7^2\lambda_0\tau\rho^{13}$. Such a correction term is given by

$$u = \lambda_3^2\tau^6\rho^4 + \lambda_3\lambda_5\tau^5\rho^6 + \lambda_3\lambda_6\tau^4\rho^7 + \lambda_5\lambda_7\tau^3\rho^{10} + (\lambda_5\lambda_8 + \lambda_6\lambda_7)\tau^2\rho^{11} + (\lambda_{11}\lambda_3 + \lambda_5\lambda_9)\tau^2\rho^{12} + (\lambda_8\lambda_7 + \lambda_7\lambda_8 + \lambda_6\lambda_9)\tau\rho^{13};$$

with this choice of u, we have $\delta(\lambda_1^2\tau^8 + u) = \lambda_7^2\lambda_0\tau\rho^{13} + O(\rho^{14})$.

(8) Here we are claiming $\delta(\lambda_3^2\tau^3 + <) = \lambda_2\lambda_3^2\tau\rho^3 + <$. Indeed, let

$$u = (\lambda_3\lambda_4 + \lambda_4\lambda_3)\tau^2\rho + \lambda_3\lambda_5\tau^2\rho^2 + \lambda_4\lambda_5\tau\rho^3;$$

then we have $\delta(\lambda_3^2\tau^3 + u) = \lambda_2\lambda_3^2\tau\rho^3 + O(\rho^4)$.

(11) Here we are claiming $\delta(\lambda_7\lambda_1\tau^8 + <) = \lambda_5\lambda_7^2\tau^2\rho^{12}$. Indeed, let

$$u = \lambda_9\lambda_7\tau^4\rho^8 + \lambda_9\lambda_{11}\tau^2\rho^{12};$$

then we have $\delta(\lambda_7\lambda_1\tau^8 + u) = \lambda_5\lambda_7^2\tau^2\rho^{12} + O(\rho^{14})$.

(15) Here we are claiming $\delta(\lambda_7\lambda_1\tau^3 + <) = \lambda_3^3\tau^2\rho^2 + <$. Indeed, let

$$u = \lambda_7\lambda_2\tau^2\rho + (\lambda_9\lambda_1 + \lambda_5\lambda_5)\tau^2\rho^2;$$

then we have $\delta(\lambda_7\lambda_1\tau^3 + <) = \lambda_3^3\tau^2\rho^2 + O(\rho^3)$.

(17) Here we are claiming $\delta(\lambda_7\lambda_1\tau^{12}+<)=\lambda_7^3\tau^5\rho^{14}+<$. Indeed, let

$$u=\lambda_{11}\lambda_3\tau^9\rho^6+(\lambda_{11}\lambda_4+\lambda_{12}\lambda_3+\lambda_7\lambda_8+\lambda_8\lambda_7)\tau^8\rho^7+\lambda_7^2\tau^9\rho^6+\lambda_9\lambda_7\tau^8\rho^8;$$

then we have $\delta(\lambda_7\lambda_1\tau^{12}+u)=\lambda_7^3\tau^5\rho^{14}+O(\rho^{15})$.

(18) Here we are claiming $\delta(\lambda_{15}\lambda_1\tau^4+<)=\lambda_7^3\tau\rho^6+<$. Indeed, let

$$u=(\lambda_{19}\lambda_3+\lambda_{11}\lambda_{11})\tau\rho^6;$$

then we have $\delta(\lambda_{15}\lambda_1\tau^4+u)=\lambda_7^3\tau\rho^6+O(\rho^7)$.

(21) Here we are claiming $\delta(\lambda_{2^b-1}\lambda_{2^c-1}\tau^{2^{b+1}(2m+1)}+<)=\lambda_{2^{b+1}-1}^2\lambda_{2^c-1}\tau^{2^{b-1}(8m+1)}\rho^{2^b3}+<$, at least provided $b-2\geq c\geq 1$. This case is intended to illustrate all the remaining cases, and is identical in form to case (2) of Proposition 4.2.7. Recall from Proposition 4.2.7 that

$$\delta(\lambda_{2^b-1}\tau^{2^{b+1}(2m+1)}+O(\rho))=\lambda_{2^{b+1}-1}^2\tau^{2^{b-1}(8m+1)}\rho^{2^b3}+O(\rho^{2^b3+1}).$$

As

$$\lambda_{2^b-1}\lambda_{2^c-1}\tau^{2^{b+1}(2m+1)}\equiv\lambda_{2^b-1}\tau^{2^{b+1}(2m+1)}\lambda_{2^c-1}\pmod{\rho},$$

it follows that

$$\begin{aligned}\delta(\lambda_{2^b-1}\lambda_{2^c-1}\tau^{2^{b+1}(2m+1)}+O(\rho))&=\delta(\lambda_{2^b-1}\tau^{2^{b+1}(2m+1)}\lambda_{2^c-1}+O(\rho))\\&=(\lambda_{2^{b+1}-1}^2\tau^{2^{b-1}(8m+1)}\rho^{2^b3}+O(\rho^{2^b3+1}))\lambda_{2^c-1}\\&=\lambda_{2^{b+1}-1}^2\lambda_{2^c-1}\tau^{2^{b-1}(8m+1)}\rho^{2^b3}+O(\rho^{2^b3+1}),\end{aligned}$$

which gives the desired relation. The remaining cases are either identical in form to this, or simpler in that they do not require one to first move τ around to reduce to a case already considered in Proposition 4.2.7. □

This produces the indecomposable classes

$$\tau^{2^{a-1}(2(16n+1)+1)}h_{a+3}^2h_a,\quad \tau^{2^a(4(4n+1)+1)}h_{a+3}^3,\quad \tau^{2^{a-1}(16n+1)}c_a,$$
$$\tau^{2^{a+1}(8n+1)}c_{a+1},\quad \tau^{2^{a-1}(2(4n+1)+1)}c_a$$

for $a,n\geq 0$, following the same recipe as employed in Definitions 4.2.6 and 4.2.9, only where one must employ θ-iterates of τ-multiples of the correction terms u given in Proposition 4.2.10.

Proposition 4.2.10 concludes the work necessary for our computation of the $\mathbb{F}_2[\rho]$-module structure of $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$. Let us now summarize in one theorem what we have learned. We wish to give a minimal generating set of $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$ whose elements are products of the indecomposable classes we have found. Before doing so, let us treat the following subtlety.

By way of example, let $x=(1/\rho^{2^b})\delta(\lambda_{2^b-1}\tau^{2^{b-1}(4m+3)})$ with $b\geq 1$, and let $\alpha\in\mathrm{Ext}_{\mathbb{R}}$ be the class represented by x. Our computation in Proposition 4.2.7 combined with the recipe of Theorem 4.1.4 would yield α as an element of a minimal generating set for $\mathrm{Ext}_{\mathbb{R}}$. Observe that x has leading term

$\lambda_{2^b-1}^2 \tau^{2^b(4m+1)}$. It follows quickly from this that x has the same leading term as the cocycle representative of $(\tau^{2^{b-1}(4m+1)} h_b)^2$ given by the product of those cocycle representatives for $\tau^{2^{b-1}(4m+1)} h_b$ given in Definition 4.2.6. However, this does not prove that $\alpha = (\tau^{2^{b-1}(4m+1)} h_b)^2$: we have not ruled out the possibility that $\alpha+\beta = (\tau^{2^{b-1}(4m+1)} h_b)^2$ for some nonzero β represented by a cycle $y < \lambda_{2^b-1}^2 \tau^{2^b(4m+1)}$. This is still sufficient to deduce that we may, if necessary, replace α with $\alpha+\beta$ in our minimal generating set in order to obtain a minimal generating set built as products of indecomposables. It turns out that no such correction is necessary.

Lemma 4.2.11 *Write $\phi\colon \mathrm{Ext}_{\mathbb{R}} \to \mathrm{Ext}_{\mathbb{C}}$ for the quotient. Fix classes α, β in $\mathrm{Ext}_{\mathbb{R}}^1$ or $\mathrm{Ext}_{\mathbb{R}}^2$, at least one of which is ρ-torsion and not both in $\mathrm{Ext}_{\mathbb{R}}^2$. Let r be minimal for which $\rho^r\alpha = 0$ or $\rho^r\beta = 0$. Fix $\gamma \in \mathrm{Ext}_{\mathbb{R}}^{\leq 3}$ not divisible by ρ and such that $\rho^r\gamma = 0$, and suppose $\phi(\alpha)\cdot\phi(\beta) = \phi(\gamma)$. Then $\alpha\cdot\beta = \gamma$.*

Proof Under the given conditions, there is in fact a unique class in the degree of $\alpha\cdot\beta$ which is not divisible by ρ and is killed by ρ^r. This may be seen by direct inspection of the propositions preceding this. □

We may now state the main theorem of this section.

Theorem 4.2.12 (1) *A minimal multiplicative generating set for $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$ as an $\mathbb{F}_2[\rho]$-algebra is given by the classes in the following table:*

multiplicative generator	*ρ-torsion exponent*
h_{a+1}	∞
c_{a+1}	∞
$\tau^{\lfloor 2^{a-1}(4n+1)\rfloor} h_a$	2^a
$\tau^{2^a(8n+1)} h_{a+2}^2$	$2^{a+1}\cdot 3$
$\tau^{\lfloor 2^{a-1}(2(16n+1)+1)\rfloor} h_{a+3}^2 h_a$	$2^a\cdot 13$
$\tau^{2^a(4(4n+1)+1)} h_{a+3}^3$	$2^a\cdot 7$
$\tau^{\lfloor 2^{a-1}(16n+1)\rfloor} c_a$	$2^a\cdot 7$
$\tau^{2^{a+1}(8n+1)} c_{a+1}$	$2^{a+2}\cdot 3$
$\tau^{\lfloor 2^{a-1}(2(4n+1)+1)\rfloor} c_a$	$2^a\cdot 3$

Here $a, n \geq 0$, and the ρ-torsion exponent of a class α is the minimal r for which $\rho^r\alpha = 0$; the classes h_{a+1} and c_{a+1} are ρ-torsion-free.

(2) *The operation Sq^0 acts on these classes by incrementing a in each row.*

(3) *The image of these classes under $\mathrm{Ext}_{\mathbb{R}} \to \mathrm{Ext}_{\mathbb{C}}$ is as their name suggests.*

(4) *A minimal $\mathbb{F}_2[\rho]$-module generating set for $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$ is given in the following table. In all cases, the ρ-torsion exponent of a given class is the minimal ρ-torsion exponent of the multiplicative generators it is written as a product of.*

$\mathbb{F}_2[\rho]$-module generator	constraints
1	
h_a	$a \geq 1$
$\tau^{\lfloor 2^{a-1}(4n+1)\rfloor} h_a$	$a, n \geq 0$
$h_a \cdot h_b$	$a \geq b \geq 1$ and $a \neq b+1$
$h_a \cdot \tau^{\lfloor 2^{b-1}(4n+1)\rfloor} h_b$	$a \geq 1$ and $b, n \geq 0$, and $a \neq b \pm 1$
$\tau^{\lfloor 2^{a-1}\rfloor} h_a \cdot \tau^{\lfloor 2^{a-1}(4n+1)\rfloor} h_a$	$a, n \geq 0$
$\tau^{2^a(8n+1)} h_{a+2}^2$	$a, n \geq 0$
$h_a \cdot h_b \cdot h_c$	$a \geq b \geq c \geq 1$ with $a \neq b+1$, $b \neq c+1$, and if $b = c$ or $a = b$ then $a \neq c+2$
$h_a \cdot h_b \cdot \tau^{\lfloor 2^{c-1}(4n+1)\rfloor} h_c$	$a \geq b \geq 1$ and $c, n \geq 0$ with $a \neq b+1$ and $c \notin \{a \pm 1, b \pm 1\}$, and if $a = b$ then $c \notin \{a-2, a, a+2\}$, and if $a = b+2$ then $c \neq a$
$h_a \cdot \tau^{\lfloor 2^{b-1}\rfloor} h_b \cdot \tau^{\lfloor 2^{b-1}(2n+1)\rfloor} h_b$	$a \geq 1$ and $b, n \geq 0$, and $a \notin \{b-2, b-1, b+1\}$
$h_0 \cdot h_0 \cdot \tau^{2n} h_0$	$n \geq 0$
$h_a \cdot \tau^{2^b(8n+1)} h_{b+2}^2$	$a \geq 1$ and $b, n \geq 0$, and either $a \leq b-1$ or $a \geq b+4$
$\tau^{\lfloor 2^{a-1}(2(16n+1)+1)\rfloor} h_{a+3}^2 h_a$	$a, n \geq 0$
$\tau^{2^a(4(4n+1)+1)} h_{a+3}^3$	$a, n \geq 0$
c_a	$a \geq 1$
$\tau^{\lfloor 2^{a-1}(16n+1)\rfloor} c_a$	$a, n \geq 0$
$\tau^{2^{a+1}(8n+1)} c_{a+1}$	$a, n \geq 0$
$\tau^{\lfloor 2^{a-1}(2(4n+1)+1)\rfloor} c_a$	$a, n \geq 0$

Proof All of this may be read off the preceding computations, using Lemma 4.2.11 with Proposition 4.2.1 if necessary to write a given class as a product of classes in the given generating set. □

We point out the following corollary.

Corollary 4.2.13 *The operation $\rho \cdot \mathrm{Sq}^0$ is injective on $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$.* □

Remark 4.2.14 As indicated in Remark 4.1.6, one may also read off our computation a description of all differentials in the ρ-Bockstein spectral sequence $\mathrm{Ext}_{\mathbb{C}}[\rho] \Rightarrow \mathrm{Ext}_{\mathbb{R}}$ emanating out of filtration at most 2. We leave this to the interested reader. ◁

4.3 Multiplicative structure

We now compute the multiplicative structure of $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$. This material is mostly not needed for our study of the 1-line of the motivic Adams spectral sequence in Section 7; the exception is that we will use the relation Proposition 4.3.4(4) in the proof of Theorem 7.4.9.

Already Lemma 4.2.11 produces a large number of relations. For example, it implies that we may always shift powers of τ around in products that do not vanish in $\mathrm{Ext}_{\mathbb{C}}$, provided it makes sense to do so, yielding relations such as

$$\tau^{\lfloor 2^{a-1}(4n+1)\rfloor}h_a \cdot \tau^{\lfloor 2^{b-1}(4m+1)\rfloor}h_b = h_a \cdot \tau^{\lfloor 2^{b-1}(4(m+2^{a-b-2}(4n+1))+1)\rfloor}h_b$$

for $a \geq b+2$. These were implicitly used in the proof of Theorem 4.2.12. The condition that the product does not vanish in $\mathrm{Ext}_{\mathbb{C}}$ is necessary; see Example 4.3.3 below.

We are left only with relations that would be realized as hidden extensions in the ρ-Bockstein spectral sequence. These arise from the possible failure of the relations $h_{a+1}h_a = 0$ and $h_{a+2}^2 h_a = 0$ to lift through $\mathrm{Ext}_{\mathbb{R}} \to \mathrm{Ext}_{\mathbb{C}}$.

Remark 4.3.1 The following computations will involve some explicit calculations with cocycle representatives. For ease of reference, we collect some important cocycle representatives here:

class	cocycle representative
h_0	λ_0
h_1	λ_1
h_2	λ_3
h_3	λ_7
c_0	$\lambda_2\lambda_3^2$
c_1	$\lambda_5\lambda_7^2$
$\tau^{2^a}h_{a+1}$	$\rho^{-2^{a+1}} \cdot \delta(\tau^{2^{a+1}}) = \theta^{a+1}(\lambda_0) = \tau^{2^a}\lambda_{2^{a+1}-1} = \lambda_{2^{a+1}-1}\tau^{2^a} + O(\rho^{2^a})$
$\tau^2 h_0$	$\lambda_0\tau^2 + \lambda_1\tau^2\rho + \lambda_3\tau\rho^3 + \lambda_4\rho^4$
$\tau^4 h_0$	$\lambda_0\tau^4 + \lambda_3\tau^3\rho^3 + \lambda_5\tau^2\rho^5 + \lambda_7\tau\rho^7 + \lambda_8\rho^8$
τh_2^2	$\lambda_3^2\tau + (\lambda_3\lambda_4 + \lambda_4\lambda_3)\rho = \tau\lambda_3^2$
$\tau^9 h_2^2$	$\lambda_3^2\tau^9 + (\lambda_3\lambda_4 + \lambda_4\lambda_3)\tau^8\rho + \lambda_5\lambda_3\tau^8\rho^2 + O(\rho^{10})$

We will use these without further comment. $\triangleleft$

We begin with some products in $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$ which lift the relation $h_{a+1}h_a = 0$.

Proposition 4.3.2 (1) $h_{a+1} \cdot \tau^{\lfloor 2^{a-1}(4(2n+1)+1)\rfloor}h_a = \rho^{2^a} \cdot \tau^{2^a}h_{a+1} \cdot \tau^{2^a(4n+1)}h_{a+1}$.

(2) $h_{a+1} \cdot \tau^{\lfloor 2^{a-1}(8n+1)\rfloor}h_a = 0$.

(3) $\tau^{2^{a+1}(4n+1)}h_{a+2} \cdot h_{a+1} = \rho^{2^{a+1}} \cdot \tau^{2^a(8n+1)}h_{a+2}^2$.

(4) $\tau^{2^a(8n+1)}h_{a+2}^2 \cdot h_{a+1} = \rho^{2^a} \cdot \tau^{\lfloor 2^{a-1}(16n+1)\rfloor}c_a$.

(5) $\tau^{2^a(8n+1)}h_{a+2}^2 \cdot \tau^{2^a(4m+1)}h_{a+1} = \rho^{2^a} \cdot \tau^{\lfloor 2^{a-1}(2(4(m+2n))+1)\rfloor}c_a$.

(6) $h_{a+3} \cdot \tau^{2^a(16n+1)}h_{a+2}^2 = 0$.

(7) $h_{a+3} \cdot \tau^{2^a(8(2n+1)+1)}h_{a+2}^2 = \rho^{2^{a+3}} \cdot \tau^{2^a(4(4n+1)+1)}h_{a+3}^2$.

Proof In each of these, we may use Sq^0 to reduce to the case $a=0$. In all cases where the product does not vanish, the claimed value of the product is the unique nonzero class in its degree which is both ρ-torsion and divisible by ρ, so it suffices to verify the product working modulo the smallest power of ρ in which the claimed value does not vanish. In doing so, we may in each case reduce to $n=m=0$. With these reductions in place, the proposition is proved by the following computations:

(1) Here we are claiming $h_1\cdot\tau^2h_0=\rho\cdot\tau h_1\cdot\tau h_1$. Indeed, we may compute

$$(\lambda_0\tau^2+\lambda_1\tau^2\rho+\lambda_3\tau\rho^3+\lambda_4\rho^4)\cdot\lambda_1=\lambda_1^2\tau^2\rho+\lambda_2\lambda_1\tau\rho^2+\lambda_2\lambda_2\rho^3+\lambda_2\lambda_3\rho^4=\rho(\lambda_1\tau+\lambda_2\rho)^2=\rho\theta(\lambda_0)^2,$$

which represents $\rho\cdot\tau h_1\cdot\tau h_1$.

(2) There are no nonzero ρ-torsion classes in this degree, so the product must vanish.

(3) Here we are claiming $h_1\cdot\tau^2h_2=\rho^2\cdot\tau h_2^2$. Indeed, we may compute

$$\lambda_1\cdot\tau^2\lambda_3=\rho^2\cdot\tau\lambda_3^2$$

on the nose.

(4) Here we are claiming $h_1\cdot\tau h_2^2=\rho\cdot c_0$. Indeed, we may compute

$$\lambda_1\cdot\tau\lambda_3^2=\lambda_2\lambda_3^2\rho$$

on the nose.

(5) Here we are claiming $\tau h_1\cdot\tau h_2^2=\rho\cdot\tau c_0$. For this, it suffices to work mod ρ^2. Here we may compute

$$\tau\lambda_1\cdot\tau\lambda_2^2=\rho\cdot\lambda_2\lambda_3^2\tau+O(\rho^2),$$

and the claim follows.

(6) Here we have reduced to $a=0$ but not yet to $n=0$. The only nonzero ρ-torsion class in this degree is $\rho^6\tau^{16n+1}c_1$, so it suffices to work mod ρ^7. In doing so, we may now reduce to $n=0$. Indeed, we have

$$\tau\lambda_3^2\cdot\lambda_7=0,$$

and the claim follows.

(7) Here we are claiming $h_3\cdot\tau^9h_2^2=\rho^8\cdot\tau^5h_3^3$. For this, it suffices to work mod ρ^9. Here we may compute

$$(\lambda_3^2\tau^9+(\lambda_4\lambda_3+\lambda_3\lambda_4)\tau^8\rho+\lambda_5\lambda_3\tau^8\rho^2)\cdot\lambda_7=\lambda_7^3\tau^5\rho^8+O(\rho^9),$$

yielding the claim. □

Example 4.3.3 We have

$$\tau^2h_2\cdot h_1^2=\rho^3c_0,\quad h_2\cdot(\tau h_1)^2=0.$$

This serves as a warning that one cannot in general freely shift around powers of τ in products. ◁

We now give some products that lift the relation $h_{a+2}^2 h_a = 0$.

Proposition 4.3.4 (1) $h_{a+2}^2 \cdot \tau^{\lfloor 2^{a-1}(16n+1)\rfloor} h_a = 0$.

(2) $h_{a+2}^2 \cdot \tau^{\lfloor 2^{a-1}(4(2n+1)+1)\rfloor} h_a = \rho^{2^{a+1}} \cdot \tau^{\lfloor 2^{a-1}(2(4n+1)+1)\rfloor} c_a$.

(3) $h_{a+2}^2 \cdot \tau^{\lfloor 2^{a-1}(8(2n+1)+1)\rfloor} h_a = \rho^{2^a 3} \cdot \tau^{2^{a+1}} h_{a+2} \cdot \tau^{2^a(8n+1)} h_{a+2}^2$.

(4) $\tau^{2^{a+2}} h_{a+3} \cdot \tau^{2^{a+2}(4n+1)} h_{a+3} \cdot h_{a+1} = \rho^{2^{a+1}3} \cdot \tau^{2^a(4(4n+1)+1)} h_{a+3}^3$.

(5) $h_{a+1} \cdot h_{a+3} \cdot \tau^{2^{a+2}(4n+1)} h_{a+3} = \rho^{2^{a+2}} \cdot \tau^{2^{a+1}(8n+1)} c_{a+1}$.

(6) $\tau^{2^{a+1}(8n+1)} h_{a+3}^2 \cdot h_{a+1} = 0$.

Proof As in the proof of Proposition 4.3.2, we may use Sq^0 to reduce to the case $a = 0$, and in all cases where the product does not vanish may reduce to $n = 0$. With these reductions in place, the proposition is proved by the following computations:

(1) There are no nonzero ρ-torsion classes in this degree, so the product must vanish.

(2) Here we are claiming $h_2^2 \cdot \tau^2 h_0 = \rho^2 \cdot \tau c_0$. For this, it suffices to work mod ρ^3. Recall that $\tau^2 h_0$ is represented by $\lambda_0 \tau^2 + \lambda_1 \tau^2 \rho + O(\rho^3)$. We may compute

$$(\lambda_0 \tau^2 + \lambda_1 \tau^2 \rho) \cdot \lambda_3^2 = \rho^2 \cdot \lambda_2 \lambda_3^2 \tau + O(\rho^3),$$

and the claim follows.

(3) Here we are claiming $h_2^2 \cdot \tau^4 h_0 = \rho^3 \cdot \tau^2 h_2 \cdot \tau h_2^2$. For this, it suffices to work mod ρ^4. Observe that

$$h_2 \cdot h_2 \cdot \tau^4 h_0 = h_2 \cdot \tau^2 h_2 \cdot \tau^2 h_0 = \tau^2 h_2 \cdot h_2 \cdot \tau^2 h_0 = \tau^2 h_2 \cdot \tau^2 h_2 \cdot h_0$$

by Lemma 4.2.11. We may now compute

$$\lambda_0 \cdot \tau^2 \lambda_3 \cdot \tau^2 \lambda_3 = \rho^3 \cdot \lambda_3^3 \tau^3 + O(\rho^4),$$

yielding the claim.

(4) Here we are claiming $\tau^4 h_3 \cdot \tau^4 h_3 \cdot h_1 = \rho^6 \cdot \tau^5 h_3^3$. For this, it suffices to work mod ρ^7. Here we may compute

$$\lambda_1 \cdot \tau^4 \lambda_7 \cdot \tau^4 \lambda_7 = \rho^6 \cdot \lambda_7^3 \tau^5 + O(\rho^7),$$

yielding the claim.

(5) Here we are claiming $h_1 \cdot h_3 \cdot \tau^4 h_3 = \rho^4 \cdot \tau^2 c_1$. For this, it suffices to work mod ρ^5. Here we may compute

$$\lambda_1 \cdot \tau^4 \lambda_7 \cdot \lambda_7 = \rho^4 \cdot \lambda_5 \lambda_7^2 \tau^2 + O(\rho^6),$$

yielding the claim.

(6) There are no nonzero ρ-torsion classes in this degree, so the product must vanish. □

The preceding propositions leave open three families of products. A complete resolution of these requires the following, which appeared as a conjecture in an earlier version of this work. We thank Dugger, Hill and Isaksen for supplying a proof.

Lemma 4.3.5 (Dugger, Hill and Isaksen) *There are relations*

(1) $\tau^{4m+1}h_1 \cdot \tau^{2l}h_0 = \tau h_1 \cdot \tau^{2(2m+l)}h_0$;

(2) $\tau^{4(4m+1)}h_3 \cdot \tau^{8l+1}h_2^2 = \tau^4 h_3 \cdot \tau^{8(2m+l)+1}h_2^2$;

(3) $\tau^{8m+1}h_2^2 \cdot \tau^{2l}h_0 = \tau h_2^2 \cdot \tau^{2(2m+l)}h_0$.

Proof These will be proved using Massey product-shuffling techniques. The Massey products we require are most easily computed using the ρ-Bockstein spectral sequence; see especially [Belmont and Isaksen 2022, Section 7.4] for a discussion of Massey products in $\mathrm{Ext}_{\mathbb{R}}$.

(1) By induction on m, it suffices to show

$$\tau^{2l}h_0 \cdot \tau^{4m+5}h_1 = \tau^{2l+4}h_0 \cdot \tau^{4m}h_1$$

for $m \geq 0$. Observe that

$$\tau^{4m+5}h_1 = \langle \rho^2, \rho^2\tau^2h_2, \tau^{4m+1}h_1 \rangle, \quad \tau^{2l+4}h_0 = \langle \tau^{2l}, \rho^2, \rho^2\tau^2h_1 \rangle$$

with no indeterminacy. We may therefore shuffle

$$\tau^{2l}h_0 \cdot \tau^{4m+5}h_1 = \tau^{2l}h_0\langle \rho^2, \rho^2\tau^2h_2, \tau^{4m+1}h_1 \rangle = \langle \tau^{2l}h_0, \rho^2, \rho^2\tau^2h_2 \rangle \tau^{4m+1}h_1 = \tau^{2l+4}h_0 \cdot \tau^{4m+1}h_1.$$

(2) By induction on m, it suffices to show

$$\tau^{8l+1}h_2^2 \cdot \tau^{4(4m+1)+16}h_3 = \tau^{8l+17}h_2^2 \cdot \tau^{4(4m+1)}h_3$$

for $m \geq 0$. Observe that

$$\tau^{4(4m+1)+16}h_3 = \langle \rho^8, \rho^8\tau^8h_4, \tau^{4(4m+1)}h_3 \rangle, \quad \tau^{8l+17}h_2^2 = \langle \tau^{8l+1}h_2^2, \rho^8, \rho^8\tau^8h_4 \rangle$$

with no indeterminacy. We may therefore shuffle

$$\begin{aligned} \tau^{8l+1}h_2^2 \cdot \tau^{4(4m+1)+16}h_3 &= \tau^{8l+1}h_2^2\langle \rho^8, \rho^8\tau^8h_4, \tau^{4(4m+1)}h_3 \rangle \\ &= \langle \tau^{8l+1}h_2^2, \rho^8, \rho^8\tau^8h_4 \rangle \tau^{4(4m+1)}h_3 = \tau^{8l+17}h_2^2 \cdot \tau^{4(4m+1)}h_3. \end{aligned}$$

(3) By induction on m, it suffices to show

$$\tau^{2l}h_0 \cdot \tau^{8m+9}h_2^2 = \tau^{2l+8}h_0 \cdot \tau^{8m+1}h_2^2$$

for $m \geq 0$. Observe that

$$\tau^{8m+9}h_2^2 = \langle \rho\tau^4h_3, \rho^7, \tau^{8m+1}h_2^2 \rangle, \quad \tau^{2l+8}h_0 = \langle \tau^{2l}h_0, \rho\tau^4h_0, \rho^7 \rangle$$

with no indeterminacy. We may therefore shuffle

$$\begin{aligned} \tau^{2l}h_0 \cdot \tau^{8m+9}h_2^2 &= \tau^{2l}h_0\langle \rho\tau^4h_3, \rho^7, \tau^{8m+1}h_2^2 \rangle \\ &= \langle \tau^{2l}h_0, \rho\tau^4h_0, \rho^7 \rangle \tau^{8m+1}h_2^2 = \tau^{2l+8}h_0 \cdot \tau^{8m+1}h_2^2. \end{aligned}$$ □

From here, we have the following.

Proposition 4.3.6 *Write* $2m+l+1=2^k(2n+1)$*. Then the following hold*:

(1) $\tau^{2^a(4m+1)}h_{a+1}\cdot\tau^{\lfloor 2^{a-1}(4l+1)\rfloor}h_a=\rho^{2^a(2^{k+1}-1)}\cdot h_{a+1}\cdot\tau^{2^{a+k}(4n+1)}h_{a+k+1}$.

(2) $\tau^{2^{a+2}(4m+1)}h_{a+3}\cdot\tau^{2^a(8l+1)}h_{a+2}^2=\rho^{2^{a+1}(2^{k+2}-3)}\cdot h_{a+1}\cdot h_{a+3}\cdot\tau^{2^{a+k+2}(4n+1)}h_{a+k+3}$.

(3) $\tau^{2^a(8m+1)}h_{a+2}^2\cdot\tau^{\lfloor 2^{a-1}(4l+1)\rfloor}h_a=\rho^{2^a(2^{k+1}-1)}\cdot h_{a+2}^2\cdot\tau^{2^{a+k}(4n+1)}h_{a+k+1}$.

Proof In each of these, we may use Sq^0 to reduce to the case $a=0$. By working modulo the smallest power of ρ in which the claimed product does not vanish, we may reduce to the case $n=0$. By Lemma 4.3.5, we may moreover reduce to the case $m=0$. The proposition is now proved by the following computations:

(1) Here we are claiming $\tau h_1\cdot\tau^{2(2^k-1)}h_0=\rho^{2^{k+1}-1}\cdot h_1\cdot\tau^{2^k}h_{k+1}$. Recall that $\tau^{2(2^k-1)}h_0$ is represented by $\rho^{-1}\delta(\tau^{2(2^k-1)+1})$. Now, the Leibniz rule implies

$$\rho^{-1}\delta(\tau^{2(2^k-1)+1})\cdot\tau\lambda_1=\rho^{-1}\delta(\tau^{2^{k+1}})\cdot\lambda_1+\rho^{-1}\tau^{2(2^k-1)+1}\cdot\delta(\tau)\cdot\lambda_1.$$

The second summand vanishes, as $\delta(\tau)\cdot\lambda_1=\rho\lambda_0\cdot\lambda_1=0$; the first represents $\rho^{2^{k+1}-1}\cdot\tau^{2^k}h_{k+1}\cdot h_1$, yielding the claimed relation.

(2) Here we are claiming $\tau^4h_3\cdot\tau^{8(2^k-1)+1}h_2^2=\rho^{2(2^{k+2}-3)}\cdot h_1\cdot h_3\cdot\tau^{2^{k+2}}h_{k+3}$. Recall that $\tau^{8(2^k-1)+1}h_2^2$ is represented by $\rho^{-6}\delta(\lambda_1\tau^{8(2^k-1)+4})$. Now, the Leibniz rule implies

$$\rho^{-6}\delta(\lambda_1\tau^{8(2^k-1)+4})\cdot\tau^4\lambda_7=\rho^{-6}\lambda_1\cdot\delta(\tau^{2^{k+3}})\cdot\lambda_7+\rho^{-6}\lambda_1\cdot\tau^{8(2^k-1)+4}\cdot\delta(\tau^4)\cdot\lambda_7.$$

The second term vanishes, as $\delta(\tau^4)\cdot\lambda_3=\tau^2\lambda_3\cdot\lambda_7=0$; the first represents $\rho^{2(2^{k+2}-3)}\cdot h_1\cdot h_3\cdot\tau^{2^{k+2}}h_{k+3}$, yielding the claimed relation.

(3) Here we are claiming $\tau h_2^2\cdot\tau^{2(2^k-1)}h_0=\rho^{2^{k+1}-1}\cdot h_2^2\cdot\tau^{2^k}h_{k+1}$. Recall that $\tau^{2(2^k-1)}h_0$ is represented by $\rho^{-1}\delta(\tau^{2(2^k-1)+1})$. Now, the Leibniz rule implies

$$\rho^{-1}\delta(\tau^{2(2^k-1)+1})\cdot\tau\lambda_3^2=\rho^{-1}\delta(\tau^{2^{k+1}})\cdot\lambda_3^2+\rho^{-1}\tau^{2(2^k-1)+1}\cdot\delta(\tau)\cdot\lambda_3^2.$$

The second term vanishes, as $\delta(\tau)\cdot\lambda_3^2=\rho\lambda_0\cdot\lambda_3^2=0$. The first summand represents $\rho^{2^{k+1}-1}h_{k+1}\cdot h_2^2$, yielding the claimed relation. □

The relations above suffice to write any product in $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$ in terms of the basis given in Theorem 4.2.12. Thus we have the following.

Theorem 4.3.7 *A full set of relations for* $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$ *is given by those visible relations which may be deduced from Lemma 4.2.11 together with the products listed in Propositions 4.3.2, 4.3.4, and 4.3.6.* □

Part II The motivic Hopf invariant one problem

5 Some homotopical preliminaries

With the algebraic computation of Section 4 out of the way, we now proceed to more homotopical considerations. This brief section collects a couple of constructions that will be used in the following sections. Explicitly, Section 5.1 will be used in our computation of $d_2(h_5)$ in Section 7, and Section 5.3 will be used in our discussion of the unstable Hopf invariant one problem in Section 6.

5.1 The Hurewicz map

The constant functor $c\colon \mathcal{S}\mathrm{p}^{\mathrm{cl}} \to \mathcal{S}\mathrm{p}^F$ has a lax symmetric monoidal right adjoint c^*, described by

$$c^*(X) = \mathcal{S}\mathrm{p}^F(S^{0,0}, X).$$

In particular, the unit of $c^*(S^{0,0})$ gives a ring map

$$S^0 \to c^*(S^{0,0}),$$

and on homotopy groups this yields a Hurewicz map

$$c\colon \pi_*^{\mathrm{cl}} \to \pi_{*,0}^F.$$

Proposition 5.1.1 *For any F, there is map*

$$c\colon \mathrm{Ext}_{\mathrm{cl}}^{s,f} \to \mathrm{Ext}_F^{s,f,0}$$

of multiplicative spectral sequences, converging to the Hurewicz map

$$c\colon \pi_*^{\mathrm{cl}} \to \pi_{*,0}^F.$$

Moreover, c commutes with Sq^0 and satisfies $c(h_0) = h_0 + \rho h_1$.

Proof Write $H\mathbb{F}_2$ for the ordinary mod 2 Eilenberg–Mac Lane spectrum and $H\mathbb{F}_2^F$ for the motivic spectrum representing mod 2 motivic cohomology. Then $c^*(H\mathbb{F}_2^F) = H\mathbb{F}_2$, thereby giving maps

$$H\mathbb{F}_2^{\otimes n} \simeq c^*(H\mathbb{F}_2^F)^{\otimes n} \to c^*((H\mathbb{F}_2^F)^{\otimes n}).$$

Thus there is a map from the canonical Adams resolution of the sphere to the restriction along c^* of the canonical Adams resolution of the F-motivic sphere. On homotopy groups, this gives a map from the cobar complex of $\mathcal{A}^{\mathrm{cl}}$ to the weight 0 portion of the cobar complex of $\mathcal{A}^F$, and passing to homology we obtain a map

$$\mathrm{Ext}_{\mathrm{cl}}^{s,f} \to \mathrm{Ext}_F^{s,f,0}$$

which is multiplicative and commutes with Sq^0, and by construction this is a map of spectral sequences converging to the Hurewicz map. That $c(h_0) = h_0 + \rho h_1$ follows as these are the classes detecting 2 (see for instance [Isaksen and Østvær 2020, Remark 6.3]). □

5.2 The Lefschetz principle

The *Lefschetz principle* asserts, informally, that "everything" which is true over $\mathbb{C}$ is true over any algebraically closed field. In this subsection, we note how one may read off a certain motivic Lefschetz principle from [Wilson and Østvær 2017].

So far, we have primarily been concerned with F-motivic homotopy theory for F a field of characteristic not equal to 2. For this subsection, we extend our notation to apply also when F is some ring in which 2 is invertible. We shall write $S^{0,0}$ for the $H\mathbb{F}_2^F$-nilpotent completion of the F-motivic sphere spectrum. When F is a field, this is the $(2,\eta)$-completion of the F-motivic sphere spectrum, and, when F is an algebraically closed field, this reduces to a 2-completion [Hu et al. 2011a; Kylling and Wilson 2019]. Let $\mathcal{S}\mathrm{p}_2^F$ denote the category of modules over this completed F-motivic sphere spectrum. In addition, let $\mathcal{S}\mathrm{p}_2^{F,\mathrm{cell}} \subset \mathcal{S}\mathrm{p}_2^F$ denote the cellular subcategory, ie the category generated by the spheres $S^{a,b}$ under colimits.

Proposition 5.2.1 *Let F be an algebraically closed field. Then there is an equivalence*

$$\mathcal{S}\mathrm{p}_2^{F,\mathrm{cell}} \simeq \mathcal{S}\mathrm{p}_2^{\mathbb{C},\mathrm{cell}}.$$

Moreover, this is compatible on Adams spectral sequences with the isomorphism $\mathrm{Ext}_F \cong \mathrm{Ext}_{\mathbb{C}}$.

Proof First suppose that F is of odd characteristic p. We follow the methods of [Wilson and Østvær 2017, Section 6]. Let $W(F)$ be the ring of Witt vectors on F, and choose an algebraically closed field L of characteristic 0 together with embeddings

$$\mathbb{C} \to L \leftarrow W(F) \to F.$$

This gives rise to base change functors

$$\mathcal{S}\mathrm{p}^{\mathbb{C}} \to \mathcal{S}\mathrm{p}^{L} \leftarrow \mathcal{S}\mathrm{p}^{W(F)} \to \mathcal{S}\mathrm{p}^{F},$$

and, in particular, maps

$$\pi_{*,*}^{\mathbb{C}} \to \pi_{*,*}^{L} \leftarrow \pi_{*,*}^{W(F)} \to \pi_{*,*}^{F}. \tag{5-1}$$

Although $W(F)$ is not a field, Wilson and Østvær [2017] show that its Steenrod algebra and Adams spectral sequence are still well behaved, and [loc. cit., Corollary 6.3] that the above maps are modeled on motivic Adams spectral sequences by a zigzag of isomorphisms

$$\mathrm{Ext}_{\mathbb{C}} \to \mathrm{Ext}_{L} \leftarrow \mathrm{Ext}_{W(F)} \to \mathrm{Ext}_{F}.$$

It follows that (5-1) is a zigzag of isomorphisms. In particular, consider the zigzag

$$\mathcal{S}\mathrm{p}_2^{\mathbb{C},\mathrm{cell}} \to \mathcal{S}\mathrm{p}_2^{L,\mathrm{cell}} \leftarrow \mathcal{S}\mathrm{p}_2^{W(F),\mathrm{cell}} \to \mathcal{S}\mathrm{p}_2^{F,\mathrm{cell}}.$$

This is a zigzag of colimit-preserving functors of compactly generated stable categories which are equivalences on subcategories of compact generators, and is thus a zigzag of equivalences. This yields the canonical equivalence $\mathcal{S}\mathrm{p}_2^{\mathbb{C},\mathrm{cell}} \simeq \mathcal{S}\mathrm{p}_2^{F,\mathrm{cell}}$.

If F is of characteristic zero, then we may apply the same argument instead to a zigzag of the form

$$\mathbb{C} \to L \leftarrow F$$

with L algebraically closed. □

5.3 Betti realization

If X is a smooth scheme over $\mathbb{C}$, then the space of complex points of X is a complex manifold. This refines to give *Betti realization* functors [Morel and Voevodsky 1999] from $\mathbb{C}$-motivic spaces to ordinary spaces, and from $\mathbb{C}$-motivic spectra to ordinary spectra, with a number of nice properties. We may use the Lefschetz principle of Proposition 5.2.1 to obtain an analogue for an arbitrary algebraically closed field F.

Let S^0 denote the 2-completed sphere spectrum, and $\mathcal{S}\mathrm{p}_2^{\mathrm{cl}}$ the category of modules thereover.

Proposition 5.3.1 *Let F be an algebraically closed field. Then there is a symmetric monoidal "Betti realization" functor*

$$\mathrm{Be}\colon \mathcal{S}\mathrm{p}_2^{F,\mathrm{cell}} \to \mathcal{S}\mathrm{p}_2^{\mathrm{cl}},$$

factoring through an equivalence from the category of modules over $S^{0,0}[\tau^{-1}]$ in $\mathcal{S}\mathrm{p}_2^{F,\mathrm{cell}}$ to $\mathcal{S}\mathrm{p}_2^{\mathrm{cl}}$, with the following properties:

1. $\mathrm{Be}(\tau)=1$. *In particular,* $\mathrm{Be}(S^{a,b})=S^a$, *so that* Be *induces a map* $\pi^F_{s,w}\to\pi^{\mathrm{cl}}_s$, *and these patch together to an isomorphism* $\pi^F_{*,*}[\tau^{-1}]\cong\pi^{\mathrm{cl}}_*[\tau^{\pm1}]$.
2. *The above isomorphism is modeled on Adams spectral sequences by the map*

$$\mathrm{Ext}_F \to \mathrm{Ext}_F[\tau^{-1}] \cong \mathrm{Ext}_{\mathrm{cl}}[\tau^{\pm1}].$$

3. *The composite* $\mathrm{Be}\circ c\colon \mathcal{S}\mathrm{p}_2^{\mathrm{cl}}\to\mathcal{S}\mathrm{p}_2^{F,\mathrm{cell}}\to\mathcal{S}\mathrm{p}_2^{\mathrm{cl}}$ *is an equivalence. In particular, the map* $c\colon\mathrm{Ext}_{\mathrm{cl}}\to\mathrm{Ext}_F$ *of Proposition 5.1.1 extends to an equivalence* $\mathrm{Ext}_{\mathrm{cl}}[\tau^{\pm1}]\to\mathrm{Ext}_F[\tau^{-1}]$.

Proof These facts are known of the Betti realization functor for $F=\mathbb{C}$ [Dugger and Isaksen 2010, Section 2], and the general case immediately follows from Proposition 5.2.1. □

Using Mandell's p-adic homotopy theory [2001], we may also produce an unstable analogue. Let F be an algebraically closed field. Note from [Hu et al. 2011b, Proposition 15] that the spectrum $H\mathbb{F}_2^F$ is cellular; moreover, $\mathrm{Be}(H\mathbb{F}_2^F)=H\mathbb{F}_2$, as can be seen by inspection of homotopy groups. Let $\mathcal{S}\mathrm{pc}(F)$ be the category of F-motivic spaces and $\mathcal{S}\mathrm{pc}_2$ be the category of 2-complete spaces.

Proposition 5.3.2 *Let F be an algebraically closed field, and define*

$$\mathrm{Be}\colon \mathcal{S}\mathrm{pc}(F)\to\mathcal{S}\mathrm{pc}_2,\quad \mathrm{Be}(X)=\mathcal{C}\mathrm{Alg}_{H\mathbb{F}_2}\big(\mathrm{Be}((H\mathbb{F}_2^F)^{X_+}),\overline{\mathbb{F}}_2\big).$$

Then $\mathrm{Be}(S^{a,b})=(S^a)^\wedge_2$, *and, at least when restricted to the full subcategory of* $\mathcal{S}\mathrm{pc}(F)$ *consisting of simply connected finite motivic cell complexes, the functor* Be *preserves finite colimits and satisfies*

$$H\mathbb{F}_2^{\mathrm{Be}(X)_+}\simeq\mathrm{Be}((H\mathbb{F}_2^F)^{X_+}).$$

Proof We begin by recalling two facts from Mandell's work [2001] on p-adic homotopy theory. Strictly speaking, Mandell states his main theorem at the level of homotopy categories; a reference explicitly treating the full homotopical version we use is [Lurie 2011, Section 3]. First, the functor

$$\mathcal{S}\mathrm{pc} \to \mathcal{C}\mathrm{Alg}_{H\overline{\mathbb{F}}_2}, \quad Y \mapsto H\overline{\mathbb{F}}_2^{Y_+},$$

is fully faithful when restricted to the full subcategory of connected 2-complete nilpotent spaces with locally finite mod 2 cohomology. In particular, if Y is a connected nilpotent space with locally finite mod 2 cohomology, then the unit map

$$Y \simeq \mathcal{S}\mathrm{pc}(*, Y) \to \mathcal{C}\mathrm{Alg}_{H\overline{\mathbb{F}}_2}(H\overline{\mathbb{F}}_2^{Y_+}, H\overline{\mathbb{F}}_2^{*_+}) \simeq \mathcal{C}\mathrm{Alg}_{H\mathbb{F}_2}(H\mathbb{F}_2^{Y_+}, H\overline{\mathbb{F}}_2)$$

realizes the target as the 2-completion of Y. Second, the functor

$$\mathcal{C}\mathrm{Alg}_{H\mathbb{F}_2}^{\mathrm{op}} \to \mathcal{S}\mathrm{pc}, \quad R \mapsto \mathcal{C}\mathrm{Alg}_{H\mathbb{F}_2}(R, H\overline{\mathbb{F}}_2),$$

lands in $\mathcal{S}\mathrm{pc}_2$ and preserves finite colimits when restricted to the full subcategory of $\mathbb{E}_\infty$-algebras R over $\mathbb{F}_2$ such that R_* is locally finite-dimensional, $R_0 = \mathbb{F}_2$, $R_1 = 0$, and the Dyer–Lashof operation Q^0 acts by the identity on R_*.

We now apply this to our situation. The stable Betti realization functor is symmetric monoidal, and thus $\mathrm{Be}((H\mathbb{F}_2^F)^{X_+})$ is indeed an $\mathbb{E}_\infty$-ring over $\mathbb{F}_2$. Moreover, as Sq^0 acts by the identity on $H^{*,*}(X)$, the Dyer–Lashof operation Q^0 acts by the identity on $\pi_* \mathrm{Be}((H\mathbb{F}_2^F)^{X_+})$. In particular, $\mathrm{Be}((H\mathbb{F}_2^F)^{S^{a,b}_+}) \simeq H\mathbb{F}_2^{S^a_+}$, and so the proposition follows by applying Mandell's theory. □

Remark 5.3.3 We have focused in this section on 2-primary motivic homotopy theory over a field F of characteristic not 2. However, our discussion applies in general to p-primary motivic homotopy theory over a field F of characteristic not p. ◁

6 The motivic Hopf invariant one problem

In this section, we formulate and discuss motivic analogues of the Hopf invariant one problem. The material in this section is not needed for Section 7.

6.1 The unstable Hopf invariant one problem

Classically, Adams' determination of the permanent cycles in $\mathrm{Ext}^1_{\mathrm{cl}}$ resolved the Hopf invariant one problem. The Hopf invariant one problem may be formulated motivically using the following.

Definition 6.1.1 Let $f : S^{2a-1,2b} \to S^{a,b}$ be an unstable map between motivic spheres; in particular, $a \geq b \geq 0$ and $a \geq 1$. Write $C(f)$ for the cofiber of f. The map f vanishes in mod 2 motivic cohomology for degree reasons, and thus there exists an isomorphism

$$H^{*,*}(C(f)_+) \cong \mathbb{M}^F\{1, x, y\}$$

of $\mathbb{M}^F$-modules, where $|x| = (-a, -b)$ and $|y| = (-2a, -2b)$. Say that f has *Hopf invariant one* if one may choose such generators x and y to satisfy

$$x^2 = y,$$

ie if $H^{*,*}(C(f)_+) \cong \mathbb{M}^F[x]/(x^3)$; otherwise $x^2 = 0$ and f has Hopf invariant zero. $\triangleleft$

The unstable motivic Hopf invariant one problem is now the following question.

Question 6.1.2 For which (a, b) does there exist a map $f : S^{2a-1,2b} \to S^{a,b}$ of Hopf invariant one? $\triangleleft$

This turns out to mostly reduce to the classical case, by way of the following.

Lemma 6.1.3 *Let $f : S^{2a-1,2b} \to S^{a,b}$ be an unstable F-motivic map. Then f has Hopf invariant one if and only if its base change to an algebraic closure of F is of Hopf invariant one.*

Proof This is immediate from the definitions. $\square$

Proposition 6.1.4 *Fix an unstable F-motivic map $f : S^{2a-1,2b} \to S^{a,b}$ of Hopf invariant one. Then the Betti realization* (see Proposition 5.3.2) *of f is an odd multiple of 2, η, ν, or σ. In particular, $a \in \{1, 2, 4, 8\}$.*

Proof By Lemma 6.1.3, we may as well suppose that F is algebraically closed. Let $C(f)$ denote the cofiber of f and $C(\mathrm{Be}(f))$ the cofiber of $\mathrm{Be}(f)$. Then $\mathrm{Be}(C(f)) = C(\mathrm{Be}(f))$ by Proposition 5.3.2, and thus $H^*\big(C(\mathrm{Be}(f))_+\big) = H^*\big(\mathrm{Be}(C(f))_+\big) = \mathbb{F}_2[x]/(x^3)$ with $|x| = -a$. In other words, the map between 2-completed spheres $\mathrm{Be}(f) : S^{2a-1} \to S^a$ has Hopf invariant one. The proposition now follows from Adams' resolution [1960] of the Hopf invariant one problem. $\square$

Proposition 6.1.4 is not a complete answer to Question 6.1.2, as we have not given any bounds on b, nor have we discussed the existence of maps of Hopf invariant one. Although we will not end up with a complete answer in general, there is more we can say. Before this, we recall what information is encoded in the 1-line of the F-motivic Adams spectral sequence.

6.2 The stable Hopf invariant one problem

Question 6.1.2 can be rephrased as asking when there exists an unstable 2-cell complex, with cells in dimension (a, b) and $(2a, 2b)$, such that in cohomology the bottom cell squares to the top cell. In the stable category, one no longer has cup squares; instead, one has Steenrod operations. Thus we may consider the stable motivic Hopf invariant one problem to be the following question.

Question 6.2.1 What $\mathcal{A}^F$-modules arise as the cohomology of 2-cell complexes? In particular, for which (a, b) does there exist a 2-cell complex, with cells in dimensions $(0, 0)$ and (a, b) and attaching map vanishing in mod 2 motivic cohomology, such that $H^{*,*}X = \mathbb{M}^F\{x, y\}$ is not split as an $\mathcal{A}^F$-module? $\triangleleft$

This is a particular case of the *realization problem* for $\mathcal{A}^F$-modules, and is exactly what the 1-line of the F-motivic Adams spectral sequence encodes. The following is standard.

Proposition 6.2.2 *Fix a class $\epsilon \in \mathrm{Ext}_F^{a-1,1,b}$ classifying an extension $0 \to \mathbb{M}^F\{y\} \to E \to \mathbb{M}^F\{x\} \to 0$ of $\mathcal{A}^F$-modules with $|x| = (0,0)$ and $|y| = (-a,-b)$. Then the following are equivalent*:

1. *There is stable 2-cell complex C with cells in dimensions $(0,0)$ and (a,b) such that $H^{*,*}C \cong E$.*
2. *The class ϵ is a permanent cycle in the F-motivic Adams spectral sequence, and thus detects a stable class $\alpha \in \pi_{a-1,b}^F$.*

Explicitly, if $\epsilon \in \mathrm{Ext}_F^{a-1,1,b}$ detects $\alpha \in \pi_{a-1,b}^F$, then the cofiber $C(\alpha)$ satisfies $H^{*,*}C(\alpha) \cong E$; and, if C is a stable 2-cell complex with $H^{*,*}C = E$, then the fiber of the inclusion $S^{0,0} \to C$ is a map $\alpha\colon S^{a-1,b} \to S^{0,0}$ detected by $\epsilon \in \mathrm{Ext}_F^{a-1,1,b}$. □

As we will see in Section 7, the 1-line of the F-motivic Adams spectral sequence is already quite rich, and strongly depends on the base field F. Thus, in considering the stable Hopf invariant one problem, one may not reduce to the case where F is algebraically closed, unlike in the unstable case.

6.3 Relation between the unstable and stable motivic Hopf invariant one problems

We may now relate the unstable and stable questions, Questions 6.1.2 and 6.2.1.

Proposition 6.3.1 *Let $f\colon S^{2a-1,2b} \to S^{a,b}$ be a map of Hopf invariant one. Then the associated stable class $\alpha \in \pi_{a-1,b}^F$ is detected by a permanent cycle in $\mathrm{Ext}_F^{a-1,1,b}$ which, after base change to the algebraic closure of F, is one of*

$$h_0,\quad h_1,\quad \tau h_1,\quad h_2,\quad \tau h_2,\quad \tau^2 h_2,\quad h_3,\quad \tau h_3,\quad \tau^2 h_3,\quad \tau^3 h_3,\quad \tau^4 h_3.$$

In particular, if $\mathrm{Ext}_F^{a-1,1,b}$ does not contain any such permanent cycle, then there is no map $f\colon S^{2a-1,2b} \to S^{a,b}$ of Hopf invariant one.

Proof By Lemma 6.1.3, we may suppose that F itself is algebraically closed. By stabilizing Proposition 6.1.4, we find that $\mathrm{Be}(\alpha)$ is detected by h_1, h_2, or h_3 in $\mathrm{Ext}_{\mathrm{cl}}^1$. Recall from Proposition 5.3.1 that Betti realization is modeled on Adams spectral sequences by the map

$$\mathrm{Ext}_F \to \mathrm{Ext}_F[\tau^{-1}] \cong \mathrm{Ext}_{\mathrm{cl}}[\tau^{\pm 1}].$$

In particular, the structure of Ext_F (see Proposition 4.2.1) implies that α must be detected by a permanent cycle in Ext_F of the form $\tau^n h_0$, $\tau^n h_1$, $\tau^n h_2$, or $\tau^n h_3$ for some $n \geq 0$. As f is an unstable map, this class must have nonnegative weight, reducing to the listed classes. □

Remark 6.3.2 Our method of relating the unstable motivic Hopf invariant one problem to the stable motivic Hopf invariant one problem, going through the "Betti realization" functors of Section 5.3, may seem somewhat roundabout. This route was taken for the following reason: if $f\colon S^{2a-1} \to S^a$ is a map of Hopf invariant one, then the fact that $H^*(C(f))$ is nonsplit as an $\mathcal{A}^{\mathrm{cl}}$-module, and thus the associated stable class $\alpha \in \pi^{\mathrm{cl}}_{a-1}$ is detected in $\mathrm{Ext}^1_{\mathrm{cl}}$, follows from the instability condition $\mathrm{Sq}^a(x) = x^2$.

Motivically, the analogous instability condition asserts that, if X is a motivic space and $x \in H^{2a,a}(X_+)$, then $\mathrm{Sq}^{2a}(x) = x^2$ [Voevodsky 2003, Lemma 9.7]. Now suppose that $f : S^{2a-1,2b} \to S^{a,b}$ is an unstable map of Hopf invariant one, and write $H^{*,*}(C(f)_+) = \mathbb{M}^F[x]/(x^3)$ with $|x| = (-a,-b)$. If a is even and $b \le \frac{1}{2}a$, then one may set $c = \frac{1}{2}a - b$ and deduce $\mathrm{Sq}^a(\tau^c x) = \tau^{2c} x^2$, so that $H^{*,*}(C(f))$ is not split as an $\mathcal{A}^F$-module. If a is odd, then one may argue by appealing to an integral motivic Hopf invariant and graded commutativity, as in the classical case. Thus, it is to rule out the possibility of a map $f : S^{2a-1,2b} \to S^{a,b}$ of Hopf invariant one with $b > \frac{1}{2}a$ that we have taken our approach. ◁

Our computations in Section 7 show, for a variety of base fields F, when Ext^1_F contains a permanent cycle whose image over the algebraic closure is one of the classes listed in Proposition 6.3.1, yielding various nonexistence results. To obtain existence results, we must recall how maps of Hopf invariant one arise.

6.4 Geometric applications

Adams' resolution of the classical Hopf invariant one problem had geometric consequences; notably, it implied that the only spheres which admit H-space structures are S^0, S^1, S^3, and S^7. It makes sense to ask for the motivic analogue of this, ie to ask which spheres $S^{a,b}$ admit H-space structures.

This question is in some sense geometric, but we can also ask for something even more concrete. The spheres $S^{a,b}$ are certain sheaves on the Nisnevich site of smooth F-schemes, and so it is reasonable to ask when $S^{a,b}$ is in fact represented by a smooth F-scheme. This question was raised and studied by Asok et al. [2017]; in particular, they produce explicit smooth affine schemes representing $S^{a,\lceil a/2\rceil}$, as well as prove that $S^{a,b}$ is not represented by a smooth scheme for $a > 2b$. Motivated by this, we are led to ask the following question.

Question 6.4.1 For what pairs (a,b) is $S^{a,b}$ a motivic H-space? Of these, when is it represented by a smooth F-scheme which admits a unital product? ◁

Classically, the connection between the H-space structures and the Hopf invariant one problem is via the *Hopf construction*. This construction may also be carried out in the motivic category, and has been studied in this context in [Dugger and Isaksen 2013]. We recall the key points.

Definition 6.4.2 [Dugger and Isaksen 2013, Definition C.1] Let X, Y, and Z be pointed spaces, and let $h : X \times Y \to Z$ be a pointed map. The *Hopf construction of h* is the map $H(h) : X \star Y \to \Sigma Z$ obtained by taking homotopy colimits of the rows of the diagram

$$\begin{tikzcd} X \arrow[d] & X \times Y \arrow[l] \arrow[r] \arrow[d] & Y \arrow[d] \\ * & Z \arrow[l] \arrow[r] & * \end{tikzcd}$$

◁

Here $\star$ is the join. Note that $S^{a,b} \star S^{c,d} \simeq S^{a+c+1,b+d}$; thus the Hopf construction may be used to construct maps between motivic spheres. Using the theory of Cayley–Dickson algebras, Dugger and Isaksen [2013, Section 4] used this to define *motivic Hopf maps* $\eta \in \pi^F_{1,1}$, $\nu \in \pi^F_{3,2}$, and $\sigma \in \pi^F_{7,4}$. As noted in [loc. cit., Remark 4.14], these motivic Hopf maps have Hopf invariant one. This is a general property of the Hopf construction, which we may summarize in the following.

Lemma 6.4.3 *If $\mu\colon S^{a-1,b} \times S^{a-1,b} \to S^{a-1,b}$ is an H-space product, then its Hopf construction $H(\mu)\colon S^{2a-1,2b} \to S^{a,b}$ has Hopf invariant one.*

Proof The proof of the analogous fact for topological spaces [Steenrod 1962, Section I.5] extends to motivic spaces. We summarize the key points.

Define the (mod 2) *degree* of a pointed map $S^{a,b} \to S^{a,b}$ of motivic spaces to be its induced map in reduced motivic cohomology. A pointed map $f\colon S^{a-1,b} \times S^{a-1,b} \to S^{a-1,b}$ of motivic spaces is said to have *degree* (α,β) if $f|_{S^{a-1,b}\times\{p_2\}}$ has degree α and $f|_{\{p_1\}\times S^{a-1,b}}$ has degree β. Since μ is an H-space product, its restrictions to $S^{a-1,b} \times \{p_2\}$ and $\{p_1\} \times S^{a-1,b}$ are homotopic to the identity, so μ has degree $(1,1)$. The lemma follows by showing that, more generally, the Hopf invariant, defined in the evident way, of the Hopf construction of a map of degree (α,β) is $\alpha\cdot\beta$.

Steenrod and Epstein's proof of [Steenrod 1962, Lemma 5.3] carries over to the motivic setting to complete the proof. The main point is that Steenrod and Epstein work with particular models of the cone, join, homotopy cofiber, and suspension in their proof, but any model would work, as all of their statements only depend on the homotopy types of the relevant spaces and homotopy classes of the relevant maps. More precisely, with notation as in their proof, one may replace E_1, E_2, E_+, and E_- by the cones on S_1, S_2, S, and S, respectively, to avoid any potential point-set issues. In particular, one regards E_1, E_2, E_+, and E_- as suspension data in the sense of [Dugger and Isaksen 2013, Remark 2.9] for the various suspensions appearing in the Hopf construction. In this language, the identifications of various pushouts in the proof of [Steenrod 1962, Lemma 5.3] are examples of induced orientations [Dugger and Isaksen 2013, Remark 2.10]. The proof carries through unchanged with these new choices of E_1, E_2, E_+, and E_-.

To be precise, their proof considers maps $S^{n-1} \times S^{n-1} \to S^{n-1}$ with $n > 1$ even and works integrally. Routine modifications extend this to arbitrary $n \geq 1$ provided one works mod 2 throughout. Classically, this is the adaption needed to incorporate the degree 2 map $S^1 \to S^1$, which is the Hopf construction of the standard product on $S^0 \cong C_2$. Motivically, this is the adaption needed for our lemma to hold for arbitrary unstable motivic spheres $S^{a-1,b}$, allowing especially for the uniform treatment of 2 and η. □

Remark 6.4.4 Under Definition 6.1.1, the map $h\colon S^{1,1} \to S^{1,1}$ represented by the squaring map on $\mathbb{G}_m$, sometimes called the "zeroth Hopf map" and stably detected by h_0, is *not* a map of Hopf invariant one. In the context of Lemma 6.4.3, this is justified by the fact that, for degree reasons, h is not the Hopf construction of an H-space structure on any motivic sphere. ◁

We can now summarize what is known in the following.

Theorem 6.4.5 *A motivic sphere is represented by a smooth F-scheme admitting a unital product if and only if it is one of*

$$S^{0,0},\quad S^{1,1},\quad S^{3,2},\quad S^{7,4}.$$

In addition to the motivic spheres listed above, the following motivic spheres admit H-space structures:

$$S^{1,0},\quad S^{3,0},\quad S^{7,0}.$$

The only other motivic spheres that could possibly admit H-space structures are

$$S^{3,1},\quad S^{7,3},\quad S^{7,2},\quad S^{7,1};$$

moreover, an H-space structure on such a sphere produces a permanent cycle in Ext_F whose image over the algebraic closure is τh_2, τh_3, $\tau^2 h_3$, or $\tau^3 h_3$, respectively.

Proof That the spheres $S^{0,0}$, $S^{1,1}$, $S^{3,2}$, and $S^{7,4}$ are represented by smooth F-schemes admitting a unital product is given by [Dugger and Isaksen 2013]. The spheres $S^{1,0}$, $S^{3,0}$, and $S^{7,0}$ are the images of S^1, S^3, and S^7, respectively, under the unstable constant functor from spaces to motivic spaces, and so inherit H-space structures from their classical structures. That all the spheres listed are the only spheres which may admit H-space structures follows from Lemma 6.4.3 and Proposition 6.3.1, as does the final claim concerning the F-motivic Adams spectral sequence. Finally, Asok et al. [2017, Proposition 2.3.1] prove that, if $S^{a-1,b}$ is represented by a smooth F-scheme, then necessarily $2b \geq a-1$, and the only possible H-spaces satisfying this are $S^{0,0}$, $S^{1,1}$, $S^{3,2}$, and $S^{7,4}$, as listed. □

We note the following special case.

Corollary 6.4.6 *Suppose there is an $\mathbb{R}$-motivic map $f\colon S^{2a-1,2b} \to S^{a,b}$ of Hopf invariant one. Then (a,b) is one of*

$$(1,0),\quad (2,1),\quad (4,2),\quad (8,4),\quad (2,0),\quad (4,0),\quad (8,0).$$

Moreover, all of these are realized, and in fact

$$S^{0,0},\quad S^{1,1},\quad S^{3,2},\quad S^{7,4},\quad S^{1,0},\quad S^{3,0},\quad S^{7,0}$$

are all the $\mathbb{R}$-motivic spheres admitting H-space structures.

Proof This is immediate from Theorem 6.4.5, either appealing to the fact that $\mathrm{Ext}_{\mathbb{R}}$ vanishes in the degrees detecting the remaining possibilities, or else noting that the real points of $S^{a,b}$ are S^{a-b}, so that, if $S^{a,b}$ is an H-space, then $a-b \in \{0,1,3,7\}$. □

7 The 1-line of the motivic Adams spectral sequence

We now analyze the 1-line of the F-motivic Adams spectral sequence. We begin in Section 7.1 by explaining how to read off the structure of Ext_F for various fields F from our computation of $\mathrm{Ext}_{\mathbb{R}}$.

After some additional preliminaries in Section 7.2, we give a direct motivic analogue of the classical differentials in Section 7.3, proving $d_2(h_{a+1}) = (h_0 + \rho h_1)h_a^2$ for $a \geq 3$ over arbitrary base fields. We then proceed to give more detailed information about the 1-line for the particular fields F of the form $\mathbb{R}$, $\mathbb{F}_q$ with q an odd prime power, $\mathbb{Q}_p$ with p any prime, and $\mathbb{Q}$.

7.1 Computing Ext_F

As a general rule, Ext_F is largely understood once $\mathbb{M}^F$ and $\mathrm{Ext}_{\mathbb{R}}$ are both understood. Rather than formulate a precise statement, let us just describe Ext_F for the various particular fields F we shall encounter, namely those described in Example 2.2.1 as well as $F = \mathbb{Q}$.

Recall from Remark 2.3.2 that, for any field F, we may view $\mathbb{M}^F$ as an $\mathcal{A}^{\mathbb{R}}$-module, and there is an isomorphism

$$\mathrm{Ext}_F \cong \mathrm{Ext}_{\mathcal{A}^{\mathbb{R}}}(\mathbb{M}^{\mathbb{R}}, \mathbb{M}^F).$$

Thus, the main point is to understand $\mathbb{M}^F$ as an $\mathcal{A}^{\mathbb{R}}$-module, and this is in fact determined by $\mathbb{M}_0^F$ as an $\mathbb{F}_2[\rho]$-module. For the examples of interest, we have the following. Abbreviate

$$\mathbb{M} = \mathbb{F}_2[\tau, \rho], \quad \mathbb{M}_{(r)} = \mathbb{M}/(\rho^r).$$

Lemma 7.1.1 *As $\mathcal{A}^{\mathbb{R}}$-modules, we have the following*:

(1) $\mathbb{M}^{\mathbb{R}} = \mathbb{M}$.

(2) *If $F = \bar{F}$ is algebraically closed, then* $\mathbb{M}^F = \mathbb{M}_{(1)}$.

(3) *If* $q \equiv 1 \pmod 4$, *then* $\mathbb{M}^{\mathbb{F}_q} = \mathbb{M}_{(1)}\{1, u\}$.

(4) *If* $q \equiv 3 \pmod 4$, *then* $\mathbb{M}^{\mathbb{F}_q} = \mathbb{M}_{(2)}$.

(5) *If* $p \equiv 1 \pmod 4$, *then* $\mathbb{M}^{\mathbb{Q}_p} = \mathbb{M}_{(1)}\{1, \pi, u, \pi u\}$.

(6) *If* $p \equiv 3 \pmod 4$, *then* $\mathbb{M}^{\mathbb{Q}_p} = \mathbb{M}_{(2)}\{1, \pi\}$.

(7) $\mathbb{M}^{\mathbb{Q}_2} = \mathbb{M}_{(3)}\{1\} \oplus \mathbb{M}_{(1)}\{u, \pi\}$.

(8) $\mathbb{M}^{\mathbb{Q}} = \mathbb{M}\{1\} \oplus \mathbb{M}_{(1)}\{[2]\} \oplus \mathbb{M}_{(1)}\{[p], a_p : p \equiv 1 \pmod 4\} \oplus \mathbb{M}_{(2)}\{u_p : p \equiv 3 \pmod 4\}$.

Proof All but the case $F = \mathbb{Q}$ may be read off the examples listed in Example 2.2.1. When $F = \mathbb{Q}$, the ring $\mathbb{M}^{\mathbb{Q}}$ is described in [Ormsby and Østvær 2013, Propositions 5.3 and 5.4], following [Milnor 1970]. Our description may be read off this upon setting $u_p = [p] + \rho$ for $p \equiv 3 \pmod 4$. □

For $r \geq 0$, define

$$\mathrm{Ext}_{(r)} = \mathrm{Ext}_{\mathcal{A}^{\mathbb{R}}}(\mathbb{M}, \mathbb{M}_{(r)}) = H_*(\Lambda^{\mathbb{R}}/(\rho^r)).$$

The $\mathbb{F}_2[\rho]$-module structure of $\mathrm{Ext}_{(r)}$ may be easily computed from $\mathrm{Ext}_{\mathbb{R}}$ via the long exact sequence associated to the cofiber of ρ^r. Even less work is necessary when $\mathrm{Ext}_{\mathbb{R}}$ has been computed by some method compatible with the ρ-Bockstein spectral sequence such as ours; see in particular Remark 4.1.5. Thus Theorem 4.2.12 allows us to read off $\mathrm{Ext}^f_{(r)}$ for $f \leq 2$, as well as the image of $\mathrm{Ext}_{\mathbb{R}} \to \mathrm{Ext}^3_{(r)}$. This does not give the entirety of $\mathrm{Ext}^3_{(r)}$; however, we at least know that whatever remains is generated by classes which appear in the ρ-Bockstein spectral sequence as $\rho^k\alpha$ with $\alpha \in \mathrm{Ext}^3_{(1)}$ and $k < r$, and this is enough information for our purposes.

Lemma 7.1.1 describes for various F how Ext_F may be written as a direct sum of copies of various $\mathrm{Ext}_{(r)}$. For example, $\mathrm{Ext}_{\mathbb{Q}_2} = \mathrm{Ext}_{(3)}\{1\} \oplus \mathrm{Ext}_{(1)}\{u, \pi\}$. We may use this to prove a Hasse principle for $\mathrm{Ext}_{\mathbb{Q}}$.

Lemma 7.1.2 *The map*

$$\mathbb{M}^{\mathbb{Q}} \to \mathbb{M}^{\mathbb{Q}_p}$$

satisfies

$$[p] \mapsto \pi, \quad a_p \mapsto u\pi, \quad u_p \mapsto \pi + \rho.$$

Here the first is relevant for $p = 2$ *or* $p \equiv 1 \pmod 4$, *the second for* $p \equiv 1 \pmod 4$, *and the third for* $p \equiv 3 \pmod 4$.

Proof The behavior of these maps is described in [Ormsby and Østvær 2013, Proposition 5.3]. Our description follows immediately; note we have defined $u_p = [p] + \rho$ for $p \equiv 3 \pmod 4$. □

Proposition 7.1.3 *The Hasse map*

$$\mathrm{Ext}_{\mathbb{Q}} \to \mathrm{Ext}_{\mathbb{R}} \times \prod_p \mathrm{Ext}_{\mathbb{Q}_p}$$

is injective.

Proof By Lemma 7.1.1, we have

$$\mathrm{Ext}_{\mathbb{Q}} = \mathrm{Ext}_{\mathbb{R}} \oplus \mathrm{Ext}_{(1)}\{[2]\} \oplus \mathrm{Ext}_{(1)}\{[p], a_p : p \equiv 1 \pmod 4\} \oplus \mathrm{Ext}_{(2)}\{u_p : p \equiv 3 \pmod 4\}.$$

The summand $\mathrm{Ext}_{\mathbb{R}}$ maps isomorphically to $\mathrm{Ext}_{\mathbb{R}}$, and the maps $\mathrm{Ext}_{\mathbb{Q}} \to \mathrm{Ext}_{\mathbb{Q}_p}$ are determined by Lemma 7.1.2. In particular, it is easily seen that the maps

$$\mathrm{Ext}_{(1)}\{[2]\} \to \mathrm{Ext}_{\mathbb{Q}_2}, \quad \mathrm{Ext}_{(1)}\{[p], a_p\} \to \mathrm{Ext}_{\mathbb{Q}_p}, \quad \mathrm{Ext}_{(2)}\{u_p\} \to \mathrm{Ext}_{\mathbb{Q}_p}$$

are all split injections, and the proposition follows. □

The preceding discussion, together with our computation of $\mathrm{Ext}_{\mathbb{R}}$, describes what we will need of Ext_F in low filtrations and arbitrary stem. So that we may rule out various higher differentials in low stems for degree reasons, we record the following.

Lemma 7.1.4 $\mathrm{Ext}_{(1)}$ *is given in stems* $s \leq 6$ *by the module*

$$\mathbb{F}_2[\tau] \otimes \left(\mathbb{F}_2\{h_0^n : n \geq 0\} \oplus \mathbb{F}_2\{h_1, h_1^2, h_1^3, h_2, h_0h_1, h_2^2\}\right) \oplus \mathbb{F}_2[\tau]/(\tau)\{h_1^4, h_1^5, h_1^6\}.$$

Proof These groups have been computed in [Dugger and Isaksen 2010]. □

7.2 Existence of Hopf elements

Our computation of the F-motivic Adams differentials $d_2(h_{a+1})$ will follow a similar pattern to Wang's computation [1967] of the corresponding classical Adams differentials (differentials which were first computed in [Adams 1960]). This is an inductive argument, beginning with information about the Hopf elements which are known to exist. We record some of this information in this subsection.

Write $\epsilon \in \pi^F_{0,0}$ for the class represented by the twist map $S^{1,1} \otimes S^{1,1} \to S^{1,1} \otimes S^{1,1}$.

Lemma 7.2.1 *Fix $\alpha \in \pi^F_{a,b}$ and $\beta \in \pi^F_{c,d}$. Then there is an identity*

$$\alpha \cdot \beta = (-1)^{(a-b)(c-d)} \epsilon^{bd} \cdot \beta \cdot \alpha.$$

Moreover, $1-\epsilon$ is detected in Ext_F by h_0 and 2 by $h_0 + \rho h_1$.

Proof The claimed graded commutativity is given in [Morel 2004, Corollary 6.1.2]; see also [Isaksen and Østvær 2020, Section 6.1] for a discussion. That $1-\epsilon$ is detected by h_0 and 2 by $h_0 + \rho h_1$ is noted in [Isaksen and Østvær 2020, Remark 6.3]. □

Lemma 7.2.2 *For any field F, the class h_a is a permanent cycle for $a \in \{0, 1, 2, 3\}$.*

Proof The class h_0 is a permanent cycle by Lemma 7.2.1. Dugger and Isaksen [2013] construct the motivic Hopf elements η, ν, and σ, and indicate [loc. cit., Remark 4.14] that these are detected by h_1, h_2, and h_3, respectively; see also our discussion in Section 6.4. Thus these classes must be permanent cycles. □

7.3 Nonexistence of Hopf elements

The purpose of this subsection is to prove the following.

Theorem 7.3.1 *For an arbitrary base field F of characteristic not equal to 2, there are differentials of the form*

$$d_2(h_{a+1}) = (h_0 + \rho h_1) h_a^2$$

in the F-motivic Adams spectral sequence, which are nonzero for $a \geq 3$. ◁

By naturality, it suffices to produce these differentials in the case where F is a prime field, ie $F = \mathbb{F}_q$ or $F = \mathbb{Q}$, and when F is algebraically closed. Moreover, by the Hasse principal given in Proposition 7.1.3, the case $F = \mathbb{Q}$ may be deduced from the cases $F = \mathbb{Q}_p$ and $F = \mathbb{R}$ combined. All of these build on the case where F is algebraically closed, which may be treated as follows.

Proposition 7.3.2 *If $F = \overline{F}$ is algebraically closed, then*

$$d_2(h_{a+1}) = h_0 h_a^2.$$

This is nonzero for $a \geq 3$.

Proof The corresponding classical differentials are known due to [Adams 1960]. The proposition could be reduced to this by appealing to Proposition 5.3.1; however, we shall instead proceed as follows.

Wang [1967, Section 3] gives another proof of the classical differentials, combining only a minimal amount of homotopical input with a good understanding of $\mathrm{Ext}_{\mathrm{cl}}$. His argument may be applied essentially verbatim to produce the claimed $\overline{F}$-motivic differentials. It is this argument that may be adapted to work for other base fields, so to motivate our later computations let us recall this argument in full.

The proof proceeds by induction on a, where only the base case requires any homotopical input.

Consider the base case $a = 3$. The class h_3 is a permanent cycle, detecting the Hopf element σ; see Lemma 7.2.2. By Lemma 7.2.1, we find that $2\sigma^2 = 0$. As 2 is detected by h_0 over algebraically closed fields, it follows that $h_0h_3^2$ cannot survive the Adams spectral sequence. The structure of $\mathrm{Ext}_{\overline{F}}$ implies that $d_2(h_4) = h_0h_3^2$ is the only way for $h_0h_3^2$ to die.

Now suppose we have produced the differential $d_2(h_a) = h_0h_{a-1}^2$ for some $n \geq 4$. The relation $h_{a+1}h_a = 0$ together with the Leibniz rule implies

$$0 = d_2(h_{a+1}h_a) = d_2(h_{a+1}) \cdot h_a + h_{a+1} \cdot d_2(h_a).$$

Applying our inductive hypothesis and the relation $h_{a+1} \cdot h_{a-1}^2 = h_a^3$, this reduces to

$$(d_2(h_{a+1}) + h_0h_a^2) \cdot h_a = 0.$$

The algebraic structure of $\mathrm{Ext}_{\overline{F}}^3$ implies that $d_2(h_{a+1}) \in \mathbb{F}_2\{h_0h_a^2\}$, so it suffices to verify that $h_0h_a^3 \neq 0$ for $a \geq 4$. This follows from Wang's computation [1967, Proposition 3.4] by comparison along the map $\mathrm{Ext}_{\overline{F}} \to \mathrm{Ext}_{\overline{F}}[\tau^{-1}] \simeq \mathrm{Ext}_{\mathrm{cl}}[\tau^{\pm 1}]$. □

The base step for the inductive argument given in Proposition 7.3.2 works for arbitrary base fields, but the inductive step falls apart. This inductive step relies on the algebraic fact that, when working over an algebraically close field, multiplication by h_a is injective on the degree of $d_2(h_{a+1})$ for $a \geq 4$. Over other base fields, this fails for $a = 4$: this degree may contain elements of the form $\omega h_1 h_4^2$ where $\omega \in \mathrm{Ext}_F^{-1,0,-1}$ is a sum of elements such as ρ, π, and u, and

$$\omega h_1 h_4^2 \cdot h_4 = \omega h_1 \cdot h_4^3 = \omega h_1 \cdot h_3^2 \cdot h_5 = 0.$$

Luckily, the inductive step fails only for $a = 4$; once we have resolved $d_2(h_5)$, the remaining differentials will follow via the same argument. To resolve this differential, we proceed as follows.

Proposition 7.3.3 *Let F be a field of the form $\mathbb{F}_q$ for q odd, $\mathbb{Q}_p$ for any p, or $\mathbb{R}$. Then there is a differential*

$$d_2(h_5) = (h_0 + \rho h_1)h_4^2$$

in the F-motivic Adams spectral sequence.

Proof When $F = \mathbb{R}$, we first make the following reduction. Observe that $\mathrm{Ext}_{\mathbb{R}}$ in the degree of $d_2(h_5)$ is given by $\mathbb{F}_2\{h_0 h_4^2, \rho h_1 h_4^2\}$, and that neither of these classes are divisible by ρ^2. Thus it is sufficient to verify this differential in the Adams spectral sequence for the cofiber of ρ^2. By [Behrens and Shah 2020, Lemma 7.8], this cofiber is a ring spectrum, and so its Adams spectral sequence is multiplicative. Having made this reduction, the remainder of the argument is uniform in the given choices of F. For brevity of notation, in the following we shall write Ext_F for the object so named when $F = \mathbb{F}_q$ or $F = \mathbb{Q}_p$, and write the same for $\mathrm{Ext}_{(2)}$ when $F = \mathbb{R}$.

First observe that, as $\tau^4 \in \mathrm{Ext}_F^0$, the class τ^{16} is a square and thus a d_2-cycle. As τ^{16} acts injectively on Ext_F^f for $f \le 3$, it suffices to show

$$d_2(\tau^{16} h_5) = (h_0 + \rho h_1)\tau^{16} h_4^2.$$

Consider the Hurewicz map $c\colon \pi_* \to \pi_{*,0}^F$. Let $\theta_4 \in \pi_{30} S^0$ be the Kervaire class, detected by h_4^2 and satisfying $2\theta_4 = 0$. By Proposition 5.1.1, we find that $c(\theta_4)$ is detected by $(\mathrm{Sq}^0)^4(h_0^2) = \tau^{16} h_4^2$. As $2 \cdot c(\theta_4) = 0$, the class $(h_0 + \rho h_1)\tau^{16} h_4^2$ cannot survive. The only possibility is that $d_2(\tau^{16} h_4) = (h_0 + \rho h_1)\tau^{16} h_4^2$, yielding the desired differential. □

Remark 7.3.4 When $F = \mathbb{R}$, the differential $d_2(h_5)$, and in fact all the differentials $d_2(h_{a+1})$, may also be produced as follows. By comparison with $\mathbb{C}$, one finds $d_2(h_5) \in h_0 h_4^2 + \mathbb{F}_2\{\rho h_1 h_4^2\}$. Thus it suffices to verify that $d_2(h_5)$ is not ρ-torsion. This is a consequence of the fact that the isomorphism $\mathrm{Ext}_{\mathbb{R}}[\rho^{-1}] \simeq \mathrm{Ext}_{\mathrm{dcl}}[\rho^{\pm 1}]$ [Dugger and Isaksen 2017b, Theorem 4.1] commutes with Adams differentials. ◁

We need just one more algebraic fact for the proof of Theorem 7.3.1.

Lemma 7.3.5 *Let $\omega \in \mathrm{Ext}_F^0$ be nonzero. Then $\omega h_1 h_a^3 \neq 0$ for all $a \ge 5$.*

Proof The class $h_0 h_{a-1}^3$ is nonzero in $\mathrm{Ext}_{\mathrm{cl}}$ for $a \ge 5$ by [Wang 1967, Proposition 3.4]. Proposition 3.2.1 gives an injection $\mathrm{Ext}_{\mathrm{dcl}} \to \mathrm{Ext}_F$, and this extends by linearity to an injection $\mathrm{Ext}_F^0 \otimes_{\mathbb{F}_2} \mathrm{Ext}_{\mathrm{dcl}} \to \mathrm{Ext}_F$, as can be seen by using Lemma 7.1.1 to reduce to the injections $\mathrm{Ext}_{(r)}^0 \otimes_{\mathbb{F}_2} \mathrm{Ext}_{\mathrm{dcl}} \to \mathrm{Ext}_{(r)}$. The class $\omega h_1 h_a^3$ is the image of $\omega \otimes h_0 h_{a-1}^3$ under this map, yielding the claim. □

We may now give the following.

Proof of Theorem 7.3.1 As discussed, it suffices to consider only the cases where F is of the form $\mathbb{F}_q$ for some q odd, $\mathbb{Q}_q$ for some q, or $\mathbb{R}$. So let F be one of these. We now induct on a, with base cases $a = 3$ and $a = 4$.

First consider the case $a = 3$. By Lemma 7.2.2, the class h_3 is a permanent cycle detecting the class σ. By Lemma 7.2.1, $2\sigma^2 = 0$, and so $(h_0 + \rho h_1)h_3^2$ must be the target of a differential. The only possibility is that $d_2(h_4) = (h_0 + \rho h_1)h_3^2$.

The case $a = 4$ was handled in Proposition 7.3.3.

Now suppose inductively that we have produced the differential $d_2(h_a) = (h_0 + \rho h_1)h_{a-1}^2$ for some $a \geq 5$. Combining the Leibniz rule with the relation $h_{a+1}h_a = 0$, we find

$$0 = d_2(h_{a+1}h_a) = d_2(h_{a+1})h_a + h_{a+1}d_2(h_a).$$

Applying our inductive hypothesis and the relation $h_{a+1}h_{a-1}^2 = h_a^2$, we find

$$(d_2(h_{a+1}) + (h_0 + \rho h_1)h_a^2)h_a = 0.$$

It follows that $d_2(h_{a+1}) = (h_0 + \rho h_1)h_a^2 + x$ where x is some class killed by h_a. The only classes in this degree are $h_0 h_a^2$ and those of the form $\omega h_1 h_a^2$ where $\omega \in \mathrm{Ext}_F^0$. By comparison with $\overline{F}$, we find that x must be zero or a nonzero class of the form $\omega h_1 h_a^2$ with $\omega \in \mathrm{Ext}_F^{-1,0,-1}$. As $a \geq 5$, Lemma 7.3.5 implies that none of the latter are killed by h_a. Thus $x = 0$, yielding the desired differential. □

This concludes our uniform analysis of differentials out of Ext_F^1. The rest of this section is dedicated to studying the 1-line in more detail for particular fields F.

7.4 The real numbers

We now study the case $F = \mathbb{R}$ in more detail. Recall from Theorem 4.2.12 that

$$\mathrm{Ext}_{\mathbb{R}}^1 = \mathbb{F}_2[\rho]\{h_a : a \geq 1\} \oplus \bigoplus_{a \geq 0} \mathbb{F}_2[\rho]/(\rho^{2^a})\{\tau^{\lfloor 2^{a-1}(4n+1)\rfloor} h_a : n \geq 0\}.$$

Here recall that $2^{a-1}(4n+1) = 2n$ for $a = 0$. Theorem 7.3.1 allows one to understand the fate of the classes in the ρ-torsion-free summand, so we turn our attention to the ρ-torsion subgroup. We shall first pin down which of these ρ-torsion classes are permanent cycles, and then by separate methods compute all d_2-differentials on these ρ-torsion classes. A comparison reveals that there must be numerous higher differentials, but determining these is outside the scope of our computation. The first point of order is the following.

Definition 7.4.1 For $a \geq 0$, write $a = c + 4d$ with $0 \leq c \leq 3$, and define $\psi(a) = 2^c + 8d$ to be the a^{th} *Radon–Hurwitz number*. ◁

Proposition 7.4.2 *The class $\rho^r \tau^{2^{a-1}(4n+1)} h_a$ is a permanent cycle if and only if $r \geq 2^a - \psi(a)$.*

The proof of Proposition 7.4.2 requires some preliminaries. We proceed by comparison with Borel C_2-equivariant stable homotopy theory. Let Ext_{BC_2} denote the E_2-page of the Borel C_2-equivariant Adams spectral sequence [Greenlees 1988]. Explicitly,

$$\mathrm{Ext}_{BC_2}^{s,f,w} = \mathrm{Ext}_{\mathcal{A}^{\mathrm{cl}}}^{s-w,f}(\mathbb{F}_2, H^* P_w^\infty);$$

this is just a combination of the ordinary Adams spectral sequences for the stable cohomotopy groups of infinite stunted projective space. By Lin's positive resolution [1980] of the Segal conjecture, this spectral sequence converges to $\pi_{*,*}^{C_2}$, the homotopy groups of the 2-completion of the C_2-equivariant sphere spectrum.

Betti realization followed by Borel completion yields a functor from the stable $\mathbb{R}$-motivic category to the Borel C_2-equivariant stable category $\mathrm{Fun}(BC_2, \mathcal{S}\mathrm{p})$, and Behrens and Shah [2020, Section 8] show that this may be understood as completing at ρ and inverting τ. Applying this to an Adams resolution, we find that

$$\mathrm{Ext}_{BC_2} = \lim_{n\to\infty} \mathrm{Ext}_{(2^n)}[\tau^{-2^n}].$$

The simple form of $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$ allows us to immediately read off $\mathrm{Ext}_{BC_2}^{\leq 3}$.

Lemma 7.4.3 *$\mathrm{Ext}_{BC_2}^{\leq 3}$ is exactly as $\mathrm{Ext}_{\mathbb{R}}^{\leq 3}$ is described in Theorem 4.2.12, except n is allowed to be negative, and in place of the map $\mathrm{Ext}_{\mathbb{R}} \to \mathrm{Ext}_{\mathbb{C}}$ is a map $\mathrm{Ext}_{BC_2} \to \mathrm{Ext}_{\mathbb{C}}[\tau^{-1}] \cong \mathrm{Ext}_{\mathrm{cl}}[\tau^{\pm 1}]$.* □

In particular,

$$\mathrm{Ext}_{BC_2}^1 = \mathbb{F}_2[\rho]\{h_a : a \geq 1\} \oplus \bigoplus_{a\geq 0} \mathbb{F}_2[\rho]/(\rho^{2^a})\{\tau^{\lfloor 2^{a-1}(4n+1)\rfloor} h_a : n \in \mathbb{Z}\}.$$

We have introduced Ext_{BC_2} in order to make the following reduction.

Lemma 7.4.4 *Write $h\colon \mathrm{Ext}_{\mathbb{R}} \to \mathrm{Ext}_{BC_2}$ for the canonical map of spectral sequences. Fix a ρ-torsion class $x \in \mathrm{Ext}_{\mathbb{R}}^1$. Then x is a permanent cycle if and only if $h(x)$ is a permanent cycle.*

Proof Clearly, if x is a permanent cycle, then the same must be true of $h(x)$. Conversely, suppose that $h(x)$ is a nontrivial permanent cycle; we claim that x is a permanent cycle.

Write Ext_{C_2} for the E_2-page of the C_2-equivariant Adams spectral sequence [Hu and Kriz 2001, Section 6], converging to the same target as Ext_{BC_2}. This splits additively as $\mathrm{Ext}_{C_2} = \mathrm{Ext}_{\mathbb{R}} \oplus \mathrm{Ext}_{\mathrm{NC}}$ for a certain summand $\mathrm{Ext}_{\mathrm{NC}}$ (see [Guillou et al. 2020, Section 2]), and h factors as $h = g \circ f : \mathrm{Ext}_{\mathbb{R}} \to \mathrm{Ext}_{C_2} \to \mathrm{Ext}_{BC_2}$, the first map being the obvious inclusion and the second map killing the summand $\mathrm{Ext}_{\mathrm{NC}}$.

As $h(x)$ is a nontrivial permanent cycle, it detects a class α in Borel Adams filtration 1. The class α must then be detected in $\mathrm{Ext}_{C_2}^{\leq 1}$. By [Belmont et al. 2021], the map $\mathrm{Ext}_{\mathbb{R}} \to \mathrm{Ext}_{C_2}$ is an isomorphism in the degrees under consideration, so the same must be true for $\mathrm{Ext}_{C_2} \to \mathrm{Ext}_{BC_2}$. As there is at most one nonzero ρ-torsion class in these degrees, the only possibility is that α is detected by $f(x)$ in $\mathrm{Ext}_{C_2}^1$, implying that $f(x)$ is a permanent cycle. As $\mathrm{Ext}_{\mathbb{R}} \to \mathrm{Ext}_{C_2}$ is the inclusion of a summand, this implies that x is a permanent cycle, as claimed. □

Thus it suffices to understand permanent cycles in $\mathrm{Ext}_{BC_2}^1$. The main point is the following.

Lemma 7.4.5 *There exists a nonzero ρ-torsion class $\alpha \in \pi_{s,w}^{C_2}$ detected in Borel Adams filtration 1 if and only if the inclusion of the bottom cell of P_{w-s-1}^{w-1} is split, where P_k^n is the Thom spectrum of the k-fold Whitney sum of the tautological line bundle over the real projective space $\mathbb{R}P^n$.*

Proof First suppose given such a map α. The structure of $\mathrm{Ext}^1_{BC_2}$ implies that α must have ρ-torsion exponent $s+1$, and so there is a lift $\overline{\alpha}$ in the diagram

$$\begin{tikzcd}
\Sigma^{s-w+1} P^{w-1}_{w-s-1} \arrow[drr, dashed, "\overline{\alpha}"] & & \\
\Sigma^{s-w} P^{\infty}_{w} \arrow[u, "\partial"] \arrow[rr, "\alpha"] & & S^0 \\
\Sigma^{s-w} P^{\infty}_{w-s-1} \arrow[u] \arrow[urr, "\rho^{s+1}\alpha=0"'] & &
\end{tikzcd}$$

As α and ∂ have Adams filtration 1, necessarily $\overline{\alpha}$ has Adams filtration 0. It follows that precomposing $\overline{\alpha}$ with the inclusion of the bottom cell $S^0 \to \Sigma^{s-w+1} P^{w-1}_{w-s-1}$ gives a map $S^0 \to S^0$ which is nonzero in mod 2 cohomology, and must therefore be an equivalence. In other words, $\overline{\alpha}$ splits off the bottom cell of P^{w-1}_{w-s-1}.

Conversely, if the inclusion of the bottom cell of P^{w-1}_{w-s-1} is split, then its splitting gives a nonzero map $\overline{\alpha}$ as above in Adams filtration 0. Let $\alpha = \overline{\alpha} \circ \partial$; we claim that α is a nonzero class detected in Adams filtration 1. Indeed, the cofibering $P^{w-1}_{w-s-1} \to P^{\infty}_{w-s-1} \to P^{\infty}_{w}$ gives an exact sequence

$$\mathrm{Ext}^0(\mathbb{F}_2, H^* P^{\infty}_{w}) \to \mathrm{Ext}^0(\mathbb{F}_2, H^* P^{\infty}_{w-s-1}) \to \mathrm{Ext}^0(\mathbb{F}_2, H^* P^{w-1}_{w-s-1}) \xrightarrow{\partial'} \mathrm{Ext}^1(\mathbb{F}_2, H^* P^{\infty}_{w}),$$

where ∂' models restriction along ∂ in the previous diagram. The first map is exactly

$$\rho^{s+1} \colon \mathrm{Ext}^{*,0,w}_{BC_2} \to \mathrm{Ext}^{*,0,w-s-1}_{BC_2}.$$

As $\mathrm{Ext}^0_{BC_2} = \mathbb{F}_2[\rho]$, we find that the kernel of ∂' consists of only that class represented by the inclusion $\mathbb{F}_2 \to H^0 P^{w-1}_{w-s-1}$. So ∂' is injective in the relevant degrees, implying that α is nonzero and of Adams filtration 1, as claimed. □

We may now give the following.

Proof of Proposition 7.4.2 By Lemma 7.4.4, it suffices to show that a class $\rho^r \tau^{\lfloor 2^{a-1}(4n+1)\rfloor} h_a \in \mathrm{Ext}^1_{BC_2}$ is a permanent cycle if and only if $r \geq 2^a - \psi(a)$. By sparseness of $\mathrm{Ext}^1_{BC_2}$, the class $\rho^r \tau^{\lfloor 2^{a-1}(4n+1)\rfloor} h_a$ is a permanent cycle if and only if there is some ρ-torsion class $\alpha \in \pi^{C_2}_{2^a-r-1,-2^{a+1}n-r}$ detected in Borel Adams filtration 1. By Lemma 7.4.5, this holds if and only if inclusion of the bottom cell of $P^{-2^{a+1}n-r-1}_{-2^{a+1}n-2^a}$ is split. By James periodicity [1958; 1959], this holds if and only if the inclusion of the bottom cell of $P^{2^N-2^{a+1}n-r-1}_{2^N-2^{a+1}-2^a}$ is split for some sufficiently large $N \gg 0$; that is, we may assume ourselves to be working with suspension spectra of honest real projective spaces. When this happens was resolved by Adams' solution [1962, Theorem 1.2] of the vector fields on spheres problem, yielding the condition claimed. □

Corollary 7.4.6 *The classes $\tau^{\lfloor 2^{a-1}(4n+1)\rfloor} h_a$ are permanent cycles for $a \leq 3$.* □

Corollary 7.4.6 could also be proved more directly, applying the technique used in the proof of Theorem 7.3.1 or Proposition 7.4.8 below to reduce to the region considered by Belmont and Isaksen.

It is worth summarizing what we have learned from the proof of Proposition 7.4.2 about the stable cohomotopy groups of projective spaces.

Theorem 7.4.7 *The subgroup of permanent cycles in* $\mathrm{Ext}^1_{BC_2}$ *is given by*

$$\mathbb{F}_2[\rho]\{h_1, h_2, h_3, \rho h_4\} \oplus \bigoplus_{a \geq 0} \mathbb{F}_2[\rho]/(\rho^{\psi(a)})\{\rho^{2^a - \psi(a)} \tau^{\lfloor 2^{a-1}(4n+1) \rfloor} h_a : n \in \mathbb{Z}\}.$$

A choice of maps $\Sigma^c P^\infty_w \to S^0$ *detected by these permanent cycles is given by the following:*

(1) *For all* $r \geq 0$, *there are maps*

$$P^\infty_{1-r} \to P^\infty_1 \xrightarrow{\eta} S^0, \quad \Sigma P^\infty_{2-r} \to \Sigma P^\infty_2 \xrightarrow{\nu} S^0, \quad \Sigma^3 P^\infty_{4-r} \to \Sigma^3 P^\infty_3 \xrightarrow{\sigma} S^0.$$

Here η, ν, *and* σ *are equivariant refinements of the Hopf maps with the same names. These composites are detected by* $\rho^r h_1$, $\rho^r h_2$, *and* $\rho^r h_3$, *respectively.*

(2) *For all* $r \geq 0$, *there is a map*

$$\Sigma^7 P^\infty_{7-r} \to \Sigma^7 P^\infty_7 \xrightarrow{\mathrm{Sq}(\sigma)} S^0,$$

where $\mathrm{Sq}(\sigma)$ *is the symmetric square of* $\sigma : S^7 \to S^0$. *This composite is detected by* $\rho^{1+r} h_4$.

(3) *For all* $a \geq 0$, $n \in \mathbb{Z}$, *and* $1 \leq r \leq \psi(a)$, *there is a map*

$$\Sigma^{2^a(2n+1)-1} P^\infty_{-2^a(2n+1)+r} \xrightarrow{\partial} \Sigma^{2^a(2n+1)} P^{-2^a(2n+1)+r-1}_{-2^a(2n+1)} \xrightarrow{s} S^0.$$

Here ∂ *is the cofiber of the map* $\Sigma^{2^a(2n+1)-1} P^\infty_{-2^a(2n+1)} \to \Sigma^{2^a(2n+1)-1} P^\infty_{-2^a(2n+1)+r}$, *and* s *is any map that splits off the bottom cell of* $P^{-2^a(2n+1)+r-1}_{-2^a(2n+1)}$. *This composite is detected by* $\rho^{2^a - r} \tau^{\lfloor 2^{a-1}(4n+1) \rfloor} h_a$.

Proof Recall that

$$\mathrm{Ext}^1_{BC_2} = \mathbb{F}_2[\rho]\{h_a : a \geq 1\} \oplus \bigoplus_{a \geq 0} \mathbb{F}_2[\rho]/(\rho^{2^a})\{\tau^{\lfloor 2^{a-1}(4n+1) \rfloor} h_a : n \in \mathbb{Z}\}.$$

We have just analyzed which classes in the ρ-torsion summand are permanent cycles, leading to exactly the claimed ρ-torsion permanent cycles with representatives as described in (3). Lemma 7.2.2 implies that h_1, h_2, and h_3 are permanent cycles, and these detect the maps described in (1). Theorem 7.3.1 shows that $\rho^n h_a$ supports a d_2-differential for $a \geq 5$ and $n \geq 0$, and that h_4 supports a d_2-differential but ρh_4 does not. We are left with verifying that ρh_4 is a permanent cycle detecting the map $\mathrm{Sq}(\sigma)$. Indeed, taking geometric fixed points yields an isomorphism $\pi^{C_2}_{*,*}[\rho^{-1}] \cong \pi^{\mathrm{cl}}_*[\rho^{\pm 1}]$ which sends $\mathrm{Sq}(\alpha)$ to α for any $\alpha \in \pi^{\mathrm{cl}}_*$. This isomorphism is modeled on Adams spectral sequences by $\mathrm{Ext}_{C_2}[\rho^{-1}] \cong \mathrm{Ext}_{\mathbb{R}}[\rho^{-1}] \cong \mathrm{Ext}_{\mathrm{cl}}[\rho^{\pm 1}]$. As ρh_4 is the only class in its degree lifting $h_3 \in \mathrm{Ext}^1_{\mathrm{cl}}$, it must be that ρh_4 detects $\mathrm{Sq}(\sigma)$. □

Proposition 7.4.2 implies that the classes $\tau^{2^{a-1}(4n+1)} h_a$ must support Adams differentials for $a \geq 4$. Although we do not compute all these differentials, we do give the following.

Proposition 7.4.8 *For all* $n \geq 0$ *and* $a \geq 3$, *there is a differential*

$$d_2(\tau^{2^a(4n+1)} h_{a+1}) = (h_0 + \rho h_1)(\tau^{2^{a-1}(4n+1)} h_a)^2.$$

Proof We give separate arguments for the case $a=3$ and $a>3$. First consider the case $a=3$. The class $\tau^{4(4n+1)}h_3$ is a permanent cycle by Corollary 7.4.6, detecting a class which we might call $\tau^{4(4n+1)}\sigma$. By Lemma 7.2.1, $2\cdot(\tau^{4(4n+1)}\sigma)^2=0$, and so $(h_0+\rho h_1)\cdot(\tau^{4(4n+1)}h_3)^2$ must die. This class is not divisible by ρ, and the only non-ρ-divisible classes that may hit it are $\tau^8 h_4$ and $\tau^8 h_4+\rho^{16}h_5$. By Theorem 7.3.1, if $d_2(\tau^8 h_4+\rho^{16}h_5)=(h_0+\rho h_1)\cdot(\tau^{4(4n+1)}h_3)^2$, then $d_2(\tau^8 h_4)=(h_0+\rho h_1)\cdot(\tau^{4(4n+1)}h_3+h_4)^2$. This is not possible as $\tau^8 h_4$ is ρ-torsion and this target is not. Thus, in fact, $d_2(\tau^8 h_4)=(h_0+\rho h_1)\cdot(\tau^{4(4n+1)}h_3)^2$, as claimed.

Next consider the case $a>3$. The ρ-torsion subgroup of $\mathrm{Ext}_{\mathbb{R}}$ in the degree of $d_2(\tau^{2^a(4n+1)}h_{a+1})$ is given by $\mathbb{F}_2\{h_0,\rho h_1\}\otimes\mathbb{F}_2\{(\tau^{2^{a-1}(4n+1)}h_a)^2\}$. These classes are not divisible by ρ^2, and so it suffices to verify the differential in the Adams spectral sequence for the cofiber of ρ^2. By [Behrens and Shah 2020, Lemma 7.8], this cofiber is a ring spectrum, so its Adams spectral sequence is multiplicative. As τ^2 is a cycle, τ^4 is a d_2-cycle, so we reduce to showing $d_2(h_{a+1})=(h_0+\rho h_1)h_a^2$. This was shown in Theorem 7.3.1. □

We may summarize what we have learned as follows.

Theorem 7.4.9 *The nontrivial d_2-differentials out of the 1-line of the $\mathbb{R}$-motivic Adams spectral sequence are exactly those given in the following table*:

source	target	constraints
h_4	$h_0h_3^2$	
$\rho^r h_a$	$\rho^r(h_0+\rho h_1)h_{a-1}^2$	$a\ge 5,\ r\ge 0$
$\rho^r\tau^{2^{a-1}(4n+1)}h_a$	$\rho^r(h_0+\rho h_1)(\tau^{2^{a-2}(4n+1)}h_{a-1})^2$	$n\ge 0,\ a\ge 4,\ 0\le r\le 2^{a-1}-1$

The 1-line of the E_3-page of the $\mathbb{R}$-motivic Adams spectral sequence has a basis given by the elements in the following table:

$\mathbb{F}_2[\rho]$-module generator	constraints	ρ-torsion exponent
h_a	$a\in\{1,2,3\}$	∞
ρh_4		∞
$\tau^{\lfloor 2^{a-1}(4n+1)\rfloor}h_a$	$n\ge 0$ and $a\in\{0,1,2,3\}$	2^a
$\rho^{2^{a-1}-1}\tau^{2^{a-1}(4n+1)}h_a$	$n\ge 0$ and $a\ge 4$	$2^{a-1}+1$

Those classes in $\mathrm{Ext}^1_{\mathbb{R}}$ which are permanent cycles are given in the following table:

$\mathbb{F}_2[\rho]$-module generator	constraints	ρ-torsion exponent	stem
h_a	$a\in\{1,2,3\}$	∞	2^a-1
ρh_4		∞	14
$\tau^{\lfloor 2^{a-1}(4n+1)\rfloor}h_a$	$n\ge 0$ and $a\in\{0,1,2,3\}$	2^a	2^a-1
$\rho^{2^a-\psi(a)}\tau^{2^{a-1}(4n+1)}h_a$	$n\ge 0,\ a\ge 4$	$\psi(a)$	$\psi(a)-1$

Proof All of this is immediate from Theorem 7.3.1, Propositions 7.4.2 and 7.4.8, Theorem 7.4.7, and the ρ-torsion exponents of the generators of $\mathrm{Ext}^3_{\mathbb{R}}$ given in Theorem 4.2.12, with the following exception: Proposition 7.4.8 produces differentials $d_2(\tau^{8(4n+1)}h_4) = (h_0+\rho h_1)(\tau^{4(4n+1)}h_3)^2$, and one must use Proposition 4.3.4(4) to check that this target has ρ-torsion exponent 7. □

7.5 Finite fields

We now study the case where F is a finite field. For the most part, this case follows by combining Theorem 7.3.1 with differentials out of Ext^0_F that may be deduced from [Kylling 2015]. By naturality, our discussion in this subsection gives information for F an arbitrary field of odd characteristic.

We will need the following definition.

Definition 7.5.1 For an integer q, let $\nu_2(q)$ denote the 2-adic valuation of q, ie

$$q = 2^{\nu_2(q)}(2n+1)$$

for some integer n, and let

$$\varepsilon(q) = \nu_2(q-1), \quad \lambda(q) = \nu_2(q^2-1). \qquad \triangleleft$$

We now split into cases based on congruence of the order of the field mod 4.

7.5.1 $q \equiv 1 \pmod 4$

Fix a prime power q such that $q \equiv 1 \pmod 4$. We work over $F = \mathbb{F}_q$. Recall that $\mathrm{Ext}_{\mathbb{F}_q} = \mathrm{Ext}_{(1)}\{1,u\}$. In particular,

$$\mathrm{Ext}^1_{\mathbb{F}_q} = \mathbb{F}_2[\tau]\{1,u\} \otimes \mathbb{F}_2\{h_a : a \geq 0\}.$$

The class u is a permanent cycle for degree reasons, and we have already computed the differential on all the classes h_a. However the story does not stop there; instead, we have the following.

Lemma 7.5.2 *There are differentials*

$$d_{\varepsilon(q)+s}(\tau^{2^s}) = u\tau^{2^s-1}h_0^{\varepsilon(q)+s}$$

for all $s \geq 0$.

Proof Kylling [2015, Lemma 4.2.1] produces identical differentials in the $\mathbb{F}_q$-motivic Adams spectral sequence for $H\mathbb{Z}$. The claimed differentials follow by naturality. □

This may be combined with Theorem 7.3.1 to easily compute all differentials out of the 1-line.

Theorem 7.5.3 *For* $q \equiv 1 \pmod 4$, *the* 1*-line of the* $\mathbb{F}_q$*-motivic Adams spectral sequence supports only the nontrivial differentials given in the following table*:

source	d_r	target	constraints
$\tau^n h_0$	$d_{\varepsilon(q)+\nu_2(n)}$	$\tau^{n-1} h_0^{\varepsilon(q)+\nu_2(n)+1}$	$n \geq 1$
$\tau^{2n+1} h_2$	d_2	$u\tau^{2n} h_2 h_0^2$	$n \geq 0$, $\varepsilon(q) = 2$
$\tau^{2n+1} h_3$	d_2	$u\tau^{2n} h_3 h_0^2$	$n \geq 0$, $\varepsilon(q) = 2$
$\tau^{2n+1} h_3$	d_3	$u\tau^{4n+1} h_3 h_0^3$	$n \geq 0$, $\varepsilon(q) = 3$
$\tau^{4n+2} h_3$	d_3	$u\tau^{4n+1} h_3 h_0^3$	$n \geq 0$, $\varepsilon(q) = 2$
$\tau^n h_b$	d_2	$\tau^n h_0 h_{b-1}^2 + d_2(\tau^n) h_b$	$n \geq 0$, $b \geq 4$
$u\tau^n h_b$	d_2	$u\tau^n h_0 h_{b-1}^2$	$n \geq 0$, $b \geq 4$

After these have been run, the 1*-line of the* E_∞*-page of the* $\mathbb{F}_q$*-motivic Adams spectral sequence has a basis given by the elements in the following table*:

class	constraints
h_0	
$\tau^n h_1$	$n \geq 0$
$\tau^n h_2$	$n \geq 0$, *where if* $\varepsilon(q) = 2$ *then* $n \equiv 0 \pmod 2$
$\tau^n h_3$	$n \geq 0$, *where if* $\varepsilon(q) = 2$ *then* $n \equiv 0 \pmod 4$, *and if* $\varepsilon(q) = 3$ *then* $n \equiv 0 \pmod 2$
$u\tau^n h_b$	$n \geq 0$, $b \in \{0, 1, 2, 3\}$

Proof The first four families of differentials follow immediately from Lemmas 7.5.2 and 7.2.2, and the remaining two by combining Lemma 7.5.2 with Theorem 7.3.1. Note in particular that $d_2(\tau^n) \equiv 0 \pmod u$, and thus $d_2(\tau^n h_b) \neq 0$ for $b \geq 4$. The second table may be easily read off the first, provided we verify that we have not missed any differentials, ie that the classes listed in the second table are indeed permanent cycles. For degree reasons, the only possible nontrivial differentials on the classes $\tau^n h_b$ with $b \in \{1, 2, 3\}$ would be of the form

(1) $d_r(\tau^n h_1) \overset{?}{=} \tau^{n-1} h_0^{r-1}$,

(2) $d_2(\tau^n h_2) \overset{?}{=} u\tau^{n-1} h_0^2 h_2$,

(3) $d_2(\tau^n h_3) \overset{?}{=} u\tau^{n-1} h_0^2 h_3$,

(4) $d_3(\tau^n h_3) \overset{?}{=} u\tau^{n-1} h_0^3 h_3$

with $n \geq 1$. The first is impossible for $n = 1$ as h_0 detects 2 and thus no power of h_0 may be killed, and is impossible for $n \geq 2$ as the class $\tau^{n-1} h_0^{r-1}$ must support the differential given the first row of the first table. The remaining three differentials may occur, and when they occur is accounted for in the given tables. □

7.5.2 $q \equiv 3 \pmod 4$ Now fix a prime power q such that $q \equiv 3 \pmod 4$. We work over $F = \mathbb{F}_q$. Recall that $\mathrm{Ext}_{\mathbb{F}_q} = \mathrm{Ext}_{(2)}$.

Lemma 7.5.4 *We may identify*

$$\mathrm{Ext}^0_{\mathbb{F}_q} = \mathbb{F}_2[\tau^2, \rho, \tau\rho]/(\rho^2 = \rho \cdot (\tau\rho) = (\tau\rho)^2 = 0),$$

and $\mathrm{Ext}^1_{\mathbb{F}_q}$ *is the tensor product of* $\mathbb{F}_2[\tau^2]$ *with*

$$\mathbb{F}_2\{h_0, \rho\tau \cdot h_0\} \oplus \mathbb{F}_2\{h_1, \rho \cdot h_1, \rho\tau \cdot h_1, \tau h_1\} \oplus \mathbb{F}_2\{h_b, \rho \cdot h_b, \rho\tau \cdot h_b : b \geq 2\}.$$

Proof This follows quickly from our computation of $\mathrm{Ext}_{\mathbb{R}}$, following the recipe of Remark 4.1.5. Alternatively, one may compute the ρ-Bockstein spectral sequence

$$\mathrm{Ext}_{(1)}[\rho]/(\rho^2) \Rightarrow \mathrm{Ext}_{(2)}$$

directly (see [Wilson and Østvær 2017]); the only relevant differential is $d_1(\tau) = \rho h_0$. □

As in the previous case, powers of τ support arbitrarily long differentials.

Lemma 7.5.5 *There are differentials*

$$d_{\lambda(q)+s}(\tau^{2^{s+1}}) = \rho\tau^{2^{s+1}-1} h_0^{\lambda(q)+s}$$

for all $s \geq 0$. *On the other hand*, $\rho\tau$ *is a permanent cycle.*

Proof The class $\rho\tau$ is a permanent cycle for degree reasons. Kylling [2015, Lemma 4.2.2] produces identical differentials in the $\mathbb{F}_q$-motivic Adams spectral sequence for $\mathbb{F}_q$-motivic $H\mathbb{Z}$. The claimed differentials follow by naturality. □

Theorem 7.5.6 *For* $q \equiv 3 \pmod 4$, *the* 1*-line of the* $\mathbb{F}_q$*-motivic Adams spectral sequence supports the differentials given in the following table*:

source	d_r	target	constraints
$\tau^{2n} h_0$	$d_{\lambda(q)+\nu_2(n)}$	$\rho\tau^{2n-1} h_0^{\lambda(q)+\nu_2(n)+1}$	$n \geq 1$
$\tau^{4n+2} h_3$	d_3	$\rho\tau^{4n+1} h_0^3 h_3$	$n \geq 0$, $\lambda(q) = 3$
$\tau^{2n} h_b$	d_2	$\tau^{2n}(h_0 + \rho h_1) h_{b-1}^2$	$n \geq 1$, $b \geq 4$
$\rho\tau^{2n+1} h_b$	d_2	$\rho\tau^{2n+1} h_0 h_{b-1}^2$	$n \geq 0$, $b \geq 4$

After the d_2*-differentials have been run, the* 1*-line of the* E_3*-page of the* $\mathbb{F}_q$*-motivic Adams spectral sequence has a basis given by the classes in the following table*:

class	constraints
h_0	
$\rho^\epsilon \cdot \tau^{2n} h_b$	$n \geq 0$, $\epsilon \in \{0, 1\}$, $b \in \{1, 2, 3\}$
$\rho\tau^{2n+1} h_b$	$n \geq 0$, $b \in \{0, 1, 2, 3\}$
$\rho^\epsilon \tau^{4n+1} h_1$	$n \geq 0$, $\epsilon \in \{0, 1\}$
$\tau^{2n} h_0$	$n \geq 1$
$\rho\tau^{2n} h_b$	$n \geq 0$, $b \geq 4$

Of these, all the classes in the first region are permanent cycles, with the exception that $\tau^{4n+2}h_3$ *supports a* d_3*-differential if* $\lambda(q)=3$. *The classes* $\tau^{2n}h_0$ *for* $n\geq 1$ *are not permanent cycles, and we leave open the fate of the classes* $\rho\tau^{2n}h_b$ *for* $n\geq 1$ *and* $b\geq 4$.

Proof The given differentials follow quickly by combining Theorem 7.3.1 with Lemma 7.5.5, and this accounts for all d_2-differentials. Note in particular that τ^2 is a d_2-cycle as $\lambda(q)\geq 3$ whenever $q\equiv 3 \pmod 4$. Thus the given E_3-page may be produced by linearly propagating the differentials of Theorem 7.3.1. Note also that $d_2(\rho\tau^{2n}h_b)=\rho\tau^{2n}(h_0+\rho h_1)h_{b-1}^2=0$ for all $n\geq 0$ and $b\geq 4$, yielding the classes in the final row of the second table.

It remains only to verify that the permanent cycles provided are indeed permanent cycles. As ρ and $\rho\tau$ are permanent cycles for degree reasons, we may reduce to considering only the classes $\tau^{2n}h_b$, $\rho\tau^{2n+1}h_0$, and $\tau^{4n+1}h_1$ for $b\in\{1,2,3\}$ and $n\geq 0$. For degree reasons, the only possible nontrivial differentials supported by these classes would be of the form

(1) $d_2(\tau^{2n}h_b)\overset{?}{=}\rho\tau^{2n-1}h_0^2h_b$ for $b\in\{2,3\}$,

(2) $d_3(\tau^{2n}h_3)\overset{?}{=}\rho\tau^{2n-1}h_0^3h_3$

with $n\geq 1$. The first does not hold, as τ^2 and h_b are d_2-cycles. The second holds only when $\lambda(q)=3$, and this is accounted for in the theorem statement. □

7.6 The p-adic rationals

We now work over $F=\mathbb{Q}_p$, the p-adic rationals. This is very similar to the case where $F=\mathbb{F}_q$, only where the additional input necessary to understand differentials out of $\mathrm{Ext}^0_{\mathbb{Q}_p}$ comes from work of Ormsby [2011] for p odd and Ormsby and Østvær [2013] for $p=2$. The case where p is odd turns out to entirely reduce to what we have already done.

Lemma 7.6.1 *There are the following differentials in the* $\mathbb{Q}_p$*-motivic Adams spectral sequence*:

(1) *If* $p\equiv 1 \pmod 4$, *then* $d_{a(q)+s}(\tau^{2^s})=u\tau^{2^s-1}h_0^{a(q)+s}$;

(2) *If* $p\equiv 3 \pmod 4$, *then* $d_{\lambda(q)+s}(\tau^{2^{s+1}})=\rho\tau^{2^{s+1}-1}h_0^{\lambda(q)+s}$.

Proof Ormsby [2011, Theorem 5.2] produces identical differentials in the $\mathbb{Q}_p$-motivic Adams spectral sequence for the Brown–Peterson spectrum $\mathrm{BP}\langle 0\rangle$. The claimed differentials follow by naturality. □

We may summarize the situation as follows.

Theorem 7.6.2 *Fix an odd prime* p, *and consider the facts outlined about the* $\mathbb{F}_p$*-motivic Adams spectral sequence in Theorems 7.5.3 and 7.5.6. The same facts hold for the* $\mathbb{Q}_p$*-motivic Adams spectral sequence upon tensoring with* $\mathbb{F}_p\{1,\pi\}$.

Proof The class π is a permanent cycle for degree reasons, and the differentials given in Lemma 7.6.1 agree with those given in Lemmas 7.5.2 and 7.5.5. All of the work carried out over $\mathbb{F}_p$ then goes through verbatim, only where everything in sight has a twin copy indexed by π. □

Remark 7.6.3 The somewhat awkward phrasing of Theorem 7.6.2 is necessary as we did not wish to repeat two verbatim copies of both Theorems 7.5.3 and 7.5.6, but we have not shown that the 1-line of the $\mathbb{Q}_p$-motivic Adams spectral sequence is a direct sum of two copies of the 1-line of the $\mathbb{F}_p$-motivic Adams spectral sequence. The possible failure of this arises from the fact that when $p \equiv 3 \pmod 4$, the classes $\rho\tau^{2n}h_b$ for $b \geq 4$ could support different higher differentials over $\mathbb{F}_p$ and $\mathbb{Q}_p$. ◁

The case where $p = 2$ requires a separate analysis. Recall that

$$\mathrm{Ext}_{\mathbb{Q}_2} = \mathrm{Ext}_{(3)}\{1\} \oplus \mathrm{Ext}_{(1)}\{u, \pi\}.$$

Lemma 7.6.4 *We may identify*

$$\mathrm{Ext}^0_{(3)} = \mathbb{F}_2(\tau^4, \rho\tau^2, \rho^2\tau, \rho^2\tau^3, \rho) \subset \mathbb{F}_2[\tau, \rho]/(\rho^3),$$

and $\mathrm{Ext}^1_{(3)}$ *is the tensor product of* $\mathbb{F}_2[\tau^4]$ *with the direct sum of the modules*

$$\mathbb{F}_2\{h_0, \tau^2h_0, \rho^2\tau h_0, \rho^2\tau^3h_0\},$$
$$\mathbb{F}_2\{1, \rho\} \otimes \mathbb{F}_2\{\tau h_1\} \oplus \mathbb{F}_2\{\rho\tau^3h_1\} \oplus \mathbb{F}_2\{1, \rho, \rho^2, \rho\tau^2, \rho^2\tau^2, \rho^2\tau^3\} \otimes \mathbb{F}_2\{h_1\},$$
$$\mathbb{F}_2\{1, \rho, \rho^2, \rho^2\tau, \rho^2\tau^3, \rho\tau^2, \rho^2\tau^2\} \otimes \mathbb{F}_2\{h_b : b \geq 2\}.$$

Proof As with Lemma 7.5.4, this follows from our computation of $\mathrm{Ext}_{\mathbb{R}}$ via the recipe in Remark 4.1.5, or via the ρ-Bockstein spectral sequence; here the relevant ρ-Bockstein differentials are $d_1(\tau) = \rho h_0$ and $d_2(\tau^2) = \rho^2\tau h_1$. □

Lemma 7.6.5 *The classes*

$$\tau^{4n+1}\rho^2, \quad \tau^{2n}\rho, \quad \tau^{4n+3}\rho^2, \quad \pi\tau^n, \quad u, \quad u\tau^{2n+1}$$

are permanent cycles. There are differentials

$$d_{4+r}(\tau^{2^{r+2}}) = \pi\tau^{2^{r+2}-1}h_0^{4+r}, \quad d_{3+r}(u\tau^{2^{r+1}}) = \rho^2\tau^{2^{r+1}-1}h_0^{3+r}, \quad d_{3+r}(\tau^{2^{r+1}}h_0) = \pi\tau^{2^{r+1}-1}h_0^{4+r}$$

for all $r \geq 0$.

Proof Ormsby and Østvær [2013, Lemma 5.7] compute differentials in the $\mathbb{Q}_2$-motivic Adams spectral sequence for $\mathrm{BP}\langle 0\rangle$. The claimed facts follow by comparison. □

Theorem 7.6.6 *The* 1*-line of the* $\mathbb{Q}_2$*-motivic Adams spectral sequence supports the following nontrivial differentials*:

source	d_r	target	constraints
$\tau^{2n}h_0$	$d_{3+\nu_2(n)}$	$\pi\tau^{2n-1}h_0^{4+\nu_2(n)}$	$n \geq 1$
$\tau^{4n}h_b$	d_2	$\tau^{4n}(h_0 + \rho h_1)h_{b-1}^2$	$n \geq 0, b \geq 4$
$\rho\tau^{2n}h_b$	d_2	$\rho^2\tau^{2n}h_1h_{b-1}^2$	$n \geq 0, b \geq 5$
$u\tau^n h_b$	d_2	$u\tau^n h_0h_{b-1}^2$	$n \geq 0, b \geq 4$
$\pi\tau^n h_b$	d_2	$\pi\tau^n h_0h_{b-1}^2$	$n \geq 0, b \geq 4$
$u\tau^{4n+2}h_3$	d_3	$\rho^2\tau^{4n+1}h_0^3h_3$	$n \geq 0$

After all the d_2-differentials have been run, the 1-line of the E_3-page of the $\mathbb{Q}_2$-motivic Adams spectral sequence has a basis given by the classes in the following table:

class	constraints
h_0	
$\rho^\delta\tau^{4n}h_b$	$n\geq 0, \delta\in\{0,1,2\}, b\in\{1,2,3\}$
$\rho^2\tau^{2n+1}h_0$	$n\geq 0$
$\rho^\epsilon\tau^{4n+1}h_1$	$n\geq 0, \epsilon\in\{0,1\}$
$\rho^{1+\epsilon}\tau^{4n+3}h_1$	$n\geq 0, \epsilon\in\{0,1\}$
$\rho^{1+\epsilon}\tau^{4n+2}h_1$	$n\geq 0, \epsilon\in\{0,1\}$
uh_0	
$u\tau^{2n+1}h_0$	$n\geq 0$
$u\tau^n h_b$	$n\geq 0, b\in\{1,2\}$
$u\tau^{2n+1}h_3$	$n\geq 0$
$u\tau^{4n}h_3$	$n\geq 0$
$\pi\tau^n h_b$	$n\geq 0, b\in\{0,1,2,3\}$
$u^\epsilon\tau^{2n}h_0$	$n\geq 1, \epsilon\in\{0,1\}$
$u\tau^{4n+2}h_3$	$n\geq 0$
$\rho^{1+\epsilon}\tau^{4n}h_4$	$n\geq 0, \epsilon\in\{0,1\}$
$\rho^2\tau^{4n}h_b$	$n\geq 0, b\geq 5$

Of these, the classes in the first region are permanent cycles, the classes $u^\epsilon\tau^{2n}h_0$ with $n\geq 1$ and $\epsilon\in\{0,1\}$, as well as $u\tau^{4n+2}h_3$ with $n\geq 0$, support higher differentials, and we leave open the fate of the classes $\rho^{1+\epsilon}\tau^{4n}h_4$ and $\rho^2\tau^{4n}h_b$ for $n\geq 0$, $\epsilon\in\{0,1\}$, and $b\geq 5$.

Proof The given differentials follow by combining Theorem 7.3.1 with Lemma 7.6.5. For example,

$$d_2(\rho\tau^{2n}h_b)=\rho\tau^{2n}\cdot d_2(h_b)=\rho\tau^{2n}\cdot(h_0+\rho h_1)h_{b-1}^2=\rho^2\tau^{2n}h_1h_{b-1}^2$$

for $b\geq 4$, which is nonzero precisely when $b\geq 5$; as another example,

$$d_3(u\tau^{4n+2}h_3)=d_3(u\tau^2)\cdot\tau^{4n}h_3=\rho^2\tau\cdot\tau^{4n}h_3=\rho^2\tau^{4n+1}h_3.$$

We must verify that all d_2-differentials are accounted for in this table; the claimed description of the E_3-page follows quickly. We must also verify that the classes we give as permanent cycles are indeed permanent cycles. It suffices to verify the latter.

We may cut down the number of classes to consider by taking into account the classes which are products of the permanent cycles given in Lemma 7.6.5 with some other class. After this reduction, degree considerations rule out all differentials except for possibly

(1) $d_r(\tau^{4n+1}h_1)\overset{?}{=}\tau^{4n}h_0^{r+1}$,

(2) $d_r(\rho\tau^{4n+3}h_1)\in\mathbb{F}_2\{u,\pi\}\otimes\mathbb{F}_2\{\tau^{4n+2}h_0^{r+1}\}$,

(3) $d_r(\rho\tau^{4n+2}h_1)\overset{?}{=}\rho^2\tau^{4n+1}h_0^{r+1}$,

(4) $d_r(u\tau^{2n}h_1) \overset{?}{=} \rho^2\tau^{2n-1}h_0^{r+1}$

with $n \geq 0$, and in the fourth case $n \geq 1$. In all cases, the possible nonzero targets are present and not boundaries in Ormsby and Østvær's computation [2013] of the Adams spectral sequence for the $\mathbb{Q}_2$-motivic $\mathrm{BP}\langle 0\rangle$, so by naturality they cannot be boundaries in the Adams spectral sequence for the sphere. Thus these possible nonzero differentials are in fact not possible, yielding the theorem. □

7.7 The rational numbers

We end by considering the case $F = \mathbb{Q}$. By naturality, this gives information over arbitrary fields of characteristic zero. Recall the functions ε and λ defined in Definition 7.5.1.

Theorem 7.7.1 *The* 1*-line of the* E_3*-page of the* $\mathbb{Q}$*-motivic Adams spectral sequence is given by a direct sum of that for the* $\mathbb{R}$*-motivic Adams spectral sequence with the classes in the following table, where* p *ranges through all primes*:

class	*constraints*
$\tau^n h_b[2]$	$n \geq 0, b \in \{0,1,2,3\}$
$h_0[p]$	$p \equiv 1 \pmod 4$
$\tau^n h_1[p]$	$p \equiv 1 \pmod 4, n \geq 0$
$\tau^{2n} h_2[p]$	$p \equiv 1 \pmod 4, n \geq 0$
$\tau^{2n+1} h_2[p]$	$p \equiv 1 \pmod 4, n \geq 0, \varepsilon(p) \geq 3$
$\tau^{4n} h_3[p]$	$p \equiv 1 \pmod 4, n \geq 0$
$\tau^{4n+2} h_3[p]$	$p \equiv 1 \pmod 4, n \geq 0, \varepsilon(p) \geq 3$
$\tau^{2n+1} h_3[p]$	$p \equiv 1 \pmod 4, n \geq 0, \varepsilon(p) \geq 4$
$\tau^n h_b a_p$	$p \equiv 1 \pmod 4, n \geq 0, b \in \{0,1,2,3\}$
$h_0 u_p$	$p \equiv 3 \pmod 4$
$\tau^{2n} h_b u_p$	$p \equiv 3 \pmod 4, n \geq 0, b \in \{1,2\}$
$\tau^{4n} h_3 u_p$	$p \equiv 3 \pmod 4, n \geq 0$
$\tau^{4n+2} h_3 u_p$	$p \equiv 3 \pmod 4, n \geq 0, \lambda(p) \geq 4$
$\rho\tau^{2n} h_b u_p$	$p \equiv 3 \pmod 4, n \geq 0, b \in \{1,2,3\}$
$\rho\tau^{2n+1} h_b u_p$	$p \equiv 3 \pmod 4, n \geq 0, b \in \{1,2,3,4\}$
$\rho^\epsilon \tau^{4n+1} h_1 u_p$	$p \equiv 3 \pmod 4, n \geq 0, \epsilon \in \{0,1\}$
$\tau^{2n} h_0[p]$	$p \equiv 1 \pmod 4, n \geq 1$
$\tau^{2n+1} h_0[p]$	$p \equiv 1 \pmod 4, n \geq 1, \varepsilon(p) \geq 3$
$\tau^{4n+2} h_3[p]$	$p \equiv 1 \pmod 4, n \geq 0, \varepsilon(p) = 2$
$\tau^{2n+1} h_3[p]$	$p \equiv 1 \pmod 4, n \geq 0, \varepsilon(p) = 3$
$\tau^{4n+2} h_3 u_p$	$p \equiv 3 \pmod 4, n \geq 0, \lambda(p) = 3$
$\tau^{2n} h_0 u_p$	$p \equiv 3 \pmod 4, n \geq 1$
$\rho\tau^{2n} h_b u_p$	$p \equiv 3 \pmod 4, n \geq 0, b \geq 4$

Moreover, we have the following information about higher differentials. The classes in the first region of this table are permanent cycles, as are the classes h_a and $\tau^{\lfloor 2^{a-1}(4n+1)\rfloor}h_a$ for $a \le 3$. The classes in the second region of this table support higher differentials, as do the classes in $\mathrm{Ext}^1_{\mathbb{R}}$, which must support higher differentials by Theorem 7.4.9. We leave open the fate of the classes in the third region of this table, as well as the possibility of exotic higher differentials on the classes ρh_4 and $\rho^{2^a-\psi(a)}\tau^{2^{a-1}(4n+1)}h_a$ for $a \ge 4$.

Proof Recall the splitting

$$\mathrm{Ext}_{\mathbb{Q}} = \mathrm{Ext}_{\mathbb{R}} \oplus \mathrm{Ext}_{(1)}\{[2]\} \oplus \mathrm{Ext}_{(1)}\{[p], a_p : p \equiv 1 \pmod 4\} \oplus \mathrm{Ext}_{(2)}\{u_p : p \equiv 3 \pmod 4\}$$

implied by Lemma 7.1.1. As in the proof of Proposition 7.1.3, each of these summands is itself either $\mathrm{Ext}_{\mathbb{R}}$ or an identifiable summand of some corresponding $\mathrm{Ext}_{\mathbb{Q}_p}$; for p odd, this summand looks like $\mathrm{Ext}_{\mathbb{F}_p}$. We may thus read the given table off the information given in Theorems 7.4.9, 7.6.2 (with Theorems 7.5.3 and 7.5.6), and 7.6.6, provided we verify the following claim: if $\alpha[p] \in \mathrm{Ext}^1_{\mathbb{Q}}$ is a class in stem $s \le 6$, then $\alpha[p]$ or αu_p is a d_r-cycle if and only if it projects to a d_r-cycle in the $\mathbb{Q}_p$-motivic Adams spectral sequence; and, likewise, if $\alpha \in \mathrm{Ext}^1_{\mathbb{R}}$ is a class in stem $s \le 7$, then α is a d_r-cycle in the $\mathbb{Q}$-motivic Adams spectral sequence if and only if it projects to a d_r-cycle in the $\mathbb{R}$-motivic Adams spectral sequence.

As in the proofs of Theorems 7.5.3, 7.5.6, and 7.6.6, differentials on the classes $\alpha[p]$ and αu_p in stems $s \le 6$ are completely determined by the structure of differentials on the classes $[p]\tau^{2^i}$ and $u_p\tau^{2^i}$ in the $\mathbb{Q}$-motivic Adams spectral sequence for $\mathrm{BP}\langle 0\rangle$, together with the fact that h_0, h_1, h_2, and h_3 are permanent cycles. The $\mathbb{Q}$-motivic Adams spectral sequence for $\mathrm{BP}\langle 0\rangle$ was computed in [Ormsby and Østvær 2013, Theorem 5.8]. We find that differentials on the classes $[p]\tau^{2^i}$ and $u_p\tau^{2^i}$ in the $\mathbb{Q}$-motivic Adams spectral sequence for $\mathrm{BP}\langle 0\rangle$ are entirely detected over $\mathbb{Q}_p$, and our first claim follows. That the classes $h_a \in \mathrm{Ext}^1_{\mathbb{R}}$ for $a \le 3$ are permanent cycles was seen in Lemma 7.2.2, and the classes $\tau^{\lfloor 2^{a-1}(4n+1)\rfloor}h_a \in \mathrm{Ext}^1_{\mathbb{R}}$ must be permanent cycles for $a \le 3$ as there is no room for exotic higher differentials. □

References

[Adams 1958] **J F Adams**, *On the structure and applications of the Steenrod algebra*, Comment. Math. Helv. 32 (1958) 180–214 MR Zbl

[Adams 1960] **J F Adams**, *On the non-existence of elements of Hopf invariant one*, Ann. of Math. 72 (1960) 20–104 MR Zbl

[Adams 1962] **J F Adams**, *Vector fields on spheres*, Ann. of Math. 75 (1962) 603–632 MR Zbl

[Andrews and Miller 2017] **M Andrews**, **H Miller**, *Inverting the Hopf map*, J. Topol. 10 (2017) 1145–1168 MR Zbl

[Asok et al. 2017] **A Asok**, **B Doran**, **J Fasel**, *Smooth models of motivic spheres and the clutching construction*, Int. Math. Res. Not. 2017 (2017) 1890–1925 MR Zbl

[Bachmann 2022] **T Bachmann**, *η-periodic motivic stable homotopy theory over Dedekind domains*, J. Topol. 15 (2022) 950–971 MR Zbl

[Bachmann and Hopkins 2020] **T Bachmann**, **M J Hopkins**, *η-periodic motivic stable homotopy theory over fields*, preprint (2020) arXiv 2005.06778

[Bachmann et al. 2022] **T Bachmann**, **H J Kong**, **G Wang**, **Z Xu**, *The Chow t-structure on the ∞-category of motivic spectra*, Ann. of Math. 195 (2022) 707–773 MR Zbl

[Balderrama 2023] **W Balderrama**, *Algebraic theories of power operations*, J. Topol. 16 (2023) 1543–1640 Zbl

[Behrens and Shah 2020] **M Behrens**, **J Shah**, *C_2-equivariant stable homotopy from real motivic stable homotopy*, Ann. K-Theory 5 (2020) 411–464 MR Zbl

[Belmont and Isaksen 2022] **E Belmont**, **D C Isaksen**, *$\mathbb{R}$-motivic stable stems*, J. Topol. 15 (2022) 1755–1793 MR Zbl

[Belmont et al. 2021] **E Belmont**, **B J Guillou**, **D C Isaksen**, *C_2-equivariant and $\mathbb{R}$-motivic stable stems, II*, Proc. Amer. Math. Soc. 149 (2021) 53–61 MR Zbl

[Bousfield et al. 1966] **A K Bousfield**, **E B Curtis**, **D M Kan**, **D G Quillen**, **D L Rector**, **J W Schlesinger**, *The mod $-$ p lower central series and the Adams spectral sequence*, Topology 5 (1966) 331–342 MR Zbl

[Brown and Gitler 1973] **E H Brown, Jr**, **S Gitler**, *A spectrum whose cohomology is a certain cyclic module over the Steenrod algebra*, Topology 12 (1973) 283–295 MR Zbl

[Bruner 1986a] **R R Bruner**, *The homotopy theory of H_∞ ring spectra*, from "H_∞ ring spectra and their applications" (R R Bruner, J P May, J E McClure, M Steinberger, editors), Lecture Notes in Math. 1176, Springer (1986) Chapter IV, 88–128 MR Zbl

[Bruner 1986b] **R R Bruner**, *The Adams spectral sequence of H_∞ ring spectra*, from "H_∞ ring spectra and their applications" (R R Bruner, J P May, J E McClure, M Steinberger, editors), Lecture Notes in Math. 1176, Springer (1986) Chapter VI, 169–214 MR Zbl

[Bruner 1988] **R Bruner**, *An example in the cohomology of augmented algebras*, J. Pure Appl. Algebra 55 (1988) 81–84 MR Zbl

[Burklund et al. 2020] **R Burklund**, **J Hahn**, **A Senger**, *Galois reconstruction of Artin–Tate $\mathbb{R}$-motivic spectra*, preprint (2020) arXiv 2010.10325

[Chen 2011] **T-W Chen**, *Determination of $\mathrm{Ext}^{5,*}_{\mathscr{A}}(\mathbb{Z}/2,\mathbb{Z}/2)$*, Topology Appl. 158 (2011) 660–689 MR Zbl

[Cohen 1985] **R L Cohen**, *The immersion conjecture for differentiable manifolds*, Ann. of Math. 122 (1985) 237–328 MR Zbl

[Culver and Quigley 2021] **D L Culver**, **J D Quigley**, *kq-resolutions, I*, Trans. Amer. Math. Soc. 374 (2021) 4655–4710 MR Zbl

[Curtis 1971] **E B Curtis**, *Simplicial homotopy theory*, Adv. Math. 6 (1971) 107–209 MR Zbl

[Curtis et al. 1987] **E B Curtis**, **P Goerss**, **M Mahowald**, **R J Milgram**, *Calculations of unstable Adams E_2 terms for spheres*, from "Algebraic topology" (H R Miller, D C Ravenel, editors), Lecture Notes in Math. 1286, Springer (1987) 208–266 MR Zbl

[Dugger and Isaksen 2010] **D Dugger**, **D C Isaksen**, *The motivic Adams spectral sequence*, Geom. Topol. 14 (2010) 967–1014 MR Zbl

[Dugger and Isaksen 2013] **D Dugger**, **D C Isaksen**, *Motivic Hopf elements and relations*, New York J. Math. 19 (2013) 823–871 MR Zbl

[Dugger and Isaksen 2017a] **D Dugger**, **D C Isaksen**, *Low-dimensional Milnor–Witt stems over $\mathbb{R}$*, Ann. K-Theory 2 (2017) 175–210 MR Zbl

[Dugger and Isaksen 2017b] **D Dugger**, **D C Isaksen**, *$\mathbb{Z}/2$-equivariant and $\mathbb{R}$-motivic stable stems*, Proc. Amer. Math. Soc. 145 (2017) 3617–3627 MR Zbl

[Gheorghe et al. 2021] **B Gheorghe**, **G Wang**, **Z Xu**, *The special fiber of the motivic deformation of the stable homotopy category is algebraic*, Acta Math. 226 (2021) 319–407 MR Zbl

[Goerss 1999] **P G Goerss**, *Hopf rings, Dieudonné modules, and $E_*\Omega^2 S^3$*, from "Homotopy invariant algebraic structures" (J-P Meyer, J Morava, W S Wilson, editors), Contemp. Math. 239, Amer. Math. Soc., Providence, RI (1999) 115–174 MR Zbl

[Goerss et al. 1986] **P G Goerss**, **J D S Jones**, **M E Mahowald**, *Some generalized Brown–Gitler spectra*, Trans. Amer. Math. Soc. 294 (1986) 113–132 MR Zbl

[Greenlees 1988] **J P C Greenlees**, *Stable maps into free G-spaces*, Trans. Amer. Math. Soc. 310 (1988) 199–215 MR Zbl

[Guillou and Isaksen 2015] **B J Guillou**, **D C Isaksen**, *The η-local motivic sphere*, J. Pure Appl. Algebra 219 (2015) 4728–4756 MR Zbl

[Guillou and Isaksen 2016] **B J Guillou**, **D C Isaksen**, *The η-inverted $\mathbb{R}$-motivic sphere*, Algebr. Geom. Topol. 16 (2016) 3005–3027 MR Zbl

[Guillou and Isaksen 2020] **B J Guillou**, **D C Isaksen**, *The Bredon–Landweber region in C_2-equivariant stable homotopy groups*, Doc. Math. 25 (2020) 1865–1880 MR Zbl

[Guillou et al. 2020] **B J Guillou**, **M A Hill**, **D C Isaksen**, **D C Ravenel**, *The cohomology of C_2-equivariant $\mathscr{A}(1)$ and the homotopy of ko_{C_2}*, Tunis. J. Math. 2 (2020) 567–632 MR Zbl

[Hill 2011] **M A Hill**, *Ext and the motivic Steenrod algebra over $\mathbb{R}$*, J. Pure Appl. Algebra 215 (2011) 715–727 MR Zbl

[Hoyois et al. 2017] **M Hoyois**, **S Kelly**, **P A Østvær**, *The motivic Steenrod algebra in positive characteristic*, J. Eur. Math. Soc. 19 (2017) 3813–3849 MR Zbl

[Hu and Kriz 2001] **P Hu**, **I Kriz**, *Real-oriented homotopy theory and an analogue of the Adams–Novikov spectral sequence*, Topology 40 (2001) 317–399 MR Zbl

[Hu et al. 2011a] **P Hu**, **I Kriz**, **K Ormsby**, *Convergence of the motivic Adams spectral sequence*, J. K-Theory 7 (2011) 573–596 MR Zbl

[Hu et al. 2011b] **P Hu**, **I Kriz**, **K Ormsby**, *Remarks on motivic homotopy theory over algebraically closed fields*, J. K-Theory 7 (2011) 55–89 MR Zbl

[Hunter and Kuhn 1999] **D J Hunter**, **N J Kuhn**, *Mahowaldean families of elements in stable homotopy groups revisited*, Math. Proc. Cambridge Philos. Soc. 127 (1999) 237–251 MR Zbl

[Isaksen 2019] **D C Isaksen**, *Stable stems*, Mem. Amer. Math. Soc. 1269, Amer. Math. Soc., Providence, RI (2019) MR Zbl

[Isaksen and Østvær 2020] **D C Isaksen**, **P A Østvær**, *Motivic stable homotopy groups*, from "Handbook of homotopy theory" (H Miller, editor), CRC Press, Boca Raton, FL (2020) 757–791 MR Zbl

[Isaksen et al. 2023] **D C Isaksen**, **G Wang**, **Z Xu**, *Stable homotopy groups of spheres: from dimension 0 to 90*, Publ. Math. Inst. Hautes Études Sci. 137 (2023) 107–243 MR Zbl

[James 1957] **I M James**, *On the suspension sequence*, Ann. of Math. 65 (1957) 74–107 MR Zbl

[James 1958] **I M James**, *Cross-sections of Stiefel manifolds*, Proc. Lond. Math. Soc. 8 (1958) 536–547 MR Zbl

[James 1959] **I M James**, *Spaces associated with Stiefel manifolds*, Proc. Lond. Math. Soc. 9 (1959) 115–140 MR Zbl

[Kylling 2015] **J I Kylling**, *Hermitian K-theory of finite fields via the motivic Adams spectral sequence*, master's thesis, University of Oslo (2015) Available at https://www.duo.uio.no/bitstream/handle/10852/45345/JonasIrgensKylling.pdf?sequence=11&isAllowed=y

[Kylling and Wilson 2019] **J I Kylling**, **G M Wilson**, *Strong convergence in the motivic Adams spectral sequence*, preprint (2019) arXiv 1901.03399

[Lin 1980] **W H Lin**, *On conjectures of Mahowald, Segal and Sullivan*, Math. Proc. Cambridge Philos. Soc. 87 (1980) 449–458 MR Zbl

[Lin 2008] **W-H Lin**, $\mathrm{Ext}_A^{4,*}(\mathbb{Z}/2,\mathbb{Z}/2)$ *and* $\mathrm{Ext}_A^{5,*}(\mathbb{Z}/2,\mathbb{Z}/2)$, Topology Appl. 155 (2008) 459–496 MR Zbl

[Lurie 2011] **J Lurie**, *Derived algebraic geometry, XIII: Rational and p-adic homotopy theory*, preprint (2011) Available at https://www.math.ias.edu/~lurie/papers/DAG-XIII.pdf

[Mahowald 1977] **M Mahowald**, *A new infinite family in ${}_2\pi_*{}^s$*, Topology 16 (1977) 249–256 MR Zbl

[Mahowald 1981] **M Mahowald**, *bo-resolutions*, Pacific J. Math. 92 (1981) 365–383 MR Zbl

[Mandell 2001] **M A Mandell**, *E_∞ algebras and p-adic homotopy theory*, Topology 40 (2001) 43–94 MR Zbl

[Massey and Peterson 1965] **W S Massey**, **F P Peterson**, *The cohomology structure of certain fibre spaces, I*, Topology 4 (1965) 47–65 MR Zbl

[Milnor 1970] **J Milnor**, *Algebraic K-theory and quadratic forms*, Invent. Math. 9 (1970) 318–344 MR Zbl

[Morel 1999] **F Morel**, *Suite spectrale d'Adams et invariants cohomologiques des formes quadratiques*, C. R. Acad. Sci. Paris Sér. I Math. 328 (1999) 963–968 MR Zbl

[Morel 2004] **F Morel**, *An introduction to $\mathbb{A}^1$-homotopy theory*, from "Contemporary developments in algebraic K-theory" (M Karoubi, A O Kuku, C Pedrini, editors), ICTP Lect. Notes 15, Abdus Salam Int. Cent. Theoret. Phys., Trieste, Italy (2004) 357–441 MR Zbl

[Morel and Voevodsky 1999] **F Morel**, **V Voevodsky**, *$\mathbb{A}^1$-homotopy theory of schemes*, Inst. Hautes Études Sci. Publ. Math. 90 (1999) 45–143 MR Zbl

[Ormsby 2011] **K M Ormsby**, *Motivic invariants of p-adic fields*, J. K-Theory 7 (2011) 597–618 MR Zbl

[Ormsby and Østvær 2013] **K M Ormsby**, **P A Østvær**, *Motivic Brown–Peterson invariants of the rationals*, Geom. Topol. 17 (2013) 1671–1706 MR Zbl

[Ormsby and Röndigs 2020] **K Ormsby**, **O Röndigs**, *The homotopy groups of the η-periodic motivic sphere spectrum*, Pacific J. Math. 306 (2020) 679–697 MR Zbl

[Palmieri 2007] **J H Palmieri**, *The lambda algebra and* Sq^0, from "Proceedings of the School and Conference in Algebraic Topology" (J Hubbuck, N H V Hung, L Schwartz, editors), Geom. Topol. Monogr. 11, Geom. Topol. Publ., Coventry (2007) 201–216 MR Zbl

[Polishchuk and Positselski 2005] **A Polishchuk**, **L Positselski**, *Quadratic algebras*, Univ. Lect. Ser. 37, Amer. Math. Soc., Providence, RI (2005) MR Zbl

[Priddy 1970] **S B Priddy**, *Koszul resolutions*, Trans. Amer. Math. Soc. 152 (1970) 39–60 MR Zbl

[Rezk 2012] **C Rezk**, *Rings of power operations for Morava E-theories are Koszul*, preprint (2012) arXiv 1204.4831

[Riou 2012] **J Riou**, *Opérations de Steenrod motiviques*, preprint (2012) arXiv 1207.3121

[Röndigs and Østvær 2016] **O Röndigs**, **P A Østvær**, *Slices of hermitian K-theory and Milnor's conjecture on quadratic forms*, Geom. Topol. 20 (2016) 1157–1212 MR Zbl

[Röndigs et al. 2019] **O Röndigs**, **M Spitzweck**, **P A Østvær**, *The first stable homotopy groups of motivic spheres*, Ann. of Math. 189 (2019) 1–74 MR Zbl

[Röndigs et al. 2021] **O Röndigs**, **M Spitzweck**, **P A Østvær**, *The second stable homotopy groups of motivic spheres* (2021) arXiv 2103.17116 To appear in Duke Math. J.

[Shimamoto 1984] **D H Shimamoto**, *An integral version of the Brown–Gitler spectrum*, Trans. Amer. Math. Soc. 283 (1984) 383–421 MR Zbl

[Steenrod 1962] **N E Steenrod**, *Cohomology operations*, Ann. of Math. Stud. 50, Princeton Univ. Press (1962) MR Zbl

[Tangora 1985] **M C Tangora**, *Computing the homology of the lambda algebra*, Mem. Amer. Math. Soc. 337, Amer. Math. Soc., Providence, RI (1985) MR Zbl

[Tangora 1993] **M C Tangora**, *Some remarks on products in* Ext, from "Algebraic topology" (M C Tangora, editor), Contemp. Math. 146, Amer. Math. Soc., Providence, RI (1993) 419–429 MR Zbl

[Tangora 1994] **M C Tangora**, *Some Massey products in* Ext, from "Topology and representation theory" (E M Friedlander, M E Mahowald, editors), Contemp. Math. 158, Amer. Math. Soc., Providence, RI (1994) 269–280 MR Zbl

[Tilson 2017] **S Tilson**, *Squaring operations in the $RO(C_2)$-graded and real motivic Adams spectral sequences*, preprint (2017) arXiv 1702.04632

[Voevodsky 2003] **V Voevodsky**, *Reduced power operations in motivic cohomology*, Publ. Math. Inst. Hautes Études Sci. 98 (2003) 1–57 MR Zbl

[Voevodsky 2011] **V Voevodsky**, *On motivic cohomology with $\mathbb{Z}/l$-coefficients*, Ann. of Math. 174 (2011) 401–438 MR Zbl

[Wang 1967] **J S P Wang**, *On the cohomology of the* mod $-$ 2 *Steenrod algebra and the non-existence of elements of Hopf invariant one*, Illinois J. Math. 11 (1967) 480–490 MR Zbl

[Wang and Xu 2016] **G Wang**, **Z Xu**, *The algebraic Atiyah–Hirzebruch spectral sequence of real projective spectra*, preprint (2016) arXiv 1601.02185

[Wang and Xu 2017] **G Wang**, **Z Xu**, *The triviality of the* 61*-stem in the stable homotopy groups of spheres*, Ann. of Math. 186 (2017) 501–580 MR Zbl

[Wilson 2016] **G M Wilson**, *Motivic stable stems over finite fields*, PhD thesis, Rutgers University (2016) Available at `https://www.proquest.com/docview/1844381634`

[Wilson 2018] **G M Wilson**, *The eta-inverted sphere over the rationals*, Algebr. Geom. Topol. 18 (2018) 1857–1881 MR Zbl

[Wilson and Østvær 2017] **G M Wilson**, **P A Østvær**, *Two-complete stable motivic stems over finite fields*, Algebr. Geom. Topol. 17 (2017) 1059–1104 MR Zbl

WB: *Department of Mathematics, University of Virginia*
Charlottesville, VA, United States
DLC: *Max-Planck-Institut für Mathematik*
Bonn, Germany
JDQ: *Department of Mathematics, Cornell University*
Ithaca, NY, United States
Current address: *Department of Mathematics, University of Virginia*
Charlottesville, VA, United States

`williamb@math.uni-bonn.de`, `dominic.culver@gmail.com`, `mbp6pj@virginia.edu`

Proposed: Marc Levine — Received: 19 July 2022
Seconded: Stefan Schwede, Haynes R Miller — Revised: 21 August 2023

Geometry & Topology **29**:3 (2025) 1571–1618
DOI: 10.2140/gt.2025.29.1571
Published: 31 May 2025

Exotic Dehn twists on sums of two contact 3-manifolds

EDUARDO FERNÁNDEZ
JUAN MUÑOZ-ECHÁNIZ

We exhibit the first examples of exotic contactomorphisms with infinite order as elements of the contact mapping class group. These are given by certain Dehn twists on the separating sphere in a connected sum of two closed contact 3-manifolds. We detect these by a combination of hard and soft techniques. We make essential use of an invariant for families of contact structures which generalizes the Kronheimer–Mrowka contact invariant in monopole Floer homology. We then exploit an h-principle for families of convex spheres in tight contact 3-manifolds, from which we establish a parametric version of Colin's decomposition theorem. As a further application, we exhibit new exotic 1-parametric phenomena in overtwisted contact 3-manifolds.

53D35, 57K33, 57R58

1 Introduction

Throughout this article all 3-manifolds are closed, oriented and connected unless otherwise noted, and all contact structures on 3-manifolds are co-oriented and positive.

1.1 Main result

A fundamental problem in contact topology is to understand the isotopy classes of contact diffeomorphisms, usually called "contactomorphisms", of a contact manifold. The following is a longstanding open question in all dimensions:

Question 1.1 *Do there exist exotic contactomorphisms with infinite order as elements in the contact mapping class group*?

In this article we answer this question in the *affirmative* in dimension *three*. Here, and throughout the article, by *exotic* we will mean *nontrivial in the contact category* but *formally trivial* (and, in particular, *trivial in the smooth category*). See Section 2.3 and below for further details. We consider a contact 3-manifold given by the connected sum of two contact 3-manifolds $(Y_\#, \xi_\#) := (Y_-, \xi_-) \# (Y_+, \xi_+)$. Recall that the connected sum is built by removing Darboux balls $B_\pm \subset Y_\pm$ and gluing the complements $Y \setminus B_\pm$ by an orientation-reversing diffeomorphism of their boundary spheres which preserves their characteristic foliations. Reparametrization of one of the spheres provides a $U(1)$ worth of choices for gluing, and thus $(Y_\#, \xi_\#)$ naturally belongs in a *family* of contact 3-manifolds

$$(Y_\#, \xi_\#) \hookrightarrow \mathcal{Y}_\# \to U(1).$$

The monodromy of this family is realized by a contactomorphism of $(Y_\#, \xi_\#)$, well-defined up to contact isotopy. Its underlying diffeomorphism is the *Dehn twist* on the separating sphere $S_\#$ in the neck of the connected sum $Y_\# = Y_- \# Y_+$. We denote this contactomorphism by $\tau_{S_\#}$ and call it the *contact Dehn twist* on $S_\#$. Unlike previous constructions of contactomorphisms, the contact Dehn twist is a *local symmetry* of an arbitrarily small neighborhood of a 2-sphere (see Section 3 for further details). As a diffeomorphism, the Dehn twist can be isotoped so that it is supported on a neighborhood $[0,1] \times S^2$ of $S_\# \simeq S^2$ on which it acts as $[0,1] \times S^2 \ni (t, p) \mapsto (t, R_{\theta(t)}(p))$, where R_φ denotes the rotation of angle φ along the z axis in $\mathbb{R}^3$, and $\theta : [0,1] \to [0, 2\pi]$ is a smooth function with $\theta \equiv 0$ near $t = 0$ and $\theta \equiv 2\pi$ near $t = 1$. Because $\pi_1 \mathrm{SO}(3) = \mathbb{Z}/2$, the 2-fold iterate $\tau_{S_\#}^2$ is *smoothly* isotopic to the identity, but it remains to be understood whether:

Question 1.2 *Is $\tau_{S_\#}^2$ contact isotopic to the identity*?

Associated to the contact structures $\xi_\pm$ we have their Kronheimer–Mrokwa contact invariants $\boldsymbol{c}(\xi_\pm) \in \widecheck{\mathrm{HM}}(-Y_\pm)$; see [Kronheimer and Mrowka 1997; Kronheimer et al. 2007]. These are canonical elements (defined up to sign) in the "to" flavor of the monopole Floer homology of $-Y_\pm$. The contact invariant was also defined in the setting of Heegaard–Floer homology by Ozsváth and Szabó [2005]. Under the isomorphism between the monopole and Heegaard–Floer homologies [Kutluhan et al. 2020; Colin et al. 2011] the contact invariants agree. Throughout this article we only consider monopole Floer homology and the contact invariant with *coefficients* in $\mathbb{Q}$, for simplicity. Our main result is the following:

Theorem 1.3 *Let $(Y_\pm, \xi_\pm)$ be* ***irreducible*** *contact 3-manifolds. Suppose that the Kronheimer–Mrowka contact invariants $\boldsymbol{c}(\xi_\pm)$ do not lie in the image of the U-map*

$$U : \widecheck{\mathrm{HM}}(-Y_\pm) \to \widecheck{\mathrm{HM}}(-Y_\pm).$$

Then:

(A) *The k-fold iterates $\tau_{S_\#}^k$ for $k \geq 1$ of the contact Dehn twist are not contact isotopic to the identity.*

(B) *If the Euler classes of $\xi_\pm$ vanish, then $\tau_{S_\#}^2$ is formally contact isotopic to the identity.*

We now explain the meaning of the assertion in Theorem 1.3(B). Given a contact 3-manifold (Y,ξ), a *formal contactomorphism* of (Y,ξ) consists of a pair (f,F) where f is a diffeomorphism of Y and $F=(F_s)$ is a homotopy through vector bundle isomorphisms $F_s\colon TY\to f^*TY$ such that $F_0=df$ and F_1 preserves ξ. Any contactomorphism f yields a formal contactomorphism, and one says that f is *formally trivial* if f can be deformed to the identity through formal contactomorphisms. A contactomorphism f of (Y,ξ) will be called *exotic* if it is formally contact isotopic to the identity but is not contact isotopic to the identity. Thus exotic contactomorphisms are those which are "geometrically" nontrivial, and not for reasons having to do with the underlying smooth or tangential structures. See Section 2.3 for further context. Thus Theorem 1.3 asserts that $\tau^2_{S_\#}$ and all its iterates are exotic.

Remark 1.4 In fact, we will establish more: the contactomorphism $\tau^2_{S_\#}$ from Theorem 1.3 has infinite order as an element in the *abelianization* of the group

$$\ker(\pi_0\,\mathrm{Cont}(Y,\xi)\to\pi_0\,\mathrm{Diff}(Y)). \tag{1}$$

Remark 1.5 For comparison with Theorem 1.3, whenever either of $(Y_\pm,\xi_\pm)$ is the tight $S^1\times S^2$ or a quotient of tight (S^3,ξ) — eg the lens spaces $L(p,q)$ or the Poincaré sphere $\Sigma(2,3,5)$ — then the squared contact Dehn twist $\tau^2_{S_\#}$ of $(Y_\#,\xi_\#)$ is contact isotopic to the identity; see Lemmas 3.13–3.15.

We also establish an analogous result for connected sums with multiple summands. Let (Y,ξ) be a *tight* 3-manifold. By the prime decomposition theorem combined with Colin's decomposition theorem [1997] (see also [Honda 2002; Ding and Geiges 2007]) we have a unique connected sum decomposition

$$(Y,\xi)\cong(Y_0,\xi_0)\,\#\cdots\#\,(Y_N,\xi_N)$$

into tight contact 3-manifolds (Y_j,ξ_j), where each piece Y_j is a prime 3-manifold. Let $n+1\le N$ be the number of prime summands (Y_j,ξ_j) such that $\boldsymbol{c}(\xi_j)\notin\operatorname{Im}U$ and the Euler class of ξ_j vanishes. Let $\mathcal{C}(Y,\xi)$ (resp. $\Xi(Y,\xi)$) be the space of contact structures (resp. co-oriented 2-plane fields) on Y in the path-component of ξ.

Theorem 1.6 *With (Y,ξ) as above, when $n\ge 1$ there is a $\mathbb{Z}^n$ subgroup in the kernel of*

$$\pi_1\mathcal{C}(Y,\xi)\to\pi_1\Xi(Y,\xi)$$

which induces a $\mathbb{Z}^n$ subgroup in the first singular homology $H_1(\mathcal{C}(Y,\xi);\mathbb{Z})$.

In particular, the exotic subgroup $\mathbb{Z}^n$ exhibited in Theorem 1.6 can be arbitrarily large in the following sense: for every $n\ge 1$ there exists a tight contact 3-manifold, in fact infinitely many, such that the kernel of the previous homomorphism contains a subgroup isomorphic to $\mathbb{Z}^n$.

Remark 1.7 The n homologically independent loops of contact structures that we detect in Theorem 1.6 yield, under the natural map

$$\pi_1\mathcal{C}(Y,\xi)\to\pi_0\,\mathrm{Cont}(Y,\xi),$$

the squared contact Dehn twists on each of the n spheres which separate the $n+1$ prime summands (Y_j,ξ_j). However, we are unable to establish that the corresponding squared contact Dehn twists are nontrivial or that they yield a subgroup $\mathbb{Z}^n \subset \pi_0\,\mathrm{Cont}(Y,\xi)$ when $n \geq 2$, but we conjecture that this should be true. See Remark 6.6.

The proofs of Theorems 1.3 and 1.6 combine rigid obstructions arising from Floer homology with flexibility results. An essential ingredient is a families generalization of the Kronheimer–Mrowka contact invariant in monopole Floer homology, introduced by the second author [Muñoz-Echániz 2024]. This obstructs the existence of sections of a natural fibration given by the *evaluation map* $\mathrm{ev}\colon \mathcal{C}(Y,\xi) \to S^2$ which sends a contact structure to its plane at p, where $p \in Y$ is some fixed point. We combine this machinery with the multiparametric convex surface theory techniques introduced by the first author together with J Martínez-Aguinaga and F Presas [Fernández et al. 2020]. In particular, we use these techniques to establish the following generalization of the much-celebrated decomposition theorem of Colin [1997], which could be of independent interest for contact topologists, and which will be crucial to the proof of Theorem 1.6.

We consider two tight contact 3-manifolds $(Y_\pm,\xi_\pm)$ equipped with Darboux balls $B_\pm \subset (Y_\pm,\xi_\pm)$. Let $\mathcal{C}(Y_\pm,\xi_\pm,B_\pm) \subset \mathcal{C}(Y_\pm,\xi_\pm)$ denote the subspace of contact structures on $Y_\pm$ that coincide with $\xi_\pm$ over $B_\pm$. We consider the evaluation maps $\mathrm{ev}_\pm\colon \mathcal{C}(Y_\pm,\xi_\pm) \to S^2$ which send a contact structure to its plane at the point $p_\pm$ given by the center of $B_\pm$. These maps are fibrations, and the inclusion of $\mathcal{C}(Y_\pm,\xi_\pm,B_\pm)$ into the fiber of $\mathrm{ev}_\pm$ induces a homotopy equivalence. We form the connected sum $(Y_\#,\xi_\#) = (Y_-,\xi_-)\,\#\,(Y_+,\xi_+)$ by carving out the balls $B_\pm$ and gluing together the boundary components thus created. Consider the evaluation map $\mathrm{ev}_\#\colon \mathcal{C}(Y_\#,\xi_\#) \to S^2$ at a point on the "neck" region. We establish the following h-principle type result, which should be regarded as a parametric version of Colin's theorem:

Theorem 1.8 *The inclusion of $\mathcal{C}(Y_-,\xi_-,B_-)\times\mathcal{C}(Y_+,\xi_+,B_+)$ into the fiber of $\mathrm{ev}_\#$ induces a homotopy equivalence. Thus there is a fibration sequence*

$$\mathcal{C}(Y_-,\xi_-,B_-)\times\mathcal{C}(Y_+,\xi_+,B_+) \hookrightarrow \mathcal{C}(Y_\#,\xi_\#) \xrightarrow{\mathrm{ev}_\#} S^2.$$

We refer to Theorem 6.1 for a more general version.

1.2 Examples

We now give examples of irreducible contact 3-manifolds (Y,ξ) such that $\boldsymbol{c}(\xi) \notin \mathrm{Im}\,U$, many of which also have vanishing Euler class.

Example 1.9 (links of singularities) The simplest example is the Brieskorn sphere

$$\Sigma(p,q,r) = \{(x,y,z) \in \mathbb{C}^3 \mid x^p + y^q + z^r = 0 \text{ and } |x|^2+|y|^2+|z|^2 = \epsilon\},$$

where $\epsilon \in \mathbb{R}_{>0}$ is small and $p,q,r \geq 1$ are integers with $1/p+1/q+1/r < 1$, equipped with the contact structure ξ_{sing} induced from the Brieskorn singularity. More generally, we could take any isolated normal surface singularity germ (X, o) and let (Y, ξ_{sing}) be the contact manifold arising as the *link* of the singularity. Neumann [1981] proved that the 3-manifold Y is irreducible. Provided that Y is also a rational homology sphere, then the following are equivalent statements, as proved by Bodnár and Plamenevskaya [2021] and Némethi [2017]:

(a) $\boldsymbol{c}(\xi_{\text{sing}}) \notin \operatorname{Im} U$.

(b) Y is not an L-space.

(c) (X, o) is not a rational singularity.

For instance, all Seifert fibered integral homology spheres excluding S^3 or the Poincaré sphere carry a contact structure ξ_{sing} with the above properties.

Example 1.10 Several surgeries on the figure eight knot are hyperbolic (and hence irreducible) and support contact structures with $c(\xi) \notin \operatorname{Im} U$. Contact structures on these manifolds have been classified by Conway and Min [2020].

Example 1.11 All but one of the $\frac{1}{2}n(n-1)$ tight contact structures supported on $-\Sigma(2,3,6n-1)$, up to isotopy, were classified by Ghiggini and Van Horn-Morris [2016].

1.3 Exotic overtwisted phenomena

Let (Y, ξ) be such that $\boldsymbol{c}(\xi) \notin \operatorname{Im} U$ and ξ has vanishing Euler class. Let $B \subset (Y, \xi)$ be a Darboux ball. From this, one can produce overtwisted contact manifolds by modifying (Y, ξ) by a Lutz twist inside B, or by taking the connected sum (using B) with an overtwisted contact manifold (M, ξ_{ot}). In either case, the squared contact Dehn twist on the boundary of B becomes isotopic to the identity in this new overtwisted manifold, by an application of Eliashberg's h-principle [1989] for overtwisted contact structures. However, this has surprising implications (see Section 7 for the precise statement).

Proposition 1.12 (A) *There exist overtwisted contact 3-manifolds that have an exotic loop of Lutz twist embeddings.*

(B) *There exist overtwisted contact 3-manifolds that have an exotic loop of standard sphere embeddings.*

In other words, (A) says that the h-principle for codimension-0 isocontact embeddings of embedded S^1-families of overtwisted disks fails in 1-parametric families; see [Gromov 1986; Eliashberg and Mishachev 2002]. To the best of our knowledge this is the first example of this nature. On the other hand, (B) says that the h-principle for standard spheres [Fernández et al. 2020] in tight contact 3-manifolds fails in the overtwisted case.

The first known exotic phenomena regarding overtwisted disks in overtwisted contact 3-manifolds are due to Vogel [2018]. He has proved that the space of overtwisted disks in certain overtwisted 3-sphere is disconnected and used this to construct an exotic loop of overtwisted contact structures. By Eliashberg's

h-principle [1989], understanding the homotopy type of the space of overtwisted disks is the only obstacle remaining in order to completely understand the homotopy type of the space of overtwisted contact structures on a 3-manifold. Thus understanding families of overtwisted disks or overtwisted objects bears special importance in 3-dimensional contact topology.

1.4 Context

1.4.1 *h*-principles As with symplectic topology, an ubiquitous theme of contact topology is the contrast between two types of behaviors: flexible (similar to differential topology) and rigid (similar to algebraic geometry). Beyond the tight–overtwisted dichotomy, 3-dimensional contact topology would seem to be dominated by *flexibility*, due to the following h-principle of Eliashberg and Mishachev:

Theorem 1.13 [Eliashberg and Mishachev 2021] *Let $(\mathbb{B}^3, \xi_{\mathrm{st}} = \ker(dz - y\,dx))$ be the standard contact unit 3-ball. Then the inclusion $\mathrm{Cont}(\mathbb{B}^3, \xi_{\mathrm{st}}) \to \mathrm{Diff}(\mathbb{B}^3)$ is a homotopy equivalence.*

Here $\mathrm{Cont}(\mathbb{B}^3, \xi)$ is the group of contactomorphisms of Y fixing a neighborhood of $\partial\mathbb{B}^3$, and likewise for the group of diffeomorphisms $\mathrm{Diff}(\mathbb{B}^3)$. This result was claimed, without a complete proof, by Eliashberg [1992], treating the 0-1 parametric case. The complete proof recently appeared in [Eliashberg and Mishachev 2021]. To give some context, the analogous statement that $\mathrm{Diff}(\mathbb{B}^3) \to \mathrm{Homeo}(\mathbb{B}^3)$ is a homotopy equivalence is equivalent to the Smale conjecture in dimension 3, a deep result proved by Hatcher [1983]. Then an argument due to Cerf [1968] shows that the Smale conjecture implies that $\mathrm{Diff}(Y) \to \mathrm{Homeo}(Y)$ is a homotopy equivalence for all 3-manifolds. Thus the exotic phenomena at the π_0-level which are exhibited in Theorems 1.3–1.6 are in sharp contrast with the above, and unexpected.

Remark 1.14 In 4-dimensional symplectic topology, the statement analogous to the h-principle of Eliashberg and Mishachev is false: for the standard symplectic $(\mathbb{R}^4, \omega = dx \wedge dy + dz \wedge dw)$, the inclusion

$$\mathrm{Symp}_c(\mathbb{R}^4, \omega) \to \mathrm{Diff}_c(\mathbb{R}^4)$$

is not a homotopy equivalence. This follows from M Gromov's [1985] result on the contractibility of $\mathrm{Symp}_c(\mathbb{R}^4, \omega)$ combined with Watanabe's recent disproof of the 4-dimensional Smale conjecture [Watanabe 2018].

1.4.2 Gompf's contact Dehn twist We will see (Section 3) that the contact Dehn twist is well defined on a (co-oriented) sphere $S \subset (Y, \xi)$ with a *tight neighborhood*. To the authors' knowledge, this contactomorphism was first considered by Gompf [1998] on the nontrivial sphere in the tight $S^1 \times S^2$. Gompf observed that τ_S and its iterates are not contact isotopic to the identity. Ding and Geiges [2010] later established that τ_S^2 generates all smoothly trivial contact mapping classes; see also [Min 2024]. Gironella [2021] has recently studied higher-dimensional analogues of Gompf's contactomorphism. However, all iterates of Gompf's τ_S and Gironella's generalizations happen to be *formally nontrivial* already, and hence *not* exotic.

1.4.3 Finite-order exotic contactomorphisms The previously known exotic 3-dimensional contactomorphisms have *finite* order, and the underlying 3-manifolds have $b_1 \geq 3$. These were detected on torus bundles by Geiges and Gonzalo [2004], who used an essentially elementary argument to reduce the problem to the Giroux–Kanda classification of tight contact structures on T^3. This was reproved using contact homology by Bourgeois [2006], who also found more exotic contactomorphisms in Legendrian circle bundles over surfaces of positive genus. In the latter case, those contactomorphisms have been shown to generate the group (1) by Geiges and Klukas [2014] and Giroux and Massot [2017]. Unlike the squared Dehn twists, these exotic contactomorphisms are all given by global symmetries. The paradigmatic example is the following:

Example 1.15 [Geiges and Gonzalo Perez 2004; Bourgeois 2006] Consider the 3-torus T^3 with the fillable contact structure $\xi_1 = \ker(\cos\theta\, dx - \sin\theta\, dy)$. By passing to n-fold covers $T^3 \to T^3$ given by $(\theta, x, y) \mapsto (n\theta, x, y)$, we obtain contact structures ξ_n on T^3. By a classical result of Giroux [1999] and Kanda [1997], the contact structures ξ_n (for $n \geq 1$) are pairwise not contactomorphic and give all the tight contact structures on T^3. When $n \geq 2$ the deck transformations of the n-fold cover $T^3 \to T^3$ generate all the exotic contactomorphisms of (T^3, ξ_n).

1.4.4 Other exotic Dehn twists Dehn twists have been a common source of exotic phenomena in topology:

(a) Let $Y_\# = Y_- \# Y_+$ be the sum of two aspherical 3-manifolds $Y_\pm$. By a result of McCullough [1990] (see also [Hatcher and Wahl 2010]) it follows that the kernel of $\pi_0 \operatorname{Diff}(Y_\#) \to \operatorname{Out}(\pi_1 Y_\#)$ is $\cong \mathbb{Z}_2$, generated by the smooth Dehn twist on the separating sphere.

(b) Seidel [1999] used Lagrangian Floer homology to detect exotic 4-dimensional symplectomorphisms with infinite order in the symplectic mapping class group, given by squared Dehn twists on Lagrangian spheres. He later generalized these results to higher dimensions [Seidel 2000; 2003]. See also the recent work of Smirnov [2020; 2022] using Seiberg–Witten gauge theory.

(c) Kronheimer and Mrowka [2020] have proved that the smooth Dehn twist on the separating sphere in the connected sum of two copies of the smooth 4-manifold underlying a $K3$ surface is not smoothly isotopic to the identity, even if it is topologically. For this they employ the Bauer–Furuta homotopical refinement of the Seiberg–Witten invariants of 4-manifolds. See also [Lin 2023].

1.5 Sketch of the proof of Theorem 1.3(A)

We outline here a proof of Theorem 1.3(A) which is simpler than the one we give later. In particular, the proof that we present now does not yield the stronger conclusion that the class of $\tau_{S_\#}^2$ is nontrivial in the abelianization of (1). We will need a stronger argument, which uses Theorem 1.8, in order to deduce both this and Theorem 1.6.

The main ideas go as follows. First, we have a *relative* version of the problem. Given a Darboux ball B in a contact 3-manifold (Y, ξ) we have a contactomorphism given by a Dehn twist $\tau_{\partial B}$ performed on an exterior sphere parallel to ∂B. This contactomorphism fixes the ball B and need not be contact isotopic to the identity *relative* to B, even if it always is globally (not fixing the ball). The problem of whether the squared Dehn twist $\tau_{\partial B}^2$ is isotopic rel B to the identity can be essentially recast as a lifting problem involving families of contact structures: if Y is aspherical (irreducible and with infinite fundamental group) then $\tau_{\partial B}^2$ is isotopic to the identity rel B precisely when the fibration given by the evaluation map $\mathrm{ev}\colon \mathcal{C}(Y,\xi) \to S^2$ admits a (homotopy) section (see Corollary 3.7). We recall that ev is defined by evaluating contact structures at a point. The key point that we exploit is that this fibration resembles a corresponding "evaluation map" pertaining to the Seiberg–Witten gauge theory of the manifold Y, and which is closely related to the U map in monopole Floer homology. As a result, an obstruction to the existence of a section was given by the second author in [Muñoz-Echániz 2024]: if $\boldsymbol{c}(\xi) \notin \operatorname{Im} U$ then no (homotopy) section exists, and thus $\tau_{\partial B}^2$ isn't isotopic to the identity rel B.

Going back to the original problem, consider two *tight* irreducible contact manifolds $(Y_\pm, \xi_\pm)$ and their sum $(Y_\#, \xi_\#)$. Let $\mathrm{CEmb}(S^2, (Y_\#, \xi_\#))_{S_\#}$ be the space of co-oriented *convex* embeddings $S^2 \hookrightarrow (Y_\#, \xi_\#)$ with standard characteristic foliation, in the isotopy class of the separating sphere $S_\#$. The group of contactomorphisms of $(Y_\#, \xi_\#)$ acts transitively on this space and yields a fibration[1]

$$\mathrm{Cont}(Y_\#, \xi_\#, S_\#) \to \mathrm{Cont}(Y_\#, \xi_\#) \to \mathrm{CEmb}(S^2, (Y_\#, \xi_\#))_{S_\#}, \quad f \mapsto f(S_\#). \tag{2}$$

From the long exact sequence of homotopy groups, a contactomorphism f of $(Y_\#, \xi_\#)$ fixing the sphere $S_\#$ is contact isotopic to the identity (not necessarily fixing $S_\#$) precisely when it arises as the monodromy in (2) of a loop of sphere embeddings. It thus becomes essential to understand the topology of the sphere embedding space. This brings us to the following h-principle-type result, which asserts that the topological complexity of this space only comes from reparametrizations of the source:

Theorem 1.16 *If $(Y_\pm, \xi_\pm)$ are irreducible and tight then the reparametrization map provides a homotopy equivalence* $U(1) \xrightarrow{\simeq} \mathrm{CEmb}(S^2, (Y_\#, \xi_\#))_{S_\#}$.

In the smooth case, the result analogous to the above was proved by Hatcher [1981]. The proof of Theorem 1.16 rests on the h-principle for standard convex spheres established by the first author together with Martínez-Aguinaga and Presas [Fernández et al. 2020], and should be regarded as an application of the h-principle of Eliashberg and Mishachev [2021].

With these ingredients in place, the proof of Theorem 1.3(A) goes as follows. The monodromy in (2) over the standard loop in $U(1)$ is given by the product of Dehn twists $\tau_{\partial B_-}\tau_{\partial B_+}$ (see Lemma 3.5). The contact Dehn twist $\tau_{S_\#}$ agrees with the image of $\tau_{\partial B_-}$ in $\pi_0 \mathrm{Cont}(Y_\#, \xi_\#)$. Because the manifolds $(Y_\pm, \xi_\pm)$ have infinite-order contact Dehn twists $\tau_{\partial B_\pm}$ rel $B_\pm$, for all $k \geq 1$ the class $\tau_{\partial B_-}^k \in \pi_0 \mathrm{Cont}(Y_\#, \xi_\#, S_\#)$ is not an iterate of $\tau_{\partial B_-}\tau_{\partial B_+}$ or its inverse. It follows that $\tau_{S_\#}$ and its iterates are not contact isotopic to the identity in $(Y_\#, \xi_\#)$.

[1] Strictly speaking, we should replace $\mathrm{Cont}(Y_\#, \xi_\#)$ with the subgroup consisting of contactomorphisms which preserve the isotopy class of the co-oriented sphere $S_\#$.

Outline The structure of the article is as follows. In Section 2 we introduce notation and present background material. In Section 3 we define the contact Dehn twist, establish various key properties and present examples where it is isotopic to the identity. In Section 4 we provide background on the families version of the Kronheimer–Mrowka contact invariant introduced in [Muñoz-Echániz 2024], which will be one of the main ingredients in the proofs of our main results. In Section 5 we review the h-principle for families of convex spheres in tight contact 3-manifolds established in [Fernández et al. 2020]. In Section 6 we use this h-principle to establish Theorem 1.8. We then complete the proofs of Theorems 1.3 and 1.6. In Section 7 we deduce exotic 1-parametric phenomena in overtwisted contact 3-manifolds.

Acknowledgements Fernández would like to acknowledge his advisor Francisco Presas for valuable conversations. Muñoz-Echániz thanks his advisor Francesco Lin for his support and encouragement, together with Hyunki Min for useful conversations. We would also like to thank the referee for the careful suggestions from which this manuscript has greatly benefited. Muñoz-Echániz was partially supported by NSF grant DMS-2203498.

2 Background

This section introduces the main players in this article: spaces of contact structures, contactomorphisms, embeddings, etc.

Remark 2.1 For convenience, throughout this article by a "fibration" we will mean a "Serre fibration". By a "homotopy equivalence" we will mean a "weak homotopy equivalence". However, the latter distinction isn't important: the various infinite-dimensional spaces that we consider are Fréchet manifolds, and hence they have the homotopy type of countable CW complexes [Palais 1966; Milnor 1959] and Whitehead's theorem applies.

2.1 Notation

Let (Y,ξ) be a closed contact 3-manifold. We always assume Y is connected and oriented, and ξ co-oriented and positive. Occasionally we will allow Y to be compact with nonempty boundary, in which case we assume that ∂Y is *convex* for the contact structure ξ and we fix a collar neighborhood $C=(-1,0]\times\partial Y$ of ∂Y. We quickly introduce here some of the spaces that will be relevant, all of which are equipped with the Whitney C^∞ topology:

- We denote by $\mathrm{Emb}(\mathbb{B}^3,Y)$ the space of orientation-preserving smooth embeddings $\phi\colon\mathbb{B}^3\hookrightarrow Y$ of the closed unit ball (avoiding the closure of C, if $\partial Y\neq\varnothing$). Let $\mathrm{Emb}((\mathbb{B}^3,\xi_{\mathrm{st}}),(Y,\xi))$ be the subspace consisting of contact embeddings of the standard contact unit ball. Such embeddings will be referred to as *Darboux balls* in (Y,ξ). Darboux's theorem asserts that for any interior point p of a contact manifold we may find such ϕ with $\phi(0)=p$. We will often abuse notation by referring to a Darboux ball only by its image $B:=\phi(\mathbb{B}^3)$.

• We denote by $\mathrm{Diff}(Y)$ the group of orientation-preserving diffeomorphisms, and by $\mathrm{Diff}(Y, B)$ the subgroup consisting of those which fix a Darboux ball B pointwise. By $\mathrm{Diff}_0(Y)$ and $\mathrm{Diff}_0(Y, B)$ we denote the subgroups consisting of those which are smoothly isotopic to the identity (rel B in the second case). We denote by $\mathrm{Cont}(Y) \subset \mathrm{Diff}$ the subgroup of co-orientation-preserving contactomorphisms of (Y, ξ), and by $\mathrm{Cont}(Y, B)$ the subgroup consisting of those which fix a Darboux ball B pointwise. By $\mathrm{Cont}_0(Y)$ and $\mathrm{Cont}_0(Y, B)$ we denote the subgroups consisting of those which are *smoothly* isotopic to the identity (rel B in the second case).

• We denote by $\mathcal{C}(Y, \xi)$ the space of contact structures on Y in the path-component of ξ. When $\partial Y \neq \varnothing$ we also require that they agree with ξ over C. Given a Darboux ball B in (Y, ξ) we denote by $\mathcal{C}(Y, \xi, B)$ the subspace consisting of contact structures ξ' for which the coordinate ball B is a Darboux ball for (Y, ξ') — ie $\xi = \xi'$ over B.

• We denote by $\mathrm{Fr}(Y)$ the principal $(\mathrm{SO}(3) \simeq)\, \mathrm{GL}^+(3)$-bundle over Y of oriented frames in TY, and by $\mathrm{CFr}(Y)$ the principal $(U(1) \simeq)\, \mathrm{CSp}^+(2, \mathbb{R})$-bundle over Y of co-oriented frames in ξ. Here $\mathrm{CSp}^+(2, \mathbb{R})$ denotes the linear conformal-symplectomorphism group. By the smooth and contact versions of the disk theorem[2] we have homotopy equivalences

$$\begin{aligned} \mathrm{Emb}(\mathbb{B}^3, Y) &\xrightarrow{\simeq} \mathrm{Fr}(Y), & \phi &\mapsto (d\phi)_0(e_1, e_2, e_3), \\ \mathrm{Emb}((\mathbb{B}^3, \xi_{\mathrm{st}}), (Y, \xi)) &\xrightarrow{\simeq} \mathrm{CFr}(Y, \xi), & \phi &\mapsto (d\phi)_0(e_1, e_2). \end{aligned} \tag{3}$$

Notice that $\mathrm{Fr}(Y) \simeq Y \times \mathrm{SO}(3)$ and, when the Euler class of ξ vanishes, $\mathrm{CFr}(Y, \xi) \simeq Y \times U(1)$.

• We denote by $\mathrm{Emb}(S^2, Y)$ the space of co-oriented embeddings of 2-spheres. By $\mathrm{CEmb}(S^2, (Y, \xi))$ we denote the subspace consisting of *convex* embeddings with *standard characteristic foliation* ("standard convex spheres" in short). Recall that a surface $\Sigma \subset (Y, \xi)$ is convex [Giroux 1991; Geiges 2008] if there exists a contact vector field on a neighborhood which is transverse to Σ. The standard characteristic foliation on S^2 is that induced from its embedding as the boundary of the Darboux ball.

• We denote by $\mathrm{Cont}(Y, \xi, S)$ the subgroup of contactomorphisms which fix a standard convex sphere S pointwise, and likewise for $\mathrm{Diff}(Y, S)$.

2.2 Standard fibrations

Next, we review how the spaces introduced above relate to each other through various natural fibrations. Some of the material from this section is treated in [Giroux and Massot 2017] in greater detail.

2.2.1 Diffeomorphisms acting on contact structures By an application of Gray's stability theorem (Moser's argument) [Geiges 2008] with parameters, one can show:

[2] The key point in the contact case is that $\varphi_t(x, y, z) := (tx, ty, t^2 z)$ is a contactomorphism of $(\mathbb{R}^3, \xi_{\mathrm{st}})$ for every $t > 0$, so the proof in the contact case follows along the same lines as in the smooth case; see [Geiges 2008, Theorem 2.6.7].

Lemma 2.2 *The action $f \mapsto f_*\xi$ of the group of diffeomorphisms on a fixed contact structure ξ gives a fibration*

$$\mathrm{Cont}_0(Y,\xi) \to \mathrm{Diff}_0(Y) \to \mathcal{C}(Y,\xi). \tag{4}$$

Similarly, there is fibration

$$\mathrm{Cont}_0(Y,\xi,B) \to \mathrm{Diff}_0(Y,B) \to \mathcal{C}(Y,\xi,B). \tag{5}$$

By (4), understanding the homotopy type of the space of contact structures $\mathcal{C}(Y,\xi)$ and the group of contactomorphisms $\mathrm{Cont}_0(Y,\xi)$ is essentially equivalent, since the homotopy type of $\mathrm{Diff}_0(Y)$ is often well understood (eg for all prime 3-manifolds by now).

2.2.2 Contactomorphisms acting on Darboux balls By an application of the contact isotopy extension theorem [Geiges 2008] with parameters, we have:

Lemma 2.3 *The action $f \mapsto f(B)$ of the group of contactomorphisms on a fixed Darboux ball $B \subset Y$ gives a fibration*

$$\mathrm{Cont}(Y,\xi,B) \to \mathrm{Cont}(Y,\xi) \to \mathrm{Emb}((\mathbb{B}^3,\xi_{\mathrm{st}}),(Y,\xi)). \tag{6}$$

Similarly, there is a fibration

$$\mathrm{Diff}(Y,B) \to \mathrm{Diff}(Y) \to \mathrm{Emb}(\mathbb{B}^3,Y). \tag{7}$$

2.2.3 Evaluation of contact structures at a point Fix a Darboux ball $B \subset Y$ with center $0 \in Y$. By regarding the 2-sphere S^2 as the space of co-oriented planes in the tangent space T_0B, we obtain the *evaluation map*

$$\mathrm{ev}_B : \mathcal{C}(Y,\xi) \to S^2, \quad \xi' \mapsto \xi'(0). \tag{8}$$

The following result is well known, but we provide a proof:

Lemma 2.4 *The evaluation map (8) is a fibration. The inclusion $\mathcal{C}(Y,\xi,B) \to (\mathrm{ev}_B)^{-1}(\xi(0))$ is a homotopy equivalence.*

Proof Let $\mathbb{B}^j$ be the unit j-disk and consider a homotopy $[0,1]\times\mathbb{B}^j \to S^2$, $(t,u)\mapsto\sigma_{t,u}$, together with a lift of the time zero map $\{0\}\times\mathbb{B}^j \to \mathcal{C}(Y,\xi)$ given by $u \mapsto \xi_u$, ie at the point $0 \in B$ we have $\xi_u(0) = \sigma_{0,u}$. We must find a family of contact structures $\xi_{t,u}$ with $\xi_{t,u}(0) = \sigma_{t,u}$ and $\xi_{0,u} = \xi_u$.

Let $v_{t,u} \in S(T_0B) = S^2$ be the unit normal (with respect to the standard flat metric on B) to the plane $\sigma_{t,u}$. Since the action of $\mathrm{SO}(3)$ on S^2 gives a fibration $\mathrm{SO}(3) \to S^2$ given by $A \mapsto Ae_3$, we may find $A_{t,u} \in \mathrm{SO}(3)$ such that $A_{t,u}e_3 = v_{t,u}$. Differentiating $A_{t,u}$ in t we get a vector field $V_{t,u}$ on $\mathbb{R}^3$. After cutting off $V_{t,u}$ outside the unit ball $B \subset Y$ we regard $V_{t,u}$ as a u-family of t-dependent vector fields on Y whose associated flows (starting at time $t=0$) we denote by ϕ_u^t. We obtain contact structures $\xi_{t,u} := (\phi_u^t)_*\xi_u$ with the desired property, which in fact agree with ξ outside $B \subset Y$.

For the second part, let $\xi_u = \ker \alpha_u$ be a family of contact structures parametrized by a disk $\mathbb{B}^j \ni u$ so that $\xi_u(0) = \xi(0)$ for all $u \in \mathbb{B}^j$ and $\xi_u(p) = \xi(p)$ for all $(u, p) \in \partial\mathbb{B}^j \times B$. We must deform relative to $\partial\mathbb{B}^j$ this family of contact structures to another family which agrees with ξ over the Darboux ball B. Denote by $i : \mathbb{B}^3 \hookrightarrow Y$ the inclusion of $B = i(\mathbb{B}^3)$. By the parametric version of Darboux's theorem we obtain a family of disk embeddings $\phi_u : \mathbb{B}^3 \hookrightarrow Y$, which are Darboux balls for ξ_u, such that $\phi_u(0) = 0 \in B$, $(d\phi_u)_0 = \mathrm{id}$ for all $u \in \mathbb{B}^j$, and $\phi_u = i$ for all $u \in \partial\mathbb{B}^j$. By (3) we may deform the family of embeddings ϕ_u to the inclusion i relative to $\partial\mathbb{B}^j$, and this deformation may be followed by an isotopy $f_{u,t} \in \mathrm{Diff}(Y)$ for $(u, t) \in \mathbb{B}^j \times [0, 1]$, with $f_{u,t} = \mathrm{id}$ for all $(u, t) \in \mathbb{B}^j \times \{0\} \cup \partial\mathbb{B}^j \times [0, 1]$. The homotopy of contact structures $(f_{u,t})_*\xi_u$ for $(u, t) \in \mathbb{B}^j \times [0, 1]$ solves the problem since the $(f_{u,1})_*\xi_u$ agree with ξ over B for all $u \in \mathbb{B}^j$. □

2.2.4 Contactomorphisms act on standard convex spheres Again, an application of the contact isotopy extension theorem gives:

Lemma 2.5 *The action $f \mapsto f(S)$ of the group of contactomorphisms on a fixed standard convex sphere $S \subset Y$ gives a fibration*

$$\mathrm{Cont}(Y, \xi, S) \to \mathrm{Cont}(Y, \xi) \to \mathrm{CEmb}(S^2, (Y, \xi)). \tag{9}$$

Similarly, there is a fibration

$$\mathrm{Diff}(Y, S) \to \mathrm{Diff}(Y) \to \mathrm{Emb}(S^2, Y). \tag{10}$$

Remark 2.6 The above statement isn't quite precise. For either (9) or (10), the downstairs projection is not surjective in general, so strictly speaking we only have a fibration over a union of connected components of the right-hand side. We will make no further comment on this point from now on.

2.3 Formal triviality and exoticness

Here we collect basic material that we need related to the notion of a formal contactomorphism. The material in this section should be well known to experts, but we did not find a convenient reference.

2.3.1 Formal contact structures and contactomorphisms For a 3-manifold Y, the flexible analogue[3] of a contact structure is a 2-*plane field*, ie a codimension-1 distribution $\xi \subset TY$. Henceforth all 2-planes in a 3-manifold are assumed to be co-oriented, as we've been assuming with contact structures. Let $\Xi(Y, \xi)$ denote the path-component of a fixed 2-plane field ξ in the space of all such. If ξ is a contact structure we have a natural inclusion map $\mathcal{C}(Y, \xi) \to \Xi(Y, \xi)$. The correct flexible analogue of a contactomorphism is:

Definition 2.7 A *formal contactomorphism* of (Y, ξ), where ξ is a 2-plane field, is a pair $(f, \{\phi^s\}_{0 \le s \le 1})$ such that $f \in \mathrm{Diff}(Y)$ and $\{\phi^s\}_{0 \le s \le 1}$ is a homotopy through vector bundle isomorphisms

$$\phi^s : TY \xrightarrow{\cong} f^*TY$$

such that $\phi^0 = df$ and ϕ^1 preserves the 2-plane field ξ.

[3]In general, if Y has dimension $2n + 1 \ge 3$, one should define $\Xi(Y, \xi)$ as the space of codimension-1 hyperplane fields in TY *equipped with* a $U(n)$ structure.

Of course, the above notion can be generalized to an arbitrary n-manifold equipped with a reduction of structure group to a subgroup $G \subset \mathrm{GL}(n,\mathbb{R})$. The group of formal contactomorphisms of (Y,ξ) is denoted by $\mathrm{FCont}(Y,\xi)$. When ξ is a contact structure there is the obvious inclusion map $\mathrm{Cont}(Y,\xi) \to \mathrm{FCont}(Y,\xi)$ given by $f \mapsto (f, df)$, where df denotes the constant homotopy at df.

A homotopy class in $\pi_j \mathrm{Cont}(Y,\xi)$ is said to be *formally trivial* if it lies in the kernel of $\pi_j \mathrm{Cont}(Y,\xi) \to \pi_j \mathrm{FCont}(Y,\xi)$. If, in addition, such a homotopy class is nontrivial in $\pi_j \mathrm{Cont}(Y,\xi)$ then we call it *exotic*. Similar terminology applies for families of contact structures.

2.3.2 A flexible analogue of (4) We introduce a flexible counterpart of the fibration (4). This is done via fibrant replacement of the map $\mathrm{Diff}_0(Y) \to \Xi(Y,\xi)$ given by $f \mapsto f^*\xi$. That is, we decompose this map as the composite of a homotopy equivalence $\mathrm{Diff}_0(Y) \xrightarrow{\simeq} \mathrm{FDiff}_0(Y)$ and a fibration $\mathrm{FDiff}_0(Y) \to \Xi(Y,\xi)$. Here $\mathrm{FDiff}(Y)$ is the topological group which consists of pairs $(f, \{\phi^t\}_{0\le t\le 1})$ where $f \in \mathrm{Diff}(Y)$ and $\{\phi^t\}_{0\le t\le 1}$ is a homotopy of vector bundle isomorphisms $\phi^t\colon TY \xrightarrow{\cong} f^*TY$ such that $\phi^0 = df$. By $\mathrm{FDiff}_0(Y)$ we denote the identity component. Clearly the inclusion induces a homotopy equivalence $\mathrm{Diff}(Y) \simeq \mathrm{FDiff}(Y)$. Define a mapping

$$\mathrm{FDiff}_0(Y) \to \Xi(Y,\xi), \quad (f, \{\phi^t\}) \mapsto \phi^1(\xi). \tag{11}$$

Lemma 2.8 *Let ξ be a 2-plane field on a compact oriented 3-manifold Y. Then the mapping (11) is a fibration with fiber $\mathrm{FCont}_0(Y,\xi)$. Thus, for a contact structure ξ, we have a commuting diagram of fibrations inducing a homotopy equivalence of total spaces*

$$\begin{tikzcd}
\mathrm{FCont}_0(Y,\xi) \arrow[r] & \mathrm{FDiff}_0(Y) \arrow[r] & \Xi(Y,\xi) \\
\mathrm{Cont}_0(Y,\xi) \arrow[r] \arrow[u] & \mathrm{Diff}_0(Y) \arrow[r] \arrow[u, "\simeq"] & \mathcal{C}(Y,\xi) \arrow[u]
\end{tikzcd}$$

We leave the proof of this lemma as an exercise.

Corollary 2.9 *Let (Y,ξ) be a contact 3-manifold. If $\beta \in \pi_j \mathcal{C}(Y,\xi)$ is formally trivial, then so is its image in $\pi_{j-1}\mathrm{Cont}_0(Y,\xi)$ under the connecting map of the fibration (4).*

The homotopy type of the space $\Xi(Y,\xi)$ is often easy to understand, unlike that of $\mathcal{C}(Y,\xi)$.

Example 2.10 Let Y be any integral homology 3-sphere, and ξ a 2-plane field on Y. Let ξ_{st} be any contact structure on S^3 (say, the tight one). By a result of Hansen [1978] there is a homotopy equivalence $\Xi(S^3,\xi_{\mathrm{st}}) \simeq \Xi(Y,\xi)$. From this one easily calculates

$$\pi_j \Xi(Y,\xi) \approx \pi_j S^2 \times \pi_{j+3} S^2.$$

3 Contact Dehn twists on spheres

In this section we define the contact Dehn twist on a sphere in several equivalent ways, establish some key properties and discuss some examples when its square is isotopic to the identity.

3.1 The contact Dehn twist

Let (Y, ξ) be a contact 3-manifold, and $S \subset Y$ be a co-oriented embedded sphere. Provided S has a tight neighborhood, we can associate to S a contactomorphism τ_S well defined in $\pi_0 \operatorname{Cont}(Y, \xi)$. We discuss this construction now.

3.1.1 Local model We start by discussing the local picture. Consider the contact 3-manifold $Y_0 = [-1,1] \times S^2$ with the tight contact structure $\xi_0 = \ker(\alpha_0)$ where $\alpha_0 = z\, ds + \frac{1}{2} x\, dy - \frac{1}{2} y\, dx$. Here s is the standard coordinate on $[-1,1]$, and x, y and z are coordinates on $\mathbb{R}^3$ restricted onto the unit sphere S^2. Consider the sphere $S_0 = \{0\} \times S^2 \subset Y_0$. We now describe the contact Dehn twist τ_{S_0} on the sphere S_0.

We choose a smooth function $\theta \colon [-1,1] \to [0, 2\pi]$ with $\theta(s) \equiv 0$ near $s = -1$ and $\theta(s) = 2\pi$ near $s = 1$. Let R_φ be the counterclockwise rotation in the xy plane with angle φ. Consider the diffeomorphism $\tilde\tau_{S_0}$ of Y_0 given by a smooth Dehn twist along S_0

$$\tilde\tau_{S_0}(s, x, y, z) = (s, R_{\theta(s)}(x, y), z).$$

Since $\pi_1 \mathrm{SO}(3) = \mathbb{Z}/2$ it follows that the squared Dehn twist $\tilde\tau_{S_0}^2$ is smoothly isotopic to the identity rel ∂Y_0. We don't quite have a contactomorphism of (Y_0, ξ_0), since

$$\tilde\tau_{S_0}^* \alpha_0 = \alpha_0 + \tfrac{1}{2} \theta'(s)(x^2 + y^2)\, ds.$$

However, consider the naive interpolation from α_0 to $\tilde\tau_{S_0}^* \alpha_0$

$$\alpha_t = \alpha_0 + t \tfrac{1}{2} \theta'(s)(x^2 + y^2)\, ds,$$

and observe that:

Lemma 3.1 *For any $t \in [0,1]$ the form α_t is a contact form.*

Proof A straightforward calculation shows $\alpha_t \wedge d\alpha_t = \alpha_0 \wedge d\alpha_0 > 0$. □

Thus, by Gray stability (Moser's argument) [Geiges 2008], the deformation of contact structures $\xi_t = \ker(\alpha_t)$ is realized by an isotopy f_t, ie $f_0 = \mathrm{id}$ and $(f_t)^* \xi_t = \xi_0$. Since the forms α_t don't depend on t near ∂Y_0 we may further assume that $f_t = \mathrm{id}$ near ∂Y_0. We then replace $\tilde\tau_{S_0}$ with $\tau_{S_0} := \tilde\tau_{S_0} \circ f_1$, and the latter is a contactomorphism of (Y_0, ξ_0). Also, the support of τ_{S_0} can be made arbitrarily close to the sphere S_0 by choosing $\theta(s)$ appropriately. Then for any $\epsilon \in (0,1]$ we have a well-defined isotopy class of contact Dehn twist

$$\tau_{S_0} \in \pi_0 \operatorname{Cont}([-\epsilon, \epsilon] \times S^2, \xi_0).$$

It is worth pointing out the following:

Lemma 3.2 *The group $\operatorname{Cont}(Y_0, \xi_0)$ is homotopy equivalent to $\Omega U(1) \simeq \mathbb{Z}$. Its π_0 is generated by the contact Dehn twist τ_{S_0}.*

Proof Gluing a Darboux ball B to (Y_0, ξ_0) gives back the standard contact ball $(\mathbb{B}^3, \xi_{\mathrm{st}})$. Thus, from the fibration (6), we have a map of fiber sequences

$$\begin{tikzcd} \mathrm{Cont}(Y_0,\xi_0) \arrow[r] & \mathrm{Cont}(\mathbb{B}^3,\xi_{\mathrm{st}}) \arrow[r] & \mathrm{Emb}((\mathbb{B}^3,\xi_{\mathrm{st}}),(\mathbb{B}^3,\xi_{\mathrm{st}})) \\ \Omega U(1) \arrow[u] \arrow[r] & \{*\} \arrow[u, "\simeq"] \arrow[r] & U(1) \arrow[u, "\simeq"] \end{tikzcd}$$

where the middle homotopy equivalence follows from Theorem 1.13 combined with Hatcher's theorem [1983]. The first assertion now follows. For the second assertion, we need to show that the generator $1 \in \pi_1 U(1)$ maps to the class of the contact Dehn twist τ_{S_0} under the connecting map.

We first describe the contact Dehn twist on S_0 more conveniently in terms of the coordinates on the ball $\mathbb{B}^3 = B \cup Y_0$. Recall that the standard contact structure on $\mathbb{B}^3$ is $\xi_{\mathrm{st}} = \ker \alpha_{\mathrm{st}}$ where $\alpha_{\mathrm{st}} = dz + \frac{1}{2}x\,dy - \frac{1}{2}y\,dx$. Choose a smooth function $\theta\colon [0,1] \to [0, 2\pi]$ with $\theta = 0$ near 0 and $\theta = 2\pi$ near 1. Let $r^2 := x^2 + y^2 + z^2$ be the radius squared function on $\mathbb{B}^3$. Then the diffeomorphism of $\mathbb{B}^3$ given by

$$\tilde{\tau}(x,y,z) := (R_{\theta(r^2)}(x,y), z)$$

does not quite preserve the contact structure, but

$$(\tilde{\tau})^*\alpha_{\mathrm{st}} = \alpha_{\mathrm{st}} + \tfrac{1}{2}(x^2+y^2)\theta'(r^2)\,d(r^2).$$

As in Lemma 3.1, the obvious interpolation that takes the second term in the above identity to zero gives a path of *contact* forms, and as in Section 3.1.2 we may canonically deform $\tilde{\tau}$ to a contactomorphism τ_{S_0} in the isotopy class of the contact Dehn twist on S_0.

Consider now a homotopy of maps $\theta_t\colon [0,1] \to [0,2\pi]$ with θ_t constant near 1 (with value 2π), such that $\theta_0 = \theta$ and θ_1 is the constant function with value 2π. We obtain an isotopy through diffeomorphisms of $\mathbb{B}^3$ (fixing a neighborhood of the boundary $\partial\mathbb{B}^3$, but not the smaller ball B!) given by

$$\tilde{\tau}_t(x,y,z) := (R_{\theta_t(r^2)}(x,y), z)$$

such that $\tilde{\tau}_0 = \tilde{\tau}$ and $\tilde{\tau}_1 = \mathrm{id}$. Again, by observing that for each t the obvious interpolation from $(\tilde{\tau}_t)^*\alpha_{\mathrm{st}}$ and α_{st} gives a path of contact forms, we may canonically deform the isotopy $\tilde{\tau}_t$ to a *contact* isotopy τ_t with $\tau_0 = \tau_{S_0}$ and $\tau_1 = \mathrm{id}$.

Now, the path of contactomorphisms τ_{1-t} from the identity to $\tau_{\partial B}$ induces a *loop* of Darboux balls $(\tau_{1-t})(B)$ in the class of the generator $1 \in \mathbb{Z} = \pi_1 \mathrm{Emb}((\mathbb{B}^3, \xi_{\mathrm{st}}), (\mathbb{B}^3, \xi_{\mathrm{st}}))$. □

Likewise, we have a firm hold on the topology of the space of standard spheres in our local model. Let $S_\pm = \{\pm\frac{1}{2}\} \times S^2 \subset Y_0$, and denote by $e_0\colon S^2 \hookrightarrow Y_0$ the embedding of $S_0 \subset Y_0$.

Lemma 3.3 *The map induced by reparametrization of e_0*

$$U(1) \to \mathrm{CEmb}(S^2, (Y_0, \xi_0)), \quad \theta \mapsto e_0 \circ r_\theta,$$

is a homotopy equivalence. Here $r_\theta(x,y,z)=(R_\theta(x,y),z)$*. Under the connecting homomorphism of the fibration* (9)*, the generator of* $\pi_1 U(1)=\mathbb{Z}$ *maps to the class*

$$(\tau_{S_-})^{-1}\tau_{S_+}\in\pi_0\,\mathrm{Cont}(Y_0,\xi_0,S_0).$$

Proof We have the following map of fiber sequences, with homotopy equivalences on the fiber and total space by Lemma 3.2:

$$\begin{tikzcd}
\mathrm{Cont}(Y_0,\xi_0,S_0) \arrow[r] & \mathrm{Cont}(Y_0,\xi_0) \arrow[r] & \mathrm{CEmb}(S^2,(Y_0,\xi_0)) \\
\Omega U(1)\times\Omega U(1) \arrow[r] \arrow[u, "\simeq"] & \Omega U(1) \arrow[r] \arrow[u, "\simeq"] & U(1) \arrow[u]
\end{tikzcd}$$

□

3.1.2 General case The robustness of our local picture allows us to consider contact Dehn twists in more general settings. We fix a 3-manifold (Y,ξ) together with a co-oriented *standard convex sphere* $S\subset Y$, ie an embedded sphere whose characteristic foliation agrees with that of $S_0\subset Y_0$ in the local model. It follows that neighborhoods of $S\subset Y$ and $S_0\subset Y_0$ are contactomorphic in a (homotopically) canonical fashion [Giroux 1991; Geiges 2008], and by making the support of τ_{S_0} sufficiently close to S_0 we may therefore implant τ_{S_0} into (Y,ξ) as a compactly supported contactomorphism τ_S, which we refer to as the *contact Dehn twist* on the co-oriented standard convex sphere $S\subset Y$. The class of τ_S in $\pi_0\,\mathrm{Cont}(Y,\xi)$ only depends on the isotopy class of S in the space of co-oriented standard convex spheres, defining a map of sets

$$\pi_0\,\mathrm{CEmb}(S^2,(Y,\xi))\to\pi_0\,\mathrm{Cont}(Y,\xi),\qquad S\mapsto\tau_S.$$

The contactomorphism τ_S makes sense more generally whenever $S\subset Y$ is a just a convex co-oriented sphere with a *tight neighborhood* U (but not necessarily having standard characteristic foliation). Indeed, by Giroux's criterion [2001] the dividing set of S is connected. Then by Giroux's realization theorem, we may find a smooth isotopy of sphere embeddings S_t whose image lies in the tight neighborhood U, $S_0=S$ and S_1 is a *standard* convex sphere, to which we associate the Dehn twist τ_{S_1} by the previous construction. A different choice of isotopy S_t' may yield a different standard convex sphere S_1'. The two spheres (S_1 and S_1') are isotopic within U as *standard* convex spheres by a result of Colin [1997, Proposition 10], so the contact Dehn twists τ_{S_1} and $\tau_{S_1'}$ are contact isotopic. Therefore we have a well-defined contact Dehn twist $\tau_S\in\pi_0\,\mathrm{Cont}(Y,\xi)$ associated to the convex sphere S with tight neighborhood U. In fact, since any *smooth* sphere can be made convex by a small isotopy [Giroux 1991], this construction defines a map

$$\pi_0\,\mathrm{Emb}_{\mathrm{tight}}(S^2,(Y,\xi))\to\pi_0\,\mathrm{Cont}(Y,\xi),\qquad S\mapsto\tau_S,$$

where $\mathrm{Emb}_{\mathrm{tight}}(S^2,(Y,\xi))$ stands for the space of *smooth* co-oriented embeddings $S^2\subset Y$ which admit a tight neighborhood. In particular, if (Y,ξ) is tight (globally) then τ_S only depends up to contact isotopy on the *smooth* isotopy class of the co-oriented sphere S.

The following particular case will play an essential role in this article, so we emphasize it now. Consider a Darboux ball $B=\phi(\mathbb{B}^3)$ in a contact manifold (Y,ξ). Associated to an exterior sphere (a sphere contained in the complement $Y \setminus B$) parallel to ∂B we have a well-defined contact Dehn twist which fixes B pointwise. By abuse of notation and for convenience we denote this contactomorphism by $\tau_{\partial B}$ even if the Dehn twist is not on the sphere ∂B. This defines a map of sets

$$\pi_0 \operatorname{Emb}((\mathbb{B}^3, \xi_{\mathrm{st}}), (Y, \xi)) \to \pi_0 \operatorname{Cont}(Y, \xi, B), \quad B \mapsto \tau_{\partial B}.$$

The following convenient description of $\tau_{\partial B}$ follows from the local calculation in the proof of Lemma 3.2.

Lemma 3.4 *The Dehn twist $\tau_{\partial B} \in \pi_0 \operatorname{Cont}(Y, \xi, B)$ agrees with the image of $1 \in \mathbb{Z}$ under the map*

$$\mathbb{Z} = \pi_1 U(1) \to \pi_1 \operatorname{Emb}((\mathbb{B}^3, \xi_{\mathrm{st}}), (Y, \xi)) \to \pi_0 \operatorname{Cont}(Y, \xi, B),$$

where the first map is induced by the reparametrization map

$$U(1) \to \operatorname{Emb}((\mathbb{B}^3, \xi_{\mathrm{st}}), (Y, \xi)), \quad \theta \mapsto \phi \circ r_\theta,$$

and the second map is the connecting map in the long exact sequence of the fibration (6).

Let $e \in \operatorname{CEmb}(S^2, (Y, \xi))$ be an embedding of a standard convex sphere. Thus the image $S = e(S^2) \subset (Y, \xi)$ is a co-oriented standard convex sphere. Let $S_\pm$ be two parallel copies of S given by pushing S forward and backward. By the local calculation in Lemma 3.3 we have:

Lemma 3.5 *The product of Dehn twists $(\tau_{S_-})^{-1} \tau_{S_+} \in \pi_0 \operatorname{Cont}(Y, \xi, S)$ agrees with the image of $1 \in \mathbb{Z}$ under the map*

$$\mathbb{Z} = \pi_1 U(1) \to \pi_1 \operatorname{CEmb}(S^2, (Y, \xi)) \to \pi_0 \operatorname{Cont}(Y, \xi, S),$$

where the first map is induced by the reparametrization map

$$U(1) \to \operatorname{CEmb}(S^2, (Y, \xi)), \quad \theta \mapsto e \circ r_\theta,$$

and the second map is the connecting map in the long exact sequence of the fibration (9).

3.2 The Dehn twist and the evaluation map

We move on to study a *relative* version of the isotopy problem for the Dehn twist. Consider the Dehn twist $\tau_{\partial B}$ on (an exterior sphere parallel to) the boundary ∂B of a Darboux ball, as in the previous section. We will now rephrase the problem of whether $\tau_{\partial B}^2$ defines the trivial class in $\pi_0 \operatorname{Cont}_0(Y, \xi, B)$ as a *lifting* problem.

3.2.1 The obstruction class The main player is the evaluation mapping $\operatorname{ev}_B : \mathcal{C}(Y, \xi) \to S^2$ defined by (8), which is a fibration (Lemma 2.4). If $\delta \colon \pi_2 S^2 \to \pi_1 \mathcal{C}(Y, \xi, B)$ is the connecting map in the homotopy long exact sequence, then we have a distinguished class

$$\mathcal{O}_\xi := \delta(1) \in \pi_1 \mathcal{C}(Y, \xi, B), \tag{12}$$

which, by construction, is the *obstruction class* to finding a homotopy section of ev_B (a map $s : S^2 \to \mathcal{C}(Y, \xi)$ such that $\operatorname{ev}_B \circ s \colon S^2 \to S^2$ has degree 1):

$$\operatorname{ev}_B \text{ admits a homotopy section if and only if } \mathcal{O}_\xi = 0.$$

Later in this section we will explicitly describe a loop of contact structures that represents the obstruction class $\mathcal{O}_\xi \in \pi_1\mathcal{C}(Y,\xi,B)$.

We now relate the problem of finding a section of ev_B to the triviality of the Dehn twist $\tau^2_{\partial B}$ as follows. Consider the connecting map $\delta'\colon \pi_1\mathcal{C}(Y,\xi,B) \to \pi_0\,\mathrm{Cont}_0(Y,\xi,B)$ of the fibration (5). The key observation is the following:

Proposition 3.6 *The class $\delta'(\mathcal{O}_\xi) \in \pi_0\,\mathrm{Cont}_0(Y,\xi,B)$ agrees with the **squared** contact Dehn twist $\tau^2_{\partial B}$.*

Proof Consider first the case when (Y,ξ) is the contact unit ball $\bigl(\mathbb{B}^3, \xi_{\mathrm{st}} = \ker\bigl(dz + \tfrac{1}{2}x\,dy - \tfrac{1}{2}y\,dx\bigr)\bigr)$ and $B \subset \mathbb{B}^3$ is a subball of smaller radius with center at 0. The fibrations from Section 2.2 fit into a commuting diagram

$$\begin{tikzcd}
\mathcal{C}(\mathbb{B}^3,\xi_{\mathrm{st}},B) \arrow[r] & \mathcal{C}(\mathbb{B}^3,\xi_{\mathrm{st}}) \arrow[r, "\mathrm{ev}_B"] & S^2 \\
\mathrm{Diff}_0(\mathbb{B}^3,B) \arrow[r] \arrow[u] & \mathrm{Diff}_0(\mathbb{B}^3) \arrow[r] \arrow[u] & \mathrm{Emb}(\mathbb{B}^3,\mathbb{B}^3) \simeq \mathrm{SO}(3) \arrow[u] \\
\mathrm{Cont}_0(\mathbb{B}^3,\xi_{\mathrm{st}},B) \arrow[r] \arrow[u] & \mathrm{Cont}_0(\mathbb{B}^3,\xi_{\mathrm{st}}) \arrow[r] \arrow[u] & \mathrm{Emb}((\mathbb{B}^3,\xi_{\mathrm{st}}),(\mathbb{B}^3,\xi_{\mathrm{st}})) \simeq U(1) \arrow[u]
\end{tikzcd}$$

In the third vertical fiber sequence the map $\pi_2 S^2 = \mathbb{Z} \to \pi_1 U(1) = \mathbb{Z}$ is multiplication by 2. From the diagram we see that the image of $\mathcal{O}_{\xi_{\mathrm{st}}} \in \pi_1\mathcal{C}(\mathbb{B}^3,\xi_{\mathrm{st}},B)$ in $\pi_0\,\mathrm{Cont}_0(\mathbb{B}^3,\xi_{\mathrm{st}},B)$ can be alternatively calculated as the image of $2 \in \mathbb{Z} = \pi_1 U(1)$ in $\pi_0\,\mathrm{Cont}_0(\mathbb{B}^3,\xi_{\mathrm{st}},B)$. From Lemma 3.4 this is the class of $\tau^2_{\partial B}$.

For an arbitrary (Y,ξ) and a Darboux ball, $B \subset Y$ the result then follows from the previous local calculation by extending the contact embedding $B \hookrightarrow Y$ to a contact embedding $B \subset \mathbb{B}^3 \hookrightarrow Y$, and considering the commuting diagram

$$\begin{tikzcd}
\pi_2 S^2 \arrow[r] \arrow[d] & \pi_1\mathcal{C}(\mathbb{B}^3,\xi_{\mathrm{st}},B) \arrow[r] \arrow[d] & \pi_0\,\mathrm{Cont}_0(\mathbb{B}^3,\xi_{\mathrm{st}},B) \arrow[d] \\
\pi_2 S^2 \arrow[r] & \pi_1\mathcal{C}(Y,\xi,B) \arrow[r] & \pi_0\,\mathrm{Cont}_0(Y,\xi,B)
\end{tikzcd}$$

□

Corollary 3.7 *Suppose Y is aspherical (irreducible and with infinite fundamental group). Then $\tau^2_{\partial B}$ is isotopic to the identity rel B if and only if the evaluation mapping (8) admits a homotopy section (the obstruction class $\mathcal{O}_\xi$ vanishes).*

Proof By the fibration (5) we have the exact sequence

$$\pi_1\,\mathrm{Diff}_0(Y,B) \to \pi_1\mathcal{C}(Y,\xi,B) \to \pi_0\,\mathrm{Cont}_0(Y,\xi,B),$$

so by Proposition 3.6 the result will follow from $\pi_1\,\mathrm{Diff}_0(Y,B) = 0$. Let us now explain why this group vanishes. By the fibration (7) we have an exact sequence

$$1 \to \pi_1\,\mathrm{Diff}_0(Y,B) \to \pi_1\,\mathrm{Diff}_0(Y) \to \pi_1\,\mathrm{Fr}(Y) \cong \pi_1 Y \times \mathbb{Z}_2.$$

Here, to have a 1 on the left we use $\pi_2 Y = 0$ (which follows from Y being aspherical). Since the homomorphism $\pi_1 \operatorname{Diff}_0(Y) \to \pi_1 Y$, and hence $\pi_1 \operatorname{Diff}_0(Y) \to \operatorname{Fr}(Y)$, is injective, it follows by exactness that $\pi_1 \operatorname{Diff}_0(Y, B) = 0$ as required. The fact that $\pi_1 \operatorname{Diff}_0(Y) \to \pi_1 Y$ is injective follows from the calculation of the homotopy type of the group $\operatorname{Diff}_0(Y)$ for all aspherical[4] 3-manifolds. More precisely, the papers [Hatcher 1976, 1981; 1983, Gabai 2001; Ivanov 1976; McCullough and Soma 2013] cover all aspherical 3-manifolds with the exception of the non-Haken infranilmanifold (see [McCullough and Soma 2013] for a nice summary). The latter consist of the nontrivial S^1-bundles over T^2, which are covered by [Bamler and Kleiner 2024]. In all these cases $\operatorname{Diff}_0(Y)$ has the homotopy type of $(S^1)^k$, where k is the rank of the center of $\pi_1 Y$ and $\pi_1 \operatorname{Diff}_0(Y) \to \pi_1 Y$ is the inclusion of the center. □

In the local model $(Y, \xi) = \big(\mathbb{B}^3, \xi_{\mathrm{st}} = \ker\big(dz + \tfrac{1}{2}x\,dy - \tfrac{1}{2}y\,dx\big)\big)$, and letting $B \subset \mathbb{B}^3$ be any concentric subball, we have the following unique characterization of the obstruction class:

Lemma 3.8 *The evaluation of contact structures on $\mathbb{B}^3$ along the radial line $\{(0,0,z) \mid z \in [0,1]\} \subset \mathbb{B}^3$ identifies the evaluation fibration on $\mathcal{C}(\mathbb{B}^3, \xi_{\mathrm{st}})$ with the path fibration on S^2:*

$$\begin{tikzcd}
\mathcal{C}(\mathbb{B}^3, \xi_{\mathrm{st}}, B) \arrow[r] \arrow[d, "\simeq"] & \mathcal{C}(\mathbb{B}^3, \xi_{\mathrm{st}}) \arrow[r, "\mathrm{ev}_B"] \arrow[d, "\simeq"] & S^2 \arrow[d, "="] \\
\Omega S^2 \arrow[r] & PS^2 \arrow[r] & S^2.
\end{tikzcd}$$

Thus the obstruction class $\mathcal{O}_{\xi_{\mathrm{st}}} \in \pi_1 \mathcal{C}(\mathbb{B}^3, \xi_{\mathrm{st}}, B)$ corresponds to the standard generator of $\pi_1 \Omega S^2 = \pi_2 S^2$.

Proof By the Eliashberg–Mishachev theorem [2021], the space $\mathcal{C}(\mathbb{B}^3, \xi_{\mathrm{st}})$ is contractible. So is the path space PS^2, so the desired result follows. □

3.2.2 Geometric description of the obstruction class It is instructive to describe an explicit loop $(\xi_\varphi)_{\varphi \in S^1}$ of contact structures on Y fixed over a Darboux ball $B \subset Y$ which represents the obstruction class $\mathcal{O}_\xi \in \pi_1 \mathcal{C}(Y, \xi, B)$; we do this now, but, won't use this construction in the remainder of the article.

By definition of the connecting map $\delta \colon \pi_2 S^2 \to \pi_1 \mathcal{C}(Y, \xi, B)$ associated to the fibration (8), a based loop $\xi_\varphi \in \mathcal{C}(Y, \xi, B)$ represents the obstruction class $\mathcal{O}_\xi$ precisely when there exists a $\mathbb{B}^2$-family of contact structures $\xi_{r,\varphi} \in \mathcal{C}(Y, \xi)$ — here, the unit 2-ball $\mathbb{B}^2$ is parametrized using polar coordinates (r, φ) — with $\xi_{r,0} = \xi$ such that $\xi_{1,\varphi} = \xi_\varphi$ and the mapping $\mathbb{B}^2 \ni (r, \varphi) \mapsto \mathrm{ev}_B(\xi_{r,\varphi}) = \xi_{r,\varphi}(p) \in S^2$ induces a *degree*-1 mapping $\mathbb{B}^2/\partial\mathbb{B}^2 \to S^2$ (note that the map out of $\mathbb{B}^2/\partial\mathbb{B}^2$ is well defined because $\xi_{1,\varphi}(p)$ is constant in φ).

Such a family $\xi_{r,\varphi}$ can be constructed as follows. First, it suffices to consider the case $(Y, \xi) = (\mathbb{B}^3, \xi_{\mathrm{st}})$, and construct a family $\xi_{r,\varphi} \in \mathcal{C}(\mathbb{B}^3, \xi, B)$ as above (and with $\xi_{r,\varphi} = \xi_{\mathrm{st}}$ near $\partial\mathbb{B}^3$). For this, choose any smooth mapping $q \colon [0,1] \times S^1 \to \mathrm{SO}(3)/U(1)$ with

$$q(0, \varphi) = q(1, \varphi) = [\mathrm{id}], \quad q(r, 0) = q(r, 2\pi) = [\mathrm{id}],$$

[4]For the irreducible 3-manifolds with finite fundamental group, the calculation of the homotopy type of $\operatorname{Diff}_0(Y)$ has also been completed [Hong et al. 2012; Bamler and Kleiner 2019, 2023]. Thus the homotopy type of $\operatorname{Diff}_0(Y)$ is known for all prime 3-manifolds.

such that the induced map $\Sigma S^1 \to \mathrm{SO}(3)/U(1)$ has degree 1. Here $\Sigma S^1 \approx S^2$ is the reduced suspension of S^1, and in what follows we regard S^1 as $\mathbb{R}/2\pi\mathbb{Z}$. Next, regarding q as a homotopy of based maps $S^1 \to \mathrm{SO}(3)/U(1)$, we may lift it along the fibration $\mathrm{SO}(3) \to \mathrm{SO}(3)/U(1)$ to produce a family of matrices $A_{r,\varphi} \in \mathrm{SO}(3)$ parametrized by $(r,\varphi) \in [0,1] \times S^1$ such that

$$[A_{r,\varphi}] = q(r,\varphi), \quad A_{0,\varphi} = \mathrm{id}, \quad A_{r,0} = A_{r,2\pi} = \mathrm{id}. \tag{13}$$

Because of the second condition in (13), $A_{r,\varphi}$ is, in fact, a family of matrices parametrized by the unit 2-ball $\mathbb{B}^2 \cong [0,1] \times S^1/0 \times S^1$. At this point, we would like to take the $\mathbb{B}^2$-family of contact structures on $\mathbb{B}^3$ given by $\xi_{r,\varphi} = (A_{r,\varphi})_*\xi_{\mathrm{st}}$. By the first condition in (13) $A_{1,\varphi} \in U(1)$, and from this it follows that $\xi_{1,\varphi}$ is a loop of contact structures on $\mathbb{B}^3$ fixed over the ball $B \subset \mathbb{B}^3$ (in fact this loop is constant everywhere on $\mathbb{B}^3$, $\xi_{1,\varphi} = \xi_{\mathrm{st}}$) and the induced map $\mathbb{B}^2/\partial\mathbb{B}^2 \to S^2$ given by $(r,\varphi) \mapsto \xi_{r,\varphi}(0)$ has degree 1, as required. However, the $\mathbb{B}^2$-family $\xi_{r,\varphi}$ is not constant near the boundary of $\mathbb{B}^3$, so we must appropriately "cut off" this family near the boundary.

We can do this as follows. Introduce a smooth cutoff function β on $\mathbb{B}^3$ which is identically 1 over $B \subset \mathbb{B}^3$ and vanishes near $\partial\mathbb{B}^3$. For each $(r,\varphi) \in [0,1] \times S^1$ we consider the following vector field supported in the interior of $\mathbb{B}^3$:

$$V_{r,\varphi}(x) = \beta(x)\frac{\partial}{\partial r}A_{r,\varphi} \cdot x.$$

We regard $V_{r,\varphi}$ as a φ-family of r-dependent vector fields on $\mathbb{B}^3$, and consider the associated φ-family of flows $\Phi^r_\varphi \colon \mathbb{B}^3 \to \mathbb{B}^3$ starting at time $r = 0$, namely

$$\frac{\partial}{\partial r}\Phi^r_\varphi(x) = V_{r,\varphi}(\Phi^r_\varphi(x)), \quad \Phi^0_\varphi(x) = x.$$

Over the ball $B \subset \mathbb{B}^3$ we have, by construction, that $\Phi^r_\varphi(x) = A_{r,\varphi} \cdot x$, and near the boundary of $\mathbb{B}^3$ we have $\Phi_{r,\varphi} = \mathrm{id}$. Hence the $\mathbb{B}^2$-family of contact structures defined by $\xi_{r,\varphi} := (\Phi^r_\varphi)_*\xi_{\mathrm{st}}$ is now constant near the boundary of $\mathbb{B}^3$, and still has the required properties.

The loop $\xi_{1,\varphi}$ of contact structures in $\mathcal{C}(\mathbb{B}^3, \xi_{\mathrm{st}}, B)$ thus constructed is an explicit representative of the obstruction class $\mathcal{O}_\xi \in \pi_1\mathcal{C}(\mathbb{B}^3, \xi_{\mathrm{st}}, B)$. This loop can then be implanted into an arbitrary contact 3-manifold (Y,ξ) along a Darboux chart $(\mathbb{B}^3, \xi_{\mathrm{st}}) \subset (Y,\xi)$ to give a representative of the obstruction class for arbitrary (Y,ξ).

3.3 Formal triviality of $\tau^2_{\partial B}$

We continue in the setting of the previous section, and we show:

Lemma 3.9 *Suppose the Euler class of ξ vanishes. Then both the loop of contact structures given by the obstruction class $\mathcal{O}_\xi \in \pi_1\mathcal{C}(Y,\xi,B)$ and the squared Dehn twist $\tau^2_{\partial B} \in \pi_0\,\mathrm{Cont}_0(Y,\xi,B)$ are formally trivial rel B.*

Proof On the space of co-oriented plane fields we have an analogous evaluation mapping (a fibration also, in fact)

$$\Xi(Y,\xi) \to S^2, \quad \xi' \mapsto \xi'(0).$$

When the Euler class of ξ vanishes we may identify $\Xi(Y,\xi)$ with the space $\mathrm{Map}_0(Y, S^2)$ of nullhomotopic smooth maps $Y \to S^2$. The evaluation mapping becomes identified with the obvious evaluation mapping on this latter space. Clearly this fibration admits a section given by the constant maps $Y \to S^2$. Hence the corresponding obstruction class vanishes, and hence

$$\mathcal{O}_\xi \in \ker(\pi_1 \mathcal{C}(Y,\xi,B) \to \pi_1 \Xi(Y,\xi,B)),$$

so $\mathcal{O}_\xi$ is formally trivial. From the rel B analogue of Corollary 2.9 it follows that $\tau^2_{\partial B}$ is formally trivial also. $\square$

3.4 Behavior of $\mathcal{O}_\xi$ under summation

We proceed by discussing how the obstruction class $\mathcal{O}_\xi$ from (12) interacts with the formation of connected sums.

First, we briefly review a convenient model for the contact connected sum [Colin 1997; Geiges 2008]. We write $(Y_0,\xi_0) = \big([-1,1]\times S^2, \ker\big(z\,ds + \frac{1}{2}x\,dy - \frac{1}{2}y\,dx\big)\big)$. Let $(Y_\pm,\xi_\pm)$ be two contact 3-manifolds with Darboux balls $B_\pm \subset Y_\pm$, and coordinates $\phi_\pm : (Y_0,\xi_0) \hookrightarrow (Y_\pm,\xi_\pm)$ around $\partial B_\pm$ such that $\phi_\pm(\{0\}\times S^2) = \partial B_\pm$ and $\phi_\pm(Y_0) \cap Y_\pm \backslash B_\pm = \phi_\pm((0,1]\times S^2)$. Consider the smaller ball $B^0_\pm = B_\pm \backslash \phi_\pm(Y_0) \subset B_\pm$. To define the contact connected sum we use the gluing contactomorphism G of (Y_0,ξ_0) given by $G(s,x,y,z) = (-s,-x,-y,-z)$.

Definition 3.10 The *connected sum* of contact manifolds

$$(Y_\#,\xi_\#) = (Y_-,\xi_-)\,\#\,(Y_+,\xi_+)$$

is defined to be $(Y_-\backslash B^0_-,\xi_-) \cup_G (Y_+\backslash B^0_+,\xi_+)$.

The connected sum of contact manifolds is well defined and independent of choices up to contactomorphism [Colin 1997].

We will fix a Darboux ball $B_\# \subset Y_\#$ inside the neck region $[-1,1]\times S^2 = \phi_-(Y_0) = \phi_+(Y_0) \subset Y_\#$. We also have natural inclusions $\mathcal{C}(Y_\pm,\xi_\pm,B_\pm) \subset \mathcal{C}(Y_\#,\xi_\#,B_\#)$. We consider their induced maps on π_1

$$(-)\,\#\,\xi_+ : \pi_1\mathcal{C}(Y_-,\xi_-,B_-) \to \pi_1\mathcal{C}(Y_\#,\xi_\#,B_\#),$$
$$\xi_-\,\#\,(-) : \pi_1\mathcal{C}(Y_+,\xi_+,B_+) \to \pi_1\mathcal{C}(Y_\#,\xi_\#,B_\#).$$

Proposition 3.11 *The obstruction class $\mathcal{O}_{\xi_\#} \in \pi_1\mathcal{C}(Y_\#,\xi_\#,B_\#)$ is given by*

$$\mathcal{O}_{\xi_\#} = (\mathcal{O}_{\xi_-}\,\#\,\xi_+)\cdot(\xi_-\,\#\,\mathcal{O}_{\xi_+}).$$

Proof It suffices to prove the corresponding statement in the local model where $(Y_\pm,\xi_\pm)=(\mathbb{B}^3,\xi_{\mathrm{st}})$, $B_\pm\subset Y_\pm$ are smaller concentric subballs and $(Y_\#,\xi_\#)=(Y_-\setminus B_-,\xi_-)\cup_{\partial B_-=-\partial B_+}(Y_+,\xi_+)$. We let $B_\#\subset Y_\#$ be a Darboux chart contained in the interior.

Consider the map $j:\mathcal{C}(Y_-,\xi_-,B_-)\times\mathcal{C}(Y_+,\xi_+,B_+)\to\mathcal{C}(Y_\#,\xi_\#,B_\#)$ given by gluing together the contact structures on the pieces $Y_\pm\setminus B_\pm$ to produce a contact structure on $Y_\#$. Choose paths $\gamma_\pm\subset Y_\#$ which connect the ball $B_\#$ to the component $\partial Y_\pm$ of the boundary of $Y_\#$. There is a commuting diagram

$$\begin{tikzcd}
\mathcal{C}(Y_\#,\xi_\#,B_\#) \arrow[r, "\mathrm{ev}_{\gamma_-}\times\mathrm{ev}_{\gamma_+}"] & (\Omega S^2)^2 \\
\mathcal{C}(Y_-,\xi_-,B_-)\times\mathcal{C}(Y_+,\xi_+,B_+) \arrow[u, "j"] \arrow[r, "\simeq"] & (\Omega S^2)^2 \arrow[u, "="']
\end{tikzcd}$$

where the bottom horizontal map is given by the homotopy equivalences of Lemma 3.8. By a similar argument to that in the proof of Lemma 3.8, one shows that the top horizontal map is a homotopy equivalence. Thus j is also a homotopy equivalence.

Next, consider the following diagram of spaces and maps, where the bottom row is two copies of the path fibration over S^2, and Δ is the diagonal map:

$$\begin{tikzcd}
\mathcal{C}(Y_\#,\xi_\#,B_\#) \arrow[r] \arrow[d, "\mathrm{ev}_{\gamma_-}\times\mathrm{ev}_{\gamma_+}"] & \mathcal{C}(Y_\#,\xi_\#) \arrow[r, "\mathrm{ev}_{B_\#}"] \arrow[d, "\mathrm{ev}_{\gamma_-}\times\mathrm{ev}_{\gamma_+}"] & S^2 \arrow[d, "\Delta"] \\
(\Omega S^2)^2 \arrow[r] & (PS^2)^2 \arrow[r] & (S^2)^2
\end{tikzcd}$$

The path fibration on S^2 can be identified with the evaluation fibration on $\mathcal{C}(Y_\pm,\xi_\pm)$ at the center of $B_\pm$, by Lemma 3.8. Under this identification, the leftmost vertical map in the second diagram becomes the homotopy inverse of the map j (which follows from the first diagram). Combining these observations and passing to the long exact sequence in homotopy groups in the second diagram yields a commutative square

$$\begin{tikzcd}
\pi_2 S^2 \arrow[r, "\delta"] \arrow[d] & \pi_1\mathcal{C}(Y_\#,\xi_\#,B_\#) \\
\pi_2 S^2\times\pi_2 S^2 \arrow[r, "\delta\times\delta"] & \pi_1\mathcal{C}(Y_-\setminus B_-,\xi_-)\times\pi_1\mathcal{C}(Y_+\setminus B_+,\xi_+) \arrow[u, "j"]
\end{tikzcd}$$

from which the desired result follows. □

Remark 3.12 In particular, it follows from Propositions 3.11 and 3.6 that we have the relation $\tau^2_{\partial B_+}\tau^2_{\partial B_-}=\tau^2_{B_\#}$ in $\pi_0\,\mathrm{Cont}_0(Y_\#,\xi_\#,\partial B_\#)$.

3.5 Examples: trivial Dehn twists

For comparison with Theorem 1.3 we now exhibit examples where the squared Dehn twist on a connected sum becomes trivial as a contactomorphism.

3.5.1 Quotients of S^3 Let Γ be a finite subgroup of $U(2)$. Then Γ preserves the standard contact structure $\xi_{\mathrm{st}} = \ker\bigl(\sum_{j=1,2} x_j\,dy_j - y_j\,dx_j\bigr)$ on the unit 3-sphere S^3, so it descends onto the quotient $M_\Gamma = S^3/\Gamma$. The M_Γ are the spherical 3-manifolds and include, among others, the lens spaces $L(p,q)$ and the Poincaré sphere $\Sigma(2,3,5)$.

Lemma 3.13 *The squared Dehn twist $\tau^2_{\partial B}$ on the boundary of a Darboux ball $B \subset M_\Gamma$ is contact isotopic to the identity rel B. Hence the squared Dehn twist $\tau^2_{S_\#}$ on the separating sphere $S_\#$ in $(Y,\xi)\,\#\,(M_\Gamma,\xi_{\mathrm{st}})$ is contact isotopic to the identity.*

Proof The center of $U(2)$ is given by the subgroup $\cong U(1)$ of diagonal matrices with diagonal (λ,λ) for some $\lambda \in U(1)$. This subgroup acts on M_Γ by contactomorphisms and thus also on the space of Darboux balls, which is homotopy equivalent to $M_\Gamma \times U(1)$ by (3). This gives a map $\pi_1 U(1) = \mathbb{Z} \to \pi_1(M_\Gamma \times U(1)) = \Gamma \times \mathbb{Z}$ which we assert is given by $1 \mapsto (e,2)$ where $e \in \Gamma$ is the identity element. From Lemma 3.4 and this assertion, the result would follow.

That the component $\mathbb{Z} \to \Gamma$ is trivial follows from $U(1)$ being the center of $U(2)$. To verify that $\mathbb{Z} \to \mathbb{Z}$ is multiplication by 2 we need to calculate the change in contact framing under the action of $U(1)$. We view S^3 as the unit sphere in the quaternions $\mathbb{H} = \mathbb{R}\langle 1,i,j,k\rangle$, so the tangent space at $q \in S^3$ is given by $T_q S^3 = \mathbb{R}\langle iq, jq, kq\rangle$ and the standard contact structure is $\xi_{\mathrm{st}}(q) = \mathbb{R}\langle jq, kq\rangle = \mathbb{C}\langle jq\rangle$. Thus the frame jq trivializes $\xi_{\mathrm{st}} \cong \mathbb{C}$ as a complex line bundle. The center subgroup $U(1) \subset U(2)$ acts on S^3 by $(\lambda,q) \mapsto \lambda q$, and the action of $U(1)$ on the frame jq is

$$\lambda \cdot jq = j\bar{\lambda}q = \lambda^2 \cdot j(\lambda q).$$

Thus the action on $\xi_{\mathrm{st}} \cong \mathbb{C}$ is by multiplication by λ^2 on the fibers. This establishes our assertion, and hence the proof is complete. □

Remark 3.14 When $\Gamma \subset \mathrm{SU}(2)$, an alternative proof of Lemma 3.13 can be obtained by instead exhibiting a section of $\mathrm{ev}_B : \mathcal{C}(M_\Gamma,\xi_{\mathrm{st}}) \to S^2$. The point is that the radial vector field $x\partial_x + y\partial_y + z\partial_z + w\partial_w$ is a Liouville vector field for each of the symplectic forms ω_u, for $u \in S^2$, in the flat hyperkähler structure of $\mathbb{R}^4$. The induced S^2-family of contact structures ξ_u on S^3 descends to the quotients M_Γ (with $\Gamma \subset \mathrm{SU}(2)$) and provides a section of ev_B.

3.5.2 $S^1 \times S^2$ Consider the unique tight contact structure on $S^1 \times S^2$, given by

$$\xi_0 = \ker\bigl(z\,d\theta + \tfrac{1}{2}x\,dy - \tfrac{1}{2}y\,dx\bigr).$$

Lemma 3.15 *The squared Dehn twist $\tau^2_{\partial B}$ on the boundary of a Darboux ball $B \subset S^1 \times S^2$ is contact isotopic to the identity rel B. Hence the squared Dehn twist $\tau^2_{S_\#}$ on the separating sphere $S_\#$ in any contact connected sum of the form $(Y,\xi)\,\#\,(S^1 \times S^2,\xi_0)$ is contact isotopic to the identity.*

Proof Let R_φ be the counterclockwise rotation in the xy plane of angle φ. By considering the subgroup $\{F_\varphi \mid \varphi \in S^1\} \simeq U(1) \subset \mathrm{Cont}(S^1 \times S^2, \xi_0)$, given by $F_\varphi(\theta,x,y,z) := (\theta, R_\varphi(x,y), z)$, one easily checks that $\pi_1(\mathrm{Cont}(S^1\times S^2,\xi_0)\to \pi_1\,\mathrm{Emb}((\mathbb{B}^3,\xi_{\mathrm{st}}),(S^1\times S^2,\xi_0))\to \pi_1 U(1)$ is surjective, so the result follows. □

Remark 3.16 In turn, the contact Dehn twist on the nontrivial sphere in $(S^1\times S^2,\xi_0)$ is nontrivial (and has infinite order). However, it is formally nontrivial already and therefore not exotic; see Section 3.6.

3.5.3 Sum with an overtwisted contact 3-manifold Let $(r,\theta,z)\in\mathbb{R}^3$ be cylindrical coordinates. Consider the contact structure ξ_{ot} in $\mathbb{R}^3$ defined by the kernel of

$$\alpha_{\mathrm{ot}}=\cos r\,dz+r\sin r\,d\theta.$$

The disk $\Delta_{\mathrm{ot}}=\{(r,\theta,z)\in\mathbb{R}^3\mid z=0,\ r\le\pi\}$ is an *overtwisted disk*.

Definition 3.17 [Eliashberg 1989] An overtwisted contact 3-manifold is a contact 3-manifold that contains an embedded overtwisted disk.

Let $\mathcal{C}(Y,\Delta_{\mathrm{ot}})$ be the space of contact structures in Y with a fixed overtwisted disk $\Delta_{\mathrm{ot}}\subset Y$. Let $\Xi(Y,\Delta_{\mathrm{ot}})$ be the space of co-oriented plane fields in Y tangent to Δ_{ot} at the point $0\in\Delta_{\mathrm{ot}}$. A foundational result of Eliashberg, generalized in higher dimensions by Borman, Eliashberg and Murphy, is:

Theorem 3.18 [Eliashberg 1989; Borman et al. 2015] *The inclusion*

$$\mathcal{C}(Y,\Delta_{\mathrm{ot}})\to\Xi(Y,\Delta_{\mathrm{ot}})$$

is a homotopy equivalence.

Remark 3.19 A relative version Eliashberg's h-principle is available. Suppose $A\subseteq Y\setminus\Delta_{\mathrm{ot}}$ is compact and $Y\setminus A$ is connected. Given a family of co-oriented plane fields $\xi^k\in\Xi(Y,\Delta_{\mathrm{ot}})$ that is contact over an open neighborhood of A, there exists a homotopy rel A from ξ^k to a family of contact structures.

Using Eliashberg's h-principle we obtain:

Lemma 3.20 *Let (Y,ξ) be a contact 3-manifold with vanishing Euler class. Then, for every overtwisted contact 3-manifold (M,ξ_{ot}), the squared contact Dehn twist $\tau^2_{S_\#}$ in $(Y,\xi)\,\#\,(M,\xi_{\mathrm{ot}})$ is contact isotopic to the identity.*

Proof Let $B\subset(Y,\xi)$ be a Darboux ball that we remove when performing the connected sum. By Lemma 3.9 $\tau^2_{\partial B}$ is formally contact isotopic to the identity rel B. It follows that $\tau^2_{S_\#}$ is formally contact isotopic to the identity on $Y\,\#\,M$, in fact relative to a small ball B_{ot} containing an overtwisted disk $\Delta_{\mathrm{ot}}\subset M$. At this point, by Eliashberg's Theorem 3.18 and Lemma 2.8 applied to the contact 3-manifold with convex boundary $(Y\,\#\,(M\setminus B_{\mathrm{ot}}),\xi\,\#\,\xi_{\mathrm{ot}})$, the group of contactomorphisms fixing Δ_{ot} is homotopy equivalent to the corresponding space of formal contactomorphisms. □

In Section 7 we will see that Lemma 3.20 implies exotic 1-parametric phenomena in overtwisted contact 3-manifolds.

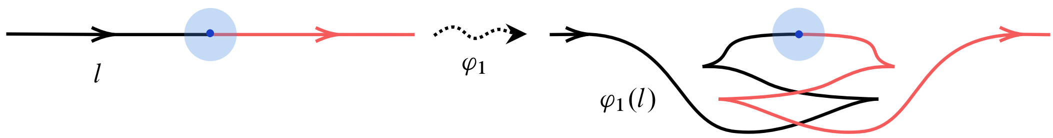


Figure 1: Front projection of l and $\hat{l}$. The shaded ball represents the small ball $B_\varepsilon \subset \mathbb{B}^3$.

3.6 The Reidemeister I move and Gompf's contactomorphism

We now describe the contact Dehn twist diagrammatically by means of front projections of Legendrian arcs. This approach is in the spirit of Gompf's description [1998] of the contact Dehn twist. For convenience we consider the unit ball $(\mathbb{B}^3, \xi = \ker(dz - y\,dx))$. Let $Y_0 = [-1,1] \times S^2$ be the complement in $\mathbb{B}^3$ of a small open ball B_ε around the origin. Consider the standard Legendrian arc $l\colon [-1,1] \to \mathbb{B}^3$ given by $t \mapsto (t,0,0)$. Perform two Reidemeister I moves to the Legendrian l to obtain a second Legendrian arc $\hat{l}$. We may assume that $\hat{l}$ coincides with l over the B_ε. The fronts of these arcs are depicted in Figure 1.

These arcs are Legendrian isotopic, so there exists a contact isotopy $\varphi_t \in \mathrm{Cont}(\mathbb{B}^3, \xi)$ with $\varphi_0 = \mathrm{id}$ and $\varphi_1 \circ l = \hat{l}$. Moreover, φ_1 can be taken to be the identity over B_ε. Therefore φ_1 gives a contactomorphism τ of the contact manifold with convex boundary (Y_0, ξ). From now on, we will denote the restrictions of l and $\hat{l}$ to the red segments in Figure 1 by the same letters for convenience. We have $\tau(l) = \hat{l}$ and the arc $\hat{l}$ is obtained in (Y_0, ξ) from l by a positive stabilization; see Figure 2. In particular,

$$\mathrm{rot}(\tau(l)) = \mathrm{rot}(l) + 1.$$

It follows that τ is not (formally) contact isotopic to the identity as a contactomorphism of (Y_0, ξ) rel ∂Y_0. This contactomorphism is contact isotopic to the contact Dehn twist as we have defined it in this section. In fact, as we will see in Lemma 3.22, since the complement of l is a tight 3-ball, any contactomorphism of (Y_0, ξ) can be described, up to contact isotopy, just in terms of its effect on l and, therefore, just by means of front projections of Legendrian arcs. First, we observe that the path-connected components of the space $\mathrm{Leg}(Y_0, \xi)$ of Legendrian embeddings of arcs that coincide with l at the endpoints can be easily understood:

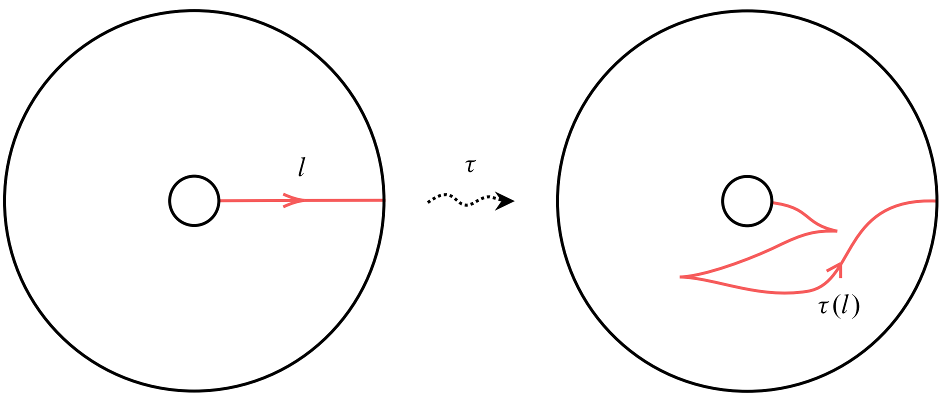


Figure 2: The image of l under τ.

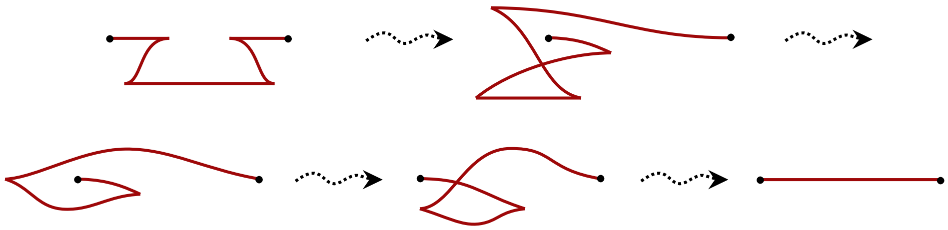


Figure 3: Legendrian isotopy from a double stabilization of l to l in (Y_0,ξ).

Lemma 3.21 *The map* $\mathrm{rot}\colon \pi_0\,\mathrm{Leg}(Y_0,\xi)\to\mathbb{Z}$ *given by* $L\mapsto\mathrm{rot}(L)$ *is an isomorphism.*

Proof Two smoothly isotopic Legendrian arcs with the same rotation number are Legendrian isotopic after adding a finite number of double stabilizations (pairs of positive and negative stabilizations) because of the Fuchs–Tabachnikov theorem [1997]. As depicted in Figure 3, this can done by a Legendrian isotopy in (Y_0,ξ). Therefore the proof follows from the 3-dimensional light bulb theorem. □

We conclude the following:

Lemma 3.22 *The map* $\mathrm{Cont}(Y_0,\xi)\to\mathrm{Leg}(Y_0,\xi)$ *given by* $f\mapsto f\circ l$ *is a homotopy equivalence. In particular,*

$$\pi_0\,\mathrm{Cont}(Y_0,\xi)\to\mathbb{Z},\qquad f\mapsto\mathrm{rot}(f\circ l),$$

is an isomorphism. Moreover, the contact Dehn twist is characterized, up to contact isotopy, by the relation

$$\mathrm{rot}(f(l))=\mathrm{rot}(l)+1.$$

Proof This follows by the previous lemma, the Eliashberg–Mishachev theorem (Theorem 1.13) and Hatcher's theorem [1983], since the fiber of $\mathrm{Cont}(Y_0,\xi)\to\mathrm{Leg}(Y_0,\xi)$ can be identified with the contactomorphism group of the complement of a neighborhood of l, and the latter is a tight 3-ball. □

4 Monopole Floer homology and families of contact structures

In this section we provide the necessary background on the Floer-theoretic ingredients that come into the proof of Theorem 1.3. Henceforth all homology groups are taken with $\mathbb{Q}$ coefficients for simplicity.

4.1 Monopole Floer homology and the contact invariant

For a quick introduction to Kronheimer and Mrowka's monopole Floer homology groups we recommend [Lin 2016; Kronheimer et al. 2007] and for a detailed treatment the monograph [Kronheimer and Mrowka 2007]. Here we just comment briefly on a few formal aspects.

Consider a 3-manifold Y together with spin-c structure $\mathfrak{s}$ (here the only spin-c structure that will be relevant is that induced by a contact structure ξ, denoted by $\mathfrak{s}_\xi$). Associated to it there are various monopole Floer homology groups ($\mathbb{Q}$-vector spaces in this article). The ones relevant to us are the "to" and "tilde" flavors: $\widecheck{\mathrm{HM}}(Y,\mathfrak{s})$ and $\widetilde{\mathrm{HM}}(Y,\mathfrak{s})$. The former arises "formally" as the S^1-equivariant Morse homology of the Chern–Simons–Dirac functional. An algebraic manifestation of this equivariant nature is that $\widecheck{\mathrm{HM}}(Y,\mathfrak{s})$ carries a module structure over the polynomial algebra $\mathbb{Q}[U]$ (the S^1-equivariant cohomology of a point, $H^\bullet_{S^1}(\text{point}) = \mathbb{Q}[U]$) and U decreases grading by two. In turn, the "tilde" flavor should be regarded as the (nonequivariant) Morse homology, and thus is an $H_\bullet(S^1) = \mathbb{Q}[\chi]/(\chi^2)$-module, with χ raising the degree by one. A standard *Gysin sequence* relates the two groups:

$$\cdots \xrightarrow{p} \widecheck{\mathrm{HM}}_\bullet(Y,\mathfrak{s}) \xrightarrow{U} \widecheck{\mathrm{HM}}_{\bullet-2}(-Y,\mathfrak{s}) \xrightarrow{j} \widetilde{\mathrm{HM}}_{\bullet-1}(-Y,\mathfrak{s}) \xrightarrow{p} \cdots .$$

The map χ is recovered from this by $\chi = jp$. A common feature of all flavors of the monopole groups is a canonical grading by the set of homotopy classes of plane fields $\pi_0\Xi(Y)$, which carries a natural $\mathbb{Z}$-action.

The *contact invariant* $\boldsymbol{c}(\xi)$ is an element of $\widecheck{\mathrm{HM}}_{[\xi]}(-Y,\mathfrak{s}_\xi)$ which is well defined up to a sign, and is canonically attached to a contact structure ξ on Y. It was defined by Kronheimer, Mrowka, Ozsváth and Szabó in [Kronheimer et al. 2007], but its definition goes back essentially to the earlier paper [Kronheimer and Mrowka 1997]. Ozsváth and Szabó [2005] gave a definition of $\boldsymbol{c}(\xi)$ in Heegaard–Floer homology. Under the isomorphism between the monopole and Heegaard–Floer groups [Kutluhan et al. 2020; Colin et al. 2011], the contact invariants are shown to agree. Some of the basic properties of $\boldsymbol{c}(\xi)$ are:

- [Mrowka and Rollin 2006] $\boldsymbol{c}(\xi) = 0$ if (Y,ξ) is overtwisted.
- [Echeverria 2020] $\boldsymbol{c}(\xi) \neq 0$ if (Y,ξ) admits a strong symplectic filling.
- [Echeverria 2020] $\boldsymbol{c}(\xi)$ is natural under symplectic cobordisms: if (W,ω) is a symplectic cobordism $(Y_1,\xi_1) \rightsquigarrow (Y_2,\xi_2)$ — here the convex end is (Y_2,ξ_2) — then
$$\widecheck{\mathrm{HM}}(-W,\mathfrak{s}_\omega)\boldsymbol{c}(\xi_2) = \boldsymbol{c}(\xi_1).$$
- $U \cdot \boldsymbol{c}(\xi) = 0$ (this is clear from the Heegaard–Floer point of view; in the monopole case this follows from Theorem 4.2).

4.2 The families contact invariant

Remark 4.1 Throughout this section we assume that $\boldsymbol{c}(\xi) \neq 0$ because it simplifies a little the exposition that follows (otherwise one should consider homologies with twisted coefficients; see [Muñoz-Echániz 2024]). We also resolve the sign ambiguity of $\boldsymbol{c}(\xi)$ by fixing one of the two. All homologies are taken with $\mathbb{Q}$ coefficients.

A version of the contact invariant for a family of contact structures was introduced by the second author in [Muñoz-Echániz 2024]. We summarize now some of those results. We have homomorphisms

$$\boldsymbol{Fc}_\bullet \colon H_\bullet(\mathcal{C}(Y,\xi)) \to \widecheck{\mathrm{HM}}_{[\xi]+\bullet}(-Y,\mathfrak{s}_\xi), \tag{14}$$

$$\widetilde{\boldsymbol{Fc}}_\bullet \colon H_\bullet(\mathcal{C}(Y,\xi,B)) \to \widetilde{\mathrm{HM}}_{[\xi]+\bullet}(-Y,\mathfrak{s}_\xi). \tag{15}$$

The invariant $\boldsymbol{Fc}_\bullet$ recovers the usual contact invariant: $H_0(\mathcal{C}(Y,\xi)) = \mathbb{Q}$ and then $\boldsymbol{Fc}_0(1) = \boldsymbol{c}(\xi)$. Their main property we exploit is the following. Associated to the fibration $\mathrm{ev}_B\colon \mathcal{C}(Y,\xi) \to S^2$ there is the Serre spectral sequence in homology. The latter collapses on the E^3 page and assembles into the *Wang* long exact sequence,

$$\cdots \to H_\bullet(\mathcal{C}(Y,\xi)) \xrightarrow{U_B} H_{\bullet-2}(\mathcal{C}(Y,\xi,B)) \to H_{\bullet-1}(\mathcal{C}(Y,\xi,B)) \xrightarrow{\iota_*} \cdots,$$

where U_B takes the intersection of cycles in the total space of the fibration with the fiber, and ι_* is the map induced by inclusion $\iota\colon \mathcal{C}(Y,\xi,B) \to \mathcal{C}(Y,\xi)$. We note that the obstruction class $\mathcal{O}_\xi$ for ev_B to admit a homotopy section arises here homologically as the image of 1 under $\mathbb{Q} = H_0(\mathcal{C}(Y,\xi,B)) \to H_1(\mathcal{C}(Y,\xi,B))$.

Theorem 4.2 [Muñoz-Echániz 2024] *There is a commutative diagram* (*up to signs*)

$$\begin{tikzcd}
\cdots \arrow[r, "p"] & \widetilde{\mathrm{HM}}_{[\xi]+\bullet}(-Y,\mathfrak{s}_\xi) \arrow[r, "U"] & \widetilde{\mathrm{HM}}_{[\xi]+\bullet-2}(-Y,\mathfrak{s}_\xi) \arrow[r, "j"] & \widetilde{\mathrm{HM}}_{[\xi]+\bullet-1}(-Y,\mathfrak{s}_\xi) \arrow[r, "p"] & \cdots \\
\cdots \arrow[r] & H_\bullet(\mathcal{C}(Y,\xi)) \arrow[r, "U_B"] \arrow[u, "\boldsymbol{Fc}_\bullet"] & H_{\bullet-2}(\mathcal{C}(Y,\xi,B)) \arrow[r] \arrow[u, "(\boldsymbol{Fc}_{\bullet-2})\circ\iota_*"] & H_{\bullet-1}(\mathcal{C}(Y,\xi,B)) \arrow[r, "\iota_*"] \arrow[u, "\widetilde{\boldsymbol{Fc}}_{\bullet-1}"] & \cdots
\end{tikzcd}$$

Some observations are in order:

- As a particular case, Theorem 4.2 recovers a property about the contact invariant $\boldsymbol{c}(\xi)$ which is well known from the Heegaard–Floer point of view: that $U\cdot \boldsymbol{c}(\xi) = 0$ and we have a canonical element $\tilde{\boldsymbol{c}}(\xi) := \widetilde{\boldsymbol{Fc}}_0(1) \in \widetilde{\mathrm{HM}}_{[\xi]}(-Y,\mathfrak{s}_\xi)$ such that $p\tilde{\boldsymbol{c}}(\xi) = \boldsymbol{c}(\xi)$. Conjecturally, the invariant $\boldsymbol{c}(\xi)$ corresponds to the Heegaard–Floer contact invariant that takes values in $\widehat{\mathrm{HF}}(-Y,\mathfrak{s}_\xi)$, which is defined in [Ozsváth and Szabó 2005].

- For 2-dimensional families, Theorem 4.2 gives us the simple formula

$$U\cdot \boldsymbol{Fc}_2(\beta) = \deg(\beta)\boldsymbol{c}(\xi),$$

where $\deg(\beta) = (\mathrm{ev}_B)_*\beta \in H_2(S^2) = \mathbb{Q}$ is the *degree* of the family $\beta \in H_2(\mathcal{C}(Y,\xi,B))$. In particular, by Theorem 4.2 we have the following:

Corollary 4.3 [Muñoz-Echániz 2024] *If $\boldsymbol{c}(\xi) \notin \mathrm{Im}\,U$ then the fibration ev_B does not admit a homotopy section and thus the obstruction class $\mathcal{O}_\xi$ is nonvanishing homologically.*

- Other statements that are easily derived from Theorem 1.3 are:

$$\boldsymbol{c}(\xi) \notin \mathrm{Im}\,U \text{ if and only if } \widetilde{\boldsymbol{Fc}}_1(\mathcal{O}_\xi) \neq 0, \quad \widetilde{\boldsymbol{Fc}}_1(\mathcal{O}_\xi) = \chi\tilde{\boldsymbol{c}}(\xi).$$

- If we define a $\mathbb{Q}[U]$-module structure on $H_\bullet(\mathcal{C}(Y,\xi))$ by setting $U := \iota_*\circ U_B$, then Theorem 4.2 asserts, in particular, that the homomorphism $\boldsymbol{Fc}_\bullet\colon H_\bullet(\mathcal{C}(Y,\xi,B)) \to \widetilde{\mathrm{HM}}_{[\xi]+\bullet}(-Y,\mathfrak{s}_\xi)$ is a map of $\mathbb{Q}[U]$-modules. Notice that we have, in fact, a $\mathbb{Q}[U]/(U^2)$-module structure on $H_2(\mathcal{C}(Y,\xi))$, ie the action of U^2 on $H_\bullet(\mathcal{C}(Y,\xi))$ vanishes. This can be regarded as a manifestation of the following geometric fact, that we have already encountered in Section 3. Consider two disjoint Darboux balls $B, B' \subset Y$. Whereas the spaces $\mathcal{C}(Y,\xi)$ and $\mathcal{C}(Y,\xi,B)$ are related in a possibly nontrivial way by the fibration ev_B, the spaces $\mathcal{C}(Y,\xi,B)$ and $\mathcal{C}(Y,\xi,B\cup B')$ are related in a straightforward way:

$$\mathcal{C}(Y,\xi,B\cup B') \simeq \Omega S^2 \times \mathcal{C}(Y,\xi,B).$$

Indeed, the evaluation map corresponding to the ball B' gives a fibration

$$\mathcal{C}(Y,\xi,B\cup B') \to \mathcal{C}(Y,\xi,B) \xrightarrow{\mathrm{ev}_{B'}} S^2,$$

but now the map $\mathrm{ev}_{B'}$ is nullhomotopic, as can be seen by dragging the evaluation point (the center of B') into the first ball B.

4.3 Summary of the construction of the families invariants

We summarize in this section the construction of the invariants $\boldsymbol{Fc}$ and $\widetilde{\boldsymbol{Fc}}$, carried out in detail by the second author in [Muñoz-Echániz 2024]. This is included here for background purposes, but will not be used.

4.3.1 The invariant $\boldsymbol{Fc}$ We begin with some general observations. Let X be a 4-manifold together with a nondegenerate 2-form ω, ie ω^2 is a volume form. We use ω^2 to orient X. Choose an almost-complex structure J compatible with ω, which by definition gives a metric $g=\omega(\cdot\,,J\cdot)$. The space of choices of J is contractible. The structure J equips X with a spin-c structure, ie a lift of the SO(4)-frame bundle of X along the map $\mathrm{Spin}^c(4)\to\mathrm{SO}(4)$. In differential-geometric terms this yields rank-2 complex hermitian bundles $S^\pm\to X$ and Clifford multiplication $\rho\colon TX\to\mathrm{Hom}(S^+,S^-)$ satisfying the "Clifford identity" $\rho(v)^*\rho(v)=g(v,v)\,\mathrm{id}$. We follow the notation and conventions from [Kronheimer and Mrowka 2007, Section 1] and we assume the reader is familiar with these.

The Clifford action of the 2-form ω on S^+ splits the bundle S^+ into $\mp 2i$ eigensubbundles of rank 1. These are given by $S^+=E\oplus EK_J^{-1}$, where K_J is the canonical bundle of (X,J) and E is a complex line bundle which is easily verified to be trivial. Choose a unit-length section Φ_0 of E. A simple calculation shows that there is a unique spin-c connection A_0 on S^+ such that $\nabla_{A_0}\Phi_0$ is a 1-form with values in the $+2i$ eigenspace EK_J^{-1}. At this point, the symplectic condition comes in through the following calculation involving the coupled Dirac operator $D_{A_0}\colon\Gamma(S^+)\to\Gamma(S^-)$:

Lemma 4.4 [Taubes 1994] *The nondegenerate 2-form ω is symplectic ($d\omega=0$) if and only if $D_{A_0}\Phi_0=0$.*

We now bring in a smoothly varying family of symplectic structures ω_u parametrized by a smooth manifold $U\ni u$, with each ω_u in the same deformation class as ω. Again, we equip the ω_u with compatible almost-complex structures J_u varying smoothly, which provide us with a family of metrics g_u. From our original Clifford bundle $(S^{\pm}, \rho)$ we canonically obtain new ones as follows. The bundles $S^{\pm}$ remain the same, but new Clifford structures ρ_u are obtained by setting $\rho_u = \rho \circ b_u$, where b_u is the canonical isometry $(TX, g_u) \xrightarrow{\cong} (TX, g)$ — the unique isometry which is positive and symmetric with respect to g_u. The Clifford action of ω_u again decomposes S^+ into eigenspaces $S^+ = E_u \oplus E_u K_{J_u}^{-1}$. Each E_u is trivializable individually, but the family $(E_u)_{u \in U}$ might give a nontrivial line bundle over $U \times X$. When U is contractible we may choose a family of trivializing sections Φ_u of E_u with unit length, and as before these determine unique spin-c connections A_u with $D_{A_u}\Phi_u = 0$. Then, associated to our family (ω_u, J_u) and the choices of Φ_u, we have a family of "deformed" Seiberg–Witten equations on X given by

$$\tfrac{1}{2}\rho_u(F_A^+) - (\Phi\Phi^*)_0 = \tfrac{1}{2}\rho_u(F_{A_u}^+) - (\Phi_u\Phi_u^*)_0, \quad D_A\Phi = D_{A_u}\Phi_u.$$

For each $u \in U$ this is an equation on the pair (A, Φ), where A is a connection on $\Lambda^2 S^+$ and Φ is a section of S^+. In this "deformed" version of the equations, the configurations (A_u, Φ_u) solve the equation for u.

We apply now the above considerations to a special case. Let (Y, ξ) be a closed contact 3-manifold with a contact form α, and let (X, ω) be the *symplectization* $X = [1, +\infty) \times Y$, with the exact symplectic form $\omega = d\left(\frac{1}{2}t^2\alpha\right)$. The structure J is chosen to be invariant under the Liouville flow, and the associated Riemannian metric on X is conical. We now bring into the picture a family of contact structures ξ_u parametrized by $U = \Delta^n$, to which we would like to associate an element in the Floer chain complex of $-Y = \partial X$. Here Δ^n is the standard n-simplex. We equip our family ξ_u with corresponding contact forms α_u. This gives a family ω_u of symplectic structures on X.

The construction now proceeds by forming a manifold Z^+ by gluing the cylinder $Z = (-\infty, 0] \times Y$ with the symplectic manifold X. We extend all metrics g_u over to Z^+ in such a way that they all agree with a fixed translation-invariant metric on the cylinder Z. Then the bundle S^+, together with its splitting $S^+ = E \oplus EK_J^{-1}$, extends over Z^+ naturally in a translation-invariant manner. The U-family of metrics and spin-c structures thus constructed on Z^+ are independent of u over Z, so we have effectively trivialized our data over the cylinder end $Z \subset Z^+$. In order to extend the Seiberg–Witten equations over Z^+ we cut off the perturbation term on the right-hand side of the equations so that it vanishes on the cylinder end Z. This way, we have a U-parametric family of Seiberg–Witten equations over Z^+, and natural boundary conditions for these equations (modulo gauge) are:

- On the cylinder Z, solutions should approach a translation-invariant solution $\mathfrak{a}$ (a generator of the "to" Floer complex $\check{C}(-Y, \mathfrak{s}_\xi)$, ie $\mathfrak{a}$ is an irreducible or boundary-stable monopole on $-Y$).
- On the symplectic end X solutions should approach the configuration (A_u, Φ_u).

This way we obtain parametrized moduli spaces of solutions

$$\pi : M([\mathfrak{a}], \Delta^n) \to \Delta^n.$$

By introducing suitable perturbations we may achieve the necessary transversality [Muñoz-Echániz 2024] and $M([\mathfrak{a}], \Delta^n)$ will be C^1-manifolds of finite dimension. At this point we note that, because of the gauge invariance of the equations, a different choice of trivializations Φ_u would yield diffeomorphic moduli spaces. The connected components of $M([\mathfrak{a}], \Delta^n)$ where the index of π is $-n$ consist of a finite number of isolated points lying over values in the interior of Δ^n, and a signed count of these points gives an integer $\#M([\mathfrak{a}], \Delta^n) \in \mathbb{Z}$. We organize these counts into a Floer chain $\psi(\Delta^n)$:

$$\psi(\Delta^n) = \sum_{[\mathfrak{a}]} \#M([\mathfrak{a}], \Delta^n) \cdot [\mathfrak{a}] \in \check{\mathrm{C}}(-Y, \mathfrak{s}_\xi).$$

The assignment $\Delta^n \mapsto \psi(\Delta^n)$ can be made into a chain map

$$\psi : C_\bullet(\mathcal{C}(Y, \xi)) \to \check{\mathrm{C}}_\bullet(-Y, \mathfrak{s}_\xi)$$

from the complex of singular chains on $\mathcal{C}(Y, \xi)$. Taking homology yields the families invariant (14). The analytic underpinnings that make all the above rigorous are discussed in [Muñoz-Echániz 2024], and are essentially no different than those of [Kronheimer and Mrowka 1997; Taubes 2000].

4.3.2 The invariant $\widetilde{\boldsymbol{Fc}}$ In terms of the "to" Floer complex $\widetilde{C}_\bullet$, the "tilde" Floer complex can be defined by taking the mapping cone of (a suitable chain level version of) the U map. We have $\widetilde{C}_\bullet(Y, \mathfrak{s}) = \check{C}_\bullet(Y, \mathfrak{s}) \oplus \check{C}_{\bullet-1}(Y, \mathfrak{s})$ with differential given by the matrix (ignoring signs)

$$\tilde{\partial} = \begin{pmatrix} \check{\partial} & 0 \\ U & \check{\partial} \end{pmatrix}.$$

If a family $\beta \in H_n(\mathcal{C}(Y, \xi))$ is in the image of $\iota_* \colon H_n(\mathcal{C}(Y, \xi, B)) \to H_n(\mathcal{C}(Y, \xi))$ then it is proved in [Muñoz-Echániz 2024] that $U \cdot \boldsymbol{Fc}(\beta) = 0$. At the chain level this is witnessed by a canonical chain homotopy θ:

$$U \cdot \psi \circ \iota_* = \check{\partial}\theta + \theta\partial. \tag{16}$$

From this we build the chain map

$$\tilde{\psi} = (\psi \circ \iota_*, \theta) \colon C_\bullet(C(Y, \xi, B)) \to \widetilde{C}_\bullet(-Y, \mathfrak{s}_\xi),$$

which, upon taking homology, gives the definition of (15). The chain homotopy θ is roughly constructed as follows. We introduce a new parameter $t \in \mathbb{R}$ and let $0 \in Y$ be the center of the ball B. Consider the moduli space

$$\mathcal{M}([\mathfrak{a}], \Delta^n) \to \mathbb{R} \times \Delta^n$$

consisting of quadruples (A, Φ, u, t) such that (A, Φ, u) solve the previous set of equations and boundary conditions subject to the further constraint that at the point $(t, 0) \in \mathbb{R} \times Y \cong Z^+$ the spinor Φ lies in the second component of the splitting $S^+ = E \oplus EK_J^{-1}$. By a simple modification of this construction one can again achieve transversality and ensure that the $\mathcal{M}([\mathfrak{a}], \Delta^n)$ are C^1-manifolds of finite dimension. Then we set

$$\theta(\Delta^n) = \sum_{[\mathfrak{a}]} \#\mathcal{M}([\mathfrak{a}], \Delta^n) \cdot [\mathfrak{a}].$$

Theorem 4.2 is established by carefully analyzing the "boundary at infinity" of the 1-dimensional components of the moduli $\mathcal{M}([\mathfrak{a}], \Delta^n)$; see [Muñoz-Echániz 2024].

5 The space of standard convex spheres in a tight contact 3-manifold

In this section we provide background on an h-principle for standard convex embeddings in tight contact 3-manifolds which was established in work of the first author with Martínez-Aguinaga and Presas [Fernández et al. 2020]. For the sake of completeness, we will provide here a detailed account which isn't quite the same as in [loc. cit.].

Throughout this section (Y, ξ) will be a tight contact 3-manifold. Recall that given a contact 3-manifold (Y, ξ), by a standard convex embedding of S^2 we mean a convex embedding $e\colon S^2 \hookrightarrow (Y, \xi)$ such that its oriented characteristic foliation $(e^*\xi) \cap TS^2$ coincides with the characteristic foliation of the sphere

$$e_0\colon S^2 \hookrightarrow \{0\} \times S^2 \subset (Y_0, \xi_0) = \bigl([-1, 1] \times S^2, \ker\bigl(z\, ds + \tfrac{1}{2} x\, dy - \tfrac{1}{2} y\, dx\bigr)\bigr).$$

In fact, by this property we obtain a (homotopically) unique contact embedding of a neighborhood of $e_0(S^2) \subset Y_0$ inside Y such that e_0 is identified with e. We recall that the north pole of e is then a positive elliptic point and the south pole a negative elliptic point. See Figure 4.

All of our arguments below work well for any other foliation of a convex sphere. The key fact is that the space of tight convex spheres with fixed characteristic foliation is C^0-dense inside the space of smooth spheres when the contact 3-manifold is tight, because of Giroux's genericity and realization theorems and Giroux's tightness criterion [1991; 2001].

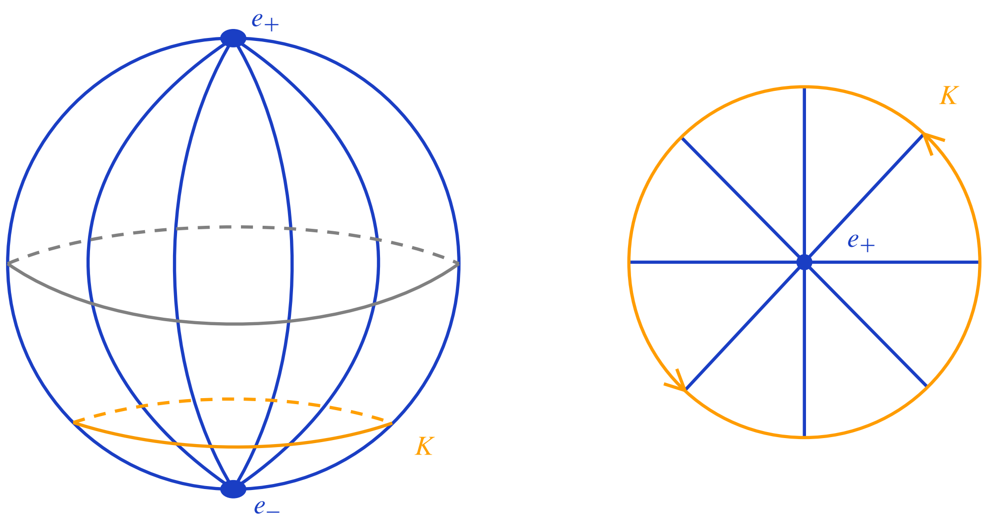


Figure 4: Schematic depiction of the standard sphere and the transverse curve K on the left. The mini-disk is depicted on the right.

Remark 5.1 The tightness condition is just required "locally", and therefore the results described in this section hold in overtwisted contact 3-manifolds if one replaces the space of smooth spheres by the space of smooth spheres with a tight neighborhood.

The main goal of this section is Theorem 5.8, which states that the space of standard embeddings of spheres into (Y,ξ) fixed near the south pole is homotopy equivalent to the corresponding space of smooth embeddings. In order to prove this result we will first study the closely related space of "mini-disks".

5.1 Mini-disks in a tight 3-manifold

Pick a small positively transverse curve $K \subseteq e(S^2)$ surrounding the negative elliptic point $e(s)$. The curve K divides the standard embedded sphere $e(S^2)$ in two disjoint disks $e(S^2)\backslash K = D^2_+ \cup D^2_-$. Here D^2_+ contains the positive elliptic point and D^2_- the negative one. In particular, we observe that the self-linking number of K is -1. The curve K is oriented as the boundary of D^2_+. Each disk $D^2_\pm$ is equipped with a natural parametrization induced by e. In particular, we will still denote by $e\colon \mathbb{D}^2 \to (Y,\xi)$ the parametrization of D^2_+. A smooth embedding of a disk with positive transverse boundary with self-linking number -1, which is convex and induces the same characteristic foliation as e, is called a *mini-disk*.[5]

We will denote by $\mathrm{CEmb}(\mathbb{D}^2,(Y,\xi))$ the space of embeddings of mini-disks which coincide with e over an open neighborhood of the boundary $\partial\mathbb{D}^2 \subset \mathbb{D}^2$. Define the space of smooth embeddings $\mathrm{Emb}(\mathbb{D}^2, Y)$ analogously. A consequence of Giroux's elimination theorem and the tightness of (Y,ξ) is the following result, which will be crucial to us:

Lemma 5.2 [Eliashberg 1993; Giroux 1991] *Let $f \in \mathrm{Emb}(\mathbb{D}^2, Y)$ be a smooth embedding. Then there exists a C^0-small isotopy of f, relative to an open neighborhood of the boundary, that makes f standard.*

This result is also explained in [Eliashberg 1993]; see also [Colin 1997]. Here the tightness condition is crucial.

We will prove the following h-principle:

Theorem 5.3 [Fernández et al. 2020] *The inclusion $\mathrm{CEmb}(\mathbb{D}^2,(Y,\xi)) \hookrightarrow \mathrm{Emb}(\mathbb{D}^2, Y)$ is a homotopy equivalence whenever (Y,ξ) is tight.*

Remark 5.4

- The π_0-surjectivity of the previous map follows from the previous lemma.
- The π_0-injectivity and also the π_1-surjectivity follow from Colin [1997], who proved this by applying his discretization trick. However, this does not quite work parametrically due to the fact that convexity is not generic among k-parametric families for $k > 0$.

[5] The terminology is due to Presas; see Figure 4. Observe that a mini-disk can be contracted inside small neighborhoods of the positive elliptic point.

Here we will use the approach of [Fernández et al. 2020] based on the notion of a *microfibration*, introduced by Gromov [1986]. We will apply the following "microfibration trick", which can also be applied to an arbitrary space of convex embeddings whenever this space is dense inside the space of smooth embeddings (Lemma 5.2) and we are able to establish a corresponding local version of the h-principle (ie in a neighborhood of a smooth embedding). These ingredients are the same as those required to effectively apply Colin's trick. Our advantage with respect to Colin is that our techniques work parametrically. However, we lose control of the geometric picture by using an algebraic construction.

Definition 5.5 A map $p\colon Y \to X$ of topological spaces is a *Serre microfibration* if for every homotopy $H\colon \mathbb{D}^k \times [0,1] \to X$ with a lift $h_0\colon \mathbb{D}^k \times \{0\} \to Y$ along p at time $t = 0$ there exists a positive real number $\varepsilon > 0$ together with an extension $h\colon \mathbb{D}^k \times [0,\varepsilon] \to Y$ of h_0 such that $p \circ h = H|_{\mathbb{D}^k \times [0,\varepsilon]}$.

A key property about microfibrations that we will use is:

Lemma 5.6 [Weiss 2005] *Every microfibration $p\colon Y \to X$ with* ***nonempty*** *and contractible fiber is a Serre fibration and*, *therefore*, *a weak homotopy equivalence.*

Proof of Theorem 5.3 Let K be a compact parameter space and $G \subset K$ a subspace. Consider $e^k \in \mathrm{Emb}(\mathbb{D}^2, Y)$ for $k \in K$ a family of smooth embeddings such that $e^k \in \mathrm{CEmb}(\mathbb{D}^2, (Y,\xi))$ for every $k \in G$. It is enough to establish the existence of a homotopy $e^k_t \in \mathrm{Emb}(\mathbb{D}^2, Y)$ such that

- $e^k_0 = e^k$,
- $e^k_t = e^k_0$ for $k \in G$,
- $e^k_1 \in \mathrm{CEmb}(\mathbb{D}^2, (Y,\xi))$.

Consider any extension of the embeddings e^k into a family of closed tubular neighborhood embeddings

$$E^k\colon \mathbb{D}^2 \times [-1,1] \hookrightarrow Y$$

such that

$$E^k|_{\mathbb{D}^2 \times \{0\}} = e^k.$$

Consider the space $\mathcal{B}$ of embeddings $E\colon \mathbb{D}^2 \times [-1,1] \hookrightarrow Y$ such that $E|_{\mathbb{D}^2 \times \{0\}}$ coincides with e over an open neighborhood of the boundary of $\mathbb{D}^2 = \mathbb{D}^2 \times \{0\}$. This space is the base of a microfibration

$$p\colon \mathcal{X} \to \mathcal{B}, \quad (E, p_t) \mapsto E,$$

where $\mathcal{X}$ is the space consisting of pairs (E, p_t) such that

- $E \in \mathcal{B}$,
- $p_t \in \mathrm{Emb}(\mathbb{D}^2, E(\mathbb{D}^2 \times [-1,1]))$ for $t \in [0,1]$ is a homotopy of proper embeddings of disks into the closed ball $E(\mathbb{D}^2 \times [-1,1])$, agreeing with the fixed embedding e near the boundary, and joining $p_0 = E|_{\mathbb{D}^2 \times \{0\}}$ with a mini-disk embedding $p_1 \in \mathrm{CEmb}(\mathbb{D}^2, (E(\mathbb{D}^2 \times [-1,1]), \xi))$.

The microfibration property is obviously satisfied. We will use Lemma 5.6 to conclude that p is, in fact, a fibration. Observe that the fiber $\mathcal{F}_E$ of p is nonempty because of Lemma 5.2. We claim that the fiber is also contractible. This is equivalent to the fact that the space of mini-disk embeddings, fixed near the boundary, into a tight 3-ball is homotopy equivalent to the space of smooth embeddings, fixed near the boundary, which is a combination of Eliashberg and Mishachev's theorem [2021] and Hatcher's theorem [1983]. Indeed, let $(\mathbb{B}^3,\xi)=(E(\mathbb{D}^2\times[-1,1]),\xi)$ and consider any mini-disk embedding $f\colon\mathbb{D}^2\to(\mathbb{B}^3,\xi)$ which coincides near the boundary with $E|_{\mathbb{D}^2\times\{0\}}$. The complement of $f(\mathbb{D}^2)$ in $\mathbb{B}^3$ is given by two tight balls $\mathbb{B}^3_\pm$. Denote by $\mathrm{CEmb}(\mathbb{D}^2,(\mathbb{B}^3,\xi))$ the corresponding space of mini-disk embeddings and by $\mathrm{Emb}(\mathbb{D}^2,\mathbb{B}^3)$ the smooth analogue. There is a map between fibrations

$$\begin{tikzcd}
\mathrm{Diff}(\mathbb{B}^3_+)\times\mathrm{Diff}(\mathbb{B}^3_-) \arrow[r] & \mathrm{Diff}(\mathbb{B}^3) \arrow[r] & \mathrm{Emb}(\mathbb{D}^2,\mathbb{B}^3) \\
\mathrm{Cont}(\mathbb{B}^3_+,\xi)\times\mathrm{Cont}(\mathbb{B}^3_-,\xi) \arrow[r] \arrow[u] & \mathrm{Cont}(\mathbb{B}^3,\xi) \arrow[r] \arrow[u] & \mathrm{CEmb}(\mathbb{D}^2,(\mathbb{B}^3,\xi)) \arrow[u]
\end{tikzcd}$$

inducing homotopy equivalences of total spaces and fibers, and thus the claim follows. Then from Lemma 5.6 we have that the map $p\colon\mathcal{X}\to\mathcal{B}$ is a Serre fibration with contractible fibers. This is enough to conclude the proof. Indeed, recall that we have built a map $j:K\to\mathcal{B}$ given by $k\mapsto E^k$, so $j^*\mathcal{X}\to K$ is a fibration with contractible nonempty fibers. We have a natural section over $G\subset K$ given by the constant homotopy:

$$s\colon G\to\mathcal{X},\quad k\mapsto(E^k,p^k_t=e^k).$$

Hence, by the contractibility of the fibers, we may extend this section over to K and obtain $\hat{s}\colon K\to\mathcal{X}$ given by $k\mapsto(E^k,p^k_t)$. The homotopy

$$e^k_t=p^k_t$$

solves our problem. □

We will need the following generalization. Let $\mathrm{CEmb}\big(\bigsqcup_j\mathbb{D}^2,(Y,\xi)\big)$ be the space of embeddings $e\colon\bigsqcup_j\mathbb{D}^2\hookrightarrow(Y,\xi)$ of n mini-disks, all of them fixed at an open neighborhood of $\bigsqcup_j\partial\mathbb{D}^2$. Denote also by $\mathrm{Emb}\big(\bigsqcup_j\mathbb{D}^2,Y\big)$ the corresponding space of smooth embeddings.

Theorem 5.7 *The natural inclusion* $\mathrm{CEmb}\big(\bigsqcup_j\mathbb{D}^2,(Y,\xi)\big)\hookrightarrow\mathrm{Emb}\big(\bigsqcup_j\mathbb{D}^2,Y\big)$ *is a homotopy equivalence whenever* (Y,ξ) *is tight.*

Proof The proof follows word by word the proof of Theorem 5.3. In this case the microfibration built is going to have as fiber the space of isotopies of n 2-disks into n disjoint tubular neighborhoods $\cong\mathbb{D}^2\times[-1,1]$. □

5.2 The space of standard spheres

As a consequence of our previous discussion we may compare the homotopy types of the space of standard spheres and the space of smooth spheres in a tight contact 3-manifold (Y,ξ). For this, consider the space of smooth embeddings $\mathrm{Emb}\big(\bigsqcup_j S^2,Y\big)$ of n-disjoint spheres and the corresponding subspace of standard spheres $\mathrm{CEmb}\big(\bigsqcup_j S^2,(Y,\xi)\big)$. Fix also an arbitrary standard embedding $e\colon\bigsqcup S^2\to(Y,\xi)$ and consider the subspaces $\mathrm{Emb}\big(\bigsqcup_j S^2, Y, \bigsqcup_j s_j\big)$ of embeddings that agree with e on an open neighborhood $\bigsqcup_j U_j$ of the south pole s_j of each sphere. Here we assume that the boundary $e|_{\bigsqcup_j \partial U_j}$ parametrizes n disjoint positively transverse knots K_j as in the previous section. Similarly, consider the analogous subspace of standard embeddings $\mathrm{CEmb}\big(\bigsqcup_j S^2, (Y,\xi), \bigsqcup_j s_j\big)$. Observe that the space $\mathrm{CEmb}\big(\bigsqcup_j S^2, (Y,\xi), \bigsqcup_j s_j\big)$ is homotopy equivalent to the space of n mini-disk embeddings into the tight contact manifold with convex boundary obtained from (Y,ξ) by removing an open neighborhood of $e\big(\bigsqcup_j U_j\big)$ whose boundary parametrizes K_j. The same observation applies to the space $\mathrm{Emb}\big(\bigsqcup_j S^2, Y, \bigsqcup_j s_j\big)$. We obtain:

Theorem 5.8 *Assume that (Y,ξ) is tight. Then*:

- *The inclusion* $\mathrm{CEmb}\big(\bigsqcup_j S^2, (Y,\xi), \bigsqcup_j s\big) \hookrightarrow \mathrm{Emb}\big(\bigsqcup_j S^2, Y, \bigsqcup_j s_j\big)$ *is a homotopy equivalence.*
- *For every $k \geq 1$ the natural homomorphism*
$$\pi_k(\mathrm{SO}(3)^n, U(1)^n) \to \pi_k\left(\mathrm{Emb}\left(\bigsqcup_j S^2, Y\right), \mathrm{CEmb}\left(\bigsqcup_j S^2, (Y,\xi)\right)\right)$$
induced by reparametrization on the source is an isomorphism.

Proof As explained above, the proof of the first assertion follows from Theorem 5.7. For the second assertion note that there is a natural map of fibrations given by the evaluation at the n south poles

$$\begin{tikzcd}
\mathrm{Emb}\big(\bigsqcup_j S^2, Y, \bigsqcup_j s_j\big) \arrow[r] & \mathrm{Emb}\big(\bigsqcup_j S^2, Y\big) \arrow[r] & \mathrm{Fr}_n(Y) \\
\mathrm{CEmb}\big(\bigsqcup_j S^2, (Y,\xi), \bigsqcup_j s_j\big) \arrow[r] \arrow[u] & \mathrm{CEmb}\big(\bigsqcup_j S^2, (Y,\xi)\big) \arrow[r] \arrow[u] & \mathrm{CFr}_n(Y,\xi) \arrow[u]
\end{tikzcd}$$

in which the vertical maps are inclusions. Here the base $\mathrm{Fr}_n(Y)$ is the space of framings over n different points of M, that is, the total space of a fiber bundle over the configuration space $\mathrm{Conf}_n(Y)$ with fiber $\approx \mathrm{GL}^+(3)^n$, and likewise for $\mathrm{CFr}_n(Y,\xi)$ but with contact frames. Observe that the map between the fibers is a homotopy equivalence because of the first assertion, so that the homomorphism induced by the evaluation map

$$\pi_k\left(\mathrm{Emb}\left(\bigsqcup_j S^2, Y\right), \mathrm{CEmb}\left(\bigsqcup_j S^2, (Y,\xi)\right)\right) \to \pi_k(\mathrm{Fr}_n(Y), \mathrm{CFr}_n(Y,\xi)) \cong \pi_k(\mathrm{SO}(3)^n, U(1)^n)$$

is an isomorphism and defines an inverse to the reparametrization map. □

5.3 Standard spheres in sums of two irreducible 3-manifolds

In this section we establish Theorem 1.16. We first discuss its smooth counterpart. The relevant reference on this topic is Hatcher's work [1981]. Let $Y_\# = Y_- \# Y_+$ with $Y_\pm$ now *irreducible*. Let $\mathrm{Emb}(S^2, Y_\#)_{S_\#} \subset \mathrm{Emb}(S^2, Y_\#)$ be the subspace of smooth co-oriented embeddings $S^2 \hookrightarrow Y_\#$ isotopic to a fixed given one $S_\#$, and let
$$\mathcal{S} = \mathrm{Emb}(S^2, Y_\#)_{S_\#} / \operatorname{Diff}(S^2)$$

be the space of *unparametrized* co-oriented nontrivial spheres. Hatcher [1981] proved that $\mathcal{S}$ is contractible. We also have a fibration

$$\mathrm{SO}(3) \simeq \mathrm{Diff}(S^2) \to \mathrm{Emb}(S^2, Y_\#)_{S_\#} \to \mathcal{S},$$

and hence

$$\mathrm{Emb}(S^2, Y_\#)_{S_\#} \simeq \mathrm{SO}(3).$$

Proof of Theorem 1.16 This is immediate from the long exact sequence of pairs associated to the horizontal maps in the commutative diagram

$$\begin{tikzcd}
\mathrm{CEmb}(S^2, (Y_\#, \xi_\#))_{S_\#} \arrow[r] & \mathrm{Emb}(S^2, Y_\#)_{S_\#} \\
U(1) \arrow[r] \arrow[u] & \mathrm{SO}(3) \arrow[u, "\simeq"]
\end{tikzcd}$$

combined with Theorem 5.8. □

6 Families of contact structures on sums of contact 3-manifolds

In this section we establish our main results, Theorems 1.3, 1.6 and 1.8, by combining the tools discussed in Sections 4 and 5.

6.1 The space of tight contact structures on a sum

Consider $n+1$ tight contact 3-manifolds (Y_j, ξ_j) for $j = 0, \ldots, n$ with $n \geq 1$. Let $(Y_\#, \xi_\#)$ be their connected sum, which we build as follows. We fix Darboux balls $B_{0-} \subset Y_0$, $B_{n+} \subset Y_n$ and for each $0 < j < n$ we fix two disjoint Darboux balls $B_{j\pm} \subset Y_j$. Then the connected sum $(Y_\#, \xi_\#)$ is formed by gluing in the order

$$(Y_0 \setminus B_{0-}) \bigcup_{\partial B_{0-} = -\partial B_{1+}} (Y_1 \setminus (B_{1+} \cup B_{1-})) \cdots \bigcup_{\partial B_{(n-1)-} = -\partial B_{n+}} (Y_n \setminus B_{n+}).$$

We will denote by $e_j \colon S^2 \hookrightarrow (Y_\#, \xi_\#)$, for $j = 1, \ldots, n$, the embedding of the j^{th} separating standard sphere given by $\partial B_{(j-1)-} = -\partial B_{j+}$ in the connected sum $(Y_\#, \xi_\#)$. Denote by s_j the south pole on the j^{th} sphere, regarded as a point in $e_j(S^2) \subset Y_\#$.

We will denote by $\mathrm{Tight}(Y, B)$ the space of tight contact structures on Y that are fixed on a Darboux ball B, and by $\mathrm{Tight}(Y, B, B')$ the subspace of $\mathrm{Tight}(Y, B)$ given by contact structures that are fixed on a second Darboux ball B' disjoint from B. A classical result of Colin [1997] asserts that the contact manifold $(Y_\#, \xi_\#)$ is tight, and we have a well-defined map

$$\#_{n+1} \colon \mathrm{Tight}(Y_0, B_{0-}) \times \prod_{j=1}^{n-1} \mathrm{Tight}(Y_j, B_{j+}, B_{j-}) \times \mathrm{Tight}(Y_n, B_{n+}) \to \mathrm{Tight}(Y_\#). \tag{17}$$

On the other hand, the evaluation map of each tight contact structure on Y at the south poles s_j defines a fibration

$$\mathrm{ev}_{n+1}\colon \mathrm{Tight}(Y_\#) \to (S^2)^n. \tag{18}$$

The fiber $\mathcal{F}$ of ev_{n+1} over $(\xi_\#(s_j))$ has the homotopy type of the space of tight contact structures on $Y_\#$ that agree with $\xi_\#$ over n disjoint Darboux balls $B_{\# j}$ around s_j. Therefore there is a natural inclusion

$$i_\#\colon \mathrm{Tight}(Y_0, B_{0-}) \times \prod_{j=1}^{n-1} \mathrm{Tight}(Y_j, B_{j+}, B_{j-}) \times \mathrm{Tight}(Y_n, B_{n+}) \hookrightarrow \mathcal{F}.$$

We establish the following stronger version of Theorem 1.8:

Theorem 6.1 *The inclusion $i_\#$ is a homotopy equivalence.*

Remark 6.2 Since S^2 is simply connected, we deduce from the long exact sequence in homotopy groups of (18) that

$$\pi_0(\mathrm{Tight}(Y_\#)) \cong \prod_{j=0}^{n} \pi_0(\mathrm{Tight}(Y_j)),$$

which is the classical result of Colin [1997].

Proof Let K be a compact parameter space and $G \subseteq K$ a subspace. It is enough to prove that if $\xi^k \in \mathcal{F}$ is a K-family of tight contact structures on $Y_\#$ that coincide with $\xi_\#$ over the n Darboux balls $B_{\# j}$ and such that $\xi^k \in \mathrm{Im}(i_\#)$ for $k \in G$, then there exists a homotopy of tight contact structures ξ^k_t for $t \in [0,1]$ such that

- $\xi^k_0 = \xi^k$,
- $\xi^k_t = \xi^k$ for $k \in G$, and
- $\xi^k_1 \in \mathrm{Im}(i_\#)$.

The key point is to observe that $\xi^k \in \mathrm{Im}(i_\#)$ if and only if the embeddings $e_j\colon S^2 \hookrightarrow (Y_\#, \xi^k)$ are standard for $j = 1, \dots, n$. For a given tight contact structure ξ, denote by

$$\mathcal{CE}_\xi := \mathrm{CEmb}\Bigl(\bigsqcup_{j=1}^{n} S^2, (Y_\#, \xi), \bigsqcup_{j=1}^{n} s_j\Bigr)$$

the space of standard embeddings of n disjoint spheres that coincide with (e_j) over a neighborhood of the south poles (s_j), and by

$$\mathcal{E} := \mathrm{Emb}\Bigl(\bigsqcup_{j=1}^{n} S^2, Y_\#, \bigsqcup_{j=1}^{n} s_j\Bigr)$$

the analogous space of smooth embeddings. Consider the space $\mathcal{X}$ of pairs (ξ, e_t) where $\xi \in \mathcal{F}$ and $e_t \in \mathcal{E}$, with $t \in [0,1]$, is a homotopy of embeddings with $e_0 = e$ and $e_1 \in \mathcal{CE}_\xi$. There is a natural forgetful map

$$p\colon \mathcal{X} \to \mathcal{F}, \quad (\xi, e_t) \mapsto \xi,$$

which is in fact a fibration because of Lemma 2.2. By Theorem 5.8 the inclusion $\mathcal{CE}_\xi \to \mathcal{E}$ is a homotopy equivalence. Therefore the fibers of the previous fibration are contractible.

This is enough to conclude the proof. Indeed, our initial family ξ^k is given by a map $j : K \to \mathcal{F}$ and the pullback fibration $j^*\mathcal{X} \to K$ has a well-defined section over $G \subseteq K$ given by the constant isotopy $e_t^k = e$ for $(k,t) \in G \times [0,1]$. Since the fiber of this fibration is contractible we can extend this section over K, obtaining a section e_t^k for $(k,t) \in K \times [0,1]$. Then we apply the smooth isotopy extension theorem to this family of embeddings to find an isotopy $\varphi_t^k \in \mathrm{Diff}(Y_\#)$ for $(k,t) \in K \times [0,1]$ such that

- $\varphi_0^k = \mathrm{id}$,
- φ_t^k is the identity over a neighborhood of the south poles (s_j),
- $\varphi_t^k \circ e = e_t^k$,
- $\varphi_t^k = \mathrm{id}$ for $(k,t) \in G \times [0,1]$.

The homotopy of contact structures $\xi_t^k = (\varphi_t^k)^* \xi^k$ solves the problem since now $e = (\varphi_1^k)^{-1} \circ e_1^k$ is standard for $(\varphi_t^k)^*\xi^k$ because e_1^k is standard for ξ^k. □

6.2 Diffeomorphisms of connected sums of two irreducible 3-manifolds

Consider $Y_\# = Y_- \# Y_+$ with $Y_\pm$ *irreducible*. Recall from Section 5.3 that Hatcher [1981] proved

$$\mathrm{Emb}(S^2, Y_\#)_{S_\#} \simeq \mathrm{SO}(3).$$

This has the following useful consequence:

Lemma 6.3 *Suppose that $Y_\pm$ are aspherical (irreducible and with infinite fundamental group). Then $\pi_1 \mathrm{Diff}(Y_\#) = 0$.*

Proof From the fibration (10) we have an exact sequence

$$\pi_1 \mathrm{Diff}(Y_-, B_-) \times \pi_1 \mathrm{Diff}(Y_+, B_+) \longrightarrow \pi_1 \mathrm{Diff}(Y_\#) \longrightarrow \mathbb{Z}_2 \longrightarrow \pi_0 \mathrm{Diff}(Y_-, B_-) \times \pi_0 \mathrm{Diff}(Y_+, B_+)$$

Under the connecting map, the nontrivial element in $\mathbb{Z}/2$ maps to

$$\tau_{\partial B_-}\tau_{\partial B_+} \in \pi_0 \mathrm{Diff}(Y_-, B_-) \times \pi_0 \mathrm{Diff}(Y_+, B_+).$$

We saw in the proof of Corollary 3.7 that the Dehn twists $\tau_{\partial B_\pm} \in \pi_0 \mathrm{Diff}(Y_\pm, B_\pm)$ are nontrivial and $\pi_1 \mathrm{Diff}(Y_\pm, B_\pm) = 0$. From this and the exact sequence above it now follows that $\pi_1 \mathrm{Diff}(Y_\#) = 0$. □

6.3 Proof of Theorem 1.3

As we've been doing so far, all homologies considered below are taken with $\mathbb{Q}$ coefficients, unless otherwise noted.

By Theorem 6.1 we have

$$\mathcal{C}(Y_\#, \xi_\#, B_\#) \simeq \mathcal{C}(Y_-, \xi_-, B_-) \times \mathcal{C}(Y_+, \xi_+, B_+),$$

and then by Proposition 3.11 the obstruction class $\mathcal{O}_{\xi_\#} \in \pi_1\mathcal{C}(Y_\#, \xi_\#, B_\#)$ to finding a homotopy section of $\mathrm{ev}_\#\colon \mathcal{C}(Y_\#, \xi_\#) \to S^2$ corresponds to

$$\mathcal{O}_{\xi_\#} \cong (\mathcal{O}_{\xi_-}, \mathcal{O}_{\xi_+}) \in \pi_1\mathcal{C}(Y_-, \xi_-, B_-) \times \pi_1\mathcal{C}(Y_+, \xi_+, B_+).$$

We recall that all homologies are taken with $\mathbb{Q}$ coefficients. A portion of the Wang long exact sequence for the fibration $\mathrm{ev}_{B_\#}$ is

$$\mathbb{Q} \xrightarrow{\delta} H_1(\mathcal{C}(Y_-, \xi_-, B_-)) \oplus H_1(\mathcal{C}(Y_+, \xi_+, B_+)) \to H_1(\mathcal{C}(Y_\#, \xi_\#)) \to 0,$$

where $\delta(1) = \mathcal{O}_{\xi_\#} = (\mathcal{O}_{\xi_-}, \mathcal{O}_{\xi_+})$. In the latter formula, and in what follows, we will incur a small abuse of notation by denoting the obstruction class and its image in homology by the same symbol.

The nontrivial input from Floer theory appears now. Because $\boldsymbol{c}(\xi_\pm) \notin \operatorname{Im} U$, by Corollary 4.3 to Theorem 4.2 the classes $\mathcal{O}_{\xi_\pm}$ are nontrivial in $H_1(\mathcal{C}(Y_\pm, \xi, B_\pm))$. It then follows from the Wang exact sequence that the class $(\mathcal{O}_{\xi_-}, 0)$ is not in the image of δ; thus the image of $(\mathcal{O}_{\xi_-}, 0)$ in $H_1(\mathcal{C}(Y_\#, \xi_\#))$ is nontrivial.

By Lemma 6.3, $\pi_1 \operatorname{Diff}(Y_\#) = 0$. Note that the hypotheses of that lemma indeed apply, because the manifolds $Y_\pm$ are aspherical. To see this, recall that an irreducible 3-manifold is aspherical precisely when it is not one of the quotients M_Γ of S^3 by a finite subgroup Γ of $\mathrm{SO}(4)$. The manifolds M_Γ have $\widecheck{\mathrm{HM}}(-M_\Gamma, \mathfrak{s}) = \mathbb{Q}[U, U^{-1}]$ in every spin-c structure $\mathfrak{s}$ because M_Γ is a rational homology sphere with a positive scalar curvature metric; see [Kronheimer and Mrowka 2007, Proposition 36.1.3]. In particular, the U map is surjective on $\widecheck{\mathrm{HM}}(-M_\Gamma, \mathfrak{s})$. The manifolds $Y_\pm$ are irreducible but can't be of the form M_Γ since $\boldsymbol{c}(\xi_\pm) \notin \operatorname{Im} U$.

Then, by the long exact sequence in homotopy groups of (4), it follows that

$$H_1(\mathcal{C}(Y_\#, \xi_\#); \mathbb{Z}) \cong \mathrm{Ab}(\pi_0 \operatorname{Cont}_0(Y_\#, \xi_\#)).$$

Under this isomorphism, the nontrivial class $(\mathcal{O}_{\xi_-}, 0)$ corresponds to the class of the squared Dehn twist $\tau_{S_\#}^2$ by Proposition 3.6. This proves that $\tau_{S_\#}^2$ is not contact isotopic to the identity. Since we've shown that $(\mathcal{O}_{\xi_-}, 0)$ is nontrivial *rationally*, it follows that all the even powers of $\tau_{S_\#}$ (and therefore all the powers) are also not contact isotopic to the identity. This completes the proof of Theorem 1.3(A).

By Lemma 3.9, the image of $\tau_{\partial B_\pm}$ in $\pi_0 \operatorname{FCont}_0(Y_\pm, \xi_\pm, B_\pm)$ is trivial. Hence so is the image of $\tau_{S_\#}^2$ in $\pi_0 \operatorname{FCont}_0(Y_\#, \xi_\#)$, proving Theorem 1.3(B). □

Remark 6.4 Working with $\mathbb{Z}$ coefficients rather than $\mathbb{Q}$, we can establish the following analogue of Theorem 1.3(A) by the same argument. If $2\boldsymbol{c}(\xi_\pm; \mathbb{Z}) \neq 0$ in $\widecheck{\mathrm{HM}}(-Y_\pm)$ and $0 \neq k \in \mathbb{Z}$ satisfies $k\boldsymbol{c}(\xi_\pm; \mathbb{Z}) \notin \operatorname{Im} U$, then the $2k$-fold iterate $\tau_{S_\#}^{2k}$ is not contact isotopic to the identity. All examples known to the authors where the latter hypothesis is satisfied for some $k \neq 0$ also satisfy the stronger $\mathbb{Q}$ version of the hypothesis. The assumption $2\boldsymbol{c}(\xi_\pm;\mathbb{Z}) \neq 0$ guarantees the orientability of the moduli spaces involved in the construction of the families contact invariant; see [Muñoz-Echániz 2024, Corollary 1.8].

6.4 Proof of Theorem 1.6

We write $(Y,\xi) = (Y_0,\xi_0)\,\#\cdots\#\,(Y_n,\xi_n)\,\#\,(Y_{n+1},\xi_{n+1})$, where $(Y_0,\xi_1),\dots,(Y_n,\xi_n)$ are those prime summands of (Y,ξ) such that $\boldsymbol{c}(\xi_j) \notin \operatorname{Im} U$ and the Euler class of ξ_j vanishes, and (Y_{n+1},ξ_{n+1}) is the sum of the remaining prime summands. We take the latter to be (S^3,ξ_{st}) if there are no prime summands remaining. We choose Darboux balls $B_{0-} \subset Y_0$ and $B_{n+1,+} \subset Y_{n+1}$, and for $j = 1,\dots,n$ we choose two Darboux balls $B_{j\pm} \subset Y_j$ disjoint from each other. We may take the connected sum (Y,ξ) to be built by gluing in the order

$$(Y_0 \setminus B_{0-}) \bigcup_{\partial B_{0-}=-\partial B_{1+}} (Y_1 \setminus (B_{1+} \cup B_{1-}))\cdots(Y_n \setminus (B_{n+} \cup B_{n-})) \bigcup_{\partial B_{n-}=-\partial B_{n+1,+}} (Y_{n+1} \setminus B_{n+1,+}),$$

with $n+1$ separating spheres. We fix $n+1$ Darboux balls $B_{\#j}$ for $j = 1,\dots,n+1$ centered at the south poles of the separating spheres ($B_{\#j}$ is centered at the south pole of the sphere which separates the pieces $Y_{j-1} \setminus B_{j-1,-}$ and $Y_{j+} \setminus B_{j+}$) and which are disjoint from each other.

Consider the evaluation map at the $n+1$ south poles of the spheres, which provides a fibration

$$\mathcal{F} \to \mathcal{C}(Y,\xi) \to (S^2)^{n+1}. \tag{19}$$

Theorem 6.1 identifies the fiber as

$$\mathcal{F} \simeq \mathcal{C}(Y_0, B_{0-}) \times \Biggl(\prod_{j=1,\dots,n+1} \mathcal{C}(Y_j, B_{j+} \cup B_{j-})\Biggr) \times \mathcal{C}(Y_{n+1}, B_{n+1,+}).$$

Recall that we have a homotopy equivalence

$$\mathcal{C}(Y_j, B_{j+} \cup B_{j-}) \simeq \Omega S^2 \times \mathcal{C}(Y_j, B_{j-}),$$

since the evaluation map $\mathrm{ev}_{B_{j+}} \colon \mathcal{C}(Y_j, B_{j-}) \to S^2$ is nullhomotopic (a nullhomotopy is obtained by dragging the evaluation point from B_{j+} into B_{j-}, and this yields the required homotopy equivalence). Thus

$$\mathcal{F} \simeq \mathcal{C}(Y_0, B_{0-}) \times \Omega S^2 \times \mathcal{C}(Y_1, B_{1-}) \times \cdots \times \Omega S^2 \times \mathcal{C}(Y_n, B_{n-}) \times \mathcal{C}(Y_{n+1}, B_{n+1,+}).$$

The connecting map in the long exact sequence in homotopy groups of the fibration (19) yields a homomorphism

$$\delta \colon \mathbb{Z}^{n+1} \to \pi_1\mathcal{C}(Y_0, B_{0-}) \times \mathbb{Z} \times \pi_1\mathcal{C}(Y_1, B_{1-}) \times \cdots \times \mathbb{Z} \times \pi_1\mathcal{C}(Y_n, B_{n-}) \times \pi_1\mathcal{C}(Y_{n+1}, B_{n+1,+}),$$

which we now calculate.

Lemma 6.5 *For $(a_1,\dots,a_{n+1}) \in \mathbb{Z}^{n+1}$ we have*

$$\delta(a_1,\dots,a_{n+1}) = (a_1 \cdot \mathcal{O}_{\xi_0}, a_1, a_2 \cdot \mathcal{O}_{\xi_1}, \cdots, a_n, a_{n+1} \cdot \mathcal{O}_{\xi_n}, a_{n+1} \cdot \mathcal{O}_{\xi_{n+1}}).$$

Proof Our argument is modeled on the proof of Proposition 3.11. It suffices to work in the local model where $(Y_j,\xi_j)=(\mathbb{B}^3,\xi_{\mathrm{st}})$ for all $j=0,\dots,n+1$. We choose $2n+2$ paths $\gamma_{0-},\gamma_{1\pm},\dots,\gamma_{n\pm},\gamma_{n+1,+}$ in Y, where each γ_{j+} goes from $B_{\#j}$ to $\partial Y_j\subset\partial Y$, and each γ_{j-} goes from $B_{\#j+1}$ to $\partial Y_j\subset\partial Y$. We consider the following commutative diagram of maps and spaces:

$$\begin{tikzcd}
\mathcal{C}(Y,\cup_{j=1}^{n+1}B_{\#j}) \arrow[r] \arrow[d, "\mathrm{ev}_\gamma"] & \mathcal{C}(Y) \arrow[r, "\mathrm{ev}_{B_\#}"] \arrow[d, "\mathrm{ev}_\gamma"] & (S^2)^{n+1} \arrow[d, "\Delta^{n+1}"] \\
(\Omega S^2)^{2n+2} \arrow[r] & (PS^2)^{2n+2} \arrow[r] & (S^2)^{2n+2}
\end{tikzcd}$$

Here $\mathrm{ev}_{\#B}$ stands for the evaluation map at the centers of the $n+1$ balls $B_{\#j}$, and the bottom row is given by $2n+2$ product of the path fibration on S^2. In particular, both rows are fibration sequences. The maps denoted by ev_γ stand for evaluation of contact structures along the $2n+2$ paths chosen above, and $\Delta\colon S^2\to(S^2)^2$ is the diagonal map.

Each of the two fibrations $\mathrm{ev}_{B_{0-}}\colon\mathcal{C}(Y_0)\to S^2$ and $\mathrm{ev}_{B_{n+1,+}}\colon\mathcal{C}(Y_{n+1})\to S^2$ is identified with the path fibration on S^2, by Lemma 3.8. Similarly, each of the n fibrations $\mathrm{ev}_{B_{j+}}\times\mathrm{ev}_{B_{j-}}\colon\mathcal{C}(Y_j)\to(S^2)^2$ with $j=1,\dots,n$ is identified with two copies of the path fibration on S^2. Using these, we identify the bottom row of the first diagram with the product of these $n+2$ fibrations.

The leftmost vertical map in the first diagram is a homotopy equivalence, which follows by an argument similar to the proof of Lemma 3.8. Consider the inclusion map

$$j\colon\mathcal{C}(Y_0,B_{0-})\times\Bigl(\prod_{j=1}^{n}\mathcal{C}(Y_j,B_{j+}\cup B_{j-})\Bigr)\times\mathcal{C}(Y_{n+1},B_{n+1,+})\to\mathcal{C}\Bigl(Y,\bigcup_{j=1}^{n+1}B_{\#j}\Bigr).$$

Under the identification of the bottom row of the first diagram with the product of the $n+2$ fibrations from the previous paragraph, the map j becomes the homotopy inverse of the leftmost vertical map in the first diagram, as in the proof of Proposition 3.11. The required result follows now from the commuting square obtained from taking homotopy groups in the first diagram:

$$\begin{tikzcd}
(\pi_2S^2)^{n+1} \arrow[r] \arrow[d, "\Delta^{n+1}"] & \pi_1\mathcal{C}(Y,\bigcup_{j=1}^{n+1}B_{\#j}) \\
(\pi_2S^2)^{2n+2} \arrow[r] & \pi_1\mathcal{C}(Y_0,B_{0-})\times\bigl(\prod_{j=1}^{n}\pi_1\mathcal{C}(Y_j,B_{j+}\cup B_{j-})\bigr)\times\pi_1\mathcal{C}(Y_{n+1},B_{n+1,+}) \arrow[u, "j"]
\end{tikzcd}$$ □

With this in place, we now look at the Serre spectral sequence of the fibration (19). From it we can assemble an exact sequence

$$\mathbb{Q}^{n+1}\xrightarrow{\delta}H_1(\mathcal{F})\to H_1(\mathcal{C}(Y,\xi))\to 0,$$

where δ is given by the same formula as in Lemma 6.5. By $\boldsymbol{c}(\xi_j)\notin\operatorname{Im}U$ and Corollary 4.3 to Theorem 4.2 we again deduce that the classes $\mathcal{O}_{\xi_j}$ for $j=0,\dots,n$ are homologically nontrivial (over $\mathbb{Q}$). Hence the n-dimensional subspace of $H_1(\mathcal{F})$ given by the elements

$$(b_1\cdot\mathcal{O}_{\xi_0},0,b_2\cdot\mathcal{O}_{\xi_1},0,\dots,0,b_n\cdot\mathcal{O}_{\xi_{n-1}},0,0,0)\quad\text{for }(b_j)\in\mathbb{Q}^n$$

injects as a subspace of $H_1(\mathcal{C}(Y,\xi))$. The proof of the formal triviality assertion is similar to the one given for Theorem 1.3. □

Remark 6.6 When Y is the sum of two aspherical 3-manifolds we have $\pi_1 \operatorname{Diff}(Y)=0$ (see Lemma 6.3). In the proof of Theorem 1.3 this allowed us to pass from a nontrivial element in $\pi_1\mathcal{C}(Y,\xi)$ to a nontrivial element in $\pi_0 \operatorname{Cont}_0(Y,\xi)$ via the fibration (4). This is a special situation. For instance, if Y is instead the sum of *at least three* aspherical 3-manifolds then it is known that $\pi_1 \operatorname{Diff}(Y)$ is not finitely generated [McCullough 1981]. A better control on $\pi_1 \operatorname{Diff}(Y)$ for general Y would allow us to understand whether the exotic loops of contact structures that we find in Theorem 1.6 yield nontrivial contactomorphisms.

7 Exotic phenomena in overtwisted contact 3-manifolds

In this final section we exhibit examples of 1-parametric exotic phenomena in *overtwisted* contact 3-manifolds.

On a heuristic level, Eliashberg's overtwisted h-principle [1989] is based on applying Gromov's h-principle for open manifolds to the complement of a 3-ball and using the overtwisted disk to fill in the ball. In the same spirit of this idea is what we call the "overtwisted escape principle", explained to us by Presas, which is a general strategy for proving an h-principle for a family of objects in a contact manifold (Y,ξ). First, perform the connected sum with an overtwisted manifold (M,ξ_{ot}), in order to apply the overtwisted h-principle [Eliashberg 1989; Borman et al. 2015] in the contact 3-manifold $(Y,\xi)\,\#\,(M,\xi_{\mathrm{ot}})$. This could be thought of as analogous to opening up the 3-manifold in the previous situation. Second, try to isotope the objects for which you want an h-principle so that they avoid ("escape") the overtwisted region $(M,\xi_{\mathrm{ot}})\setminus B$, where B is a Darboux ball. However, there could be obstructions to carrying out this second step. There are two scenarios: if these obstructions can be sorted out then our initial problem satisfies an h-principle; if not these obstructions should give rise to an exotic phenomenon in the overtwisted contact manifold $(Y,\xi)\,\#\,(M,\xi_{\mathrm{ot}})$. In [Casals et al. 2021] the authors successfully carry out this procedure to prove an existence h-principle for codimension-2 isocontact embeddings. Next, we will instead start out with a problem in (Y,ξ) which we know is geometrically obstructed a priori, and from this deduce an exotic overtwisted phenomenon.

Let $e\colon S^2\to(Y,\xi)$ be a standard embedding into a contact manifold (Y,ξ). A *formal standard embedding* of a sphere into (Y,ξ) is a pair (f,F^s) for $s\in[0,1]$ such that $f\in\operatorname{Emb}(S^2,Y)$ is a smooth embedding and $F^s\colon TS^2\to f^*TY$ is a homotopy of vector bundle injections with $F^0=df$ and $(F^1)^*\xi=e^*\xi\subset TS^2$. We will denote by $\operatorname{FCEmb}(S^2,(Y,\xi))$ the space of formal standard embeddings and by $\operatorname{FCEmb}(S^2,(Y,\xi),s)$ the subspace of formal standard embedding that coincide with e over an open neighborhood U of the south pole $s\in S^2$.

Let (M,ξ_{ot}) be an overtwisted contact 3-manifold. Consider the overtwisted contact 3-manifold $(Y_\#,\xi_\#)=(Y,\xi)\#(M,\xi_{\mathrm{ot}})$. We will consider the spaces $\operatorname{CEmb}(S^2,(Y_\#,\xi_\#),s)$ and $\operatorname{FCEmb}(S^2,(Y_\#,\xi_\#),s)$ as pointed spaces with basepoint given by the separating sphere $e\colon S^2 \hookrightarrow (Y_\#, \xi_\#)$. We have a natural inclusion $\mathrm{CEmb}(S^2, (Y_\#, \xi_\#), s) \hookrightarrow \mathrm{FCEmb}(S^2, (Y_\#, \xi_\#), s)$. From our previous discussion and the theory developed in this article we deduce the following:

Corollary 7.1 *Assume that (Y, ξ) is irreducible, ξ has vanishing Euler class and $\boldsymbol{c}(\xi) \notin \operatorname{Im} U$. Then there exists an element with infinite order in*

$$\ker(\pi_1 \mathrm{CEmb}(S^2, (Y_\#, \xi_\#), s) \to \pi_1 \mathrm{FCEmb}(S^2, (Y_\#, \xi_\#), s)).$$

Remark 7.2

- This should be compared with Theorem 5.8, which in particular asserts that this type of phenomenon does not happen when the underlying contact manifold is tight.
- Under the same assumptions, our proof also yields an element with infinite order in

$$\ker(\pi_1 \mathrm{CEmb}(S^2, (Y_\#, \xi_\#)) \to \pi_1 \mathrm{FCEmb}(S^2, (Y_\#, \xi_\#))).$$

Proof Denote by $S_\# = e(S^2)$ the standard separating sphere. Consider the squared Dehn twist $\tau^2_{S^+_\#}$ along a parallel copy $S^+_\#$ of $S_\#$, where we assume that $S^+_\#$ is contained in $(Y, \xi) \backslash B$, where B is the Darboux ball used to perform the connected sum. By the vanishing of the Euler class of ξ there exists a homotopy through formal contactomorphisms joining the identity with $\tau^2_{S^+_\#}$ (Lemma 3.9). It follows from Eliashberg's Theorem 3.18 combined with Lemma 2.8 that we can deform this homotopy (through formal contactomorphisms) to a homotopy φ_t through contactomorphisms with $\varphi_0 = \mathrm{id}$ and $\varphi_1 = \tau^2_{S^+_\#}$. This process can be done relative to an open neighborhood of the south pole $e(s) \in (Y \# M, \xi \# \xi_{\mathrm{ot}})$; see Remark 3.19. The loop of standard spheres $\varphi_t \circ e$ is formally trivial by construction but geometrically nontrivial. Indeed, by the contact isotopy extension theorem, the triviality of this loop would imply that $\tau^2_{S^+_\#}$, regarded as a contactomorphism of (Y, ξ), is contact isotopic to the identity rel B, which is in contradiction with Corollaries 3.7 and 4.3. □

Given a contact 3-manifold (Y, ξ) and a transverse knot $K \subset (Y, \xi)$, one can replace a small tubular neighborhood of K by a *Lutz twist* $(\mathrm{LT} = \mathbb{D}^2 \times S^1, \xi_{\mathrm{ot}})$ to obtain an *overtwisted* contact manifold (Y, ξ_K). Intuitively, the Lutz twist $(\mathrm{LT}, \xi_{\mathrm{ot}})$ is an *embedded* S^1-family of overtwisted disks; see [Geiges 2008] for precise definitions. We will denote by $\mathrm{LT}(Y, \xi_K)$ the space of contact embeddings $e\colon (\mathrm{LT}, \xi_{\mathrm{ot}}) \hookrightarrow (Y, \xi_K)$, regarded as a based space with basepoint the standard one, and by $\mathrm{FLT}(Y, \xi_K)$ the corresponding space of formal contact embeddings. As before, there is an inclusion map $\mathrm{LT}(Y, \xi_K) \to \mathrm{FLT}(Y, \xi_K)$. The following can be deduced using the same strategy as above:

Corollary 7.3 *Let (Y, ξ) be an irreducible contact 3-manifold with vanishing Euler class and such that $\boldsymbol{c}(\xi) \notin \operatorname{Im} U$. Consider a Darboux ball $B \subset (Y, \xi)$ and a transverse knot $K \subset B$. Then there exists an element with infinite order in*

$$\ker(\pi_1 \mathrm{LT}(Y, \xi_K) \to \pi_1 \mathrm{FLT}(Y, \xi_K)).$$

References

[Bamler and Kleiner 2019] **R H Bamler**, **B Kleiner**, *Ricci flow and contractibility of spaces of metrics*, preprint (2019) arXiv 1909.08710

[Bamler and Kleiner 2023] **R H Bamler**, **B Kleiner**, *Ricci flow and diffeomorphism groups of 3-manifolds*, J. Amer. Math. Soc. 36 (2023) 563–589 MR Zbl

[Bamler and Kleiner 2024] **R H Bamler**, **B Kleiner**, *Diffeomorphism groups of prime 3-manifolds*, J. Reine Angew. Math. 806 (2024) 23–35 MR Zbl

[Bodnár and Plamenevskaya 2021] **J Bodnár**, **O Plamenevskaya**, *Heegaard Floer invariants of contact structures on links of surface singularities*, Quantum Topol. 12 (2021) 411–437 MR Zbl

[Borman et al. 2015] **M S Borman**, **Y Eliashberg**, **E Murphy**, *Existence and classification of overtwisted contact structures in all dimensions*, Acta Math. 215 (2015) 281–361 MR Zbl

[Bourgeois 2006] **F Bourgeois**, *Contact homology and homotopy groups of the space of contact structures*, Math. Res. Lett. 13 (2006) 71–85 MR Zbl

[Casals et al. 2021] **R Casals**, **D M Pancholi**, **F Presas**, *The Legendrian Whitney trick*, Geom. Topol. 25 (2021) 3229–3256 MR Zbl

[Cerf 1968] **J Cerf**, *Sur les difféomorphismes de la sphère de dimension trois* ($\Gamma_4 = 0$), Lecture Notes in Math. 53, Springer (1968) MR Zbl

[Colin 1997] **V Colin**, *Chirurgies d'indice un et isotopies de sphères dans les variétés de contact tendues*, C. R. Acad. Sci. Paris Sér. I Math. 324 (1997) 659–663 MR Zbl

[Colin et al. 2011] **V Colin**, **P Ghiggini**, **K Honda**, *Equivalence of Heegaard Floer homology and embedded contact homology via open book decompositions*, Proc. Natl. Acad. Sci. USA 108 (2011) 8100–8105 MR Zbl

[Conway and Min 2020] **J Conway**, **H Min**, *Classification of tight contact structures on surgeries on the figure-eight knot*, Geom. Topol. 24 (2020) 1457–1517 MR Zbl

[Ding and Geiges 2007] **F Ding**, **H Geiges**, *A unique decomposition theorem for tight contact 3-manifolds*, Enseign. Math. 53 (2007) 333–345 MR Zbl

[Ding and Geiges 2010] **F Ding**, **H Geiges**, *The diffeotopy group of $S^1 \times S^2$ via contact topology*, Compos. Math. 146 (2010) 1096–1112 MR Zbl

[Echeverria 2020] **M Echeverria**, *Naturality of the contact invariant in monopole Floer homology under strong symplectic cobordisms*, Algebr. Geom. Topol. 20 (2020) 1795–1875 MR Zbl

[Eliashberg 1989] **Y Eliashberg**, *Classification of overtwisted contact structures on 3-manifolds*, Invent. Math. 98 (1989) 623–637 MR Zbl

[Eliashberg 1992] **Y Eliashberg**, *Contact 3-manifolds twenty years since J Martinet's work*, Ann. Inst. Fourier (Grenoble) 42 (1992) 165–192 MR Zbl

[Eliashberg 1993] **Y Eliashberg**, *Legendrian and transversal knots in tight contact 3-manifolds*, from "Topological methods in modern mathematics" (L R Goldberg, A V Phillips, editors), Publish or Perish, Houston, TX (1993) 171–193 MR Zbl

[Eliashberg and Mishachev 2002] **Y Eliashberg**, **N Mishachev**, *Introduction to the h-principle*, Graduate Studies in Math. 48, Amer. Math. Soc., Providence, RI (2002) MR Zbl

[Eliashberg and Mishachev 2021] **Y Eliashberg**, **N Mishachev**, *The space of tight contact structures on $\mathbb{R}^3$ is contractible*, preprint (2021) arXiv 2108.09452

[Fernández et al. 2020] **E Fernández**, **J Martínez-Aguinaga**, **F Presas**, *The homotopy type of the contactomorphism groups of tight contact* 3*-manifolds*, *I*, preprint (2020) arXiv 2012.14948

[Fuchs and Tabachnikov 1997] **D Fuchs**, **S Tabachnikov**, *Invariants of Legendrian and transverse knots in the standard contact space*, Topology 36 (1997) 1025–1053 MR Zbl

[Gabai 2001] **D Gabai**, *The Smale conjecture for hyperbolic* 3*-manifolds:* $\mathrm{Isom}(M^3) \simeq \mathrm{Diff}(M^3)$, J. Differential Geom. 58 (2001) 113–149 MR Zbl

[Geiges 2008] **H Geiges**, *An introduction to contact topology*, Cambridge Stud. Adv. Math. 109, Cambridge Univ. Press (2008) MR Zbl

[Geiges and Gonzalo Perez 2004] **H Geiges**, **J Gonzalo Perez**, *On the topology of the space of contact structures on torus bundles*, Bull. London Math. Soc. 36 (2004) 640–646 MR Zbl

[Geiges and Klukas 2014] **H Geiges**, **M Klukas**, *The fundamental group of the space of contact structures on the 3-torus*, Math. Res. Lett. 21 (2014) 1257–1262 MR Zbl

[Ghiggini and Van Horn-Morris 2016] **P Ghiggini**, **J Van Horn-Morris**, *Tight contact structures on the Brieskorn spheres* $-\Sigma(2,3,6n-1)$ *and contact invariants*, J. Reine Angew. Math. 718 (2016) 1–24 MR Zbl

[Gironella 2021] **F Gironella**, *Examples of contact mapping classes of infinite order in all dimensions*, Math. Res. Lett. 28 (2021) 707–727 MR Zbl

[Giroux 1991] **E Giroux**, *Convexité en topologie de contact*, Comment. Math. Helv. 66 (1991) 637–677 MR Zbl

[Giroux 1999] **E Giroux**, *Une infinité de structures de contact tendues sur une infinité de variétés*, Invent. Math. 135 (1999) 789–802 MR Zbl

[Giroux 2001] **E Giroux**, *Structures de contact sur les variétés fibrées en cercles audessus d'une surface*, Comment. Math. Helv. 76 (2001) 218–262 MR Zbl

[Giroux and Massot 2017] **E Giroux**, **P Massot**, *On the contact mapping class group of Legendrian circle bundles*, Compos. Math. 153 (2017) 294–312 MR Zbl

[Gompf 1998] **R E Gompf**, *Handlebody construction of Stein surfaces*, Ann. of Math. 148 (1998) 619–693 MR Zbl

[Gromov 1985] **M Gromov**, *Pseudo holomorphic curves in symplectic manifolds*, Invent. Math. 82 (1985) 307–347 MR Zbl

[Gromov 1986] **M Gromov**, *Partial differential relations*, Ergebnisse der Math. (3) 9, Springer (1986) MR Zbl

[Hansen 1978] **V L Hansen**, *The homotopy type of the space of maps of a homology* 3*-sphere into the* 2*-sphere*, Pacific J. Math. 76 (1978) 43–49 MR Zbl

[Hatcher 1976] **A Hatcher**, *Homeomorphisms of sufficiently large* P^2*-irreducible* 3*-manifolds*, Topology 15 (1976) 343–347 MR Zbl

[Hatcher 1981] **A Hatcher**, *On the diffeomorphism group of* $S^1 \times S^2$, Proc. Amer. Math. Soc. 83 (1981) 427–430 MR Zbl

[Hatcher 1983] **A E Hatcher**, *A proof of the Smale conjecture,* $\mathrm{Diff}(S^3) \simeq \mathrm{O}(4)$, Ann. of Math. 117 (1983) 553–607 MR Zbl

[Hatcher and Wahl 2010] **A Hatcher**, **N Wahl**, *Stabilization for mapping class groups of* 3*-manifolds*, Duke Math. J. 155 (2010) 205–269 MR Zbl

[Honda 2002] **K Honda**, *Gluing tight contact structures*, Duke Math. J. 115 (2002) 435–478 MR Zbl

[Hong et al. 2012] **S Hong**, **J Kalliongis**, **D McCullough**, **J H Rubinstein**, *Diffeomorphisms of elliptic 3-manifolds*, Lecture Notes in Math. 2055, Springer (2012) MR Zbl

[Ivanov 1976] **N V Ivanov**, *Groups of diffeomorphisms of Waldhausen manifolds*, Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. 66 (1976) 172–176 MR Zbl In Russian; translated in J. Soviet Math. 12 (1979) 115–118

[Kanda 1997] **Y Kanda**, *The classification of tight contact structures on the 3-torus*, Comm. Anal. Geom. 5 (1997) 413–438 MR Zbl

[Kronheimer and Mrowka 1997] **P B Kronheimer**, **T S Mrowka**, *Monopoles and contact structures*, Invent. Math. 130 (1997) 209–255 MR Zbl

[Kronheimer and Mrowka 2007] **P Kronheimer**, **T Mrowka**, *Monopoles and three-manifolds*, New Math. Monogr. 10, Cambridge Univ. Press (2007) MR Zbl

[Kronheimer and Mrowka 2020] **P B Kronheimer**, **T S Mrowka**, *The Dehn twist on a sum of two $K3$ surfaces*, Math. Res. Lett. 27 (2020) 1767–1783 MR Zbl

[Kronheimer et al. 2007] **P Kronheimer**, **T Mrowka**, **P Ozsváth**, **Z Szabó**, *Monopoles and lens space surgeries*, Ann. of Math. 165 (2007) 457–546 MR Zbl

[Kutluhan et al. 2020] **Ç Kutluhan**, **Y-J Lee**, **C H Taubes**, *HF = HM, I: Heegaard Floer homology and Seiberg–Witten Floer homology*, Geom. Topol. 24 (2020) 2829–2854 MR Zbl

[Lin 2016] **F Lin**, *Lectures on monopole Floer homology*, from "Proceedings of the Gökova geometry–topology conference 2015" (S Akbulut, D Auroux, T Önder, editors), GGT, Gökova (2016) 39–80 MR Zbl

[Lin 2023] **J Lin**, *Isotopy of the Dehn twist on $K3 \# K3$ after a single stabilization*, Geom. Topol. 27 (2023) 1987–2012 MR Zbl

[McCullough 1981] **D McCullough**, *Homotopy groups of the space of self-homotopy-equivalences*, Trans. Amer. Math. Soc. 264 (1981) 151–163 MR Zbl

[McCullough 1990] **D McCullough**, *Topological and algebraic automorphisms of 3-manifolds*, from "Groups of self-equivalences and related topics" (R A Piccinini, editor), Lecture Notes in Math. 1425, Springer (1990) 102–113 MR Zbl

[McCullough and Soma 2013] **D McCullough**, **T Soma**, *The Smale conjecture for Seifert fibered spaces with hyperbolic base orbifold*, J. Differential Geom. 93 (2013) 327–353 MR Zbl

[Milnor 1959] **J Milnor**, *On spaces having the homotopy type of a CW-complex*, Trans. Amer. Math. Soc. 90 (1959) 272–280 MR Zbl

[Min 2024] **H Min**, *The contact mapping class group and rational unknots in lens spaces*, Int. Math. Res. Not. 2024 (2024) 11315–11342 MR Zbl

[Mrowka and Rollin 2006] **T Mrowka**, **Y Rollin**, *Legendrian knots and monopoles*, Algebr. Geom. Topol. 6 (2006) 1–69 MR Zbl

[Muñoz-Echániz 2024] **J Muñoz-Echániz**, *A monopole invariant for families of contact structures*, Adv. Math. 439 (2024) art. id. 109483 MR Zbl

[Némethi 2017] **A Némethi**, *Links of rational singularities, L-spaces and LO fundamental groups*, Invent. Math. 210 (2017) 69–83 MR Zbl

[Neumann 1981] **W D Neumann**, *A calculus for plumbing applied to the topology of complex surface singularities and degenerating complex curves*, Trans. Amer. Math. Soc. 268 (1981) 299–344 MR Zbl

[Ozsváth and Szabó 2005] **P Ozsváth**, **Z Szabó**, *Heegaard Floer homology and contact structures*, Duke Math. J. 129 (2005) 39–61 MR Zbl

[Palais 1966] **R S Palais**, *Homotopy theory of infinite dimensional manifolds*, Topology 5 (1966) 1–16 MR Zbl

[Seidel 1999] **P Seidel**, *Lagrangian two-spheres can be symplectically knotted*, J. Differential Geom. 52 (1999) 145–171 MR Zbl

[Seidel 2000] **P Seidel**, *Graded Lagrangian submanifolds*, Bull. Soc. Math. France 128 (2000) 103–149 MR Zbl

[Seidel 2003] **P Seidel**, *A long exact sequence for symplectic Floer cohomology*, Topology 42 (2003) 1003–1063 MR Zbl

[Smirnov 2020] **G Smirnov**, *Seidel's theorem via gauge theory*, preprint (2020) arXiv 2010.03361

[Smirnov 2022] **G Smirnov**, *Symplectic mapping class groups of $K3$ surfaces and Seiberg–Witten invariants*, Geom. Funct. Anal. 32 (2022) 280–301 MR Zbl

[Taubes 1994] **C H Taubes**, *The Seiberg–Witten invariants and symplectic forms*, Math. Res. Lett. 1 (1994) 809–822 MR Zbl

[Taubes 2000] **C H Taubes**, SW $\Rightarrow$ Gr: *from the Seiberg–Witten equations to pseudo-holomorphic curves*, from "Seiberg Witten and Gromov invariants for symplectic 4-manifolds" (R Wentworth, editor), First Int. Press Lect. Ser. 2, International, Somerville, MA (2000) 1–97 MR Zbl

[Vogel 2018] **T Vogel**, *Non-loose unknots, overtwisted discs, and the contact mapping class group of S^3*, Geom. Funct. Anal. 28 (2018) 228–288 MR Zbl

[Watanabe 2018] **T Watanabe**, *Some exotic nontrivial elements of the rational homotopy groups of* $\mathrm{Diff}(S^4)$, preprint (2018) arXiv 1812.02448

[Weiss 2005] **M Weiss**, *What does the classifying space of a category classify?*, Homology Homotopy Appl. 7 (2005) 185–195 MR Zbl

EF: *Mathematics Department, University of Georgia*
Athens, GA, United States

JM-E: *Department of Mathematics, Columbia University*
New York, NY, United States

Current address: *Simons Center for Geometry and Physics, State University of New York, Stony Brook*
Stony Brook, NY, United States

eduardofernandez@uga.edu, jmunozechaniz@scgp.stonybrook.edu

Proposed: Yakov Eliashberg — Received: 3 July 2023
Seconded: Ciprian Manolescu, Leonid Polterovich — Revised: 9 April 2024

Geometry & Topology 29:3 (2025) 1619–1691
DOI: 10.2140/gt.2025.29.1619
Published: 31 May 2025

On boundedness and moduli spaces of K-stable Calabi–Yau fibrations over curves

KENTA HASHIZUME
MASAFUMI HATTORI

We show boundedness of polarized Calabi–Yau fibrations over curves only with fixed volumes of general fibers and Iitaka volumes. As its application, we construct a separated coarse moduli space of K-stable Calabi–Yau fibrations over curves in an adiabatic sense (Hattori 2022) and show that all members (resp. smooth members) of the moduli are simultaneously uniformly K-stable (resp. have cscK metrics) for a certain choice of polarizations.

14J10; 14J17, 14J27, 14J40

1 Introduction

1.1 Moduli problem

Classification of higher-dimensional algebraic varieties by their geometries is one of the most important problems in algebraic geometry. Moduli spaces, that are parameter spaces of specific classes of varieties, are effective tools to classify varieties.

The moduli spaces of stable curves were constructed by Deligne and Mumford [1969] as Deligne–Mumford stacks, and they are compactifications of the moduli spaces of smooth curves of general type of fixed genus; see also Mumford, Fogarty and Kirwan [Mumford et al. 1994]. After [Deligne and Mumford 1969], the moduli spaces of stable curves and the moduli spaces of canonically polarized surfaces with only canonical singularities have been constructed by Mumford [1977] and Gieseker [1977], respectively. For the construction, Mumford's geometric invariant theory (GIT, for short, see [Mumford et al. 1994]) was used, and the GIT-stability of those varieties was studied to apply the GIT. However, it is very difficult to detect the GIT-stability — more precisely, the asymptotic Chow stability, see [Mumford 1977] — of other kinds of polarized varieties. As an other strategy, Kollár and Shepherd-Barron [1988] (see also [Kollár 1990] and [Alexeev 1996b]) used the minimal model theory to construct the moduli spaces of stable surfaces. By their works and [Alexeev 1996a], semi-log-canonical models turned out to be a suitable higher-dimensional analog of stable curves to construct the moduli space. Their moduli spaces, called KSBA-moduli, have been completed as a full generalization of the moduli of stable curves. For details, see [Kollár 2023]. The recent developments of the minimal model theory [Birkar et al. 2010; Hacon and Xu 2013; Hacon et al. 2018] are indispensable for the theory of KSBA-moduli. The construction in

[Gieseker 1977] using GIT does not work for the KSBA-moduli because there is a klt variety with the ample canonical divisor that is asymptotically Chow unstable; see [Odaka 2012] for example.

We need a polarization when we discuss moduli theory of varieties whose canonical divisor is neither ample nor antiample. See [Seiler 1987] and [Viehweg 1995] for the study of the moduli theory of good minimal models with polarizations. In general, the moduli theory for noncanonically polarized varieties is much more complicated. For example, we cannot directly apply the theories as above to construct separated moduli of all polarized rational elliptic surfaces. The GIT-stability of rational Weierstrass fibrations [Miranda 1981] and Halphen pencils [Miranda 1980; Zanardini 2023; Hattori and Zanardini 2022] was investigated to consider the moduli of rational elliptic surfaces from the viewpoint of GIT. Seiler [1987] constructed the moduli space of some polarized elliptic surfaces by applying the GIT. He treated not only elliptic surfaces with nef canonical divisors but also rational elliptic surfaces whose fibers are reduced or of ${}_mI_n$-type. However, Seiler did not study the Chow stability of all polarized rational elliptic surfaces. By [Hattori 2022, Corollary 5.7], the moduli constructed by Seiler does not contain a polarized smooth rational elliptic surface (X, L) of index two with a unique constant scalar curvature Kähler (cscK, for short) metric. This (X, L) is asymptotically Chow stable; see [Donaldson 2001]. Thus, we naturally expect the existence of a moduli space parametrizing more polarized rational elliptic surfaces.

1.2 K-stability and K-moduli

K-stability was introduced by Tian [1997] and Donaldson [2002] in the context of the Kähler geometry to detect the existence of cscK metrics, and the notion is closely related to the GIT. Odaka [2012] found a relationship between the K-stability and the minimal model theory, and he proved that semi-log-canonical models are K-stable. This implies that the KSBA-moduli is a kind of moduli of K-polystable varieties. A moduli space parametrizing all K-polystable varieties is called a *K-moduli*. Odaka [2010, Conjecture 5.2] proposed the following conjecture.

Conjecture 1.1 (K-moduli conjecture) *There exists a quasiprojective moduli scheme parametrizing all polarized K-polystable varieties with fixed some numerical data* (*eg genera of curves*, *or volumes of polarizations*).

This conjecture was motivated by the work of Fujiki and Schumacher [1990] on the construction and a partial projectivity of moduli spaces of some projective manifolds with unique cscK metrics.

On the other hand, Dervan and Naumann [2018] constructed the moduli spaces of projective manifolds admitting cscK metrics and nondiscrete automorphism groups. As the Yau–Tian–Donaldson conjecture predicts that the K-polystability is equivalent to the existence of cscK metrics, the K-moduli can be thought of as an algebrogeometric generalization of the moduli in [Fujiki and Schumacher 1990] and [Dervan and Naumann 2018].

For the K-stability of log Fano pairs, algebraic geometers and differential geometers have made remarkable progress and they have constructed the projective moduli space of all K-polystable log Fano pairs; see Li, Wang and Xu [Li et al. 2019], Alper, Blum, Halpern-Leistner and Xu [Alper et al. 2020], Xu and Zhuang [2020] and Liu, Xu and Zhuang [Liu et al. 2022]. Moreover, quasiprojective moduli schemes of polarized K-polystable Calabi–Yau varieties have already been constructed in several cases [Odaka 2021]. Thus, it seems that we can make use of the K-stability to construct moduli spaces of polarized varieties. The following problems are keys to constructing the desired moduli spaces as Deligne–Mumford stacks.

(I) **Boundedness** Are polarized K-stable varieties parametrized by a scheme of finite type over $\mathbb{C}$?

(II) **Openness** Do polarized K-stable varieties form an open subset of the Hilbert scheme?

(III) **Separatedness** Let C be a smooth curve, $0 \in C$ and $(\mathcal{X}^\circ, \mathcal{L}^\circ) \to C^\circ$ be a family of polarized K-stable varieties over $C^\circ = C \setminus \{0\}$. Let $(\mathcal{X}, \mathcal{L}) \to C$ and $(\mathcal{X}', \mathcal{L}') \to C$ be two extensions of this family over C. If $(\mathcal{X}_0, \mathcal{L}_0)$ and $(\mathcal{X}'_0, \mathcal{L}'_0)$ are K-stable, does it hold that $(\mathcal{X}, \mathcal{L}) \cong (\mathcal{X}', \mathcal{L}')$?

In the case of (uniformly) K-stable $\mathbb{Q}$-Fano varieties, these problems had been settled. More precisely, (I) was solved by Jiang [2020] (see also [Xu and Zhuang 2021]), (II) was solved by Blum and Liu [2022], and (III) was solved by Blumand Xu [2019]. Their proofs are based on the work of Blum and Jonsson [2020], which shows that the δ-invariant introduced by Fujita and Odaka [2018] completely detects uniform K-stability of log Fano pairs. However, there are few criteria of the K-stability for other kinds of polarized varieties, and we do not know whether (I)–(III) hold or not.

1.3 Adiabatic K-stability and moduli

Adiabatic K-stability was introduced by the second author [Hattori 2024b] and it was inspired by the works of Fine [2004; 2007] and Dervan and Sektnan [2021b; 2021a] on the existence problem of cscK metrics of fibrations. Frankly speaking, uniform adiabatic K-stability [Hattori 2022, Definition 2.6] is designed to be "uniform" K-stability of fiber spaces when their polarizations are very close to ample line bundles on the base spaces. Such K-stability and cscK metrics on fiber spaces when their polarizations are very close to ample line bundles on the base are studied in [Fine 2004; 2007] and [Dervan and Sektnan 2021b; 2021a]. On the other hand, Dervan and Ross [2019] point out that there is a relationship between adiabatic K-stability and "K-stability" of the base. More precisely, they show that adiabatic K-semistability implies twisted K-semistability of the base. Recently, by replacing log twisted K-stability with twisted K-stability, the second author [Hattori 2022] showed for klt–trivial fibrations over curves that uniform adiabatic K-stability are equivalent to log-twisted K-stability of the base. Moreover, he showed the existence of cscK metrics corresponding to the uniform adiabatic K-stability for klt–trivial fibrations over curves. Using this criterion, the uniform adiabatic K-stability of elliptic surfaces is closely related to the GIT-stability of rational Weierstrass fibrations and Halphen pencils; see [Hattori 2022, Section 5] and [Hattori and Zanardini 2022, Remark 4.3]. Moreover, elliptic surfaces treated by Seiler are uniformly adiabatically K-stable, and the result in [Hattori 2022] (cf Definition 2.23) gave a useful characterization of the uniform adiabatic K-stability for klt-trivial fibrations over curves. Quite recently, (III) was also proved by the second author [Hattori 2024a] over $\mathbb{C}$, thus we may expect a variant of Conjecture 1.1 for the uniform adiabatic K-stability in an appropriate formulation.

The main purpose of this paper is to prove the following result.

Theorem 1.2 *There exists a moduli space parametrizing uniformly adiabatically K-stable klt-trivial fibrations over curves as a separated algebraic space of finite type.*

To state the result more precisely, we prepare some notation. Let d be a positive integer, v a positive rational number, and u a rational number. We set $\mathfrak{Z}_{d,v,u}$ to be

$$\left\{ f\colon (X,\Delta=0,A)\to C \;\middle|\; \begin{array}{ll} \text{(i)} & f \text{ is a uniformly adiabatically K-stable polarized klt-trivial fibration} \\ & \text{over a curve } C, \\ \text{(ii)} & \dim X = d, \\ \text{(iii)} & K_X \equiv u f^*H \text{ for some line bundle } H \text{ on } C \text{ such that } \deg H = 1, \\ \text{(iv)} & A \text{ is an } f\text{-ample line bundle on } X \text{ such that } (H\cdot A^{d-1}) = v. \end{array} \right\}$$

When $u \neq 0$, the boundedness result by Birkar [2023] implies the effectivity of the klt-trivial fibrations; see Lemma 3.1. More precisely, there exists a positive integer r, depending only on d, u and v, such that for any element $f\colon (X,0,A)\to C$ of $\mathfrak{Z}_{d,v,u}$, erK_X is a basepoint-free Cartier divisor and the linear system $|erK_X|$ defines f, where $e := u/|u|$. We can write the precise statement of Theorem 1.2 with this notation.

Theorem 1.3 *We fix $d\in\mathbb{Z}_{>0}$, $u\in\mathbb{Q}_{<0}$, $v\in\mathbb{Q}_{>0}$ and $r\in\mathbb{Z}_{>0}$ as above. For any locally Noetherian scheme S over $\mathbb{C}$, we define $\mathscr{M}_{d,v,u,r}(S)$ to be a groupoid whose objects are*

$$\left\{ \begin{tikzcd} (\mathcal{X},\mathscr{A}) \arrow[r, "f"] \arrow[dr, "\pi_{\mathcal{X}}"'] & \mathcal{C} \arrow[d] \\ & S \end{tikzcd} \;\middle|\; \begin{array}{ll} \text{(i)} & \pi_{\mathcal{X}} \text{ is a flat projective morphism and } \mathcal{X} \text{ is a scheme}, \\ \text{(ii)} & \mathscr{A}\in \mathrm{Pic}_{\mathcal{X}/S}(S) \text{ (see Section 2.2) such that } \mathscr{A}_{\overline{s}} \text{ is } f_{\overline{s}}\text{-ample for any} \\ & \text{geometric point } \overline{s}\in S, \\ \text{(iii)} & \omega^{[r]}_{\mathcal{X}/S} \text{ (see Definition 2.21) exists as a line bundle}, \\ \text{(iv)} & \pi_{\mathcal{X}*}\omega^{[-lr]}_{\mathcal{X}/S} \text{ is locally free and it generates } H^0(\mathcal{X}_s,\mathcal{O}_{\mathcal{X}_s}(-lrK_{\mathcal{X}_s})) \\ & \text{for any point } s\in S \text{ and any } l\in\mathbb{Z}_{>0}, \\ \text{(v)} & f \text{ is the ample model of } \omega^{[-r]}_{\mathcal{X}/S} \text{ over } S \text{ and } (\mathcal{X}_{\overline{s}},0,\mathscr{A}_{\overline{s}})\to \mathcal{C}_{\overline{s}}\in \mathfrak{Z}_{d,v,u} \\ & \text{for any geometric point } \overline{s}\in S. \end{array} \right\}$$

Here, we define an isomorphism $\alpha\colon (f\colon(\mathcal{X},\mathscr{A})\to\mathcal{C}) \to (f'\colon(\mathcal{X}',\mathscr{A}')\to\mathcal{C}')$ of any two objects of $\mathscr{M}_{d,v,u,r}(S)$ to be an S-isomorphism $\alpha\colon\mathcal{X}\to\mathcal{X}'$ such that there exists $\mathscr{B}\in\mathrm{Pic}_{\mathcal{C}/S}(S)$ satisfying that $\alpha^*\mathscr{A}' = \mathscr{A}\otimes f^*\mathscr{B}$ as elements of $\mathrm{Pic}_{\mathcal{X}/S}(S)$.

Then $\mathscr{M}_{d,v,u,r}$ is a separated Deligne–Mumford stack of finite type over $\mathbb{C}$ with a coarse moduli space.

We emphasize that $\mathscr{A}_{\bar{s}}$ are not assumed to be ample or the volumes of $\mathscr{A}_{\bar{s}}$ in $\mathscr{M}_{d,v,u,r}(S)$ are not bounded from above. Theorem 1.3 is the combination of the conditions (I)–(III) for the uniform adiabatic K-stability ([Hattori 2022, Theorem B], Corollary 3.8, Theorem 4.2, and Theorem 4.6) and Theorems 1.4 and 1.5 below, which are also key ingredients.

The first ingredient (Theorem 1.4 below) is the existence of a separated coarse moduli space that parametrizes $f\colon (X, 0, A) \to C \in \mathfrak{Z}_{d,v,u}$ such that f is uniformly adiabatically K-stable and A is an ample line bundle whose volume is bounded from above. We set

$$\mathfrak{Z}_{d,v,u,w} := \{f\colon (X,0,A) \to C \in \mathfrak{Z}_{d,v,u} \mid A \text{ is (globally) ample and } \mathrm{vol}(A) \le w\}$$

for any positive rational number w. Then the following holds.

Theorem 1.4 (see Theorem 5.1) *We fix $d \in \mathbb{Z}_{>0}$, $u \in \mathbb{Q}_{\neq 0}$ with $e := u/|u|$, $v \in \mathbb{Q}_{>0}$, $w \in \mathbb{Q}_{>0}$ and $r \in \mathbb{Z}_{>0}$ as above. For any locally Noetherian scheme S over $\mathbb{C}$, we define $\mathscr{M}_{d,v,u,w,r}(S)$ to be a groupoid whose objects are*

$$\left\{ \begin{tikzcd} (\mathcal{X},\mathscr{A}) \arrow[rr, "f"] \arrow[dr, "\pi_{\mathcal{X}}"'] & & \mathcal{C} \arrow[dl] \\ & S & \end{tikzcd} \;\middle|\; \begin{array}{rl} \text{(i)} & \pi_{\mathcal{X}} \text{ is a flat projective morphism and } \mathcal{X} \text{ is a scheme,} \\ \text{(ii)} & \mathscr{A} \in \mathrm{Pic}_{\mathcal{X}/S}(S) \text{ such that } \mathscr{A}_{\bar{s}} \text{ is ample for any geometric point } \bar{s} \in S, \\ \text{(iii)} & \omega_{\mathcal{X}/S}^{[r]} \text{ exists as a line bundle,} \\ \text{(iv)} & \pi_{\mathcal{X}*}\omega_{\mathcal{X}/S}^{[ler]} \text{ is locally free and it generates } H^0(\mathcal{X}_s, \mathcal{O}_{\mathcal{X}_s}(lerK_{\mathcal{X}_s})) \text{ for any} \\ & \text{point } s \in S \text{ and any } l \in \mathbb{Z}_{>0}, \\ \text{(v)} & f \text{ is the ample model of } \omega_{\mathcal{X}/S}^{[er]} \text{ over } S \text{ and } (\mathcal{X}_{\bar{s}}, 0, \mathscr{A}_{\bar{s}}) \to \mathcal{C}_{\bar{s}} \in \mathfrak{Z}_{d,v,u,w} \\ & \text{for any geometric point } \bar{s} \in S. \end{array} \right\}$$

Here, we define an isomorphism $\alpha\colon (f\colon (\mathcal{X},\mathscr{A}) \to \mathcal{C}) \to (f'\colon (\mathcal{X}',\mathscr{A}') \to \mathcal{C}')$ of any two objects of $\mathscr{M}_{d,v,u,w,r}(S)$ to be an S-isomorphism $\alpha\colon \mathcal{X} \to \mathcal{X}'$ such that $\alpha^*\mathscr{A}' = \mathscr{A}$ as elements of $\mathrm{Pic}_{\mathcal{X}/S}(S)$.

Then $\mathscr{M}_{d,v,u,w,r}$ is a separated Deligne–Mumford stack of finite type over $\mathbb{C}$ with a coarse moduli space.

When $u > 0$, Theorem 1.4 shows the existence of the moduli of the Iitaka fibrations from klt good minimal models of Iitaka dimension one; see [Birkar 2022] for the related topic.

We note that the isomorphisms in $\mathscr{M}_{d,v,u,w,r}$ are those in $\mathscr{M}_{d,v,u,r}$, but the converse is not necessarily true. The choice of r in Lemma 3.1 is not unique and the stacks $\mathscr{M}_{d,v,u,r}$ in Theorem 1.3 and $\mathscr{M}_{d,v,u,w,r}$ in Theorem 1.4 depend on the choice of r, however, their reduced structures are independent of r. For details, see Remark 5.7.

The second ingredient (Theorem 1.5 below) is the boundedness of $\mathscr{M}_{d,v,u,r}(\mathrm{Spec}\,\mathbb{C})$. In fact, we prove the following much stronger assertion. Let d be a positive integer, $\Theta \subset \mathbb{Q}$ a DCC set, v a positive rational number, and u a rational number. We set

$$\mathfrak{D}_{d,\Theta,v,u} := \left\{ f\colon (X,\Delta,A)\to C \;\middle|\; \begin{array}{rl} \text{(i)} & f\colon (X,\Delta)\to C \text{ is a klt-trivial fibration over a curve } C \text{ such that} \\ & K_X+\Delta\equiv uf^*H \text{ with a line bundle } H \text{ of degree one,} \\ \text{(ii)} & \dim X = d, \\ \text{(iii)} & \text{the coefficients of } \Delta \text{ belong to } \Theta, \\ \text{(iv)} & A \text{ is an } f\text{-ample } \mathbb{Q}\text{-Cartier Weil divisor such that } (H\cdot A^{d-1})=v. \end{array} \right\} \Big/ \sim,$$

where $\sim$ means relative linear equivalence of A over C. Let $[f\colon (X,\Delta,A)\to C]$ denote the equivalence class. Consider also for any $w>0$,

$$\mathfrak{G}_{d,\Theta,v,u,w} := \left\{ f\colon (X,\Delta,A)\to C \;\middle|\; \begin{array}{l} f \text{ satisfies the conditions (i)–(iv) such that } A \text{ is (globally) ample} \\ \text{and } \mathrm{vol}(A)\le w. \end{array} \right\}$$

Theorem 1.5 (boundedness) *Fix $d\in\mathbb{Z}_{>0}$, a DCC set $\Theta\subset\mathbb{Q}$, $v\in\mathbb{Q}_{>0}$ and $u\in\mathbb{Q}$. With notation as above, the following hold.*

(1) *The set of klt pairs (X,Δ) appearing in $\mathfrak{D}_{d,\Theta,v,u}$ is log bounded.*

(2) *There exists $w\in\mathbb{Q}_{>0}$, depending only on d, Θ, v and u, such that the natural map*

$$\mathfrak{G}_{d,\Theta,v,u,w}\to\mathfrak{D}_{d,\Theta,v,u}$$

is surjective.

After the paper was completed, Birkar informed the authors that he and Hacon obtained Theorem 1.5(1) independently. Theorem 1.5 not only shows the boundedness of $\mathfrak{Z}_{d,v,u}$ but also asserts that $\mathscr{M}_{d,v,u,r}$ is of finite type in Theorem 1.3. For the other statement of the boundedness, see Proposition 6.1.

We also study special K-stability, which was introduced by the second author [Hattori 2024a]. This is a stronger condition than uniform K-stability. By [Hattori 2024a], there exists an explicit criterion of the special K-stability without using test configurations, and the CM minimization conjecture, a numerical and stronger assertion than (III), holds for the spacial K-stability. We note that a uniformly adiabatically K-stable klt-trivial fibration over a curve is specially K-stable for a certain polarization [Hattori 2024a, Theorem 3.12]. We show that all members of $\mathfrak{G}_{d,\Theta,v,u,w}$ are simultaneously specially K-stable for a certain choice of polarizations as follows.

Theorem 1.6 (uniformity of adiabatic K-stability) *Let $d\in\mathbb{Z}_{>0}$, $\Theta\subset\mathbb{Q}$, $u\in\mathbb{Q}$ and $v\in\mathbb{Q}_{>0}$ be as in Theorem 1.5 and w be a positive rational number. Then, there exists an $\epsilon_0\in\mathbb{Q}_{>0}$, depending only on d, Θ, u, w and v, such that $(X,\Delta,\epsilon A+f^*H)$ is specially K-stable for any rational number $\epsilon\in(0,\epsilon_0)$, line bundle H on C of $\deg H=1$, and $f\colon (X,\Delta,A)\to C\in\mathfrak{G}_{d,\Theta,v,u,w}$.*

Furthermore, there exists $\alpha > 0$ *such that*

$$M_{\Delta}^{\mathrm{NA}}(\mathcal{X}, \mathcal{M}) \geq \alpha (\mathcal{J}^{\epsilon A + f^* H})^{\mathrm{NA}}(\mathcal{X}, \mathcal{M})$$

for any $f : (X, \Delta, A) \to C \in \mathfrak{G}_{d,\Theta,v,u,w}$ *with a line bundle* H *on* C *of* $\deg H = 1$, *normal semiample test configuration* $(\mathcal{X}, \mathcal{M})$ *for* $(X, \epsilon A + f^* H)$, *and rational number* $\epsilon \in (0, \epsilon_0)$.

It is known by [Zhang 2024] that every specially K-stable smooth polarized manifold (X, L) has a cscK metric in the first Chern class $\mathrm{c}_1(L)$. By Theorems 1.5 and 1.6, we have the following corollary on the "uniform" existence of cscK metrics.

Corollary 1.7 *Let* $d \in \mathbb{Z}_{>0}$, $\Theta \subset \mathbb{Q}$, $u \in \mathbb{Q}$ *and* $v \in \mathbb{Q}_{>0}$ *be as in Theorem 1.5. Then there exists a* $w > 0$, *depending only on* d, Θ, u *and* v, *satisfying the following: For any representative* $f : (X, \Delta, A) \to C$ *of any element of* $\mathfrak{D}_{d,\Theta,v,u}$ *with a general fiber* F *of* f, *if* $\mathrm{vol}(A + tF) \geq w$ *for some* $t \in \mathbb{Q}$ *then* $(X, \Delta, A + tF)$ *is specially K-stable.*

Furthermore, if X *is smooth and* $\Delta = 0$, *then* X *has a cscK metric* ω *in* $\mathrm{c}_1(A + tF)$.

Furthermore, Corollary 1.7 states that there is a "universal" family $\mathscr{U}'$ over $\mathscr{M}_{d,v,u,r}$ in Theorem 1.3 with a polarization $\mathscr{A}_{\mathscr{U}'}$ whose geometric fibers are specially K-stable varieties. Here, the word "universal" comes from the construction; see Remark 6.5.

1.4 Structure of this paper and overview of proof

The contents of this paper are as follows.

In Section 2, we collect notation and definitions in birational geometry, Hilbert schemes and stacks. To discuss the $\mathbb{Q}$-Gorensteinness of families, we explain the universal hull of coherent sheaves introduced by Kollár [2023]. We also collect basic facts of K-stability and some results on J-stability and uniform adiabatic K-stability [Hattori 2022], and we introduce a characterization of the uniform adiabatic K-stability of klt-trivial fibrations over curves (Definition 2.23). We make use of this characterization to construct our moduli spaces.

In Section 3, we prove Theorem 1.5. The idea is as follows: With notation in Theorem 1.5, we first give an upper bound n of the Cartier indices of the log canonical divisors (see Lemma 3.1) and we reduce Theorem 1.5 to the case where $\Theta = (1/n)\mathbb{Z} \cap [0, 1]$. We also know that (X, Δ) are $(1/n)$-lc. By using the boundedness of singularities, we next find a lower bound of the α-invariants of $A|_F$ with respect to $(F, \Delta|_F)$ for the general fibers F of f (Lemma 3.2). Since $(F, \Delta|_F)$ are $(1/n)$-lc pairs polarized by $A|_F$, the existence of the lower bound is a consequence of Birkar's result [2021b]. By using this lower bound and the semipositivity theorem by Fujino [2018, Theorem 1.11], we find an $m \in \mathbb{Z}$ such that $A + mF$ is ample and $\mathrm{vol}(A + mF)$ is universally bounded from above (Proposition 3.4, which is a special case of

Theorem 1.5(2)). Here m can be negative. From this result and [Birkar 2023], we obtain Theorem 1.5. The boundedness problem (I) will be solved in this section. For the construction of our moduli, we also prove a result on the finiteness of the Hilbert polynomials (Corollary 3.8).

In Section 4, we prove two important properties, ie the openness (Theorem 4.2) and the separatedness (Theorem 4.6) of uniformly adiabatically K-stable klt-trivial fibrations over curves. The openness is a direct consequence of the lower semicontinuity of the δ-invariants of the log twisted bases, which will be proved in Theorem 4.2. Note that we cannot apply [Blum and Liu 2022] since the case of families of polarized log pairs was studied in their paper. The separatedness has already been known by the second author [Hattori 2024a] when the varieties are over $\mathbb{C}$. In Theorem 4.6, we will give an alternative proof of the separatedness, which is an enhancement of [Hattori 2024a, Corollary 3.22] and works for any algebraically closed field of characteristic zero. These two results directly imply (II) and (III), respectively. Thus, we can obtain all the key conditions (I)–(III) for the uniform adiabatic K-stability in Sections 3 and 4. We also discuss the invariance of (anti)plurigenera used in the construction of our moduli spaces; see Theorem 4.8.

In Section 5, we prove Theorem 1.4; in other words, we construct the moduli space by using tools proved in Sections 3 and 4. We also show Theorem 1.3 by applying Proposition 3.4.

In Section 6, we prove Theorem 1.6 and Corollary 1.7. For this, we first show that there exist finitely many log $\mathbb{Q}$-Gorenstein families parametrizing polarized klt-trivial fibrations over curves (Proposition 6.1). Compared to Section 5, we deal with klt-trivial fibrations whose boundary divisors are not necessarily zero. However, these log $\mathbb{Q}$-Gorenstein families can be constructed by a similar argument to the proof of Theorem 1.4. In the case of nef log canonical divisors, Theorem 1.6 follows from a simple observation of J-stability in [Hattori 2021] (Theorem 6.2). For other case, we prove that the uniform "convergence of the δ-invariant" (cf [Hattori 2022, Theorem D]) holds for members of a family of klt-trivial fibrations (Proposition 6.3). Theorem 1.6 follows from these results, and Corollary 1.7 follows from Theorems 1.5 and 1.6.

Acknowledgements This work was partially supported by JSPS KAKENHI 22K13887 for Hashizume and JSPS KAKENHI 22J20059 (Grant-in-Aid for JSPS Fellows DC1) for Hattori. The authors are grateful to Professors Makoto Enokizono and Yuji Odaka and Doctors Dai Imaike, Kentaro Inoue, and Kazuhiro Ito for fruitful discussions, which improved Theorem 1.4. The authors thank Professors János Kollár and Shou Yoshikawa for answering questions on the book [Kollár 2023]. The authors thank Professors Caucher Birkar, Yoshonori Hashimoto, and Gang Tian and Doctor Eiji Inoue for useful comments.

2 Preliminaries

Throughout this paper, we work over an algebraically closed field $\Bbbk$ of characteristic zero unless otherwise stated.

Notation and conventions

We collect notation and conventions used in this paper.

(1) A *scheme* means a locally Noetherian scheme over $\Bbbk$. For a scheme X, we denote the induced reduced scheme by X_{red}. A *variety* means an integral separated scheme of finite type over $\Bbbk$. A *curve* means a smooth variety of dimension one.

A *geometric point* of X is a morphism $\operatorname{Spec}\Omega \to X$, where Ω is an algebraically closed field. For a point $x \in X$, $\bar{x}$ denotes the geometric point of X which maps the unique point of $\operatorname{Spec}\Omega$ to x. We simply denote it by $\bar{x} \in X$.

(2) For any scheme S and positive integer d, we denote $\mathbb{P}^d_{\Bbbk} \times_{\operatorname{Spec}\Bbbk} S$ by $\mathbb{P}^d_S$. We simply write $\mathbb{P}^d$ if there is no risk of confusion, for example if $S = \operatorname{Spec}\Bbbk$ or S is a geometric point of a scheme. Let $p\colon \mathbb{P}^d_S \to \mathbb{P}^d_{\Bbbk}$ be the projection. For any $m \in \mathbb{Z}$, we often denote $p^*\mathcal{O}_{\mathbb{P}^d_{\Bbbk}}(m)$ by $\mathcal{O}(m)$, and we sometimes think $\mathcal{O}(m)$ of a Cartier divisor on $\mathbb{P}^d_S$ if there is no risk of confusion.

(3) A morphism $f\colon X \to Y$ of schemes is called a *contraction* if f is projective and $f_*\mathcal{O}_X \cong \mathcal{O}_Y$. For a morphism $f\colon X \to Y$ of schemes and a (geometric) point $y \in Y$, the fiber of f over y is denoted by X_y. For a $\mathbb{Q}$-divisor D on X, we denote the restriction of D to $f^{-1}(y)$ by D_y when it is well-defined; for example, when D is $\mathbb{Q}$-Cartier at every codimension-one point of $f^{-1}(y)$ and $\operatorname{Supp} D \not\supset f^{-1}(y)$. We note that D_y does not coincide with the scheme-theoretic fiber in general.

(4) Let X be a smooth variety, D an snc divisor on X, and $f\colon X \to Z$ a morphism to a scheme Z. We say that (X, D) is *log smooth over* Z or $f\colon (X, D) \to Z$ is *log smooth* if f is a smooth surjective morphism and for any stratum T of (X, D), the restriction $f|_T\colon T \to Z$ is also a smooth surjective morphism.

(5) We say that a subset of $\mathbb{R}$ satisfies the *descending chain condition* (*DCC*, for short) if the subset does not contain any strictly decreasing infinite sequence. We say that a subset of $\mathbb{R}$ satisfies the *ascending chain condition* (*ACC*, for short) if the subset does not contain any strictly increasing infinite sequence. A subset of $\mathbb{R}$ is called a *DCC set* (resp. an *ACC set*) if the subset satisfies the DCC (resp. ACC).

(6) Let a be a real number. Then we define $\lceil a \rceil$ to be the unique integer satisfying $\lceil a \rceil - 1 < a \le \lceil a \rceil$. Let X be a normal variety and let D be an $\mathbb{R}$-divisor on X. Let $D = \sum_i d_i D_i$ be the prime decomposition. Then we define $\lceil D \rceil := \sum_i \lceil d_i \rceil D_i$. We say that D is a *Weil divisor* if every coefficients of D is an integer, in other words, $D = \lceil D \rceil$ holds. We define the reduced divisor D_{red} of D to be $\sum_i D_i$.

(7) Let X be a normal variety. For a line bundle (resp. $\mathbb{Q}$-line bundle, $\mathbb{R}$-line bundle) L on X, we often think L of a Cartier (resp. $\mathbb{Q}$-Cartier, $\mathbb{R}$-Cartier) divisor on X. When L is a line bundle on X, we often denote $L^{\otimes m} \otimes \mathcal{O}_X(D)$ by $\mathcal{O}_X(mL + D)$ for every Weil divisor D on X.

(8) Let $f\colon X \to Y$ be a morphism of schemes. Let L_1 and L_2 be line bundles on X. We say that L_1 and L_2 are *linearly equivalent over* Y, denoted by $L_1 \sim_Y L_2$, if there is a line bundle L on Y such that

$L_1 \cong L_2 \otimes f^*L$. When Y is a point, we simply say that L_1 and L_2 are *linearly equivalent* and we write $L_1 \sim L_2$.

Suppose that X is a normal variety. Let D_1 and D_2 be $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisors on X. We say that D_1 and D_2 are *$\mathbb{Q}$-linearly equivalent over Y*, denoted by $D_1 \sim_{\mathbb{Q},Y} D_2$, if there exists a positive integer m such that both mD_1 and mD_2 are Cartier and $\mathcal{O}_X(mD_1) \sim_Y \mathcal{O}_X(mD_2)$. This definition is not standard. However, the definition coincides with the usual definition of the relative $\mathbb{Q}$-linear equivalence when Y is a variety (eg f is a contraction). When Y is a point, we simply say that D_1 and D_2 are *$\mathbb{Q}$-linearly equivalent* and we write $D_1 \sim_{\mathbb{Q}} D_2$.

(9) Let X be a projective scheme over $\Bbbk$, let A be a Cartier divisor of X, and let $\phi_{|A|}$ be a rational map $X \dashrightarrow \mathbb{P}^{h^0(X,\mathcal{O}_X(A))-1}$ defined by the linear system $|A|$. If A is semiample, $\phi_{|mA|}$ induces a contraction for every sufficiently large and divisible $m > 0$, and this is a kind of ample model defined in [Birkar et al. 2010, Lemma 3.6.5(3)]. Similarly, for a projective morphism $\pi\colon \mathcal{X} \to S$ of schemes and a π-semiample line bundle $\mathscr{A}$ on $\mathcal{X}$, we call a morphism $f\colon \mathcal{X} \to \mathbf{Proj}_S\big(\bigoplus_{l \geq 0} \pi_*\mathscr{A}^{\otimes l}\big)$ the *ample model* of $\mathscr{A}$ over S.

(10) Let $f_1\colon X_1 \to Y_1$ and $f_2\colon X_2 \to Y_2$ be morphisms of schemes over a scheme S. Then the induced morphism $X_1 \times_S X_2 \to Y_1 \times_S Y_2$ from f_1 and f_2 is denoted by $f_1 \times_S f_2$. When $S = \operatorname{Spec} \Bbbk$, we simply write $f_1 \times f_2\colon X_1 \times X_2 \to Y_1 \times Y_2$.

(11) For any morphisms $f\colon \mathcal{X} \to S$ and $h\colon T \to S$, we denote $\mathcal{X} \times_S T$ by $\mathcal{X}_T$ and the base change $\mathcal{X}_T \to T$ by f_T. For any coherent sheaf $\mathscr{A}$ on $\mathcal{X}$, we denote $(h \times_S \mathrm{id}_{\mathcal{X}})^*\mathscr{A}$ by $\mathscr{A}_T$. For an f-ample line bundle H on $\mathcal{X}$ and a polynomial p, if $\chi(\mathscr{A}_s(tH_s)) = p(t)$ for every $t \in \mathbb{Z}$, then we say that *$\mathscr{A}_s$ has the Hilbert polynomial p with respect to H.*

(12) Let $f\colon Y \to C$ be a contraction from a normal variety to a curve and D a $\mathbb{Q}$-divisor on Y. Then we can decompose D into $D_{\mathrm{vert}} + D_{\mathrm{hor}}$, where the support of D_{hor} is flat over C and the support of D_{vert} has zero-dimensional image in C.

Definition 2.1 Let S be a Noetherian scheme and let $S_1, \dots, S_l$ be locally closed subschemes of S that are disjoint in each other and $\bigsqcup_{1=1}^{l} S_i = S$ set-theoretically. Then we call the natural inclusion $\bigsqcup_{1=1}^{l} S_i \to S$ a *locally closed decomposition*. A subset $F \subset S$ is called a *constructible subset* if F is a finite union of locally closed subsets.

Lemma 2.2 *Let S be a scheme of finite type over $\Bbbk$. Suppose that $F \subset S$ is a constructible subset. Then F is closed if and only if the following holds.*

- *For any morphism $\varphi\colon C \to S$ from an affine curve C, if $\varphi^{-1}(F)$ is dense, then $\varphi(C) \subset F$.*

Proof The assertion is local and we may assume that S is affine. Suppose that the condition holds. Let $\overline{F}$ be the Zariski closure and assume that there exists a point $s \in \overline{F} \setminus F$. Take an irreducible component Z of $\overline{F}$ containing s. It is easy to see that F contains a nonempty open subset of Z; cf [Matsumura 1980, 6.C].

On the other hand, since $\overline{\{s\}} \cap F$ is not dense in $\overline{\{s\}}$, there is a closed point $s_0 \in \overline{\{s\}} \setminus F$. By [Mumford 2008, Section 6, Lemma], there exists a morphism $\varphi\colon C \to Z$ from an affine curve such that $\varphi^{-1}(F)$ is dense and $s_0 \in \varphi(C)$. Thus, $s_0 \in F$ by the condition and this is a contradiction. □

2.1 Birational geometry

In this subsection, we collect definitions concerned with singularities of pairs, klt-trivial fibration, and boundedness.

Definition 2.3 (singularities of pairs) A *subpair* (X, Δ) consists of a proper normal variety X and a $\mathbb{Q}$-divisor Δ on X such that $K_X + \Delta$ is $\mathbb{Q}$-Cartier. A subpair is called a *pair* if the coefficients of Δ are positive. Let F be a prime divisor over X and take $\pi\colon Y \to X$ a proper birational morphism from a normal variety such that F appears as a divisor on Y. Then we define the *log discrepancy* of F with respect to (X, Δ) by

$$A_{(X,\Delta)}(F) := 1 + \mathrm{ord}_F(K_Y - \pi^*(K_X + \Delta)),$$

where ord_F is the divisorial valuation associated to F with $\mathrm{ord}_F(F) = 1$. It is easy to see that $A_{(X,\Delta)}(F)$ is independent of π. A (sub)pair (X, Δ) is called *(sub)klt* (resp. *(sub)lc*, ϵ-*(sub)lc*) if $A_{(X,\Delta)}(F) > 0$ (resp. ≥ 0, $\geq \epsilon$) for every prime divisor F over X. We say that a proper normal variety V is a *klt variety* if $(V, 0)$ is a klt pair.

For an effective $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor M on a normal variety X, the *log canonical threshold* of M with respect to a subpair (X, Δ), denoted by $\mathrm{lct}(X, \Delta; M)$, is defined as follows: If there exists a t such that $(X, B + tM)$ is sublc, then

$$\mathrm{lct}(X, \Delta; M) := \sup\{t \in \mathbb{Q} \mid (X, \Delta + tM) \text{ is sublc}\},$$

and otherwise we set $\mathrm{lct}(X, \Delta; M) := -\infty$.

Definition 2.4 (Iitaka volume) Let X be a normal projective variety and let D be a $\mathbb{Q}$-Cartier divisor on X such that the Iitaka dimension $\kappa(X, D)$ is nonnegative. Then the *Iitaka volume* of D, denoted by $\mathrm{Ivol}(D)$, is defined by

$$\mathrm{Ivol}(D) := \limsup_{m\to\infty} \frac{\dim H^0(X, \mathcal{O}_X(\lfloor mD \rfloor))}{m^{\kappa(X,D)}/\kappa(X, D)!}.$$

When D is big, the Iitaka volume of D coincides with the usual volume. By definition, we can easily check that $\mathrm{Ivol}(rD) = r^{\kappa(X,D)} \cdot \mathrm{Ivol}(D)$ for every $r \in \mathbb{Z}_{>0}$.

Definition 2.5 (klt-trivial fibration) Let (X, Δ) be a klt pair, and let $f\colon X \to C$ be a contraction of normal projective varieties. Then $f\colon (X, \Delta) \to C$ is called a *klt-trivial fibration* if $K_X + \Delta \sim_{\mathbb{Q},C} 0$.

For a klt-trivial fibration $f\colon (X, \Delta) \to C$, we define the *discriminant* $\mathbb{Q}$*-divisor* B_C and the *moduli* $\mathbb{Q}$*-divisor* M_C on C as follows: For every prime divisor P on C, let b_P be the largest real number such that after shrinking C around the generic point η of P, the pair $(X, \Delta + b_P f^* P)$ is lc. Note that b_P is well-defined since P is Cartier at η. Then we define the discriminant $\mathbb{Q}$-divisor B_C by

$$B_C := \sum_P (1 - b_P) P,$$

where P runs over prime divisors on C. Next, fix a $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor L on C such that $K_X + \Delta \sim_{\mathbb{Q}} f^* L$. Then the moduli $\mathbb{Q}$-divisor M_C is defined by

$$M_C := L - (K_C + B_C).$$

Note that M_C is only defined up to $\mathbb{Q}$-linear equivalence class. We call

$$K_X + \Delta \sim_{\mathbb{Q}} f^*(K_C + B_C + M_C)$$

the *canonical bundle formula*.

In general, we can define klt-trivial fibrations for subpairs and contractions, cf [Ambro 2004]. However, for simplicity we always assume that (X, Δ) in klt-trivial fibrations are klt pairs.

We make use of the following fundamental fact.

Theorem 2.6 [Ambro 2004, Theorem 0.1] *If* $\dim C = 1$, *then* M_C *is a semiample* $\mathbb{Q}$*-Cartier* $\mathbb{Q}$*-divisor.*

Definition 2.7 (discriminant $\mathbb{Q}$-divisor with respect to contraction) By extending the notion of the discriminant $\mathbb{Q}$-divisors in Definition 2.5, for every sublc pair (X, Δ) with a contraction $f : X \to Z$ of normal varieties, we define the *discriminant* $\mathbb{Q}$*-divisor with respect to* $f : (X, \Delta) \to Z$ as follows: For each prime divisor P on Z, we define

$$\mu_P := \sup\{\gamma \in \mathbb{R} \mid (X, \Delta + \gamma f^* P) \text{ is sublc over the generic point of } P\}.$$

We may assume that $f^* P$ is well-defined since we may shrink Z around the generic point of P. Define

$$B := \sum_P (1 - \mu_P) P,$$

where P runs over prime divisors on Z.

It is easy to see that this definition coincides with the discriminant $\mathbb{Q}$-divisor in Definition 2.5 when $f : (X, \Delta) \to Z$ is a klt-trivial fibration.

Definition 2.8 (boundedness) We say a set $\mathfrak{Q}$ of normal projective varieties is *bounded* if there exist finitely many projective morphisms $V_i \to T_i$ of varieties such that for each $X \in \mathfrak{Q}$ there exist an index i, a closed point $t \in T_i$, and an isomorphism $\phi : (V_i)_t \to X$.

A *couple* (X, S) consists of a normal projective variety X and a reduced divisor S on X. We use the term couple because $K_X + S$ is not assumed to be $\mathbb{Q}$-Cartier. We say that a set $\mathfrak{P}$ of couples is *bounded* if there exist finitely many projective morphisms $V_i \to T_i$ of varieties and a reduced divisor C on each V_i such that for any $(X,S)\in\mathfrak{P}$ there exist an index i, a closed point $t\in T_i$, and an isomorphism $\phi\colon (V_i)_t\to X$ such that $((V_i)_t, C_t)$ is a couple and $C_t=\phi_*^{-1}S$.

Finally, we say that a set $\mathfrak{R}$ of projective pairs (X,Δ) is *log bounded* if the set of the corresponding couples $(X,\operatorname{Supp}\Delta)$ is bounded.

2.2 Hilbert schemes

Let $f:\mathcal{X}\to S$ be a proper morphism of schemes and $\mathscr{A}$ an f-ample line bundle on $\mathcal{X}$. Then $\operatorname{Hilb}\mathcal{X}/S^{p,\mathscr{A}}$ denotes the scheme representing the following functor $\mathfrak{Hilb}^{p,\mathscr{A}}_{\mathcal{X}/S}$. For any morphism $T\to S$, we attain

$$\mathfrak{Hilb}^{p,\mathscr{A}}_{\mathcal{X}/S}(T):=\left\{\mathcal{Z}\subset\mathcal{X}_T\;\middle|\;\begin{array}{l}\mathcal{Z}\text{ is a closed subscheme of }\mathcal{X}_T\text{ flat over }T\text{ whose fibers have}\\ \text{the same Hilbert polynomial }p\text{ with respect to }\mathscr{A}.\end{array}\right\}$$

We remark that $\operatorname{Hilb}\mathcal{X}/S^{p,\mathscr{A}}$ exists as a locally projective scheme over S. Indeed, it is well-known that if $\mathscr{A}$ is further f-very ample, then $\operatorname{Hilb}\mathcal{X}/S^{p,\mathscr{A}}$ is projective over S entirely; cf [Fantechi et al. 2005, Section 5]. Therefore, for any quasicompact open subset $U\subset S$, by taking $m>0$ such that $\mathscr{A}^{\otimes m}|_{\mathcal{X}_U}$ is f_U-very ample, we see that

$$\operatorname{Hilb}\mathcal{X}_U/U^{p,\mathscr{A}|_{\mathcal{X}_U}}=\operatorname{Hilb}\mathcal{X}_U/U^{q,\mathscr{A}^{\otimes m}|_{\mathcal{X}_U}}$$

exists as a projective scheme over S, where $q(n)=p(mn)$ for any $n\in\mathbb{Z}$. By patching $\operatorname{Hilb}\mathcal{X}_U/U^{p,\mathscr{A}|_{\mathcal{X}_U}}$ together over S, we obtain a unique locally projective scheme $\operatorname{Hilb}\mathcal{X}/S^{p,\mathscr{A}}$ over S up to isomorphism. In this paper, we call $\operatorname{Hilb}\mathcal{X}/S^{p,\mathscr{A}}$ the *Hilbert scheme*. When $S=\operatorname{Spec}\Bbbk$, we simply write $\operatorname{Hilb}\mathcal{X}^{p,\mathscr{A}}$. We write $\bigsqcup_p\operatorname{Hilb}\mathcal{X}/S^{p,\mathscr{A}}$ by $\operatorname{Hilb}\mathcal{X}/S$, where p runs over polynomials.

Next, we assume that f is flat and it has geometrically connected and normal fibers. Let $g:\mathcal{Y}\to S$ be another proper morphism of schemes and $\mathscr{B}$ a g-ample line bundle on $\mathcal{Y}$ such that the fibers of g have the Hilbert polynomial p with respect to $\mathscr{B}$. For any S-scheme T, we set

$$\mathfrak{Isom}_S(\mathcal{X},\mathcal{Y})(T)=\{h\colon\mathcal{X}_T\to\mathcal{Y}_T\text{ is a }T\text{-isomorphism}\},$$
$$\mathfrak{Isom}_S((\mathcal{X},\mathscr{A}),(\mathcal{Y},\mathscr{B}))(T)=\{h\in\mathfrak{Isom}_S(\mathcal{X},\mathcal{Y})(T)\mid h^*\mathscr{B}_T\sim_T\mathscr{A}_T\}.$$

By [Fantechi et al. 2005, Theorem 5.23] and Corollary 2.20, which we will treat later, the functor $\mathfrak{Isom}_S(\mathcal{X},\mathcal{Y})$ (resp. $\mathfrak{Isom}_S((\mathcal{X},\mathscr{A}),(\mathcal{Y},\mathscr{B}))$) is represented by a locally closed subscheme $\operatorname{Isom}_S(\mathcal{X},\mathcal{Y})$ (resp. $\operatorname{Isom}_S((\mathcal{X},\mathscr{A}),(\mathcal{Y},\mathscr{B}))$) of $\operatorname{Hilb}\mathcal{X}\times_S\mathcal{Y}/S$ (resp. $\operatorname{Hilb}^{q,\,p_1^*\mathscr{A}\otimes p_2^*\mathscr{B}}_{\mathcal{X}\times_S\mathcal{Y}/S}$), where q is the polynomial defined by $q(m)=p(2m)$ and $p_1:\mathcal{X}\times_S\mathcal{Y}\to\mathcal{X}$ (resp. $p_2:\mathcal{X}\times_S\mathcal{Y}\to\mathcal{Y}$) is the first (resp. second) projection. Thus, we see that $\operatorname{Isom}_S((\mathcal{X},\mathscr{A}),(\mathcal{Y},\mathscr{B}))$ is locally quasiprojective over S. If $(\mathcal{X},\mathscr{A})=(\mathcal{Y},\mathscr{B})$, we set

$$\operatorname{Aut}_S(\mathcal{X},\mathscr{A}):=\operatorname{Isom}_S((\mathcal{X},\mathscr{A}),(\mathcal{X},\mathscr{A})).$$

For details, we refer to [Fantechi et al. 2005, Section 5.6]. We note that if we define

$$\mathfrak{Aut}(\mathbb{P}^n_{\mathbb{Z}})(T):=\{T\text{-automorphisms of }\mathbb{P}^n_T\}$$

for any locally Noetherian scheme T, then it is well-known that the functor $\mathfrak{Aut}(\mathbb{P}^n_{\mathbb{Z}})$ is represented by $\mathrm{PGL}(n+1,\mathbb{Z})$ (cf [Mumford et al. 1994, Section 0.5(b)]).

On the other hand, we define a presheaf $\mathfrak{Pic}_{\mathcal{X}/S}$ as follows: For any morphism $T \to S$, we attain the following set

$$\mathfrak{Pic}_{\mathcal{X}/S}(T) = \{L\colon \text{a line bundle on } \mathcal{X}_T\}/{\sim_T}.$$

In other words, $\mathfrak{Pic}_{\mathcal{X}/S}(T)$ is the set of all relative linear equivalence classes of line bundles on $\mathcal{X}_T$ over T. In general, $\mathfrak{Pic}_{\mathcal{X}/S}$ is not an étale sheaf. However, it is well-known (see [Mumford et al. 1994, Section 0.5] and [Fantechi et al. 2005, Section 9]) that under the same assumption on $\mathcal{X} \to S$ as the previous paragraph, there exists a separated scheme $\mathrm{Pic}_{\mathcal{X}/S}$ locally of finite type over S such that there exist the maps for all $T \to S$,

$$\mathfrak{Pic}_{\mathcal{X}/S}(T) \hookrightarrow \mathrm{Hom}_S(T, \mathrm{Pic}_{\mathcal{X}/S}),$$

that are injective and induce the étale sheafification $\mathfrak{Pic}_{\mathcal{X}/S} \to \mathrm{Hom}_S(\bullet, \mathrm{Pic}_{\mathcal{X}/S})$ of $\mathfrak{Pic}_{\mathcal{X}/S}$. Moreover, if $\mathcal{X} \to S$ has a section, then $\mathfrak{Pic}_{\mathcal{X}/S}$ coincides with $\mathrm{Hom}_S(\bullet, \mathrm{Pic}_{\mathcal{X}/S})$.

2.3 Stacks

We refer to [Olsson 2016, Sections 3 and 5] and [Stacks 2005–] for the notation of fibered categories and algebraic spaces. Let **Sets** be the category of sets and $\mathbf{Sch}_{/S}$ the category of (locally Noetherian) schemes over S. If $S = \operatorname{Spec} \Bbbk$, we write $\mathbf{Sch}_{/\Bbbk}$. For any scheme S, we endow $\mathbf{Sch}_{/S}$ with the étale topology. We recall the definition of stacks; see [Olsson 2016, Proposition 4.6.2].

Definition 2.9 (stacks) Let $p\colon \mathscr{C} \to \mathbf{Sch}_{/S}$ be a category fibered in groupoids. $\mathscr{C}$ is called a *stack* over S if the following two conditions hold (cf [Olsson 2016, Definition 4.6.1]).

(1) For any S-scheme X and any two objects $x, y \in \mathscr{C}(X) := p^{-1}(X)$, the presheaf $\mathfrak{Isom}_X(x, y)$, defined by

$$\mathfrak{Isom}_X(x,y)(f\colon Y \to X) := \mathrm{Isom}_Y(f^*x, f^*y),$$

where the right-hand side is the set of all isomorphisms g such that $p(g) = \mathrm{id}_Y$, is an étale sheaf.

(2) For any étale covering in $\mathbf{Sch}_{/S}$, all descent data with respect to the covering are effective; cf [Olsson 2016, Definition 4.2.6].

Remark 2.10 In the situation of Definition 2.9, we consider the following condition.

- For any set of S-schemes $\{X_i\}_{i\in I}$, the natural functor

$$\mathscr{C}\Big(\bigsqcup_{i\in I} X_i\Big) \to \prod_{i\in I} \mathscr{C}(X_i)$$

is an equivalence of categories.

We note that all stacks we treat in this paper satisfy this condition. Here, we explain the definition of descent data when $\mathscr{C}$ satisfies the above condition. We say that a surjective morphism $f\colon X' \to X$ is an *étale covering* if $X' := \bigsqcup_{i \in I} X_i$ and $f|_{X_i}\colon X_i \to X$ is étale for any $i \in I$. For any étale covering f, let

$$p_1, p_2\colon X' \times_X X' \to X' \quad \text{and} \quad p_{12}, p_{23}, p_{13}\colon X' \times_X X' \times_X X' \to X' \times_X X'$$

be the projections. A pair $(u' \in \mathscr{C}(X'), \sigma)$ is called a *descent datum* with respect to $f\colon X' \to X$ if $\sigma \in \mathrm{Isom}_{X' \times_X X'}(p_1^* u', p_2^* u')$ such that $p_{23}^*\sigma \circ p_{12}^*\sigma = p_{13}^*\sigma$. Note that for any $u \in \mathscr{C}(X)$, there exists a canonical isomorphism $\sigma_{\mathrm{can}} \in \mathrm{Isom}_{X' \times_X X'}(p_1^* f^* u, p_2^* f^* u)$ such that $(f^*u, \sigma_{\mathrm{can}})$ is a descent datum. If there is $u \in \mathscr{C}(X)$ such that $\sigma \circ p_1^* g = p_2^* g \circ \sigma_{\mathrm{can}}$ for some $g \in \mathrm{Isom}(f^*u, u')$, then we call (u', σ) an *effective descent datum*. We see that our definition and the original definition [Olsson 2016, Definition 4.2.6] of descent data are the same in this situation by [Olsson 2016, Lemma 4.2.7].

Definition 2.11 (Artin stacks, Deligne–Mumford stacks) Let $\mathscr{C}$ be a stack over $\Bbbk$. $\mathscr{C}$ is called a *Deligne–Mumford* (*resp. Artin*) *stack* if the following hold.

(i) The diagonal $\Delta\colon \mathscr{C} \to \mathscr{C} \times \mathscr{C}$ is representable, ie for any morphism $U \to \mathscr{C} \times \mathscr{C}$ from a scheme, $U \times_{\mathscr{C} \times \mathscr{C}} \mathscr{C}$ is an algebraic space.

(ii) There exists an étale (resp. smooth) surjective morphism $\pi\colon X \to \mathscr{C}$ from a scheme.

If $\mathscr{C}$ is a Noetherian Artin stack, we can define coherent sheaves on $\mathscr{C}$ in the way of [Olsson 2016, Definition 9.1.14]. If $\mathscr{L}$ is a coherent sheaf on $\mathscr{C}$ and there exists a smooth surjection $g\colon T \to \mathscr{C}$ such that $g^*\mathscr{L}$ is a line bundle, we say that $\mathscr{L}$ is a *line bundle* on $\mathscr{C}$; see also [Olsson 2016, Section 9.3].

Example 2.12 It is well-known that $\mathbf{Sch}_{/S}$ has the natural stack structure over $\Bbbk$ for any scheme S; see [Deligne and Mumford 1969, page 97]. We simply denote this stack by S. For any scheme T, we know that $\mathbf{Sch}_{/S}(T) = \mathrm{Hom}(T, S)$. We denote this by $S(T)$.

Example 2.13 Let X be a scheme of finite type over $\Bbbk$ and G be a linear algebraic group over $\Bbbk$. Then there exists a quotient stack $[X/G]$ defined as [Olsson 2016, Example 8.1.12]. We remark that $[X/G]$ is an Artin stack of finite type over $\Bbbk$; cf [Stacks 2005–, Tag 036O]. Note that X is quasicompact. Moreover, for any G-equivariant line bundle L on X, we can find a line bundle $\mathscr{L}$ on $[X/G]$ such that $\pi^*\mathscr{L} = L$, where $\pi\colon X \to [X/G]$ is the canonical projection; see [Olsson 2016, Exercise 9.H]. This $\mathscr{L}$ is unique up to isomorphism.

The following category will be used in Section 5.

Definition 2.14 Let $\mathfrak{Pol}$ be the category such that the collection of objects is

$$\left\{ f\colon (\mathcal{X}, \mathscr{A}) \to S \;\middle|\; \begin{array}{l} f \text{ is a surjective proper flat morphism of schemes whose geometric fibers} \\ \text{are normal and connected, and } \mathscr{A} \in \mathrm{Pic}_{\mathcal{X}/S}(S) \text{ such that there exists an étale} \\ \text{covering } S' \to S \text{ by which the pullback of } \mathscr{A} \text{ to } \mathcal{X} \times_S S' \text{ is represented by a} \\ \text{relatively ample line bundle over } S', \end{array} \right\}$$

and an arrow $(g,\alpha)\colon (f\colon(\mathcal{X},\mathscr{A})\to S)\to(f'\colon(\mathcal{X}',\mathscr{A}')\to S')$ is defined in the way that $\alpha\colon S\to S'$ is a morphism and $g\colon\mathcal{X}\to\mathcal{X}'\times_{S'}S$ is an isomorphism such that $g^*\alpha^*_{\mathcal{X}'}\mathscr{A}'=\mathscr{A}$ as elements of $\mathrm{Pic}_{\mathcal{X}'/S'}(S)$.

It is easy to see that there exists a natural functor $p\colon\mathfrak{Pol}\to\mathbf{Sch}_{/\Bbbk}$ such that $\mathfrak{Pol}$ is a category fibered in groupoids. We further show the following.

Lemma 2.15 *$\mathfrak{Pol}$ is a stack over $\Bbbk$.*

Proof It suffices to check the conditions (1) and (2) of Definition 2.9 for $\mathfrak{Pol}$. We first treat (1). Take objects $f\colon(\mathcal{X},\mathscr{A})\to S$ and $f'\colon(\mathcal{X}',\mathscr{A}')\to S$. Then $\mathfrak{Isom}_S(\mathcal{X},\mathcal{X}')$ is represented by a locally closed subscheme of $\mathrm{Hilb}\,\mathcal{X}\times_S\mathcal{X}'/S$; see [Fantechi et al. 2005, Section 5.6]. Since $\mathrm{Pic}_{\mathcal{X}/S}$ is separated, we see that $\mathfrak{Isom}_S(f,f')\hookrightarrow\mathfrak{Isom}_S(\mathcal{X},\mathcal{X}')$ is a closed immersion. Therefore $\mathfrak{Isom}_S(f,f')$ is represented by a scheme. In particular, it is an étale sheaf. Hence, (1) holds.

Next, we treat (2). One can check that for any set of schemes $\{X_i\}_{i\in I}$, the natural functor

$$\mathfrak{Pol}\Big(\coprod_{i\in I}X_i\Big)\to\prod_{i\in I}\mathfrak{Pol}(X_i)$$

is an equivalence of categories. By Remark 2.10, it suffices to show the following: For any étale covering $S'\to S$ with the projections

$$p_1,\ p_2\colon S'\times_S S'\to S'\quad\text{and}\quad p_{12},\ p_{23},\ p_{13}\colon S'\times_S S'\times_S S'\to S'\times_S S',$$

any descent datum $(f'\colon(\mathcal{X}',\mathscr{A}')\to S',\sigma)$ is effective. Here, $\sigma\in\mathrm{Isom}_{S'\times_S S'}(p_1^*f',p_2^*f')$. If the pullback of (f',σ) by an étale covering $T\to S$ is effective, then so is (f',σ) by the condition (1). From this fact, by replacing S' with a scheme T admitting an étale covering $T\to S'$, we may assume that $\mathscr{A}'$ is an f'-ample line bundle.

By the f'-ampleness, there exists $m\in\mathbb{Z}_{>0}$ such that $H^i(\mathcal{X}'_s,\mathscr{A}'^{\otimes m}_s)=0$ for every $s\in S'$ and $i>0$ and the natural morphism $\mathcal{X}'\to\mathbb{P}_{S'}(f'_*\mathscr{A}'^{\otimes m})$ is a closed immersion. We note that for any flat morphism $g\colon T\to S$, we have the natural isomorphism

$$f'_{T*}g^*_{\mathcal{X}}\mathscr{A}'^{\otimes m}\cong g^*f'_*\mathscr{A}'^{\otimes m}$$

by [Hartshorne 1977, III, Proposition 9.3]. Thus, we may identify $f'_{T*}g^*_{\mathcal{X}}\mathscr{A}'^{\otimes m}$ with $g^*f'_*\mathscr{A}'^{\otimes m}$. Furthermore, $f'_*\mathscr{A}'^{\otimes m}$ is locally free by [Hartshorne 1977, III, Theorem 12.11]. On the other hand, there exist a line bundle $\mathcal{M}$ and an isomorphism

$$h\colon\sigma^*p^*_{2,\mathcal{X}'}\mathscr{A}'\cong p^*_{1,\mathcal{X}'}\mathscr{A}'\otimes(p_1^*f')^*\mathcal{M},$$

where $p_{1,\mathcal{X}'}\colon\mathcal{X}'\times_S S'\to\mathcal{X}'$ (resp. $p_{2,\mathcal{X}'}\colon\mathcal{X}'\times_S S'\to\mathcal{X}'$) is the morphism induced from base change of p_1 (resp. p_2) by the canonical morphism $\mathcal{X}'\to S'$, and p_1^*f' (resp. p_2^*f') is the base change of f' by p_1 (resp. p_2). Then h induces the isomorphism

$$\varphi\colon\mathbb{P}_{S'\times_S S'}(p_1^*f'_*\mathscr{A}'^{\otimes m})=\mathbb{P}_{S'\times_S S'}(p_1^*f'_*\mathscr{A}'^{\otimes m}\otimes\mathcal{M}^{\otimes m})\xrightarrow{\cong}\mathbb{P}_{S'\times_S S'}(p_2^*f'_*\mathscr{A}'^{\otimes m}).$$

Here, the first equality is via the canonical isomorphism in [Hartshorne 1977, II, Lemma 7.9].

The following easy claim implies that φ is independent of the choice of h.

Claim *Let $p\colon \mathcal{Y} \to T$ be a proper flat surjective morphism whose geometric fibers are connected and normal, and let $\mathscr{L}$ be a line bundle on $\mathcal{X}$. Suppose that $p_*\mathscr{L}$ is locally free. Any isomorphism $\tau\colon \mathscr{L} \to \mathscr{L}$ induces the identity morphism of $\mathbb{P}_T(p_*\mathscr{L})$.*

Proof Now $\tau \in \mathrm{Hom}(\mathscr{L},\mathscr{L})$, and $\mathrm{Hom}(\mathscr{L},\mathscr{L}) \cong H^0(\mathcal{Y},\mathcal{O}_{\mathcal{Y}}) \cong H^0(T,\mathcal{O}_T)$ by [Fantechi et al. 2005, 9.3.11]. Since τ is an isomorphism, we may regard τ as an element of $H^0(T,\mathcal{O}_T^*)$. Then the argument in the proof of [Hartshorne 1977, II, Lemma 7.9] works without any change. $\square$

We continue to prove Lemma 2.15. We will show that φ defines a descent datum, ie $p_{23}^*\varphi \circ p_{12}^*\varphi = p_{13}^*\varphi$. Clearly, this is equivalent to that $(p_{13}^*\varphi)^{-1} \circ p_{23}^*\varphi \circ p_{12}^*\varphi$ is the identity. We note that if the base change of the morphism σ in the decent datum (f',σ) by p_{12} (resp. p_{23}, p_{13}) is denoted by $p_{12}^*\sigma$ (resp. $p_{23}^*\sigma$, $p_{13}^*\sigma$), then $p_{23}^*\sigma \circ p_{12}^*\sigma = p_{13}^*\sigma$ holds. This follows from the definition of descent data. We also note that the relative linear equivalence p_{ij}^*h induces $p_{ij}^*\varphi$ for any $1 \le i < j \le 3$. Thus, $(p_{13}^*\varphi)^{-1} \circ p_{23}^*\varphi \circ p_{12}^*\varphi$ is induced from the linear equivalence over $S' \times_S S' \times_S S'$

$$\begin{aligned} p_{12,\mathcal{X}'\times_S S'}^* p_{1,\mathcal{X}'}^*\mathscr{A}' &= (p_{13}^*\sigma)^*((p_{23}^*\sigma)^{-1})^*((p_{12}^*\sigma)^{-1})^* p_{12,\mathcal{X}'\times_S S'}^* p_{1,\mathcal{X}'}^*\mathscr{A}' \\ &\sim_{S'\times_S S'\times_S S'} (p_{13}^*\sigma)^*((p_{23}^*\sigma)^{-1})^* p_{12,S'\times_S \mathcal{X}'}^* p_{2,\mathcal{X}'}^*\mathscr{A}' \\ &= (p_{13}^*\sigma)^*((p_{23}^*\sigma)^{-1})^* p_{23,\mathcal{X}'\times_S S'}^* p_{1,\mathcal{X}'}^*\mathscr{A}' \\ &\sim_{S'\times_S S'\times_S S'} (p_{13}^*\sigma)^* p_{13,S'\times_S \mathcal{X}'}^* p_{2,\mathcal{X}'}^*\mathscr{A}' \\ &\sim_{S'\times_S S'\times_S S'} p_{12,\mathcal{X}'\times_S S'}^* p_{1,\mathcal{X}'}^*\mathscr{A}', \end{aligned}$$

where $p_{12,\mathcal{X}'\times_S S'}\colon \mathcal{X}' \times_S S' \times_S S' \to \mathcal{X}' \times_S S'$ is the base change of p_{12} by the canonical morphism $\mathcal{X}' \times_S S' \to S' \times_S S'$, and $p_{23,\mathcal{X}'\times_S S'}\colon \mathcal{X}' \times_S S' \times_S S' \to \mathcal{X}' \times_S S'$ and $p_{13,\mathcal{X}'\times_S S'}\colon \mathcal{X}' \times_S S' \times_S S' \to \mathcal{X}' \times_S S'$ are defined similarly. By Claim, it immediately follows that $(p_{13}^*\varphi)^{-1} \circ p_{23}^*\varphi \circ p_{12}^*\varphi$ is the identity morphism. Thus $p_{23}^*\varphi \circ p_{12}^*\varphi = p_{13}^*\varphi$.

On the other hand, $-K_{\mathbb{P}_{S'}(f'_*\mathscr{A}'^{\otimes m})/S'}$ is relatively ample over S. Hence, applying [Olsson 2016, Proposition 4.4.12] to $\mathbb{P}_{S'}(f'_*\mathscr{A}'^{\otimes m})$ and $-K_{\mathbb{P}_{S'}(f'_*\mathscr{A}'^{\otimes m})/S'}$, we may find a scheme $\mathcal{P}$ and a projective flat surjective morphism $\mathcal{P} \to S$ that canonically defines a descent datum isomorphic to $(\mathbb{P}_{S'}(f'_*\mathscr{A}'^{\otimes m}),\varphi)$. Note that $\mathcal{P}$ is not a projective bundle but every geometric fiber over S is a projective space. By applying [Olsson 2016, Proposition 4.4.3] to the closed immersion $\mathcal{X}' \hookrightarrow \mathbb{P}_{S'}(f'_*\mathscr{A}'^{\otimes m})$, we obtain a closed immersion $\mathcal{X} \hookrightarrow \mathcal{P}$ whose base change by $S' \to S$ coincides with $\mathcal{X}' \hookrightarrow \mathbb{P}_{S'}(f'_*\mathscr{A}'^{\otimes m})$. On the other hand, by the definition of the Picard scheme, there exists a unique element $\mathscr{A} \in \mathrm{Pic}_{\mathcal{X}/S}(S)$ such that the pullback of $\mathscr{A}$ to $\mathcal{X} \times_S S'$ coincides with $\mathscr{A}'$.

From the above facts, (f',σ) is effective. We finish the proof of Lemma 2.15. $\square$

Remark 2.16 Let $f\colon \mathcal{X} \to S$ be a proper surjective flat morphism of schemes whose geometric fibers are normal and connected. We fix $\mathscr{A} \in \mathrm{Pic}_{\mathcal{X}/S}$. Then $(\mathcal{X},\mathscr{A}) \to S$ is an object of $\mathfrak{Pol}$ if and only if $\mathscr{A}_{\bar{s}}$ is ample for any geometric point $\bar{s} \in S$. Indeed, we may replace f by $\mathcal{X} \times_S S' \to S'$ for some étale covering $S' \to S$, thus we may assume that $\mathscr{A}$ is a line bundle. Then $\mathscr{A}$ is f-ample if and only if $\mathscr{A}_{\overline{s}}$ is ample for any geometric point $\overline{s} \in S$; see [Kollár and Mori 1998, Proposition 1.41]. The converse is easy.

The following theorem is well-known to experts and holds since we assume that $\mathrm{char}(\mathbb{k}) = 0$; cf [Mumford 2008, Section 11, Theorem].

Theorem 2.17 [Olsson 2016, Remark 8.3.4; Keel and Mori 1997] *Let $\mathscr{C}$ be an Artin stack of finite type over $\mathbb{k}$. If the diagonal morphism $\Delta\colon \mathscr{C} \to \mathscr{C} \times \mathscr{C}$ is finite, then $\mathscr{C}$ is a separated Deligne–Mumford stack. Furthermore, there exists a separated coarse moduli space of finite type over $\mathbb{k}$.*

Remark 2.18 The authors in [Olsson 2016] and [Stacks 2005–] treat the category of schemes that are not necessarily locally Noetherian, but our theory works even if we treat $\mathbf{Sch}_{/\mathbb{k}}$. For example, we can extend $\mathfrak{Pol}$ to a stack over the category of all schemes, including schemes that are not locally Noetherian; see [Vakil 2010, 28.2.12]. We can also apply Theorem 2.17 to $[N/\mathrm{PGL}(d_1) \times \mathrm{PGL}(d_2) \times \mathrm{PGL}(d_3)]$, which is defined on the category of all schemes, in the proof of Theorem 5.1.

2.4 Universal hull and $\mathbb{Q}$-Gorenstein family

For any scheme X and coherent sheaf $\mathscr{F}$ on X of pure dimension, we can define the *S_2-hull* of $\mathscr{F}$, which we denote by $\mathscr{F}^{[**]}$. For details, we refer to [Huybrechts and Lehn 1997, Section 1.1]. If X is a normal variety of dimension d and $\mathscr{F}$ is of pure dimension d, then $\mathscr{F}^{[**]} = \mathcal{H}om_{\mathcal{O}_X}(\mathcal{H}om_{\mathcal{O}_X}(\mathscr{F}, \mathcal{O}_X), \mathcal{O}_X)$.

Let $f\colon \mathcal{X} \to S$ be a flat projective surjective morphism between locally Noetherian schemes such that the relative dimension of f is d and all geometric fibers of f are normal and connected. Then there is a closed reduced subscheme $Z \subset \mathcal{X}$ such that f is smooth on $\mathcal{X} \setminus Z$ and the fiber Z_s over any $s \in S$ satisfies $\mathrm{codim}_{\mathcal{X}_s}(Z_s) \geq 2$. Let $\mathscr{F}$ be a coherent sheaf on X such that $\mathscr{F}|_{\mathcal{X}\setminus Z}$ is an invertible sheaf. We define a (*universal*) *hull* of $\mathscr{F}$, which we also denote by $\mathscr{F}^{[**]}$, to be a coherent sheaf with the following properties (cf [Kollár 2008] or [Kollár 2023, Section 9]):

- $\mathscr{F}^{[**]}$ is flat over S.
- There exists a morphism $q\colon \mathscr{F} \to \mathscr{F}^{[**]}$ that is an isomorphism outside Z.
- For any point $s \in S$, the morphism $\mathscr{F}^{[**]}|_{\mathcal{X}_s} \to \mathscr{F}_s^{[**]}$ induced by q is an isomorphism, where $\mathscr{F}_s^{[**]}$ is the S_2-hull of $\mathscr{F}_s := \mathscr{F}|_{\mathcal{X}_s}$.

A universal hull does not always exist for the sheaf $\mathscr{F}$ as above, but if it exists then $\mathscr{F}^{[**]} \cong j_*(\mathscr{F}|_{\mathcal{X}\setminus Z})$, where $j\colon \mathcal{X}\setminus Z \hookrightarrow \mathcal{X}$ is the inclusion. Indeed, for any $p \in \mathcal{X}$ and any affine open neighborhood $U \subset \mathcal{X}$ of p, let $I_Z \subset \mathcal{O}_{\mathcal{X}}$ be the ideal sheaf corresponding to Z. Take a regular sequence $\overline{a}, \overline{b} \in I_Z \otimes \mathcal{O}_{U\cap\mathcal{X}_s}(U \cap \mathcal{X}_s)$ of $\mathscr{F}_s^{[**]}$ for $s := f(p)$, ie $\overline{a}$ is a nonzero divisor in $\mathscr{F}_s^{[**]}$ and $\overline{b}$ is a nonzero divisor in $\mathscr{F}_s^{[**]}/\overline{a}\mathscr{F}_s^{[**]}$. If $\overline{a}, \overline{b}$ are the restrictions of $a, b \in I_Z(U)$, then a, b is also a regular sequence of $\mathscr{F}^{[**]}$ around $\mathcal{X}_s \cap U$ by [Matsumura 1980, (20.E)]. Thus, shrinking U if necessary, we may assume that there exists a regular sequence $a, b \in I_Z(U)$ of $\mathscr{F}^{[**]}$. Now the natural map $\mathscr{F}^{[**]} \to j_*(\mathscr{F}|_{\mathcal{X}\setminus Z})$ is injective over U, and the surjectivity can be proved as follows: Let $m \in j_*(\mathscr{F}|_{\mathcal{X}\setminus Z})$ be a local section over U. By the assumption, there exists two sections $m_a, m_b \in \mathscr{F}^{[**]}(U)$ such that $m = m_a/a = m_b/b$. Here, we applied [Matsumura 1980, Theorem 27] and assumed that b is also a nonzero divisor by shrinking U. Thus, $bm_a = am_b$ as elements of $\mathscr{F}^{[**]}(U)$. From this and the fact that b is a nonzero divisor in $(\mathscr{F}^{[**]}/a\mathscr{F}^{[**]})(U)$, we have $m \in \mathscr{F}^{[**]}(U)$.

Hence, if a universal hull of $\mathscr{F}$ exists, then $\mathscr{F}^{[**]} \cong j_*(\mathscr{F}|_{\mathcal{X}\setminus Z})$, and furthermore we have $(\mathscr{F}^{[**]})_T = (\mathscr{F}_T)^{[**]}$ for any morphism $g\colon T \to S$. We denote this by $\mathscr{F}_T^{[**]}$.

By applying Kollár's theory [2023] to our setup, we obtain the following theorem.

Theorem 2.19 (cf [Kollár 2023, Theorem 9.40]) *Let $f\colon \mathcal{X} \to S$ be a flat projective surjective morphism between schemes of finite type over $\Bbbk$ such that the relative dimension of f is d and the geometric fibers of f are normal and connected. Let $Z \subset \mathcal{X}$ be a closed subset such that f is smooth on $\mathcal{X} \setminus Z$ and the fiber Z_s over any $s \in S$ satisfies $\operatorname{codim}_{\mathcal{X}_s}(Z_s) \geq 2$. Let $\mathscr{F}$ be a coherent sheaf on X such that $\mathscr{F}|_{\mathcal{X}\setminus Z}$ is an invertible sheaf on $\mathcal{X} \setminus Z$. Let H be an f-ample line bundle on $\mathcal{X}$.*

Then there exist finitely many distinct polynomials $p_1, \dots, p_l$ with corresponding locally closed subschemes $S_1, \dots, S_l$ of S satisfying the following:

- *$S = \bigsqcup_{i=1}^{l} S_i$ set-theoretically.*
- *For each $1 \leq i \leq l$, there exists the universal hull $\mathscr{F}_{S_i}^{[**]}$ of $\mathscr{F}_{S_i}$ such that the Hilbert polynomial of $\mathscr{F}_s^{[**]}$ with respect to H is p_i for all $s \in S_i$.*
- *For any morphism $g\colon T \to S$ from a locally Noetherian scheme T, if $\mathscr{F}_T$ has a universal hull $\mathscr{F}_T^{[**]}$ such that all fibers $\mathscr{F}_t^{[**]}$ have the same Hilbert polynomial p with respect to H_t, then $p = p_i$ and g factors through S_i for some i.*

The following result was proved by Hassett and Kovács [2004, Theorem 3.11] when the fibers are Cohen–Macaulay, and Kollár [2023, Proposition 9.42] proved a more general statement. Thus, we omit the proof.

Corollary 2.20 *Let $f\colon \mathcal{X} \to S$, $\mathscr{F}$, and H be as in Theorem 2.19. For any line bundle L on $\mathcal{X}$, there exists a locally closed subscheme $S^u \subset S$ such that a morphism $g\colon T \to S$ from a locally Noetherian scheme T factors through $S^u \hookrightarrow S$ if and only if the universal hull $\mathscr{F}_T^{[**]}$ exists and $L_T \otimes f_T^* M \cong \mathscr{F}_T^{[**]}$ for some line bundle M on T.*

From now on, we deal with the relative dualizing sheaf. Let $f\colon \mathcal{X} \to S$ be a flat projective surjective morphism of schemes of finite type over $\Bbbk$ whose geometric fibers are normal and connected, and let $U \subset \mathcal{X}$ be the largest open subscheme such that f is smooth at every point of U. Let $\omega_{\mathcal{X}/S}$ be the relative dualizing sheaf. Then $\omega_{\mathcal{X}/S}^{\otimes m}$ is a coherent sheaf and $\omega_{\mathcal{X}/S}^{\otimes m}|_U$ is an invertible sheaf for every $m \in \mathbb{Z}$.

Hence, we may use the previous results to $\omega_{\mathcal{X}/S}^{\otimes m}$. For each $m \in \mathbb{Z}$, if the universal hull of $\omega_{\mathcal{X}/S}^{\otimes m}$ exists, then $\omega_{\mathcal{X}/S}^{[m]}$ denotes the (universal) hull. We also have $\omega_{\mathcal{X}_T/T}^{[m]} = (h \times_S \mathrm{id}_{\mathcal{X}})^* \omega_{\mathcal{X}/S}^{[m]}$ for every morphism $h: T \to S$ since $\omega_{U/S} = \omega_{\mathcal{X}/S}|_U$ is a line bundle that commutes with the base change.

Definition 2.21 ($\mathbb{Q}$-Gorenstein family, log $\mathbb{Q}$-Gorenstein family) Let $f: \mathcal{X} \to S$ be a flat projective surjective morphism of schemes of finite type over $\Bbbk$ whose geometric fibers are normal and connected. We say that $f: \mathcal{X} \to S$ is a *$\mathbb{Q}$-Gorenstein family* over S if there exists $m \in \mathbb{Z}_{>0}$ such that $\omega_{\mathcal{X}/S}^{[m]}$ exists as a line bundle.

For any $f: \mathcal{X} \to S$ as above, if S is normal, then $\mathcal{X}$ is also normal, $\omega_{\mathcal{X}/S}$ is reflexive and $\omega_{\mathcal{X}/S} = \mathcal{O}_{\mathcal{X}}(K_{\mathcal{X}/S})$ for some Weil divisor $K_{\mathcal{X}/S}$ on $\mathcal{X}$; see [Patakfalvi et al. 2018, Proposition A10] and [Codogni and Patakfalvi 2021, Section 2].

Let $f: \mathcal{X} \to S$ be as above. Suppose that S is normal. Let Δ be an effective $\mathbb{Q}$-divisor on $\mathcal{X}$ such that the support of Δ contains no fiber of f. We say that $f: (\mathcal{X}, \Delta) \to S$ is a *log $\mathbb{Q}$-Gorenstein family* if $K_{\mathcal{X}/S} + \Delta$ is $\mathbb{Q}$-Cartier.

Remark 2.22 Let $f: \mathcal{X} \to S$ be as above. Suppose that S is normal.

- Let $U \subset \mathcal{X}$ be the open locus on which f is smooth. If $f: (\mathcal{X}, \Delta) \to S$ is a log $\mathbb{Q}$-Gorenstein family, then $\omega_{\mathcal{X}/S}|_U = \omega_{U/S}$ is an invertible sheaf (cf [Stacks 2005–, 0E9Z]), and thus $\Delta|_U$ is $\mathbb{Q}$-Cartier. For any morphism $h: T \to S$ from a normal variety T and the induced morphism $\sigma: \mathcal{X}_T \to \mathcal{X}$, we define Δ_T as a unique extension of $\sigma^*(\Delta|_U)$ on $U \times_S T$. Then we can check that
$$K_{\mathcal{X}_T/T} + \Delta_T = \sigma^*(K_{\mathcal{X}/S} + \Delta)$$
by applying [Conrad 2000, Theorem 3.6.1] to $U \times_S T \to U$. See also [Codogni and Patakfalvi 2021, Section 2].
- Let $\mathcal{D}$ be an effective Weil divisor on $\mathcal{X}$ such that $\mathcal{D}$ is flat over S as a scheme and it has only geometrically integral fibers over S. Then the scheme-theoretic fiber $\mathcal{D}_s$ for any $s \in S$ is also a Weil divisor, $\mathcal{O}_{\mathcal{X}}(-\mathcal{D})$ is also flat and the restriction $\mathcal{O}_{\mathcal{X}}(-\mathcal{D})|_U$ is locally free by [Huybrechts and Lehn 1997, Lemma 2.1.7].
- Let Δ be an effective $\mathbb{Q}$-divisor on $\mathcal{X}$ such that the support of Δ contains no fiber of f. Here, we do not assume that $f: (\mathcal{X}, \Delta) \to S$ is a log $\mathbb{Q}$-Gorenstein family. Let $j: U \hookrightarrow \mathcal{X}$ be the open immersion. We fix $m \in \mathbb{Z}_{>0}$ such that $m\Delta$ is a Weil divisor on $\mathcal{X}$. If a universal hull of $\mathcal{O}_{\mathcal{X}}(m(K_{\mathcal{X}/S} + \Delta))$ exists, then the S_2 condition of $\mathcal{O}_{\mathcal{X}}(m(K_{\mathcal{X}/S} + \Delta))$ implies
$$\mathcal{O}_{\mathcal{X}}(m(K_{\mathcal{X}/S} + \Delta)) = j_*\mathcal{O}_U(m(K_{U/S} + \Delta|_U)) = \mathcal{O}_{\mathcal{X}}(m(K_{\mathcal{X}/S} + \Delta))^{[**]}.$$
Moreover, if any irreducible component of Δ is flat over S as a reduced scheme and it has only geometrically integral fibers over S, then $\mathcal{O}_{\mathcal{X}}(m(K_{\mathcal{X}/S} + \Delta))|_U$ is locally free. Then, we can apply Corollary 2.20 to $\mathcal{O}_{\mathcal{X}}(m(K_{\mathcal{X}/S} + \Delta))$ and any line bundle L on $\mathcal{X}$ to construct a locally closed subscheme $S^u \subset S$ satisfying the property of Corollary 2.20.

2.5 K-stability

In this subsection, we collect some definitions and known results on K-stability.

A *polarized variety* (X, L) consists of a proper normal variety X and an ample $\mathbb{Q}$-line bundle L on it. The notation of polarized varieties and subpairs are the same, however, we adopt this notation because both are standard. We will mainly deal with subpairs in Sections 3, 4 and 6, whereas we will deal with polarized varieties in Section 5.

Let Δ be a $\mathbb{Q}$-divisor such that (X, Δ) is a pair. We call (X, Δ, L) a *polarized pair*. We denote the algebraic group

$$\{g \in \mathrm{Aut}(X) \mid g_*\Delta = \Delta,\ g^*L \sim_{\mathbb{Q}} L\}$$

by $\mathrm{Aut}(X, \Delta, L)$. This is a closed subscheme of $\{g \in \mathrm{Aut}(X) \mid g^*L \sim_{\mathbb{Q}} L\}$, which is a group scheme of finite type over $\Bbbk$ since $\chi(X, L^{\otimes m} \otimes g^*L^{\otimes n}) = \chi(X, L^{\otimes m+n})$ for every sufficiently divisible m and $n \in \mathbb{Z}_{>0}$; see [Fantechi et al. 2005, Section 5.6]. Hence, the above algebraic group is also of finite type over $\Bbbk$. We can check that $\mathrm{Aut}(X, \Delta, L)$ is a linear algebraic group. Indeed, for any sufficiently divisible $m \in \mathbb{Z}_{>0}$, since there exists a well-defined closed immersion $G_m \hookrightarrow \mathrm{PGL}(h^0(X, L^{\otimes m}))$, the group scheme

$$G_m := \{g \in \mathrm{Aut}(X) \mid g_*\Delta = \Delta,\ g^*L^{\otimes m} \sim L^{\otimes m}\}$$

is affine. Since $\mathrm{Aut}(X, \Delta, L)$ is an algebraic group and

$$\mathrm{Aut}(X, \Delta, L) = \bigcup_{m:\ \text{sufficiently divisible}} G_m$$

as sets, we have $\mathrm{Aut}(X, \Delta, L) = G_m$ for some m. Hence, $\mathrm{Aut}(X, \Delta, L)$ is affine.

We say that $f : (X, \Delta, A) \to C$ is a *polarized klt-trivial fibration over a curve* if $f : (X, \Delta) \to C$ is a klt-trivial fibration over a proper curve and A is an f-ample $\mathbb{Q}$-line bundle on X.

We give the following ad hoc definition of uniform adiabatic K-stability of f.

Definition 2.23 (uniform adiabatic K-stability) We say that a polarized klt-trivial fibration over a curve $f : (X, \Delta, A) \to C$ is *uniformly adiabatically K-stable* if one of the following hold:

- $K_X + \Delta \sim_{\mathbb{Q}} f^*(K_C + B_C + M_C)$ is nef, or
- $C = \mathbb{P}^1$, $K_X + \Delta \sim_{\mathbb{Q}} uf^*(\mathcal{O}(1))$ for some $u < 0$, and $\max_{p \in \mathbb{P}^1} \mathrm{ord}_p(B_C) < 1 + \frac{1}{2}u$, where B_C is the discriminant $\mathbb{Q}$-divisor with respect to f.

Here, B_C and M_C are $\mathbb{Q}$-divisors defined in Definition 2.5.

We note that the uniform adiabatic K-stability is a condition of (C, B_C, M_C), which we call a *log-twisted pair*, rather than f.

Next, we recall the definition of K-stability, but we do not need it except in Section 6.

Definition 2.24 (K-stability) Let (X, Δ, L) be a polarized log pair of dimension d. We say that $\pi\colon (\mathcal{X}, \mathcal{L}) \to \mathbb{A}^1$ is a *(semi)ample test configuration* if the following hold.

- $\pi\colon \mathcal{X} \to \mathbb{A}^1$ is a proper and flat morphism of schemes.
- $\mathcal{L}$ is a (semi)ample $\mathbb{Q}$-line bundle on $\mathcal{X}$.
- $\mathbb{G}_m$ acts on $(\mathcal{X}, \mathcal{L})$ so that π is $\mathbb{G}_m$-equivariant where $\mathbb{G}_m$ acts on $\mathbb{A}^1$ by multiplication.
- $(\pi^{-1}(1), \mathcal{L}|_{\pi^{-1}(1)}) \cong (X, L)$.

We will write $\pi\colon (\mathcal{X}, \mathcal{L}) \to \mathbb{A}^1$ by $(\mathcal{X}, \mathcal{L})$ for simplicity. In this paper, we only treat test configurations $(\mathcal{X}, \mathcal{L})$ such that $\mathcal{X}$ is normal. A test configuration $(\mathcal{X}, \mathcal{L})$ is *trivial* if $\mathcal{X}$ is $\mathbb{G}_m$-equivariantly isomorphic to $X \times \mathbb{A}^1$ and we denote $\mathcal{X}$ by $X_{\mathbb{A}^1}$ in this case. Let $p\colon X_{\mathbb{A}^1} \to X$ be the canonical projection. It is well-known that for any semiample test configuration $(\mathcal{X}, \mathcal{L})$, there is a normal semiample test configuration $(\mathcal{Y}, \sigma^*\mathcal{L})$ together with two $\mathbb{G}_m$-equivariant morphisms $\sigma\colon \mathcal{Y} \to \mathcal{X}$ and $\rho\colon \mathcal{Y} \to X_{\mathbb{A}^1}$ that are the identity morphisms over $\mathbb{A}^1 \setminus \{0\}$. Let H be an $\mathbb{R}$-line bundle on X and $\mathcal{D}$ be the closure of $\Delta \times \mathbb{G}_m \subset \mathcal{X}$. Then we define the non-Archimedean Mabuchi functional and the non-Archimedean J^H-functional by

$$M_\Delta^{\mathrm{NA}}(\mathcal{X}, \mathcal{L}) := (K_{\overline{\mathcal{X}}/\mathbb{P}^1} + \overline{\mathcal{D}} + \mathcal{X}_{0,\mathrm{red}} - \mathcal{X}_0) \cdot \overline{\mathcal{L}}^d - \frac{d(K_X + \Delta) \cdot L^{d-1}}{(d+1)L^d} \cdot \overline{\mathcal{L}}^{d+1},$$

$$(\mathcal{J}^H)^{\mathrm{NA}}(\mathcal{X}, \mathcal{L}) := (p \circ \rho)^* H \cdot \sigma^* \overline{\mathcal{L}}^d - \frac{dH \cdot L^{d-1}}{(d+1)L^d} \cdot \overline{\mathcal{L}}^{d+1}.$$

Here the bar denotes the canonical compactification; cf [Boucksom et al. 2017, Sections 3 and 7]. It is easy to see that $M_\Delta^{\mathrm{NA}}(\mathcal{X}, \mathcal{L}) = M_\Delta^{\mathrm{NA}}(\mathcal{Y}, \sigma^*\mathcal{L})$ and $(\mathcal{J}^H)^{\mathrm{NA}}(\mathcal{X}, \mathcal{L}) = (\mathcal{J}^H)^{\mathrm{NA}}(\mathcal{Y}, \sigma^*\mathcal{L})$. Hence, the functionals are well-defined. We say that (X, B, L) is *uniformly K-stable* (resp. (X, L) is *uniformly* J^H*-stable*) if there exists a positive constant $\epsilon > 0$ such that

$$M_\Delta^{\mathrm{NA}}(\mathcal{X}, \mathcal{L}) \geq \epsilon (\mathcal{J}^L)^{\mathrm{NA}}(\mathcal{X}, \mathcal{L}) \quad (\text{resp. } (\mathcal{J}^H)^{\mathrm{NA}}(\mathcal{X}, \mathcal{L}) \geq \epsilon (\mathcal{J}^L)^{\mathrm{NA}}(\mathcal{X}, \mathcal{L}))$$

for any normal semiample test configuration.

We note that $(\mathcal{J}^L)^{\mathrm{NA}}(\mathcal{X}, \mathcal{L}) \geq 0$, and $(\mathcal{J}^L)^{\mathrm{NA}}(\mathcal{X}, \mathcal{L}) = 0$ if and only if $(\mathcal{X}, \mathcal{L})$ is trivial for any ample normal test configuration (cf [Boucksom et al. 2017, Proposition 7.8]). In [Boucksom et al. 2017], $(\mathcal{J}^L)^{\mathrm{NA}}$ is introduced and denoted by $I^{\mathrm{NA}} - J^{\mathrm{NA}}$. This coincides with the minimum norm independently introduced in [Dervan 2016].

Definition 2.25 Let (X, Δ, L) be a klt polarized pair. Let r be a positive integer such that rL is a line bundle. For any $m \in \mathbb{Z}_{>0}$, a $\mathbb{Q}$-divisor D_{mr} is called a *mr-basis type divisor* of L if

$$D_{mr} = \frac{1}{mrh^0(X, \mathcal{O}_X(mrL))} \sum_{i=1}^{h^0(X, \mathcal{O}_X(mrL))} E_i$$

such that the E_i form a basis of $H^0(X, \mathcal{O}_X(mrL))$. We define δ_{mr} and δ-invariants as

$$\delta_{mr,(X,\Delta)}(L) := \inf_{D_{mr}} \operatorname{lct}(X, \Delta; D_{mr}) \quad \text{and} \quad \delta_{(X,\Delta)}(L) := \lim_{m\to\infty} \delta_{mr,(X,\Delta)}(L),$$

where D_{mr} runs over all mr-basis type divisors; cf [Fujita and Odaka 2018] and [Blum and Jonsson 2020]. By [Blum and Jonsson 2020], the above limit exists.

For any prime divisor F over X with a projective birational morphism $\pi: Y \to X$ such that F appears as a prime divisor on Y, we define

$$S_L(F) := \frac{1}{L^n} \int_0^\infty \operatorname{vol}(L - tF)\, dt,$$

where $\operatorname{vol}(L - tF)$ denotes $\operatorname{vol}(\pi^* L - tF)$ by abuse of notation. We set

$$S_{mr,L}(F) := \max_{D_{mr}} \operatorname{ord}_F(D_{mr}) = \frac{\sum_{i\geq 1} h^0(Y, \mathcal{O}_Y(mr\pi^* L - iF))}{mrh^0(X, \mathcal{O}_X(mrL))},$$

where D_{mr} runs over all mr-basis type divisors; cf [Fujita and Odaka 2018, Lemma 2.2]. It is well-known [Blum and Jonsson 2020, Lemma 2.9] that

$$\lim_{m\to\infty} S_{mr,L}(F) = S_L(F).$$

Furthermore, we have

$$\delta_{(X,\Delta)}(L) = \inf_F \frac{A_{(X,\Delta)}(F)}{S_L(F)}$$

by [Blum and Jonsson 2020], where F runs over all prime divisors over X.

Definition 2.26 (α-invariant) Let (X, Δ, L) be a klt polarized pair. We define the *α-invariant*, denoted by $\alpha_{(X,\Delta)}(L)$, by

$$\alpha_{(X,\Delta)}(L) := \inf\{\operatorname{lct}(X, \Delta; D) \mid D \in |L|_{\mathbb{Q}}\} = \inf\{\operatorname{lct}(X, \Delta; D) \mid D \in |L|_{\mathbb{R}}\}.$$

This notion was introduced by Tian [1987] (and restated in [Tian 1990]) to obtain a sufficiency condition for the existence of Kähler–Einstein metrics on Fano manifolds.

The following fact is well-known.

Lemma 2.27 (cf [Boucksom et al. 2017, Theorem 9.14; Fujita 2019, Proposition 2.1, Lemma 2.2]) *Let (X, Δ, L) be a d-dimensional klt polarized pair. Then*

$$0 < \frac{d+1}{d}\alpha_{(X,\Delta)}(L) \leq \delta_{(X,\Delta)}(L) \leq (d+1)\alpha_{(X,\Delta)}(L).$$

Example 2.28 When X is a curve, we can easily compute $\delta_{(X,\Delta)}(L)$ as follows: Since every prime divisor over X is a point $P \in X$, we have

$$S_L(P) = \frac{1}{\deg L} \int_0^{\deg L} (\deg L - t)\, dt = \frac{\deg L}{2}.$$

Thus, we have
$$\delta_{(X,\Delta)}(L) = \frac{2}{\deg L} \inf_P A_{(X,\Delta)}(P) = 2\frac{(1-\max_{P\in X} \mathrm{ord}_P(\Delta))}{\deg L}.$$
In this case we have
$$\alpha_{(X,\Delta)}(L) = \frac{1-\max_{P\in X} \mathrm{ord}_P(\Delta)}{\deg L} = \frac{1}{2}\delta_{(X,\Delta)}(L).$$

The following notion from [Hattori 2024a] will also be used in this paper.

Definition 2.29 (special K-stability) We say that a klt polarized pair (X, Δ, L) is *specially K-stable* if $\delta_{(X,\Delta)}(L)L + K_X + \Delta$ is ample and (X, L) is uniformly $\mathrm{J}^{\delta_{(X,\Delta)}(L)L+K_X+\Delta}$-stable.

Note that the special K-stability depends only on the numerical class of L since so do $\delta_{(X,\Delta)}(L)$ and the uniform $\mathrm{J}^{\delta_{(X,\Delta)}(L)L+K_X+\Delta}$-stability.

By the following, we know that the special K-stability implies the uniform K-stability.

Theorem 2.30 [Hattori 2024a, Corollary 3.21] *Let (X, Δ, L) be a klt polarized variety and $(\mathcal{X}, \mathcal{L})$ be a normal semiample test configuration for (X, L). Then,*
$$M_\Delta^{\mathrm{NA}}(\mathcal{X}, \mathcal{L}) \geq (\mathcal{J}^{\delta_{(X,\Delta)}(L)L+K_X+\Delta})^{\mathrm{NA}}(\mathcal{X}, \mathcal{L}).$$

Over $\mathbb{C}$, there exists an intrinsic criterion for J-stability and special K-stability without using test configurations.

Theorem 2.31 *Let (X, L) be a polarized variety over $\mathbb{C}$ of dimension d, and let H be an ample $\mathbb{R}$-line bundle on X. Then (X, L) is uniformly J^H-stable if and only if there exists $\epsilon > 0$ such that*
$$\left(d\frac{H\cdot L^{d-1}}{L^d}L - pH\right)\cdot L^{p-1}\cdot V \geq \epsilon(d-p)L^p\cdot V$$
for any p-dimensional subvariety $V \subset X$ with $0 < p < d$. In particular, if (X, Δ, L) is a polarized klt pair and $H := \delta_{(X,\Delta)}(L)L + K_X + \Delta$ is ample, then the specially K-stability of (X, Δ, L) is equivalent to the existence of $\epsilon > 0$ such that the above inequality holds for any subvariety $V \subset X$.

The above theorem was first proved in the case of Kähler manifolds by Chen [2021], but currently the theorem holds for all polarized varieties by Hattori [2021, Theorem 8.12]. For polarized (resp. Kähler) manifolds, it was shown by Datar and Pingali [2021] (resp. Song [2020]) that uniform J^H-stability is equivalent to a certain weaker condition.

Roughly speaking, the uniform adiabatic K-stability of $f\colon (X, \Delta, A) \to C$ was originally defined to be the uniform K-stability of $(X, \Delta, \epsilon A + L)$ with fixed some ample $\mathbb{Q}$-line bundle L on C for any sufficiently small $\epsilon > 0$. The original definition [Hattori 2022, Definition 2.6] and the ad hoc definition coincide by the following theorems. Furthermore, we do not have to fix L on C by what we stated after Definition 2.29.

Theorem 2.32 [Hattori 2021, Theorem 8.15] *Let (X,Δ,A) be a klt polarized pair over $\mathbb{C}$. If $K_X+\Delta$ is nef, then there exists a real number $C>0$, depending only on the intersection numbers $A^i\cdot(K_X+\Delta)^{\dim X-i}$ for $0\le i\le \dim X$, such that $(X,\epsilon A+K_X+\Delta)$ is uniformly $\mathrm{J}^{K_X+\Delta+C\epsilon(\epsilon A+K_X+\Delta)}$-stable for every $\epsilon>0$. Furthermore, there exist real numbers $\epsilon_0>0$ and $\alpha>0$ such that $(X,\Delta,\epsilon A+K_X+\Delta)$ is specially K-stable and*

$$M_\Delta^{\mathrm{NA}}(\mathcal{X},\mathcal{M})\ge\alpha(\mathcal{J}^{\epsilon A+K_X+\Delta})^{\mathrm{NA}}(\mathcal{X},\mathcal{M})$$

for any $0<\epsilon<\epsilon_0$ and normal semiample test configuration $(\mathcal{X},\mathcal{M})$ for $(X,\epsilon A+K_X+\Delta)$.

Theorem 2.33 [Hattori 2022, Theorem B] *Let $f\colon(X,\Delta,A)\to(\mathbb{P}^1,\mathcal{O}(1))$ be a polarized klt trivial fibration over $\mathbb{C}$ such that $K_X+\Delta\sim_{\mathbb{Q}}uf^*(\mathcal{O}(1))$ for some $u<0$.*

Then, f is uniformly adiabatically K-stable if and only if there exist real numbers $\epsilon_0>0$ and $\alpha>0$ such that $(X,\Delta,\epsilon A-(K_X+\Delta))$ is specially K-stable and

$$M_\Delta^{\mathrm{NA}}(\mathcal{X},\mathcal{M})\ge\alpha(\mathcal{J}^{\epsilon A-(K_X+\Delta)})^{\mathrm{NA}}(\mathcal{X},\mathcal{M})$$

for any $0<\epsilon<\epsilon_0$ and for any normal semiample test configuration $(\mathcal{X},\mathcal{M})$ for $(X,\epsilon A-(K_X+\Delta))$.

We remark that Theorem 2.32 follows from the proof of [Hattori 2021, Theorem 8.15]. On the other hand, the equalities

$$\lim_{\epsilon\to 0}\delta_{(X,\Delta)}(\epsilon A-(K_X+\Delta))=2\frac{(\max_{p\in\mathbb{P}^1}\mathrm{ord}_p(B_{\mathbb{P}^1})-1)}{u}=\delta_{(\mathbb{P}^1,B_{\mathbb{P}^1})}(-K_{\mathbb{P}^1}-B_{\mathbb{P}^1}-M_{\mathbb{P}^1})$$

are key steps to show Theorem 2.33; cf [Hattori 2022, Theorem D]. We will show Theorem 1.6 in Section 6 with them in mind.

3 Boundedness

In this section, we prove results of the boundedness of certain classes.

Let d be a positive integer, v a positive rational number, $u\ne 0$ a rational number, and $\Theta\subset\mathbb{Q}_{\ge 0}$ a DCC set. We set $e:=u/|u|$ (thus $eu=|u|$). We consider the following sets:

$$\mathfrak{F}_{d,\Theta,v}:=\left\{ f\colon(X,\Delta)\to C \;\middle|\; \begin{array}{rl} \text{(i)} & f \text{ is a klt-trivial fibration over a curve } C,\\ \text{(ii)} & \dim X=d,\\ \text{(iii)} & \text{the coefficients of } \Delta \text{ belong to } \Theta,\\ \text{(iv)} & \text{there is an } f\text{-ample } \mathbb{Q}\text{-Cartier Weil divisor } A \text{ on } X \text{ such}\\ & \text{that } \mathrm{vol}(A|_F)=v, \text{ where } F \text{ is a general fiber of } f, \end{array} \right\}$$

$$\mathfrak{G}_{d,\Theta,v,u}:=\left\{ f\colon(X,\Delta)\to C\in\mathfrak{F}_{d,\Theta,v} \;\middle|\; \begin{array}{l} K_X+\Delta\equiv uf^*H \text{ for some Cartier divisor } H \text{ with}\\ \deg H=1. \end{array} \right\}$$

Lemma 3.1 below is crucial for the boundedness and the lemma will be used in Section 5.

Lemma 3.1 *There exists a positive integer r, depending only on d, Θ, v and u, such that for any element $f\colon (X,\Delta)\to C$ of $\mathfrak{G}_{d,\Theta,v,u}$, we have $er(K_X+\Delta)\sim f^*D$ for some very ample Cartier divisor D on C. In particular, $er(K_X+\Delta)$ is a basepoint-free Cartier divisor and the linear system $|er(K_X+\Delta)|$ defines f. Furthermore, there are only finitely many possibilities of $\dim H^0(X,\mathcal{O}_X(er(K_X+\Delta)))$.*

Proof We fix an element $f\colon (X,\Delta)\to C$ of $\mathfrak{G}_{d,\Theta,v,u}$, and we pick a Weil divisor A on X as in (iv) of $\mathfrak{F}_{d,\Theta,v}$.

Let F be a general fiber of f, and pick a Cartier divisor H on C with $\deg H=1$. By applying [Birkar 2023, Corollary 1.4] to $(F,\Delta|_F)$ and $A|_F$, we can find $m\in\mathbb{Z}_{>0}$, depending only on d and Θ, such that $H^0(F,\mathcal{O}_X(mA|_F))\neq 0$. Then there is a sufficiently large positive integer t such that

$$E\sim mA+tf^*H$$

for some effective Weil divisor E on X. By construction, we have $\mathrm{vol}(E|_F)=m^{d-1}v$. By applying [Birkar 2021a, Lemma 7.4] to $(X,\Delta)\to C$ and E, we can find a positive integer q, depending only on d, Θ and v, such that we can write

$$q(K_X+\Delta)\sim qf^*(K_C+B+M),$$

where B (resp. M) is the discriminant part (resp. moduli part) of the canonical bundle formula, such that qM is Cartier. Then we have $\deg(K_C+B+M)\le eu$. By definition of the discriminant part of the canonical bundle formula and the ACC for lc thresholds [Hacon et al. 2014, Theorem 1.1], we see that the coefficients of B belong to a DCC set of $\mathbb{Q}_{>0}$ depending only on d and Θ, which we denote by Ψ. Let q' be the smallest positive integer such that $q'u$ is an integer and q divides q'. Since $\deg K_C\ge -2$ and $\deg M\in(1/q)\mathbb{Z}_{\ge 0}$ by Theorem 2.6, we see that

$$\deg(q'B)\in\{0,1,\dots,eq'u+2q'\}.$$

We define $\delta:=\inf\Psi$, which is a positive rational number because Ψ satisfies the DCC. Since $\deg B\le eu+2$, the number of components of B is not greater than $(eu+2)/\delta$. Thus, all the coefficients of B belong to the set

$$\Psi':=\left\{a_0-\sum_{i=1}^{l}a_i \;\middle|\; a_0\in\frac{1}{q'}\mathbb{Z}\cap[0,eu+2],\ a_i\in\Psi,\ l\le\frac{eu+2}{\delta}\right\}.$$

We can easily check that Ψ' satisfies the ACC because $(1/q')\mathbb{Z}\cap[0,eu+2]$ is a finite set and Ψ satisfies the DCC. Hence $\Psi\cap\Psi'$ satisfies the ACC and the DCC, which implies that $\Psi\cap\Psi'$ is a finite set.

From these facts, we can find q'', depending only on $\Psi\cap\Psi'$, such that $q''B$ is a Weil divisor. By construction, $\Psi\cap\Psi'$ depends only on d, Θ, v and u. Since q divides q' by construction, we have

$$q'q''(K_X+\Delta)\sim q'q''f^*(K_C+B+M),$$

and the right-hand side is Cartier. The integer $q'q''$ depends only on d, Θ, v, and u.

Since $eq'q''(K_C+B+M)$ is ample and Cartier, there is a positive integer r, depending only on d, Θ, v and u, such that r is divided by $q'q''$ and $er(K_C+B+M)$ is very ample. Then this r is the desired positive integer. We put $D := er(K_C+B+M)$. The finiteness of $\dim H^0(X,\mathcal{O}_X(er(K_X+\Delta)))$ follows from

$$\begin{aligned}0 &\le \dim H^0(X,\mathcal{O}_X(er(K_X+\Delta))) = \dim H^0(C,\mathcal{O}_C(D))\\ &= \dim H^1(C,\mathcal{O}_C(D)) + \deg D + \chi(C,\mathcal{O}_C)\\ &= eru + 1 + \big(\dim H^0(C,\mathcal{O}_C(K_C-D)) - \dim H^0(C,\mathcal{O}_C(K_C))\big)\\ &\le eru+1\end{aligned}$$

by the Riemann–Roch theorem. □

Let n be a positive integer. We define

$$\mathfrak{F}_{d,n,v} := \left\{ f\colon (X,\Delta)\to C \;\middle|\; \begin{array}{rl} \text{(i)} & f \text{ is a klt-trivial fibration over a curve } C,\\ \text{(ii)} & \dim X = d,\\ \text{(iii)} & n\Delta \text{ is a Weil divisor},\\ \text{(iv)} & \text{there is an } f\text{-ample } \mathbb{Q}\text{-Cartier Weil divisor } A \text{ on } X \text{ such that}\\ & \operatorname{vol}(A|_F)=v, \text{ where } F \text{ is a general fiber of } f. \end{array} \right\}$$

Then Lemma 3.1 shows the existence of an $n\in\mathbb{Z}_{>0}$ such that $\mathfrak{G}_{d,\Theta,v,u}$ is a subset of the set

$$\mathfrak{G}_{d,n,v,u} := \{ f\colon (X,\Delta)\to C \in \mathfrak{F}_{d,n,v} \mid K_X+\Delta \equiv uf^*H \text{ for some Cartier divisor } H \text{ with } \deg H = 1\}.$$

Moreover, there exists a positive integer r, depending only on d, n, v and u, such that for any element $f\colon (X,\Delta)\to C$ of $\mathfrak{G}_{d,n,v,u}$, the divisor $er(K_X+\Delta)$ is a basepoint-free Cartier divisor and the linear system $|er(K_X+\Delta)|$ defines f.

In the rest of this section, we will deal with $\mathfrak{F}_{d,n,v}$ and $\mathfrak{G}_{d,n,v,u}$ for the fixed $d,n\in\mathbb{Z}_{>0}$, $v\in\mathbb{Q}_{>0}$ and $u\in\mathbb{Q}_{\neq 0}$.

The following lemma gives a lower bound of the α-invariants for general fibers of the elements of $\mathfrak{F}_{d,n,v}$.

Lemma 3.2 *There exists a positive integer N, depending only on d, n and v, such that for any element $f\colon (X,\Delta)\to C$ of $\mathfrak{F}_{d,n,v}$ and any $\mathbb{Q}$-Cartier Weil divisor A on X as in* (iv) *of $\mathfrak{F}_{d,n,v}$, we have the inequality*

$$\alpha_{(F,\Delta|_F)}(A|_F) \ge \frac{1}{N},$$

where F is a general fiber of f.

Proof We fix $f\colon (X,\Delta)\to C\in\mathfrak{F}_{d,n,v}$ and A as in the final condition of $\mathfrak{F}_{d,n,v}$. Let F be a general fiber of f. We put $\Delta_F = \Delta|_F$.

By applying [Birkar 2023, Corollary 1.4] to (F, Δ_F) and $A|_F$, we can find an $m \in \mathbb{Z}_{>0}$, depending only on d and n, such that

$$mA|_F \sim E_F$$

for some effective Weil divisor E_F on F. Then we have $\mathrm{vol}(E_F) = m^{d-1}v$. By applying [Birkar 2023, Corollary 1.6] to (F, Δ_F) and E_F, we see that $(F, \mathrm{Supp}(\Delta_F + E_F))$ belongs to a bounded family depending only on d, n, v, and m. By [Birkar 2019, Lemma 2.24], there exists $m' \in \mathbb{Z}_{>0}$, depending only on d, n, v and m, such that $m'E_F \sim m'mA|_F$ is Cartier. Then $m'm$ depends only on d, n and v.

Because the divisor $m'mA|_F - (K_F + \Delta_F)$ is ample, we may apply the effective basepoint-freeness [Kollár 1993, Theorem 1.1] and the effective very ampleness [Fujino 2017, Lemma 7.1]. Hence, there exists $m'' \in \mathbb{Z}_{>0}$, depending only on d, n and v, such that $m''A|_F$ is very ample. Taking a small $\mathbb{Q}$-factorialization of X and applying the length of extremal rays, we see that $K_F + 3dm''A|_F$ is the pushdown of a big divisor; cf [Birkar 2019, Lemma 2.46]. Because we have

$$3dm''A|_F - \Delta_F \sim_{\mathbb{Q}} 3dm''A|_F + K_F,$$

we see that $3dm''A|_F - \Delta_F$ is the pushdown of a big divisor. We also have

$$\mathrm{vol}(3dm''A|_F) = (3dm'')^{d-1}v.$$

By the ACC for numerically trivial pairs [Hacon et al. 2014, Theorem 1.5], we can find a positive real number $\delta > 0$, depending only on $d-1$ and n, such that (F, Δ_F) is δ-lc.

From the above discussion, we may apply [Birkar 2021b, Theorem 1.8] to (F, Δ_F) and $A|_F$, and we can find $\epsilon \in \mathbb{R}_{>0}$, depending only on $d-1$, δ and $(3dm'')^{d-1}v$, such that

$$\alpha_{(F,\Delta|_F)}(A|_F) \geq \epsilon.$$

Construction of m'' and δ implies that ϵ depends only on d, n and v. Finally, we define N to be the minimum positive integer satisfying $\epsilon > 1/N$. Then N satisfies the condition of Lemma 3.2. □

Definition 3.3 Let N be the positive integer in Lemma 3.2. We define

$$\alpha := d(4N + \lceil euN \rceil)v.$$

Note that α depends only on d, n, v and u.

The following result is a crucial step for the boundedness and a special form of Theorem 1.5(2).

Proposition 3.4 *For any element $f: (X, \Delta) \to C$ of $\mathfrak{G}_{d,n,v,u}$ and any $\mathbb{Q}$-Cartier Weil divisor A on X as in* (iv) *of $\mathfrak{F}_{d,n,v}$, there exists a Cartier divisor D on C such that $A + f^*D$ is ample and $\mathrm{vol}(A + f^*D) \leq \alpha$.*

Proof We fix an element $f: (X, \Delta) \to C$ of $\mathfrak{G}_{d,n,v,u}$, and we pick a Weil divisor A on X as in (iv) of $\mathfrak{F}_{d,n,v}$. Let H be a Cartier divisor on C such that $\deg H = 1$.

We define τ to be the smallest integer such that $A+\tau f^*H$ is big. Note that τ is well-defined since H is not numerically trivial and $A+tf^*H$ is ample for all $t\gg 0$. We fix an effective $\mathbb{Q}$-divisor

$$A'\sim_{\mathbb{Q}} A+\tau f^*H.$$

Let $N\in\mathbb{Z}_{>0}$ be as in Lemma 3.2. By the property of N in Lemma 3.2, there is a nonempty open subset $U\subset C$ such that $(X,\Delta+(1/N)A')$ is lc on $f^{-1}(U)$.

Because A' is ample over C and $K_X+\Delta\sim_{\mathbb{Q},C}0$, we can find a positive integer l such that

$$\Phi:=l\Big(K_X+\Delta-f^*K_C+\frac{1}{N}A'\Big)$$

is Cartier and basepoint-free over C and Φ defines an embedding into $\mathbb{P}_C(f_*\mathcal{O}_X(\Phi))$. By applying [Fujino 2018, Theorem 1.11] to $f\colon X\to C$ and Φ, we see that $f_*\mathcal{O}_X(\Phi)$ is nef. In other words, the Cartier divisor corresponding to $\mathcal{O}_{\mathbb{P}_C(f_*\mathcal{O}_X(\Phi))}(1)$ is nef. Because $\mathcal{O}_X(\Phi)$ coincides with the pullback of $\mathcal{O}_{\mathbb{P}_C(f_*\mathcal{O}_X(\Phi))}(1)$ to X, it follows that Φ is nef.

Since $f\colon(X,\Delta)\to C$ is an element of $\mathfrak{G}_{d,n,v,u}$, it follows that the divisor

$$euf^*H-(K_X+\Delta)$$

is nef. Since $\deg H=1$, we see that $2H+K_C$ is nef. Therefore, the divisor

$$\begin{aligned}\Big(2+eu+\frac{\tau}{N}\Big)f^*H+\frac{1}{N}A &\sim_{\mathbb{Q}}(2+eu)f^*H+\frac{1}{N}A'\\ &=\Big(K_X+\Delta-f^*K_C+\frac{1}{N}A'\Big)-(K_X+\Delta)+f^*K_C+(2+eu)f^*H\\ &=\frac{1}{l}\Phi+f^*(2H+K_C)+(euf^*H-(K_X+\Delta))\end{aligned}$$

is nef. Thus, $A+(N(2+eu)+\tau)f^*H$ is nef. Since A is f-ample, we see that

$$A+(3N+euN+\tau)f^*H=A+(N(2+eu)+\tau)f^*H+Nf^*H$$

is ample.

By definition of τ, the divisor $A+(\tau-1)f^*H$ is not big. Hence, we have

$$\operatorname{vol}(A+(\tau-N)f^*H)=0.$$

Since $K_X+\Delta\equiv uf^*H$, by the canonical bundle formula, we have $eu\ge\deg K_C$. Then $(4N+\lceil euN\rceil)H$ is very ample.

We put $A'':=A+(\tau-N)f^*H$ and $N':=4N+\lceil euN\rceil$. Let $G\in|N'f^*H|$ be a member consisting of N' general fibers. Then $\operatorname{vol}(A'')=0$ and $A''|_G=A|_G$ by definition. For each $m\in\mathbb{Z}_{>0}$ and $0<k\le m$, we consider the exact sequence

$$0\to H^0(X,\mathcal{O}_X(mA''+(k-1)N'f^*H))\to H^0(X,\mathcal{O}_X(mA''+kN'f^*H))\to H^0(G,\mathcal{O}_G(mA|_G))$$

induced by

$$0 \to \mathcal{O}_X(mA'' + kN' f^* H - G) \to \mathcal{O}_X(mA'' + kN' f^* H) \to \mathcal{O}_G(mA|_G) \to 0.$$

By similar arguments to [Jiang 2018, Proof of Lemma 2.5], we have

$$\dim H^0(X, \mathcal{O}_X(mA'' + mN' f^* H)) \le \dim H^0(X, \mathcal{O}_X(mA'')) + m \cdot \dim H^0(G, \mathcal{O}_G(mA|_G)).$$

Since G consists of N' general fibers, taking the limit $m \to \infty$ we have

$$\mathrm{vol}(A'' + N' f^* H) \le \mathrm{vol}(A'') + dN' \cdot \mathrm{vol}(A|_F) = 0 + dN' v.$$

Here, we used that $\mathrm{vol}(A'') = 0$ and $\mathrm{vol}(A|_F) = v$. We put

$$D = (3N + \lceil euN \rceil + \tau)H.$$

By recalling the definitions of A'', N' and α (see Definition 3.3), we obtain

$$\mathrm{vol}(A + f^* D) = \mathrm{vol}(A'' + N' f^* H) \le dN' v = \alpha.$$

Thus D is the desired Cartier divisor on C. □

Definition 3.5 Let α be the positive real number in Definition 3.3. For any element $f\colon (X, \Delta) \to C$ of $\mathfrak{G}_{d,n,v,u}$ and any $\mathbb{Q}$-Cartier Weil divisor A on X as in (iv) of $\mathfrak{F}_{d,n,v}$, we pick a Cartier divisor H on C with $\deg H = 1$ and we define

$$m_{(f,A)} := \max\{m \in \mathbb{Z} \mid A + mf^* H \text{ is ample, } \mathrm{vol}(A + mf^* H) \le \alpha\}.$$

Note that $m_{(f,A)}$ is well-defined by Proposition 3.4. We may have $m_{(f,A)} \le 0$. We define

$$L_{(f,A)} := A + m_{(f,A)} f^* H.$$

Now we are ready to prove the boundedness.

Theorem 3.6 (boundedness) *The set of klt pairs (X, Δ) appearing in $\mathfrak{G}_{d,n,v,u}$ is log bounded. Furthermore, there exists a positive integer I_0, depending only on d, n, v and u, such that $I_0 L_{(f,A)}$ is an ample Cartier divisor on X. In particular, $I_0 A$ is Cartier.*

Proof We pick $f\colon (X, \Delta) \to C \in \mathfrak{G}_{d,n,v,u}$ and a $\mathbb{Q}$-Cartier Weil divisor A on X as in (iv) of $\mathfrak{F}_{d,n,v}$. By Lemma 3.1, we can find a positive integer r, depending only on d, n, v and u, such that (X, Δ) is $(1/r)$-lc. By [Birkar et al. 2010, Corollary 1.4.3], there is a small $\mathbb{Q}$-factorization $\phi\colon X' \to X$ of X. Then $(X', 0)$ is an $(1/r)$-lc pair and $3d\phi^* L_{(f,A)} - K_{X'}$ is big. By applying [Birkar 2023, Theorem 1.1] to X' and $3d\phi^* L_{(f,A)}$, we can find a positive integer m, depending only on d, n, v and u, such that

$$H^0(X', \mathcal{O}_{X'}(m\phi^* L_{(f,A)})) \ne 0.$$

Thus, we can find an effective Weil divisor $E \sim mL_{(f,A)}$. We have

$$\mathrm{vol}(E) \le m^d \alpha.$$

For each $f\colon (X,\Delta)\to C\in\mathfrak{G}_{d,n,v,u}$, we fix $E\sim mL_{(f,A)}$ as above, and we prove that the set of $(X,\mathrm{Supp}(\Delta+E))$ is bounded. If $u<0$, then we have $eu\le 2$ by the canonical bundle formula; see cf Theorem 2.6. By taking a reduced divisor G on X consisting of three general fibers of f, we get a klt Calabi–Yau pair $(X,\Delta+\frac{1}{3}euG)$. By applying [Birkar 2023, Corollary 1.6] to $(X,\Delta+\frac{1}{3}euG)$ and E, we see that the set of such couples $(X,\mathrm{Supp}(\Delta+G+E))$ is bounded. In particular, the set of $(X,\mathrm{Supp}(\Delta+E))$ for some $f\colon (X,\Delta)\to C\in\mathfrak{G}_{d,n,v,u}$ is bounded. If $u>0$, then we pick a Cartier divisor H on C such that $\deg H=1$ and $K_X+\Delta\equiv uf^*H$. Since the volume depends only on the numerical class, we have

$$\begin{aligned}\mathrm{vol}(K_X+\Delta+E)=(K_X+\Delta+E)^d&=d(uf^*H\cdot E^{d-1})+(E^d)\\&=duf^*H\cdot(mL_{(f,A)})^{d-1}+\mathrm{vol}(E)\\&\le dum^{d-1}v+m^d\alpha.\end{aligned}$$

Since (X,Δ) is $(1/r)$-lc by Lemma 3.1 and $n\Delta$ is a Weil divisor by definition, we may apply [Birkar 2023, Theorem 1.5] to (X,Δ) and E, and the set of $(X,\mathrm{Supp}(\Delta+E))$ is bounded. By these arguments, we obtain the boundedness of the set of $(X,\mathrm{Supp}(\Delta+E))$.

The first statement of Theorem 3.6 immediately follows from the above discussion. Moreover, [Birkar 2019, Lemma 2.24] implies the existence of a positive integer I', depending only on d, n, v and u, such that $I'E$ is Cartier. Set $I_0:=I'm$. Then I_0 depends only on d, n, v and u, and $I_0L_{(f,A)}$ is Cartier. □

Remark 3.7 We define

$$\begin{aligned}\mathfrak{G}_{d,n,v,0}&:=\{f\colon (X,\Delta)\to C\in\mathfrak{F}_{d,n,v}\mid K_X+\Delta\equiv 0\},\\ \mathfrak{V}_{d,n,v}&:=\{(X,\Delta)\mid (X,\Delta)\text{ is a klt pair and } f\colon (X,\Delta)\to C\in\mathfrak{G}_{d,n,v,0}\text{ for some } f\}.\end{aligned}$$

Then the same argument as in this section implies that $\mathfrak{V}_{d,n,v}$ is log bounded. Indeed, applying the argument in Lemma 3.1, we can find a positive integer r, depending only on d, n and v, such that $r(K_X+\Delta)\sim 0$. We define $\alpha:=4dNv$; see Definition 3.3. By the same argument as in Proposition 3.4, for any element $f\colon (X,\Delta)\to C$ of $\mathfrak{G}_{d,n,v,0}$ and any $\mathbb{Q}$-Cartier Weil divisor A on X as in (iv) of $\mathfrak{F}_{d,n,v}$, there exists a Cartier divisor D on C such that $A+f^*D$ is ample and $\mathrm{vol}(A+f^*D)\le\alpha$. Then we can define the line bundle $L_{(f,A)}$ as in Definition 3.5. Then the argument in Theorem 3.6 implies that $\mathfrak{V}_{d,n,v}$ is log bounded. Moreover, there exists a positive integer I_0, depending only on d, n and v, such that $I_0L_{(f,A)}$ is an ample Cartier divisor; see Theorem 3.6.

Proof of Theorem 1.5 By Lemma 3.1, we may assume that $\Theta=(1/n)\mathbb{Z}\cap[0,1]$ for some $n\in\mathbb{Z}_{>0}$. Then the assertion immediately follows from Theorem 3.6, Remark 3.7, and the existence of $L_{(f,A)}$ as in Definition 3.5. □

We make use of the following result to construct moduli spaces in Section 5.

Corollary 3.8 *Fix $d \in \mathbb{Z}_{>0}$, a DCC subset $\Theta \subset \mathbb{Q} \cap [0,1]$ and rational numbers $u, v \in \mathbb{Q}$, where $v > 0$. For any $w \in \mathbb{Q}_{>0}$, consider the set*

$$\mathfrak{G}_{d,\Theta,v,u,w} := \left\{ f\colon (X,\Delta,A) \to C \;\middle|\; \begin{array}{rl} \text{(i)} & f \text{ is a klt-trivial fibration over a curve } C \text{ such that} \\ & K_X + \Delta \equiv u f^* H \text{ with a line bundle } H \text{ on } C \text{ of} \\ & \deg H = 1, \\ \text{(ii)} & \dim X = d, \\ \text{(iii)} & \text{the coefficients of } \Delta \text{ belong to } \Theta, \\ \text{(iv)} & A \text{ is an ample } \mathbb{Q}\text{-Cartier Weil divisor on } X \text{ such} \\ & \text{that } (H \cdot A^{d-1}) = v \text{ and } \mathrm{vol}(A) \le w. \end{array} \right\}$$

We also fix $w' \in \mathbb{Q}_{>0}$. Then, there exist

- *a positive integer I, depending only on d, Θ, u, v and w, and*
- *finitely many polynomials $P_1, \dots, P_l$, depending only on d, Θ, u, v, w and w',*

satisfying the following. For any $f\colon (X,\Delta,A) \to C \in \mathfrak{G}_{d,\Theta,v,u,w}$ and nef Cartier divisor M on X,

- *$IA + M$ is very ample,*
- *$H^j(X, \mathcal{O}_X(m(IA+M))) = 0$ for every $j > 0$ and $m \in \mathbb{Z}_{>0}$, and*
- *if $\mathrm{vol}(IA+M) \le w'$, then there is $1 \le i \le l$ such that*

$$\chi(X, \mathcal{O}_X(m(IA+M))) = P_i(m) \quad \text{for every } m \in \mathbb{Z}_{>0}.$$

Before the proof, we show the following criterion for very ampleness and finiteness of Hilbert polynomials.

Lemma 3.9 *Fix $d \in \mathbb{Z}_{>0}$ and $w \in \mathbb{R}_{>0}$. Then there are finitely many polynomials $P_1, \dots, P_l$, depending only on d and w, such that for any d-dimensional projective klt pair (X,Δ), very ample Cartier divisor A on X, and nef Cartier divisor M on X, if $A - (K_X + \Delta)$ is nef and big and $\mathrm{vol}((d+2)A + M) \le w$, then*

- *$(d+2)A + M$ is very ample,*
- *$H^j(X, \mathcal{O}_X(m((d+2)A+M))) = 0$ for every $j > 0$ and $m \in \mathbb{Z}_{>0}$, and*
- *there is a $1 \le i \le l$ such that*

$$\chi(X, \mathcal{O}_X(m((d+2)A+M))) = P_i(m) \quad \text{for every } m \in \mathbb{Z}_{>0}.$$

Proof Put $A' = (d+2)A + M$. By the Kawamata–Viehweg vanishing theorem, we have

$$H^j(X, \mathcal{O}_X(mA' - kA)) = 0$$

for every $m \in \mathbb{Z}_{>0}$, $0 \le k \le d+1$ and $j > 0$. By [Fantechi et al. 2005, Lemma 5.1], $A' - A$ is globally generated, and

$$\chi(X, \mathcal{O}_X(mA')) = \dim H^0(X, \mathcal{O}_X(mA')) \quad \text{for every } m \in \mathbb{Z}_{>0}.$$

Since A is very ample, so is $A' = (A' - A) + A$. Furthermore, we can check the following claim:

Claim *For every $m \in \mathbb{Z}_{>0}$, we have*

$$\dim H^0(X, \mathcal{O}_X(mA')) \le d + m^d w.$$

Proof Let $Y \sim mA'$ be a general hyperplane section. Then we have

$$0 \to \mathcal{O}_X \to \mathcal{O}_X(mA') \to \mathcal{O}_Y(mA'|_Y) \to 0,$$

hence we see that

$$H^0(X, \mathcal{O}_X(mA')) \le 1 + H^0(Y, \mathcal{O}_Y(mA'|_Y))).$$

This relation implies the case of $\dim X = 1$ in the claim, and the general case follows from the relation and induction on the dimension of X. □

Thus, if we put $P(m) = \chi(X, \mathcal{O}_X(mA'))$, there are only finitely many possibilities of $P(1), \dots, P(d+1)$ depending only on d and w. In particular, they do not depend on M. Lemma 3.9 follows from this fact. □

Proof of Corollary 3.8 By Lemma 3.1, we can find r, depending only on d, Θ, v and u, such that $\Theta = [0,1] \cap (1/r)\mathbb{Z}$ and $r(K_X + \Delta)$ is Cartier for all $f\colon (X, \Delta, A) \to C \in \mathfrak{G}_{d,\Theta,v,u,w}$. By Theorem 3.6 and Remark 3.7, there exists I_0, depending only on d, Θ, v and u, such that $I_0 A$ is Cartier for all $f\colon (X, \Delta, A) \to C \in \mathfrak{G}_{d,\Theta,v,u,w}$. Note that $3dI_0A + K_X + \Delta$ is ample by [Kollár and Mori 1998, Theorem 3.7]. Set

$$A' := 3drI_0A + r(K_X + \Delta) \quad \text{and} \quad A'' := A' + 3drI_0A.$$

Then A', $A' - (K_X + \Delta)$, A'', $A'' - (K_X + \Delta)$ are ample. By the effective basepoint-freeness [Kollár 1993, Theorem 1.1] and the effective very ampleness [Fujino 2017, Lemma 7.1], there exists $I_1 \in \mathbb{Z}_{>0}$, depending only on d, such that I_1A' and I_1A'' are very ample. Now

$$\begin{aligned}\operatorname{vol}((d+2)I_1(A' + A'')) &= ((d+2)I_1)^d \operatorname{vol}(A' + A'') \\ &\le ((d+2)I_1)^d \big((6drI_0)^d w + (6drI_0)^{d-1} drvu\big),\end{aligned}$$

hence Lemma 3.9 implies that there are only finitely many possibilities of the Hilbert polynomials

$$m \mapsto \chi\big(X, \mathcal{O}_X\big(m((d+2)I_1(A' + A''))\big)\big)$$

for the elements $f\colon (X, \Delta, A) \to C \in \mathfrak{G}_{d,\Theta,v,u,w}$. Similarly, Lemma 3.9 implies that there are only finitely many possibilities of the Hilbert polynomials

$$m \mapsto \chi\big(X, \mathcal{O}_X\big(m((d+2)I_1A')\big)\big) \quad \text{and} \quad m \mapsto \chi\big(X, \mathcal{O}_X\big(m((d+2)I_1A'')\big)\big)$$

for the elements $f\colon (X, \Delta, A) \to C \in \mathfrak{G}_{d,\Theta,v,u,w}$. In particular, there exist positive integers N_1 and N_2, depending only on d, Θ, u, v and w, such that

$$\dim H^0(X, \mathcal{O}_X((d+2)I_1A')) \le N_1 \quad \text{and} \quad \dim H^0(X, \mathcal{O}_X((d+2)I_1A'')) \le N_2$$

for every $f\colon (X,\Delta,A)\to C\in\mathfrak{G}_{d,\Theta,v,u,w}$. From this fact, there exists a closed immersion $\varphi\colon X\hookrightarrow\mathbb{P}^{N_1}\times\mathbb{P}^{N_2}$ such that

$$\varphi^* p_1^*\mathcal{O}_{\mathbb{P}^{N_1}}(1)\cong\mathcal{O}_X((d+2)I_1A') \quad\text{and}\quad \varphi^* p_2^*\mathcal{O}_{\mathbb{P}^{N_2}}(1)\cong\mathcal{O}_X((d+2)I_1A''),$$

where $p_1\colon\mathbb{P}^{N_1}\times\mathbb{P}^{N_2}\to\mathbb{P}^{N_1}$ and $p_2\colon\mathbb{P}^{N_1}\times\mathbb{P}^{N_2}\to\mathbb{P}^{N_2}$ are the projections. By the theory of Hilbert schemes, there exist a scheme S of finite type over $\Bbbk$ and a closed subscheme $\mathcal{X}\subset\mathbb{P}^{N_1}\times\mathbb{P}^{N_2}\times S$, which is flat over S, such that for any $f\colon (X,\Delta,A)\to C\in\mathfrak{G}_{d,\Theta,v,u,w}$, there is a closed point $s\in S$ satisfying $X\cong\mathcal{X}_s$ and the condition that the immersion $\mathcal{X}_s\subset\mathbb{P}^{N_1}\times\mathbb{P}^{N_2}$ coincides with φ.

By the definitions of A' and A'' and shrinking S if necessary, we may assume that

$$\bigl(p_1^*\mathcal{O}_{\mathbb{P}^{N_1}}(-1)\otimes p_2^*\mathcal{O}_{\mathbb{P}^{N_2}}(1)\bigr)\big|_{\mathcal{X}}$$

is ample over S; see [Kollár and Mori 1998, Corollary 1.41]. Then there exists a positive integer N', depending only on d, Θ, u, v and w, such that the line bundle $\bigl(p_1^*\mathcal{O}_{\mathbb{P}^{N_1}}(-N'-2)\otimes p_2^*\mathcal{O}_{\mathbb{P}^{N_2}}(N'+1)\bigr)\big|_{\mathcal{X}}$ is ample over S. This fact and the definitions of A' and A'' imply that

$$\begin{aligned}-(N'+2)(d+2)I_1A'+(N'+1)(d+2)I_1A' &= (d+2)I_1(-A'+(N'+1)\cdot 3drI_0A)\\ &= (d+2)I_1(3drI_0N'A-r(K_X+\Delta))\\ &= (d+2)I_1r(3dI_0N'A-(K_X+\Delta))\end{aligned}$$

is ample for all $f\colon (X,\Delta,A)\to C\in\mathfrak{G}_{d,\Theta,v,u,w}$, therefore $3dI_0N'A-(K_X+\Delta)$ is ample.

Recall from the definition of I_0 (see Theorem 3.6 and Remark 3.7) that I_0A is Cartier. By the effective basepoint-freeness [Kollár 1993, Theorem 1.1] and the effective very ampleness [Fujino 2017, Lemma 7.1], there exists $I_2\in\mathbb{Z}_{>0}$, depending only on d, such that I_2A is very ample for every $f\colon (X,\Delta,A)\to C\in\mathfrak{G}_{d,\Theta,v,u,w}$. Now define

$$I := (d+2)3dI_0N'I_2,$$

which depends only on d, Θ, u, v and w. Then $(1/(d+2))IA$ is a very ample Cartier divisor and $(1/(d+2))IA-(K_X+\Delta)$ is ample for all $f\colon (X,\Delta,A)\to C\in\mathfrak{G}_{d,\Theta,v,u,w}$. By Lemma 3.9, we see that this I is the desired positive integer. □

4 Tools for construction of the moduli spaces

In this section we prove some results to construct the moduli spaces in this paper.

4.1 Openness

In this subsection, we prove the openness of uniformly adiabatically K-stable klt-trivial fibrations.

Lemma 4.1 *Let $\mathcal{X} \to S$ and $\mathcal{Z} \to S$ be flat projective surjective morphisms of normal varieties such that all the geometric fibers of the morphisms are normal and connected. Let $f\colon \mathcal{X} \to \mathcal{Z}$ be a contraction over S. Let $(\mathcal{X}, \mathcal{D})$ be a pair such that $K_{\mathcal{X}} + \mathcal{D} \sim_{\mathbb{Q},\mathcal{Z}} 0$, $\operatorname{Supp}\mathcal{D}$ does not contain any fiber of $\mathcal{X} \to S$, and $(\mathcal{X}_{\bar{s}}, \mathcal{D}_{\bar{s}})$ is a klt pair for every geometric point $\bar{s} \in S$. Let $\bar{\eta} \in S$ be the geometric generic point.*

Then there exists an open subset $U \subset S$ such that for every closed point $t \in U$, the discriminant $\mathbb{Q}$-divisors $\mathcal{B}$, $B_{\bar{\eta}}$ and B_t with respect to $f\colon (\mathcal{X}, \mathcal{D}) \to \mathcal{Z}$, $f_{\bar{\eta}}\colon (\mathcal{X}_{\bar{\eta}}, \mathcal{D}_{\bar{\eta}}) \to \mathcal{Z}_{\bar{\eta}}$ and $f_t\colon (\mathcal{X}_t, \mathcal{D}_t) \to \mathcal{Z}_t$ respectively satisfy

$$\max_{\bar{P}} \operatorname{coeff}_{\bar{P}}(B_{\bar{\eta}}) = \max_{P} \operatorname{coeff}_{P}(\mathcal{B}|_{\mathcal{Z}\times_S U}) = \max_{P'} \operatorname{coeff}_{P'}(B_t),$$

where $\bar{P}$ (resp. P, P') runs over prime divisors on $\mathcal{Z}_{\bar{\eta}}$ (resp. $\mathcal{Z} \times_S U$, $\mathcal{Z}_t$). Furthermore, $\mathcal{B}|_{\mathcal{Z}_t}$ is well-defined and $\mathcal{B}|_{\mathcal{Z}_t} = B_t$ for any closed point $t \in U$.

Proof First, note that we may shrink S whenever we focus on an open subset of S. Moreover, as in [Ambro 2004, Lemma 5.1], we see that $\max_P \operatorname{coeff}_P(\mathcal{B}|_{\mathcal{Z}\times_S U})$ is not changed for any $U \subset S$ even if we replace $(\mathcal{X}, \mathcal{D}) \to \mathcal{Z} \to S$ with the base change by any étale surjective morphism $S' \to S$. Thus, in the rest of the proof, we will freely shrink S and take the base change of $(\mathcal{X}, \mathcal{D}) \to \mathcal{Z} \to S$ by an étale surjective morphism if necessary.

By shrinking S, we may assume that S is smooth, $\operatorname{Supp}\mathcal{B}$ does not contain any fiber of $\mathcal{Z} \to S$ and the codimension of $\operatorname{Sing}(\mathcal{Z}) \cap \mathcal{Z}_s$ in $\mathcal{Z}_s$ is at least two for every $s \in S$. In particular, we can define $\mathcal{B}_{\bar{\eta}}$, and we can also define $\mathcal{B}_s$ for every closed point $s \in S$.

In this paragraph, we show the first equality of Lemma 4.1. We denote the morphism $\bar{\eta} \to S$ by τ. By shrinking S, we can find a finite morphism $\varphi\colon S' \to S$ and a morphism $\psi\colon \bar{\eta} \to S'$ such that $\tau = \varphi \circ \psi$ and for any component $\bar{Q}$ of $B_{\bar{\eta}}$, there is a prime divisor Q' on $\mathcal{Z} \times_S S'$ whose pullback to $\mathcal{Z}_{\bar{\eta}}$ is $\bar{Q}$. By shrinking S, we may assume that φ is étale. By replacing $(\mathcal{X}, \mathcal{D}) \to \mathcal{Z} \to S$ with the base change by φ, we may assume that for any component $\bar{Q}$ of $B_{\bar{\eta}}$, there is a prime divisor Q on S such that $\tau^*Q = \bar{Q}$. Let $B_{\bar{\eta}}$ and $\mathcal{B}$ be $\mathbb{Q}$-divisors as in Lemma 4.1. By shrinking S and replacing $(\mathcal{X}, \mathcal{D}) \to \mathcal{Z} \to S$ with an étale base change, we may assume $\tau^*(K_{\mathcal{Z}} + \mathcal{B}) = K_{\mathcal{Z}_{\bar{\eta}}} + B_{\bar{\eta}}$. Then $\mathcal{B}_{\bar{\eta}} = B_{\bar{\eta}}$. Shrinking S, we may further assume that any component of $\mathcal{B}$ dominates S. Then

$$\max_{P} \operatorname{coeff}_{P}(\mathcal{B}) = \max_{\bar{P}} \operatorname{coeff}_{\bar{P}}(B_{\bar{\eta}}),$$

where P (resp. $\bar{P}$) runs over prime divisors on $\mathcal{Z}$ (resp. $\mathcal{Z}_{\bar{\eta}}$).

From now on, we show the second equality of Lemma 4.1. We construct a diagram of projective morphisms

$$\begin{array}{ccc} \mathcal{Y} & \xrightarrow{g} & \mathcal{X} \\ {\scriptstyle f'}\downarrow & & \downarrow{\scriptstyle f} \\ \mathcal{W} & \xrightarrow{h} & \mathcal{Z} \end{array}$$

where $\mathcal{Y}$ and $\mathcal{W}$ are smooth varieties, and snc divisors Σ on $\mathcal{W}$ and Ξ on $\mathcal{Y}$ such that

- h is birational and g is a log resolution of $(\mathcal{X}, \mathcal{D})$,

- f' is a contraction,
- $\Xi \supset f'^{*}\Sigma \cup g_*^{-1}\mathcal{D} \cup \mathrm{Ex}(g)$ and the vertical part of Ξ with respect to f' maps into Σ, and
- $(\mathcal{Y}, \Xi)$ is log smooth over $\mathcal{W} \setminus \Sigma$, in other words, the restriction of $f' \colon (\mathcal{Y}, \Xi) \to \mathcal{W}$ over $\mathcal{W} \setminus \Sigma$ is log smooth.

By shrinking S, we may assume that for every closed point $t \in S$, the restricted diagram

$$\begin{tikzcd} (\mathcal{Y}_t, \Xi_t) \arrow[r, "g_t"] \arrow[d, "f'_t"'] & (\mathcal{X}_t, \mathcal{D}_t) \arrow[d, "f_t"] \\ (\mathcal{W}_t, \Sigma_t) \arrow[r, "h_t"] & \mathcal{Z}_t \end{tikzcd}$$

satisfies the same conditions as stated above. We define $\mathcal{D}_{\mathcal{Y}}$ by $K_{\mathcal{Y}} + \mathcal{D}_{\mathcal{Y}} = g^*(K_{\mathcal{X}} + \mathcal{D})$ and $g_*\mathcal{D}_{\mathcal{Y}} = \mathcal{D}$. Let Γ be the discriminant $\mathbb{Q}$-divisor with respect to $f' \colon (\mathcal{Y}, \mathcal{D}_{\mathcal{Y}}) \to \mathcal{W}$. For each closed point $t \in S$, let G_t be the discriminant $\mathbb{Q}$-divisor with respect to $f'_t \colon (\mathcal{Y}_t, \mathcal{D}_{\mathcal{Y}_t}) \to \mathcal{W}_t$. Then $h_*\Gamma = \mathcal{B}$ and $h_{t*}G_t = B_t$, where B_t is the discriminant $\mathbb{Q}$-divisor with respect to the klt-trivial fibration $f_t \colon (\mathcal{X}_t, \mathcal{D}_t) \to \mathcal{Z}_t$. Shrinking S, we may assume $\Gamma_t = G_t$ for every closed point $t \in S$. Then $\mathcal{B}_t = h_{t*}\Gamma_t = h_{t*}G_t = B_t$, and the snc condition of Σ_t implies that

$$\max_{P} \operatorname{coeff}_P(\mathcal{B}) \max_{P'} \operatorname{coeff}_{P'}(B_t),$$

where P (resp. P') runs over prime divisors on $\mathcal{Z}$ (resp. $\mathcal{Z}_t$).

By the above discussion, Lemma 4.1 holds. □

Theorem 4.2 (openness of uniform adiabatic K-stability) *Let S be a normal variety, $\pi \colon (\mathcal{X}, \mathcal{D}) \to S$ a log $\mathbb{Q}$-Gorenstein family, and let $f \colon \mathcal{X} \to \mathcal{P}$ be a contraction over S, where $\mathcal{P}$ is a scheme that is projective and smooth over S. Let $\mathcal{H}$ be an f-ample $\mathbb{Q}$-divisor on $\mathcal{X}$, and let $\mathcal{L}$ be a Cartier divisor on $\mathcal{P}$. Suppose that there exists an integer $m > 0$ such that $(\mathcal{P}_{\bar{s}}, \mathcal{L}_{\bar{s}}) = (\mathbb{P}^1, \mathcal{O}(m))$ for any geometric point $\bar{s} \in S$. Assume that $-(K_{\mathcal{X}/S} + \mathcal{D}) \sim_{\mathbb{Q},S} (u/m) f^*\mathcal{L}$ for some $u \in \mathbb{Q}_{>0}$ and all the geometric fibers of π are klt.*

Then the function

$$h \colon S \ni s \mapsto \max_{P_{\bar{s}}} \operatorname{coeff}_{P_{\bar{s}}}(B_{\bar{s}}),$$

where $\overline{P}_{\bar{s}}$ runs over prime divisors on $\mathbb{P}^1_{\bar{s}}$, is constructible and upper semicontinuous. In particular, the subset

$$W = \{s \in S \mid f_{\bar{s}} \colon (\mathcal{X}_{\bar{s}}, \mathcal{D}_{\bar{s}}, \mathcal{H}_{\bar{s}}) \to \mathcal{P}_{\bar{s}} \text{ is uniformly adiabatically K-stable}\}$$

is open and there exists a positive real number v such that

$$\delta_{(\mathbb{P}^1, B_{\bar{s}})}(-K_{\mathbb{P}^1} - B_{\bar{s}} - M_{\bar{s}}) \geq 1 + v$$

for every geometric point $\bar{s} \in W$, where $B_{\bar{s}}$ and $M_{\bar{s}}$ are the discriminant $\mathbb{Q}$-divisor and the moduli $\mathbb{Q}$-divisor with respect to $f_{\bar{s}}$, respectively.

Proof We first reduce Theorem 4.2 to the case where $\mathcal{P} \cong \mathbb{P}^1_S$, $m = 1$ and $\mathcal{L} = \mathcal{O}_{\mathbb{P}^1_S}(1)$. For every closed point $s \in S$, there exists an étale morphism $g^s \colon T^s \to S$ such that $s \in g^s(T^s)$ and $\mathcal{P}_{T^s} \to T^s$ has a section $\iota^s \colon T^s \to \mathcal{P}_{T^s}$; see [Olsson 2016, Corollary 1.3.10]. By considering $T = \bigsqcup_{s_i} T^{s_i}$ for some finitely many closed points $s_i \in S$, we obtain an étale surjective morphism $g \colon T \to S$ such that $h \colon \mathcal{P}_T \to T$ has a section $\iota \colon T \to \mathcal{P}_T$. Then $\iota(T)$ is a Cartier divisor on $\mathcal{P}_T$; see [Fantechi et al. 2005, Lemma 9.3.4] and [Kollár 2023, Definition–Lemma 4.20]. By [Hartshorne 1977, III, Corollary 12.9], the sheaf $h_*\mathcal{O}_{\mathcal{P}_T}(\iota(T))$ is locally free of rank two and

$$h^*h_*\mathcal{O}_{\mathcal{P}}(\iota(T)) \to H^0(\mathcal{P}_t, \mathcal{O}_{\mathcal{P}_t}(\iota(T)|_{\mathcal{P}_t}))$$

is surjective. Therefore, we obtain a morphism $\mathcal{P}_T \to \mathbb{P}_T(h_*\mathcal{O}_{\mathcal{P}_T}(\iota(T)))$. Then the right-hand side is a $\mathbb{P}^1$-bundle, and the morphism is an isomorphism. Since g is open and surjective, if Theorem 4.2 holds for T, then Theorem 4.2 also holds for S. Thus, we may assume that $\mathcal{P}$ is a $\mathbb{P}^1$-bundle over S. Since the problem is local, we may assume that $\mathcal{P} = \mathbb{P}^1_S$ by shrinking S. Then $\mathcal{L} \sim_S \mathcal{O}_{\mathbb{P}^1_S}(m)$. It is easy to see that we may replace $\mathcal{L}$ by $\mathcal{O}_{\mathbb{P}^1_S}(m)$. In this way, we may assume that $\mathcal{L} = \mathcal{O}_{\mathbb{P}^1_S}(m)$. By replacing u with u/m, we may assume $m = 1$.

Next, we show that h is constructible. By [Matsumura 1980, (6,C)], it suffices to show that $h^{-1}(w)$ contains a nonempty open subset of S under the assumption that S is a variety and that $h^{-1}(w)$ is dense for every $w \in \mathbb{Q}_{>0}$. We pick an open subset $V \subset S \setminus \mathrm{Sing}(S)$. Since the fibers of $\pi \colon \mathcal{X} \to S$ are normal, we have $K_{\pi^{-1}(V)/V} = K_{\pi^{-1}(V)} - (\pi|_{\pi^{-1}(V)})^* K_V$. Thus, $K_{\pi^{-1}(V)} + \mathcal{D}|_{\pi^{-1}(V)}$ is $\mathbb{Q}$-Cartier. Let η be the generic point of S. By Lemma 4.1 and shrinking V if necessary, we may assume

$$\max_{\overline{P}} \operatorname{coeff}_{\overline{P}}(B_{\overline{\eta}}) = \max_{P} \operatorname{coeff}_P(B_t)$$

for every closed point $t \in V$, where $\overline{P}$ (resp. P) runs over prime divisors on $\mathbb{P}^1_{\overline{\eta}}$ (resp. $\mathbb{P}^1_t$). For any point $s \in V$, by applying Lemma 4.1 to $\overline{\{s\}} \cap V$, we see that $\max_{P_{\overline{s}}} \operatorname{coeff}_{P_{\overline{s}}}(B_{\overline{s}})$ are determined by $f_t \colon (\mathcal{X}_t, \mathcal{D}_t) \to \mathbb{P}^1_t$ for general closed points $t \in \overline{\{s\}} \cap V$. This means that h is constant on V. Thus the constructibility holds.

From now on we prove the upper semicontinuity. The constructibility of h implies that h takes only finitely many values. We fix $w \in \mathbb{Q}_{>0}$. By Lemma 2.2, we may assume that S is a curve and $h(s) \geq w$ for every general point $s \in S$. Then S is smooth, and hence we may write $K_{\mathcal{X}/S} = K_{\mathcal{X}} - \pi^* K_S$. Thus $K_{\mathcal{X}} + \mathcal{D}$ is $\mathbb{Q}$-Cartier. Let $\mathcal{B}$ be the discriminant $\mathbb{Q}$-divisor with respect to $f \colon (\mathcal{X}, \mathcal{D}) \to \mathbb{P}^1_S$. By Lemma 4.1, we can find an open subset $V \subset S$ such that

$$\max_{Q} \operatorname{coeff}_Q(\mathcal{B}) = \max_{P} \operatorname{coeff}_P(B_t)$$

for every closed point $t \in V$, where Q (resp. P) runs over prime divisors on $\mathbb{P}^1_S$ (resp. $\mathbb{P}^1_t$). Since $h(s) \geq w$ for general points $s \in S$, we have

$$\max_{Q} \operatorname{coeff}_Q(\mathcal{B}) \geq u.$$

In our situation, the klt property of the geometric fibers of π and the inversion of adjunction [Kawakita 2007] imply that $(\mathcal{X}, \mathcal{D} + \mathcal{X}_s)$ is lc for every closed point $s \in S$. Therefore, every component of $\mathcal{B}$ is horizontal over S. By our assumption, there exists a component T of $\mathcal{B}$ such that

$$\mathrm{coeff}_T(\mathcal{B}) = \max_Q \mathrm{coeff}_Q(\mathcal{B}) \geq u.$$

Since every component of $\mathcal{B}$ dominates S, by [Ambro 2004, Lemma 5.1], this fact is preserved even if we take any finite base change of $(\mathcal{X}, \mathcal{D}) \to \mathbb{P}^1_S \to S$. Let $\psi\colon T^\nu \to S$ be the natural morphism, where T^ν is the normalization of T. We consider the base change of $(\mathcal{X}, \mathcal{D}) \to \mathbb{P}^1_S \to S$ by ψ, which we denote by $(\mathcal{X}_{T^\nu}, \mathcal{D}_{T^\nu}) \to \mathbb{P}^1_{T^\nu} \to T^\nu$, with the morphism $\psi_{\mathbb{P}^1}\colon \mathbb{P}^1_{T^\nu} \to \mathbb{P}^1_S$. By construction, $\psi^*_{\mathbb{P}^1} T$ has a component isomorphic to T^ν. Since we only need to deal with closed points of S, we may replace $(\mathcal{X}, \mathcal{D}) \to \mathbb{P}^1_S \to S$ with $(\mathcal{X}_{T^\nu}, \mathcal{D}_{T^\nu}) \to \mathbb{P}^1_{T^\nu} \to T^\nu$. By this replacement, we may assume that $T \to S$ is an isomorphism. We put $\gamma = 1 - \mathrm{coeff}_T(\mathcal{B})$. Then there is a prime divisor $\mathcal{E}$ over $\mathcal{X}$ such that $\mathcal{E}$ maps onto T and $A_{(\mathcal{X}, \mathcal{D} + u' f^* T)}(\mathcal{E}) < 0$ for any real number $u' > \gamma$. Since T dominates S, for every closed point $c \in S$, the pair $(\mathcal{X}, \mathcal{D} + u' f^* T + \mathcal{X}_c)$ is not lc around $\mathcal{X}_c$. By the inversion of adjunction [Kawakita 2007], the pair $(\mathcal{X}_c, \mathcal{D}_c + u' f_c^* T|_{\mathbb{P}^1_c})$ is not lc for any $u' > \gamma$. Since $\mathbb{P}^1_S$ is smooth and $T \to S$ is an isomorphism, $T|_{\mathbb{P}^1_c}$ is a prime divisor on $\mathbb{P}^1_c$. Thus, the discriminant $\mathbb{Q}$-divisor B_c with respect to $f_c\colon (\mathcal{X}_c, \mathcal{D}_c) \to \mathbb{P}^1_c$ has a component whose coefficient is at least $1 - \gamma$. This shows that for every closed point $c \in S$,

$$u \leq \max_Q \mathrm{coeff}_Q(\mathcal{B}) = 1 - \gamma \leq \max_{P'} \mathrm{coeff}_{P'}(B_c),$$

where P' runs over prime divisors on $\mathbb{P}^1_c$. Thus the upper semicontinuity of h holds. The final statement of Theorem 4.2 follows from this fact and Example 2.28. □

4.2 Separatedness

In this subsection we show the separatedness of the moduli spaces that we will construct in Section 5.

Notation 4.3 Let C be an affine curve. We say that $f\colon (X, \Delta, L) \to C$ is a *polarized* $\mathbb{Q}$*-Gorenstein family* if $f\colon (X, \Delta) \to C$ is a log $\mathbb{Q}$-Gorenstein family over C and L is an f-ample line bundle. Let $0 \in C$ be a closed point and $C^\circ = C \setminus \{0\}$ the punctured curve. We put

$$(X, \Delta, L) \times_C C^\circ = (X \times_C C^\circ, \Delta \times_C C^\circ, L|_{X \times_C C^\circ}).$$

For another polarized $\mathbb{Q}$-Gorenstein family $f'\colon (X', \Delta', L') \to C$, we define

$$g\colon (X, \Delta, L) \to (X', \Delta', L')$$

to be a C-isomorphism $g\colon X \to X'$ such that $f' \circ g = f$, $g_* \Delta = \Delta'$ and $g^* L' \sim_C L$. We define C°-isomorphisms between $(X, \Delta, L) \times_C C^\circ$ and $(X', \Delta', L') \times_C C^\circ$ similarly.

Let $f\colon (X, \Delta, L) \to C$ and $f'\colon (X', \Delta', L') \to C$ be polarized $\mathbb{Q}$-Gorenstein families. For contractions $\pi\colon (X, \Delta, L) \to (\mathbb{P}^1_C, \mathcal{O}(1))$ and $\pi'\colon (X', \Delta', L') \to (\mathbb{P}^1_C, \mathcal{O}(1))$ over C, we define $(\alpha, \beta)\colon \pi \to \pi'$ as a pair of C-isomorphisms $\alpha\colon (X, \Delta, L) \to (X', \Delta', L')$ and $\beta\colon (\mathbb{P}^1_C, \mathcal{O}(1)) \to (\mathbb{P}^1_C, \mathcal{O}(1))$ such that $\pi' \circ \alpha = \beta \circ \pi$.

The following was shown by Boucksom when $\Delta = 0$, but we write here the proof for the sake of completeness.

Proposition 4.4 [Boucksom 2014, Theorem 1.1; Blum and Xu 2019, Theorem 3.1 and Remark 3.6] *Let C be an affine curve. Let $f\colon (X,\Delta,L)\to C$ and $f'\colon (X',\Delta',L')\to C$ be two polarized $\mathbb{Q}$-Gorenstein families. Suppose that there exists a C°-isomorphism*

$$g^\circ\colon (X,\Delta,L)\times_C C^\circ \to (X',\Delta',L')\times_C C^\circ$$

and both $K_X+\Delta$ and $K_{X'}+\Delta'$ are nef over C. If (X_0,Δ_0) is klt and (X'_0,Δ'_0) is lc, then g° can be extended to a C-isomorphism $g\colon (X,\Delta,L)\to(X',\Delta',L')$.

Proof Let $g\colon X \dashrightarrow X'$ be the birational map induced by g°. It is sufficient to prove that g is a C-isomorphism.

We first show that g and g^{-1} do not contract any divisor. We apply the argument in [Boucksom 2014]. By the inversion of adjunction [Kawakita 2007] and shrinking C around $0\in C$, we may assume that $(X,\Delta+X_0)$ is plt and $(X',\Delta'+X'_0)$ is lc. Take a common log resolution $\pi: Y\to X$ and $\pi': Y\to X'$ of g. By construction, $g^\circ\circ\pi|_{Y\times_C C^\circ}$ coincides with $\pi'|_{Y\times_C C^\circ}$. Let Γ be the sum of $\pi|_{Y\times_C C^\circ}$-exceptional prime divisors, and let $\overline{\Gamma}$ be the closure in Y. Then $\overline{\Gamma}$ is π-exceptional and also π'-exceptional. By the log canonicity of $(X,\Delta+X_0)$, the $\mathbb{Q}$-divisor

$$E := K_Y+\pi_*^{-1}\Delta+\overline{\Gamma}+Y_{0,\mathrm{red}}-\pi^*(K_X+\Delta+X_0)$$

is effective and π-exceptional. Similarly, we see that

$$E' := K_Y+\pi_*^{-1}\Delta+\overline{\Gamma}+Y_{0,\mathrm{red}}-\pi'^*(K_{X'}+\Delta'+X'_0)$$

is effective and π'-exceptional. Since $K_X+\Delta$ is nef over C, by applying the negativity lemma to π' and $E-E'$, we see that $E-E'$ is effective. Similarly, we see that $E'-E$ is effective. These facts imply that $E=E'$ and hence

$$\pi^*(K_X+\Delta+X_0)=\pi'^*(K_{X'}+\Delta'+X'_0).$$

Therefore, $A_{(X,\Delta+X_0)}(F)=A_{(X',\Delta'+X'_0)}(F)$ for every prime divisor F on Y. Recalling that $(X,\Delta+X_0)$ is plt, we see that $A_{(X,\Delta+X_0)}(F)=0$ if and only if $F=\pi_*^{-1}X_0$. Now

$$\begin{aligned} A_{(X,\Delta+X_0)}(\pi'^{-1}_*X'_0) &= A_{(X',\Delta'+X'_0)}(\pi'^{-1}_*X'_0)\\ &= 1-\mathrm{coeff}_{X'_0}(\Delta'+X'_0)=0. \end{aligned}$$

From these facts, we have $\pi_*\pi'^{-1}_*X'_0=X_0$. Since X_0 (resp. X'_0) is the fiber of f (resp. f') over $0\in C$, we see that g and g^{-1} do not contract any divisor.

We now prove that g is a C-isomorphism. Consider L' as a Cartier divisor on X', and put $D=g_*^{-1}L'$. By our hypothesis, we have $L|_{X\times_C C^\circ}\sim_{C^\circ}(g^\circ)^*L|_{X'\times_C C^\circ}$. Since g does not contract any divisor, we have $D \sim_C L$; see [Hartshorne 1977, II, Proposition 6.5]. Thus, by Serre's S_2-condition, g induces

$$X = \mathbf{Proj}_C\Big(\bigoplus_{m\ge 0} f_*L^{\otimes m}\Big) = \mathbf{Proj}_C\Big(\bigoplus_{m\ge 0} f_*\mathcal{O}_X(mD)\Big) \cong \mathbf{Proj}_C\Big(\bigoplus_{m\ge 0} f'_*L'^{\otimes m}\Big) = X'.$$

This shows that g is indeed a C-isomorphism. □

Corollary 4.5 *Let (X, Δ, L) be a polarized klt pair such that $K_X + \Delta$ is nef. Then $\operatorname{Aut}(X, \Delta, L)$ is finite.*

Proof It follows from Proposition 4.4 as [Blum and Xu 2019, Corollary 3.5]. □

We are ready to prove the main theorem of this subsection.

Theorem 4.6 (separatedness) *Let $\pi\colon (X, \Delta, H) \to C$ and $\pi'\colon (X', \Delta', H') \to C$ be two polarized $\mathbb{Q}$-Gorenstein families over a curve such that (X_0, Δ_0) is klt and (X'_0, Δ'_0) is lc. Let*

$$g\colon (X, \Delta, H) \to (\mathbb{P}^1_C, \mathcal{O}(1)) \quad\text{and}\quad g'\colon (X', \Delta', H') \to (\mathbb{P}^1_C, \mathcal{O}(1))$$

be contractions over C such that $K_X + \Delta \sim_{\mathbb{Q},\mathbb{P}^1_C} 0$ and $K_{X'} + \Delta' \sim_{\mathbb{Q},\mathbb{P}^1_C} 0$. Let $0 \in C$ be a closed point, and let $g_0\colon X_0 \to \mathbb{P}^1$ (resp. $g'_0\colon X'_0 \to \mathbb{P}^1$) be the restriction of g (resp. g') to $0 \in C$. Suppose that there exists an isomorphism $(\alpha^\circ, \beta^\circ)\colon g|_{X\times_C C^\circ} \cong g'|_{X'\times_C C^\circ}$ over C° such that

- *$g_0\colon (X_0, \Delta_0, H_0) \to \mathbb{P}^1$ is uniformly adiabatically K-stable,*
- *for the discriminant $\mathbb{Q}$-divisor B'_0 and the moduli $\mathbb{Q}$-divisor M'_0 with respect to $g'_0\colon (X'_0, \Delta'_0) \to \mathbb{P}^1$, we have $\delta_{(\mathbb{P}^1, B'_0)}(-K_{\mathbb{P}^1} - B'_0 - M'_0) \ge 1$, and*
- *we have $-K_X - \Delta \sim_{C,\mathbb{Q}} wg^*\mathcal{O}_{\mathbb{P}^1_C}(1)$ and $-K_{X'} - \Delta' \sim_{C,\mathbb{Q}} w'g'^*\mathcal{O}_{\mathbb{P}^1_C}(1)$ for some positive rational numbers w and w'.*

Then $(\alpha^\circ, \beta^\circ)$ can be extended to an isomorphism $(\alpha, \beta)\colon g \to g'$ over C.

Proof We will reduce the theorem to Proposition 4.4 as in [Blum and Xu 2019, Theorem 3.1]. By our hypothesis of $(\alpha^\circ, \beta^\circ)$, it is easy to see that $w = w'$. Denote the bases of g and g' by $\mathcal{C}$ and $\mathcal{C}'$, respectively. Note that $\mathcal{C}$ and $\mathcal{C}'$ are isomorphic to $\mathbb{P}^1_C$. Let $\mathcal{L}$ (resp. $\mathcal{L}'$) be the line bundle on $\mathcal{C}$ (resp. $\mathcal{C}'$) isomorphic to $\mathcal{O}(1)$. We denote the structure morphisms $\mathcal{C} \to C$ and $\mathcal{C}' \to C$ by η and η', respectively.

Replacing $\mathcal{L}$ by $\mathcal{L} + d\mathcal{C}_0$ for some sufficiently large $d \in \mathbb{Z}_{>0}$, we may assume that the birational map $(\mathcal{C}, \mathcal{L}) \dashrightarrow (\mathcal{C}', \mathcal{L}')$ over C induces an inclusion $\eta'_*\mathcal{L}' \subset \eta_*\mathcal{L}$ as sheaves of $\mathcal{O}_C$-modules. Now $\mathcal{O}_{C,0}$ is a divisorial valuation ring and both $\eta_*\mathcal{L} \otimes \mathcal{O}_{C,0}$ and $\eta'_*\mathcal{L}' \otimes \mathcal{O}_{C,0}$ are free $\mathcal{O}_{C,0}$-modules of rank two. Hence, by fixing a generator t of the maximal ideal of $\mathcal{O}_{C,0}$, we may find free bases $\{s, u\} \subset \eta_*\mathcal{L} \otimes \mathcal{O}_{C,0}$ and $\{s', u'\} \subset \eta'_*\mathcal{L}' \otimes \mathcal{O}_{C,0}$ such that $s' = t^\lambda s$ and $u' = t^\mu u$ for some $\lambda, \mu \in \mathbb{Z}_{\ge 0}$.

By shrinking C around 0, we may assume that t, s, u, s', and u' are global sections. By definition, s and u correspond to prime divisors E_1 and E_2 on $\mathcal{C}$ respectively such that $E_1|_{\mathcal{C}_0} \ne E_2|_{\mathcal{C}_0}$. Similarly,

s' and u' correspond to prime divisors E_1' and E_2' on $\mathcal{C}'$ respectively such that $E_1'|_{\mathcal{C}_0'} \neq E_2'|_{\mathcal{C}_0'}$. We put $D = \frac{1}{2}w(E_1 + E_2)$ and $D' = \frac{1}{2}w(E_1' + E_2')$. Then D is the strict transform of D' since they coincide over C°. Note that $(X_0, \Delta_0 + g_0^* D_0)$ is klt and $(X_0', \Delta_0' + g_0'^* D_0')$ is lc. Indeed, $(\mathbb{P}^1, B_0' + D_0')$ is lc since

$$\alpha_{(\mathbb{P}^1, B_0')}(-K_{\mathbb{P}^1} - B_0' - M_0') = \tfrac{1}{2}\delta_{(\mathbb{P}^1, B_0')}(-K_{\mathbb{P}^1} - B_0' - M_0') \geq \tfrac{1}{2}$$

by Example 2.28 and the second assumption of Theorem 4.6. Then the log canonicity of $(X_0', \Delta_0' + g_0'^* D_0')$ follows from the log canonicity of $(\mathbb{P}^1, B_0' + D_0')$ in the same way as [Ambro 2004, Theorem 3.1]. Similarly, we have

$$\alpha_{(\mathbb{P}^1, B_0)}(-K_{\mathbb{P}^1} - B_0 - M_0) = \tfrac{1}{2}\delta_{(\mathbb{P}^1, B_0)}(-K_{\mathbb{P}^1} - B_0 - M_0) > \tfrac{1}{2}$$

by Example 2.28 and the first assumption of Theorem 4.6, where B_0 (resp. M_0) is the discriminant $\mathbb{Q}$-divisor (resp. moduli $\mathbb{Q}$-divisor) with respect to the klt-trivial fibration g_0. We see that $(X_0, \Delta_0 + g_0^* D_0)$ is klt in the same way. We also have the relations $K_X + \Delta + g^* D \sim_{C,\mathbb{Q}} 0$ and $K_{X'} + \Delta' + g'^* D' \sim_{C,\mathbb{Q}} 0$. Thus, we have a unique extension

$$\alpha\colon (X, \Delta + g^* D, H) \cong (X', \Delta' + g'^* D', H')$$

of α° over C by Proposition 4.4. Here, $\alpha^* g'^* D' = g^* D$. Thus, $\alpha^* g'^* E_1' = g^* E_1$ and $\alpha^* g'^* E_2' = g^* E_2$ hold and they generate the pencils defining the two contractions $g' \circ \alpha$ and g. Therefore, we have a natural extension $\beta\colon (\mathcal{C}, \mathcal{L}) \cong (\mathcal{C}', \mathcal{L}')$ of β°. It is easy to see that (α, β) is an isomorphism from g to g' over C. □

To construct our moduli spaces, we only need Theorem 4.6 for uniformly adiabatically K-stable g_0'. In this case, Theorem 4.6 follows from [Hattori 2024a, Corollary 3.22] when $\Bbbk = \mathbb{C}$. Since we do not know whether $(X_0', \Delta_0', \epsilon H_0' + g_0'^* \mathcal{O}(1))$ is specially K-semistable or not, we cannot apply [Hattori 2024a, Corollary 3.22] directly. Theorem 4.6 is applicable to the case when g_0' is an adiabatically K-semistable klt-trivial fibration; see [Hattori 2022, Theorem A].

The following is also important for construction of our moduli spaces.

Corollary 4.7 (finiteness of stabilizers) *Let $f\colon (X, \Delta, H) \to \mathbb{P}^1$ be a polarized uniformly adiabatically K-stable klt-trivial fibration such that $-(K_X + \Delta)$ is nef and not numerically trivial. Then* $\mathrm{Aut}\,(f\colon (X, \Delta, H) \to (\mathbb{P}^1, \mathcal{O}_{\mathbb{P}^1}(1)))$ *is a finite group.*

Proof Since $\mathrm{Aut}\,(f\colon (X, \Delta, H) \to (\mathbb{P}^1, \mathcal{O}_{\mathbb{P}^1}(1)))$ is represented by a closed subgroup of a linear algebraic group $\mathrm{Aut}\,(X, \Delta, H) \times \mathrm{Aut}\,(\mathbb{P}^1, \mathcal{O}_{\mathbb{P}^1}(1))$ [Fantechi et al. 2005, Section 5.6], it is sufficient to show that $\mathrm{Aut}\,(f\colon (X, \Delta, H) \to (\mathbb{P}^1, \mathcal{O}_{\mathbb{P}^1}(1)))$ is proper. Let C be an arbitrary affine curve and fix a closed point $0 \in C$. Set $C^\circ := C \setminus \{0\}$ and take an arbitrary isomorphism

$$\varphi^\circ \in \mathrm{Aut}\,(f|_{(X,\Delta,H)\times C^\circ}\colon (X, \Delta, H) \times C^\circ \to (\mathbb{P}^1, \mathcal{O}_{\mathbb{P}^1}(1)) \times C^\circ).$$

By Theorem 4.6, we extend φ° to φ over C entirely. □

4.3 Invariance of plurigenera

In this subsection we prove a result on the invariance of plurigenera, which is a generalization of [Nakayama 1986] and a key statement to construct our moduli spaces.

Theorem 4.8 *Let $f: (X, \Delta) \to S$ be a log $\mathbb{Q}$-Gorenstein family such that S is a normal variety. Suppose that there is $e \in \{1, -1\}$ such that for every geometric point $\overline{s} \in S$, $(X_{\overline{s}}, \Delta_{\overline{s}})$ is a klt pair and $e(K_{X_{\overline{s}}} + \Delta_{\overline{s}})$ is semiample. Let r be a positive integer such that $r(K_{X/S} + \Delta)$ is Cartier. Then, for every positive integer n, the function*

$$S \ni t \mapsto \dim H^0(X_t, \mathcal{O}_{X_t}(enr(K_{X_t} + \Delta_t)))$$

is constant.

First, we treat the case when S is a curve.

Proposition 4.9 *Let $f: (X, \Delta) \to C$ be a log $\mathbb{Q}$-Gorenstein family such that C is a curve and all closed fibers of f are klt pairs. Let D be a Cartier divisor on X such that $D - (K_X + \Delta)$ is semiample over C. Then, for every closed fiber F of f, the natural morphism $f_*\mathcal{O}_X(D) \to H^0(F, \mathcal{O}_F(D|_F))$ is surjective.*

Proof Note that X is a normal variety. The klt property of the closed fibers of f and the inversion of adjunction [Kawakita 2007] imply that (X, Δ) is klt. Let $g: Y \to X$ be a log resolution of (X, Δ). We can write

$$K_Y + \Delta_Y = g^*(K_X + \Delta) + E$$

for some effective $\mathbb{Q}$-divisors Δ_Y and E which have no common component. Then $(Y, \Delta_Y + \lceil E \rceil - E)$ is a log smooth klt pair. Let $c \in C$ be an arbitrary closed point with the fiber $F := f^*c$. Then

$$(g^*D + \lceil E \rceil - g^*F) - (K_Y + (\Delta_Y + \lceil E \rceil - E)) = g^*(D - (K_X + \Delta - F)).$$

Thus, $(g^*D + \lceil E \rceil - g^*F) - (K_Y + (\Delta_Y + \lceil E \rceil - E))$ is nef and big over X, and the divisor is semiample over C because $D - (K_X + \Delta)$ is semiample over C by the hypothesis. By the Kawamata–Viehweg vanishing theorem, we have $R^q g_*\mathcal{O}_Y(g^*D + \lceil E \rceil - g^*F) = 0$ for every $q > 0$. Thus, the Leray spectral sequence implies

$$R^1(f \circ g)_*\mathcal{O}_Y(g^*D + \lceil E \rceil - g^*F) \cong R^1 f_*(g_*\mathcal{O}_Y(g^*(D - F) + \lceil E \rceil)) = R^1 f_*\mathcal{O}_X(D - F),$$

where the last equality follows from that E is effective and g-exceptional. By applying the torsion-free theorem [Fujino 2011, Theorem 6.3(i)] to $f \circ g: Y \to C$, $g^*D + \lceil E \rceil - g^*F$ and $(Y, \Delta_Y + \lceil E \rceil - E)$, we see that $R^1 f_*\mathcal{O}_X(D - F)$ is torsion free. Now consider the exact sequence

$$f_*\mathcal{O}_X(D) \to f_*\mathcal{O}_F(D|_F) \xrightarrow{\delta} R^1 f_*\mathcal{O}_X(D - F)$$

which is induced by

$$0 \to \mathcal{O}_X(D - F) \to \mathcal{O}_X(D) \to \mathcal{O}_F(D|_F) \to 0.$$

Since $R^1 f_*\mathcal{O}_X(D-F)$ is torsion free and $f_*\mathcal{O}_F(D|_F)$ is zero outside c, we see that δ is the zero map. This implies that

$$f_*\mathcal{O}_X(D) \to f_*\mathcal{O}_F(D|_F) = H^0(F, \mathcal{O}_F(D|_F))$$

is surjective. $\square$

Proof of Theorem 4.8 It is sufficient to prove that the equality

$$\dim H^0(X_s, \mathcal{O}_{X_s}(enr(K_{X_s}+\Delta_s))) = \dim H^0(X_{s'}, \mathcal{O}_{X_{s'}}(enr(K_{X_{s'}}+\Delta_{X_{s'}})))$$

holds for any two closed points s, $s' \in S$. Let $C \subset S$ be a connected (but not necessarily irreducible or smooth) curve passing through s and s'; see [Mumford 2008, Section 6, Lemma]. Replacing S with the normalization of any component of C, we may assume that S is a curve.

By the hypothesis, $e(K_X+\Delta)$ is f-nef. Since the restriction of $e(K_X+\Delta)$ to the geometric generic fiber is semiample, $e(K_X+\Delta)$ is f-abundant [Fujino 2012, Definition 4.1]. By [Fujino 2012, Theorem 1.1], $e(K_X+\Delta)$ is semiample over S. Then $enr(K_X+\Delta)-(K_X+\Delta)$ is also semiample over S for every positive integer n. By Proposition 4.9, for every closed fiber F, the morphism

$$f_*\mathcal{O}_X(enr(K_X+\Delta)) \to H^0(F, \mathcal{O}_F(enr(K_F+\Delta|_F)))$$

is surjective. By the cohomology and base change theorem, $\dim H^0(X_t, \mathcal{O}_{X_t}(enr(K_{X_t}+\Delta_t)))$ is independent of $t \in S$. $\square$

5 Construction of moduli

In this section, we construct the moduli of uniformly adiabatically K-stable polarized klt-trivial fibrations over curves such that the canonical divisor is not numerically trivial. Throughout this section, we fix $d \in \mathbb{Z}_{>0}$, $u \in \mathbb{Q}_{\neq 0}$ with $e := u/|u|$, $v \in \mathbb{Q}_{>0}$ and $w \in \mathbb{Q}_{>0}$. We define

$$\mathfrak{Z}_{d,v,u,w} := \left\{ f\colon (X, \Delta=0, A) \to C \;\middle|\; \begin{array}{ll} \text{(i)} & f \text{ is a uniformly adiabatically K-stable polarized klt-trivial fibration over a curve } C, \\ \text{(ii)} & \dim X = d, \\ \text{(iii)} & K_X \equiv u f^* H \text{ for some line bundle } H \text{ on } C \text{ such that } \deg H = 1, \\ \text{(iv)} & A \text{ is an ample line bundle on } X \text{ such that } (K_X \cdot A^{d-1}) = uv \text{ and } \mathrm{vol}(A) \le w. \end{array} \right\}$$

Then it is not difficult to check that if $f\colon (X, \Delta=0, A) \to C$ is an element of $\mathfrak{Z}_{d,v,u,w}$ then $(X,0) \to C \in \mathfrak{G}_{d,\{0\},v,u}$, where $\mathfrak{G}_{d,\{0\},v,u}$ is the set $\mathfrak{G}_{d,\Theta,v,u}$ in Section 3 with $\Theta = \{0\}$. By Lemma 3.1, there exists an $r \in \mathbb{Z}_{>0}$, depending only on d, u and v, such that for any element $f\colon (X,0) \to C$ of $\mathfrak{G}_{d,\{0\},v,u}$, we have $erK_X \sim f^*D$ for some very ample Cartier divisor D on C.

The following theorem is the main result of this paper.

Theorem 5.1 *We fix $d \in \mathbb{Z}_{>0}$, $u \in \mathbb{Q}_{\neq 0}$ with $e := u/|u|$, $v \in \mathbb{Q}_{>0}$, $w \in \mathbb{Q}_{>0}$ and $r \in \mathbb{Z}_{>0}$ in Lemma 3.1 for $\mathfrak{G}_{d,\{0\},v,u}$. Let $\mathscr{M}_{d,v,u,w,r}$ be a full subcategory of $\mathfrak{Pol}$ such that for any locally Noetherian scheme S over $\Bbbk$, we define $\mathscr{M}_{d,v,u,w,r}(S)$ to be a groupoid whose objects are*

$$\left\{ \begin{tikzcd} (\mathcal{X},\mathscr{A}) \arrow[rr, "f"] \arrow[dr, "\pi_{\mathcal{X}}"'] & & \mathcal{C} \arrow[dl] \\ & S & \end{tikzcd} \;\middle|\; \begin{array}{ll} \text{(i)} & \pi_{\mathcal{X}} \text{ is a flat projective morphism and } \mathcal{X} \text{ is a scheme,} \\ \text{(ii)} & \mathscr{A} \in \mathrm{Pic}_{\mathcal{X}/S}(S) \text{ such that } \mathscr{A}_{\overline{s}} \text{ is ample for any geometric point } \overline{s} \in S, \\ \text{(iii)} & \omega_{\mathcal{X}/S}^{[r]} \text{ exists as a line bundle,} \\ \text{(iv)} & \pi_{\mathcal{X}*}\omega_{\mathcal{X}/S}^{[ler]} \text{ is locally free and generates } H^0(\mathcal{X}_s, \mathcal{O}_{\mathcal{X}_s}(lerK_{\mathcal{X}_s})) \text{ for any} \\ & \text{point } s \in S \text{ and any } l \in \mathbb{Z}_{>0}, \\ \text{(v)} & f \text{ is the ample model of } \omega_{\mathcal{X}/S}^{[er]} \text{ over } S \text{ and } f_{\overline{s}}\colon (\mathcal{X}_{\overline{s}}, 0, \mathscr{A}_{\overline{s}}) \to \mathcal{C}_{\overline{s}} \in \mathfrak{Z}_{d,v,u,w} \\ & \text{for any geometric point } \overline{s} \in S. \end{array} \right\}$$

Then $\mathscr{M}_{d,v,u,w,r}$ is a separated Deligne–Mumford stack of finite type over $\Bbbk$. Furthermore, there exists a coarse moduli space of $\mathscr{M}_{d,v,u,w,r}$.

Remark 5.2 For any S-isomorphism $g\colon \mathcal{X} \to \mathcal{X}'$ as above, we have a unique S-isomorphism $h\colon \mathcal{C} \to \mathcal{C}'$ such that $f' \circ g = h \circ f$. This is the reason why we do not consider morphisms between $\mathcal{C}$ and $\mathcal{C}'$.

In this section, for every object $(\mathcal{X},\mathscr{A}) \to \mathcal{C} \in \mathscr{M}_{d,v,u,w,r}(S)$, the structure morphism $(\mathcal{X},\mathscr{A}) \to S$ is denoted by $\pi_{\mathcal{X}}$ unless otherwise stated. When an object $(\mathcal{X}_T,\mathscr{A}_T) \to \mathcal{C}_T$ of $\mathscr{M}_{d,v,u,w,r}(T)$ is the base change of $(\mathcal{X},\mathscr{A}) \to \mathcal{C}$ by $T \to S$, the morphism $\pi_{\mathcal{X}_T}$ is nothing but $(\pi_{\mathcal{X}})_T$ as in (11) in Notation and conventions.

Lemma 5.3 *$\mathscr{M}_{d,v,u,w,r}$ is a stack.*

Proof We first check that $\mathscr{M}_{d,v,u,w,r}$ is a category fibered in groupoids. It suffices to show that for any $\pi_{\mathcal{X}}\colon (\mathcal{X},\mathscr{A}) \to \mathcal{C} \to S \in \mathscr{M}_{d,v,u,w,r}(S)$ and any morphism $h\colon T \to S$ of schemes, the base change $\pi_{\mathcal{X}_T}\colon (\mathcal{X}_T,\mathscr{A}_T) \to \mathcal{C}_T \to T$ is the pullback of π along h in the sense of [Olsson 2016, Definition 3.1.1]. By the conditions (iv) and (v) in the definition of $\mathscr{M}_{d,v,u,w,r}$ and the theorem of cohomology and base change, we see that

$$\mathcal{C}_T := \mathcal{C} \times_S T = \mathbf{Proj}_S\Biggl(\bigoplus_{l \geq 0} \pi_{\mathcal{X}*}\omega_{\mathcal{X}/S}^{[ler]}\Biggr) \times_S T \cong \mathbf{Proj}_T\Biggl(\bigoplus_{l \geq 0} \pi_{\mathcal{X}_T*}\omega_{\mathcal{X}_T/T}^{[ler]}\Biggr).$$

This shows $(\mathcal{X}_T,\mathscr{A}_T) \to \mathcal{C}_T \in \mathscr{M}_{d,v,u,w,r}(T)$. Hence, $\mathscr{M}_{d,v,u,w,r}$ is indeed a category fibered in groupoids.

From now on, we check that $\mathscr{M}_{d,v,u,w,r}$ is a stack. Since Definition 2.9(1) has been already checked in Lemma 2.15, it suffices to show the condition of Definition 2.9 (2) for $\mathscr{M}_{d,v,u,w,r}$. We note that $\mathscr{M}_{d,v,u,w,r}$ satisfies the condition of Remark 2.10. Let $g\colon S' \to S$ be an étale covering and $(f'\colon (\mathcal{X}',\mathscr{A}') \to \mathcal{C}',\sigma)$ a descent datum with the structure morphism $\pi_{\mathcal{X}'}\colon (\mathcal{X}',\mathscr{A}') \to S'$. We will show that (f',σ) is effective.

By Lemma 2.15, $(\pi_{\mathcal{X}'}\colon(\mathcal{X}',\mathscr{A}')\to S',\sigma)$ is a descent datum in $\mathfrak{Pol}$. Therefore, the datum comes from some element $\pi\colon(\mathcal{X},\mathscr{A})\to S\in\mathfrak{Pol}(S)$. By the functoriality of $\omega^{[r]}_{\mathcal{X}'/S'}$ and [Fantechi et al. 2005, Theorem 4.23], there exists a line bundle $\mathscr{L}$ on $\mathcal{X}$ such that

$$g_{\mathcal{X}}^{*}\mathscr{L}=\omega^{[r]}_{\mathcal{X}'/S'},$$

and there exists a morphism $\omega^{\otimes r}_{\mathcal{X}/S}\to\mathscr{L}$ whose pullback $g_{\mathcal{X}}^{*}\omega^{\otimes r}_{\mathcal{X}/S}\to g_{\mathcal{X}}^{*}\mathscr{L}$ coincides with the natural morphism $\omega^{\otimes r}_{\mathcal{X}'/S'}\to\omega^{[r]}_{\mathcal{X}'/S'}$. From these facts, we have that $\mathscr{L}=\omega^{[r]}_{\mathcal{X}/S}$. By the faithful flatness of g and the flat base change theorem [Hartshorne 1977, III, Proposition 9.3], the condition (iv) of $\mathscr{M}_{d,v,u,w,r}$ holds for π. Thus, $\omega^{[er]}_{\mathcal{X}'/S'}$ is relatively semiample, and if we set $\mathcal{C}:=\mathbf{Proj}_S\big(\bigoplus_{l\geq 0}\pi_{*}\omega^{[ler]}_{\mathcal{X}/S}\big)$, then

$$\mathcal{C}\times_S S'\cong\mathbf{Proj}_{S'}\Bigg(\bigoplus_{l\geq 0}\pi_{\mathcal{X}'*}\omega^{[ler]}_{\mathcal{X}'/S'}\Bigg)=\mathcal{C}'$$

by [Hartshorne 1977, III, Theorem 12.11]. Let $f\colon\mathcal{X}\to\mathcal{C}$ be the canonical morphism. Then the base change of f by $S'\to S$ is isomorphic to f'. From this, (v) of $\mathscr{M}_{d,v,u,w,r}$ holds for $f\colon(\mathcal{X},\mathscr{A})\to\mathcal{C}$. This shows $f\colon(\mathcal{X},\mathscr{A})\to\mathcal{C}\in\mathscr{M}_{d,v,u,w,r}(S)$, and hence (f',σ) is an effective descent datum. □

Note that the set of all klt-trivial fibrations over $\Bbbk$ belonging to $\mathfrak{Z}_{d,v,u,w}$ coincides with the set of isomorphic classes of $\mathscr{M}_{d,v,u,w,r}(\operatorname{Spec}\Bbbk)$. From now on, we fix $I\in\mathbb{Z}_{>0}$ as in Corollary 3.8 for $\mathfrak{G}_{d,\{0\},v,u,w}$. Note that $\mathfrak{Z}_{d,v,u,w}\subset\mathfrak{G}_{d,\{0\},v,u,w}$.

Lemma 5.4 *For any $d_1,d_2,d_3\in\mathbb{Z}_{>0}$ and $h\in\mathbb{Q}[t]$, let $\mathscr{M}_{d_1,d_2,d_3,h}$ be a full subcategory of $\mathscr{M}_{d,v,u,w,r}$ such that for any locally Noetherian scheme S over $\Bbbk$, we define a groupoid $\mathscr{M}_{d_1,d_2,d_3,h}(S)$ whose objects are*

$$\left\{ f\colon(\mathcal{X},\mathscr{A})\to\mathcal{C}\in\mathscr{M}_{d,v,u,w,r}(S) \;\middle|\; \begin{array}{l}\text{for every geometric point } \bar{s}\in S,\\ \bullet\ h^0(\mathcal{X}_{\bar{s}},\mathcal{O}_{\mathcal{X}_{\bar{s}}}(I\mathscr{A}_{\bar{s}}))=d_1,\\ \bullet\ h^0(\mathcal{X}_{\bar{s}},\mathcal{O}_{\mathcal{X}_{\bar{s}}}((I+1)\mathscr{A}_{\bar{s}}))=d_2,\\ \bullet\ h^0(\mathcal{X}_{\bar{s}},\mathcal{O}_{\mathcal{X}_{\bar{s}}}(erK_{\mathcal{X}_{\bar{s}}}))=d_3,\\ \bullet\ \text{the Hilbert polynomial of } \mathcal{X}_{\bar{s}} \text{ with respect to}\\ \quad (2I+1)\mathscr{A}_{\bar{s}}+erK_{\mathcal{X}_{\bar{s}}} \text{ is } h.\end{array}\right\}$$

Then $\mathscr{M}_{d_1,d_2,d_3,h}$ is an open and closed substack of $\mathscr{M}_{d,v,u,w,r}$. Furthermore, there are only finitely many $d_1,d_2,d_3\in\mathbb{Z}_{>0}$ and $h\in\mathbb{Q}[t]$ such that $\mathscr{M}_{d_1,d_2,d_3,h}$ is not an empty stack.

Proof By Theorem 4.8, any scheme S and $f\colon(\mathcal{X},\mathscr{A})\to\mathcal{C}\in\mathscr{M}_{d,v,u,w,r}(S)$ satisfy the property that $h^0(\mathcal{X}_{\bar{s}},\mathcal{O}_{\mathcal{X}_{\bar{s}}}(I\mathscr{A}_{\bar{s}}))$, $h^0(\mathcal{X}_{\bar{s}},\mathcal{O}_{\mathcal{X}_{\bar{s}}}((I+1)\mathscr{A}_{\bar{s}}))$, $h^0(\mathcal{X}_{\bar{s}},\mathcal{O}_{\mathcal{X}_{\bar{s}}}(erK_{\mathcal{X}_{\bar{s}}}))$, and the Hilbert polynomial of $\mathcal{X}_{\bar{s}}$ with respect to $(2I+1)\mathscr{A}_{\bar{s}}+erK_{\mathcal{X}_{\bar{s}}}$ are locally constant on $s\in S$. The first assertion follows from this fact. The second assertion follows from Lemma 3.1 and Corollary 3.8. □

The invariants $h^0(\mathcal{X}_{\bar{s}}, \mathcal{O}_{\mathcal{X}_{\bar{s}}}(I\mathscr{A}_{\bar{s}}))$ and $h^0(\mathcal{X}_{\bar{s}}, \mathcal{O}_{\mathcal{X}_{\bar{s}}}((I+1)\mathscr{A}_{\bar{s}}))$ in Lemma 5.4 are used to determine $\mathscr{A} \in \mathrm{Pic}_{\mathcal{X}/S}(S)$.

Notation 5.5 For each $d_1, d_2, d_3 \in \mathbb{Z}_{>0}$ and $h \in \mathbb{Q}[t]$, we set

$$H := \mathrm{Hilb}\,{\mathbb{P}^{d_1-1} \times \mathbb{P}^{d_2-1} \times \mathbb{P}^{d_3-1}}^{h,\, p_1^*\mathcal{O}(1)\otimes p_2^*\mathcal{O}(1)\otimes p_3^*\mathcal{O}(1)}.$$

Let $\tilde{\pi}:\mathcal{U} \to H$ be the morphism from the universal family $\mathcal{U}$. We set $p_i:\mathcal{U} \to \mathbb{P}_H^{d_i-1}$ as the morphism induced by the projections $\mathbb{P}_H^{d_1-1} \times_H \mathbb{P}_H^{d_2-1} \times_H \mathbb{P}_H^{d_3-1} \to \mathbb{P}_H^{d_i-1}$. We remark that H is of finite type over $\Bbbk$.

For any morphism $T \to H$, the morphism $\tilde{\pi}_T:\mathcal{U}_T \to T$ denotes the base change of $\tilde{\pi}$ by $T \to H$.

Proposition 5.6 *Fix $I \in \mathbb{Z}_{>0}$ of Corollary 3.8. For all d_1, d_2, $d_3 \in \mathbb{Z}_{>0}$ and $h \in \mathbb{Q}[t]$, the following $\mathfrak{H}$: $(\mathbf{Sch}_{/\Bbbk})^{\mathrm{op}} \to$ **Sets** is a well-defined functor and $\mathfrak{H}$ is represented by a locally closed subscheme $N_{d_1,d_2,d_3,h} \subset H$: For a scheme S, define*

$$\mathfrak{H}(S) := \left\{ (f:(\mathcal{X},\mathscr{A}) \to \mathcal{C}, \rho_1, \rho_2, \rho_3) \;\middle|\; \begin{array}{l} f \in \mathscr{M}_{d_1,d_2,d_3,h}(S) \text{ is such that } \mathscr{A} \text{ is} \\ \text{represented by a line bundle, and} \\ \rho_1: \mathbb{P}_S(\pi_{\mathcal{X}*}\mathscr{A}^{\otimes I}) \to \mathbb{P}_S^{d_1-1}, \\ \rho_2: \mathbb{P}_S(\pi_{\mathcal{X}*}\mathscr{A}^{\otimes I+1}) \to \mathbb{P}_S^{d_2-1} \text{ and} \\ \rho_3: \mathbb{P}_S(\pi_{\mathcal{X}*}\omega_{\mathcal{X}/S}^{[er]}) \to \mathbb{P}_S^{d_3-1} \text{ are} \\ \text{isomorphisms} \end{array} \right\} \Big/ \sim,$$

where $(f: (\mathcal{X},\mathscr{A}) \to \mathcal{C}, \rho_1, \rho_2, \rho_3) \sim (f': (\mathcal{X}',\mathscr{A}') \to \mathcal{C}', \rho_1', \rho_2', \rho_3')$ if and only if there exists an isomorphism $\alpha: (\mathcal{X},\mathscr{A}) \to (\mathcal{X}',\mathscr{A}')$ of $\mathscr{M}_{d,v,u,w,r}(S)$ (see the definition of $\mathfrak{Pol}$) such that the induced isomorphisms

$$\begin{aligned} \alpha_1&: \mathbb{P}_S(\pi_{\mathcal{X}*}\mathscr{A}^{\otimes I}) \to \mathbb{P}_S(\pi_{\mathcal{X}'*}\mathscr{A}'^{\otimes I}), \\ \alpha_2&: \mathbb{P}_S(\pi_{\mathcal{X}*}\mathscr{A}^{\otimes I+1}) \to \mathbb{P}_S(\pi_{\mathcal{X}'*}\mathscr{A}'^{\otimes I+1}), \\ \alpha_3&: \mathbb{P}_S(\pi_{\mathcal{X}*}\omega_{\mathcal{X}/S}^{[er]}) \to \mathbb{P}_S(\pi_{\mathcal{X}'*}\omega_{\mathcal{X}'/S}^{[er]}), \end{aligned}$$

satisfy $\rho_i' \circ \alpha_i = \rho_i$ for $i = 1, 2, 3$. Here, the structure morphisms $\mathcal{X} \to S$ and $\mathcal{X}' \to S$ are denoted by $\pi_{\mathcal{X}}$ and $\pi_{\mathcal{X}'}$ respectively, and the line bundle representing $\mathscr{A}$ is denoted by $\mathscr{A}$ by abuse of notation.

In particular, $N_{d_1,d_2,d_3,h}$ inherits the $\mathrm{PGL}(d_1) \times \mathrm{PGL}(d_2) \times \mathrm{PGL}(d_3)$ action on H.

Proof We first note that $\mathbb{P}_S(\pi_{\mathcal{X}*}\mathscr{A}^{\otimes I})$ and $\mathbb{P}_S(\pi_{\mathcal{X}*}\mathscr{A}^{\otimes I+1})$ are independent of a representative of $\mathscr{A}$; see the claim in the proof of Lemma 2.15. The well-definedness of $\mathfrak{H}$ follows from the fact that we can define the pullback of $(f, \rho_1, \rho_2, \rho_3) \in \mathfrak{H}(S)$ by any morphism $S' \to S$ by using [Hartshorne 1977, III, Theorem 12.11] and the condition (iv) of $\mathscr{M}_{d,v,u,w,r}$. Indeed, by the properties of I (see Corollary 3.8), we have

$$h^i(\mathcal{X}_{\bar{s}}, \mathcal{O}_{\mathcal{X}_{\bar{s}}}(I\mathscr{A}_{\bar{s}})) = h^i(\mathcal{X}_{\bar{s}}, \mathcal{O}_{\mathcal{X}_{\bar{s}}}((I+1)\mathscr{A}_{\bar{s}})) = 0$$

for every $i > 0$ and geometric point $\overline{s} \in S$. Thus,

$$\begin{aligned}\mathbb{P}_{S'}(\pi_{\mathcal{X}_{S'}*}\mathscr{A}_{S'}^{\otimes I}) &\cong \mathbb{P}_S(\pi_{\mathcal{X}*}\mathscr{A}^{\otimes I}) \times_S S',\\ \mathbb{P}_{S'}(\pi_{\mathcal{X}_{S'}*}\mathscr{A}_{S'}^{\otimes I+1}) &\cong \mathbb{P}_S(\pi_{\mathcal{X}*}\mathscr{A}^{\otimes I+1}) \times_S S',\\ \mathbb{P}_{S'}(\pi_{\mathcal{X}_{S'}*}\omega^{[er]}_{\mathcal{X}_{S'}/S'}) &\cong \mathbb{P}_S(\pi_{\mathcal{X}*}\omega^{[er]}_{\mathcal{X}/S}) \times_S S'.\end{aligned}$$

We will prove the proposition in several steps.

Step 1 In this step, we introduce a claim and give an explanation of the claim.

We will consider the following claim, which will be proved in Step 3.

Claim 1 *There exists a locally closed subscheme N of H such that a morphism $T \to H$ factors through $N \hookrightarrow H$ if and only if there exists a $\tilde{\pi}_T$-ample line bundle $\mathscr{A}'$ on $\mathcal{U}_T$ such that $\mathscr{A}'$ and $\tilde{\pi}_T : \mathcal{U}_T \to T$ satisfy the following.*

(a) *Any geometric fiber of $\tilde{\pi}_T$ is connected and normal.*

(b) *$p_{1,\overline{t}}$ and $p_{2,\overline{t}}$ are closed immersions for any geometric point $\overline{t} \in T$.*

(c) *$\mathscr{A}'^{\otimes I} \sim_T p_{1,T}^*\mathcal{O}_{\mathbb{P}_T^{d_1-1}}(1)$ and $\mathscr{A}'^{\otimes I+1} \sim_T p_{2,T}^*\mathcal{O}_{\mathbb{P}_T^{d_2-1}}(1)$.*

(d) *For any point $t \in T$, the morphisms $\mathcal{O}_T^{\oplus d_1} \to H^0(\mathcal{U}_t, \mathscr{A}_t'^{\otimes I})$ and $\mathcal{O}_T^{\oplus d_2} \to H^0(\mathcal{U}_t, \mathscr{A}_t'^{\otimes I+1})$ are surjective, $h^0(\mathcal{U}_t, \mathscr{A}_t'^{\otimes I}) = d_1$ and $h^0(\mathcal{U}_t, \mathscr{A}_t'^{\otimes I+1}) = d_2$.*

(e) *$\omega^{[er]}_{\mathcal{U}_T/T} \sim_T p_{3,T}^*\mathcal{O}(1)$.*

(f) *$\mathcal{U}_{\overline{t}}$ is a klt variety for any geometric point $\overline{t} \in T$.*

(g) *$\tilde{\pi}_{T*}\omega^{[ler]}_{\mathcal{U}_T/T} \to H^0(\mathcal{U}_t, \mathcal{O}_{\mathcal{U}_t}(lerK_{\mathcal{U}_t}))$ is surjective for any point $t \in T$ and any $l \in \mathbb{Z}_{>0}$.*

(h) *$\mathcal{O}_T^{\oplus d_3} \to H^0(\mathcal{U}_t, \mathcal{O}_{\mathcal{U}_t}(erK_{\mathcal{U}_t}))$ is surjective and $h^0(\mathcal{U}_t, \mathcal{O}_{\mathcal{U}_t}(erK_{\mathcal{U}_t})) = d_3$ for any point $t \in T$.*

(i) *$(\mathcal{U}_{\overline{t}}, 0, \mathscr{A}'_{\overline{t}}) \to \tilde{C}_{\overline{t}} \in \mathfrak{Z}_{d,v,u,w}$ for any geometric point $\overline{t} \in T$, where $\mathcal{U}_T \to \tilde{C}_T$ is the ample model of $\omega^{[er]}_{\mathcal{U}_T/T}$.*

Here, the morphism $\mathcal{O}_T^{\oplus d_1} \to H^0(\mathcal{U}_t, \mathscr{A}_t'^{\otimes I})$ in (d) is defined to be the composition

$$\mathcal{O}_T^{\oplus d_1} \longrightarrow \tilde{\pi}_{T*}p_{3,T}^*\mathcal{O}_{\mathbb{P}_T^{d_3-1}}(1) \longrightarrow H^0(\mathcal{U}_t, p_{3,t}^*\mathcal{O}_{\mathbb{P}^{d_3-1}}(1)) \xrightarrow{\cong} H^0(\mathcal{U}_t, \mathscr{A}_t'^{\otimes I}),$$

where the last isomorphism is induced by $\mathscr{A}'^{\otimes I} \sim_T p_{1,T}^*\mathcal{O}_{\mathbb{P}_T^{d_1-1}}(1)$ in (c), and the other morphisms in (d) and (h) are defined similarly.

We give a few words about the conditions (a)–(i). The roles of these conditions are as follows:

- (e), (g), and (i) are related to the conditions of $\mathscr{M}_{d,v,u,w,r}(T)$,
- (c), (d), (e) and (h) are utilized to prove the representability of $\mathfrak{H}$, and
- (a), (b) and (f) are extra and written just for the convenience of the proof.

More precisely, (a), (b) and (f) immediately follow from (d), (i), and the properties of I in Corollary 3.8, and there are the following correspondences:

- (e) implies (iii) of $\mathscr{M}_{d,v,u,w,r}(T)$,
- (g) corresponds to (iv) of $\mathscr{M}_{d,v,u,w,r}(T)$,
- (i) corresponds to (v) of $\mathscr{M}_{d,v,u,w,r}(T)$.

Thus, every morphism $(\mathcal{U}_T, \mathscr{A}') \to \widetilde{\mathcal{C}}_T$ satisfying (a)–(i) is an object of $\mathscr{M}_{d,v,u,w,r}(T)$.

Step 2 In this step, we prove Proposition 5.6 assuming the existence of N in Claim 1.

Let N be the scheme in Claim 1 and let $\mathcal{U}_N \subset \mathbb{P}_N^{d_1-1} \times_N \mathbb{P}_N^{d_2-1} \times_N \mathbb{P}_N^{d_3-1}$ be the universal subscheme. We fix $\widetilde{\mathscr{A}}_N$ as in Claim 1. By (c) and (e), we can find line bundles $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_3$ on N such that

$$p_{1,N}^*\mathcal{O}(1) \sim \widetilde{\pi}_N^*\mathcal{M}_1 \otimes \widetilde{\mathscr{A}}_N^{\otimes I}, \quad p_{2,N}^*\mathcal{O}(1) \sim \widetilde{\pi}_N^*\mathcal{M}_2 \otimes \widetilde{\mathscr{A}}_N^{\otimes I+1}, \quad p_{3,N}^*\mathcal{O}(1) \sim \widetilde{\pi}_N^*\mathcal{M}_3 \otimes \omega_{\mathcal{U}_N/N}^{[er]}.$$

By (d) and (h) and applying [Mumford 1966, Lecture 7, Corollary 2] to the natural morphisms

$$\begin{aligned}
\mathcal{O}_N^{\oplus d_1} &\to \widetilde{\pi}_{N*}p_{1,N}^*\mathcal{O}(1) \cong \mathcal{M}_1 \otimes \widetilde{\pi}_{N*}\widetilde{\mathscr{A}}_N^{\otimes I},\\
\mathcal{O}_N^{\oplus d_2} &\to \widetilde{\pi}_{N*}p_{2,N}^*\mathcal{O}(1) \cong \mathcal{M}_2 \otimes \widetilde{\pi}_{N*}\widetilde{\mathscr{A}}_N^{\otimes I+1},\\
\mathcal{O}_N^{\oplus d_3} &\to \widetilde{\pi}_{N*}p_{3,N}^*\mathcal{O}(1) \cong \mathcal{M}_3 \otimes \widetilde{\pi}_{N*}\omega_{\mathcal{U}_N/N}^{[er]},
\end{aligned}$$

we see that

$$\mathcal{O}_N^{\oplus d_1} \cong \mathcal{M}_1 \otimes \widetilde{\pi}_{N*}\widetilde{\mathscr{A}}_N^{\otimes I}, \quad \mathcal{O}_N^{\oplus d_2} \cong \mathcal{M}_2 \otimes \widetilde{\pi}_{N*}\widetilde{\mathscr{A}}_N^{\otimes I+1}, \quad \mathcal{O}_N^{\oplus d_3} \cong \mathcal{M}_3 \otimes \widetilde{\pi}_{N*}\omega_{\mathcal{U}_N/N}^{[er]}.$$

From these relations, we obtain isomorphisms

$$\begin{aligned}
\widetilde{\rho}_1\colon\ \mathbb{P}_N(\widetilde{\pi}_{N*}\widetilde{\mathscr{A}}_N^{\otimes I}) &\xrightarrow{\cong} \mathbb{P}_N^{d_1-1},\\
\widetilde{\rho}_2\colon \mathbb{P}_N(\widetilde{\pi}_{N*}\widetilde{\mathscr{A}}_N^{\otimes I+1}) &\xrightarrow{\cong} \mathbb{P}_N^{d_2-1},\\
\widetilde{\rho}_3\colon \mathbb{P}_N(\widetilde{\pi}_{N*}\omega_{\mathcal{U}_N/N}^{[er]}) &\xrightarrow{\cong} \mathbb{P}_N^{d_3-1}.
\end{aligned}$$

This fact and the universal property of N show that there exists an injective map

$$\eta(S)\colon \mathrm{Hom}(S, N) \hookrightarrow \mathfrak{H}(S)$$

which maps $\gamma\colon S \to N$ to $(f_S\colon (\mathcal{U}_S, \widetilde{\mathscr{A}}_S) \to \widetilde{\mathcal{C}}_S, \widetilde{\rho}_{1,S}, \widetilde{\rho}_{2,S}, \widetilde{\rho}_{3,S})$, where $\widetilde{\rho}_{i,S}$ is the base change of $\widetilde{\rho}_i$ by S. Therefore we obtain a morphism $\eta\colon \mathrm{Hom}(\bullet, N) \to \mathfrak{H}$.

It suffices to prove the surjectivity of η. In general, for two locally free sheaves $\mathcal{E}$ and $\mathcal{E}'$ on S with an S-isomorphism $g\colon \mathbb{P}_S(\mathcal{E}) \to \mathbb{P}_S(\mathcal{E}')$, we have $g^*\mathcal{O}_{\mathbb{P}_S(\mathcal{E}')}(1) \sim_S \mathcal{O}_{\mathbb{P}_S(\mathcal{E})}(1)$. Indeed, we put $\mathcal{F} := g^*\mathcal{O}_{\mathbb{P}_S(\mathcal{E}')}(1) \otimes \mathcal{O}_{\mathbb{P}_S(\mathcal{E})}(-1)$. Then $\mathcal{F}$ is locally trivial over S by [Mumford et al. 1994, Section 0.5(b)]. Thus, the pushforward of $\mathcal{F}$ to S is an invertible sheaf. From this fact and the global generation of $\mathcal{F}$ over S, we have $g^*\mathcal{O}_{\mathbb{P}_S(\mathcal{E}')}(1) \sim_S \mathcal{O}_{\mathbb{P}_S(\mathcal{E})}(1)$. By using this fact, for any object $(f\colon (\mathcal{X}, \mathscr{A}) \to \mathcal{C}, \rho_1, \rho_2, \rho_3)$ of $\mathfrak{H}(S)$ with the canonical morphisms

$$f_1\colon \mathcal{X} \to \mathbb{P}_S(\pi_{\mathcal{X}*}\mathscr{A}^{\otimes I}), \quad f_2\colon \mathcal{X} \to \mathbb{P}_S(\pi_{\mathcal{X}*}\mathscr{A}^{\otimes I+1}) \quad \text{and} \quad f_3\colon \mathcal{X} \to \mathbb{P}_S(\pi_{\mathcal{X}*}\omega_{\mathcal{X}/S}^{[er]}),$$

we have

(I) $\quad (\rho_1\circ f_1)^*\mathcal{O}(1)\sim_S \mathscr{A}^{\otimes I},\quad (\rho_2\circ f_2)^*\mathcal{O}(1)\sim_S \mathscr{A}^{\otimes I+1}\quad$ and $\quad (\rho_3\circ f_3)^*\mathcal{O}(1)\sim_S \omega^{[er]}_{\mathcal{X}/S}.$

By the properties of I in Corollary 3.8 and the condition (iv) of $\mathscr{M}_{d,v,u,w,r}$ with the aid of [Hartshorne 1977, III, Theorem 12.11], we see that the fibers of $\pi_{\mathcal{X}*}\mathscr{A}^{\otimes I}$, $\pi_{\mathcal{X}*}\mathscr{A}^{\otimes I+1}$ and $\pi_{\mathcal{X}*}\omega^{[er]}_{\mathcal{X}/S}$ coincide with $H^0(\mathcal{X}_{\bar{s}},\mathscr{A}_{\bar{s}}^{\otimes I})$, $H^0(\mathcal{X}_{\bar{s}},\mathscr{A}_{\bar{s}}^{\otimes I+1})$ and $H^0(\mathcal{X}_{\bar{s}},\mathcal{O}_{\mathcal{X}_{\bar{s}}}(erK_{\mathcal{X}_{\bar{s}}}))$, respectively, over every geometric point $\bar{s}\in S$. Then the three linear equivalences in (I) induce the surjective morphisms

(II) $\quad \mathcal{O}_S^{\oplus d_1}\to H^0(\mathcal{X}_s,\mathscr{A}_s^{\otimes I}),\quad \mathcal{O}_S^{\oplus d_2}\to H^0(\mathcal{X}_s,\mathscr{A}_s^{\otimes I+1})\quad$ and $\quad \mathcal{O}_S^{\oplus d_3}\to H^0(\mathcal{X}_s,\mathcal{O}_{\mathcal{X}_s}(erK_{\mathcal{X}_s}))$

for any point $s\in S$.

We set $p_i:=\rho_i\circ f_i$. By the properties of I in Corollary 3.8, $p_{1,\bar{s}}$ is a closed immersion for every geometric point $\bar{s}\in S$. Thus,

$$p_1\times p_2\times p_3\colon \mathcal{X}\hookrightarrow \mathbb{P}_S^{d_1-1}\times_S\mathbb{P}_S^{d_2-1}\times_S\mathbb{P}_S^{d_3-1}$$

is a closed immersion. The morphism $\gamma\colon S\to H$ corresponding to $p_1\times p_2\times p_3$ factors through N since (I) (resp. (II)) corresponds to (c) and (e) (resp. (d) and (h)). Then it immediately follows that $\eta(S)$ is surjective and hence η is an isomorphism.

Therefore, $\mathfrak{H}$ is represented by N and hence Proposition 5.6 holds if Claim 1 holds. We finish this step.

Step 3 In this final step, we prove Claim 1. To prove Claim 1, it suffices to check that (a)–(i) are locally closed conditions.

We first deal with (a) and (b). By [Grothendieck 1966, Théorème (12.2.1) and (12.2.4)], the subset

$$U_1:=\{s\in H\mid \mathcal{U}_s \text{ is geometrically connected and geometrically normal}\}$$

is open. By [Görtz and Wedhorn 2010, Proposition 12.93], the subset

$$U_2:=\{s\in U_1\mid p_{1,s}\colon\mathcal{U}_s\to\mathbb{P}_s^{d_1-1} \text{ and } p_{2,s}\colon\mathcal{U}_s\to\mathbb{P}_s^{d_2-1} \text{ are closed immersions}\}\subset U_1$$

is also open.

Next, we treat (c). We put

$$\widetilde{\mathscr{A}}=p^*_{2,U_2}\mathcal{O}_{\mathbb{P}^{d_2-1}_{U_2}}(1)\otimes p^*_{1,U_2}\mathcal{O}_{\mathbb{P}^{d_1-1}_{U_2}}(-1).$$

Then the condition (c) implies that

$$\mathscr{A}'\sim_T p^*_{1,T}\mathcal{O}_{\mathbb{P}^{d_1-1}_T}(-1)\otimes p^*_{2,T}\mathcal{O}_{\mathbb{P}^{d_2-1}_T}(1)=\widetilde{\mathscr{A}}_T.$$

Hence, the existence of $\mathscr{A}'$ satisfying (c) is equivalent to the $\tilde{\pi}_T$-ampleness of $\widetilde{\mathscr{A}}_T$ and the relations

$$\widetilde{\mathscr{A}}_T^{\otimes I}\sim_T p^*_{1,T}\mathcal{O}_{\mathbb{P}^{d_1-1}_T}(1)\quad\text{and}\quad \widetilde{\mathscr{A}}_T^{\otimes I+1}\sim_T p^*_{2,T}\mathcal{O}_{\mathbb{P}^{d_2-1}_T}(1).$$

By Corollary 2.20, there exists a locally closed subscheme

$$U_3 \subset U_2$$

such that a morphism $T \to U_2$ factors through $U_3 \hookrightarrow U_2$ if and only if the relations

$$\widetilde{\mathscr{A}}_T^{\otimes I} \sim_T p_{1,T}^* \mathcal{O}_{\mathbb{P}_T^{d_1-1}}(1) \quad \text{and} \quad \widetilde{\mathscr{A}}_T^{\otimes I+1} \sim_T p_{2,T}^* \mathcal{O}_{\mathbb{P}_T^{d_2-1}}(1)$$

hold true. Since p_{1,U_3} is a closed immersion, $\widetilde{\mathscr{A}}_{U_3}$ is $\tilde{\pi}_{U_3}$-ample.

For (d), set

$$U_4 := \left\{ s \in U_3 \;\middle|\; \begin{array}{l} \bullet\ \ h^0(\mathcal{U}_s, \widetilde{\mathscr{A}}_s^{\otimes I}) = d_1, \\ \bullet\ \ h^0(\mathcal{U}_s, \widetilde{\mathscr{A}}_s^{\otimes I+1}) = d_2, \text{ and} \\ \bullet\ \ \text{both } \mathcal{O}_{U_3}^{\oplus d_1} \to H^0(\mathcal{U}_s, \widetilde{\mathscr{A}}_s^{\otimes I}) \text{ and } \mathcal{O}_{U_3}^{\oplus d_2} \to H^0(\mathcal{U}_s, \widetilde{\mathscr{A}}_s^{\otimes I+1}) \text{ are surjective.} \end{array} \right\}$$

Then U_4 is open. Indeed, pick a point $s \in U_4$. We take a line bundle $\mathcal{M}$ on U_3 such that

$$\tilde{\pi}_{U_3}^* \mathcal{M} \otimes \widetilde{\mathscr{A}}_{U_3}^{\otimes I} \sim p_{1,\mathcal{U}_{U_3}}^* \mathcal{O}_{\mathbb{P}}^{d_1-1}(1).$$

By the third condition in U_4 and the construction of $\mathcal{O}_{U_3}^{\oplus d_1} \to H^0(\mathcal{U}_s, \widetilde{\mathscr{A}}_s^{\otimes I})$, we have that

$$\tilde{\pi}_{U_3 *} p_{1,U_3}^* \mathcal{O}(1) \cong \mathcal{M} \otimes_{\mathcal{O}_{U_3}} \tilde{\pi}_{U_3 *} \widetilde{\mathscr{A}}_{U_3}^{\otimes I} \to H^0(\mathcal{U}_s, \widetilde{\mathscr{A}}_s^{\otimes I})$$

is surjective. By [Hartshorne 1977, III, Theorem 12.11], $\mathcal{M} \otimes_{\mathcal{O}_{U_3}} \tilde{\pi}_{U_3 *} \widetilde{\mathscr{A}}_{U_3}^{\otimes I}$ is locally free near s and

$$\mathcal{M} \otimes_{\mathcal{O}_{U_3}} \tilde{\pi}_{U_3 *} \widetilde{\mathscr{A}}_{U_3}^{\otimes I} \otimes_{\mathcal{O}_{U_3}} \mathcal{O}_{U_3,s'}/\mathfrak{m}_{s'} \to H^0(\mathcal{U}_{s'}, \widetilde{\mathscr{A}}_{s'}^{\otimes I})$$

is an isomorphism for every point $s' \in U_3$ on some neighborhood of s, where $\mathfrak{m}_{s'}$ is the maximal ideal of $\mathcal{O}_{U_3,s'}$. Therefore we have that $h^0(\mathcal{U}_{s'}, \widetilde{\mathscr{A}}_{s'}^{\otimes I}) = d_1$ for every point s' on some neighborhood of s, and the third condition on U_4 implies that $\mathcal{O}_{U_3}^{\oplus d_1} \to \mathcal{M} \otimes_{\mathcal{O}_{U_3}} \tilde{\pi}_{U_3 *} \widetilde{\mathscr{A}}_{U_3}^{\otimes I}$ is surjective at s. Then the morphism

$$\mathcal{O}_{U_3}^{\oplus d_1} \to \tilde{\pi}_{U_3 *} p_{1,U_3}^* \mathcal{O}(1) \cong \mathcal{M} \otimes_{\mathcal{O}_{U_3}} \tilde{\pi}_{U_3 *} \widetilde{\mathscr{A}}_{U_3}^{\otimes I} \to H^0(\mathcal{U}_{s'}, \widetilde{\mathscr{A}}_{s'}^{\otimes I})$$

is surjective for every point s' on some neighborhood of s. By the same argument, we see that $h^0(\mathcal{U}_{s'}, \widetilde{\mathscr{A}}_{s'}^{\otimes I}) = d_2$ and $\mathcal{O}_{U_3}^{\oplus d_2} \to H^0(\mathcal{U}_{s'}, \widetilde{\mathscr{A}}_{s'}^{\otimes I+1})$ is surjective for every point s' on some neighborhood of s. In this way, if $s \in U_4$ then a neighborhood of s is contained in U_4, which implies the openness of U_4.

In this paragraph, we discuss the conditions (e) and (f). By Corollary 2.20, we may find a locally closed subscheme

$$U_5 \subset U_4$$

such that a morphism $T \to U_4$ factors through U_5 if and only if $\omega_{\mathcal{U}_T/T}^{[er]} \sim_T p_{3,T}^* \mathcal{O}(1)$. Furthermore, the subset

$$U_6 = \{t \in U_5 \mid \mathcal{U}_{\bar{t}} \text{ is klt}\}$$

is open since $\omega_{\mathcal{U}_T/T}^{[r]}$ is a line bundle; see [Kollár 2013, Corollary 4.10].

We will discuss condition (g) for three paragraphs. We fix an $l \in \mathbb{Z}_{>0}$, and we will discuss the surjectivity of $\tilde{\pi}_{T*}\omega^{[ler]}_{\mathcal{U}_T/T} \to H^0(\mathcal{U}_t, \mathcal{O}_{\mathcal{U}_t}(lerK_{\mathcal{U}_t}))$ for every $t \in T$, which we call condition $(\mathrm{g})_l$. For any Noetherian affine scheme $\bar{U}$ with a morphism $\bar{U} \to U_8$, we define functors $F^0_{\bar{U}}$ and $F^1_{\bar{U}}$ that send an affine scheme U' over $\bar{U}$ to $\tilde{\pi}_{U'*}\omega^{[ler]}_{\mathcal{U}_{U'}/U'}$ and $R^1\tilde{\pi}_{U'*}\omega^{[ler]}_{\mathcal{U}_{U'}/U'}$, respectively. These are the same functors as discussed in [Hartshorne 1977, III, Section 12], and the functor $F^0_{\bar{U}}$ is always left exact by the flatness of $\tilde{\pi}_{U_6*}\omega^{[ler]}_{\mathcal{U}_{U_6}/U_6}$. Pick an affine open subset $U \subset U_6$. We pick a Grothendieck complex $(K^\bullet, d^\bullet)$ for $\omega^{[ler]}_{\mathcal{U}_U/U}$. This is the same complex as in [Mumford 2008, Section 5, Lemma 1]; see also [Hartshorne 1977, III, Proposition 12.2]. We define a coherent sheaf

$$W^1 := \mathrm{Coker}\,(d^0 : K^0 \to K^1)$$

on U. For any affine morphism $g\colon T \to U$, the pullback of the complex $(g^*K^\bullet, g^*d^\bullet)$ is a Grothendieck complex with respect to $\omega^{[ler]}_{\mathcal{U}_T/T}$ and

$$g^*W^1 = \mathrm{Coker}\,(g^*d^0\colon g^*K^0 \to g^*K^1).$$

By applying Theorem 4.8 and [Mumford 2008, Section 5, Corollary 2] to the normalization of U_6, we see that $\tilde{\pi}_{U_{6,\mathrm{red}}}$ satisfies $(\mathrm{g})_l$. By [Hartshorne 1977, III, Corollary 12.6 and Proposition 12,10], $F^0_{U_{\mathrm{red}}}$ is exact and hence $F^1_{U_{\mathrm{red}}}$ is left exact. By [Hartshorne 1977, III, Proposition 12,4], $W^1 \otimes_{\mathcal{O}_U} \mathcal{O}_{U_{\mathrm{red}}}$ is flat for any choice of $(K^\bullet, d^\bullet)$. By this fact and the flattening stratification for W^1, we get a closed subscheme $Z_U \subset U$ such that a morphism $g\colon T \to U$ factors through $Z_U \hookrightarrow U$ if and only if g^*W^1 is flat, which is equivalent to the exactness of F^0_T by [Hartshorne 1977, III, Proposition 12.4]. From this, we can check that Z_U is independent of the choice of $(K^\bullet, d^\bullet)$ and hence $Z_U|_{U'} = Z_{U'}$ for any affine open embedding $U' \hookrightarrow U$. By these facts, we can construct a closed subscheme

$$U_7^{(l)} \subset U_6$$

by gluing all Z_U for affine open subsets $U \subset U_6$.

By construction, a morphism $g'\colon T' \to U_6$ factors through $U_7^{(l)} \hookrightarrow U_6$ if and only if $g'^*_V W^1$ is flat for any affine open subsets $V \subset T'$ and $U \subset U_6$ with the induced morphism $g'_V\colon V \to U$. By [Hartshorne 1977, III, Proposition 12.4], $g'^*_V W^1$ is flat if and only if F^1_V is left exact. By recalling that F^0_V is always left exact, we see that F^1_V is left exact if and only if F^0_V is exact. Thus, $g'\colon T' \to U_6$ factors through $U_7^{(l)} \hookrightarrow U_6$ if and only if F^0_V is exact for any affine open subset $V \subset T'$. By the argument of cohomology and base change [Hartshorne 1977, III, Proposition 12.5, Corollary 12.6 and Proposition 12.10], the exactness of F^0_V for every $V \subset T'$ is equivalent to the condition $(\mathrm{g})_l$. In this way, $g'\colon T' \to U_6$ factors through $U_7^{(l)} \hookrightarrow U_6$ if and only if the morphism $\tilde{\pi}_{T'}\colon \mathcal{U}_{T'} \to T'$ satisfies $(\mathrm{g})_l$.

By the above argument, we have a sequence of closed subschemes of U_6

$$U_6 \supset U_7^{(1)} \supset U_7^{(1)} \cap U_7^{(2)} := U_7^{(1)} \times_{U_6} U_7^{(2)} \supset \cdots,$$

and the Noetherian property of U_6 implies that the above sequence is stationary and

$$U_7 := \bigcap_{l \in \mathbb{Z}_{>0}} U_7^{(l)} \subset U_6$$

is well-defined as a closed subscheme. By the construction of $U_7^{(l)}$, a morphism $T \to U_7$ factors through $U_7 \hookrightarrow U_6$ if and only if $\mathcal{U}_T \to T$ satisfies (g). We have finished discussing (g).

By Theorem 4.8 and applying the same argument as in the construction of U_4, we see that

$$U_8 := \{s \in U_7 \mid \mathcal{O}_{U_7}^{\oplus d_3} \to H^0(\mathcal{U}_s, \mathcal{O}_{\mathcal{U}_s}(erK_{\mathcal{U}_s})) \text{ is surjective, } h^0(\mathcal{U}_{\bar{s}}, \mathcal{O}_{\mathcal{U}_{\bar{s}}}(erK_{\mathcal{U}_{\bar{s}}})) = d_3\}$$

is open. This corresponds to (h).

Finally, we discuss the condition (i). We put

$$\widetilde{\mathcal{C}} := \mathbf{Proj}_{U_8}\biggl(\bigoplus_{l \geq 0} \widetilde{\pi}_{U_8 *}\omega_{\mathcal{U}_{U_8}/U_8}^{[ler]}\biggr).$$

Let $f : \mathcal{U}_{U_8} \to \widetilde{\mathcal{C}}$ be the induced morphism. Now the sheaf $\omega_{\mathcal{U}_T/T}^{[er]}$ is $\widetilde{\pi}_T$-semiample by the construction of U_5. Furthermore,

$$\widetilde{\mathcal{C}}_T = \mathbf{Proj}_T\biggl(\bigoplus_{l \geq 0} \widetilde{\pi}_{T*}\omega_{\mathcal{U}_T/T}^{[ler]}\biggr)$$

for any morphism $T \to U_8$ by the condition (g) and [Hartshorne 1977, Theorem 12.11]. Thus, $f_{\bar{s}} : \mathcal{U}_{\bar{s}} \to \widetilde{\mathcal{C}}_{\bar{s}}$ is the contraction induced by $eK_{\mathcal{U}_{\bar{s}}}$ for any geometric point $\bar{s} \in U_8$. We consider the set

$$U_9 := \left\{ s \in U_8 \;\middle|\; \begin{array}{l} (p_{3,\bar{s}}^*\mathcal{O}(1)^2 \cdot \widetilde{\mathscr{A}}_{\bar{s}}^{d-2}) = 0,\ (p_{3,\bar{s}}^*\mathcal{O}(1) \cdot \widetilde{\mathscr{A}}_{\bar{s}}^{d-1}) = eruv, \\ \mathrm{vol}(\widetilde{\mathscr{A}}_{\bar{s}}) \leq w \text{ and } \mathrm{Ivol}(erK_{\mathcal{U}_{\bar{s}}}) = eru. \end{array} \right\}$$

We note that if the Iitaka dimension of $eK_{\mathcal{U}_{\bar{s}}}$ is one, we have $\mathrm{Ivol}(erK_{\mathcal{U}_{\bar{s}}}) = r \cdot \mathrm{Ivol}(eK_{\mathcal{U}_{\bar{s}}})$. This fact and (e) show that a point $s \in U_8$ is contained U_9 if and only if $(\mathcal{U}_{\bar{s}}, \widetilde{\mathscr{A}}_{\bar{s}}) \to \widetilde{\mathcal{C}}_{\bar{s}}$ satisfies (ii)–(iv) of $\mathfrak{Z}_{d,v,u,w}$. We will check that U_9 is open. By applying Theorem 4.8 to the normalization of U_8, we see that the function

$$U_8 \ni s \mapsto h^0(\mathcal{U}_{\bar{s}}, \mathcal{O}_{\mathcal{U}_{\bar{s}}}(emrK_{\mathcal{U}_{\bar{s}}}))$$

is locally constant for every $m \in \mathbb{Z}_{>0}$. We also see that

$$U_8 \ni s \mapsto \bigl((p_{3,s}^*\mathcal{O}(1)^2 \cdot \widetilde{\mathscr{A}}_s^{d-2}), (p_{3,s}^*\mathcal{O}(1) \cdot \widetilde{\mathscr{A}}_s^{d-1}), \mathrm{vol}(\widetilde{\mathscr{A}}_s)\bigr) \in \mathbb{Q}^3$$

is locally constant by the flatness. Therefore, we see that U_9 is open. Now it suffices to show the uniform adiabatic K-stability of $f_{\bar{s}}$ for any geometric point $\bar{s} \in U_9$. If $u > 0$, every $f_{\bar{s}}$ is uniformly adiabatically K-stable and hence we may set $N := U_9$. If $u < 0$, then we apply Theorem 4.2 and Example 2.28 to the normalization U_9^ν of U_9 and $\mathcal{U}_{U_9^\nu} \to \widetilde{\mathcal{C}}_{U_9^\nu}$, and we obtain an open subset

$$W \subset U_9^\nu$$

such that $s \in U_9^\nu$ is contained in W if and only if $f_{\bar{s}}$ is uniformly adiabatically K-stable. Let $\nu : U_9^\nu \to U_9$ be the morphism of the normalization. Since ν is surjective and closed and $W = \nu^{-1}(\nu(W))$, the set

$$N = \nu(W)$$

is open. Moreover, a geometric point $\bar{s} \in U_9$ is a point of N if and only if $f_{\bar{s}}$ is uniformly adiabatically K-stable. Thus, we have finished discussing the condition (i).

By the above argument, a morphism $T \to H$ factors through $N \hookrightarrow H$, if and only if there exists $\mathscr{A}'$ as in Claim 1 such that $\tilde{\pi}_T : \mathcal{U}_T \to T$ and $\mathscr{A}'$ satisfy (a)–(i). We have finished the proof of Claim 1.

This completes the proof of Proposition 5.6. □

Now we are ready to prove Theorem 5.1.

Proof of Theorem 5.1 By Lemma 5.3, $\mathscr{M}_{d,v,u,w,r}$ is a stack. We have

$$\mathscr{M}_{d,v,u,w,r} = \bigsqcup_{d_1,\,d_2,\,d_3,\,h} \mathscr{M}_{d_1,d_2,d_3,h}$$

as stacks. Thanks to Lemma 5.4, it suffices to check that $\mathscr{M}_{d_1,d_2,d_3,h}$ is a separated Deligne–Mumford stack of finite type over $\Bbbk$ for the fixed d_1, d_2, d_3 and h.

Fix d_1, d_2, d_3 and h. We put $N := N_{d_1,d_2,d_3,h}$, where $N_{d_1,d_2,d_3,h}$ is in Proposition 5.6, and let $\pi_N : (\mathcal{U}_N, \widetilde{\mathscr{A}_N}) \to N$ be the universal family in Proposition 5.6. We check that

$$\mathscr{M}_{d_1,d_2,d_3,h} \cong [N/\mathrm{PGL}(d_1) \times \mathrm{PGL}(d_2) \times \mathrm{PGL}(d_3)].$$

We first construct a morphism from $\mathscr{M}_{d_1,d_2,d_3,h}$ to $[N/\mathrm{PGL}(d_1) \times \mathrm{PGL}(d_2) \times \mathrm{PGL}(d_3)]$. By regarding N and $\mathscr{M}_{d_1,d_2,d_3,h}$ as stacks and using $(\mathcal{U}_N, \widetilde{\mathscr{A}_N}) \to N$, we get a morphism $N \to \mathscr{M}_{d_1,d_2,d_3,h}$ between stacks. Take a scheme S and $g : (\mathcal{X}, \mathscr{A}) \to \mathcal{C} \in \mathscr{M}_{d_1,d_2,d_3,h}(S)$. By the 2-Yoneda lemma [Olsson 2016, Proposition 3.2.2] and regarding S and $\mathscr{M}_{d_1,d_2,d_3,h}$ as stacks, we can find a morphism $S \to \mathscr{M}_{d_1,d_2,d_3,h}$ that corresponds to $g : (\mathcal{X}, \mathscr{A}) \to \mathcal{C}$. For any étale covering $S' \to S$ such that the pullback of $\mathscr{A}$ is represented by a $\pi_{\mathcal{X}_{S'}}$-ample line bundle $\mathscr{A}'$, the definition of $\mathscr{M}_{d,v,u,w,r}$ and [Hartshorne 1977, III, Theorem 12.11] imply that $\pi_{\mathcal{X}_{S'}*}\mathscr{A}'^{\otimes I}$, $\pi_{\mathcal{X}_{S'}*}\mathscr{A}'^{\otimes I+1}$ and $\pi_{\mathcal{X}_{S'}*}\omega^{[er]}_{\mathcal{X}_{S'}/S'}$ are locally free sheaves of ranks d_1, d_2 and d_3, respectively. Since N is the scheme representing $\mathfrak{H}$ (Proposition 5.6), $S' \times_{\mathscr{M}_{d_1,d_2,d_3,h}} N$ is represented by

$$\begin{aligned} \mathrm{V}_{(\mathcal{X}_{S'},\mathscr{A}_{S'})} := \mathrm{Isom}_{S'}(\mathbb{P}_{S'}(\pi_{\mathcal{X}_{S'}*}\mathscr{A}'^{\otimes I}), \mathbb{P}^{d_1-1}_{S'}) \times_{S'} \mathrm{Isom}_{S'}(\mathbb{P}_{S'}(\pi_{\mathcal{X}_{S'}*}\mathscr{A}'^{\otimes I+1}), \mathbb{P}^{d_2-1}_{S'}) \\ \times_{S'} \mathrm{Isom}_{S'}(\mathbb{P}_{S'}(\pi_{\mathcal{X}_{S'}*}\omega^{[er]}_{\mathcal{X}_{S'}/S'}), \mathbb{P}^{d_3-1}_{S'}). \end{aligned}$$

Thus, we can think $S' \times_{\mathscr{M}_{d_1,d_2,d_3,h}} N$ of a principal $\mathrm{PGL}(d_1) \times \mathrm{PGL}(d_2) \times \mathrm{PGL}(d_3)$-bundle over S'. In particular, $S' \times_{\mathscr{M}_{d_1,d_2,d_3,h}} N$ is affine over S'. Hence, $S \times_{\mathscr{M}_{d_1,d_2,d_3,h}} N$ is represented by an affine scheme over S [Olsson 2016, Proposition 4.4.9]. Then $S \times_{\mathscr{M}_{d_1,d_2,d_3,h}} N$ is a principal $\mathrm{PGL}(d_1) \times \mathrm{PGL}(d_2) \times \mathrm{PGL}(d_3)$-bundle over S [Fantechi et al. 2005, Proposition 2.36], and the natural morphism $S \times_{\mathscr{M}_{d_1,d_2,d_3,h}} N \to N$ is $\mathrm{PGL}(d_1) \times \mathrm{PGL}(d_2) \times \mathrm{PGL}(d_3)$-equivariant by Proposition 5.6. For each scheme S, by considering the map

$$(S \to \mathscr{M}_{d_1,d_2,d_3,h}) \mapsto (S \times_{\mathscr{M}_{d_1,d_2,d_3,h}} N \to N)$$

and using the 2-Yoneda lemma on the left-hand side, we obtain a morphism

$$\xi : \mathscr{M}_{d_1,d_2,d_3,h} \to [N/\mathrm{PGL}(d_1) \times \mathrm{PGL}(d_2) \times \mathrm{PGL}(d_3)]$$

between stacks.

In this paragraph we prove that ξ is an isomorphism. By [Olsson 2016, Proposition 3.1.10], it suffices to show the full faithfulness and the essential surjectivity of

$$\xi(S) := \xi|_{\mathscr{M}_{d_1,d_2,d_3,h}(S)} : \mathscr{M}_{d_1,d_2,d_3,h}(S) \to [N/\mathrm{PGL}(d_1)\times\mathrm{PGL}(d_2)\times\mathrm{PGL}(d_3)](S)$$

for a fixed scheme S. To prove the full faithfulness, we pick objects $g\colon(\mathcal{X},\mathscr{A})\to\mathcal{C}$ and $g'\colon(\mathcal{X}',\mathscr{A}')\to\mathcal{C}'$ of $\mathscr{M}_{d_1,d_2,d_3,h}(S)$. Taking an étale covering of S, we may assume that $\mathscr{A}$ and $\mathscr{A}'$ are line bundles. By construction, $\xi(S)$ defines a map

$$\begin{aligned}\mathrm{Isom}_S((\mathcal{X},\mathscr{A}),(\mathcal{X}',\mathscr{A}')) &\to \mathrm{Hom}(\mathrm{V}_{(\mathcal{X},\mathscr{A})},\mathrm{V}_{(\mathcal{X}',\mathscr{A}')}),\\ &\phi\mapsto \big((\rho_1,\rho_2,\rho_3)\mapsto(\rho_1\circ\phi^{-1},\rho_2\circ\phi^{-1},\rho_3\circ\phi^{-1})\big).\end{aligned}$$

From this, the full faithfulness of $\xi(S)$ follows. For the essential surjectivity, we pick any object $\alpha\colon\mathcal{P}\to N$ of $[N/\mathrm{PGL}(d_1)\times\mathrm{PGL}(d_2)\times\mathrm{PGL}(d_3)](S)$. Here, $\mathcal{P}$ is a principal $\mathrm{PGL}(d_1)\times\mathrm{PGL}(d_2)\times\mathrm{PGL}(d_3)$-bundle over S and α is $\mathrm{PGL}(d_1)\times\mathrm{PGL}(d_2)\times\mathrm{PGL}(d_3)$-equivariant. By [Olsson 2016, Corollary 1.3.10], there exists an étale covering $\beta\colon S'\to S$ such that $S'\times_S\mathcal{P}$ has a section $S'\to S'\times_S\mathcal{P}$. Let $\sigma\colon S'\to S'\times_S\mathcal{P}\to N$ be the composition of the section and the natural morphism, and let $(\mathcal{U}_{S'},\widetilde{\mathscr{A}}_{S'})\to\widetilde{\mathcal{C}}_{S'}$ be an object of $\mathscr{M}_{d_1,d_2,d_3,h}(S')$ defined by the pullback of the universal family $(\mathcal{U}_N,\widetilde{\mathscr{A}}_N)\to\widetilde{\mathcal{C}}_N$ via σ. Then there exists an $S'\times N$-isomorphism $S'\times_S\mathcal{P}\to S'\times_{\mathscr{M}_{d_1,d_2,d_3,h}} N$, where $S'\to\mathscr{M}_{d_1,d_2,d_3,h}$ is the morphism corresponding to $(\mathcal{U}_{S'},\widetilde{\mathscr{A}}_{S'})\to\widetilde{\mathcal{C}}_{S'}$. By this discussion and the full faithfulness, $\xi(S)$ is essentially surjective. Thus ξ is an isomorphism. In this way, $\mathscr{M}_{d_1,d_2,d_3,h}$ is categorically equivalent to $[N/\mathrm{PGL}(d_1)\times\mathrm{PGL}(d_2)\times\mathrm{PGL}(d_3)]$. Thus, $\mathscr{M}_{d_1,d_2,d_3,h}$ is an Artin stack of finite type over $\Bbbk$.

In the rest of the proof, we will prove that $\mathscr{M}_{d_1,d_2,d_3,h}$ is a separated Deligne–Mumford stack with a coarse moduli space. By Theorem 2.17, it suffices to prove that the diagonal morphism is finite. By Corollaries 4.5 and 4.7, we only need to prove that the diagonal morphism is proper. For any scheme S and $g_i\colon(\mathcal{X}_i,\mathscr{A}_i)\to\mathcal{C}_i\in\mathscr{M}_{d_1,d_2,d_3,h}(S)$ for $i=1,2$, it suffices to show that the scheme $\mathscr{I}:=\mathrm{Isom}_S((\mathcal{X}_1,\mathscr{A}_1),(\mathcal{X}_2,\mathscr{A}_2))$ is proper over S. By [Fantechi et al. 2005, Proposition 2.36] and taking an étale covering of S, we may assume that $\mathscr{A}_1$ and $\mathscr{A}_2$ are represented by line bundles. By abuse of notation, we denote them by $\mathscr{A}_1$ and $\mathscr{A}_2$, respectively. Then $\mathscr{I}$ is locally quasiprojective over S (Section 2.2).

Since the problem is local, by shrinking S, we may assume that $\mathscr{I}$ is quasiprojective over S and there exist morphisms $\gamma_i\colon S\to N$ for $i=1,2$ such that $\gamma_i^*(f_N\colon(\mathcal{U}_N,\widetilde{\mathscr{A}}_N)\to\widetilde{\mathcal{C}}_N)=g_i$, where f_N is the canonical morphism of the ample model of $\omega^{[er]}_{\mathcal{U}_T/T}$. By the morphism $S\to N\times N$ naturally induced from γ_1 and γ_2, we obtain

$$\mathscr{I}=\mathrm{Isom}_{N\times N}\big((\mathcal{U}_1,\mathscr{A}_1),(\mathcal{U}_2,\mathscr{A}_2)\big)\times_{N\times N}S,$$

where $\mathcal{U}_1$ (resp. $\mathcal{U}_2$) is the base change $\mathcal{U}_N\times_N(N\times N)$ by the first (resp. second) projection $N\times N\to N$, and $\mathscr{A}_1$ (resp. $\mathscr{A}_2$) is the pullback of $\widetilde{\mathscr{A}}$. From this, we may replace S by $N\times N$. Hence, we may assume that S is of finite type over $\Bbbk$. By [Görtz and Wedhorn 2010, Corollary 13.101], it suffices to prove that the natural morphism $\mathscr{I}\times_S(S\times\mathbb{A}^n)\to S\times\mathbb{A}^n$, denoted by φ, is a closed map for every $n\in\mathbb{Z}_{>0}$.

We pick a closed subset $Z \subset \mathscr{I} \times \mathbb{A}^n$. Note that $\varphi(Z)$ is constructible. By Lemma 2.2, to prove the closedness of $\varphi(Z)$ it suffices to show that for any morphism $q\colon C \to \overline{\varphi(Z)}$ from a curve C such that $q^{-1}(\varphi(Z))$ is dense in C, we have $q(C) \subset \varphi(Z)$. Consider the scheme

$$\mathscr{J} := (\mathscr{I} \times \mathbb{A}^n) \times_{S\times\mathbb{A}^n} C \cong \mathscr{I} \times_S C.$$

Since $\mathscr{J}$ is quasiprojective, we can find a curve D with a morphism $D \to \mathscr{J}$ such that the composition $D \to \mathscr{J} \to C$ is a dominant morphism. By considering a compactification of D over C, we obtain a curve $\overline{D}$ such that $\overline{D} \to C$ is surjective and D is an open subset of $\overline{D}$. Then $\mathscr{I} \times_S D$ is an open subset of $\mathscr{I} \times_S \overline{D}$, and $\mathscr{I} \times_S D \cong \mathscr{J} \times_C D \to D$ has a section $D' \subset \mathscr{I} \times_S D$. This section can be extended to a section $\overline{D}'$ of $(\mathscr{I} \times \mathbb{A}^n) \times_{S\times\mathbb{A}^n} \overline{D} \cong \mathscr{I} \times_S \overline{D} \to \overline{D}$. Indeed, this fact follows from Proposition 4.4 if $u > 0$. If $u < 0$, then $\mathcal{C}_{\overline{D}}$ of any object $f\colon (X, A) \to \mathcal{C}_{\overline{D}}$ of $\mathscr{M}_{d_1,d_2,d_3,h}(\overline{D})$ is a $\mathbb{P}^1$-bundle over $\overline{D}$ by Tsen's theorem; see [Hartshorne 1977, V Section 2]. Thus, we can apply Theorem 4.6 and obtain $\overline{D}'$. Then

$$p(C) = \mathrm{Im}(\overline{D} \to C \to S \times \mathbb{A}^n) = \mathrm{Im}\big(\overline{D}' \hookrightarrow (\mathscr{I} \times \mathbb{A}^n) \times_{S\times\mathbb{A}^n} \overline{D} \to \mathscr{I} \times \mathbb{A}^n \xrightarrow{\varphi} S \times \mathbb{A}^n\big) \subset \varphi(Z).$$

Hence, φ is a closed map, which implies that the diagonal morphism is proper. It follows from this that $\mathscr{M}_{d_1,d_2,d_3,h}$ is separated.

By the above argument, $\mathscr{M}_{d_1,d_2,d_3,h}$ is a separated Deligne–Mumford stack of finite type over $\Bbbk$ with the coarse moduli space. □

Now we prove Theorem 1.3. More specifically, we prove that $\mathscr{M}_{d,v,u,r}$ is an open and closed substack of $\mathscr{M}_{d,v,u,r,w}$ for some $w \in \mathbb{Z}_{>0}$.

Proof of Theorem 1.3 We will freely use the notation $\mathfrak{Z}_{d,v,u}$ and $\mathfrak{Z}_{d,v,u,w}$ in Section 1.

We first check that for any scheme S and $f\colon (\mathcal{X}, \mathscr{A}) \to \mathcal{C} \in \mathscr{M}_{d,v,u,r}(S)$, there exists $\mathscr{L} \in \mathrm{Pic}_{\mathcal{C}/S}(S)$ such that $\mathscr{L}_{\overline{s}} = \mathcal{O}_{\mathbb{P}^1}(1)$ for any geometric point $\overline{s} \in S$. Since $u < 0$, the morphism $\mathcal{C} \to S$ is a smooth morphism whose geometric fibers are $\mathbb{P}^1$; see Theorem 2.6. Thus, there is an étale covering $S' \to S$ such that $\mathcal{C} \times_S S' \cong \mathbb{P}^1_{S'}$; see the first paragraph of the proof of Theorem 4.2. Then $\mathcal{O}_{\mathbb{P}^1_{S'}}(1)$ satisfies

$$p^*_{1,\mathbb{P}^1_{S'}} \mathcal{O}_{\mathbb{P}^1_{S'}}(1) \cong p^*_{2,\mathbb{P}^1_{S'}} \mathcal{O}_{\mathbb{P}^1_{S'}}(1),$$

where $p_1\colon S' \times_S S' \to S'$ (resp. $p_2\colon S' \times_S S' \to S'$) is the first (resp. second) projection. Since $\mathrm{Pic}_{\mathcal{C}/S}$ is an étale sheaf, there exists $\mathscr{L} \in \mathrm{Pic}_{\mathcal{C}/S}(S)$ that corresponds to $\mathcal{O}_{\mathbb{P}^1_{S'}}(1)$ under the canonical injection $\mathrm{Pic}_{\mathcal{C}/S}(S) \to \mathrm{Pic}_{\mathcal{C}/S}(S')$. By definition, it is easy to check that $\mathscr{L}_{\overline{s}} = \mathcal{O}_{\mathbb{P}^1}(1)$ for any geometric point $\overline{s} \in S$.

By the same argument as in the proof of Lemma 5.3, it follows that $\mathscr{M}_{d,v,u,r}$ is a category fibered in groupoids. By Proposition 3.4, we can find a positive integer w', depending only on d, v and u, such that for any $f\colon (X, 0, A) \to C \in \mathfrak{Z}_{d,v,u}$ together with a general fiber F of f, the divisor $A + t_A F$ is ample and $\mathrm{vol}(A + t_A F) \le w'$ for some integer t_A. Then

$$f\colon (X, 0, A + t_A F) \to C \in \mathfrak{Z}_{d,v,u,w'}.$$

Pick integers $w_1, \dots, w_k \in (w', w' + dv]$ such that for each $1 \le i \le k$ there is an object $f_i : (X_i, 0, A_i) \to C_i \in \mathfrak{Z}_{d,v,u,w'+dv}$ such that $\mathrm{vol}(A_i) = w_i$. Note that $\mathrm{vol}(A') \in \mathbb{Z}$ for every $f' : (X', 0, A') \to C' \in \mathfrak{Z}_{d,v,u,w'+dv}$ because A' is a line bundle. Hence, we have

$$\mathrm{vol}(A') = w_i$$

for some i if $\mathrm{vol}(A') > w'$.

For each $1 \le i \le k$, we set $\mathscr{M}_{w_i}$ as the open and closed substack of $\mathscr{M}_{d,v,u,w'+dv,r}$ that parametrizes $f : (\mathcal{X}, \mathscr{A}) \to \mathcal{C}$ such that $\mathrm{vol}(\mathscr{A}_{\overline{s}}) = w_i$ for all geometric points $\overline{s} \in S$. Then there is a natural morphism

$$\gamma : \mathscr{M}_{w_1} \sqcup \cdots \sqcup \mathscr{M}_{w_k} \longrightarrow \mathscr{M}_{d,v,u,r}$$

between categories fibered in groupoids. If γ is an isomorphism, then $\mathscr{M}_{d,v,u,r}$ is an open and closed substack of $\mathscr{M}_{d,v,u,w'+dv,r}$, and Theorem 1.3 immediately follows from Theorem 5.1. Therefore, it suffices to prove that γ is an isomorphism.

For a scheme S, we regard $f : (\mathcal{X}, \mathscr{A}) \to \mathcal{C} \in \mathscr{M}_{w_j}(S)$ as an object of $\mathscr{M}_{d,v,u,r}(S)$. It suffices to show the full faithfulness and the essential surjectivity of

$$\gamma(S) := \gamma|_{\mathscr{M}_{w_1} \sqcup \cdots \sqcup \mathscr{M}_{w_k}(S)}$$

for any scheme S. Firstly, we prove the full faithfulness. By the definitions of $\mathscr{M}_{d,v,u,r}$ and $\mathscr{M}_{d,v,u,w'+dv,r}$, we see that $\gamma(S)$ is faithful. To show the fullness, take two objects $f : (\mathcal{X}, \mathscr{A}) \to \mathcal{C}$ and $f' : (\mathcal{X}', \mathscr{A}') \to \mathcal{C}'$ of $\mathscr{M}_{w_1} \sqcup \cdots \sqcup \mathscr{M}_{w_k}(S)$ and an isomorphism $\alpha : f \to f'$ in $\mathscr{M}_{d,v,u,r}(S)$. This means that $\alpha : \mathcal{X} \to \mathcal{X}'$ is an S-isomorphism and there exists an element $\mathscr{B} \in \mathrm{Pic}_{\mathcal{C}/S}(S)$ such that $\alpha^* \mathscr{A}' = \mathscr{A} \otimes f^* \mathscr{B}$. To show that α comes from an isomorphism in $\mathscr{M}_{w_1} \sqcup \cdots \sqcup \mathscr{M}_{w_k}$, it suffices to prove $\mathscr{B} = 0$ as an element of $\mathrm{Pic}_{\mathcal{C}/S}(S)$. For any geometric point $\overline{s} \in S$, we have $\mathscr{B}_{\overline{s}} \sim \mathcal{O}_{\mathbb{P}^1}(m)$ for some $m \in \mathbb{Z}$ and

$$\mathrm{vol}(\mathscr{A}'_{\overline{s}}) = \mathrm{vol}(\mathscr{A}_{\overline{s}}) + dmv.$$

By the property of $w_1, \dots, w_k$, there exist two indices i and j such that $w_i = \mathrm{vol}(\mathscr{A}_{\overline{s}})$ and $w_j = \mathrm{vol}(\mathscr{A}'_{\overline{s}})$. Since $|w_i - w_j| < dv$, we have that $m = 0$. This implies $\mathscr{B}_{\overline{s}} \sim \mathcal{O}_{\mathbb{P}^1}$. By the proof of [Mumford et al. 1994, Section 0.5(b)], we see that $\mathscr{B} = 0$ as an element of $\mathrm{Pic}_{\mathcal{C}/S}(S)$. From this, it follows that $\gamma(S)$ is a fully faithful functor.

Secondly, we prove the essential surjectivity of $\gamma(S)$. We fix an object $f : (\mathcal{X}, \mathscr{A}) \to \mathcal{C}$ of $\mathscr{M}_{d,v,u,r}(S)$ and a geometric point $\overline{s} \in S$. As in the third paragraph of this proof, there is $p \in \mathbb{Z}$ such that $\mathscr{A}_{\overline{s}} \otimes f_{\overline{s}}^* \mathscr{L}_{\overline{s}}^{\otimes p}$ is ample and

$$\mathrm{vol}(\mathscr{A}_{\overline{s}} \otimes f_{\overline{s}}^* \mathscr{L}_{\overline{s}}^{\otimes p}) \le w'$$

for every geometric point $\overline{s} \in S$. Then there exists an open neighborhood U of $\overline{s}$ such that $\mathscr{A}_{\overline{t}} \otimes f_{\overline{t}}^* \mathscr{L}_{\overline{t}}^{\otimes p}$ is ample for any geometric point $\overline{t} \in U$. By shrinking U, we may assume that $\mathrm{vol}(\mathscr{A}_{\overline{t}} \otimes f_{\overline{t}}^* \mathscr{L}_{\overline{t}}^{\otimes p})$ is independent of $\overline{t} \in U$. This is because the function

$$U \ni t \mapsto \mathrm{vol}(\mathscr{A}_{\overline{t}} \otimes f_{\overline{t}}^* \mathscr{L}_{\overline{t}}^{\otimes p}) \in \mathbb{Z}_{>0}$$

is locally constant. Take a positive integer q such that $\mathscr{A}_{\bar{t}} \otimes f_{\bar{t}}^* \mathscr{L}_{\bar{t}}^{\otimes p+q}$ is ample and

$$w' < \mathrm{vol}(\mathscr{A}_{\bar{t}} \otimes f_{\bar{t}}^* \mathscr{L}_{\bar{t}}^{\otimes p+q}) = \mathrm{vol}(\mathscr{A}_{\bar{t}} \otimes f_{\bar{t}}^* \mathscr{L}_{\bar{t}}^{\otimes p}) + dqv \le w' + dv$$

for any geometric point $\bar{t} \in U$. Then we see that $\mathrm{vol}(\mathscr{A}_{\bar{t}} \otimes f_{\bar{t}}^* \mathscr{L}_{\bar{t}}^{\otimes p+q}) = w_i$ for some i. The above argument shows that for any geometric point $\bar{s} \in S$, there exists an open neighborhood U of $\bar{s}$ such that $f|_U$ comes from $\mathscr{M}_{w_1} \sqcup \cdots \sqcup \mathscr{M}_{w_k}(U)$. More precisely, there exists a set of open subsets $\{U_\lambda \subset S\}_{\lambda \in \Lambda}$ with $q_\lambda \in \mathbb{Z}$ such that there exists an integer $i \in [1,k]$ such that $\mathscr{A}_{\bar{s}} \otimes f_{\bar{s}}^* \mathscr{L}_{\bar{s}}^{\otimes q_\lambda}$ is ample and $\mathrm{vol}(\mathscr{A}_{\bar{s}} \otimes f_{\bar{s}}^* \mathscr{L}_{\bar{s}}^{\otimes q_\lambda}) = w_i$ for any geometric point $\bar{s} \in U_\lambda$. For any $\lambda, \lambda' \in \Lambda$, if $U_\lambda \cap U_{\lambda'} \neq \varnothing$, then

$$\mathrm{vol}(\mathscr{A}_{\bar{s}} \otimes f_{\bar{s}}^* \mathscr{L}_{\bar{s}}^{\otimes q_\lambda}) = \mathrm{vol}(\mathscr{A}_{\bar{s}} \otimes f_{\bar{s}}^* \mathscr{L}_{\bar{s}}^{\otimes q_{\lambda'}}) + d(q_\lambda - q_{\lambda'})v$$

for any geometric point $\bar{s} \in U_\lambda \cap U_{\lambda'}$. Since we have $|d(q_\lambda - q_{\lambda'})v| < dv$ by construction, we have $q_\lambda = q_{\lambda'}$. Then we can glue $\mathscr{A} \otimes f^* \mathscr{L}^{\otimes q_\lambda}|_{\pi_{\mathcal{X}}^{-1}(U_\lambda)}$, and we obtain $\mathscr{A}' \in \mathrm{Pic}_{\mathcal{X}/S}(S)$. By construction, there exists an element $\mathscr{B} \in \mathrm{Pic}_{\mathcal{C}/S}(S)$ such that $\mathscr{A} \otimes f^* \mathscr{B} = \mathscr{A}'$ and $f : (\mathcal{X}, \mathscr{A}') \to \mathcal{C} \in \mathscr{M}_{w_1} \sqcup \cdots \sqcup \mathscr{M}_{w_k}(S)$. This means that $\gamma(S)$ is essentially surjective. Thus, we conclude that γ is an isomorphism. □

Remark 5.7 In Theorem 5.1, we have constructed the moduli spaces depending on the choice of r of Lemma 3.1. However, the reduced structures of these moduli stacks are independent of r. To see this, it suffices to show the following:

(∗) Fix r as in Lemma 3.1 and $l \in \mathbb{Z}_{>0}$. For every reduced scheme S, every object $f : (\mathcal{X}, \mathscr{A}) \to \mathcal{C} \in \mathscr{M}_{d,v,u,w,lr}(S)$ is an object of $\mathscr{M}_{d,v,u,w,r}(S)$.

To show (∗), we only need to check (iii) and (iv) of $\mathscr{M}_{d,v,u,w,r}(S)$ for $f : (\mathcal{X}, \mathscr{A}) \to \mathcal{C} \in \mathscr{M}_{d,v,u,w,lr}(S)$ as above.

We first check (iii). Note that the function $S \ni s \mapsto \chi(\mathcal{X}_s, \mathcal{O}_{\mathcal{X}_s}(lerK_{\mathcal{X}_s} + m\mathscr{A}_s))$ is locally constant for every $m \in \mathbb{Z}$. By the basepoint-freeness of $erK_{\mathcal{X}_{\bar{s}}}$ and considering the exact sequence

$$0 \to \mathcal{O}_{\mathcal{X}_{\bar{s}}}(kerK_{\mathcal{X}_{\bar{s}}} + m\mathscr{A}_{\bar{s}}) \to \mathcal{O}_{\mathcal{X}_{\bar{s}}}((k+1)erK_{\mathcal{X}_{\bar{s}}} + m\mathscr{A}_{\bar{s}}) \to \mathcal{O}_D(m\mathscr{A}_{\bar{s}}|_D) \to 0$$

for every $k \ge 0$, where $D \in |erK_{\mathcal{X}_{\bar{s}}}|$ is a general member, we have

$$\chi(\mathcal{X}_s, \mathcal{O}_{\mathcal{X}_s}(kerK_{\mathcal{X}_s} + m\mathscr{A}_s)) = \chi(\mathcal{X}_s, \mathcal{O}_{\mathcal{X}_s}(m\mathscr{A}_s)) + k\chi(D, \mathcal{O}_D(m\mathscr{A}_s|_D))$$

for every $m \in \mathbb{Z}$. Considering the case $k = l$, we see that $\chi(D, \mathcal{O}_D(m\mathscr{A}_s|_D))$ is locally constant on $s \in S$. Therefore, the function

$$S \ni s \mapsto \chi(\mathcal{X}_s, \mathcal{O}_{\mathcal{X}_s}(kerK_{\mathcal{X}_s} + m\mathscr{A}_s))$$

is locally constant for every $m \in \mathbb{Z}$ and $k \in \mathbb{Z}_{\ge 0}$. By Theorem 2.19, there is the universal hull $\omega_{\mathcal{X}/S}^{[r]}$. By the definition of r in Lemma 3.1, the sheaf $\omega_{\mathcal{X}_{\bar{s}}}^{[r]}$ is invertible for any geometric point $\bar{s} \in S$. From this, $\omega_{\mathcal{X}/S}^{[r]}$ exists as a line bundle. Therefore (iii) of $\mathscr{M}_{d,v,u,w,r}(S)$ is satisfied.

Next, we check (iv) of $\mathscr{M}_{d,v,u,w,r}(S)$. By applying Theorem 4.8 to the normalization of S, the function

$$S \ni s \mapsto \dim H^0(\mathcal{X}_s, \mathcal{O}_{\mathcal{X}_s}(kerK_{\mathcal{X}_s}))$$

is locally constant for every $k \geq 0$. By [Mumford 2008, Section 5, Corollary 2] and the reducedness of S, we see that (iv) of $\mathscr{M}_{d,v,u,w,r}(S)$ is satisfied.

From the above discussion, we have $f\colon(\mathcal{X},\mathscr{A})\to\mathcal{C}\in\mathscr{M}_{d,v,u,w,r}(S)$. Thus, we see that $(*)$ holds and the reduced structures of $\mathscr{M}_{d,v,u,w,lr}$ and $\mathscr{M}_{d,v,u,w,r}$ are the same. As we saw in the proof of Theorem 1.3, $\mathscr{M}_{d,v,u,r}$ is an open and closed substack of $\mathscr{M}_{d,v,u,w,r}$ and hence its reduced structure does not depend on the choice of r.

6 Uniformity of adiabatic K-stability

This section is devoted to show Theorem 1.6 and Corollary 1.7. Throughout this section, we work over the field of complex numbers $\mathbb{C}$, and we will use $\mathfrak{F}_{d,n,v}$ and $\mathfrak{G}_{d,n,v,u}$ in Section 3. We fix d, $n\in\mathbb{Z}_{>0}$, $u\in\mathbb{Q}$ and $v\in\mathbb{Q}_{>0}$. For any $w\in\mathbb{Q}_{>0}$ we consider the set

$$\mathfrak{G}^{\mathrm{Car}}_{d,n,v,u,w} := \left\{ f\colon(X,\Delta,A)\to C \;\middle|\; \begin{array}{l} f\colon(X,\Delta)\to C\in\mathfrak{G}_{d,n,v,u} \text{ and } A \text{ is an ample Cartier} \\ \text{divisor satisfying (iv) of } \mathfrak{F}_{d,n,v} \text{ such that } \mathrm{vol}(A)\leq w. \end{array} \right\}$$

Recall from by Theorem 3.6 and Remark 3.7 that there exists $m\in\mathbb{Z}_{>0}$, depending only on d, n, u and v, such that for any element $f\colon(X,\Delta,A)\to C$ of $\mathfrak{G}_{d,(1/n)\mathbb{Z}\cap[0,1],v,u,w}$, mA is Cartier. Thus, we can regard $f\colon(X,\Delta,mA)\to C$ as an element of $\mathfrak{G}^{\mathrm{Car}}_{d,n,m^{d-1}v,u,m^d w}$.

First, we parametrize all elements of $\mathfrak{G}^{\mathrm{Car}}_{d,n,v,u,w}$ when $u\neq 0$.

Proposition 6.1 *Fix d, n, u, v and w as above such that $u\neq 0$. Then there exist a log $\mathbb{Q}$-Gorenstein family $f\colon(\mathcal{X},\mathcal{D})\to S$, an f-ample Carter divisor $\mathscr{A}$ on $\mathcal{X}$, and an S-morphism $g\colon\mathcal{X}\to\mathcal{C}$ such that S is a normal scheme of finite type over $\mathbb{C}$, $\mathcal{C}$ is a normal scheme which is smooth and projective over S, and moreover,*

- *we have $g_s\colon(\mathcal{X}_s,\mathcal{D}_s,\mathscr{A}_s)\to\mathcal{C}_s\in\mathfrak{G}^{\mathrm{Car}}_{d,n,v,u,w}$ for any closed point $s\in S$, and*
- *for any element $h\colon(X,\Delta,A)\to C\in\mathfrak{G}^{\mathrm{Car}}_{d,n,v,u,w}$, there exist a closed point $s\in S$ and isomorphisms $\alpha\colon(\mathcal{X}_s,\mathcal{D}_s)\to(X,\Delta)$ and $\beta\colon\mathcal{C}_s\to C$ satisfying $h\circ\alpha=\beta\circ g_s$ and $\alpha^*A\sim\mathscr{A}_s$.*

Proof The case when the boundary Δ is zero in the proposition easily follows from Proposition 5.6. In the general case, the proposition holds true by the standard argument of the boundedness and the idea in the proof of Proposition 5.6. □

The following is the key step to showing Theorem 1.6 for the case when $u>0$.

Theorem 6.2 *Let $f:(\mathcal{X},\mathcal{D})\to S$ be a log $\mathbb{Q}$-Gorenstein family and $\mathcal{A}$ an f-ample line bundle on $\mathcal{X}$. Suppose that $K_{\mathcal{X}/S}+\mathcal{D}$ is nef over S and the fiber $(\mathcal{X}_s,\mathcal{D}_s)$ over any closed point $s\in S$ is klt.*

Then there exists a positive rational number ϵ_0 such that $(\mathcal{X}_s,\mathcal{D}_s,\epsilon\mathcal{A}_s+K_{\mathcal{X}_s}+\mathcal{D}_s)$ is specially K-stable for the fiber $(\mathcal{X}_s,\mathcal{D}_s,\mathcal{A}_s)$ over any closed point $s\in S$ and any rational number $\epsilon\in(0,\epsilon_0)$. Furthermore, there exists a positive rational number α such that

$$M^{\mathrm{NA}}_{\mathcal{D}_s}(\mathcal{Y},\mathcal{M})\geq\alpha(\mathcal{J}^{\epsilon\mathcal{A}_s+K_{\mathcal{X}_s}+\mathcal{D}_s})^{\mathrm{NA}}(\mathcal{Y},\mathcal{M})$$

for any rational number $\epsilon\in(0,\epsilon_0)$, closed point $s\in S$, and normal semiample test configuration $(\mathcal{Y},\mathcal{M})$ for $(\mathcal{X}_s,\epsilon\mathcal{A}_s+K_{\mathcal{X}_s}+\mathcal{D}_s)$.

Proof First, we note that $\mathcal{A}+K_{\mathcal{X}/S}+\mathcal{D}$ is f-ample. By [Blum and Liu 2022, Proposition 5.3], there exists $\delta_0>0$ such that $\alpha_{(\mathcal{X}_s,\mathcal{D}_s)}(\mathcal{A}_s+K_{\mathcal{X}_s}+\mathcal{D}_s)\geq\delta_0$ for any closed point $s\in S$. We also have

$$\alpha_{(\mathcal{X}_s,\mathcal{D}_s)}(\mathcal{A}_s+K_{\mathcal{X}_s}+\mathcal{D}_s)\leq\alpha_{(\mathcal{X}_s,\mathcal{D}_s)}(\epsilon\mathcal{A}_s+K_{\mathcal{X}_s}+\mathcal{D}_s)$$

for $\epsilon\in(0,1)$. We put d as the relative dimension of f. By Lemma 2.27, we have

$$\begin{aligned}\delta_{(\mathcal{X}_s,\mathcal{D}_s)}(\epsilon\mathcal{A}_s+K_{\mathcal{X}_s}+\mathcal{D}_s)&\geq\frac{d+1}{d}\alpha_{(\mathcal{X}_s,\mathcal{D}_s)}(\epsilon\mathcal{A}_s+K_{\mathcal{X}_s}+\mathcal{D}_s)\\&\geq\frac{d+1}{d}\alpha_{(\mathcal{X}_s,\mathcal{D}_s)}(\mathcal{A}_s+K_{\mathcal{X}_s}+\mathcal{D}_s)\geq\frac{d+1}{d}\delta_0.\end{aligned}\tag{6-1}$$

By Theorem 2.32, there exists a positive rational number C, which depends only on the numbers $(K_X+\Delta)^{d-i}\cdot A^i$ for $0\leq i\leq d$, such that $(X,\epsilon A+K_X+\Delta)$ is uniformly $\mathrm{J}^{K_X+\Delta+C\epsilon(\epsilon A+K_X+\Delta)}$-stable for any $\epsilon>0$ and d-dimensional polarized klt pair (X,Δ,L) such that $K_X+\Delta$ is nef. By the flatness of f, the $(K_{\mathcal{X}_s}+\mathcal{D}_s)^{d-i}\cdot\mathcal{A}_s^i$ are independent of s, hence we can choose C so that $(\mathcal{X}_s,\epsilon\mathcal{A}_s+K_{\mathcal{X}_s}+\mathcal{D}_s)$ is $\mathrm{J}^{K_{\mathcal{X}_s}+\mathcal{D}_s+C\epsilon(\epsilon\mathcal{A}_s+K_{\mathcal{X}_s}+\mathcal{D}_s)}$-stable for any $s\in S$. By taking ϵ_0 such that $0<\epsilon_0<\min\{(d+1)\delta_0/(dC),1\}$, we have

$$M^{\mathrm{NA}}_{\mathcal{D}_s}(\mathcal{Y},\mathcal{M})\geq\left(\frac{(d+1)\delta_0}{d}-C\epsilon_0\right)(\mathcal{J}^{\epsilon\mathcal{A}_s+K_{\mathcal{X}_s}+\mathcal{D}_s})^{\mathrm{NA}}(\mathcal{Y},\mathcal{M})$$

for any rational number $\epsilon\in(0,\epsilon_0)$, closed point $s\in S$ and normal semiample test configuration $(\mathcal{Y},\mathcal{M})$ for $(\mathcal{X}_s,\epsilon\mathcal{A}_s+K_{\mathcal{X}_s}+\mathcal{D}_s)$. ◻

Now we assume $u<0$. In this case, we need to show that the uniform "convergence of the δ-invariant" (cf [Hattori 2022, Theorem D]) holds for all polarized klt-trivial fibrations belonging to one family.

Proposition 6.3 *Let S be a normal variety and $f:\mathcal{X}\to\mathbb{P}^1_S$ be a contraction of normal varieties over S. Suppose that $\pi:(\mathcal{X},\mathcal{D})\to S$ is a log $\mathbb{Q}$-Gorenstein family such that any geometric fiber is a klt pair. Let $\mathcal{H}$ be a π-ample Cartier divisor on $\mathcal{X}$. Suppose further that there exists a positive real number u such that*

$$\lim_{\epsilon\to 0}\delta_{(\mathcal{X}_{\overline{s}},\mathcal{D}_{\overline{s}})}(\epsilon\mathcal{H}_{\overline{s}}+f_{\overline{s}}^*\mathcal{O}(1))\geq u$$

for any geometric fiber $f_{\overline{s}}:(\mathcal{X}_{\overline{s}},\mathcal{D}_{\overline{s}},\mathcal{H}_{\overline{s}})\to\mathbb{P}^1$ over $\overline{s}\in S$.

Then for any $\delta_0 > 0$, *there exists a positive real number* ϵ_0 *such that*

$$\delta_{(\mathcal{X}_{\overline{s}},\mathcal{D}_{\overline{s}})}(\epsilon\mathcal{H}_{\overline{s}} + f_{\overline{s}}^{*}\mathcal{O}(1)) \geq u - \delta_0$$

for any rational number $\epsilon \in (0, \epsilon_0)$ *and geometric point* $\overline{s} \in S$.

Proof By [Blum and Liu 2022, Theorem 6.6], if

$$\delta_{(\mathcal{X}_s,\mathcal{D}_s)}(\epsilon\mathcal{H}_s + f_s^{*}\mathcal{O}(1)) \geq u - \delta_0$$

holds for any closed point $s \in S$, then

$$\delta_{(\mathcal{X}_{\overline{s}},\mathcal{D}_{\overline{s}})}(\epsilon\mathcal{H}_{\overline{s}} + f_{\overline{s}}^{*}\mathcal{O}(1)) \geq u - \delta_0$$

also holds for any geometric point $\overline{s} \in S$ since the set of all closed points is Zariski dense. Thus, it suffices to show the assertion for all closed points of S.

First, we note that to show the assertion, we may freely shrink S or replace S by S' with an étale morphism $S' \to S$. Indeed, if we can prove Proposition 6.3 for a nonempty open subset $U \subset S$, then the assertion for S follows from Noetherian induction. Thus, we may assume as in the proof of Lemma 4.1 that S is smooth and there exists a diagram

$$\begin{tikzcd} \mathcal{Y} \arrow[r, "g"] \arrow[d, "f'"'] & \mathcal{X} \arrow[d, "f"] \\ \mathcal{W} \arrow[r, "h"] & \mathbb{P}^1_S \end{tikzcd} \tag{6-2}$$

of projective morphisms, where $\mathcal{Y}$ and $\mathcal{W}$ are smooth varieties, with snc divisors Σ on $\mathcal{W}$ and Ξ on $\mathcal{Y}$, respectively, such that

(i) h is birational and g is a log resolution of $(\mathcal{X}, \mathcal{D})$,

(ii) f' is a contraction,

(iii) $\Xi \supset f'^{*}\Sigma \cup \mathrm{Supp}(g_*^{-1}\mathcal{D}) \cup \mathrm{Ex}(g)$ and the vertical part of $\Sigma_{\mathcal{Y}}$ with respect to f' maps into Σ,

(iv) the restriction of $f' \colon (\mathcal{Y}, \Xi) \to \mathcal{W}$ over $\mathcal{W} \setminus \Sigma$ is log smooth.

Since $\mathcal{W}$ is isomorphic to $\mathbb{P}^1_S$ over any codimension one point of $\mathbb{P}^1_S$, we may shrink S and assume that $\mathcal{W} \cong \mathbb{P}^1_S$. Taking a suitable étale morphism $T \to S$ and replacing S by T, we may further assume that

(v) $(\mathcal{Y}, \Xi)$ and $(\mathbb{P}^1_S, \Sigma)$ are log smooth over S and any stratum of Ξ or Σ has geometrically integral fibers over S.

Then we apply Lemma 4.1 to S and we conclude by shrinking S that $\mathcal{B}_s = B_s$ for any closed point $s \in S$, where $\mathcal{B}$ (resp. B_s) is the discriminant divisor with respect to f (resp. f_s). We may also assume by shrinking S that there exists a section S' of $\mathbb{P}^1_S \to S$ disjoint from Σ. Then, we show the following claim.

Claim 2 *There exist positive real numbers c and ϵ'_0 that satisfy the following for any closed point $s \in S$ and rational number $\epsilon \in (0, \epsilon'_0)$.*

(i) *We have $\sup_p \mathrm{mult}_p(g_s^* D)_{\mathrm{hor}} \le c\epsilon$ for any effective $\mathbb{Q}$-divisor D that is $\mathbb{Q}$-linearly equivalent to $\epsilon\mathcal{H}_s + f_s^*\mathcal{O}(1)$* (see (12) in Notation and conventions), *where $p \in \mathcal{Y}_s$ runs over all closed points. Here, we set $\mathrm{mult}_p(D') = \mathrm{ord}_E(\mu^* D' - \mu_*^{-1} D')$ for any $\mathbb{Q}$-divisor D', where μ is the blowup of $\mathcal{Y}_s$ at p and E is the exceptional divisor.*

(ii) *There exists a positive integer m_ϵ such that $m_\epsilon\epsilon \in \mathbb{Z}$ and such that for any irreducible component G_s of a fiber F_s of $f_s \circ g_s$, and $m \in \mathbb{Z}_{>0}$ such that $m_\epsilon | m$,*

$$S_{m,\epsilon\mathcal{H}_s + f_s^*\mathcal{O}(1)}(G_s) \le \tfrac{1}{2} T_{f_s^*\mathcal{O}(1)}(G_s) + c\epsilon + c(m\epsilon)^{-1}.$$

Here, we set $T_{f_s^*\mathcal{O}(1)}(G_s) := \mathrm{ord}_{G_s}(F_s)$.

Proof Let $\Sigma = \sum_j F_j$ and $(f \circ g)^* F_j = \sum a_{ij} G_i^{(j)}$ be the irreducible decompositions. Note that $G_i^{(j)}$ has only smooth and irreducible fibers over S by condition (v) of the diagram (6-2). Note also that we may shrink S freely by the same reason as in the second paragraph of the proof of Proposition 6.3. We may further assume that S is quasiprojective and hence there exists a very ample line bundle $\mathcal{A}$ on $\mathcal{Y}$. Set d as the relative dimension of f throughout this proof.

First, we deal with (i). We take $m'_1 \in \mathbb{Z}_{>0}$ such that $m'_1\mathcal{A} - K_{\mathcal{Y}/S} - G'$ is ample, where G' is an irreducible component of $(f \circ g)^* F_j$ or $(f \circ g)^*(S')$. Since all smooth fibers of $f_s \circ g_s$ are linearly equivalent to $((f \circ g)^*(S'))|_{\mathcal{Y}_s}$, $m'_1\mathcal{A}_s - K_{\mathcal{Y}_s} - G_s$ is also ample for any irreducible component G_s of a fiber of $f_s \circ g_s$ and for any $s \in S$. On the other hand, we obtain $m''_1 \in \mathbb{Z}_{>0}$ such that for any closed point $p \in \mathcal{Y}$ and the blowup $\mu \colon \widetilde{\mathcal{Y}} \to \mathcal{Y}$ at p with the exceptional divisor E, $\mu^*(m''_1\mathcal{A}) - E$ is ample by applying [Lazarsfeld 2004, 1.4.14] to a suitable projective compactification of $(\mathcal{Y}, \mathcal{A})$. Let $s = \pi \circ g(p)$ and $\mathcal{Z}$ be an irreducible component of $\widetilde{\mathcal{Y}}_s$ such that $\mu|_{\mathcal{Z}} \colon \mathcal{Z} \to \mathcal{Y}_s$ is the blowup at $p \in \mathcal{Y}_s$. Then $E \cap \mathcal{Z}$ is the $\mu|_{\mathcal{Z}}$-exceptional divisor. Now we fix an irreducible component G_s of a fiber of $f_s \circ g_s$ such that $p \in G_s$. Restricting divisors to $\mathcal{Z}$, we can check that $m_1 := m'_1 + (d-1)m''_1$ satisfies that for any closed points $s \in S$ and $p \in \mathcal{Y}_s$, both $\mu|_{\mathcal{Z}}^*(m_1\mathcal{A}_s - K_{\mathcal{Y}_s} - G_s) - (d-1)(E \cap \mathcal{Z})$ and $\mu|_{\mathcal{Z}}^*(m_1\mathcal{A}_s) - (E \cap \mathcal{Z})$ are ample. Let $\widetilde{G}_s := (\mu|_{\mathcal{Z}})_*^{-1} G_s$ and note that $\widetilde{G}_s$ is smooth. Hence, there exists a positive integer $m_2 > 0$ depending only on d such that $m_2(m_1\mu|_{\mathcal{Z}}^*\mathcal{A}_s - E \cap \mathcal{Z})|_{\widetilde{G}_s}$ is globally generated by applying the effective basepoint-freeness [Kollár 1993, Theorem 1.1] to $\widetilde{G}_s$ since

$$(\mu|_{\mathcal{Z}}^*(m_1\mathcal{A}_s - K_{\mathcal{Y}_s} - G_s) - (d-1)(E \cap \mathcal{Z}))|_{\widetilde{G}_s} = (m_1\mu|_{\mathcal{Z}}^*\mathcal{A}_s - (E \cap \mathcal{Z}))|_{\widetilde{G}_s} - K_{\widetilde{G}_s}$$

is ample. Then for any D effective $\mathbb{Q}$-divisor $\mathbb{Q}$-linearly equivalent to $\epsilon\mathcal{H}_s + f_s^*\mathcal{O}(1)$, let $\gamma' := \bigcap_{i=1}^{d-2} D_i$ and $\gamma := (\mu|_{\mathcal{Z}})_*\gamma'$ for sufficiently general $D_i \in |m_2(m_1\mu|_{\mathcal{Z}}^*\mathcal{A}_s - (E \cap \mathcal{Z}))|_{\widetilde{G}_s}|$ such that $\gamma \not\subset (g_s^* D)_{\mathrm{hor}}$ and $\gamma' \not\subset E \cap \mathcal{Z}$. Since $m_1\mu|_{\mathcal{Z}}^*\mathcal{A}_s - E \cap \mathcal{Z}$ is ample, γ' intersects with $E \cap \mathcal{Z}$. Hence, γ passes through p.

On the other hand, we see that

$$\begin{aligned}\gamma \cdot (g_s^* D)_{\mathrm{hor}} &= \gamma' \cdot (\mu|_{\mathcal{Z}})^*((g_s^* D)_{\mathrm{hor}}) \\ &= (m_2(m_1(\mu|_{\mathcal{Z}})^*\mathcal{A}_s - (E \cap \mathcal{Z})))^{d-2} \cdot \widetilde{G}_s \cdot (\mu|_{\mathcal{Z}})^*((g_s^* D)_{\mathrm{hor}}) \\ &= (m_2 m_1 \mathcal{A}_s)^{d-2} \cdot G_s \cdot (g_s^* D)_{\mathrm{hor}} \\ &\le (m_2 m_1 \mathcal{A}_s)^{d-2} \cdot (T_{f_s^*\mathcal{O}(1)}(G_s) g_s^* f_s^* \mathcal{O}(1)) \cdot (g_s^* D)_{\mathrm{hor}} \\ &\le (m_2 m_1 \mathcal{A}_s)^{d-2} \cdot (T_{f_s^*\mathcal{O}(1)}(G_s) g_s^* f_s^* \mathcal{O}(1)) \cdot g_s^* D \\ &= \epsilon (m_1 m_2)^{d-2} T_{f_s^*\mathcal{O}(1)}(G_s)(\mathcal{A}_s^{d-2} \cdot g_s^*\mathcal{H}_s \cdot g_s^* f_s^* \mathcal{O}(1)).\end{aligned}$$

Let $T = \max_{G_s} T_{f_s^*\mathcal{O}(1)}(G_s)$. Then we see that T is independent of $s \in S$ by the conditions of the diagram (6-2). Let $M := (m_1 m_2)^{d-2} T(\mathcal{A}_s^{d-2} \cdot g_s^*\mathcal{H}_s \cdot g_s^* f_s^* \mathcal{O}(1)) > 0$. Then, we see that this M is independent of $s \in S$ and that $\mathrm{mult}_p((g_s^* D)_{\mathrm{hor}}) \le M\epsilon$ for any closed points $s \in S$ and $p \in \mathcal{Y}_s$, and any effective $\mathbb{Q}$-divisor D that is $\mathbb{Q}$-linearly equivalent to $\epsilon\mathcal{H}_s + f_s^*\mathcal{O}(1)$. Indeed, we saw in the above argument that there exists a curve $\gamma \not\subset (g_s^* D)_{\mathrm{hor}}$ passing through p such that $\gamma \cdot (g_s^* D)_{\mathrm{hor}} \le M\epsilon$. Then, we have

$$\mathrm{mult}_p((g_s^* D)_{\mathrm{hor}}) \le \gamma' \cdot \mu^*((g_s^* D)_{\mathrm{hor}}) = \gamma \cdot (g_s^* D)_{\mathrm{hor}} \le M\epsilon.$$

Thus, we obtain the assertion (i).

Next, we deal with (ii). Fix $\tau > 0$ so that $\mathcal{A}_s^{d-1} \cdot g_s^*(\mathcal{H}_s - \tau f_s^*\mathcal{O}(1)) < 0$ for some closed point $s \in S$. Then, $\mathcal{H}_s - \tau f_s^*\mathcal{O}(1)$ is not pseudoeffective for any $s \in S$. Furthermore, take $m_0 \in \mathbb{Z}_{>0}$ such that $m_0\mathcal{H} - (K_{\mathcal{X}/S} + \mathcal{D})$ is π-ample.

We first treat the case when G_s is a smooth fiber of $f_s \circ g_s$ for any closed point $s \in S$. Then, it follows from [Fujita and Odaka 2018, Lemma 2.2] that

$$\begin{aligned}&m h^0(\mathcal{X}_s, m\epsilon\mathcal{H}_s + m f_s^*\mathcal{O}(1)) S_{m,\epsilon\mathcal{H}_s + f_s^*\mathcal{O}(1)}(G_s) \\ &\qquad \le \sum_{k=1}^{\infty} h^0(\mathcal{Y}_s, g_s^*(m\epsilon\mathcal{H}_s + m f_s^*\mathcal{O}(1)) - kG_s) \\ &\qquad \le \sum_{k=1}^{m} h^0(\mathcal{Y}_s, g_s^*(m\epsilon\mathcal{H}_s + (m-k) f_s^*\mathcal{O}(1))) + \sum_{k=1}^{\lceil m\epsilon\tau \rceil} h^0(\mathcal{Y}_s, g_s^*(m\epsilon\mathcal{H}_s)).\end{aligned}$$

Recall that $m\epsilon\mathcal{H} - K_{\mathcal{X}/S} - \mathcal{D}$ is π-ample for any $m \in \mathbb{Z}_{>0}$ such that $m\epsilon \in \mathbb{Z}$ and $m\epsilon \ge m_0$. Then

$$h^0(\mathcal{Y}_s, g_s^*(m\epsilon\mathcal{H}_s + k f_s^*\mathcal{O}(1))) = \chi(\mathcal{X}_s, m\epsilon\mathcal{H}_s + k f_s^*\mathcal{O}(1))$$

by the Kawamata–Viehweg vanishing theorem for any $k \ge 0$ and any closed point $s \in S$. Hence, $h^0(\mathcal{Y}_s, g_s^*(m\epsilon\mathcal{H}_s + k f_s^*\mathcal{O}(1)))$ is independent of $s \in S$. We further have that

$$\chi(\mathcal{X}_s, m\epsilon\mathcal{H}_s + k f_s^*\mathcal{O}(1)) = \chi(\mathcal{X}_s, m\epsilon\mathcal{H}_s) + k\chi(G_s, m\epsilon\mathcal{H}_s|_{G_s}).$$

Here, we note that $\chi(G_s, m\epsilon\mathcal{H}_s|_{G_s})$ is also independent of $s \in S$. Thus,

$$h^0(\mathcal{Y}_s, g_s^*(m\epsilon\mathcal{H}_s + mf_s^*\mathcal{O}(1))) = \frac{m^d\epsilon^{d-1}}{(d-1)!}((\mathcal{H}_s^{d-1}\cdot f_s^*\mathcal{O}(1)) + O(\epsilon) + O((m\epsilon)^{-1})). \tag{6-3}$$

Furthermore, we have by [Kollár and Mori 1998, Theorem 1.36] that for sufficiently large $m\epsilon$ and ϵ^{-1},

$$\begin{aligned}\sum_{k=1}^{m} h^0(\mathcal{Y}_s, g_s^*(m\epsilon\mathcal{H}_s + (m-k)f_s^*\mathcal{O}(1))) &+ \sum_{k=1}^{\lceil m\epsilon\tau\rceil} h^0(\mathcal{Y}_s, g_s^*(m\epsilon\mathcal{H}_s)) \\ &= \frac{m(m-1)}{2}\chi(G_s, m\epsilon\mathcal{H}_s|_{G_s}) + (m+\lceil m\epsilon\tau\rceil)\chi(\mathcal{X}_s, m\epsilon\mathcal{H}_s) \\ &= \frac{m^{d+1}\epsilon^{d-1}}{2(d-1)!}((\mathcal{H}_s^{d-1}\cdot f_s^*\mathcal{O}(1)) + O(\epsilon) + O((m\epsilon)^{-1})).\end{aligned} \tag{6-4}$$

By (6-3) and (6-4), we see that there exist positive real numbers C_0, C_0' and C_0'' such that

$$S_{m,\epsilon\mathcal{H}_s + f_s^*\mathcal{O}(1)}(G_s) \le \tfrac{1}{2} + C_0'\epsilon + C_0''(m\epsilon)^{-1} \tag{6-5}$$

for any rational number $0 < \epsilon < C_0^{-1}$, positive integer m such that $m\epsilon \in \mathbb{Z}$ and $m\epsilon > \max\{C_0, m_0\}$, closed point $s \in S$, and smooth fiber G_s of $f_s \circ g_s$.

Next, we deal with the case when G_s is an irreducible component of a singular fiber of $f_s \circ g_s$ for three paragraphs. Recall that $(f\circ g)^*F_j = \sum_{i=0}^{r_j} a_{ij}G_i^{(j)}$ is the irreducible decomposition. Then, we see that $G_s = (G_i^{(j)})|_{\mathcal{Y}_s}$ for some j and i. By renumbering $G_i^{(j)}$, we may assume that $G = G_0^{(j)}$. We note that a matrix $(G_{k,s}^{(j)}\cdot G_{l,s}^{(j)}\cdot \mathcal{A}_s^{d-2})_{1\le k,l\le r_j}$ is negative definite [Li and Xu 2014, Lemma 1]. Thus, there exists a Cartier divisor $F' = \sum_{i=1}^{r_j} e_iG_i^{(j)}$ such that

$$G_{i,s}^{(j)}\cdot\mathcal{A}_s^{d-2}\cdot(g_s^*\mathcal{H}_s + F_s') < 0$$

for $i > 0$ and for some closed point $s \in S$. For some $b \in \mathbb{Z}_{>0}$, $F'' := F' + b(f\circ g)^*F_j$ is effective. Then the inequality

$$G_{i,s}^{(j)}\cdot\mathcal{A}_s^{d-2}\cdot(g_s^*\mathcal{H}_s + F_s'') < 0 \tag{6-6}$$

also holds for F'' and for any closed point $s \in S$. Fix a positive integer a such that $a(f\circ g)^*F_j - F''$ is effective. We claim that for any closed point $s \in S$, $m \in \mathbb{Z}_{>0}$ such that $m\epsilon \in \mathbb{Z}$ and $k \ge 0$,

$$\mathcal{Y}_s, m\epsilon F_s'' + g_s^*(m\epsilon\mathcal{H}_s + mf_s^*\mathcal{O}(1)) - ka_{0j}G_s) = h^0(\mathcal{Y}_s, m\epsilon F_s'' + g_s^*(m\epsilon\mathcal{H}_s + (m-k)f_s^*\mathcal{O}(1))). \tag{6-7}$$

Indeed, note that

$$m(f_s\circ g_s)^*F_{j,s} - ka_{0j}G_s = (m-k)(f_s\circ g_s)^*F_{j,s} + k\sum_{i>0} a_{ij}G_{i,s}^{(j)},$$

and we claim that $k\sum_{i>0} a_{ij}G_{i,s}^{(j)}$ is contained in the fixed part of the linear system

$$|m\epsilon F_s'' + g_s^*(m\epsilon\mathcal{H}_s + mf_s^*\mathcal{O}(1)) - ka_{0j}G_s|.$$

For this, it suffices to show for any nonzero effective divisor $M=\sum_{i>0} m_i G^{(j)}_{i,s}$ and $l \in \mathbb{Z}_{\geq 0}$ that the fixed part of the linear system $|m\epsilon F''_s + g_s^*(m\epsilon\mathcal{H}_s + l f_s^*\mathcal{O}(1)) + M|$ contains some $G^{(j)}_{i,s}$ such that $m_i>0$. Here, we see that

$$G^{(j)}_{i,s}\cdot\mathcal{A}_s^{d-2}\cdot(m\epsilon F''_s + g_s^*(m\epsilon\mathcal{H}_s + l f_s^*\mathcal{O}(1)) + M) < G^{(j)}_{i,s}\cdot\mathcal{A}_s^{d-2}\cdot M$$

by (6-6), and $G^{(j)}_{i,s}\cdot\mathcal{A}_s^{d-2}\cdot M<0$ for some $i>0$ such that $m_i>0$ by [Li and Xu 2014, Lemma 1]. This means that $G^{(j)}_{i,s}$ is contained in the fixed part and thus we obtain the equality (6-7).

By equation (6-7) and the fact that $a(f\circ g)^*F_j - F''$ is effective, we have

$$\begin{aligned}
&mh^0(\mathcal{X}_s, m\epsilon\mathcal{H}_s + mf_s^*\mathcal{O}(1))S_{m,\epsilon\mathcal{H}_s+f_s^*\mathcal{O}(1)}(G_s)\\
&\quad\leq \sum_{k=1}^{\infty}\sum_{l=0}^{a_{0j}-1} h^0(\mathcal{Y}_s, g_s^*(m\epsilon\mathcal{H}_s + mf_s^*\mathcal{O}(1)) - (ka_{0j}-l)G_s)\\
&\quad\leq a_{0j}\sum_{k=1}^{\infty} h^0(\mathcal{Y}_s, g_s^*(m\epsilon\mathcal{H}_s + mf_s^*\mathcal{O}(1)) - ka_{0j}G_s)\\
&\quad\leq a_{0j}\sum_{k=0}^{\infty} h^0(\mathcal{Y}_s, m\epsilon F''_s + g_s^*(m\epsilon\mathcal{H}_s + (m-k)f_s^*\mathcal{O}(1)))\\
&\quad\leq a_{0j}\sum_{k=0}^{m} h^0(\mathcal{Y}_s, g_s^*(m\epsilon\mathcal{H}_s + (m-k+m\epsilon a)f_s^*\mathcal{O}(1)))\\
&\qquad\qquad + a_{0j}\sum_{k=1}^{\lceil m\epsilon(\tau+a)\rceil} h^0(\mathcal{Y}_s, g_s^*(m\epsilon(\mathcal{H}_s + af_s^*\mathcal{O}(1))) - kg_s^*f_s^*\mathcal{O}(1)).
\end{aligned}$$

By (6-3) and estimating the right-hand side of the above inequality as (6-4), we see that there exist positive real numbers $C_{G_0^{(j)}}, C'_{G_0^{(j)}}$ and $C''_{G_0^{(j)}}>0$ such that

$$S_{m,\epsilon\mathcal{H}_s+f_s^*\mathcal{O}(1)}(G^{(j)}_{0,s}) \leq \tfrac{1}{2}a_{0j} + C'_{G_0^{(j)}}\epsilon + C''_{G_0^{(j)}}(m\epsilon)^{-1} \tag{6-8}$$

for any rational number $0<\epsilon<C^{-1}_{G_0^{(j)}}$, positive integer m such that $m\epsilon\in\mathbb{Z}$ and $m\epsilon>\max\{C_{G_0^{(j)}}, m_0\}$ and closed point $s\in S$.

Since there exist only finitely many possibilities of $G_i^{(j)}$, we see by the inequalities (6-5) and (6-8) that there exist positive real numbers ϵ'_0 and c such that

$$S_{m,\epsilon\mathcal{H}_s+f_s^*\mathcal{O}(1)}(G_s) \leq \tfrac{1}{2}T_{f_s^*\mathcal{O}(1)}(G_s) + c\epsilon + c(m\epsilon)^{-1}$$

for any rational number $0<\epsilon<\epsilon'_0$, closed point $s\in S$, irreducible component G_s of a fiber of $f_s\circ g_s$ and $m\in\mathbb{Z}_{>0}$ such that $m\epsilon>\epsilon_0'^{-1}$ and $m\epsilon\in\mathbb{Z}$. Thus, we obtain the assertion (ii) by taking $m_\epsilon\in\mathbb{Z}$ for any rational number $0<\epsilon<\epsilon'_0$ such that $m_\epsilon\epsilon>\epsilon_0'^{-1}$ and $m_\epsilon\epsilon\in\mathbb{Z}$. □

Finally, we show that Claim 2 implies Proposition 6.3. Take an arbitrary constant $0<\delta_0<u$. Let c and ϵ'_0 be as in Claim 2 and take $m_\epsilon\in\mathbb{Z}_{>0}$ for any rational number $0<\epsilon<\epsilon'_0$ as (ii). Let D_m be an

m-basis type divisor of $\epsilon\mathcal{H}_s + f_s^*\mathcal{O}(1)$ for any $m \in \mathbb{Z}_{>0}$ such that $m_\epsilon | m$ and Γ be a $\mathbb{Q}$-divisor such that $g_{s*}\Gamma = \mathcal{D}_s$ and

$$K_{\mathcal{Y}_s} + \Gamma = g_s^*(K_{\mathcal{X}_s} + \mathcal{D}_s)$$

for any closed point $s \in S$. Now, we have

$$A_{(\mathcal{X}_s,\mathcal{D}_s)}(G_s) \geq \tfrac{1}{2} u T_{f_s^*\mathcal{O}(1)}(G_s)$$

by [Hattori 2022, Theorem D] for any irreducible component G_s of any fiber F_s of $f_s \circ g_s$. We note that there exists a positive integer r such that $r(K_{\mathcal{X}/S} + \mathcal{D})$ is Cartier. Then we see that for any geometric point $\overline{s} \in S$, $(\mathcal{X}_{\overline{s}}, \mathcal{D}_{\overline{s}})$ is $(1/r)$-lc. Let

$$\epsilon_0'' := \min\left\{\epsilon_0', \frac{\delta_0}{8c(u-\delta_0)}\right\} \quad \text{and} \quad \delta_0' := \min\left\{\frac{\delta_0}{4}, \frac{1}{r}\right\}.$$

For any $0 < \epsilon < \epsilon_0''$ and $m \in \mathbb{Z}_{>0}$ such that $m_\epsilon | m$ and $m\epsilon > \epsilon_0''^{-1}$, we claim that $(\mathcal{Y}_s, \Gamma + (u-\delta_0)(g_s^* D_m)_{\mathrm{vert}})$ is log smooth and δ_0'-sublc. Indeed, this follows from the conditions of the diagram (6-2), [Kollár and Mori 1998, Corollary 2.31] and

$$\begin{aligned} A_{\mathcal{Y}_s}(G_s) - \mathrm{ord}_{G_s}(\Gamma + (u-\delta_0)(g_s^* D_m)_{\mathrm{vert}}) &\geq A_{\mathcal{Y}_s}(G_s) - \mathrm{ord}_{G_s}(\Gamma) - (u-\delta_0) S_{m,\epsilon\mathcal{H}_s + f_s^*\mathcal{O}(1)}(G_s) \\ &\geq A_{(\mathcal{X}_s,\mathcal{D}_s)}(G_s) - (u-\delta_0)\Big(\frac{T_{f_s^*\mathcal{O}(1)}(G_s)}{2} + c\epsilon + c(m\epsilon)^{-1}\Big) \\ &\geq \tfrac{1}{2}\delta_0 T_{f_s^*\mathcal{O}(1)}(G_s) - (u-\delta_0)(c\epsilon + c(m\epsilon)^{-1}) > \tfrac{1}{4}\delta_0. \end{aligned}$$

Let $\Theta = (\Gamma + (g_s^* D_m)_{\mathrm{vert}})_{\mathrm{red}}$. Since $(1-\delta_0')\Theta \geq \Gamma + (u-\delta_0)(g_s^* D_m)_{\mathrm{vert}}$ and $(\mathcal{Y}_s, \Theta)$ is a log smooth pair, we further have as the argument of [Boucksom et al. 2017, Step 2 in Proof of 9.14] that

$$A_{\mathcal{Y}_s}(E) - \mathrm{ord}_E(\Gamma + (u-\delta_0)(g_s^* D_m)_{\mathrm{vert}}) \geq \delta_0' \, A_{\mathcal{Y}_s}(E)$$

for any prime divisor E over $\mathcal{X}_s$. On the other hand, Claim 2(i) and Skoda's theorem [loc. cit., Step 1] show that

$$c\epsilon \, A_{\mathcal{Y}_s}(E) \geq \mathrm{ord}_E((g_s^* D_m)_{\mathrm{hor}}).$$

Let $\epsilon_0 := \min\{\delta_0'/((u-\delta_0)c), \epsilon_0''\}$. Then we have for any prime divisor E over $\mathcal{X}_s$ and $0 < \epsilon < \epsilon_0$,

$$\begin{aligned} A_{(\mathcal{X}_s,\mathcal{D}_s+(u-\delta_0)D_m)}(E) &= A_{\mathcal{Y}_s}(E) - \mathrm{ord}_E(\Gamma + (u-\delta_0)((g_s^* D_m)_{\mathrm{vert}} + (g_s^* D_m)_{\mathrm{hor}})) \\ &\geq (\delta_0' - (u-\delta_0)c\epsilon) A_{\mathcal{Y}_s}(E) > 0. \end{aligned}$$

In other words, we conclude that $(\mathcal{X}_s, \mathcal{D}_s + (u-\delta_0)D_m)$ is klt for any closed point $s \in S$, rational number $\epsilon \in (0, \epsilon_0)$, $m \in \mathbb{Z}$ such that $m_\epsilon | m$ and $m > (\epsilon\epsilon_0'')^{-1}$, and m-basis type divisor D_m of $\epsilon\mathcal{H}_s + f_s^*\mathcal{O}(1)$. Thus, we obtain that

$$\delta_{(\mathcal{X}_s,\mathcal{D}_s)}(\epsilon\mathcal{H}_s + f_s^*\mathcal{O}(1)) = \lim_{l\to\infty} \delta_{lm_\epsilon,(\mathcal{X}_s,\mathcal{D}_s)}(\epsilon\mathcal{H}_s + f_s^*\mathcal{O}(1)) \geq u - \delta_0$$

for any $\epsilon \in \mathbb{Q} \cap (0, \epsilon_0)$ and closed point $s \in S$. □

Now, we are ready to show Theorem 1.6 for the case when $u < 0$.

Theorem 6.4 *Let $\pi\colon(\mathcal{X},\mathcal{D})\to S$ be a log $\mathbb{Q}$-Gorenstein family and $f\colon\mathcal{X}\to\mathcal{P}$ a contraction over S, where $\mathcal{P}$ is a scheme that is projective and smooth over S. Let $\mathcal{H}$ be a π-ample $\mathbb{Q}$-divisor on $\mathcal{X}$ and $\mathcal{L}$ a Cartier divisor on $\mathcal{P}$. Suppose that there exists an $m\in\mathbb{Z}_{>0}$ such that $(\mathcal{P}_{\overline{s}},\mathcal{L}_{\overline{s}})=(\mathbb{P}^1,\mathcal{O}(m))$ for any geometric point $\overline{s}\in S$. Assume that $-(K_{\mathcal{X}/S}+\mathcal{D})\sim_{\mathbb{Q},S}(u/m)f^*\mathcal{L}$ for some $u\in\mathbb{Q}_{>0}$, and that $f_s\colon(\mathcal{X}_s,\mathcal{D}_s,\mathcal{H}_s)\to(\mathbb{P}^1,\mathcal{O}(1))$ is uniformly adiabatically K-stable for any closed point $s\in S$.*

Then there exists a positive rational number ϵ_0 such that $(\mathcal{X}_s,\mathcal{D}_s,\epsilon\mathcal{H}_s+f_s^*\mathcal{O}(1))$ is specially K-stable for any closed point $s\in S$ and rational number $\epsilon\in(0,\epsilon_0)$. Furthermore, there exists a positive rational number α such that

$$M^{\mathrm{NA}}_{\mathcal{D}_s}(\mathcal{Y},\mathcal{M})\geq\alpha(\mathcal{J}^{\epsilon\mathcal{H}_s+f_s^*\mathcal{O}(1)})^{\mathrm{NA}}(\mathcal{Y},\mathcal{M})$$

for any rational number $\epsilon\in(0,\epsilon_0)$, closed point $s\in S$, and normal semiample test configuration $(\mathcal{Y},\mathcal{M})$ for $(\mathcal{X}_s,\epsilon\mathcal{H}_s+f_s^*\mathcal{O}(1))$.

Proof As the argument in the second paragraph of the proof of Proposition 6.3, we may freely shrink or replacing S by S' with an étale morphism $S'\to S$. Thus, by the argument as in the first paragraph of the proof of Theorem 4.2, we may assume that $\mathcal{P}=\mathbb{P}^1_S$, $\mathcal{L}=\mathcal{O}(1)$ and $m=1$.

By Theorem 4.2 and [Hattori 2022, Theorem 4.4], there exists a positive rational number δ_0 such that

$$\delta_{(\mathbb{P}^1,B_s)}(\mathcal{O}(1))=\lim_{\epsilon\to 0}\delta_{(\mathcal{X}_s,\mathcal{D}_s)}(\epsilon\mathcal{H}_s+f_s^*\mathcal{O}(1))\geq u+\delta_0$$

for any closed point $s\in S$, where B_s is the discriminant divisor of $f_s\colon(\mathcal{X}_s,\mathcal{D}_s)\to\mathbb{P}^1$. From now on, we set $\mathscr{L}_{\epsilon,s}:=\epsilon\mathcal{H}_s+f_s^*\mathcal{O}(1)$ and $\delta(\epsilon,s):=\delta_{(\mathcal{X}_s,\mathcal{D}_s)}(\mathscr{L}_{\epsilon,s})$ for any closed point $s\in S$. By Proposition 6.3, there exists a positive rational number ϵ_0, which is independent of $s\in S$, such that

$$\delta(\epsilon,s)\geq u+\tfrac{1}{2}\delta_0$$

for any $\epsilon\in(0,\epsilon_0)$. Then this ϵ_0 satisfies the assertion of Theorem 6.4. Indeed, for any normal semiample test configuration $(\mathcal{Y},\mathcal{M})$ for $(\mathcal{X}_s,\mathscr{L}_{\epsilon,s})$, we have

$$M^{\mathrm{NA}}_{\mathcal{D}_s}(\mathcal{Y},\mathcal{M})\geq(\mathcal{J}^{K_{\mathcal{X}_s}+\mathcal{D}_s+\delta(\epsilon,s)\mathscr{L}_{\epsilon,s}})^{\mathrm{NA}}(\mathcal{Y},\mathcal{M})$$

by Theorem 2.30. We also have

$$K_{\mathcal{X}_s}+\mathcal{D}_s+\left(u+\tfrac{1}{2}\delta_0\right)\mathscr{L}_{\epsilon,s}\sim_{\mathbb{Q}}\tfrac{1}{2}\delta_0\mathscr{L}_{\epsilon,s}+u\epsilon\mathcal{H}_s.$$

By the ampleness of $\mathcal{H}_s$ and the argument in the proof of [Hattori 2022, Lemma 4.5], we see that $(\mathcal{X}_s,\mathscr{L}_{\epsilon,s})$ is $\mathrm{J}^{\mathcal{H}_s}$-semistable. Thus, we obtain

$$(\mathcal{J}^{K_{\mathcal{X}_s}+\mathcal{D}_s+\delta(\epsilon,s)\mathscr{L}_{\epsilon,s}})^{\mathrm{NA}}(\mathcal{Y},\mathcal{M})\geq\tfrac{1}{2}\delta_0(\mathcal{J}^{\mathscr{L}_{\epsilon,s}})^{\mathrm{NA}}(\mathcal{Y},\mathcal{M})$$

for any rational number $\epsilon\in(0,\epsilon_0)$. □

Proof of Theorem 1.6 We first treat the case when $u\neq 0$. By Lemma 3.1, Theorem 3.6 and Remark 3.7, it is sufficient to discuss the assertion of Theorem 1.6 for $\mathfrak{G}^{\mathrm{Car}}_{d,n,m^{d-1}v,u,m^d w}$ for some n and m instead of $\mathfrak{G}_{d,\Theta,v,u,w}$. In this situation, the assertion of Theorem 1.6 immediately follows from Theorem 4.2, Proposition 6.1 and Theorems 6.2 and 6.4.

Next, we deal with the case when $u=0$. As in the previous paragraph, it is enough to show the assertion for $\mathfrak{G}^{\mathrm{Car}}_{d,\Theta,v,0,w}$. There exist $I\in\mathbb{Z}_{>0}$ and $r\in\mathbb{Z}_{>0}$ in Corollary 3.8 and Remark 3.7 respectively, such that for any element $f\colon(X,\Delta,A)\to C\in\mathfrak{G}^{\mathrm{Car}}_{d,\Theta,v,0,w}$ and any line bundle H on C of degree one, $r(K_X+\Delta)$ is Cartier and $I(A+f^*H)$ is very ample. In particular, (X,Δ) is $(1/r)$-lc. Taking a small $\mathbb{Q}$-factorialization of X and applying the length of extremal rays, we see that $K_X+3dI(A+f^*H)$ is the pushdown of a big divisor; see [Birkar 2019, Lemma 2.46]. Because we have

$$3dI(A+f^*H)-\Delta\sim_{\mathbb{Q}}3dI(A+f^*H)+K_X,$$

we see that $3dI(A+f^*H)-\Delta$ is the pushdown of a big divisor. We also have

$$\mathrm{vol}(3dI(A+f^*H))\le(3dI)^d(w+dv).$$

Thus, we may apply [Birkar 2021b, Theorem 1.8] to (X,Δ) and $3dI(A+f^*H)$, and we obtain $\delta_0>0$, depending only on d, Θ, v and w, such that

$$\alpha_{(X,\Delta)}(A+f^*H)\ge\delta_0.$$

Here, we use the fact that $\alpha_{(X,\Delta)}(3dI(A+f^*H))=(3dI)^{-1}\alpha_{(X,\Delta)}(A+f^*H)$. By the inequality (6-1) in the proof of Theorem 6.2, we obtain

$$\delta_{(X,\Delta)}(\epsilon A+f^*H)\ge\frac{d+1}{d}\delta_0\quad\text{for any }\ 0<\epsilon<1.$$

By Theorem 2.30 and the fact that $K_X+\Delta\sim_{\mathbb{Q}}0$, for any element $f\colon(X,\Delta,A)\to C\in\mathfrak{G}^{\mathrm{Car}}_{d,\Theta,v,0,w}$, $(X,\Delta,\epsilon A+f^*H)$ is specially K-stable and

$$M^{\mathrm{NA}}_{\Delta}(\mathcal{X},\mathcal{M})\ge\frac{d+1}{d}\delta_0(\mathcal{J}^{\epsilon A+f^*H})^{\mathrm{NA}}(\mathcal{X},\mathcal{M})$$

for any normal semiample test configuration $(\mathcal{X},\mathcal{M})$ for $(X,\epsilon A+f^*H)$, degree-one line bundle H on C, and rational number $\epsilon\in(0,1)$. □

Proof of Corollary 1.7 By Theorem 1.5(2), there exists a positive integer w', depending only on d, v and u, such that for any $f\colon(X,\Delta,A)\to C$ as assumption and the general fiber F of f, $A+t'_AF$ is ample and $\mathrm{vol}(A+t'_AF)\le w'$ for some integer t'_A. For some positive rational number t''_A, we have $\mathrm{vol}(A+(t'_A+t''_A)F)=w'$. We set

$$t_A:=t'_A+t''_A\in\mathbb{Q}_{>0}.$$

Note that $A+t'_AF$ is a $\mathbb{Q}$-Cartier Weil divisor. Then Theorem 1.6 shows that there exists $\epsilon_0>0$, depending only on d, v and u, such that $(X,\Delta,(\epsilon/(1+\epsilon t''_A))(A+t'_AF)+F)$ is specially K-stable for any $0<\epsilon\le\epsilon_0$ and any f. Then so is $(X,\Delta,A+(t_A+\epsilon^{-1})F)$. Here, we use the fact that

$$\frac{\epsilon}{1+\epsilon t''_A}(A+t'_AF)+F=\frac{\epsilon}{1+\epsilon t''_A}(A+(t_A+\epsilon^{-1})F).$$

Since

$$w := \mathrm{vol}(A + (t_A + \epsilon_0^{-1})F) = w' + d\epsilon_0^{-1}v$$

is independent of f, it follows that $(X, \Delta, A + tF)$ is specially K-stable if $\mathrm{vol}(A + tF) \geq w$.

As noted in the proof of [Hattori 2022, Theorem 4.6], the last assertion follows from the special K-stability and [Chen 2021, Theorem 1.1; Zhang 2024, Corollary 5.2; Chen and Cheng 2021, Theorem 1.3]. □

Finally, we remark that Corollary 1.7 states that $\mathscr{M}_{d,v,u,r}$ has a family whose geometric fibers are specially K-stable in the following sense.

Remark 6.5 As in [Deligne and Mumford 1969, Section 5], we can construct the *universal family* $(\mathscr{U}, \mathscr{A}_{\mathscr{U}})$ over $\mathscr{M}_{d,v,u,w,r}$ as follows. In the proof of Theorem 5.1, $\mathscr{M}_{d,v,u,w,r}$ is the disjoint union of finitely many $\mathscr{M}_{d_1,d_2,d_3,h}$. Thus, we can construct a (Deligne–Mumford) stack

$$\mathscr{U} := \bigsqcup_{d_1,d_2,d_3,h} [\mathcal{U}_{N_{d_1,d_2,d_3,h}}/\mathrm{PGL}(d_1) \times \mathrm{PGL}(d_2) \times \mathrm{PGL}(d_3)],$$

where $\mathcal{U}_{N_{d_1,d_2,d_3,h}}$ is the universal family of $N_{d_1,d_2,d_3,h}$. Let $\widetilde{\mathscr{A}}_{N_{d_1,d_2,d_3,h}}$ be the line bundle on $\mathcal{U}_{N_{d_1,d_2,d_3,h}}$ as in the proof of Claim 1. Note that $\widetilde{\mathscr{A}}_{N_{d_1,d_2,d_3,h}}$ is ample over $N_{d_1,d_2,d_3,h}$. By construction, we see that $\widetilde{\mathscr{A}}_{N_{d_1,d_2,d_3,h}}$ is $\mathrm{PGL}(d_1) \times \mathrm{PGL}(d_2) \times \mathrm{PGL}(d_3)$-equivariant. By Example 2.13, we can construct $\mathscr{A}_{\mathscr{U}}$ as the line bundle on $\mathscr{U}$ whose restriction to $[\mathcal{U}_{N_{d_1,d_2,d_3,h}}/\mathrm{PGL}(d_1) \times \mathrm{PGL}(d_2) \times \mathrm{PGL}(d_3)]$ corresponds to $\widetilde{\mathscr{A}}_{N_{d_1,d_2,d_3,h}}$.

In the proof of Theorem 1.3, we have proved that $\mathscr{M}_{d,v,u,r}$ is an open and closed substack of $\mathscr{M}_{d,v,u,w'+dv,r}$ for some $w' > 0$. Using this fact, we obtain a family

$$(\mathscr{U}', \mathscr{A}_{\mathscr{U}'}) := (\mathscr{U} \times_{\mathscr{M}_{d,v,u,w'+dv,r}} \mathscr{M}_{d,v,u,r}, \mathscr{A}_{\mathscr{U}}|_{\mathscr{U}'})$$

over $\mathscr{M}_{d,v,u,r}$. If we take w' larger than w in Corollary 1.7, then all geometric fibers of $(\mathscr{U}', \mathscr{A}_{\mathscr{U}'})$ over $\mathscr{M}_{d,v,u,r}$ are specially K-stable.

References

[Alexeev 1996a] **V Alexeev**, *Log canonical singularities and complete moduli of stable pairs*, preprint (1996) arXiv alg-geom/9608013

[Alexeev 1996b] **V Alexeev**, *Moduli spaces $M_{g,n}(W)$ for surfaces*, from "Higher-dimensional complex varieties", de Gruyter, Berlin (1996) 1–22 MR

[Alper et al. 2020] **J Alper**, **H Blum**, **D Halpern-Leistner**, **C Xu**, *Reductivity of the automorphism group of K-polystable Fano varieties*, Invent. Math. 222 (2020) 995–1032 MR Zbl

[Ambro 2004] **F Ambro**, *Shokurov's boundary property*, J. Differential Geom. 67 (2004) 229–255 MR Zbl

[Birkar 2019] **C Birkar**, *Anti-pluricanonical systems on Fano varieties*, Ann. of Math. 190 (2019) 345–463 MR

[Birkar 2021a] **C Birkar**, *Boundedness and volume of generalised pairs*, preprint (2021) arXiv 2103.14935

[Birkar 2021b] **C Birkar**, *Singularities of linear systems and boundedness of Fano varieties*, Ann. of Math. 193 (2021) 347–405 MR

[Birkar 2022] **C Birkar**, *Moduli of algebraic varieties*, preprint (2022) arXiv 2211.11237

[Birkar 2023] **C Birkar**, *Geometry of polarised varieties*, Publ. Math. Inst. Hautes Études Sci. 137 (2023) 47–105 MR

[Birkar et al. 2010] **C Birkar**, **P Cascini**, **C D Hacon**, **J McKernan**, *Existence of minimal models for varieties of log general type*, J. Amer. Math. Soc. 23 (2010) 405–468 MR Zbl

[Blum and Jonsson 2020] **H Blum**, **M Jonsson**, *Thresholds, valuations, and K-stability*, Adv. Math. 365 (2020) art. id. 107062 MR Zbl

[Blum and Liu 2022] **H Blum**, **Y Liu**, *Openness of uniform K-stability in families of $\mathbb{Q}$-Fano varieties*, Ann. Sci. École Norm. Sup. 55 (2022) 1–41 MR Zbl

[Blum and Xu 2019] **H Blum**, **C Xu**, *Uniqueness of K-polystable degenerations of Fano varieties*, Ann. of Math. 190 (2019) 609–656 MR Zbl

[Boucksom 2014] **S Boucksom**, *Remarks on minimal models of degenerations*, unpublished note (2014) Available at `http://sebastien.boucksom.perso.math.cnrs.fr/notes/Remarks_minimal_models.pdf`

[Boucksom et al. 2017] **S Boucksom**, **T Hisamoto**, **M Jonsson**, *Uniform K-stability, Duistermaat–Heckman measures and singularities of pairs*, Ann. Inst. Fourier (Grenoble) 67 (2017) 743–841 MR Zbl

[Chen 2021] **G Chen**, *The J-equation and the supercritical deformed Hermitian–Yang–Mills equation*, Invent. Math. 225 (2021) 529–602 MR Zbl

[Chen and Cheng 2021] **X Chen**, **J Cheng**, *On the constant scalar curvature Kähler metrics, II: Existence results*, J. Amer. Math. Soc. 34 (2021) 937–1009 MR Zbl

[Codogni and Patakfalvi 2021] **G Codogni**, **Z Patakfalvi**, *Positivity of the CM line bundle for families of K-stable klt Fano varieties*, Invent. Math. 223 (2021) 811–894 MR Zbl

[Conrad 2000] **B Conrad**, *Grothendieck duality and base change*, Lecture Notes in Math. 1750, Springer (2000) MR Zbl

[Datar and Pingali 2021] **V V Datar**, **V P Pingali**, *A numerical criterion for generalised Monge–Ampère equations on projective manifolds*, Geom. Funct. Anal. 31 (2021) 767–814 MR Zbl

[Deligne and Mumford 1969] **P Deligne**, **D Mumford**, *The irreducibility of the space of curves of given genus*, Inst. Hautes Études Sci. Publ. Math. 36 (1969) 75–109 MR Zbl

[Dervan 2016] **R Dervan**, *Uniform stability of twisted constant scalar curvature Kähler metrics*, Int. Math. Res. Not. 2016 (2016) 4728–4783 MR Zbl

[Dervan and Naumann 2018] **R Dervan**, **P Naumann**, *Moduli of polarised manifolds via canonical Kähler metrics*, preprint (2018) arXiv 1810.02576

[Dervan and Ross 2019] **R Dervan**, **J Ross**, *Stable maps in higher dimensions*, Math. Ann. 374 (2019) 1033–1073 MR Zbl

[Dervan and Sektnan 2021a] **R Dervan**, **L M Sektnan**, *Moduli theory, stability of fibrations and optimal symplectic connections*, Geom. Topol. 25 (2021) 2643–2697 MR Zbl

[Dervan and Sektnan 2021b] **R Dervan**, **L M Sektnan**, *Optimal symplectic connections on holomorphic submersions*, Comm. Pure Appl. Math. 74 (2021) 2132–2184 MR Zbl

[Donaldson 2001] **S K Donaldson**, *Scalar curvature and projective embeddings, I*, J. Differential Geom. 59 (2001) 479–522 MR Zbl

[Donaldson 2002] **S K Donaldson**, *Scalar curvature and stability of toric varieties*, J. Differential Geom. 62 (2002) 289–349 MR Zbl

[Fantechi et al. 2005] **B Fantechi**, **L Göttsche**, **L Illusie**, **S L Kleiman**, **N Nitsure**, **A Vistoli**, *Fundamental algebraic geometry: Grothendieck's FGA explained*, Math. Surv. Monogr. 123, Amer. Math. Soc., Providence, RI (2005) MR Zbl

[Fine 2004] **J Fine**, *Constant scalar curvature Kähler metrics on fibred complex surfaces*, J. Differential Geom. 68 (2004) 397–432 MR Zbl

[Fine 2007] **J Fine**, *Fibrations with constant scalar curvature Kähler metrics and the CM-line bundle*, Math. Res. Lett. 14 (2007) 239–247 MR Zbl

[Fujiki and Schumacher 1990] **A Fujiki**, **G Schumacher**, *The moduli space of extremal compact Kähler manifolds and generalized Weil–Petersson metrics*, Publ. Res. Inst. Math. Sci. 26 (1990) 101–183 MR Zbl

[Fujino 2011] **O Fujino**, *Fundamental theorems for the log minimal model program*, Publ. Res. Inst. Math. Sci. 47 (2011) 727–789 MR Zbl

[Fujino 2012] **O Fujino**, *Basepoint-free theorems: saturation, b-divisors, and canonical bundle formula*, Algebra Number Theory 6 (2012) 797–823 MR Zbl

[Fujino 2017] **O Fujino**, *Effective basepoint-free theorem for semi-log canonical surfaces*, Publ. Res. Inst. Math. Sci. 53 (2017) 349–370 MR Zbl

[Fujino 2018] **O Fujino**, *Semipositivity theorems for moduli problems*, Ann. of Math. 187 (2018) 639–665 MR Zbl

[Fujita 2019] **K Fujita**, *Uniform K-stability and plt blowups of log Fano pairs*, Kyoto J. Math. 59 (2019) 399–418 MR Zbl

[Fujita and Odaka 2018] **K Fujita**, **Y Odaka**, *On the K-stability of Fano varieties and anticanonical divisors*, Tohoku Math. J. 70 (2018) 511–521 MR Zbl

[Gieseker 1977] **D Gieseker**, *Global moduli for surfaces of general type*, Invent. Math. 43 (1977) 233–282 MR Zbl

[Görtz and Wedhorn 2010] **U Görtz**, **T Wedhorn**, *Algebraic geometry, I: Schemes with examples and exercises*, Vieweg + Teubner, Wiesbaden, Germany (2010) MR Zbl

[Grothendieck 1966] **A Grothendieck**, *Éléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, III*, Inst. Hautes Études Sci. Publ. Math. 28 (1966) 5–255 MR Zbl

[Hacon and Xu 2013] **C D Hacon**, **C Xu**, *Existence of log canonical closures*, Invent. Math. 192 (2013) 161–195 MR Zbl

[Hacon et al. 2014] **C D Hacon**, **J McKernan**, **C Xu**, *ACC for log canonical thresholds*, Ann. of Math. 180 (2014) 523–571 MR Zbl

[Hacon et al. 2018] **C D Hacon**, **J McKernan**, **C Xu**, *Boundedness of moduli of varieties of general type*, J. Eur. Math. Soc. 20 (2018) 865–901 MR Zbl

[Hartshorne 1977] **R Hartshorne**, *Algebraic geometry*, Grad. Texts in Math. 52, Springer (1977) MR Zbl

[Hassett and Kovács 2004] **B Hassett**, **S J Kovács**, *Reflexive pull-backs and base extension*, J. Algebraic Geom. 13 (2004) 233–247 MR Zbl

[Hattori 2021] **M Hattori**, *A decomposition formula for J-stability and its applications* (2021) arXiv 2103.04603 To appear in Michigan Math. J.

[Hattori 2022] **M Hattori**, *On K-stability of Calabi–Yau fibrations*, preprint (2022) arXiv 2203.11460

[Hattori 2024a] **M Hattori**, *Minimizing CM degree and specially K-stable varieties*, Int. Math. Res. Not. 2024 (2024) 5728–5772 MR Zbl

[Hattori 2024b] **M Hattori**, *On fibration stability after Dervan–Sektnan and singularities*, Osaka J. Math. 61 (2024) 195–228 MR Zbl

[Hattori and Zanardini 2022] **M Hattori**, **A Zanardini**, *On the GIT stability of linear systems of hypersurfaces in projective space*, preprint (2022) arXiv 2212.09364

[Huybrechts and Lehn 1997] **D Huybrechts**, **M Lehn**, *The geometry of moduli spaces of sheaves*, Aspects Math. E31, Vieweg & Sohn, Braunschweig, Germany (1997) MR Zbl

[Jiang 2018] **C Jiang**, *On birational boundedness of Fano fibrations*, Amer. J. Math. 140 (2018) 1253–1276 MR Zbl

[Jiang 2020] **C Jiang**, *Boundedness of $\mathbb{Q}$-Fano varieties with degrees and alpha-invariants bounded from below*, Ann. Sci. École Norm. Sup. 53 (2020) 1235–1248 MR Zbl

[Kawakita 2007] **M Kawakita**, *Inversion of adjunction on log canonicity*, Invent. Math. 167 (2007) 129–133 MR Zbl

[Keel and Mori 1997] **S Keel**, **S Mori**, *Quotients by groupoids*, Ann. of Math. 145 (1997) 193–213 MR Zbl

[Kollár 1990] **J Kollár**, *Projectivity of complete moduli*, J. Differential Geom. 32 (1990) 235–268 MR Zbl

[Kollár 1993] **J Kollár**, *Effective base point freeness*, Math. Ann. 296 (1993) 595–605 MR Zbl

[Kollár 2008] **J Kollár**, *Hulls and husks*, preprint (2008) arXiv 0805.0576

[Kollár 2013] **J Kollár**, *Singularities of the minimal model program*, Cambridge Tracts in Math. 200, Cambridge Univ. Press (2013) MR Zbl

[Kollár 2023] **J Kollár**, *Families of varieties of general type*, Cambridge Tracts in Math. 231, Cambridge Univ. Press (2023) MR Zbl

[Kollár and Mori 1998] **J Kollár**, **S Mori**, *Birational geometry of algebraic varieties*, Cambridge Tracts in Math. 134, Cambridge Univ. Press (1998) MR Zbl

[Kollár and Shepherd-Barron 1988] **J Kollár**, **N I Shepherd-Barron**, *Threefolds and deformations of surface singularities*, Invent. Math. 91 (1988) 299–338 MR Zbl

[Lazarsfeld 2004] **R Lazarsfeld**, *Positivity in algebraic geometry, I: Classical setting: line bundles and linear series*, Ergebnisse der Math. 48, Springer (2004) MR Zbl

[Li and Xu 2014] **C Li**, **C Xu**, *Special test configuration and K-stability of Fano varieties*, Ann. of Math. 180 (2014) 197–232 MR Zbl

[Li et al. 2019] **C Li**, **X Wang**, **C Xu**, *On the proper moduli spaces of smoothable Kähler–Einstein Fano varieties*, Duke Math. J. 168 (2019) 1387–1459 MR Zbl

[Liu et al. 2022] **Y Liu**, **C Xu**, **Z Zhuang**, *Finite generation for valuations computing stability thresholds and applications to K-stability*, Ann. of Math. 196 (2022) 507–566 MR Zbl

[Matsumura 1980] **H Matsumura**, *Commutative algebra*, 2nd edition, Math. Lect. Note Ser. 56, Benjamin Cummings, Reading, MA (1980) MR Zbl

[Miranda 1980] **R Miranda**, *On the stability of pencils of cubic curves*, Amer. J. Math. 102 (1980) 1177–1202 MR Zbl

[Miranda 1981] **R Miranda**, *The moduli of Weierstrass fibrations over $\mathbb{P}^1$*, Math. Ann. 255 (1981) 379–394 MR Zbl

[Mumford 1966] **D Mumford**, *Lectures on curves on an algebraic surface*, Ann. of Math. Stud. 59, Princeton Univ. Press (1966) MR Zbl

[Mumford 1977] **D Mumford**, *Stability of projective varieties*, Enseign. Math. 23 (1977) 39–110 MR Zbl

[Mumford 2008] **D Mumford**, *Abelian varieties*, corrected reprint of 2nd edition, Tata Inst. Fund. Res. Stud. Math. 5, Tata Inst. Fund. Res., Bombay (2008) MR Zbl

[Mumford et al. 1994] **D Mumford**, **J Fogarty**, **F Kirwan**, *Geometric invariant theory*, 3rd edition, Ergebnisse der Math. 34, Springer (1994) MR Zbl

[Nakayama 1986] **N Nakayama**, *Invariance of the plurigenera of algebraic varieties under minimal model conjectures*, Topology 25 (1986) 237–251 MR Zbl

[Odaka 2010] **Y Odaka**, *On the GIT stability of polarized varieties: a survey*, lecture notes, Kinosaki Algebraic Geometry Symposium (2010) Available at `https://researchmap.jp/7000009334/presentations/37010386/attachment_file.pdf`

[Odaka 2012] **Y Odaka**, *The Calabi conjecture and K-stability*, Int. Math. Res. Not. 2012 (2012) 2272–2288 MR Zbl

[Odaka 2021] **Y Odaka**, *On log minimality of weak K-moduli compactifications of Calabi–Yau varieties*, preprint (2021) arXiv 2108.03832

[Olsson 2016] **M Olsson**, *Algebraic spaces and stacks*, Amer. Math. Soc. Colloq. Publ. 62, Amer. Math. Soc., Providence, RI (2016) MR Zbl

[Patakfalvi et al. 2018] **Z Patakfalvi**, **K Schwede**, **W Zhang**, *F-singularities in families*, Algebr. Geom. 5 (2018) 264–327 MR Zbl

[Seiler 1987] **W K Seiler**, *Global moduli for polarized elliptic surfaces*, Compos. Math. 62 (1987) 187–213 MR Zbl

[Song 2020] **J Song**, *Nakai–Moishezon criterions for complex Hessian equations*, preprint (2020) arXiv 2012.07956

[Stacks 2005–] **P Belmans**, **A J de Jong**, et al., *The Stacks project*, electronic reference (2005–) Available at `http://stacks.math.columbia.edu`

[Tian 1987] **G Tian**, *On Kähler–Einstein metrics on certain Kähler manifolds with $C_1(M) > 0$*, Invent. Math. 89 (1987) 225–246 MR Zbl

[Tian 1990] **G Tian**, *On Calabi's conjecture for complex surfaces with positive first Chern class*, Invent. Math. 101 (1990) 101–172 MR Zbl

[Tian 1997] **G Tian**, *Kähler–Einstein metrics with positive scalar curvature*, Invent. Math. 130 (1997) 1–37 MR Zbl

[Vakil 2010] **R Vakil**, *The rising sea: foundations of algebraic geometry* (2010) Book manuscript to be published by Princeton Univ. Press. Available at `https://math.stanford.edu/~vakil/216blog/FOAGaug2922public.pdf`

[Viehweg 1995] **E Viehweg**, *Quasi-projective moduli for polarized manifolds*, Ergebnisse der Math. 30, Springer (1995) MR Zbl

[Xu and Zhuang 2020] **C Xu**, **Z Zhuang**, *On positivity of the CM line bundle on K-moduli spaces*, Ann. of Math. 192 (2020) 1005–1068 MR Zbl

[Xu and Zhuang 2021] **C Xu**, **Z Zhuang**, *Uniqueness of the minimizer of the normalized volume function*, Camb. J. Math. 9 (2021) 149–176 MR Zbl

[Zanardini 2023] **A Zanardini**, *Stability of pencils of plane sextics and Halphen pencils of index two*, Manuscripta Math. 172 (2023) 353–374 MR Zbl

[Zhang 2024] **K Zhang**, *A quantization proof of the uniform Yau–Tian–Donaldson conjecture*, J. Eur. Math. Soc. 26 (2024) 4763–4778 MR Zbl

Department of Mathematics, Faculty of Science, Niigata University
Niigata, Japan

Department of Mathematics, Graduate School of Science, Kyoto University
Kyoto, Japan

hkenta@math.sc.niigata-u.ac.jp, hattori.masafumi.47z@st.kyoto-u.ac.jp

Proposed: Mark Gross
Seconded: Dan Abramovich, Sándor J Kovács

Received: 3 August 2023
Revised: 7 June 2024

Guidelines for Authors

Submitting a paper to Geometry & Topology

Papers must be submitted using the upload page at the GT website. You will need to choose a suitable editor from the list of editors' interests and to supply MSC codes.

The normal language used by the journal is English. Articles written in other languages are acceptable, provided your chosen editor is comfortable with the language and you supply an additional English version of the abstract.

Preparing your article for Geometry & Topology

At the time of submission you need only supply a PDF file. Once accepted for publication, the paper must be supplied in LaTeX, preferably using the journal's class file. More information on preparing articles in LaTeX for publication in GT is available on the GT website.

`arXiv` papers

If your paper has previously been deposited on the `arXiv`, we will need its `arXiv` number at acceptance time. This allows us to deposit the DOI of the published version on the paper's `arXiv` page.

References

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited at least once in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used, but will be converted to the house style (see a current issue for examples).

Figures

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Fuzzy or sloppily drawn figures will not be accepted. For labeling figure elements consider the pinlabel LaTeX package, but other methods are fine if the result is editable. If you're not sure whether your figures are acceptable, check with production by sending an email to graphics@msp.org.

Proofs

Page proofs will be made available to authors (or to the designated corresponding author) in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Geometry & Topology

Volume 29 Issue 3 (pages 1115–1691) 2025