Algebraic & Geometric Topology

Volume 26 (2026)

Issue 5 (pages 1597–1963)

msp

ALGEBRAIC & GEOMETRIC TOPOLOGY

msp.org/agt

EDITORS

PRINCIPAL ACADEMIC EDITORS

John Etnyre etnyre@math.gatech.edu Georgia Institute of Technology	Vesna Stojanoska vesna@illinois.edu University of Illinois at Urbana-Champaign

BOARD OF EDITORS

Julie Bergner	University of Virginia jeb2md@eservices.virginia.edu
Steven Boyer	Université du Québec à Montréal cohf@math.rochester.edu
Tara E Brendle	University of Glasgow tara.brendle@glasgow.ac.uk
Indira Chatterji	CNRS & Univ. Côte d’Azur (Nice) indira.chatterji@math.cnrs.fr
Octav Cornea	Université’ de Montreal cornea@dms.umontreal.ca
Alexander Dranishnikov	University of Florida dranish@math.ufl.edu
Tobias Ekholm	Uppsala University, Sweden tobias.ekholm@math.uu.se
Mario Eudave-Muñoz	Univ. Nacional Autónoma de México mario@matem.unam.mx
David Futer	Temple University dfuter@temple.edu
John Greenlees	University of Warwick john.greenlees@warwick.ac.uk
Matthew Hedden	Michigan State University mhedden@math.msu.edu
Kristen Hendricks	Rutgers University kristen.hendricks@rutgers.edu
Hans-Werner Henn	Université Louis Pasteur henn@math.u-strasbg.fr
Kathryn Hess	École Polytechnique Féd. de Lausanne kathryn.hess@epfl.ch
Daniel Isaksen	Wayne State University isaksen@math.wayne.edu
Thomas Koberda	University of Virginia thomas.koberda@virginia.edu
Markus Land	JGU Mainz mland@uni-mainz.de
Christine Lescop	Université Joseph Fourier lescop@ujf-grenoble.fr
Norihiko Minami	OCAMI (Osaka Central Adv. Math. Inst.) norihikominami@gmail.com
Andrés Navas	Universidad de Santiago de Chile andres.navas@usach.cl
Jessica S Purcell	Monash University jessica.purcell@monash.edu
Birgit Richter	Universität Hamburg birgit.richter@uni-hamburg.de
Jérôme Scherer	École Polytech. Féd. de Lausanne jerome.scherer@epfl.ch
Zoltán Szabó	Princeton University szabo@math.princeton.edu
Maggy Tomova	University of Iowa maggy-tomova@uiowa.edu
Daniel T Wise	McGill University, Canada daniel.wise@mcgill.ca
Lior Yanovski	Hebrew University of Jerusalem lior.yanovski@gmail.com

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

The subscription price for 2026 is US $795/year for the electronic version, and $1170/year (+$80, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP. Algebraic & Geometric Topology is indexed by Mathematical Reviews, Zentralblatt MATH, Current Mathematical Publications and the Science Citation Index.

Algebraic & Geometric Topology (ISSN 1472-2747 printed, 1472-2739 electronic) is published 9 times per year and continuously online, by Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004. Periodical rate postage paid at Oakland, CA 94615-9651, and additional mailing offices. POSTMASTER: send address changes to Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004.

AGT peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY
mathematical sciences publishers
nonprofit scientific publishing
https://msp.org/

© 2026 Mathematical Sciences Publishers

Algebraic & Geometric Topology **26**:5 (2026) 1597–1634
DOI: 10.2140/agt.2026.26.1597
Published: May 29, 2026

On homology concordance in contractible manifolds and two-bridge links

HUGO ZHOU

Let $\widehat{\mathcal{C}}_{\mathbb{Z}}$ be the group which consists of manifold-knot pairs (Y, K) modulo homology concordance, where Y is an integer homology sphere bounding an integer homology ball, and let $\mathcal{C}_{\mathbb{Z}}$ be the subgroup consisting of pairs (S^3, K). Dai, Hom, Stoffregen and Truong showed that the quotient group $\widehat{\mathcal{C}}_{\mathbb{Z}}/\mathcal{C}_{\mathbb{Z}}$ admits a $\mathbb{Z}^\infty$-summand. In this paper, we improve the result by showing that there exists a family $\{(Y, K_m)\}_{m>1}$ generating a $\mathbb{Z}^\infty$-summand where Y is the boundary of a smooth contractible 4-manifold. In fact, we give a $\mathbb{Z}$-count of such families.

The examples are constructed using a family of knots obtained by blowing down a component of a two-bridge link. They are studied in Jonathan Hales's thesis. Using the algorithm due to Ozsváth, Szabó and Hales we give a classification of the knot Floer homology of a larger family of such knots that might be of independent interest.

1 Introduction

The integer homology concordance group $\widehat{\mathcal{C}}_{\mathbb{Z}}$ consists of pairs (Y, K) where Y is an integer homology sphere bounding an integer homology ball and K is a knot in Y, where the group operation is induced by the connected sum. Two classes (Y_1, K_1) and (Y_2, K_2) are equivalent if and only if there exists a pair $\partial(W, \Sigma) = (Y_1, K_1) \sqcup -(Y_2, K_2)$, where W is an integer homology cobordism and Σ a smoothly embedded cylinder in W. The subgroup $\mathcal{C}_{\mathbb{Z}}$ consists of pairs (S^3, K). A class (Y, K) is nonzero in $\widehat{\mathcal{C}}_{\mathbb{Z}}/\mathcal{C}_{\mathbb{Z}}$ if and only if K is not concordant to any knot in S^3 in any homology cobordism, or equivalently if and only if K does not bound any PL-disk in any homology ball with boundary Y.

The first nontrivial class $(Y, K) \in \widehat{\mathcal{C}}_{\mathbb{Z}}/\mathcal{C}_{\mathbb{Z}}$ was found by Levine [13], building on Akbulut's work [1]. Hom, Levine and Lidman [11] proved that $\widehat{\mathcal{C}}_{\mathbb{Z}}/\mathcal{C}_{\mathbb{Z}}$ is infinitely generated and admits a $\mathbb{Z}$-subgroup. Using the infinite family found in [20], Dai, Hom, Stoffregen and Truong [6] showed that $\widehat{\mathcal{C}}_{\mathbb{Z}}/\mathcal{C}_{\mathbb{Z}}$ admits a $\mathbb{Z}^\infty$-summand.

Among each infinite family of manifold-knot pairs which gives rise to the $\mathbb{Z}^\infty$-summand in the literature so far, the manifolds are always of the form $M_n \# -M_n$, such that they bound homology balls, but not necessarily smooth contractible manifolds. This paper strengthens the existing result by giving an infinite family of knots in the boundary of a smooth contractible manifold that generates a $\mathbb{Z}^\infty$-summand in $\widehat{\mathcal{C}}_{\mathbb{Z}}/\mathcal{C}_{\mathbb{Z}}$. In fact, we give a $\mathbb{Z}$-count of such families.

Imposing that the homology spheres bound smooth contractible manifolds guarantees that every knot in the boundary has trivial image in the fundamental group of the 4-manifold under the inclusion map.

MSC2020: 57K10, 57K18.

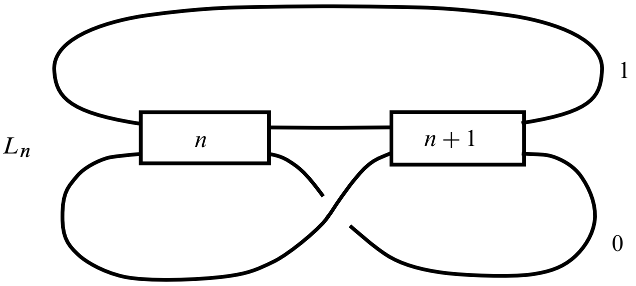


Figure 1: The two-bridge link L_n whose $(1,0)$-framed surgery yields the integer homology sphere Y_n that bounds a contractible manifold. The number in the box indicates the number of right-handed full-twists.

Therefore, there is no homotopic obstruction for the knots to bound PL-disks, making it a more interesting and more difficult question.

Consider the Mazur-type manifold Y_n with $n \geq 1$ (see [3]) as depicted in Figure 1, obtained from a $(1,0)$-framed surgery on a two-bridge link L_n. From a standard argument that switches the 0-framed two handle to a dotted circle, we see that Y_n bounds a contractible 4-manifold. (Both components of L_n are unknotted and they intersect algebraically once.) Let $K_n \subset S^3$ be the knot obtained from L_n by blowing down the $+1$-framed unknot component. It follows that $S^3_{-1}(K_n) = Y_n$ bounds a contractible 4-manifold.

Let $\mu^{(2n)}_{m,1}$ be the image of the $(m,1)$-cable of the meridian in the -1-surgery on K_{2n}.

Theorem 1.1 *The family* $\{(Y_{2n}, \mu^{(2n)}_{m,1})\}_{n>0,m>1}$ *generates a* $\mathbb{Z}^{\mathbb{Z}\oplus\mathbb{Z}}$ *summand in* $\widehat{\mathcal{C}}_{\mathbb{Z}}/\mathcal{C}_{\mathbb{Z}}$. *In particular, for each fixed* $n > 0$, $\{(Y_{2n}, \mu^{(2n)}_{m,1})\}_{m>1}$ *generates a* $\mathbb{Z}^{\mathbb{Z}}$ *summand in* $\widehat{\mathcal{C}}_{\mathbb{Z}}/\mathcal{C}_{\mathbb{Z}}$.

As abelian groups, $\mathbb{Z}^{\mathbb{Z}\oplus\mathbb{Z}}$ is of course isomorphic to $\mathbb{Z}^{\mathbb{Z}}$. Here by a $\mathbb{Z}^{\mathbb{Z}\oplus\mathbb{Z}}$ summand we would like to emphasize the existence of a (natural) two-parameter family of linear independent, surjective homomorphisms from $\widehat{\mathcal{C}}_{\mathbb{Z}}/\mathcal{C}_{\mathbb{Z}}$ to $\mathbb{Z}$. In this case the two parameters are given by n and m. (See Lemma 2.6 for the homomorphisms.) Note that the $\mathbb{Z}^{\mathbb{Z}}$ summand in [11; 20] is generated by knots living in infinitely many different manifolds, while here each family of knots $\{\mu^{(2n)}_{m,1}\}_{m>1}$ lives in the same 3-manifold Y_{2n}, which is the boundary of a smooth contractible 4-manifold.

Theorem 1.1 is the direct consequence of the two following theorems. Denote by C_k the complex isomorphic to $\mathrm{CFK}^\infty(S^3, T_{2,2k+1})$. First, using the filtered mapping cone formula [21], and with the help of the concordance homomorphisms defined in [6], we show the following:

Theorem 1.2 *Suppose a family of knots* $\{J_{2k}\}_{k>0}$ *satisfies that* $\mathrm{CFK}^\infty(S^3, J_{2k}) \cong C_{2k} \oplus A$, *where* $H_*(A) = 0$. *Then the family* $\{(S^3_{-1}(J_{2k}), \mu^{(2k)}_{m,1})\}_{k>0,m>1}$ *generates a* $\mathbb{Z}^{\mathbb{Z}\oplus\mathbb{Z}}$ *summand in* $\widehat{\mathcal{C}}_{\mathbb{Z}}/\mathcal{C}_{\mathbb{Z}}$, *where* $\mu^{(2k)}_{m,1}$ *is the image of the* $(m,1)$*-cable of the meridian in the* -1*-surgery on* J_{2k}. *In particular, for each fixed* $k > 0$, $\{(S^3_{-1}(J_{2k}), \mu^{(2k)}_{m,1})\}_{m>1}$ *generates a* $\mathbb{Z}^{\mathbb{Z}}$ *summand in* $\widehat{\mathcal{C}}_{\mathbb{Z}}/\mathcal{C}_{\mathbb{Z}}$.

Next, the knot Floer complex of the knot family K_n can be explicitly computed as follows. Ozsváth and Szabó [16, Section 6.2] give a description for a genus-one doubly pointed Heegaard diagram of any knot that results from blowing down the ±1-surgery on one component of a two-bridge link. The

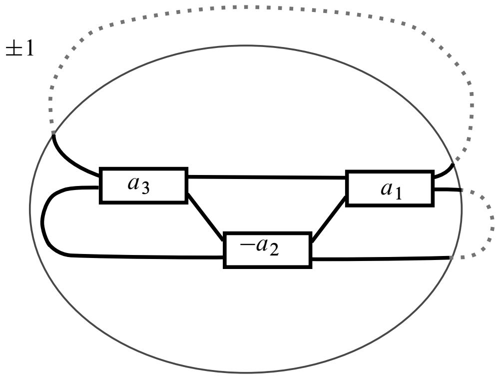


Figure 2: Blowing down one component of the closure of a rational tangle $[a_1, a_2, a_3]$ with a_1 and a_3 even yields the knot $K^{\pm}[a_1, a_2, a_3]$. The numbers in the boxes indicate the number of right-handed half-twists.

family of knots K_n is studied carefully in Jonathan Hales's thesis [8, Section 3]. Building on Ozsváth and Szabó's description, Hales develops a simple algorithm that outputs a $(1,1)$ diagram of all the knots that arise from blowing down the ± 1-surgery on one component of a two-bridge link. We review the details of this algorithm in Section 3. In particular, he proves the following theorem.

Theorem 1.3 [8] *For* $n > 0$, $\mathrm{CFK}^{\infty}(S^3, K_n) \cong C_n \oplus A$, *where* $H_*(A) = 0$.

See the $k = 1$ case of Proposition 4.27 for a more precise statement of Theorem 1.3. It is clear that Theorem 1.1 is the consequence of Theorems 1.2 and 1.3.

Remark 1.4 The family of knots K_n we consider here is the same family used to prove the infinite generation in [11], where the d-invariants of $1/p$ surgeries on $\mu_{1,1}^{(n)}$ are computed, using the fact that for certain p, the resulting manifolds are Brieskorn spheres. This is also similar to the approach employed in [2]. Interestingly, with the current computation we do not recover the infinite generation result with $\mu_{1,1}^{(n)}$ proved in [11], as it turns out that $\mathrm{CFK}^{\infty}(S^3_{-1}(K_n), \mu_{1,1}^{(n)})$ is locally equivalent to the trivial complex. In order to recover their result, one needs to in addition consider the nontrivial flip map over $\mathrm{CFK}^{\infty}(S^3_{-1}(K_n), \mu_{1,1}^{(n)})$.

1.1 Blowing down two-bridge links

We also explore the algorithm due to Ozsváth, Szabó and Hales. For a rational tangle $[a_1, \dots, a_\ell]$ whose closure is a two-component link, we can always arrange such that ℓ is odd and each a_i is even when i is odd. (See Lemma 3.2 and the discussion before that.) Denote by $K^{\pm}([a_1, \dots, a_\ell])$ the knot obtained from blowing down the ± 1-framed upper component. See Figure 2. It turns out that we can completely determine the knot Floer homology of $K^{\pm}([a_1, \dots, a_\ell])$ when $\ell \leq 3$. We include this classification result in the paper, with the expectation that these families of knots will generate more future applications.

The computation of the invariants has always been a main theme in Heegaard Floer homology. Despite significant development of the theory, there are still only a limited number of knot families whose knot

Floer complexes can be explicitly determined: L-space knots [17], thin knots [15; 18], $(1,1)$ almost L-space knots [4] (all determined by the Alexander polynomial) and certain cables of specific knots (using say [9], where the computations are already difficult).

Theorem 1.5 *The knot Floer complex of any knot* $K^{\pm}([a_1,\dots,a_\ell])$ *with* $\ell = 1, 3$ *and* a_i *even for odd* i *is classified*, *including Maslov gradings and filtration levels.*

The knot Floer chain homotopy types arise from our examples are novel. Consider the complexes D_s in Definition 4.20, the complexes $C_{n,k}$ in Definition 4.25 and the complexes $C'_{n,k}$ in Definition 4.30. Recall $C_n \cong \mathrm{CFK}^\infty(S^3, T_{2,2n+1})$ and let C_0 be the complex generated by a single element. We will show that the knot Floer complexes of various knots $K^{\pm}[n, 1, n+k]$ with $n > 0$ and $k \geq 0$ consists of direct summands of the above complexes. (For detailed statements, see Propositions 4.27 and 4.31.) As mentioned earlier, the case for K^+ when $k = 1$ has already been done by Jonathan Hales. We include it into a framework suitable for a slightly larger family.

Moreover, this classification involves interesting techniques. For a general sequence $[a_1, b, a_3]$, we observe that changing b by 2 amounts to performing a full Dehn-twist around two basepoints. Concretely, if we let γ be an arc between the two base points, the Dehn-twist is performed along the closed loop $\gamma \cup_{S^0} \overline{\gamma}$, where $\overline{\gamma}$ is a copy of γ with the reversed orientation. See Figure 7 on page 1614. This is a local transformation in a neighborhood of γ, so we can opt to perform it in the very end. It turns out that each full Dehn-twist amounts to adding a box summand at the "closest" generator near γ. See Figure 8 on page 1616 for the effect on CFK^∞ and see Definition 4.8 for a precise definition of the "closest" generator. It follows that in order to recover any general sequence $[a_1, b, a_3]$, we need only consider the case $b = \pm 1$ and 0 with a *marked basis*. See Proposition 4.11 for more details.

This method allows us to reduce the full classification to a handful of cases. For instance, in Section 4.3 we will utilize it to show that the knot Floer complex corresponding to the sequence $[a_1, b, a_3]$ when b is even is isomorphic to $\mathrm{CFK}^\infty(S^3, T_{n,n+1})$ for certain n with a certain number of box summands added to some marked generators.

Remark 1.6 When $\ell > 3$, even though a closed formula seems unlikely, using similar methods as in this paper it is still very much feasible to determine the full knot Floer complex for singular examples of $K^{\pm}([a_1,\dots,a_\ell])$. Moreover, if one opts to compute partial invariants such as the local equivalence class or the τ invariant, then I believe the computation should be practical for a larger family of knots, and I expect the result to be interesting as well.

Organization

We perform the filtered mapping cone computations and prove Theorem 1.2 in Section 2. We review the algorithm by Ozsváth, Szabó and Hales in Section 3 and prove our classification result in Section 4. More detailedly, in Section 4.1, we prove the length-one case; in Section 4.2, we discuss the length-three case in general and prove some helpful lemmas; in Sections 4.3 through 4.7, we deal with the remaining cases with length-three.

2 Knot Floer homology of the cables of the knot meridian

In this section we prove Theorem 1.2. We will make use of the concordance invariants $\varphi_{i,j}$ defined in [5]. We refer the reader to the original paper for the detailed definitions. See [5, Section 3.2] for a worked out example. See [21, Example 8.1] for yet another example. Since $\varphi_{i,j}$ factors through the group of knotlike complexes over local equivalence [5, Proposition 4.42], instead of considering -1-surgery on the knot J_{2k} with $\mathrm{CFK}^\infty(S^3, J_{2k}) \cong \mathrm{CFK}^\infty(S^3, T_{4k+1}) \oplus A$, where $H_*(A) = 0$, it suffices to consider the -1-surgery on T_{4k+1}.

Theorem 1.2 follows from a computational result, Proposition 2.5. The computational tool is the filtered mapping cone formula [21], which combines [10; 19] to give a description of the knot Floer complex of the $(m, 1)$-cable of the meridian in the image of a surgery along a knot. We include a brief review of this filtered mapping cone formula in the following subsection.

2.1 Preliminaries on the filtered mapping cone formula

Let $K \subset S^3$ be a knot with genus equal to g. For a given positive integer m, let $\mu_{m,1}$ denote the $(m, 1)$-cable of the meridian of K in the -1-surgery on K. According to [21, Theorem 1.9], the knot Floer complex $\mathrm{CFK}^\infty(S^3_{-1}(K), \mu_{m,1})$ is filtered chain homotopy equivalent to the doubly filtered chain complex $X^\infty_m(K)$, defined to be the mapping cone of

$$\bigoplus_{s=-g+1}^{g+m-1} A^\infty_s \xrightarrow{v^\infty_s + h^\infty_s} \bigoplus_{s=-g}^{g+m-1} B^\infty_s, \tag{1}$$

where each A^∞_s and B^∞_s are isomorphic to $\mathrm{CFK}^\infty(S^3, K)$. The map $v^\infty_s : A_s \to B_s$ is the identity and the map $h^\infty_s : A_s \to B_{s-1}$ is the reflection along $i = j$ precomposed with U^s. Note that in general h^∞_s is a graded filtered chain homotopy equivalence with respect to j-filtration on the domain and i-filtration on the range. For knots in S^3, since $\mathrm{HF}^-(S^3)$ is one-dimensional, h^∞_s is unique up to filtered chain homotopy equivalence and therefore we may take it to be the reflection map. See [10, Lemma 2.18].

Let $\mathcal{I}$ and $\mathcal{J}$ be the double filtrations on the filtered mapping cone complex $X^\infty_m(K)$. We have for $[\boldsymbol{x}, i, j] \in A_s$,

$$\mathcal{I}([\boldsymbol{x}, i, j]) = \max\{i, j - s\}, \tag{2}$$

$$\mathcal{J}([\boldsymbol{x}, i, j]) = \max\{i - m, j - s\} - ms + \tfrac{1}{2}m(m+1), \tag{3}$$

and for $[\boldsymbol{x}, i, j] \in B_s$,

$$\mathcal{I}([\boldsymbol{x}, i, j]) = i, \tag{4}$$

$$\mathcal{J}([\boldsymbol{x}, i, j]) = i - m - ms + \tfrac{1}{2}m(m+1). \tag{5}$$

It is straightforward to check that for $s < -g + 1$, the map h_s induces an isomorphism on the homology; for $s > g + m - 1$, the map $v_s(K)$ induces an isomorphism on the homology, which justifies the truncation of the mapping cone.

The general strategy for computation involves finding a *reduced* basis for $X_m^\infty(K)$, where every term in the differential strictly lowers at least one of the filtrations. This can be achieved through a cancellation process (see, for example, [14, Proposition 11.57]) as follows: Suppose $\partial x_i = y_i +$ lower filtration terms, where the double filtration of y_i is the same as x_i. Then the subcomplex of $X_m^\infty(K)$ generated by all such $\{x_i, \partial x_i\}$ is acyclic, and $X_m^\infty(K)$ quotient by this complex is reduced. Alternatively, one can view the above process as a change of basis that splits off acyclic summands which individually lie entirely in one double-filtration level.

2.2 Computation

Recall that it suffices to consider the $(m,1)$-cable of the knot meridian in the -1-surgery on T_{4k+1}. For $k \geq 1$, the complex $C_{2k} = \mathrm{CFK}^\infty(S^3, T_{2,4k+1})$ is generated by a_i for $i = 1, \dots, 2k$ with coordinate $(0, -2k+2i-1)$ and b_i for $i = 1, \dots, 2k+1$ with coordinate $(0, -2k+2i-2)$. The Maslov grading is supported in b_{2k+1} and the differentials are given by

$$\partial a_i = Ub_{i+1} + b_i \quad \text{for } i = 1, \dots, 2k.$$

Via the isomorphism with $\mathrm{CFK}^\infty(S^3, T_{2,4k+1})$, denote the generators in A_s by $a_i^{(s)}$ for $i = 1, \dots, 2k$ and $b_i^{(s)}$ for $i = 1, \dots, 2k+1$ and the generators in B_s by $a_i'^{(s)}$ and $b_i'^{(s)}$, where $s = -2k+1, \dots, 2k+m-1$. The differential on the mapping cone is given by

$$\begin{aligned}\partial a_i^{(s)} &= b_i^{(s)} + Ub_{i+1}^{(s)} + a_i'^{(s)} + U^{s+2k+1-2i}(a_{2k-i+1}'^{(s-1)}),\\ \partial b_i^{(s)} &= b_i'^{(s)} + U^{s+2k+2-2i}(b_{2k-i+2}'^{(s-1)}),\\ \partial a_i'^{(s)} &= b_i'^{(s)} + Ub_{i+1}'^{(s)}.\end{aligned}$$

We first look to choose a reduced basis for the complex of $X_m^\infty(T_{2,4k+1})$. As a subcomplex of $X_m^\infty(T_{2,4k+1})$, each B_s is one dimensional. Indeed, quotienting out $\{a_i'^{(s)}, \partial a_i'^{(s)}\}_{1\leq i\leq 2k}$ leaves us with the sole generator $b_{2k+1}'^{(s)}$; define $\beta_s = b_{2k+1}'^{(s)}$ after the change of basis.

Next, for the complex A_s, observe that $a_i^{(s)}$ and $Ub_{i+1}^{(s)}$ are in the same coordinate if $-2k+2i-1 \geq s$, and similarly $a_i^{(s)}$ and $b_i^{(s)}$ are in the same coordinate if $-2k+2i-1 \leq s-m$. So we may quotient out $\{a_i^{(s)}, \partial a_i^{(s)}\}$ for $i \geq k + \frac{s+1}{2}$ and $i \leq k + \frac{s-m+1}{2}$. We have obtained a reduced model of $X_m^\infty(T_{2,4k+1})$. As a notational shorthand, let us define

$$f(m,s) = \frac{m(m+1)}{2} - ms, \quad i_{m,s}^{(t)} = \min\left\{k + \left\lceil \frac{s-1}{2} \right\rceil, 2k\right\}, \quad i_{m,s}^{(b)} = \max\left\{k + 1 + \left\lceil \frac{s-m}{2} \right\rceil, 1\right\}$$

such that in the reduced model each complex A_s is generated by $a_i^{(s)}$ with $i_{m,s}^{(b)} \leq i \leq i_{m,s}^{(t)}$ and $b_i^{(s)}$ with $i_{m,s}^{(b)} \leq i \leq i_{m,s}^{(t)} + 1$. The induced differentials on the chain complex are

$$\partial a_i^{(s)} = b_i^{(s)} + Ub_{i+1}^{(s)}, \tag{6}$$

$$\partial b_i^{(s)} = U^{2k+1-i}\beta_s + U^{2k+1-i+s}\beta_{s-1}, \tag{7}$$

and the filtration level of the generators are

$$\mathcal{J}(a_i^{(s)}) = f(m,s) - 2k - s - 1 + 2i, \tag{8}$$

$$\mathcal{J}(b_i^{(s)}) = f(m,s) - 2k - s - 2 + 2i, \tag{9}$$

$$\mathcal{J}(\beta_s) = f(m,s+1), \tag{10}$$

$$\mathcal{I}(a_i^{(s)}) = \mathcal{I}(b_i^{(s)}) = \mathcal{I}(\beta_s) = 0 \tag{11}$$

for $-2k+1 \le s \le 2k+m-1$ and suitable i as discussed above. For the purpose of computing concordance invariants, we will show the mapping cone can be further truncated. Let

$$X_m^\infty(T_{2,4k+1})\langle \ell \rangle = \bigoplus_{s=-\ell+m}^{\ell} A_s^\infty \xrightarrow{v_s^\infty + h_s^\infty} \bigoplus_{s=-\ell+m-1}^{\ell} B_s^\infty.$$

Note that under this notation $X_m^\infty(T_{2,4k+1}) = X_m^\infty(T_{2,4k+1})\langle 2k+m-1 \rangle$.

Lemma 2.1 *For any $k \ge 1$ and $m \ge 2$, the filtered complex $X_m^\infty(T_{2,4k+1})\langle 2k+m-1 \rangle$ is isomorphic to $X_n^\infty(T_{2,4k+1})\langle m-1 \rangle \oplus D$ up to a filtered change of basis, where $H_*(D) = 0$.*

Proof We will show that for $m \le \ell \le 2k+m-1$, the complex $X_m^\infty(T_{2,4k+1})\langle \ell \rangle$ is isomorphic to $X_m^\infty(T_{2,4k+1})\langle \ell - 1 \rangle \oplus D'$ up to a change of basis, where $H_*(D') = 0$. For every such ℓ, in $X_m^\infty(T_{2,4k+1})\langle \ell \rangle$ perform a change of basis

$$\beta_{-\ell+m-1} \mapsto \beta_{-\ell+m-1} + U^{\ell-m}\beta_{-\ell+m}, \quad \beta_\ell \mapsto \beta_\ell + U^\ell \beta_{\ell-1}.$$

By (7), as a result the complexes given by

$$A_{-\ell+m}^\infty \xrightarrow{h_{-\ell+m}} B_{-\ell+m-1}^\infty, \quad A_\ell^\infty \xrightarrow{v_\ell} B_\ell^\infty$$

both become summands under the new basis. The change of basis is clearly filtered with respect to the $\mathcal{I}$–filtration. For the $\mathcal{J}$–filtration, we compute

$$\begin{aligned}\mathcal{J}(\beta_{-\ell+m-1}) - \mathcal{J}(U^{\ell-m}\beta_{-\ell+m}) &= \mathcal{J}(\beta_{-\ell+m-1}) - (\mathcal{J}(\beta_{-\ell+m}) - \ell + m) \\ &= f(m,-\ell+m) - f(m,-\ell+m+1) + \ell - m = \ell \ge 0, \\ \mathcal{J}(\beta_\ell) - \mathcal{J}(U^\ell \beta_{\ell-1}) = \mathcal{J}(\beta_\ell) - \mathcal{J}(\beta_{\ell-1}) + \ell &= f(m,\ell+1) - f(m,\ell) + \ell = \ell - m \ge 0.\end{aligned}$$

Therefore the change of basis is filtered. □

We also record the filtration shift between the generators in the complex.

Definition 2.2 Suppose $U^c y$ is a nontrivial term in ∂x, where c is some constant and x, y are both generators. Define

$$\Delta_{\mathcal{I},\mathcal{J}}(x,y) = (\mathcal{I},\mathcal{J})(x) - (\mathcal{I},\mathcal{J})(U^c y)$$

and similarly define $\Delta_{\mathcal{I}}$ and $\Delta_{\mathcal{J}}$.

Clearly, $\Delta_{\mathcal{I},\mathcal{J}}(a_i^{(s)}, b_i^{(s)}) = (0,1)$ and $\Delta_{\mathcal{I},\mathcal{J}}(a_i^{(s)}, b_{i+1}^{(s)}) = (1,0)$ for each $i_{m,s}^{(b)} \le i \le i_{m,s}^{(t)}$. We also have:

Lemma 2.3 *For* $-2k+1 \le s \le 2k+m-1$ *and* $i^{(b)}_{m,s} \le i \le i^{(t)}_{m,s}+1$,

$$\Delta_{\mathcal{I},\mathcal{J}}(b_i^{(s)}, \beta_{s-1}) = (2k+1-i+s, i-1), \tag{12}$$

$$\Delta_{\mathcal{I},\mathcal{J}}(b_i^{(s)}, \beta_s) = (2k+1-i, m-s+i-1). \tag{13}$$

Proof By (7), (9), (10) and (11), we compute

$$\begin{aligned}
\Delta_{\mathcal{I},\mathcal{J}}(b_i^{(s)}, \beta_{s-1}) &= (\mathcal{I},\mathcal{J})(b_i^{(s)}) - (\mathcal{I},\mathcal{J})(U^{2k+1-i+s}\beta_{s-1}) \\
&= (0, f(m,s)-2k-s-2+2i) - (0, f(m,s)) + (2k+1-i+s, 2k+1-i+s) \\
&= (2k+1-i+s, i-1), \\
\Delta_{\mathcal{I},\mathcal{J}}(b_i^{(s)}, \beta_s) &= (\mathcal{I},\mathcal{J})(b_i^{(s)}) - (\mathcal{I},\mathcal{J})(U^{2k+1-i}\beta_s) \\
&= (0, f(m,s)-2k-s-2+2i) - (0, f(m,s+1)) + (2k+1-i, 2k+1-i) \\
&= (2k+1-i, m-s+i-1).
\end{aligned}$$

□

We have all the ingredients to calculate the concordance invariants $\varphi_{i,j}$. This will be done in two steps. First, we translate the complex $X_n^\infty(T_{2,4k+1})\langle m-1\rangle$ into the ring $\mathbb{F}[U,V]$. Then we further translate it into the ring $\mathbb{X}$ and perform a change of basis, resulting a standard complex, from which the invariants $\varphi_{i,j}$ can be readily read off.

Lemma 2.4 *Over the ring* $\mathbb{F}[U,V]$, *the complex* $X_n^\infty(T_{2,4k+1})\langle m-1\rangle$ *is generated by*

$$\{\beta_s \mid 0 \le s \le m-1\} \cup \{a_i^{(s)} \mid 1 \le s \le m-1,\ i^{(b)}_{m,s} \le i \le i^{(t)}_{m,s}\} \cup \{b_i^{(s)} \mid 1 \le s \le m-1,\ i^{(b)}_{m,s} \le i \le i^{(t)}_{m,s}+1\}$$

with differentials

$$\partial a_i^{(s)} = U b_{i+1}^{(s)} + V b_i^{(s)}, \tag{14}$$

$$\partial b_i^{(s)} = U^{\Delta_{\mathcal{I}}(b_i^{(s)},\beta_{s-1})} V^{\Delta_{\mathcal{J}}(b_i^{(s)},\beta_{s-1})} \beta_{s-1} + U^{\Delta_{\mathcal{I}}(b_i^{(s)},\beta_s)} V^{\Delta_{\mathcal{J}}(b_i^{(s)},\beta_s)} \beta_s. \tag{15}$$

Proof This follows trivially from (6), (7) and the definition of $\Delta_{\mathcal{I}}$ and $\Delta_{\mathcal{J}}$. □

We are ready to prove the following proposition. Compare with [21, Proposition 1.2].

Proposition 2.5 *We have*

$$\varphi_{i,j}(X_n^\infty(T_{2,4k+1})\langle m-1\rangle) = \begin{cases} -\sum_1^{s-1}(i^{(t)}_{m,s} - i^{(b)}_{m,s} + 1) & \text{if } (i,j) = (1,0), \\ 1 & \text{if } (i,j) \in \{\Delta_{k,m}(s) \mid 1 \le s \le m-1\}, \\ 0 & \text{otherwise,} \end{cases}$$

where $\Delta_{k,m}(s)$ *is given by*

$$\Delta_{k,m}(s) = \begin{cases} \left(k - \lceil \frac{s-1}{2} \rceil, m+k-\lfloor \frac{s+1}{2} \rfloor\right) & s \le 2k, \\ (0, 2k+m-s) & s \ge 2k. \end{cases}$$

Proof Continuing from Lemma 2.4, we can further translate the complex $X_n^\infty(T_{2,4k+1})\langle m-1\rangle$ into the ring

$$\mathbb{X}=\frac{\mathbb{F}[U_B,\{W_{B,i}\}_{i\in\mathbb{Z}},V_T,\{W_{T,i}\}_{i\in\mathbb{Z}}]}{(U_BV_T,U_BW_{B,i}-W_{B,i+1},V_TW_{T,i}-W_{T,i+1})}$$

using the maps

$$U\mapsto U_B+W_{T,0},\quad V\mapsto V_T+W_{B,0}.$$

The differentials now reads

$$\begin{aligned}\partial a_i^{(s)}&=(U_B+W_{T,0})b_{i+1}^{(s)}+(W_{B,0}+V_T)b_i^{(s)},\\ \partial b_i^{(s)}&=(U_B^{\Delta_{\mathcal{I}}(b_i^{(s)},\beta_{s-1})}W_{B,0}^{\Delta_{\mathcal{J}}(b_i^{(s)},\beta_{s-1})}+V_T^{\Delta_{\mathcal{J}}(b_i^{(s)},\beta_{s-1})}W_{B,0}^{\Delta_{\mathcal{I}}(b_i^{(s)},\beta_{s-1})})\beta_{s-1}\\ &\qquad+(U_B^{\Delta_{\mathcal{I}}(b_i^{(s)},\beta_s)}W_{B,0}^{\Delta_{\mathcal{J}}(b_i^{(s)},\beta_s)}+V_T^{\Delta_{\mathcal{J}}(b_i^{(s)},\beta_s)}W_{T,0}^{\Delta_{\mathcal{I}}(b_i^{(s)},\beta_s)})\beta_s.\end{aligned}$$

Note that each term in the previous coefficient in $\mathbb{F}[U,V]$ becomes two terms in the above expression, one in the ideal (U_B) and one in the ideal (V_T). We perform the change of basis

$$\tilde{b}_i^{(s)}=\begin{cases}b_i^{(s)}+U_B^{-1}W_{B,0}(b_{i-1}^{(s)}) & \text{if } i=i_{m,s}^{(t)}+1,\\ b_i^{(s)}+U_B^{-1}W_{B,0}(b_{i-1}^{(s)})+V_T^{-1}W_{T,0}(b_{i+1}^{(s)}) & \text{if } i_{m,s}^{(b)}<i<i_{m,s}^{(t)}+1,\\ b_i^{(s)}+V_T^{-1}W_{T,0}(b_{i+1}^{(s)}) & \text{if } i=i_{m,s}^{(b)},\end{cases}\tag{16}$$

$$\tilde{\beta}_s=\begin{cases}\beta_s+V_T^{-s}W_{T,0}^{m-s}\beta_{s-1} & \text{if } s=0,\\ \beta_s+V_T^{-s}W_{T,0}^{m-s}\beta_{s-1}+U_B^{s-m}W_{B,0}^{s}\beta_{s+1} & \text{if } 1\le s\le m-2,\\ \beta_s+U_B^{s-m}W_{B,0}^{s}\beta_{s+1} & \text{if } s=m-1,\end{cases}\tag{17}$$

which simplifies the differentials to

$$\partial a_i^{(s)}=U_B\tilde{b}_{i+1}^{(s)}+V_T\tilde{b}_i^{(s)},\tag{18}$$

$$\partial\tilde{b}_i^{(s)}=\begin{cases}V_T^{\Delta_{\mathcal{J}}(b_i^{(s)},\beta_s)}W_{T,0}^{\Delta_{\mathcal{I}}(b_i^{(s)},\beta_s)}\tilde{\beta}_s & \text{if } i=i_{m,s}^{(t)}+1,\\ U_B^{\Delta_{\mathcal{I}}(b_i^{(s)},\beta_{s-1})}W_{B,0}^{\Delta_{\mathcal{J}}(b_i^{(s)},\beta_{s-1})}\tilde{\beta}_{s-1} & \text{if } i=i_{m,s}^{(b)},\\ 0 & \text{otherwise.}\end{cases}\tag{19}$$

Note that $\Delta_{\mathcal{I},\mathcal{J}}(b_i^{(s)},\beta_{s-1})-\Delta_{\mathcal{I},\mathcal{J}}(b_i^{(s)},\beta_s)=(s,s-m)$ by (12) and (13) which justifies the change of basis (17). We have obtained a standard complex [5, Definition 5.1], where the sequence of vector length of each V-edges (starting from $\tilde{\beta}_0$) is

$$\underbrace{(\overbrace{-(1,0),\dots,-(1,0)}^{(i_{m,s}^{(t)}-i_{m,s}^{(b)}+1)\text{ copies}},\Delta_{\mathcal{I},\mathcal{J}}(b_{i_{m,s}^{(t)}+1}^{(s)},\beta_s)}_{1\le s\le m-1},\dots).$$

To finish the proof, it suffices to show $\Delta_{k,m}(s) = \Delta_{\mathcal{I},\mathcal{J}}(b^{(s)}_{i^{(t)}_{m,s}+1}, \beta_s)$ for $1 \le s \le m-1$. Recall that $i^{(t)}_{m,s} = \min\{k + \lceil \frac{s-1}{2} \rceil, 2k\}$. Therefore

$$i^{(t)}_{m,s} + 1 = \begin{cases} k + \lceil \frac{s+1}{2} \rceil & \text{if } s \le 2k, \\ 2k+1 & \text{if } s \ge 2k. \end{cases}$$

By (13), we compute

$$\begin{aligned} \Delta_{\mathcal{I},\mathcal{J}}(b^{(s)}_{i^{(t)}_{m,s}+1}, \beta_s) &= \left(2k+1-k-\left\lceil \frac{s+1}{2} \right\rceil, m-s+k+\left\lceil \frac{s+1}{2} \right\rceil - 1\right) \\ &= \left(k - \left\lceil \frac{s-1}{2} \right\rceil, m+k-\left\lfloor \frac{s+1}{2} \right\rfloor\right) \quad \text{if } s \le 2k; \\ \Delta_{\mathcal{I},\mathcal{J}}(b^{(s)}_{i^{(t)}_{m,s}+1}, \beta_s) &= (2k+1-(2k+1), m-s+2k+1-1) \\ &= (0, 2k+m-s) \quad \text{if } s \ge 2k. \end{aligned}$$

□

Lemma 2.6 *Suppose* $\mathrm{CFK}^\infty(S^3, J_{2k}) \cong \mathrm{CFK}^\infty(S^3, T_{4k+1}) \oplus A$ *for* $k > 0$*, where* $H_*(A) = 0$*. Let* $\mu^{(2k)}_{m,1}$ *denote the image of the* $(m,1)$*-cable of the meridian in the* -1*-surgery on* J_{2k}*. Then we have*

$$\varphi_{i,j}(S^3_{-1}(J_{2k}), \mu^{(2k)}_{m,1}) = \begin{cases} 1 & \text{if } (i,j) = (k, m+k-1), \\ 0 & \text{if } i > k \text{ or } j > m+k-1. \end{cases}$$

Proof Since $\mathrm{CFK}^\infty(S^3, J_{2k}) \cong \mathrm{CFK}^\infty(S^3, T_{4k+1}) \oplus A$, by the filtered mapping cone formula, we have $\mathrm{CFK}^\infty(S^3_{-1}(J_{2k}), \mu^{(2k)}_{m,1}) \cong \mathrm{CFK}^\infty(S^3_{-1}(T_{4k+1}), \mu'^{(2k)}_{m,1}) \oplus A'$. By Lemma 2.1

$$\varphi_{i,j}(S^3_{-1}(J_{2k}), \mu^{(2k)}_{m,1}) = \varphi_{i,j}(X^\infty_n(T_{2,4k+1})\langle m-1 \rangle).$$

According to Proposition 2.5, the value of $\varphi_{i,j}$ is determined by the sequence $\Delta_{k,m}(s)$ for $1 \le s \le m-1$. Note that $\Delta_{k,m}(1) = (k, m+k-1)$, and as s increases by 1 either the first or the second entry of $\Delta_{k,m}(s)$ decreases by 1. □

Proof of Theorem 1.2 Lemma 2.6 shows the $\varphi_{k,m+k-1}$ are linear independent, and the homomorphism

$$\bigoplus_{k>0,m>1} \varphi_{k,m+k-1} : \hat{\mathcal{C}}_{\mathbb{Z}}/\mathcal{C}_{\mathbb{Z}} \to \bigoplus_{k>0,m>1} \mathbb{Z}$$

is surjective. In particular, for each fixed $k > 0$ the homomorphism

$$\bigoplus_{m>1} \varphi_{k,m+k-1} : \hat{\mathcal{C}}_{\mathbb{Z}}/\mathcal{C}_{\mathbb{Z}} \to \bigoplus_{m>1} \mathbb{Z}$$

is surjective. □

3 Blowing down two-bridge links

We now turn to the algorithm due to Ozsváth, Szabó and Hales. Let $R = R_1 \cup R_2$ be a two-bridge link, where R_1 and R_2 are the two link components. One can view R as two arcs A_1 and A_2 on a fixed S^2 with two trivial over-arcs C_1 and C_2 each intersecting S^2 transversely at end points, such that $R_i = A_i \cup C_i$.

Since R_2 is unknotted, ϵ-framed surgery on R_2 with $\epsilon = \pm 1$ recovers S^3. Let Σ be the genus-one surface obtained from S^2 by attaching a one-handle along the arc C_2, and let μ_2 be the meridian of the one-handle. Ozsváth and Szabó give a description for a genus-one doubly pointed Heegaard diagram for the knot R_1 in $S^3_\epsilon(R_2) = S^3$ as follows.

Proposition 3.1 [16, Proposition 6.3] *For $\epsilon = \pm 1$, let $\alpha = C_2 \cup (\text{trivial arc in } S^2)$, $\beta = A_2 \cup C_2 \cup k\mu_2$, where k is chosen such that $\alpha \cdot \beta = \epsilon$, and put z and w basepoints at the two end points of C_1. Then $(\Sigma, \alpha, \beta, z, w)$ represents $(S^3_\epsilon(R_2), R_1)$.*

Proof Clearly (Σ, α, β) represents $S^3_\epsilon(R_2)$. Connecting z to w in the complement of α traces out C_1 while connecting w to z in the complement of β traces out A_1. □

In particular, it follows that $(S^3_\epsilon(R_2), R_1)$ is a $(1,1)$ knot. Next, view the two-bridge link R as the numerator closure of a rational tangle p/q, with p even and q odd. We follow the convention in [12]. Every rational tangle can be obtained from the trivial tangle by performing

- the vertical right-handed half-twist τ and its reverse τ^{-1};
- the horizontal left-handed half-twist σ and its reverse σ^{-1}.

Specifically in the case when the numerator closure is a link, τ^2 and σ (and their reverses) suffice to generate the rational tangle.[1] This can be seen from a simple lemma regarding continued fractions, as follows. For integers $a_1, \dots, a_\ell$, denote by $[a_1, \dots, a_\ell]$ the continued fraction

$$a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_{\ell-1} + \cfrac{1}{a_\ell}}}}.$$

Lemma 3.2 *If p is even and q is odd, we can arrange such that $p/q = [a_1, \dots, a_\ell]$ where ℓ is odd, and a_i is even when i is odd.*

Proof Set $(p_0, q_0) = (p, q)$, and for each $i \geq 1$, we will recursively choose integers (a_i, p_i, q_i) with p_i and q_i coprime, such that

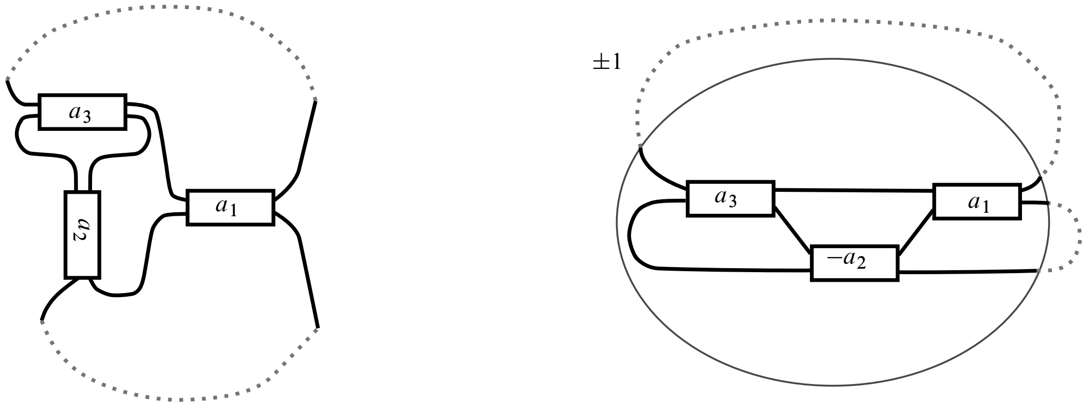
$$\frac{p_{i-1}}{q_{i-1}} = a_i + \frac{q_i}{p_i}. \tag{20}$$

Since p_{i-1}/q_{i-1} is between two consecutive integers, we choose a_i simply to be the closest even integer (resp. the closest integer) when i is odd (resp. when i is even) and choose p_i, q_i according to (20). It follows that $|p_i| = |q_{i-1}|$ and

$$p_{i-1} = a_i q_{i-1} + q_i \mod 2.$$

---

[1] In fact τ^2 and σ^2 and their reverses suffice, but at the expense of increasing the length of the continued fraction. For our purpose we will use τ^2 and σ.



(a) (b)

Figure 3: On the left, a rational tangle in the standard representation [12] corresponding to $p/q = [a_1, a_2, a_3]$, where positive numbers indicate crossings of the type ╳ and negative numbers vice versa. The diagram is alternating if all a_i are of the same sign. On the right, the same tangle in the 3-strand-braid representation [12]. The numbers in the boxes indicate the number of right-handed half-twists. The numerator closure of the tangle, indicated by the dotted lines, is a two-bridge link. For even a_1 and a_3, the knot $K^{\pm}[a_1, a_2, a_3]$ is obtained by doing ±1 surgery on one of the components.

From these one can inductively show that for all $0 \leq i \leq n$ we have

$$(p_i, q_i) = \begin{cases} (\text{even}, \text{odd}) & \text{if } i \text{ is even}, \\ (\text{odd}, \text{even}) & \text{if } i \text{ is odd}. \end{cases}$$

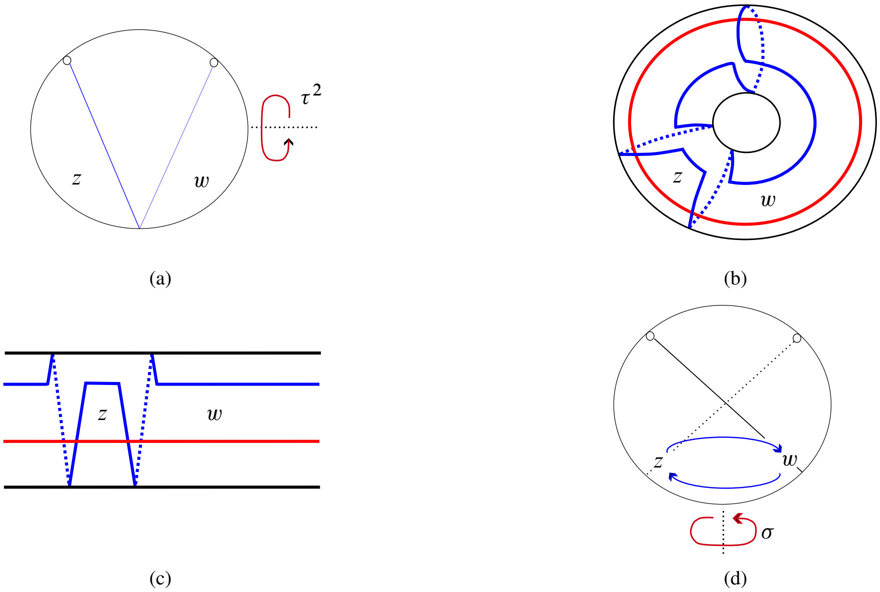
Note that $|q_i| < |p_i| = |q_{i-1}|$, and therefore this process will terminate in finite number of steps. Also the last term a_n must have n odd, since $q_n = 0$ which is an even number. □

Definition 3.3 Denote by $K^{\pm}([a_1, \dots, a_\ell])$ the knot obtained by doing ±1 surgery on the upper strand of the numerator closure of tangle $p/q = [a_1, \dots, a_\ell]$, respectively. One can also write $K^{\pm}(p/q)$. See Figure 3(b).

3.1 A diagrammatic algorithm

The following algorithm is due to Jonathan Hales [8]. Let us consider the effect of action τ^2 and σ on the Heegaard diagram $(\Sigma, \alpha, \beta, z, w)$ inside $\epsilon = \pm1$ surgery. We depict the result of these actions in Figure 4. Note that in Figure 4(b) and (c), in order to preserve the framing of β curve, for each added full twist we need to "double-back" one time such that the intersection points cancel in pairs with sign, and the framing of the β curve remains to be ϵ.

In order to compute the Heegaard Floer homology of a $(1,1)$ knot, the standard treatment is to lift the genus-one Heegaard diagram $(\Sigma, \alpha, \beta, z, w)$ to the universal cover $\mathbb{R}^2$, where the bigon counts are explicit. Therefore we would like to study the actions of τ^{2n} and σ on the universal cover of Σ. The following is interpreted from [8, Lemma 3.2.3]:



(a) (b) (c) (d)

Figure 4: (a) and (b) demonstrate the result of applying τ^2 on the four-punctured sphere and $\Sigma = S^3 \cup$ (one-handle) supposing $\epsilon = +1$, respectively. In (d) the two black arcs illustrate the effect of σ.

First note that all of the following transformations keep the lifts of α unchanged, and deform only the lifts of β.

For $n \in \mathbb{Z}$, τ^{2n} has the following effect: In the covering $S^1 \times \mathbb{R} = [0,1] \times \mathbb{R}/\{(0,r) \sim (1,r) \mid r \in \mathbb{R}\}$, where the lifts of α are identified with $S^1 \times \mathbb{Z}$, arrange such that the lifts of the z basepoint lie on a vertical line $\ell_z := \{1/4\} \times \mathbb{R}$ and the lifts of the w basepoint lie on a vertical line $\ell_w := \{3/4\} \times \mathbb{R}$. Fixing $\widetilde{\alpha}$, z and w basepoints, perform an ambient isotopy in a small neighborhood of ℓ_z given by the following: fix a small $\varepsilon > 0$,

$$H_t^{(z)}(x,r) = \begin{cases} \left(x, r + \left(1 - \frac{\left|\frac{1}{4} - x\right|}{\varepsilon}\right) nt\right) & \text{if } x \in \left[\frac{1}{4} - \varepsilon, \frac{1}{4} + \varepsilon\right], \\ (x,r) & \text{otherwise,} \end{cases} \tag{21}$$

for $0 \le t \le 1$. In the process, a small neighborhood of $\tilde{\beta}$ near each intersection point $\tilde{\beta} \cap \ell_z$ is shifted vertically by $|n|$ units and crosses $|n|$ of the z basepoints. Equivalently, τ^{2n} can also be interpreted as follows. Fixing $\widetilde{\alpha}$, z and w basepoints, perform an ambient isotopy in a small neighborhood of ℓ_w given by the following: fix a small $\varepsilon > 0$,

$$H_t^{(w)}(x,r) = \begin{cases} \left(x, r - \left(1 - \frac{\left|\frac{3}{4} - x\right|}{\varepsilon}\right) nt\right) & \text{if } x \in \left[\frac{3}{4} - \varepsilon, \frac{3}{4} + \varepsilon\right], \\ (x,r) & \text{otherwise,} \end{cases} \tag{22}$$

for $0 \le t \le 1$. In the process, a small neighborhood of $\tilde{\beta}$ near each intersection point $\tilde{\beta} \cap \ell_w$ is shifted vertically by $|n|$ units and crosses $|n|$ of the w basepoints. Further lift this to the universal cover $\mathbb{R}^2$. For an example, see Figure 6 in Example 4.2 and Figure 11.

On the other hand, note that σ is a local action. See Figure 4(d). In the universal cover, σ (resp. σ^{-1}) corresponds to performing a clockwise (resp. counterclockwise) half-Dehn twist around each lift of z and w basepoints. See the middle step of Figure 11. As a convention in this paper, after each action of σ we also switch the role of z and w basepoints, such that z is on the left and w is on the right. The switching induces a chain homotopy equivalence of the resulting Heegaard Floer complex. We include the switching such that the z basepoint is consistently on the left and w basepoint is consistently on the right in each lift.

According to Lemma 3.2, each rational tangle whose numerator closure is a two-component link can be obtained from the trivial tangle by applying τ^2 and σ iteratively. Therefore we have described an algorithm that produces the $(1,1)$ diagram of any knot that arises from blowing down a two-bridge link. Here by a $(1,1)$ diagram, we mean the universal cover of a genus-one doubly pointed Heegaard diagram, where we fix a preferred parametrization. We have outlined a proof for the following theorem:

Theorem 3.4 [8] *The actions τ^2 and σ provide an explicit description of the $(1,1)$ diagram of any knot that arises from blowing down a two-bridge link.*

4 Classifying $K^{\pm}(p/q)$ with continued fraction length ≤ 3

Using the algorithm due to Ozsváth, Szabó and Hales, the goal of this section is to give a complete classification of $\mathrm{CFK}^{\infty}(S^3, K^{\pm}(p/q))$ for $p/q = [a_1, \dots, a_\ell]$ where $\ell = 1$ or 3 and a_i is even for each odd i. Note that

$$K^{-}(p/q) = -K^{+}(-p/q) = -K^{+}([-a_1, \dots, -a_\ell]). \tag{23}$$

As a road map for this section, we will first consider the case $\ell = 1$ in the next subsection. Then, in Section 4.2, we will discuss some general facts about the $\ell = 3$ case, reducing it to five subcases. Finally we will prove each one of these subcases in Sections 4.3 through 4.7, completing the classification. All five subcases and their corresponding conclusions are recorded in Proposition 4.13 for the reader's convenience.

4.1 Case $K^{\pm}([a_1])$

The following proposition together with (23) classify all the $K^{\pm}([a_1])$.

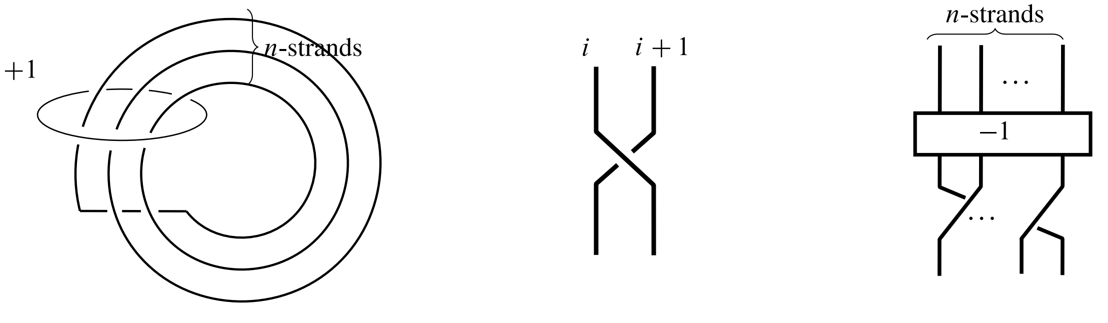
Proposition 4.1 *For $n > 0$,*

$$K^{+}([2n]) = -T_{n-1,n}, \tag{24}$$

$$K^{+}([-2n]) = -T_{n,n+1} \tag{25}$$

and $K^{+}([0])$ is the unknot.

Proof We will only prove (24). The other case is similar and left for the reader. Recall that the torus knot $-T_{n,n-1}$ is the braid closure of $(\omega_{n-1} \cdots \omega_1)^{n-1}$ where ω_i is the braid group element that exchanges the i-th and $(i+1)$-th strand, with the crossing convention given by Figure 5(b). Note that a left-handed full twist of the first k strands has a presentation of $(\omega_{k-1} \cdots \omega_1)^k$.



(a) the knot $K^+([2n])$ (b) the braid group element ω_i (c) a braid presentation of $K^+([2n])$

Figure 5: Illustrations for the proof of Proposition 4.1.

As depicted in Figure 5(c), the knot $K^+([2n])$ is the braid closure of $\omega_1\omega_2\cdots\omega_{n-1}(\omega_{n-2}\cdots\omega_1)^{n-1}$. Therefore it suffices to show that as elements of the braid group

$$(\omega_{n-1}\cdots\omega_1)^{n-1} = \omega_1\omega_2\cdots\omega_{n-1}(\omega_{n-2}\cdots\omega_1)^{n-1}. \tag{26}$$

From the braid relation

$$\omega_i\omega_{i+1}\omega_i = \omega_{i+1}\omega_i\omega_{i+1}, \tag{27}$$

it is straightforward to see that for $1 \le k < i \le k+j$,

$$\omega_i(\omega_k\omega_{k+1}\cdots\omega_{k+j}) = (\omega_k\omega_{k+1}\cdots\omega_{k+j})\omega_{i-1}. \tag{28}$$

We proceed by induction. For each $1 \le i \le n-1$, we claim that

$$(\omega_{n-1}\cdots\omega_1)^{n-1} = (\omega_{n-1}\cdots\omega_1)^{n-1-i}(\omega_{n-i}\cdots\omega_{n-1})(\omega_{n-2}\cdots\omega_1)^i. \tag{29}$$

This is obviously true for $i=1$. Suppose this is true for some $1 \le i \le n-2$. By using (28), we have

$$\begin{aligned}
&(\omega_{n-1}\cdots\omega_1)^{n-1}\\
&\quad= (\omega_{n-1}\cdots\omega_1)^{n-1-i}(\omega_{n-i}\cdots\omega_{n-1})(\omega_{n-2}\cdots\omega_1)^i\\
&\quad= (\omega_{n-1}\cdots\omega_1)^{n-2-i}\omega_{n-1}\cdots\omega_{n-i}(\omega_{n-i-1}\omega_{n-i}\cdots\omega_{n-1})\omega_{n-i-2}\cdots\omega_1(\omega_{n-2}\cdots\omega_1)^i\\
&\quad= (\omega_{n-1}\cdots\omega_1)^{n-2-i}(\omega_{n-i-1}\omega_{n-i}\cdots\omega_{n-1})(\omega_{n-2}\cdots\omega_{n-i-1})\omega_{n-i-2}\cdots\omega_1(\omega_{n-2}\cdots\omega_1)^i\\
&\quad= (\omega_{n-1}\cdots\omega_1)^{n-1-(i+1)}(\omega_{n-i-1}\cdots\omega_{n-1})(\omega_{n-2}\cdots\omega_1)^{i+1}.
\end{aligned}$$

Thus we have proved (29). Taking $i=n-1$ in (29) yields (26). □

Example 4.2 Consider $K^+([-2n])$ for $n>0$. Applying the algorithm by Ozsváth, Szabó and Hales, starting from a $(1,1)$ diagram for the unknot $K^+([0])$, where the β curve has slope -1, we equivariantly perform downwards finger moves of n units such that a small neighborhood of $\tilde{\beta}$ near each intersection point $\tilde{\beta} \cap \ell_z$ crosses n of the z basepoints. The resulting $(1,1)$ diagram is depicted in Figure 6. For a chosen lift $\tilde{\alpha}$ of α, we mark the intersection points of $\tilde{\alpha} \cap \tilde{\beta}$ from left to right by $x_1, x_2, \ldots$ in order.

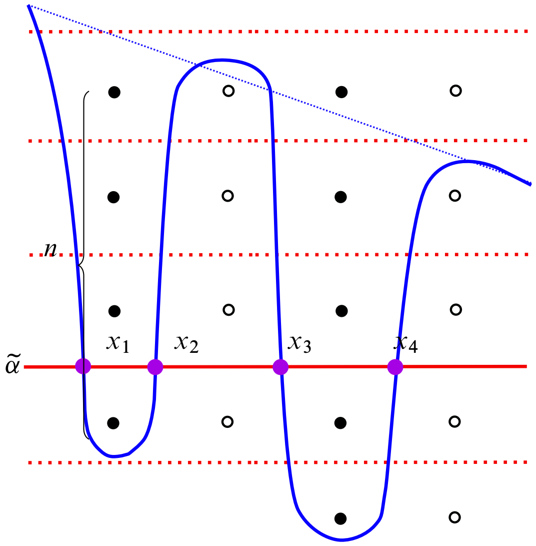


Figure 6: The $(1,1)$ diagram for $K^+([-2n])$ with $n > 0$, where the solid dots indicate (lifts of) the z basepoint and the hollow dots indicate (lifts of) the w basepoint.

There are $2n-1$ intersection points in total, and for each $i \in \{1, \dots, n-1\}$, there is a bigon from x_{2i-1} to x_{2i} with i copies of z and a bigon from x_{2i+1} to x_{2i} with $n-i$ copies of w; there are no other bigons. Therefore we conclude that $\mathrm{CFK}^\infty(S^3, K^+([-2n]))$ is generated by $x_1, \dots, x_{2n-1}$ with the differentials

$$\partial x_{2i-1} = x_{2i-2} + x_{2i}$$

and the filtration shifts

$$\Delta_{\mathcal{I},\mathcal{J}}(x_{2i-1}, x_{2i-2}) = (n-i+1, 0),$$
$$\Delta_{\mathcal{I},\mathcal{J}}(x_{2i-1}, x_{2i}) = (0, i)$$

for $i = 0, \dots, n$, where we take $x_0 = x_{2n} = 0$. For the definition of $\Delta_{\mathcal{I},\mathcal{J}}$, see Definition 2.2.

From Proposition 4.1 we already know $K^+([-2n]) = -T_{n,n+1}$. So this provides another way of computing the knot Floer complex of $T_{n,n+1}$ torus knots independent of the structure theorem of L-space knots [17] and the computation of their Alexander polynomials.

In the other direction, without knowing $K^+([-2n]) = -T_{n,n+1}$, just by looking at their $(1,1)$ diagrams in Figure 6, one can also show they are L-space knots via [7, Theorem 1.2] (the $(1,1)$ diagram is coherent) and this gives a shortcut for computing their Alexander polynomials.

4.2 General results on the case $\ell = 3$

From now on we fix the length of the continued fraction to be 3. Note that $K^{\pm}([2n_1, 0, 2n_2]) = K^{\pm}([2n_1 + 2n_2])$, and therefore the results here also apply to the knots in the previous section. We will end up reducing the classification to the case when $a_2 = \pm 1$ or 0, so for the following lemma we consider some relations between this subclass of knots.

Lemma 4.3 *For any integer* b, n, n_1 *and* n_2, *we have*

$$K^{\pm}([2n_1, b, 2n_2]) = K^{\pm}([2n_2, b, 2n_1]), \tag{30}$$

$$K^{\pm}([2n_1, 1, 2n_2]) = K^{\pm}([2(n_1+1), -1, 2(n_2+1)]), \tag{31}$$

$$K^{\pm}([0, \pm 1, 2n]) = K^{\pm}([2n]), \tag{32}$$

$$K^{\pm}([2, -1, 2n]) = K^{\pm}([2n-2]), \tag{33}$$

$$K^{\pm}([-2, 1, 2n]) = K^{\pm}([2n+2]), \tag{34}$$

$$K^{+}([2n_1, 1, 2n_2]) = -K^{-}([-2n_1, -1, -2n_2]). \tag{35}$$

Proof The relation (30) can be seen by rotating the paper plane by 180 degrees along a vertical axis. To show (31), we compute

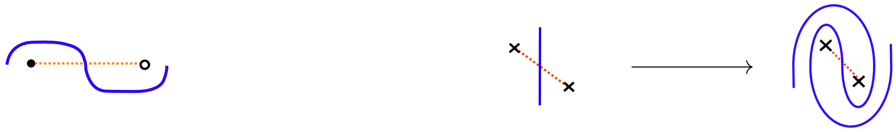
$$\begin{aligned}
[2(n_1+1), -1, 2(n_2+1)] &= 2(n_1+1) + \cfrac{1}{-1 + \cfrac{1}{2(n_2+1)}} \\
&= \frac{-4(n_1+1)(n_2+1) + 2(n_1+1) + 2(n_2+1)}{1 - 2(n_2+1)} = \frac{4n_1n_2 + 2n_1 + 2n_2}{2n_2+1} \\
&= [2n_1, 1, 2n_2].
\end{aligned}$$

The relation (32) is clear from the link diagram. Relations (33) and (34) are straightforward from the continued fraction. Relation (35) comes from mirroring. □

4.2.1 Full Dehn twists Let $\mathcal{H}^{\pm}([2n_1, b, 2n_2])$ denote the lifted Heegaard diagram in $S^1 \times \mathbb{R}$ obtained by applying the action $\tau^{2n_2}\sigma^b\tau^{2n_1}$ over the ∓ 1 sloped curve $\tilde{\beta}$. By default, we also fix a preferred lift $\tilde{\alpha}$ of α and $\tilde{\beta}$ of β. According to the algorithm discussed in Section 3.1, with the above data, $\mathcal{H}^{\pm}([2n_1, b, 2n_2])$ induces a basis B for the knot Floer complex $\mathrm{CFK}^{\infty}(S^3, K^{\pm}([2n_1, b, 2n_2]))$

When $b = \pm 1$ or 0, after performing τ^{2n_1}, connect each pair of the z and w basepoint with a horizontal line segment γ. After potentially pulling tight $\tilde{\beta}$, we see that γ intersects $\tilde{\beta}$ at most once. See Figure 7(a) for the case when $b = 1$; the other two cases are similar. The action of σ^2 is given by the local transformation depicted in Figure 7(b). Since this transformation is defined inside a small neighborhood of γ, we may reverse the order the operations, i.e., $\tau^{2n_2}\sigma_{\gamma}^2\tau^{2n_1} = \sigma_{\gamma'}^2\tau^{2n_2}\tau^{2n_1}$. Here we write σ_{γ} to stress the region where we perform the local transformation, and γ' is the image of γ under the ambient isotopy τ^{2n_2} given by (21) or (22). In other words, after performing τ^{2n_1}, instead of performing σ_{γ}^2 first, we can apply the ambient isotopy τ^{2n_2} to $\tilde{\beta}$ and γ at the same time, then apply $\sigma_{\gamma'}^2$ in a small neighborhood of γ'. This has the same effect as performing σ_{γ}^2 first, followed by τ^{2n_2}. Once the regions of the operations are understood, we drop them from the indices of σ.

Definition 4.4 For an arc γ, define ρ_+ to be the conformal transformation in a neighborhood of γ depicted in Figure 7(b). This is to be understood up to rotation in $\mathbb{R}^2$. We can similarly define ρ_-, given by a reflection of ρ_+ along a vertical axis. The action $\rho_{\pm}$ is trivial if $\tilde{\beta} \cap \gamma_{\pm} = \varnothing$.



(a) the horizontal segment γ

(b) the transformation ρ_+

(c) the result after performing ρ_+^m

(d) a counterclockwise finger move

Figure 7: In the above diagrams, the × symbol is used to indicate either basepoint.

Remark 4.5 A useful perspective is to view ρ_+ as the reverse of the counterclockwise finger move depicted in Figure 7(d), where we move both the basepoint and the $\tilde{\beta}$ curve. Equivalently one can also perform a counterclockwise finger move on the other basepoint to achieve the same result. Similar statements are true for ρ_-, with the only difference being that the finger move is clockwise.

We stress that $\rho_\pm$ preserves the angle between $\tilde{\beta}$ and γ. For example $\rho_- \circ \rho_+$ is not well defined. Nevertheless, this allows us to reduce any $\mathcal{H}^\pm([2n_1, b, 2n_2])$ to the case when $b = \pm 1$ or 0.

For the following, set $\epsilon = \operatorname{sgn}(b)$, $m = \lfloor \epsilon b/2 \rfloor$ and $b' = b - 2\epsilon m$. To understand $\mathcal{H}^\pm([2n_1, b, 2n_2])$, it suffices to understand $\mathcal{H}^\pm([2n_1, b', 2n_2])$ and the action ρ_ϵ^m over certain arcs (the image of all the γ under τ^{2n_2}).

Definition 4.6 When $b' = 1$ or 0, define γ_ϵ to be a straight line segment of slope $-2n_2$ from a z basepoint to a w basepoint in $\mathcal{H}^\pm([2n_1, 1, 2n_2])$.

When $b' = -1$, define γ_- to be a straight line segment of slope $-2(n_2 - 1)$ from a z basepoint to a w basepoint in $\mathcal{H}^\pm([2(n_1 - 1), 1, 2(n_2 - 1)])$.

We abuse the notation and let $\gamma_\pm$ also denote the line segments with the same slope in $\mathcal{H}^\pm([2n_1, b, 2n_2])$.

Lemma 4.7 *Given n_1, b and n_2, for m and b' as above,*

- *when $b \geq 0$, $\mathcal{H}^\pm([2n_1, b, 2n_2])$ is obtained from $\mathcal{H}(K^\pm([2n_1, b', 2n_2]))$ by performing the local transformation ρ_+^m over all γ_+;*
- *when $b \leq 0$, $\mathcal{H}^\pm([2n_1, b, 2n_2])$ is obtained from $\mathcal{H}(K^\pm([2n_1, b', 2n_2]))$ by performing the local transformation ρ_-^m over all γ_-.*

Proof The case when $b' = 0$ or 1 follows from the fact that $\tau^{2n_2}\sigma^b\tau^{2n_1} = \sigma^{2m}\tau^{2n_2}\sigma^{b'}\tau^{2n_1}$ and the image of γ under τ^{2n_2} is γ_+. Consider the case when $b' = -1$. We already know that

$$K^\pm([2n_1, -1, 2n_2]) = K^\pm([2(n_1 - 1), 1, 2(n_2 - 1)]).$$

In fact, one can check that after pulling tight $\tilde{\beta}$, $\mathcal{H}^{\pm}([2n_1, -1, 2n_2])$ and $\mathcal{H}^{\pm}([2(n_1-1), 1, 2(n_2-1)])$ represent the same diagram. This justifies the definition of γ_- in this case and the rest of the proof follows. □

The maps defined in Lemma 4.7 are equivariant; denote them by $\sigma^{\pm 2m}$, respectively. It turns out that they induce maps on the knot Floer chain complex together with certain *marked basis*, which we will now define.

We describe the process of assigning $+$ markings; the $-$ case is parallel.

Definition 4.8 For a given γ_+, after pulling tight $\tilde{\beta}$, let $a_+ = \gamma_+ \cap \tilde{\alpha}$ and $b_+ = \gamma_+ \cap \tilde{\beta}$. Both a_+ and b_+ are unique if they exist, as can be seen from the local picture Figure 7(a). From Figure 7(d) we can see that ρ_+ does not affect the knot Floer complex if $\tilde{\alpha} \cap \gamma_+ = \varnothing$. We give a $+$ marking to the "closest" intersection point p of $\tilde{\alpha} \cap \tilde{\beta}$ to a_+ or b_+. Concretely, p is the unique intersection point that forms a (α, β, γ) triangle with a_+ and b_+. We call p a $+$ marked point.

For the chain complex $C = \mathrm{CFK}^{\infty}(S^3, K^{\pm}([2n_1, b, 2n_2]))$ with the basis B induced by $\tilde{\alpha}$ and $\tilde{\beta}$, assign markings for every γ_+. This results in a $+$ marked basis B_+. Note that it is possible for a basis element to be assigned more than one $+$ marking.

Similarly define intersection points a_-, b_- as the $-$ marked points and B_- as a $-$ marked basis.

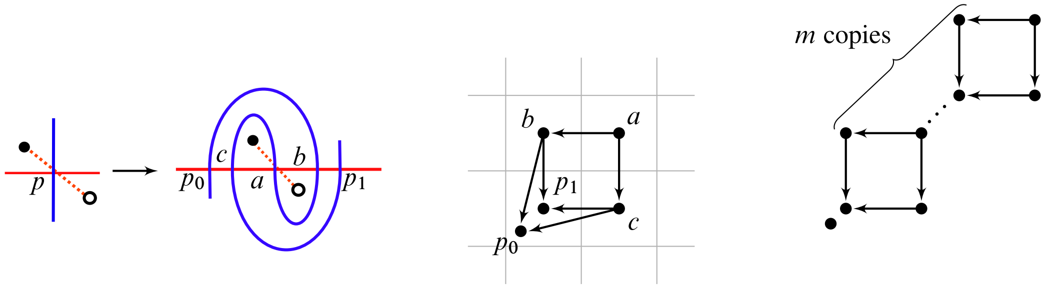
Let D_1 be the filtered chain complex over $\mathbb{F}[U, U^{-1}]$ generated by x_0, x_1, x_2 and y with differentials (each with length one) as below

$$\begin{tikzcd}
x_0 \arrow[d] & x_1 \arrow[l] \arrow[d] \\
y & x_2 \arrow[l]
\end{tikzcd}$$

Namely, D_1 is a length-one box summand.

Lemma 4.9 *Given a_1, a_3 even and $a_2 \geq 0$, we can obtain a model of $\big(\mathrm{CFK}^{\infty}(K^{\pm}([a_1, a_2+2, a_3])), B'_+\big)$ from $\big(\mathrm{CFK}^{\infty}(K^{\pm}([a_1, a_2, a_3])), B_+\big)$ as follows. For each $+$ marked point p, add ℓ copies of D_1 summands, where ℓ is the number of $+$ markings of p, such that for each D_1 summand, y and p share the same filtration level and Maslov grading. Remove all the previous markings and assign a $+$ marking to x_1 of each added D_1 summand. This gives the new marked basis B'_+.*

Proof Given a intersection point $b_+ = \gamma_+ \cap \tilde{\beta}$ for some γ_+, let p be the $+$ marked point induced by γ_+. In a small neighborhood of γ_+, the transformation ρ_+ is given by Figure 8(a), up to rotation in $\mathbb{R}^2$. In the resulting diagram, label the intersection points by p_0, c, a, b and p_1 from left to right (as we will see, this assignment also only matters up to the symmetry from the middle). They generate a complex depicted in Figure 8(b). In particular, we can perform a change of basis $\{a, b, c, p_0, p_1\} \mapsto \{a, b, c, p_0 + p_1, p_i\}$ where $i = 0$ or 1. Both choice of the change of basis splits off a summand generated by $\{a, b, c, p_0 + p_1\}$ isomorphic to D_1. We then identify either p_0 or p_1 with p. For $i = 0, 1$, clearly $p_0 + p_1$ share the same filtration level and Maslov grading with p_i; moreover observe that p_i inherits all the bigons of p (either incoming or outgoing, with the same filtration shifts). Adopting the perspective of Figure 7(d),



(a) (b) (c) the effect of the map σ^{2m}

Figure 8: The effect on CFK^∞.

performing a counterclockwise finger move will undo the transformation ρ_+ on the diagram level, and remove a box summand on the chain complex level. The choice of basepoint with which we perform the finger move corresponds to identifying p with p_0 or p_1.

Over the diagram $\mathcal{H}^\pm([a_1, a_2+2, a_3])$, following $\tilde{\beta}$ and record the sequence of b_+ in order. Perform the finger move in Figure 7(d) near each b_+ (this amounts to removing an interval of $\tilde{\beta}$ near b_+ and gluing it back). The resulting diagram is $\mathcal{H}^\pm([a_1, a_2, a_3])$. This proves the statement regarding the chain complex. For the new marked basis B'_+, simply notice that in the local picture Figure 8(a), the intersection point a forms a triangle with a_+ and b_+. □

Lemma 4.10 *Given a_1, a_3 even and $a_2 \le 0$, we can obtain a model of $\big(\mathrm{CFK}^\infty(K^\pm([a_1, a_2+2, a_3])), B'_-\big)$ from $\big(\mathrm{CFK}^\infty(K^\pm([a_1, a_2, a_3])), B_-\big)$ as follows. For each $+$ marked point p, add ℓ copies of D_1 summands, where ℓ is the number of $-$ markings of p, such that for each D_1 summand, x_1 and p share the same filtration level and Maslov grading. Remove all the previous markings and assign a $-$ marking to y of each added D_1 summand. This gives the new marked basis B'_-.*

Proof This is parallel to Lemma 4.9. □

Practically, we only need to consider the complex with marked basis $\big(\mathrm{CFK}^\infty(K^\pm([2n_1, b', 2n_2])), B_\pm\big)$ with $b' = \pm 1$ and 0 and consider the map $\sigma^{\pm 2m}$. In general it is not difficult to determine $B_\pm$ using Definitions 4.6 and 4.8. In particular, we have proved the following.

Proposition 4.11 *Up to chain homotopy equivalence,*

$$\mathrm{CFK}^\infty(S^3, K^\pm([2n_1, b, 2n_2])) \cong \mathrm{CFK}^\infty(S^3, K^\pm([2n_1, b', 2n_2])) \oplus \bigoplus_{i=1}^{N} D_1,$$

where $b' = b - 2\epsilon m$ for $\epsilon = \mathrm{sgn}(b)$, $m = \lfloor \epsilon b/2 \rfloor$, N is the number of markings in the marked basis B_ϵ and D_1 is the length-one box complex.

For any n_1 and n_2, by Definition 4.6, γ_+ and γ_- coincide for $\mathrm{CFK}^\infty(K^\pm(S^3, [2n_1, 0, 2n_2]))$, so $B_+ = B_-$. In this case it is easy to check that the map σ^2 in Lemma 4.9 and the map σ^{-2} in Lemma 4.10 induce the same map on the level of filtered chain homotopy type, even though they induce different basis in the image. In other words, we have the following.

Lemma 4.12 *Up to filtered chain homotopy equivalence, for* $n_1, n_2, b \in \mathbb{Z}$,

$$\mathrm{CFK}^\infty(K^\pm(S^3, [2n_1, 2b, 2n_2])) \cong \mathrm{CFK}^\infty(K^\pm(S^3, [2n_1, -2b, 2n_2])).$$

Note that the marked bases B_+ and B_- are generally different for $\mathrm{CFK}^\infty(K^\pm([2n_1, 1, 2n_2]))$ since γ_+ and γ_- have different slopes.

In view of Lemmas 4.3, 4.9, 4.10 and 4.12 we can reduce the $\ell = 3$ case to some subcases as follows. Suppose we are given $[a_1, a_2, a_3]$ with $a_1, a_3 \in 2\mathbb{Z}$ and $a_2 \in \mathbb{Z}$. If a_2 is even, then it suffice to consider $a_2 > 0$, and by mirroring we only need to consider K^+. If a_2 is odd, by mirroring we can guarantee that $a_1 > 0$. Next reducing a_2 to the case $a_2 = \pm 1$, we can further restrict to the case $a_2 = 1$ and consider both B_+ and B_- on $\mathrm{CFK}^\infty(S^3, K^\pm([a_1, 1, a_3]))$. Finally note that we can require $a_3 \neq -2$.

In summary, we have the following five cases, which are the topics of the next five subsections, respectively. We record the conclusion of each subsection here for the reader's convenience. This, together with the $\ell = 1$ case completes the proof of Theorem 1.5.

Proposition 4.13 *According to the discussion above, the case* $\ell = 3$ *is fully classified by the following cases: for* $n_1, n_2 \in \mathbb{Z}$ *and integer* $b > 0$,

(1) $K^+([2n_1, 2b, 2n_2])$, *Corollary 4.17*;

for $n_1, n_2 > 0$,

(2) $K^+([2n_1, 1, 2n_2])$, *Proposition 4.27*;

(3) $K^-([2n_1, 1, 2n_2])$, *Proposition 4.31*;

and for $n_1 > 0, n_2 > 1$,

(4) $K^+([2n_1, 1, -2n_2])$, *Proposition 4.32*;

(5) $K^-([2n_1, 1, -2n_2])$, *Proposition 4.33.*

4.3 Case $K^+([2n_1, 2b, 2n_2])$ with $b > 0$

Since $K^+([2n_1, 0, 2n_2]) = K^+([2n])$ for $n = n_1 + n_2$, with $K^+([2n])$ already classified in Section 4.1, the only extra data we require is a set of markings on $\mathrm{CFK}^\infty(S^3, K^+([2n]))$ (which as complexes of L-space knots, admits a unique basis).

4.3.1 Marked basis for $K^+([-2n])$ with $n > 0$ Fix $-n = n_1 + n_2$ and let $n_2 = -k$ for $k \in \mathbb{Z}$. We seek to decide the marked basis corresponding to $K^+([-2(n-k), 2b, -2k])$ for $b > 0$. By Definition 4.6 γ_+ is of slope k. Revisit Figure 6: in a $S^1 \times \mathbb{R}$ slice, denote the lifts of α that intersect a chosen lift $\tilde{\beta}$ of β by α_1 through α_n from bottom to top. Fix a lift $\tilde{\alpha}$ of α. It takes n iterations to cover all the intersections of $\tilde{\alpha} \cap \tilde{\beta}$, with $\tilde{\alpha}$ being identified with α_ℓ for $1 \leq \ell \leq n$ in each iteration. In total there are $2n-1$ generators in the complex $\mathrm{CFK}^\infty(S^3, K^+([-2n])) = \mathrm{CFK}^\infty(S^3, T_{n,n+1})$. Every iteration covers 2 intersection points except the last one, which covers 1 intersection point.

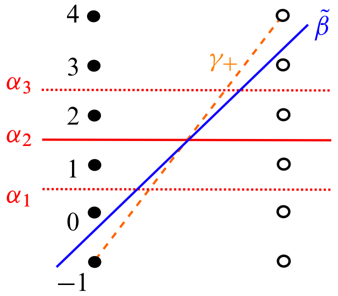


Figure 9: The γ_+ line segment with slope $n+1$ intersects the n sloped $\tilde{\beta}$ line segment once.

The only intersection points that are marked are the ones in the middle of each $S^1 \times \mathbb{R}$ slice. We focus on the portion of $\tilde{\beta}$ that travels between z and w basepoints, which is isotopic to a slope n line segment. The question of how many markings each intersection point receives is a completely combinatorial one: for each α_ℓ we need only count the number of line segments of slope k that intersect both line segments of slope n and slope 0.

In the below proposition, note that $\mathrm{CFK}^\infty(S^3, T_{n,n+1})$ admits a unique basis, and by the symmetry it does not matter from which end we start counting the generators.

Proposition 4.14 *For $n, b > 0$ and $k \in \mathbb{Z} \setminus \{0, n\}$, corresponding to $K^+([-2(n-k), 2b, -2k])$, the (2ℓ)-th generator in $\mathrm{CFK}^\infty(S^3, T_{n,n+1})$ receives $m(n,k,\ell)$ markings for $1 \le \ell \le n-1$, where*

$$m(n,k,\ell) = \begin{cases} k-n & \text{if } k \ge n+1, \\ n-\ell+\min\{\ell-k,0\}+\min\{k+\ell-n,0\} & \text{if } 1 \le k \le n-1, \\ -k & \text{if } k \le -1. \end{cases} \tag{36}$$

Proof For $1 \le \ell \le n-1$, in the ℓ-th iteration, α_ℓ is identified with $\tilde{\alpha}$ and the intersection point depicted in Figure 9 is the (2ℓ)-th generator. Label the height of the z and w basepoints by j, such that α_ℓ is between $j = \ell$ and $j = \ell - 1$.

When $k \ge n+1$, γ_+ intersects the line segment of slope n if and only if it starts from $j \le -1$ and ends at $j \ge n$. There are $k-1-n+1 = k-n$ of such line segments in total.

When $k \le -1$, γ_+ intersects the line segment of slope 0 if and only if it starts from $j \ge \ell$ and ends at $j \le \ell - 1$. There are $\ell - 1 - (\ell + k) + 1 = -k$ of such line segments in total.

When $1 \le k \le n-1$, γ_+ intersects both the line segments if only if it starts between $0 \le j \le \ell-1$ and ends between $\ell \le j \le n-1$. In other words, we need only find the length of the interval $[k, k+\ell-1] \cap [\ell, n-1]$, which is given by

$$\min\{k+\ell-1, n-1\} - \max\{\ell, k\} + 1 = n - \ell + \min\{\ell-k, 0\} + \min\{k+\ell-n, 0\}. \qquad \square$$

Note that the expression $m(n,k,\ell)$ is symmetric under the transformations

$$(n,k,\ell) \mapsto (n,k,n-\ell), \quad (n,k,\ell) \mapsto (n,n-k,\ell).$$

4.3.2 Marked basis for $K^+([2n])$ with $n > 0$ This is similar to the previous case. Fix $n = n_1 + n_2$ and let $n_2 = -k$ for $k \in \mathbb{Z}$. We seek to decide the marked basis corresponding to $K^+([2(n+k), 2b, -2k])$ for $b > 0$. Similarly label the lifts of α by α_1 through α_{n-1} from bottom to top. It takes $n-1$ iterations to cover all the intersections of $\tilde{\alpha} \cap \tilde{\beta}$, with $\tilde{\alpha}$ being identified with α_ℓ for $1 \le \ell \le n-1$ in each iteration. In total there are $2n-3$ generators in the complex $\mathrm{CFK}^\infty(S^3, K^+([2n])) = \mathrm{CFK}^\infty(S^3, T_{n,n-1})$. Every iteration covers 2 intersection points except the last one, which covers 1 intersection point.

The only intersection points are the ones in the middle of each $S^1 \times \mathbb{R}$ slice and the portion of $\tilde{\beta}$ that travels between z and w basepoints is of slope $-n$. By Definition 4.6 γ_+ is of slope k. The proof of the next proposition is similar to the previous case and left to the reader as an exercise.

Proposition 4.15 *For $n, b > 0$ and $k \in \mathbb{Z} \setminus \{0, -n\}$, corresponding to $K^+([2(n+k), 2b, -2k])$, the $(2\ell-1)$-th generator in $\mathrm{CFK}^\infty(S^3, T_{n,n-1})$ receives $m'(n,k,\ell)$ markings for $1 \le \ell \le n-1$, where*

$$m'(n,k,\ell) = \begin{cases} -k-n & \text{if } k \le -n-1, \\ n-\ell+\max\{\ell-n-k, 0\}+\max\{k+\ell, 0\} & \text{if } -n+1 \le k \le -1, \\ k & \text{if } k \ge 1. \end{cases} \tag{37}$$

4.3.3 Marked basis for $K^+([0])$ Again let $n_2 = -k$ for $k \in \mathbb{Z}$.

Definition 4.16 Let C_0 be the complex generated by one element.

We seek to decide the marked basis in C_0 corresponding to $K^+([2k, 2b, -2k])$ for $b > 0$. Similar analysis as from the previous sections shows that the unique generator in C_0 is assigned $|k|$ markings for $k \in \mathbb{Z}$. This concludes the discussion of the case $K^+([2n_1, 2b, 2n_2])$ with $b > 0$. In particular, Propositions 4.14–4.15 and the above imply the following.

Corollary 4.17 *For any $n_1, n_2 \in \mathbb{Z}_{\neq 0}$ and $b > 0$, up to chain homotopy equivalence, the complex $\mathrm{CFK}^\infty\big(S^3, K^+([2n_1, 2b, 2n_2])\big)$ is given by,*

- *if $n_1 + n_2 = 0$,*
$$C_0 \oplus b|n_2|D_1;$$
- *if $n_1 + n_2 > 0$, defining $n = n_1 + n_2$ and $k = -n_2$,*
$$\mathrm{CFK}^\infty(S^3, T_{n,n-1}) \oplus b\bigg(\sum_{\ell=1}^{n-1} m'(n,k,\ell)\bigg)D_1,$$
where $m'(n,k,\ell)$ is given by (37);
- *if $n_1 + n_2 < 0$, defining $n = -(n_1 + n_2)$ and $k = -n_2$,*
$$\mathrm{CFK}^\infty(S^3, T_{n,n+1}) \oplus b\bigg(\sum_{\ell=1}^{n-1} m(n,k,\ell)\bigg)D_1,$$
where $m(n,k,\ell)$ is given by (36).

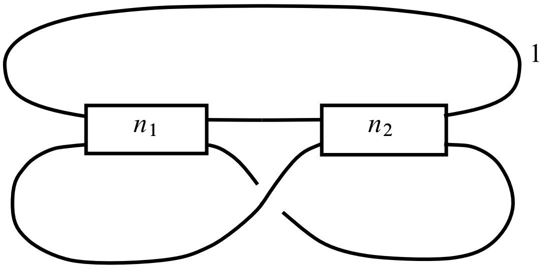


Figure 10: The knot $K^+_{n_1,n_2} = K^+([2n_1, 1, 2n_2])$. Numbers in the boxes indicate the number of right-handed full-twists.

4.4 Case $K^+([2n_1, 1, 2n_2])$ with $n_1, n_2 > 0$

From now on let us write $K^+_{n_1,n_2}$ for the knot $K^+([2n_1, 1, 2n_2])$ for the simplicity of the notation. See Figure 10 for a depiction of the knot $K^+_{n_1,n_2}$. Since the corresponding rational tangle has a presentation $[2n_1, 1, 2n_2]$, by Jonathan Hales's algorithm we shall consider the action $\tau^{2n_2}\sigma\tau^{2n_1}$ on the $(1, 1)$ diagram. The entire procedure is shown in Figure 11.

Fix a lift $\tilde{\alpha}$ (resp. $\tilde{\beta}$) of the α (resp. β) curve. Consider the final diagram in Figure 11 and suppose $\tilde{\beta}$ travels from left to right. Observe that in one iteration $\tilde{\beta}$ pass through $n_1 + n_2 - 1$ consecutive lifts of α (after isotoping away bigons without basepoint if necessary); denote them $\alpha_1, \dots, \alpha_{n_1+n_2-1}$ from bottom to top. To include all intersection points of $\tilde{\alpha} \cap \tilde{\beta}$, $n_1 + n_2 - 1$ iterations are needed. Following $\tilde{\beta}$, denote the diagram of each iteration by $H(s)$ for $1 \le s \le n_1 + n_2 - 1$; call it the s-th *block*. Note that in $H(s)$, $\tilde{\alpha}$ is identified with α_s.

To further determine the marked points, by Definition 4.6 each γ_+ is slope $-2n_2$, and viewing $\mathcal{H}^+([2n_1, 1, 2n_2])$ as $\mathcal{H}^+([2(n_1 + 1), -1, 2(n_2 + 1)])$, each γ_- is of slope $-2(n_2 + 1)$.

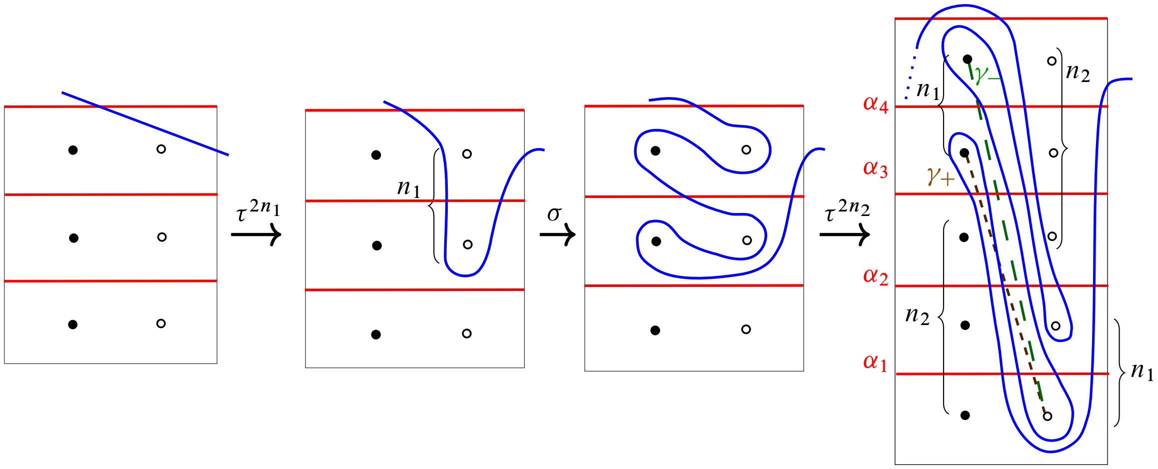
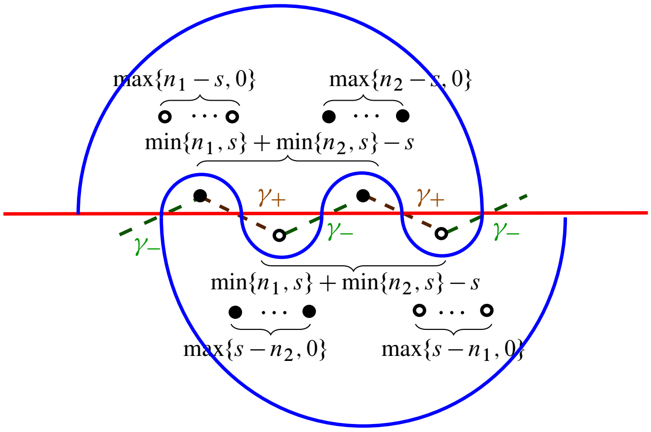


Figure 11: The action $\tau^{2n_2}\sigma\tau^{2n_1}$ when $n_1 = 2$ and $n_2 = 3$, starting from the slope -1. The solid dots indicate (lifts of) the z basepoint and the hollow dots indicate (lifts of) the w basepoint. The slope of γ_+ is $-n_2$ and the slope of γ_- is $-(n_2 + 1)$.

(a) the first block $H(1)$

(b) the last block $H(n_1+n_2-1)$



(c) the s-th block $H(s)$ with $2 \le s \le n_1+n_2-2$

Figure 12: Illustrations of blocks $H(s)$.

Lemma 4.18 *For* $1 \le s \le n_1+n_2-1$, *each block* $H(s)$ *corresponds to the diagram depicted in Figure 12.*

Proof This can be readily read off from the final diagram in Figure 11. □

Remark 4.19 To obtain Figure 12, we are allowed to isotope the curves in the universal cover in the complement of the basepoints and other curves, and to move the basepoints around as long as they do not cross the curves. Since both these operations preserve the bigons, Figure 12 can be used to compute the knot Floer complex. We need to pay some extra attention to keep track of γ_+ and γ_-.

We now abuse the notation and denote the complex generated by $H(s)$ in $\mathrm{CFK}^\infty(S^3, K^+_{n_1,n_2})$ also by $H(s)$, for $1 \le s \le n_1+n_2-1$. Define the chain complex $G(1) = \mathrm{CFK}^\infty(S^3, K^+_{n_1,n_2})$, and $G(i+1) = G(i)/(H(i) \cup H(n_1+n_2-i))$ for $1 \le i \le \lfloor \frac{n_1+n_2-1}{2} \rfloor$. Observe that $H(s)$ and $H(n_1+n_2-s)$ are both subcomplexes in $G(s)$.

We note that although suppressed from the notation, both $G(s)$ and $H(s)$ depend on s, n_1, n_2 and k.

Definition 4.20 Define a filtered chain complex D_s over $\mathbb{F}[U,U^{-1}]$ for $s \ge 1$ to be the complex generated by x_i with $0 \le i \le 2s$ and y, where differentials are given by

$$\begin{aligned} \partial x_{2i+1} &= x_{2i} + x_{2i+2}, && 0 \le i \le s-1, \\ \partial x_{2i} &= y, && 0 \le i \le s, \end{aligned}$$

with the filtration shifts

$$\begin{aligned}\Delta_{\mathcal{I},\mathcal{J}}(x_{2i+1},x_{2i})&=(1,0),\\ \Delta_{\mathcal{I},\mathcal{J}}(x_{2i+1},x_{2i+2})&=(0,1), \qquad 0\le i\le s-1,\\ \Delta_{\mathcal{I},\mathcal{J}}(x_{2i},y)&=(i,s-i), \quad 0\le i\le s.\end{aligned}$$

We say D_s is supported in Maslov grading a and filtration level (i,j) if y is supported in Maslov grading a and filtration level (i,j).

Alternatively, D_s is the complex C_s with the addition of a generator y and the differentials from x_{2i} to y for $0\le i\le s$ with above filtration shifts. The complex $D(3)$ is shown below as an example:

$$\begin{tikzcd}
x_0 \arrow[ddd] & x_1 \arrow[l] \arrow[d] & & \\
& x_2 \arrow[ddl] & x_3 \arrow[l] \arrow[d] & \\
& & x_4 \arrow[dll] & x_5 \arrow[l] \arrow[d] \\
y & & & x_6 \arrow[lll]
\end{tikzcd}$$

Lemma 4.21 *When $s\le\min\{n_1,n_2,\lfloor\frac{n_1+n_2-1}{2}\rfloor\}$, there exists a filtered change of basis T of $G(s)$, such that in the image of T, $H(s)$ becomes a summand, and moreover $H(s)\cong D_s$.*

Proof Consider two consecutive blocks $H(s)$ and $H(s+1)$ in $G(s)$ for $1\le s\le\min\{n_1,n_2,\lfloor\frac{n_1+n_2-1}{2}\rfloor\}$. See Figure 13. Denote the intersection points of $\tilde\beta\cap\tilde\alpha$ by $x_0,\dots,x_{2s},y,x'_0,\dots,x'_{2s'},y'$ in order, where $s'=\min\{n_1,s+1\}+\min\{n_2,s+1\}-s-1$, which is the number of small inner arcs in each half plane in $H(s+1)$. There are two cases to discuss: when $s=\min\{n_1,n_2\}$ or $s<\min\{n_1,n_2\}$.

- Suppose $s=\min\{n_1,n_2\}$. Then $s'=s$. Observe that the only arrows connected to $H(s)$ are given by

$$\partial x'_{2i}=y+y', \quad 0\le i\le s.$$

Performing the change of basis

$$x'_i\mapsto x'_i+x_i, \quad 0\le i\le 2s,$$

splits off $H(s)$ as a summand. Since the filtration shifts are

$$\Delta_{\mathcal{I},\mathcal{J}}(x_{2i},y)=(i,s-i), \tag{38}$$

$$\Delta_{\mathcal{I},\mathcal{J}}(x'_{2i},y)=(i,s-i)+(\max\{n_1-s-1,0\},\max\{n_2-s-1,0\}) \tag{39}$$

for $0\le i\le s$, this change of basis is filtered.

- Suppose $s<\min\{n_1,n_2\}$. Then $s'=s+1$. Similarly, the only arrows connected to $H(s)$ are given by

$$\partial x'_{2i}=y+y', \quad 0\le i\le s+1.$$

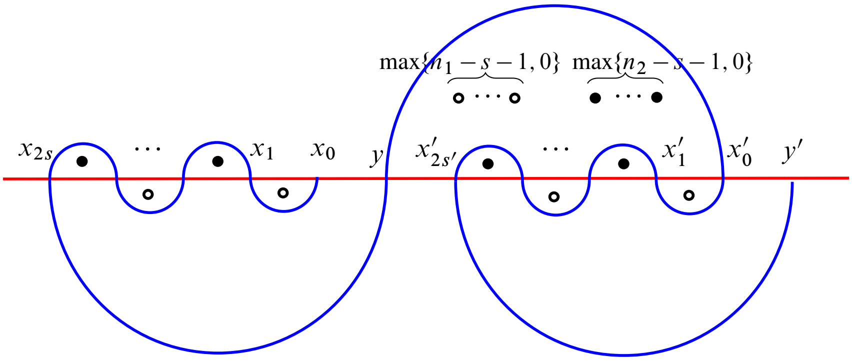


Figure 13: Two consecutive blocks $H(s)$ (on the left) and $H(s+1)$ (on the right) in $G(s)$ for $1 \le s \le \min\{n_1, n_2, \lfloor \frac{n_1+n_2-1}{2} \rfloor\}$. There could be other basepoints (at most more of each kind) in $H(s+1)$ in the lower half plane.

Performing the change of basis

$$x'_i \mapsto x'_i + x_i, \quad 0 \le i \le 2s,$$
$$x'_{2(s+1)} \mapsto x'_{2(s+1)} + x_{2s}$$

splits off $H(s)$ as a summand. Since the filtration shifts are

$$\Delta_{\mathcal{I},\mathcal{J}}(x_{2i}, y) = (i, s-i) \quad \text{for } 0 \le i \le s, \tag{40}$$
$$\Delta_{\mathcal{I},\mathcal{J}}(x'_{2i}, y) = (i, s+1-i) + (n_1-s-1, n_2-s-1) \quad \text{for } 0 \le i \le s+1, \tag{41}$$

this change of basis is filtered.

And it is clear from the diagram that each $H(s)$ generates a D_s summand (after quotienting out the top outmost arc). □

Lemma 4.22 *When $s \le \min\{n_1, n_2, \lfloor \frac{n_1+n_2-1}{2} \rfloor\}$, there exists a filtered change of basis T' of $G(s)$, such that in the image of T', $H(n_1+n_2-s)$ becomes a summand, and moreover $H(n_1+n_2-s) \cong D_s$.*

Proof This follows from Lemma 4.21 and the symmetry of the knot Floer complex. Equivalently one can run the similar argument as in the proof of Lemma 4.21 again. □

Lemma 4.23 *When $s \le \min\{n_1, n_2\} - 1$, there exists a filtered change of basis such that*

$$G(s) \cong G(s+1) \oplus H(s) \oplus H(n_1+n_2-s) \cong G(s+1) \oplus 2D_s.$$

Proof When $s \le \min\{n_1, n_2\} - 1$, we have $H(s) \cap H(n_1+n_2-1) = \varnothing$. By Lemmas 4.21 and 4.22, one simply applies the change of basis $T' \circ T$, and in the resulting complex $H(s)$ and $H(n_1+n_2-1)$ both become summands. □

These lemmas allow us to completely determine the complex $\mathrm{CFK}^\infty(S^3, K^+_{n_1,n_2})$. Due to (30), we may assume $n_1 \le n_2$. Therefore from now on we only consider knots of the form $K^+_{n,n+k}$ with $n > 0$ and $k \ge 0$.

Proposition 4.24 *For $n > 0$ and $k \geq 0$, we have*

$$\mathrm{CFK}^\infty(S^3, K^+_{n,n+k}) \cong G(n) \oplus 2\Bigl(\bigoplus_{s=1}^{n-1} D_s\Bigr).$$

Proof This follows from Lemma 4.23. Starting from $G(1) \cong \mathrm{CFK}^\infty(S^3, K^+_{n_1,n_2})$, we keep splitting off pairs of summands of D_s when $s \leq \min\{n_1, n_2\} - 1$. □

Therefore it suffices to determine the quotient complex $G(n) \cong \bigcup_{j=0}^k H(n+j)$ in $\mathrm{CFK}^\infty(S^3, K^+_{n,n+k})$. Note that the top outmost arc in $H(n)$ and the bottom outmost arc in $H(n+k)$ are quotiented out. We discuss several cases for different k.

- When $k = 0$, $G(n) \cong C_n$.
- When $k = 1$, we can use either Lemma 4.21 or 4.22 to split off a D_n summand. The remaining summand $G(n)/D_n$ is isomorphic to C_n.
- When $k = 2$, we can use Lemmas 4.21 and 4.22 to split off a pair of D_n summands. The remaining summand $G(n+1)$ is isomorphic to C_n.
- When $k > 3$, we again can use Lemmas 4.21 and 4.22 to split off a pair of D_n summands. The remaining summand is $G(n+1)$.

Definition 4.25 For $n > 0$, $k \geq 2$, define $C_{n,k}$ to be the complex generated by

$$\{x_i^{(j)} \mid 1 \leq j \leq k-1,\ 0 \leq i \leq 2n\} \cup \{y_j \mid 1 \leq i \leq k-2\}.$$

For simplicity, let $y_0 = y_{k-1} = 0$. For $1 \leq j \leq k-1$, the differentials are given by

$$\begin{aligned} \partial x_{2i+1}^{(j)} &= x_{2i}^{(j)} + x_{2i+2}^{(j)}, & 0 \leq i \leq n-1, \\ \partial x_{2i}^{(j)} &= y_j + y_{j-1}, & 0 \leq i \leq n, \end{aligned}$$

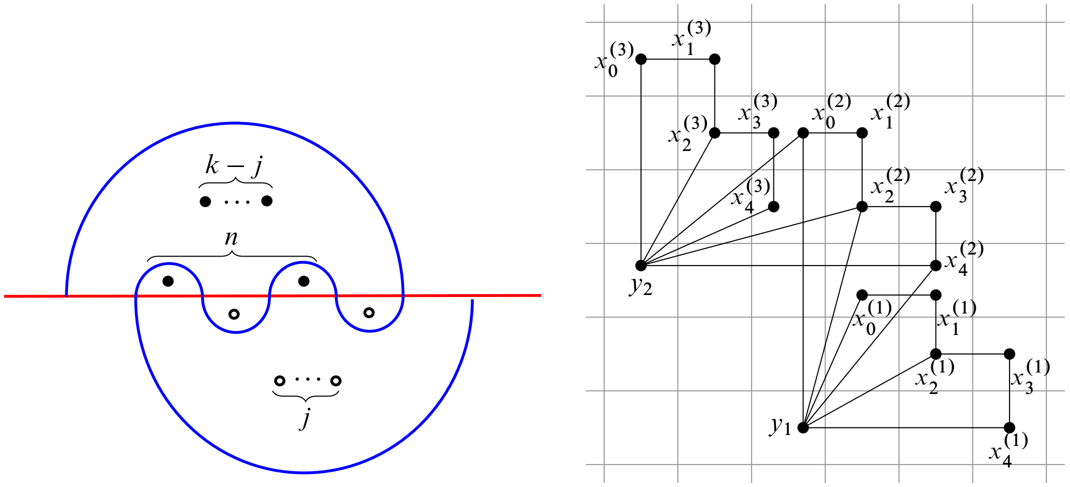
with the filtration shifts

$$\begin{aligned} \Delta_{\mathcal{I},\mathcal{J}}(x_{2i+1}^{(j)}, x_{2i}^{(j)}) &= (1,0), & \\ \Delta_{\mathcal{I},\mathcal{J}}(x_{2i+1}^{(j)}, x_{2i+2}^{(j)}) &= (0,1), & 0 \leq i \leq n-1, \\ \Delta_{\mathcal{I},\mathcal{J}}(x_{2i}^{(j)}, y_j) &= (i+j, n-i), & \\ \Delta_{\mathcal{I},\mathcal{J}}(x_{2i}^{(j)}, y_{j-1}) &= (i, k-j+n-i), & 0 \leq i \leq n. \end{aligned}$$

Note that $C_{n,2} = C_n$. See Figure 14(b) for an example when $n = 2$ and $k = 4$.

Lemma 4.26 *When $k \geq 2$, in $\mathrm{CFK}^\infty(S^3, K^+_{n,n+k})$ the quotient complex $G(n+1)$ is isomorphic to $C_{n,k}$.*

Proof The quotient complex $G(n+1) \cong \bigcup_{j=1}^{k-1} H(n+j)$ in $\mathrm{CFK}^\infty(S^3, K^+_{n,n+k})$. We have drawn the block $H(n+j)$ for $1 \leq j \leq k-1$ in Figure 14(a). Again the top outmost arc in $H(n+1)$ as well as the bottom outmost arc in $H(n+k-1)$ are quotiented out. Denoting the generators in $H(n+j)$ by $x_{2n}^{(j)}, \dots, x_0^{(j)}, y_j$ from left to right in order, note that for each j the $2n+1$ generators $x_{2n}^{(j)}, \dots, x_0^{(j)}$ form a



(a) (b)

Figure 14: On the left, the block $H(n+j)$ with $1 \le j \le k-1$. On the right, the complex $C_{2,4}$.

staircase, and for each $0 \le i \le n$ the generator $x_{2i}^{(j)}$ has a differential to y_j and y_{j-1} (taking $y_0 = y_{k-1} = 0$). The filtration shifts can be readily computed by counting the number of basepoints in the bigons. □

The following proposition describes the knot Floer complex $\mathrm{CFK}^\infty(S^3, K^+_{n,n+k})$ for $n > 0$, $k \ge 0$. The $k = 1$ case of Proposition 4.27 is the content of [8, Thoerem 3.3.14], which in turn implies Theorem 1.3. (To be precise, [8, Thoerem 3.3.14] describes the mirror of $K^+_{n,n+1}$.)

We recall that $K^+_{n_1,n_2} = K^+([2n_1, 1, 2n_2])$, $C_n \cong \mathrm{CFK}^\infty(S^3, T_{2,2n+1})$, the complexes D_s are defined in Definition 4.20 and the complexes $C_{n,k}$ in Definition 4.25.

Proposition 4.27 *For integers $n > 0$, $k \ge 0$, up to homotopy equivalence the knot Floer complex $\mathrm{CFK}^\infty(S^3, K^+_{n,n+k})$ is given by the following.*

- *When $k = 0$,*
$$\mathrm{CFK}^\infty(S^3, K^+_{n,n}) \cong C_n \oplus 2\Bigl(\bigoplus_{s=1}^{n-1} D_s\Bigr).$$
- *When $k = 1$,*
$$\mathrm{CFK}^\infty(S^3, K^+_{n,n+1}) \cong C_n \oplus D_n \oplus 2\Bigl(\bigoplus_{s=1}^{n-1} D_s\Bigr).$$
- *When $k = 2$,*
$$\mathrm{CFK}^\infty(S^3, K^+_{n,n+2}) \cong C_n \oplus 2\Bigl(\bigoplus_{s=1}^{n} D_s\Bigr).$$
- *When $k \ge 2$,*
$$\mathrm{CFK}^\infty(S^3, K^+_{n,n+k}) \cong C_{n,k} \oplus 2\Bigl(\bigoplus_{s=1}^{n} D_s\Bigr).$$

Moreover, there exists a choice of basis, such that when $k = 0$, *for each* $1 \leq s \leq n-1$ *both* D_s *are supported in filtration level* $(f_{n,s}, f_{n,s})$. *When* $k > 0$, *for* $1 \leq s \leq n$, *each pair of* D_s *is supported in filtration levels*

$$(f_{n,s}, f_{n,s}) + \left(\frac{(k-1)(k-2)}{2}, -(n-s+1)k+1\right),$$
$$(f_{n,s}, f_{n,s}) + \left(-(n-s+1)k+1, \frac{(k-1)(k-2)}{2}\right),$$

respectively, except for when $k = 1$ *the single copy of* D_n *is supported in* $(0,0)$, *where*

$$f_{n,s} = -\frac{(n-s)(n-s-1)}{2}. \tag{42}$$

Under this basis, each summand D_s *is supported in Maslov grading* -1; *each* 0*-graded generator has one* $-$ *marking and each* $+1$*-graded generator has one* $+$ *marking.*

Proof The statements regarding the homotopy equivalence type of $\mathrm{CFK}^\infty(S^3, K^+_{n,n+k})$ follow from Proposition 4.24, Lemma 4.26 and the discussion between. We are left with the statements regarding the Maslov grading, the marked basis and the filtration levels, which can be proved by examining the process of splitting off D_s summands more closely.

Since we have that for $s \leq n$,

$$\Delta_{\mathcal{I},\mathcal{J}}(x^{(s)}_{2i}, y_{s-1}) = (i, s+1-i) + (n-s, n+k-s), \quad \Delta_{\mathcal{I},\mathcal{J}}(x^{(s)}_{2i}, y_s) = (i, s+1-i),$$

the filtration shift from y_s to y_{s-1} (which supports D_s and D_{s-1}, respectively) is

$$-(n-s, n+k-s), \quad s \leq n. \tag{43}$$

For a fixed basis B of a complex C, we say B is supported in the filtration level (a,b) if the generator with the lowest i (resp. j) filtration in B is in filtration level a (resp. b).

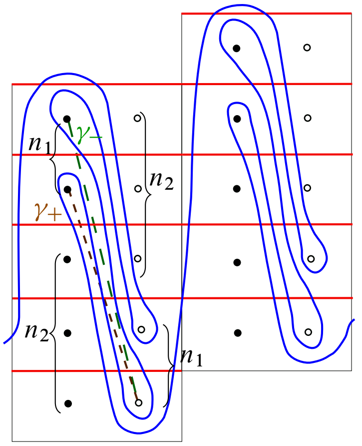
When $k = 0$, fix a basis such that the summand C_n is supported in filtration level $(0,0)$. By (39) both copies of D_{n-1} are supported in $(0,0)$. By (43), for $1 \leq s \leq n-1$ the copy of D_s that comes from $H(s)$ is supported in filtration

$$\left(-\sum_{\ell=0}^{n-s-1} \ell, -\sum_{\ell=0}^{n-s-1} \ell\right) = (f_{n,s}, f_{n,s}).$$

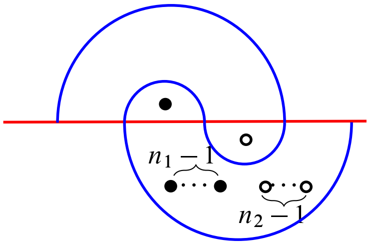
The filtration levels of the remaining complex follows from the symmetry of the knot Floer complex. When $k = 1$, fix a basis such that the single copy of D_n is supported in $(0,0)$. (The summand C_n is supported in $(0,0)$ as well.) The filtration levels in this case follows in the exact same way as above.

When $k \geq 2$, fix a basis such that the summand $C_{n,k}$ is supported in filtration level $(0,0)$. It is straightforward to check that $x^{(1)}_0$ is supported in the filtration level

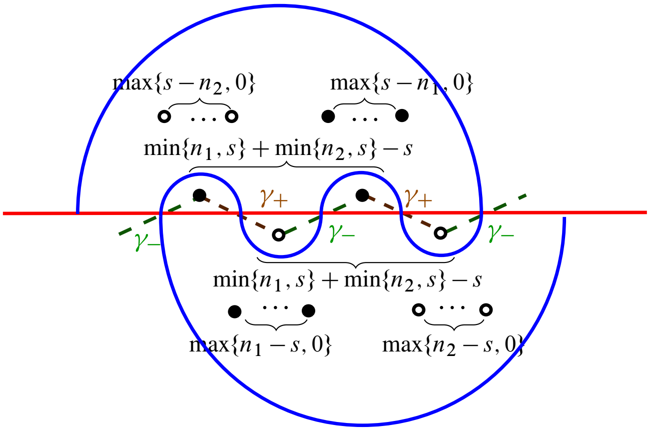
$$\left(\sum_{\ell=1}^{k-2} \ell, n\right) = \left(\frac{(k-1)(k-2)}{2}, n\right).$$



(a) the $(1,1)$ diagram of $K^-_{n_1,n_2}$



(b) the first block $H'(1)$



(c) the s-th block $H'(s)$ with $1 \le s \le n_1+n_2-1$

Figure 15: Illustrations of blocks $H'(s)$.

Then by (39) the copy of D_n that comes from $H(n)$ is supported in $\left(\frac{(k-1)(k-2)}{2}, -k+1\right)$. The filtration levels of the remaining complex follows in the same way as above.

For the statement regarding the Maslov grading, observe (for example, from Figure 13) that every y_s generator in some $H(s)$ is supported in the same Maslov grading t, and every $x_0^{(s)}$ generator in some $H(s)$ is supported in the same Maslov grading $t+1$. On the other hand, the homogeneous element $\sum_s x_0^{(s)}$ is a generator of $H_*(\mathrm{CFK}^\infty(S^3, K^+_{n,n+k})) \cong HF^\infty(S^3)$ and therefore is supported in Maslov grading 0. It follows that $t=-1$. The statement regarding the marked basis can be readily read off by Figure 12(c), using Definition 4.8. □

4.5 Case $K^-([2n_1, 1, 2n_2])$ with $n_1, n_2 > 0$

Similarly let us write $K^-_{n_1,n_2}$ for the knot $K^-([2n_1,1,2n_2])$. We still consider the action $\tau^{2n_2}\sigma\tau^{2n_1}$ on the $(1,1)$ diagram, with the difference being the starting slope is $+1$. The resulting $(1,1)$ diagram is depicted in Figure 15(a). Compare with the final diagram in Figure 11. The process for determining $\mathrm{CFK}^\infty(S^3, K^-_{n_1,n_2})$ is almost identical to the process described in the previous section, so we will only include the key steps, being much less elaborate.

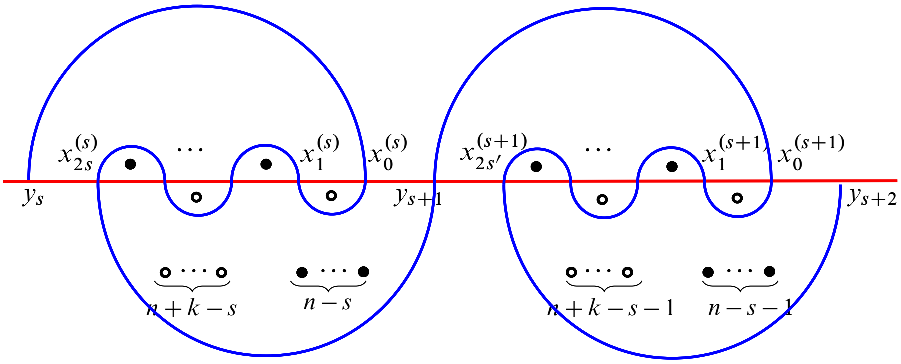


Figure 16: Two consecutive blocks $H'(s)$ (left) and $H'(s+1)$ (right) in $G'(s)$ for $1 \le s \le n$. There is either one basepoint or one of each kind in the upper half plane of $H'(n+1)$ when $s=n$.

To start, we can similarly define the s-th block $H'(s)$ for $1 \le s \le n_1+n_2-1$; we also define the chain complex $G'(1) = \mathrm{CFK}^\infty(S^3, K^-_{n_1,n_2})$, and $G'(i+1) = G'(i)/(\widetilde{H}'(i) \cup \widetilde{H}'(n_1+n_2-i))$, where $\widetilde{H}'(i) = H'(i) \setminus \{\text{right end point}\}$ and $\widetilde{H}'(n_1+n_2-i) = H'(n_1+n_2-i) \setminus \{\text{left end point}\}$ for $1 \le i \le \lfloor \frac{n_1+n_2-1}{2} \rfloor$. We add the apostrophe to differentiate from the complexes defined in the previous section.

Lemma 4.28 *For $1 \le s \le n_1+n_2-1$, each block $H'(s)$ corresponds to the diagram depicted in Figure 15(c).*

Again due to (30), we have $K^-_{n_1,n_2} = K^-_{n_2,n_1}$. Therefore from now on we only consider knots of the form $K^-_{n,n+k}$ with $n>0$, $k \ge 0$. Similarly as before, we split off D_s summands iteratively until we obtain a small complex. This process in the case of $K^-_{n,n+k}$ is somewhat more straightforward than that in the previous section. We depicted the two consecutive blocks in Figure 16, but in fact we need only to consider the first block $H'(s)$.

Lemma 4.29 *For $n>0$, $k \ge 0$,*

$$\mathrm{CFK}^\infty(S^3, K^-_{n,n+k}) \cong \begin{cases} G(n-1) \oplus 2\big(\bigoplus_{s=1}^{n-1} D_s\big) & \text{if } k=0, \\ G(n) \oplus 2\big(\bigoplus_{s=1}^{n} D_s\big) & \text{if } k \ge 1. \end{cases}$$

Proof For $1 \le s \le n$, label the generators in $H'(s)$ from left to right by $y_s, x_{2s}^{(s)}, \dots, x_0^{(s)}, y_{s+1}$ where $y_{s+1} = H'(s) \cap H'(s+1)$. The differentials are

$$\begin{aligned} \partial x_{2i+1}^{(s)} &= x_{2i}^{(s)} + x_{2i+2}^{(s)}, & 0 \le i \le s-1, \\ \partial x_{2i}^{(s)} &= y_s + y_{s+1}, & 0 \le i \le s, \end{aligned}$$

with filtration shifts

$$\begin{aligned} \Delta_{\mathcal{I},\mathcal{J}}(x_{2i}^{(s)}, y_s) &= (i, s-i), \\ \Delta_{\mathcal{I},\mathcal{J}}(x_{2i}^{(s)}, y_{s+1}) &= (i, s-i) + (n+k-s, n-s) \end{aligned}$$

for $0 \le i \le s$. Therefore performing the filtered change of basis

$$y_s \mapsto y_s + y_{s+1}$$

splits off $\widetilde{H}'(s) = H'(s) \setminus \{y_{s+1}\} \cong D_s$ as a summand. We can split off $\widetilde{H}'(s)$ and $\widetilde{H}'(2n+k-s)$ for $1 \le s \le n-1$ at the same time since $\widetilde{H}'(s) \cap \widetilde{H}'(2n+k-s) = \varnothing$. Due to the symmetry of the knot Floer complex, $\widetilde{H}'(2n+k-s) \cong D_s$. When $k \ge 1$, we can further split off $\widetilde{H}'(n) \cup \widetilde{H}'(n+k) \cong 2D_n$. □

Before stating the theorem, we first clarify some notation. Recall that C_0 is the complex generated by one element and the complexes D_s are defined in Definition 4.20.

Definition 4.30 For $n > 0$, $k \ge 2$, let $C'_{n,k}$ be the complex generated by

$$\{x_i^{(j)} \mid 1 \le j \le k-1,\ 0 \le i \le 2n\} \cup \{y_j \mid 1 \le j \le k\}$$

with differentials

$$\begin{aligned} \partial x_{2i+1}^{(j)} &= x_{2i}^{(j)} + x_{2i+2}^{(j)}, & 0 \le i \le n-1,\\ \partial x_{2i}^{(j)} &= y_j + y_{j+1}, & 0 \le i \le n, \end{aligned}$$

with the filtration shifts

$$\begin{aligned} \Delta_{\mathcal{I},\mathcal{J}}(x_{2i+1}^{(j)}, x_{2i}^{(j)}) &= (1,0), & \\ \Delta_{\mathcal{I},\mathcal{J}}(x_{2i+1}^{(j)}, x_{2i+2}^{(j)}) &= (0,1), & 0 \le i \le n-1,\\ \Delta_{\mathcal{I},\mathcal{J}}(x_{2i}^{(j)}, y_j) &= (i, n-i+j), & \\ \Delta_{\mathcal{I},\mathcal{J}}(x_{2i}^{(j)}, y_{j+1}) &= (i+k-j, n-i), & 0 \le i \le n. \end{aligned}$$

The complex $C'_{1,3}$ is shown below as an example:

$$\begin{tikzcd}
 & x_0^{(2)} \arrow[dl] \arrow[ddd, "3"] & x_1^{(2)} \arrow[l] \arrow[d] & & \\
y_3 & & x_2^{(2)} \arrow[ll] \arrow[ddl] & & \\
 & & & x_0^{(1)} \arrow[dll] \arrow[dd] & x_1^{(1)} \arrow[l] \arrow[d] \\
 & y_2 & & & x_2^{(1)} \arrow[lll, "3"] \arrow[dl] \\
 & & & y_1 &
\end{tikzcd}$$

Proposition 4.31 *For integers $n > 0$, $k \geq 0$, up to homotopy equivalence the knot Floer complex* $\mathrm{CFK}^\infty(S^3, K^-_{n,n+k})$ *is given by the following.*

- *When* $k = 0$,
$$\mathrm{CFK}^\infty(S^3, K^+_{n,n}) \cong C_0 \oplus D_n \oplus 2\Bigl(\bigoplus_{s=1}^{n-1} D_s\Bigr).$$
- *When* $k = 1$,
$$\mathrm{CFK}^\infty(S^3, K^+_{n,n+1}) \cong C_0 \oplus 2\Bigl(\bigoplus_{s=1}^{n} D_s\Bigr).$$
- *When* $k \geq 2$,
$$\mathrm{CFK}^\infty(S^3, K^+_{n,n+k}) \cong C'_{n,k} \oplus 2\Bigl(\bigoplus_{s=1}^{n} D_s\Bigr).$$

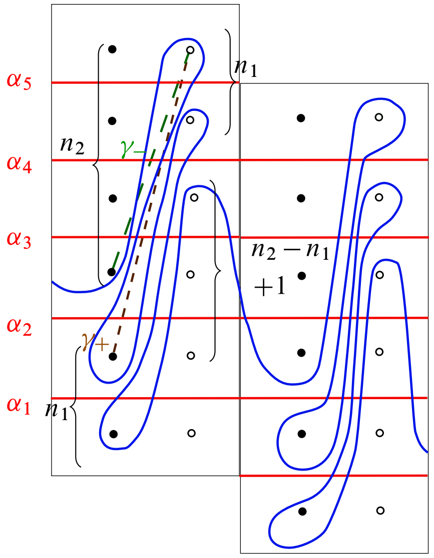
Moreover, there exists a choice of basis, such that C_0 *is always supported in the filtration level* $(0,0)$ *and each pair of* D_s *for* $1 \leq s \leq n$ *is supported in filtration levels*
$$\biggl(\Bigl(\frac{k+1}{2}+n-s\Bigr)k+\frac{(n-s)(n-s+1)}{2}, \frac{(n-s)(n-s+1)}{2}\biggr),$$
$$\biggl(\frac{(n-s)(n-s+1)}{2}, \frac{(n-s)(n-s+1)}{2}+\Bigl(\frac{k+1}{2}+n-s\Bigr)k\biggr),$$
respectively. Under this basis, each C_0 *and* D_s *summand are supported in Maslov grading* 0; *each* $+1$*-graded generator has one* $-$ *marking and each* $+2$*-graded generator has one* $+$ *marking.*

Proof From the proof of Lemma 4.29 we see that y_s and y_{s+1} have a filtration difference of
$$(n+k-s, n-s) \quad \text{for } 1 \leq s \leq n. \tag{44}$$
Each D_s is supported by y_s for $1 \leq s \leq n$. For the first two cases, fix a basis such that C_0 is supported in $(0,0)$. For the last case, fix a basis such that $C'_{n,k}$ is supported in $(0,0)$, and then it is straightforward to check that $y_1 \in C'_{n,k}$ has filtration level $((k-1)k/2, 0)$. By (44) we can determine the support of all the D_s that come from $\widetilde{H}'(s)$ for $1 \leq s \leq n$. The filtration levels of the remaining complex follows from the symmetry of the knot Floer complex.

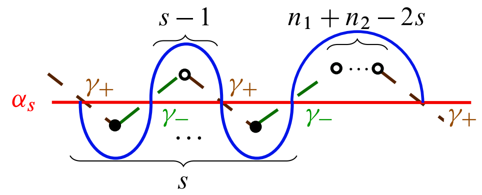
Each y_s generator has the same Maslov grading t, for some integer i, each $x^{(s)}_{2i}$ generator has the same Maslov grading $t+1$ and each $x^{(s)}_{2i+1}$ generator has the same Maslov grading $t+2$. Since the homology is supported in some y_s, we have $t = 0$. The statement regarding the marked basis can be read off from Figure 15(c). □

4.6 Case $K^+([2n_1, 1, -2n_2])$ with $n_1 > 0, n_2 > 1$

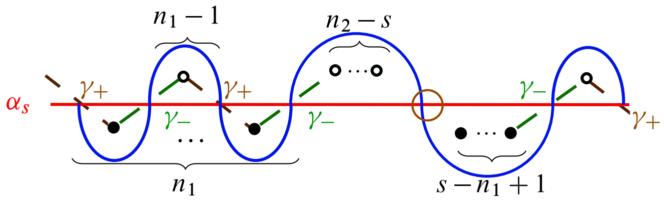
For suitable orientations of $\tilde{\alpha}$ and $\tilde{\beta}$ curve in the $(1,1)$ diagram, the induced orientation for each bigon is the same. See, for example, Figures 17 and 18. By [7, Theorem 1.2], $K^+([2n_1, 1, -2n_2])$ is a (negative)



(a) the (1, 1) diagram when $n_1 = 2, n_2 = 4$



(b) when $1 \le s \le n_1 - 1$



(c) when $n_1 \le s \le n_2 - 1$

Figure 17: The case $n_2 \ge n_1$. The only generators without a marking are given by the circled intersection point in (c).

(1, 1) L-space knot. Therefore in order to pin down its knot Floer complex, it suffices to record the length of (say) horizontal arrows.

4.6.1 When $n_2 \ge n_1$ The (1, 1) diagram in this case is depicted in Figure 17. We proceed as before. For a fixed lift $\tilde{\beta}$, label all the lifts of α which intersect $\tilde{\beta}$ by $\alpha_1, \dots, \alpha_{n_1+n_2-1}$. Identify $\tilde{\alpha}$ with α_s for $1 \le s \le n_1 + n_2 - 1$ in the s-th block. There is an ambiguity in defining the end point of each block. We define the end point of each block to be the first intersection point after $\tilde{\beta}$ travels above a w basepoint in the next block. See the right-hand side of Figure 17.

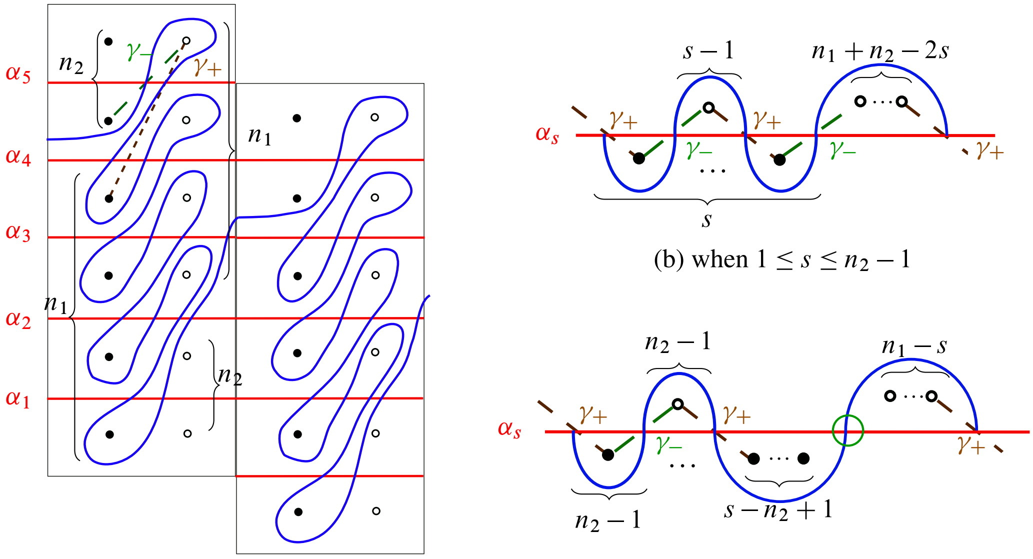
By Definition 4.6, the slope of γ_+ is n_2 and the slope of γ_- is $-(-n_2+1) = n_2 - 1$. Observe also that aside from the intersection point circled in Figure 17(c), every intersection point is marked exactly once, where the sign of the marking depends on the parity of the Maslov grading.

We simply count the number of basepoints in the bigons in the upper half plane. This is given by

$$\overbrace{1, \dots, 1}^{s-1}, n_1 + n_2 - 2s \quad \text{for } 1 \le s \le n_1 - 1 \qquad \text{and} \qquad \overbrace{1, \dots, 1}^{n_1-1}, n_2 - s, 1 \quad \text{for } n_1 \le s \le n_2 - 1.$$

When $n_1 \ne n_2$, we have $n_2 - 1 > (n_1 + n_2 - 1)/2$, therefore by the symmetry we can fill out the remaining complex. Specifically, the horizontal-vertical arrows of length $(n_2 - n_1, 1)$ in the (n_1)-th block is identified after reflection with the horizontal-vertical arrows of length $(1, n_2 - n_1)$ in the (n_2-1)-th block. It follows that the remaining horizontal arrows are all of length 1, and there are in total

$$\left(\sum_{i=1}^{n_1} i\right) - 1 = \frac{(n_1+2)(n_1-1)}{2}$$



(b) when $1 \le s \le n_2 - 1$

(c) when $n_2 \le s \le n_1 - 1$

(a) the $(1,1)$ diagram when $n_1 = 4$, $n_2 = 2$

Figure 18: The case $n_2 < n_1$. The only generators without a marking are given by the intersection point circled in (c).

of them. In summary, we have shown that the sequence of lengths of horizontal arrows in the knot Floer complex is given by

$$\underbrace{\overbrace{1,\dots,1}^{s-1}, n_1+n_2-2s,}_{\text{for } 1\le s\le n_1-1}\underbrace{\overbrace{1,\dots,1}^{n_1-1}, n_2-s, 1,}_{\text{for } n_1\le s\le n_2-1}\overbrace{1,\dots,1}^{\frac{(n_1+2)(n_1-1)}{2}}.$$

Moreover, the only generators without a marking are those whose outgoing horizontal arrows are of length $n_2 - s$ in the s-th block for $n_1 \le s \le n_2 - 1$. The above analysis does not cover the case when $n_1 = n_2$, but it is straightforward to check that the conclusion also holds there. (When $n_1 = n_2$, the (n_1)-th block consists of $n_1 - 1$ bigons in each half plane, where each bigon has exactly one basepoint, and the rest of the complex follows from symmetry.)

4.6.2 When $n_2 < n_1$ The $(1,1)$ diagram in this case is depicted in Figure 18. This case is parallel to the previous case, and one can similarly work out the sequence of length of horizontal arrows using the right-hand side diagrams in Figure 18. Putting together the discussion on both cases, we have proved the following.

Proposition 4.32 *For $n_1 > 0$, $n_2 > 1$, $K^+([2n_1, 1, -2n_2])$ is a negative L-space knot, with the length of horizontal arrows given in order by the follows.*

- *When $n_2 \ge n_1$,*

$$\underbrace{\overbrace{1,\dots,1}^{s-1}, n_1+n_2-2s,}_{\text{for } 1\le s\le n_1-1}\underbrace{\overbrace{1,\dots,1}^{n_1-1}, \overline{n_2-s}, 1,}_{\text{for } n_1\le s\le n_2-1}\overbrace{1,\dots,1}^{\frac{(n_1+2)(n_1-1)}{2}}.$$

Given a basis, each generator with odd Maslov grading has one − marking. Each generator whose outgoing horizontal arrow is marked by the overline has no marking. Apart from them, each generator with even Maslov grading has one + marking.

- *When* $n_2 < n_1$,

$$\underbrace{\overbrace{1,\dots,1}^{s-1},n_1+n_2-2s}_{\text{for }1\le s\le n_2-1},\underbrace{\overbrace{1,\dots,1}^{n_2-1},\overline{n_1-s}}_{\text{for }n_2\le s\le n_1-1},\overbrace{1,\dots,1}^{\frac{(n_2+2)(n_2-1)}{2}}.$$

Given a basis, each generator whose incoming horizontal arrow is marked by the overline has no marking. Apart from them, each generator with odd Maslov grading has one − marking. Each generator with even Maslov grading has one + marking.

4.7 Case $K^-([2n_1, 1, -2n_2])$ with $n_1 > 0, n_2 > 1$

By [7, Theorem 1.2], $K^-([2n_1, 1, -2n_2])$ is a positive $(1,1)$ L-space knot. It suffices to record (say) the length of vertical arrows. The process to determine such a sequence is entirely parallel to the process in Section 4.6. Therefore we will skip the proof, giving only the conclusion, as follows.

Proposition 4.33 *For* $n_1 > 0$, $n_2 > 1$, $K^-([2n_1, 1, -2n_2])$ *is a positive L-space knot, with the length of vertical arrows given in order by the follows.*

- *When* $n_2 \geq n_1 + 2$,

$$\underbrace{\overbrace{1,\dots,1}^{s-1},n_1+n_2-2s}_{\text{for }1\le s\le n_1},\underbrace{\overbrace{1,\dots,1}^{n_1},\overline{n_2-s-1}}_{\text{for }n_1+1\le s\le n_2-2},\overbrace{1,\dots,1}^{\frac{n_1(n_1+3)}{2}}.$$

Given a basis, each generator with even Maslov grading has one − marking. Each generator whose outgoing vertical arrow is marked by the overline has no marking. Apart from them, each generator with odd Maslov grading has one + marking.

- *When* $n_2 \leq n_1 + 1$,

$$\underbrace{\overbrace{1,\dots,1}^{s-1},n_1+n_2-2s}_{\text{for }1\le s\le n_2-2},\overbrace{1,\dots,1}^{n_2-2},\overline{n_1-n_2+1},\underbrace{\overbrace{1,\dots,1}^{n_2-1},\overline{n_1-s+1}}_{\text{for }n_2\le s\le n_1},\overbrace{1,\dots,1}^{\frac{n_2(n_2-1)}{2}}.$$

Given a basis, each generator whose incoming vertical arrow is marked by the overline has no marking. Apart from them, each generator with even Maslov grading has one − marking. Each generator with odd Maslov grading has one + marking.

Acknowledgements

I would like to thank my advisor Jen Hom for encouragement and support. I am grateful to Adam Levine, Fraser Binns and Tye Lidman for helpful conversations. I also want to express my appreciation for the amazing thesis [8] by Jonathan Hales. I thank an anonymous referee for helpful suggestions.

References

[1] **S Akbulut**, *A solution to a conjecture of Zeeman*, Topology 30:3 (1991) 513–515 MR

[2] **S Akbulut**, **Ç Karakurt**, *Heegaard Floer homology of some Mazur type manifolds*, Proc. Amer. Math. Soc. 142:11 (2014) 4001–4013 MR

[3] **S Akbulut**, **R Kirby**, *Mazur manifolds*, Michigan Math. J. 26:3 (1979) 259–284 MR

[4] **F Binns**, **H Zhou**, $(1,1)$ *almost L-space knots*, Indiana Univ. Math. J. 74:5 (2025) 1257–1291 MR

[5] **I Dai**, **J Hom**, **M Stoffregen**, **L Truong**, *More concordance homomorphisms from knot Floer homology*, Geom. Topol. 25:1 (2021) 275–338 MR

[6] **I Dai**, **J Hom**, **M Stoffregen**, **L Truong**, *Homology concordance and knot Floer homology*, Math. Ann. 390:4 (2024) 6111–6186 MR

[7] **J E Greene**, **S Lewallen**, **F Vafaee**, $(1,1)$ *L-space knots*, Compos. Math. 154:5 (2018) 918–933 MR

[8] **J Hales**, *Exotic four-manifolds, corks, and Heegaard Floer homology*, PhD thesis, State University of New York at Stony Brook (2013) MR Available at https://www.proquest.com/docview/1432009810

[9] **J Hanselman**, **L Watson**, *Cabling in terms of immersed curves*, Geom. Topol. 27:3 (2023) 925–952 MR

[10] **M Hedden**, **A S Levine**, *A surgery formula for knot Floer homology*, Quantum Topol. 15:2 (2024) 229–336 MR

[11] **J Hom**, **A S Levine**, **T Lidman**, *Knot concordance in homology cobordisms*, Duke Math. J. 171:15 (2022) 3089–3131 MR

[12] **L H Kauffman**, **S Lambropoulou**, *Classifying and applying rational knots and rational tangles*, from "Physical knots: knotting, linking, and folding geometric objects in $\mathbb{R}^3$" (Las Vegas, NV, 2001) (J A Calvo, K C Millett, E J Rawdon, editors), Contemp. Math. 304, Amer. Math. Soc., Providence, RI (2002) 223–259 MR

[13] **A S Levine**, *Nonsurjective satellite operators and piecewise-linear concordance*, Forum Math. Sigma 4 (2016) art. id. e34 MR

[14] **R Lipshitz**, **P S Ozsvath**, **D P Thurston**, *Bordered Heegaard Floer homology*, Mem. Amer. Math. Soc. 1216, Amer. Math. Soc., Providence, RI (2018) MR

[15] **P Ozsváth**, **Z Szabó**, *Heegaard Floer homology and alternating knots*, Geom. Topol. 7 (2003) 225–254 MR

[16] **P Ozsváth**, **Z Szabó**, *Holomorphic disks and knot invariants*, Adv. Math. 186:1 (2004) 58–116 MR

[17] **P Ozsváth**, **Z Szabó**, *On knot Floer homology and lens space surgeries*, Topology 44:6 (2005) 1281–1300 MR

[18] **I Petkova**, *Cables of thin knots and bordered Heegaard Floer homology*, Quantum Topol. 4:4 (2013) 377–409 MR

[19] **L Truong**, *A refinement of the Ozsváth–Szabó large integer surgery formula and knot concordance*, Proc. Amer. Math. Soc. 149:4 (2021) 1757–1771 MR

[20] **H Zhou**, *Homology concordance and an infinite rank free subgroup*, J. Topol. 14:4 (2021) 1369–1395 MR

[21] **H Zhou**, *A filtered mapping cone formula for cables of the knot meridian* (2022) arXiv 2208.11289

HUGO ZHOU hugozhou@umich.edu
Department of Mathematics, University of Michigan, Ann Arbor, MI, United States

Received: July 21, 2023 Revised: October 19, 2024

msp

Algebraic & Geometric Topology **26**:5 (2026) 1635–1653
DOI: 10.2140/agt.2026.26.1635
Published: May 29, 2026

On the mapping class groups of simply connected smooth 4-manifolds

DAVID BARAGLIA

The mapping class group $M(X)$ of a smooth manifold X is the group of smooth isotopy classes of orientation-preserving diffeomorphisms of X. We prove a number of results about the mapping class groups of compact, simply connected, smooth 4-manifolds. For example, we prove that $M(X)$ is nonfinitely generated for $X = 2n\mathbb{CP}^2 \# 10n\overline{\mathbb{CP}}^2$, where $n \geq 3$ is odd. Let $\Gamma(X)$ denote the group of automorphisms of the intersection lattice of X that can be realised by diffeomorphisms. Then $M(X)$ is an extension of $\Gamma(X)$ by $T(X)$, the Torelli group of isotopy classes of diffeomorphisms that act trivially in cohomology. We prove this extension is split for connected sums of $\mathbb{CP}^2$, but is not split for $2\mathbb{CP}^2 \# n\overline{\mathbb{CP}}^2$, where $n \geq 11$. We prove that the Nielsen realisation problem fails for certain finite subgroups of $M(p\mathbb{CP}^2 \# q\overline{\mathbb{CP}}^2)$ whenever $p + q \geq 4$. Lastly we study the extension $M_1(X) \to M(X)$, where $M_1(X)$ is the group of isotopy classes of diffeomorphisms of X which fix a neighbourhood of a point. When $X = K3$ or $K3 \# (S^2 \times S^2)$ we prove that $M_1(X) \to M(X)$ is a nontrivial extension of $M(X)$ by $\mathbb{Z}_2$. Moreover, we completely determine the extension class of $M_1(K3) \to M(K3)$.

1 Introduction

Let X be a compact, oriented, smooth, simply connected 4-manifold. Define the mapping class group $M(X)$ to be the group of smooth isotopy classes of orientation-preserving diffeomorphisms of X. There is considerable interest in the groups $M(X)$, although little is known about their structure. In this paper we will prove a number of new results concerning the structure of mapping class groups of smooth 4-manifolds.

Recall that the second cohomology group $L_X = H^2(X;\mathbb{Z})$ of X equipped with its intersection form is a unimodular lattice. We let $\mathrm{Aut}(L_X)$ denote the automorphism group of the lattice L_X. The group of orientation-preserving diffeomorphisms of X acts on L_X via $f \mapsto (f^{-1})^*$. This action depends only on this isotopy class and so defines a homomorphism $M(X) \to \mathrm{Aut}(L_X)$. Denoting the image of this map by $\Gamma(X)$ and the kernel by $T(X)$, we obtain a short exact sequence

$$1 \to T(X) \to M(X) \to \Gamma(X) \to 1. \tag{1-1}$$

We call $T(X)$ the *Torelli group* of X. It is the group of isotopy classes of diffeomorphisms of X that act trivially in cohomology. By a result of Quinn, $T(X)$ can also be defined as the group of isotopy classes of diffeomorphisms which are continuously isotopic to the identity [23]. The group $\Gamma(X)$ is the group of automorphisms of L_X that can be realised by diffeomorphisms of X.

Understanding the group $M(X)$ necessitates an understanding of the groups $T(X)$, $\Gamma(X)$ and the extension (1-1). The group $\Gamma(X)$ is known for some classes of 4-manifolds. In particular, a theorem

---

MSC2020: 57K41, 57R50.

of Wall implies that $\Gamma(X) = \mathrm{Aut}(L_X)$ for a large class of 4-manifolds [31]. In contrast, the Torelli group $T(X)$ is poorly understood. Ruberman [25] showed that $T(X)$ is not finitely generated for certain X. However this does not imply that $M(X)$ is not finitely generated, since a finitely generated group can have subgroups which are not finitely generated. Our first main result confirms that $M(X)$ is not finitely generated for certain simply connected 4-manifolds.

Theorem 1.1 *Let $X = 2n\mathbb{CP}^2 \# 10n\overline{\mathbb{CP}^2}$, where $n \geq 3$ is odd. Then $M(X)$ is not finitely generated. More precisely, the following holds:*

1. *There is an index-2 subgroup $M_+(X)$ of $M(X)$ and a surjective homomorphism $\Phi : M_+(X) \to \mathbb{Z}^\infty$ from $M_+(X)$ to $\mathbb{Z}^\infty$, where $\mathbb{Z}^\infty$ denotes a free abelian group of countably infinite rank.*
2. *The mod-2 reduction of Φ extends to a surjective homomorphism $\Phi : M(X) \to \mathbb{Z}_2^\infty$.*

As this paper was nearing completion we received a preprint by Hokuto Konno [15] which also proves that the mapping class groups of simply connected 4-manifolds can be nonfinitely generated. Konno's proof uses essentially the same method as ours, however we obtained our proofs completely independently.

Remark 1.2 It is interesting to contrast Theorem 1.1 with finiteness results for mapping class groups in other dimensions. Let X be a compact, simply connected smooth manifold of dimension d and $M(X) = \pi_0(\mathrm{Diff}(X))$ the mapping class group. If $d \neq 4$, then $M(X)$ is finitely generated. For $d \leq 3$, finite generation holds for any compact oriented manifold (see [7] for $d = 2$ and [13] for $d = 3$). If $d \geq 5$ then $M(X)$ is finitely generated [6, Theorem 2.6]. Note that Theorem 2.6 of [6] is only stated for $d \geq 6$, but when X is simply connected, the proof carries over to $d = 5$. In the proof of [6, Theorem 2.6], dimension 6 only enters in the point (i) in the proof, but Cerf's theorem says that in the simply connected case $\pi_0(C^{\mathrm{Diff}}(X)) = 0$, and in [6, Proposition 2.7] where it is not necessary. This follows from specialising Triantafillou [30] to simply connected manifolds, where none of the oversights mentioned in [6, §2.2] cause a problem.[1]

One may also consider the larger group $M'(X)$ consisting of isotopy classes of diffeomorphisms which are not necessarily orientation-preserving. Since $M(X)$ has finite index in $M'(X)$, it follows from Schreier's lemma [28] that if $M(X)$ is not finitely generated then neither is $M'(X)$.

Remark 1.3 Let X be a compact, simply connected smooth 4-manifold and $M^{\mathrm{top}}(X) = \pi_0(\mathrm{Homeo}(X))$ be the topological mapping class group. By work of Freedman [10] and Quinn [23], the natural map $M^{\mathrm{top}}(X) \to \mathrm{Aut}(H^2(X;\mathbb{Z}))$ to the group of automorphisms of the intersection lattice $H^2(X;\mathbb{Z})$ is an isomorphism. By a result of Siegel [29], the automorphism group of any lattice is finitely generated. Hence $M^{\mathrm{top}}(X)$ is finitely generated.

In contrast, we do not know whether the group $\Gamma(X) \subseteq \mathrm{Aut}(H^2(X;\mathbb{Z}))$ is always finitely generated, although we conjecture that it is.

Our next result concerns the question of whether or not the sequence (1-1) admits a splitting.

---

[1] I thank Alexander Kupers for explaining why [6, Theorem 2.6] works for simply connected 5-manifolds.

Theorem 1.4 (1) *Let* $X = n\mathbb{CP}^2$, *where* $n \geq 1$. *Then there exists a splitting* $\Gamma(X) \to M(X)$.

(2) *Let* $X = (S^2 \times S^2) \# X'$, *where* $b_+(X') = 1$, $b_-(X') \geq 10$. *Then there does not exist a splitting* $\Gamma(X) \to M(X)$.

More precise information about the failure of a splitting in case (2) is provided by Theorem 5.1.

Remark 1.5 In [5] it is shown when X is a $K3$ surface, there is a splitting $\Gamma(X) \to M(X)$. It is also easy to see that splittings exist for $S^2 \times S^2$ and $\mathbb{CP}^2 \# \overline{\mathbb{CP}^2}$.

Our next result concerns the Nielsen realisation problem. Recall that the Nielsen realisation problem for a smooth manifold X asks whether a subgroup G of the mapping class group of X can be lifted to a subgroup of $\mathrm{Diff}(X)$. Recent results of Baraglia and Konno [5], Farb and Looijenga [9], and Konno [14] show that Nielsen realisation fails for many simply connected spin 4-manifolds. Arabadji and Baykur showed that there are many nonspin 4-manifolds with finite nontrivial fundamental group for which Nielsen realisation fails [1] and Konno, Miyazawa, and Taniguchi gave examples with simply connected indefinite nonspin 4-manifolds [17].

Theorem 1.6 *Let* $X = X' \# p\mathbb{CP}^2 \# q\overline{\mathbb{CP}^2}$ *where* X' *is a compact*, *smooth*, *simply connected* 4*-manifold and* $p + q \geq 4$. *Then* $M(X)$ *contains a subgroup isomorphic to* $\mathbb{Z}_2^4$ *which cannot be lifted to* $\mathrm{Diff}(X)$.

In particular, Nielsen realisation fails for $n\mathbb{CP}^2$ for $n \geq 4$. As far as we are aware, these are the first examples of definite, simply connected 4-manifolds where Nielsen realisation fails.

Our last main result concerns a certain extension of $M(X)$. Let $X^{(1)}$ be obtained from X by removing an open ball and let $\mathrm{Diff}(X^{(1)}, \partial X^{(1)})$ denote the group of diffeomorphisms of $X^{(1)}$ which are the identity in a neighbourhood of the boundary. Let $M_1(X) = \pi_0(\mathrm{Diff}(X^{(1)}, \partial X^{(1)}))$ denote the group of components of $\mathrm{Diff}(X^{(1)}, \partial X^{(1)})$. It is known that the map $M_1(X) \to M(X)$ is surjective and that the kernel (which is either trivial or has order 2) is generated by a Dehn twist on the boundary (see Section 7 for more details).

In general it is difficult to determine whether the kernel of $M_1(X) \to M(X)$ is trivial or nontrivial, or equivalently, whether the boundary Dehn twist is trivial or nontrivial. The extension is known to be trivial when X is a connected sum of copies of $S^2 \times S^2$. In contrast we have:

Theorem 1.7 *Let* X' *be a compact*, *smooth*, *simply connected* 4*-manifold which is homeomorphic to* $K3$. *Let* X *be* X' *or* $X' \# (S^2 \times S^2)$. *Then the boundary Dehn twist is nontrivial. Moreover*, *the extension* $1 \to \mathbb{Z}_2 \to M_1(X) \to M(X) \to 1$ *does not split.*

If $M_1(X) \to M(X)$ is a nontrivial extension, then it is given by an extension class $\xi_X \in H^2(M(X); \mathbb{Z}_2)$ and the above theorem says that $\xi_X \neq 0$ when X is of the stated form. Our final result completely determines ξ_X in the case that X is homeomorphic to $K3$. Let L_X be the intersection lattice of X and $\mathrm{Aut}(L_X)$ the group of automorphisms. Over the classifying space $B\mathrm{Aut}(L_X)$ we have the tautological flat bundle $H = E\mathrm{Aut}(L_X) \times_{\mathrm{Aut}(L_X)} L_X$. Let $H^+ \to B\mathrm{Aut}(L_X)$ be a maximal positive subbundle. This defines a characteristic class $w_2(H^+) \in H^2(\mathrm{Aut}(L_X); \mathbb{Z}_2)$.

Theorem 1.8 *Let X be a smooth 4-manifold which is homeomorphic to $K3$. Then the extension class $\xi_X \in H^2(M(X);\mathbb{Z}_2)$ is the pullback of $w_2(H^+) \in H^2(\mathrm{Aut}(L_X);\mathbb{Z}_2)$ under the map $M(X) \to \mathrm{Aut}(L_X)$.*

1.1 Structure of the paper

The structure of the paper is as follows. In Section 2 we review the Seiberg–Witten invariants for the Torelli group (as in [3; 24; 26]) and show how these invariants can be assembled into cohomology classes on the mapping class group. In Section 3 we use these cohomology classes to show that $M(X)$ is not finitely generated for certain X. In Section 4 we construct a splitting $\Gamma(X) \to M(X)$ when X is a connected sum of copies of $\mathbb{CP}^2$. In Section 5 we prove the nonexistence of splittings $\Gamma(X) \to M(X)$ for certain 4-manifolds. The proof uses families Seiberg–Witten theory and more specifically the main result of [2]. In Section 6 we prove Theorem 1.6. Finally, in Section 7 we study boundary Dehn twists and the extension $M_1(X) \to M(X)$ and we prove Theorems 1.7 and 1.8.

2 Seiberg–Witten invariants for the mapping class group

In this section we define Seiberg–Witten invariants for the mapping class group, extending the Seiberg–Witten invariants on the Torelli group which have previously been considered in [3; 24; 26]. These invariants will be used to show that certain simply connected 4-manifolds have nonfinitely generated mapping class group.

Let X be a compact, smooth, simply connected 4-manifold and let $\mathfrak{s}$ be a spin^c-structure with $d(\mathfrak{s}) = -1$, where

$$d(\mathfrak{s}) = \tfrac{1}{4}(c(\mathfrak{s})^2 - \sigma(X)) - b_+(X) - 1$$

is the expected dimension of the Seiberg–Witten moduli space for $\mathfrak{s}$. Let $\mathcal{S}(X)$ denote the set of all isomorphism classes of spin^c-structures on X for which $d(\mathfrak{s}) = -1$. Since X is assumed to be simply connected, $\mathcal{S}(X)$ can be identified with the set of characteristic elements $c \in L = H^2(X;\mathbb{Z})$ for which $(c^2 - \sigma(X))/4 - b_+(X) = 0$.

Let Π denote the space of pairs (g,η) where g is a Riemannian metric on X and η is a 2-form which is self-dual with respect to g. For any $h \in \Pi$ and any $\mathfrak{s} \in \mathcal{S}(X)$ we may consider the Seiberg–Witten equations on X with respect to the metric g, spin^c-structure $\mathfrak{s}$ and 2-form perturbation η. Let $\mathcal{M}(X,\mathfrak{s},h)$ denote the moduli space of gauge equivalence classes of solutions to the Seiberg–Witten equations for $(X,\mathfrak{s},h)$. Assume $b_+(X) > 2$. We will say that $h \in \Pi$ is regular if $\mathcal{M}(X,\mathfrak{s},h)$ is empty for all $\mathfrak{s} \in S(X)$. Since $b_+(X) > 0$ and the expected dimension of $\mathcal{M}(X,\mathfrak{s},h)$ is negative, the regular elements form a subset of Π of Baire second category with respect to the $\mathcal{C}^\infty$ topology. Let $\Pi^{\mathrm{reg}} \subseteq \Pi$ denote the set of regular elements.

Suppose that $h_0, h_1 \in \Pi^{\mathrm{reg}}$. If $h : [0,1] \to \Pi$ is a path in Π from h_0 to h_1, we can consider the families moduli space, which is the union over $t \in [0,1]$ of the Seiberg–Witten moduli spaces for each $h_t \in \Pi$. For a sufficiently generic path h_t, the moduli space is a compact, smooth, 0-dimensional manifold. A choice of orientation on a maximal positive definite subspace of $H^2(X;\mathbb{R})$ determines an orientation on the moduli space and hence we can count with sign the number of points in the moduli space. Fix a choice of such an orientation. It can be shown [24] that the number of solutions depends on the endpoints h_0, h_1, but not on the choice of generic path h_t. Hence we may denote by $\mathrm{SW}_{\mathfrak{s}}(h_0, h_1) \in \mathbb{Z}$ the signed count of points in the moduli space. From the definition it is clear that this count of points satisfies the following properties:

(1) $\mathrm{SW}_{\mathfrak{s}}(h_0, h_1) + \mathrm{SW}_{\mathfrak{s}}(h_1, h_2) = \mathrm{SW}_{\mathfrak{s}}(h_0, h_2)$.

(2) $\mathrm{SW}_{\mathfrak{s}}(h_0, h_1) = \mathrm{sgn}_+(f)\, \mathrm{SW}_{f(\mathfrak{s})}(f(h_0), f(h_1))$ for any orientation-preserving diffeomorphism f.

In (2), $\mathrm{sgn}_+(f)$ is defined as follows. The space of oriented, maximal positive definite subspaces of $H^2(X;\mathbb{R})$ has two connected components. For an isometry φ of $H^2(X;\mathbb{R})$ we let $\mathrm{sgn}_+(\varphi)$ equal 1 or -1 according to whether φ preserves or exchanges the two components. If f is an orientation-preserving diffeomorphism of X, then $\mathrm{sgn}_+(f)$ denotes $\mathrm{sgn}_+(f_*)$, where $f_* = (f^{-1})^*$ is the isometry of $H^2(X;\mathbb{R})$ induced by f.

Property (1) follows by concatenating a path from h_0 to h_1 with a path from h_1 to h_2. Property (2) follows from diffeomorphism invariance of the Seiberg–Witten equations. In addition, $\mathrm{SW}_{\mathfrak{s}}(h_0, h_1)$ obeys, with respect to charge conjugation, the symmetry

(3) $\mathrm{SW}_{\mathfrak{s}}(h_0, h_1) = (-1)^{b_+(X)/2+1}\, \mathrm{SW}_{\bar{\mathfrak{s}}}(\bar{h}_0, \bar{h}_1)$,

where $\bar{\mathfrak{s}}$ denotes the charge conjugate of $\mathfrak{s}$ and for $h = (g, \eta) \in \Pi$, we set $\bar{h} = (g, -\eta)$. Property (3) is an immediate consequence of the charge conjugation symmetry of the Seiberg–Witten equations.

Let $f \in T(X)$ be an element of the Torelli group. Fix a spin^c-structure $\mathfrak{s}$ with $d(\mathfrak{s}) = -1$. The mapping cylinder of f defines a smooth family $E \to S^1$ over S^1 with fibres diffeomorphic to X. Since f acts trivially on cohomology, it preserves the isomorphism class of $\mathfrak{s}$. It follows easily that there is a unique spin^c-structure on the vertical tangent bundle of E which restricts to $\mathfrak{s}$ on each fibre. Since $b_+(X) > 2$, there is a single chamber for the families Seiberg–Witten equations for E. Furthermore, the families moduli space is oriented and so we obtain an integer-valued invariant $\mathrm{SW}_{\mathfrak{s}}(f) \in \mathbb{Z}$ which depends only on $(X, \mathfrak{s})$ and the isotopy class of f (see [3] for more details). From the definition of $\mathrm{SW}_{\mathfrak{s}}(f)$, it is easy to see that

$$\mathrm{SW}_{\mathfrak{s}}(f) = \mathrm{SW}_{\mathfrak{s}}(h, f(h))$$

for any $h \in \Pi^{\mathrm{reg}}$. It is instructive to see why $\mathrm{SW}_{\mathfrak{s}}(f)$ is independent of the choice of $h \in \Pi^{\mathrm{reg}}$. Let $h' \in \Pi^{\mathrm{reg}}$. Then

$$\begin{aligned} \mathrm{SW}_{\mathfrak{s}}(h', f(h')) &= -\mathrm{SW}_{\mathfrak{s}}(h, h') + \mathrm{SW}_{\mathfrak{s}}(h, f(h)) + \mathrm{SW}_{\mathfrak{s}}(f(h), f(h')) \\ &= -\mathrm{SW}_{\mathfrak{s}}(h, h') + \mathrm{SW}_{\mathfrak{s}}(h, f(h)) + \mathrm{SW}_{f^{-1}(\mathfrak{s})}(h, h') \\ &= \mathrm{SW}_{\mathfrak{s}}(h, f(h)), \end{aligned}$$

where the last line follows from $f^{-1}(\mathfrak{s}) = \mathfrak{s}$, which holds since $f \in T(X)$.

For $f, g \in T(X)$, we have that $\mathrm{SW}_{\mathfrak{s}}(f \circ g) = \mathrm{SW}_{\mathfrak{s}}(f) + \mathrm{SW}_{\mathfrak{s}}(g)$ (this is a special case of Proposition 2.1, proven below). Therefore $\mathrm{SW}_{\mathfrak{s}}$ defines a homomorphism

$$\mathrm{SW}_{\mathfrak{s}} : T(X) \to \mathbb{Z}$$

or equivalently, a cohomology class $\mathrm{SW}_{\mathfrak{s}} \in H^1(T(X);\mathbb{Z})$. These cohomology classes generally do not extend to the full mapping class group $M(X)$, because $\Gamma(X)$ acts nontrivially on the set of spin^c-structures.

Recall that the compactness of the Seiberg–Witten moduli space follows from a priori bounds. These bounds depend on the pair $h \in \Pi$, but not on the spin^c-structure. This argument also works for families over a compact base space, hence for fixed $f \in T(X)$, $\mathrm{SW}_{\mathfrak{s}}(f)$ is nonzero for only finitely many $\mathfrak{s} \in S(X)$. Therefore, we can collect the homomorphisms $\mathrm{SW}_{\mathfrak{s}}$ into a single invariant

$$\mathrm{SW} : T(X) \to \bigoplus_{\mathfrak{s}\in S(X)} \mathbb{Z}, \quad f \mapsto \bigoplus_{\mathfrak{s}} \mathrm{SW}_{\mathfrak{s}}(f).$$

In what follows, we will see that SW can be extended from $T(X)$ to the full mapping class group $M(X)$ as a cohomology class valued in a certain $\Gamma(X)$-module.

Recall that each $\mathfrak{s} \in S(X)$ is determined by the corresponding characteristic element $c(\mathfrak{s}) \in L$. Therefore the group $\Gamma(X)$ acts on $\mathcal{S}(X)$ and hence on $\mathbb{Z}[\mathcal{S}(X)]$, the free abelian group with basis $\mathcal{S}(X)$. Let $\widehat{\mathbb{Z}}$ denote $\mathbb{Z}$ equipped with the action of $\Gamma(X)$ such that $f \in \Gamma(X)$ acts as multiplication by $\mathrm{sgn}_+(f)$. Let $\widehat{\mathbb{Z}}[\mathcal{S}(X)] = \widehat{\mathbb{Z}} \otimes_{\mathbb{Z}} \mathbb{Z}[\mathcal{S}(X)]$. It will be convenient to regard $\widehat{\mathbb{Z}}[\mathcal{S}(X)]$ as the group of functions $\phi : \mathcal{S}(X) \to \mathbb{Z}$ with finite support. Then the action of $f \in \Gamma(X)$ is given by $(f\phi)(\mathfrak{s}) = \mathrm{sgn}_+(f)\phi(f^{-1}(\mathfrak{s}))$. We will show that the families Seiberg–Witten invariant for 1-dimensional families (where $b_+(X) > 2$) can be viewed as an element of $H^1(M(X);\widehat{\mathbb{Z}}[\mathcal{S}(X)])$.

Recall that for a group G and a G-module M, the group $H^1(G;M)$ can be viewed as the set of equivalence classes of twisted homomorphisms $G \to M$. A twisted homomorphism is a map $\phi : G \to M$ such that $\phi(gh) = \phi(g) + g\phi(h)$. A trivial twisted homomorphism is a twisted homomorphism of the form $\phi(g) = gm - m$ for some $m \in M$. Two twisted homomorphisms are considered equivalent if they differ by a trivial twisted homomorphism.

Let $h \in \Pi^{\mathrm{reg}}$. Define a map $\phi_h : \mathrm{Diff}_+(X) \to \widehat{\mathbb{Z}}[\mathcal{S}(X)]$ from the group of orientation-preserving diffeomorphisms to $\widehat{\mathbb{Z}}[\mathcal{S}(X)]$ by setting

$$(\phi_h(f))(\mathfrak{s}) = \mathrm{SW}_{\mathfrak{s}}(h, f(h)).$$

Suppose that $f_0, f_1 \in \mathrm{Diff}_+(X)$ are isotopic. Choose an isotopy f_t. Then

$$\begin{aligned}\phi_h(f_1) &= \mathrm{SW}_{\mathfrak{s}}(h, f_1(h)) \\ &= \mathrm{SW}_{\mathfrak{s}}(h, f_0(h)) + \mathrm{SW}_{\mathfrak{s}}(f_0(h), f_1(h)) \\ &= \phi_h(f_0) + \mathrm{SW}_{\mathfrak{s}}(f_0(h), f_1(h)).\end{aligned}$$

Consider the path $h_t = f_t(h)$ from $f_0(h)$ to $f_1(h)$. By diffeomorphism invariance of the Seiberg–Witten equations, the Seiberg–Witten moduli space for h_t is empty for each $t \in [0,1]$, and hence $\mathrm{SW}_{\mathfrak{s}}(f_0(h), f_1(h)) = 0$ and $\phi_h(f_1) = \phi_h(f_2)$. This shows that ϕ only depends on the underlying isotopy class and so we may view it as a map $\phi_h : M(X) \to \widehat{\mathbb{Z}}[\mathcal{S}(X)]$.

Proposition 2.1 *The map $\phi_h : M(X) \to \widehat{\mathbb{Z}}[\mathcal{S}(X)]$ is a twisted homomorphism. Furthermore, the underlying cohomology class $[\phi_h] \in H^1(M(X);\widehat{\mathbb{Z}}[\mathcal{S}(X)])$ does not depend on the choice of $h \in \Pi^{\mathrm{reg}}$.*

Proof Let $f, g \in M(X)$. Then

$$\begin{aligned}\phi_h(gf)(\mathfrak{s}) &= \mathrm{SW}_{\mathfrak{s}}(h, gf(h)) \\ &= \mathrm{SW}_{\mathfrak{s}}(h, g(h)) + \mathrm{SW}_{\mathfrak{s}}\big(g(h), g(f(h))\big) \\ &= \mathrm{SW}_{\mathfrak{s}}(h, g(h)) + \mathrm{sgn}_+(g)\,\mathrm{SW}_{g^{-1}\mathfrak{s}}(h, f(h)) \\ &= \phi_h(g)(\mathfrak{s}) + \mathrm{sgn}_+(g)\phi_h(f)(g^{-1}\mathfrak{s}) \\ &= (\phi_h(g) + (g\phi_h)(f))(\mathfrak{s}).\end{aligned}$$

Hence ϕ_h is a twisted homomorphism. Next we show that the underlying cohomology class of ϕ_h does not depend on the choice of h. Let $h' \in \Pi^{\mathrm{reg}}$ be another generic pair. Choose a path h_t from $h = h_0$ to $h' = h_1$. Then

$$\begin{aligned}\phi_{h'}(f)(\mathfrak{s}) &= \mathrm{SW}_{\mathfrak{s}}(h_1, f(h_1)) \\ &= \mathrm{SW}_{\mathfrak{s}}(h_0, f(h_0)) - \mathrm{SW}_{\mathfrak{s}}(h_0, h_1) + \mathrm{SW}_{\mathfrak{s}}(f(h_0), f(h_1)) \\ &= \phi_h(f)(\mathfrak{s}) + \mathrm{sgn}_+(f)\,\mathrm{SW}_{f^{-1}\mathfrak{s}}(h_0, h_1) - \mathrm{SW}_{\mathfrak{s}}(h_0, h_1) \\ &= \phi_h(f)(\mathfrak{s}) + (fm - m)(\mathfrak{s}),\end{aligned}$$

where $m(\mathfrak{s}) = \mathrm{SW}_{\mathfrak{s}}(h_0, h_1)$. Hence ϕ_h and $\phi_{h'}$ define the same cohomology class. □

Definition 2.2 We denote by

$$\mathrm{SW} \in H^1(M(X); \widehat{\mathbb{Z}}[\mathcal{S}(X)])$$

the cohomology class $\mathrm{SW} = [\phi_h]$ for any $h \in \Pi^{\mathrm{reg}}$.

Let $M_+(X)$ denote the subgroup of $M(X)$ consisting of all $f \in M(X)$ for which $\mathrm{sgn}_+(f) = 1$. Then $M_+(X)$ has index 1 or 2 in $M(X)$. Observe that $\widehat{\mathbb{Z}}|_{M_+(X)} = \mathbb{Z}$; thus $\mathrm{SW}\,|_{M_+(X)} \in H^1(M(X); \mathbb{Z}[\mathcal{S}(X)])$. From $\mathrm{SW}\,|_{M_+(X)}$ we can extract $\mathbb{Z}$-valued cohomology classes as follows: let $\mathcal{O} \subseteq \mathcal{S}(X)$ be a $\Gamma(X)$-invariant subset of $\mathcal{S}(X)$. Then we obtain a morphism $p_{\mathcal{O}} : \mathbb{Z}[\mathcal{S}(X)] \to \mathbb{Z}$ given by $\phi \mapsto \sum_{\mathfrak{s} \in \mathcal{O}} \phi(\mathfrak{s})$. We define $\mathrm{SW}_{\mathcal{O}} \in H^1(M_+(X); \mathbb{Z})$ by setting $\mathrm{SW}_{\mathcal{O}} = p_{\mathcal{O}}(\mathrm{SW}\,|_{M_+(X)})$. From this definition it follows that

$$\mathrm{SW}_{\mathcal{O}}\,|_{T(X)} = \sum_{\mathfrak{s} \in \mathcal{O}} \mathrm{SW}_{\mathfrak{s}}\,.$$

Furthermore, for any $f \in M_+(X)$, we have

$$\mathrm{SW}_{\mathcal{O}}(f) = \sum_{\mathfrak{s} \in \mathcal{O}} \mathrm{SW}_{\mathfrak{s}}(h, f(h)),$$

where $h \in \Pi^{\mathrm{reg}}$.

Remark 2.3 Ruberman [26] defined an invariant $\mathrm{SW}^{\mathrm{tot}}_{\mathfrak{s}}$ which is similar to the definition of $\mathrm{SW}_{\mathcal{O}}$ given above. Namely $\mathrm{SW}^{\mathrm{tot}}_{\mathfrak{s}} : M_+(X) \to \mathbb{Z}$ is given by $\mathrm{SW}^{\mathrm{tot}}_{\mathfrak{s}}(f) = \sum_{\mathfrak{s}'} \mathrm{SW}_{\mathfrak{s}'}(h, f(h))$ where the sum is over all spin^c-structures $\mathfrak{s}'$ such that $\mathfrak{s}' = f^m(\mathfrak{s})$ for some $m \in \mathbb{Z}$. However this invariant is not a group homomorphism and behaves in a complicated manner with respect to composition (see [26, Theorem 3.4]). For this reason, we find it more useful to work with the invariants $\mathrm{SW}_{\mathcal{O}}$.

Let $\mathfrak{s} \in \mathcal{S}(X)$ be a spin^c-structure and let $f \in M(X)$. Suppose that f preserves $\mathfrak{s}$. If $\mathrm{sgn}_+(f) = 1$, then the families Seiberg–Witten moduli space for the mapping cylinder of f with spin^c-structure $\mathfrak{s}$ is oriented, and we obtain an integer families Seiberg–Witten invariant $\mathrm{SW}_{\mathfrak{s}}(f) \in \mathbb{Z}$. It is given by $\mathrm{SW}_{\mathfrak{s}}(f) = \mathrm{SW}_{\mathfrak{s}}(h, f(h))$, for any $h \in \Pi^{\mathrm{reg}}$. If $\mathrm{sgn}_+(f) = -1$ then there is no natural choice of orientation on the families moduli space, hence we only get a mod-2 invariant $\mathrm{SW}_{\mathfrak{s}}(f) \in \mathbb{Z}_2$ which is given by $\mathrm{SW}_{\mathfrak{s}}(f) = \mathrm{SW}_{\mathfrak{s}}(h, f(h)) \pmod 2$, for any $h \in \Pi^{\mathrm{reg}}$ (the value of $\mathrm{SW}_{\mathfrak{s}}(h, f(h))$ depends on h, but its mod-2 reduction does not).

We will make use of the following special case of the gluing formula of [3]:

Proposition 2.4 *Suppose that $X = X' \# (S^2 \times S^2)$, where $b_+(X') > 1$. Let $\mathfrak{s}'$ be a spin^c-structure on X' with $d(\mathfrak{s}') = 0$ and let $\mathfrak{s}_0$ denote the spin structure on $S^2 \times S^2$. Let $\mathfrak{s} = \mathfrak{s}' \# \mathfrak{s}_0$. Let f' be a diffeomorphism on X' that preserves $\mathfrak{s}'$ and ρ a diffeomorphism of $S^2 \times S^2$. Suppose that f' is trivial in a neighbourhood of a point $x_1 \in X'$ and that ρ is trivial in a neighbourhood of a point $x_2 \in S^2 \times S^2$. Set $f = f' \# \rho$, where the connected sum is performed by removing balls around x_1 and x_2 and identifying boundaries. Then*:

1. *If $\mathrm{sgn}_+(\rho) = 1$, then $\mathrm{SW}_{\mathfrak{s}}(f) = 0 \pmod 2$.*
2. *If $\mathrm{sgn}_+(\rho) = -1$, then $\mathrm{SW}_{\mathfrak{s}}(f) = \mathrm{SW}(X', \mathfrak{s}') \pmod 2$, where $\mathrm{SW}(X', \mathfrak{s}')$ denotes the ordinary Seiberg–Witten invariant of $(X', \mathfrak{s}')$.*

3 Nonfinitely generated mapping class groups

In this section we prove that $M(X)$ is not finitely generated for certain 4-manifolds.

Theorem 3.1 *Let $X = 2n\mathbb{CP}^2 \# 10n\overline{\mathbb{CP}^2}$, where $n \geq 3$ is odd. Then $M(X)$ is not finitely generated. More precisely, the following holds*:

1. *There is a surjective homomorphism $\Phi : M_+(X) \to \mathbb{Z}^\infty$ from $M_+(X)$ to $\mathbb{Z}^\infty$, where $\mathbb{Z}^\infty$ denotes a free abelian group of countably infinite rank.*
2. *The mod-2 reduction of Φ extends to a homomorphism $\Phi : M(X) \to \mathbb{Z}_2^\infty$.*
3. *$M_+(X)$ has index 2 in $M(X)$.*
4. *The short exact sequence $1 \to M_+(X) \to M(X) \to \mathbb{Z}_2 \to 0$ splits.*

Proof First note that $X = X' \# (S^2 \times S^2)$, where $X' = (2n-1)\mathbb{CP}^2 \# (10n-1)\overline{\mathbb{CP}^2}$. It follows from [31] that $\Gamma(X) = \mathrm{Aut}(H^2(X; \mathbb{Z}))$. Observe that $d(\mathfrak{s}) = (c(\mathfrak{s})^2 + 8n)/4 - 2n - 1 = c(\mathfrak{s})^2/4 - 1$. Hence $d(\mathfrak{s}) = -1$ if and only if $c(\mathfrak{s})^2 = 0$. For each $k \geq 1$, let $\mathcal{O}_k \subset \mathcal{S}(X)$ denote the set of spin^c-structures whose characteristic c satisfies $c \neq 0$, $c^2 = 0$, and c is k times a primitive element. We will show that the homomorphism

$$\Phi = \bigoplus_{q=1}^{\infty} \mathrm{SW}_{\mathcal{O}_{nq-q-1}} : M_+(X) \to \mathbb{Z}^\infty$$

surjects to a subgroup of $\mathbb{Z}^\infty$ of countably infinite rank.

The decomposition $X = X' \# (S^2 \times S^2)$ yields an orthogonal decomposition $L = L' \oplus H$, where L, L' are the intersection forms of X, X' and $H = H^2(S^2 \times S^2; \mathbb{Z})$ is the hyperbolic lattice. Any characteristic $c \in L$ decomposes as $c = (c_1, c_2)$, where $c_1 \in L', c_2 \in H$ are characteristic. The intersection form L' is odd, hence $c_1 \neq 0$.

We will partition $\mathcal{O}_k$ into two types of subsets:

(1) subsets $\{\mathfrak{s}, \bar{\mathfrak{s}}\}$, where $\mathfrak{s} = \mathfrak{s}' \# \mathfrak{s}_0$ and $c(\mathfrak{s}_0) = 0$,

(2) subsets $\{\mathfrak{s}_1, \mathfrak{s}_2, \bar{\mathfrak{s}}_1, \bar{\mathfrak{s}}_2\}$, where $\mathfrak{s}_1 = \mathfrak{s}' \# \mathfrak{s}''$, $\mathfrak{s}_2 = \mathfrak{s}' \# \bar{\mathfrak{s}}''$ and where $c(\mathfrak{s}'') \neq 0$.

Since $b_+(X) = 2n = 2 \pmod 4$, we have

$$\mathrm{SW}_{\mathfrak{s}}(t) = \mathrm{SW}_{\bar{\mathfrak{s}}}(t)$$

for all $t \in T(X)$. Hence a subset of type of $\mathcal{O}_k$ of type (1) will contribute $2\,\mathrm{SW}_{\mathfrak{s}}(t)$ to $\mathrm{SW}_{\mathcal{O}_k}(t)$ and a subset of type (2) will contribute $2(\mathrm{SW}_{\mathfrak{s}_1}(t) + \mathrm{SW}_{\mathfrak{s}_2}(t))$.

Let $E(n)_q$ be the elliptic surface obtained from $E(n)$ by performing a logarithmic transform of multiplicity $q \geq 1$. Since n is odd, $E(n)_q$ is not spin and its intersection form is diagonal of signature $(2n-1, 10n-1)$. Hence $E(n)_q$ has the same intersection lattice as X'. Furthermore, we have that $E(n)_q \# (S^2 \times S^2)$ is diffeomorphic to $2n\mathbb{CP}^2 \# 10n\overline{\mathbb{CP}}^2 = X = X' \# (S^2 \times S^2)$ [12, page 320]. So there is an orientation-preserving diffeomorphism $\psi_q : E(n)_q \# (S^2 \times S^2) \to X' \# (S^2 \times S^2)$. We can choose ψ_q so that it respects the decomposition $H^2(E(n)_q; \mathbb{Z}) \oplus H^2(S^2 \times S^2; \mathbb{Z}) \to H^2(X'; \mathbb{Z}) \oplus H^2(S^2 \times S^2; \mathbb{Z})$. To see this, first let $\psi'_q : E(n)_q \# (S^2 \times S^2) \to X' \# (S^2 \times S^2)$ be any diffeomorphism. Then by [31], every isometry of $H^2(X' \# (S^2 \times S^2); \mathbb{Z})$ can be realised by a diffeomorphism. Hence, composing ψ'_q with a suitable diffeomorphism of X, we obtain the desired diffeomorphism ψ_q.

Let ρ be a diffeomorphism of $S^2 \times S^2$ which acts as -1 on $H^2(S^2 \times S^2; \mathbb{Z})$ and is trivial in a neighbourhood of some point. Such diffeomorphisms can easily be constructed, for example, take the product $r \times r$ of two copies of a reflection on S^2 and then isotopy it to act trivially in a neighbourhood of a point. Define a diffeomorphism $f_0 \in M(X)$ by setting $f_0 = \mathrm{id}_{X'} \# \rho$, where the connected sum is performed by removing a ball of $(S^2 \times S^2)$ on which ρ acts trivially. For each $q \geq 1$, define a diffeomorphism $f_q \in M(X)$ by setting $f_q = \psi_q \circ (\mathrm{id}_{E(n)_q} \# \rho) \circ \psi_q^{-1}$. Note that $\mathrm{sgn}_+(f_0) = \mathrm{sgn}_+(f_q) = -1$. Also $t_q = f_q \circ f_0 \in T(X)$.

We claim that

$$\mathrm{SW}_{\mathcal{O}_{nq-q-1}}(t_q) = 2 \pmod 4 \quad \text{and} \quad \mathrm{SW}_{\mathcal{O}_{nq'-q'-1}}(t_q) = 0 \pmod 4$$

for all $q' > q$. This implies that the elements $\{\Phi(t_q)\}_{q \geq 1}$ are linearly independent. To see this, first note that $\Phi(t_q) \in 2\mathbb{Z}^\infty$ by charge conjugation symmetry and that the image of $\{\Phi(t_q)/2\}_{q \geq 1}$ is a basis for $\mathbb{Z}_2^\infty$, by the above claim. Now suppose $n_1\Phi(t_1) + n_2\Phi(t_2) + \cdots + n_r\Phi(t_r) = 0$ for some $n_1, \ldots, n_r \in \mathbb{Z}$, not all zero. Without loss of generality we can assume that $\gcd(n_1, \ldots, n_r) = 1$. Then $n_1\Phi(t_1)/2 + \cdots + n_r\Phi(t_r)/2 = 0$. But $\{\Phi(t_q)/2\}_{q \geq 1}$ are linearly independent in $\mathbb{Z}_2^\infty$, so $n_1, \ldots, n_r$ are all even, a contradiction.

Now we prove the claim. Let $q, q' \geq 1$ and set $k = nq' - q' - 1$. By partitioning $\mathcal{O}_k$ into subsets of type (1) and (2) as described above, we can then write $\mathrm{SW}_{\mathcal{O}_k}(t_q)$ as a sum of contributions from sets of type (1) and (2). Consider a contribution from a subset $\{\mathfrak{s}, \bar{\mathfrak{s}}\}$ of type (1). So $\mathfrak{s} = \mathfrak{s}' \# \mathfrak{s}_0$. The contribution is $2\,\mathrm{SW}_{\mathfrak{s}}(t_q)$. Since f_q and f_0 both preserve $\mathfrak{s}$, we have that

$$\begin{aligned}\mathrm{SW}_{\mathfrak{s}}(t_q) &= \mathrm{SW}_{\mathfrak{s}}(f_q \circ f_0)\\ &= \mathrm{SW}_{\mathfrak{s}}(f_q) + \mathrm{SW}_{\mathfrak{s}}(f_0) \pmod 2\\ &= \mathrm{SW}(E(n)_q, \mathfrak{s}') \pmod 2,\end{aligned}$$

where the last equality is due to Proposition 2.4. Let $f \in H^2(E(n)_q; \mathbb{Z})$ denote the class of the multiple fibre. Then f is nonzero, primitive and $f^2 = 0$. From the well-known calculation of the Seiberg–Witten invariants for elliptic surfaces [21, Chapter 3], we have that $\mathrm{SW}(E(n)_q, \mathfrak{s}') = 0$ unless $c(\mathfrak{s}')$ is a multiple of f. More precisely, $\mathrm{SW}(E(n)_q, \mathfrak{s}')$ is zero unless $c(\mathfrak{s}') = 2(qk + a)f - (nq - q - 1)f$, where $0 \leq k \leq n-2$ and $0 \leq a < q$. In such a case $\mathrm{SW}(E(n)_q, \mathfrak{s}') = (-1)^k \binom{n-2}{k}$. In particular, $\mathrm{SW}(E(n)_q, \mathfrak{s}') = \pm 1$ if $c(\mathfrak{s}') = (nq - q - 1)f$ and $\mathrm{SW}(E(n)_q, \mathfrak{s}') = 0$ if $c(\mathfrak{s}) = uf$, $u > nq - q - 1$. Now suppose that $q' \geq q$. We have that $\mathrm{SW}(E(n)_q, \mathfrak{s}') = 0$ unless $c(\mathfrak{s})$ is a multiple of f. But since $\mathfrak{s} = \mathfrak{s}' \# \mathfrak{s}_0 \in \mathcal{O}_k$, this can only happen if $c(\mathfrak{s}') = \pm kf$ (recall that $\mathcal{O}_k$ is the set of spin^c-structures whose characteristic c satisfies $c \neq 0$, $c^2 = 0$, and c is k times a primitive element). Hence if $q' > q$, then every pair $\{\mathfrak{s}, \bar{\mathfrak{s}}\}$ of type (1) contributes 0 mod 4 to $\mathrm{SW}_{\mathcal{O}_k}(t_q)$. If $q' = q$, then there is exactly one pair $\{\mathfrak{s}, \bar{\mathfrak{s}}\}$ of type (1) which contributes 2 mod 4 to $\mathrm{SW}_{\mathcal{O}_k}(t_q)$ and all other pairs are 0 mod 4.

Now consider the contribution from a set $\{\mathfrak{s}_1, \mathfrak{s}_2, \bar{\mathfrak{s}}_1, \bar{\mathfrak{s}}_2\}$ of type (2), where $\mathfrak{s}_1 = \mathfrak{s}' \# \mathfrak{s}''$, $\mathfrak{s}_2 = \mathfrak{s}' \# \bar{\mathfrak{s}}''$ and where $c(\mathfrak{s}'') \neq 0$. As seen above, the contribution is $2(\mathrm{SW}_{\mathfrak{s}_1}(t_q) + \mathrm{SW}_{\mathfrak{s}_2}(t_q))$. We will show that $\mathrm{SW}_{\mathfrak{s}_1}(t_q) + \mathrm{SW}_{\mathfrak{s}_2}(t_q) = 0 \pmod 2$, hence all subsets of type (2) contribute 0 mod 4 and this will prove the claim.

Observe that $\mathfrak{s}_2 = f_q(\mathfrak{s}_1) = f_0(\mathfrak{s}_1)$. Let $h \in \Pi^{\mathrm{reg}}$. Then

$$\begin{aligned}\mathrm{SW}_{\mathfrak{s}_1}(t_q) &= \mathrm{SW}_{\mathfrak{s}_1}(h, t_q(h))\\ &= \mathrm{SW}_{\mathfrak{s}_1}(h, f_q f_0(h))\\ &= \mathrm{SW}_{\mathfrak{s}_1}(h, f_q(h)) + \mathrm{SW}_{\mathfrak{s}_1}(f_q(h), f_q f_0(h))\\ &= \mathrm{SW}_{\mathfrak{s}_1}(h, f_q(h)) - \mathrm{SW}_{\mathfrak{s}_2}(h, f_0(h)).\end{aligned}$$

Similarly, $\mathrm{SW}_{\mathfrak{s}_2}(t_q) = \mathrm{SW}_{\mathfrak{s}_2}(h, f_q(h)) - \mathrm{SW}_{\mathfrak{s}_1}(h, f_0(h))$. Hence

$$\begin{aligned}&\mathrm{SW}_{\mathfrak{s}_1}(t_q) + \mathrm{SW}_{\mathfrak{s}_2}(t_q)\\ &\qquad= \big(\mathrm{SW}_{\mathfrak{s}_1}(h, f_q(h)) + \mathrm{SW}_{\mathfrak{s}_2}(h, f_q(h))\big) - \big(\mathrm{SW}_{\mathfrak{s}_1}(h, f_0(h)) + \mathrm{SW}_{\mathfrak{s}_2}(h, f_0(h))\big)\\ &\qquad= \big(\mathrm{SW}_{\mathfrak{s}_1}(h, f_q(h)) - \mathrm{SW}_{\mathfrak{s}_1}(f_q(h), f_q^2(h))\big) - \big(\mathrm{SW}_{\mathfrak{s}_1}(h, f_0(h)) - \mathrm{SW}_{\mathfrak{s}_1}(f_0(h), f_0^2(h))\big)\\ &\qquad= \mathrm{SW}_{\mathfrak{s}_1}(h, f_q^2(h)) - \mathrm{SW}_{\mathfrak{s}_1}(h, f_0^2(h))\\ &\qquad= \mathrm{SW}_{\mathfrak{s}_1}(f_q^2) - \mathrm{SW}_{\mathfrak{s}_1}(f_0^2)\\ &\qquad= 0 \pmod 2,\end{aligned}$$

where the last line follows from Proposition 2.4. This proves the claim. Hence we have proven that there exists a surjective homomorphism $M_+(X) \to \mathbb{Z}^\infty$.

The fact the mod-2 reduction of Φ extends to $M(X)$ follows by noting that $\frac{1}{2}\,\mathrm{SW}_{\mathcal{O}_k} \in H^1(M_+(X);\mathbb{Z})$ extends to $H^1(M(X);\widehat{\mathbb{Z}})$ and then applying the mod-2 reduction map $\widehat{\mathbb{Z}} \to \mathbb{Z}_2$.

The fact that $M_+(X)$ has index 2 in $M(X)$ follows immediately from $\mathrm{sgn}_+(f_0) = -1$. Furthermore, it is easy to see that f_0^2 is isotopic to a Dehn twist on the neck of the connected sum $X' \# (S^2 \times S^2)$. By using the circle action on $S^2 \times S^2$, it follows that this is isotopic to the identity. So f_0 defines a splitting $\mathbb{Z}_2 \to M(X)$ of the sequence $1 \to M_+(X) \to M(X) \to \mathbb{Z}_2 \to 0$. □

4 Split extensions

Let $L = \mathbb{Z}^n$ denote the standard diagonal lattice of rank n with orthonormal basis $e_1, \dots, e_n$. The isometry group of L is the hyperoctahedral group H_n, which is also the Coxeter group of type BC_n. This group is easily seen to be generated by permutations of $e_1, \dots, e_n$ and the reflections $r_1, \dots, r_n$ in the hyperplanes orthogonal to $e_1, \dots, e_n$. The reflections generate a normal subgroup isomorphic to $\mathbb{Z}_2^n$ and H_n is the semidirect product $H_n = S_n \ltimes \mathbb{Z}_2^n$.

Let $X = n\mathbb{CP}^2$ be the connected sum of n copies of $\mathbb{CP}^2$. Then $H^2(X;\mathbb{Z})$ is isomorphic to L with $e_1, \dots, e_n$ corresponding to generators of $H^2(\mathbb{CP}^2;\mathbb{Z})$ for the n summands of X. It is not hard to see that $\Gamma(X)$ is equal to the full isometry group of L. This will also follow from the construction given below.

Theorem 4.1 *For $X = n\mathbb{CP}^2$, there is a splitting $\Gamma(X) \to M(X)$.*

Proof We will construct a smooth fibre bundle $\pi : E \to B$ with fibres diffeomorphic to X and such that the monodromy of the local system $R^2\pi_*\mathbb{Z}$ yields an isomorphism $\rho : \pi_1(B) \to \mathrm{Aut}(L)$. The geometric monodromy of the family defines a lift $\tilde{\rho} : \pi_1(B) \to \pi_0(\mathrm{Diff}(X)) = M(X)$ of ρ to $M(X)$. Then $\tilde{\rho} \circ \rho^{-1} : \mathrm{Aut}(L) \to M(X)$ is the desired splitting (this also proves that $\Gamma(X) \cong \mathrm{Aut}(L)$).

Let C_m be the space of m-tuples of distinct points on S^4. Clearly C_1 is diffeomorphic to S^4. For $m > 1$ there is a natural map $C_m \to C_{m-1}$ given by forgetting the m-th point. This map gives C_m the structure of a fibre bundle over C_{m-1} with fibre F_{m-1} the 4-sphere with $m-1$ points removed. Since $\pi_1(F_{m-1}) = \pi_0(F_{m-1}) = 1$, the long exact sequence in homotopy yields an isomorphism $\pi_1(C_m) \cong \pi_1(C_m)$. Since $\pi_1(C_1) = \pi_1(S^4) = 1$, it follows by induction that $\pi_1(C_n) = 1$ for all n.

Fix an orientation on S^4. Let $\widetilde{C}_n$ denote the space consisting of an n-tuple $(x_1, \dots, x_n)$ of distinct points of S^4 together with an n-tuple $(I_1, \dots, I_n)$, where I_j is a complex structure on $T_{x_j}S^4$ which induces the given orientation. The forgetful map $\widetilde{C}_n \to C_n$ which forgets the complex structures $I_1, \dots, I_n$ gives $\widetilde{C}_n$ the structure of a fibre bundle over C_n. Since the space of complex structures on $\mathbb{R}^4$ compatible with a given orientation is isomorphic to $\mathrm{SO}(4)/U(2) \cong S^2$, it follows that the fibres of $\widetilde{C}_n \to C_n$ are isomorphic to $(S^2)^n$. The long exact sequence in homotopy implies that $\pi_1(\widetilde{C}_n) = 1$.

Consider the trivial family $\widetilde{E}_0 = \widetilde{C}_n \times S^4 \to \widetilde{C}_n$. This family is equipped with n sections $s_1, \dots, s_n$, where $s_j((x_1, \dots, x_n), (I_1, \dots, I_n)) = x_j$. The normal bundle of s_j is $N_j = T_{x_j}S^4$. The complex structure I_j gives N_j the structure of a complex rank 2 vector bundle. Therefore, we can form a family $\widetilde{E}_n$

by blowing up $\widetilde{E}_0$ along the sections $s_1, \dots, s_n$. More precisely, consider the fibre bundle over $\widetilde{C}_n$ with fibre $\mathbb{CP}^2$ given by the projective bundle $\mathbb{P}(\mathbb{C} \oplus N_j)$. This bundle has a natural section t_j corresponding to the 1-dimensional subbundle $\mathbb{C} \subset \mathbb{C} \oplus N_j$. The normal bundle of t_j is isomorphic to N_j. Since s_j and t_j have isomorphic normal bundles, we can attach $\mathbb{P}(\mathbb{C} \oplus N_j)$ to $\widetilde{E}_0$ by removing tubular neighbourhoods of s_j and t_j and identifying the boundaries.

The hyperoctahedral group $H_n = S_n \ltimes \mathbb{Z}_2^n$ acts on $\widetilde{C}_n$ as follows. The permutation group S_n acts by permuting the points $x_1, \dots, x_n$ as well as the corresponding complex structures $I_1, \dots, I_n$. The group $\mathbb{Z}_2^n$ is generated by reflections $r_1, \dots, r_n$. We let r_j act by fixing $x_1, \dots, x_n$, sending I_j to $-I_j$ and fixing the remaining complex structures. The action of H_n is free and we let $B = \widetilde{C}_n / H_n$ be the quotient. It follows that $\pi_1(B) \cong H_n$. The action of H_n on $\widetilde{C}_n$ lifts to an action on $\widetilde{E}_0 = \widetilde{C}_n \times S^4$ which acts trivially on the S^4 factor. It is not hard to see that $\widetilde{E}_n$ can be constructed in such a way that the action of H_n extends to it. Now let $E = \widetilde{E}_n / H_n$. This is a family $\pi : E \to B$ over B with fibres diffeomorphic to $n\mathbb{CP}^2$. The monodromy of the local system $R^2\pi_*\mathbb{Z}$ is easily seen to yield an isomorphism $\rho : \pi_1(B) \to \operatorname{Aut}(L)$. As explained above, this yields a splitting $\Gamma(X) \to M(X)$. □

5 Nonsplit extensions

Let X be a compact, simply connected, smooth 4-manifold with intersection form $L = H^2(X;\mathbb{Z})$. Let Σ be a compact surface (orientable or nonorientable). Suppose that $\rho : \pi_1(\Sigma) \to \Gamma(X)$ is a homomorphism. Letting $\Gamma(X)$ act on $L_{\mathbb{R}} = \mathbb{R} \otimes_{\mathbb{Z}} L$, we obtain a flat vector bundle $H_\rho \to \Sigma$ which has a covariantly constant bilinear form of signature $(b_+(X), b_-(X))$. Let H_ρ^+ denote a maximal positive definite subbundle of H_ρ. The choice of subbundle H_ρ^+ is not unique, but all such subbundles are isomorphic. In particular the Stiefel–Whitney classes $w_j(H_\rho^+) \in H^j(\Sigma;\mathbb{Z}_2)$ depend only on ρ.

Theorem 5.1 *Let X be a compact, simply connected, smooth 4-manifold with $b_+(X) = 2$ and let $\rho : \pi_1(\Sigma) \to \Gamma(X)$ be a homomorphism. Suppose that $w_2(H_\rho^+) \neq 0$ and suppose there exists a characteristic $c \in L$ which is ρ-invariant and satisfies $c^2 > \sigma(X)$. Then ρ does not lift to a homomorphism $\widetilde{\rho} : \Gamma(X) \to M(X)$.*

Proof Consider first the case that Σ is orientable of genus g. Recall that $\pi_1(\Sigma)$ admits a presentation

$$\pi_1(\Sigma) = \langle a_1, b_1, \dots, a_g, b_g \mid [a_1, b_1] \cdots [a_g, b_g] \rangle.$$

Suppose that ρ admits a lift $\widetilde{\rho} : \pi_1(\Sigma) \to M(X)$. Let α_j be a diffeomorphism of X whose isotopy class is $\widetilde{\rho}(a_j)$ and let β_j be a diffeomorphism of X whose isotopy class is $\widetilde{\rho}(b_j)$. Then $[\alpha_1, \beta_1] \cdots [\alpha_g, \beta_g]$ is isotopic to the identity. The surface Σ can be constructed from a wedge of $2g$ circles by attaching a 2-cell whose attaching map represents $[a_1, b_1] \cdots [a_g, b_g]$ in $\pi_1\big(\bigvee_{i=1}^{2g} S^1\big)$. We will construct a smooth family $\pi : E \to \Sigma$ whose fibres are diffeomorphic to X as follows. Over the 1-skeleton $\bigvee_{i=1}^{2g} S^1$, we take the wedge sum of mapping cylinders associated to the diffeomorphisms $\alpha_1, \beta_1, \dots, \alpha_g, \beta_g$. A choice of isotopy from $[\alpha_1, \beta_1] \cdots [\alpha_g, \beta_g]$ to the identity allows us to extend this family over the 2-cell and in this way we obtain the family $\pi : E \to \Sigma$. By construction, the local system $R^2\pi_*\mathbb{Z}$ has monodromy ρ. Now suppose that $w_2(H^+_\rho) \neq 0$ and that there exists a characteristic $c \in L$ which is ρ-invariant and satisfies $c^2 > \sigma(X)$. This contradicts [2, Theorem 1.1], hence ρ does not lift to $M(X)$.

The case that Σ is nonorientable is similar. Recall that $\pi_1(\Sigma)$ admits a presentation

$$\pi_1(\Sigma) = \langle a_1, \dots, a_k \mid a_1^2 \cdots a_k^2 \rangle,$$

where Σ has Euler characteristic $1-k$. If ρ lifts to a homomorphism $\widetilde{\rho} : \pi_1(\Sigma) \to M(X)$, then we choose diffeomorphisms $\alpha_1, \dots, \alpha_k$ where the isotopy class of α_j is $\widetilde{\rho}(a_j)$. Then $\alpha_1^2 \cdots \alpha_k^2$ is isotopic to the identity. A choice of such an isotopy allows us to construct a smooth family $\pi : E \to \Sigma$ with fibres diffeomorphic to X and such that the monodromy of $R^2\pi_*\mathbb{Z}$ is ρ. As before, this contradicts [2, Theorem 1.1], hence ρ does not lift to $M(X)$. □

Remark 5.2 A similar argument was used in [16] to prove the nontriviality of the group $T(X)$ for the manifold $X = 2\mathbb{CP} \# n\overline{\mathbb{CP}^2}$, $n \geq 11$.

Corollary 5.3 *Let* $X = (S^2 \times S^2) \# X'$, *where* $b_+(X') = 1$, $b_-(X') \geq 10$. *Then there does not exist a splitting* $\Gamma(X) \to M(X)$.

Proof Let $L = H^2(X;\mathbb{Z})$ and $L' = H^2(X';\mathbb{Z})$ denote the intersection lattices of X and X' and let $H = H^2(S^2 \times S^2\ \mathbb{Z})$. So $L \cong H \oplus L'$. Since $X = (S^2 \times S^2) \# X'$, we have that $\Gamma(X) = \mathrm{Aut}(H^2(X;\mathbb{Z}))$ by [31]. Let $x, y \in H$ be a basis with $x^2 = y^2 = 0$, $\langle x, y \rangle = 1$. Since $b_+(X') = 1$ and $\sigma(X') < 0$, it follows that X' is not spin and it follows that $L' \cong H' \oplus E_8 \oplus L''$, where H' has basis x', y', $(x')^2 = (y')^2 = 0$, $\langle x', y' \rangle = 1$, E_8 is the negative definite E_8 lattice and L'' is a diagonal lattice with basis $e_1, \dots, e_m$, where $m = b_-(X') - 9$, with $e_i^2 = -1$ for all i, $\langle e_i, e_j \rangle = 0$ for $i \neq j$.

Let $u = x + y$, $v = x' + y'$. Then $u^2 = v^2 = 2$, $\langle u, v \rangle = 0$. Let r_u be the reflection $r_u(x) = x + \langle x, u \rangle u$ and define r_v similarly. Consider the isometry $f(x) = r_u r_v(x)$. Then $f \in \Gamma(X)$ and $f^2 = 1$. Hence we obtain a homomorphism $\rho : \pi_1(\mathbb{RP}^2) \to \Gamma(X)$ which sends the generator of $\pi_1(\mathbb{RP}^2)$ to f. Since f acts as -1 on the maximal positive definite subspace of $H^2(X;\mathbb{R})$ spanned by u and v, we have that $w_2(H^+_\rho) \neq 0$. Let $c = e_1 + \cdots + e_m$. Then c is a characteristic that $c^2 > \sigma(X)$ and $\langle c, u \rangle = \langle c, v \rangle = 0$. Then $r_u(c) = r_v(c) = c$ and hence $f(c) = c$. Then Theorem 5.1 implies that ρ does not lift to $M(X)$. Hence the subgroup $\langle f \rangle \subseteq \Gamma(X)$ does not lift to $M(X)$, in particular, there does not exist a splitting $\Gamma(X) \to M(X)$. □

Remark 5.4 In the above proof $u \in L$ can be realised by an embedded 2-sphere in X, namely the diagonal $S^2 \subset S^2 \times S^2$. By a result of Seidel [27], it follows that r_u can be lifted to an element $\hat{r}_u \in M(X)$ of order 2. Since $\Gamma(X) = \mathrm{Aut}(L)$, it follows that there is a diffeomorphism of X sending u to v. It follows that v can also be realised by an embedded 2-sphere and hence r_v can be lifted to an element $\hat{r}_v \in M(X)$ of order 2. Then $\hat{r}_u \hat{r}_v$ is a lift of f to $M(X)$. If u, v could be represented by *disjoint* embedded 2-spheres, then $\hat{r}_u$, $\hat{r}_v$ commute (since $\hat{r}_u$, $\hat{r}_v$ can be constructed to have disjoint supports) and then $\hat{r}_u \hat{r}_v$ would be an involutive lift of f, contradicting the corollary above. We deduce that u, v can be represented by embedded spheres, but they cannot be represented by disjoint embedded spheres even though $\langle u, v \rangle = 0$.

6 Nielsen realisation

As explained in the introduction, the following result shows that the Nielsen realisation problem fails for $X = X' \# p\mathbb{CP}^2 \# q\overline{\mathbb{CP}^2}$ whenever $p+q \geq 4$.

Theorem 6.1 *Let $X = X' \# p\mathbb{CP}^2 \# q\overline{\mathbb{CP}^2}$ where X' is a compact, smooth, simply connected 4-manifold and $p+q \geq 4$. Then $M(X)$ contains a subgroup isomorphic to $\mathbb{Z}_2^4$ which cannot be lifted to* Diff(X).

Proof To each summand of $\mathbb{CP}^2$ or $\overline{\mathbb{CP}^2}$ in X, there is a corresponding embedded 2-sphere of self-intersection ± 1. Let $E_1, \dots, E_4$ be any four of them. Let $t_1, \dots, t_4 \in M(X)$ be the corresponding Dehn twists around these spheres. Then $t_1, \dots, t_4$ are involutions [27] and they commute since $E_1, \dots, E_4$ are disjoint. Hence the group $G \subseteq M(X)$ generated by $t_1, \dots, t_4$ is isomorphic to $\mathbb{Z}_2^4$. Now suppose that G can be lifted to Diff(X). Hence we can find commuting diffeomorphisms $\sigma_1, \dots, \sigma_4$ such that the isotopy class of σ_i is t_i.

Consider the fixed point set F of σ_1. Since σ_1 acts on $H^2(X;\mathbb{Z})$ as a reflection in a ± 1 sphere, it follows from [8, Proposition 2.4] that F consists of a single copy of $\mathbb{RP}^2$, together with some isolated points and some 2-spheres. Let G_0 be the subgroup of G generated by $\sigma_2, \sigma_3, \sigma_4$. Since $\sigma_2, \sigma_3, \sigma_4$ commute with σ_1, they act on F and in particular must send the copy of $\mathbb{RP}^2$ to itself. Hence G_0 acts on $\mathbb{RP}^2$. We claim that the action is effective. To see this, suppose $f \in G_0$ fixes $\mathbb{RP}^2$ pointwise. Since f is an orientation-preserving involution, it must act on the normal bundle of $\mathbb{RP}^2$ in X as either the identity or multiplication by -1. Hence either f or $\sigma_1 f$ fixes $\mathbb{RP}^2$ pointwise and acts trivially on the normal bundle. For a diffeomorphism of finite order, this can only happen if the diffeomorphism is the identity. Hence f or $\sigma_1 f$ is the identity, but $f \in G_0$, so $f \neq \sigma_1$ and it must be that f is the identity.

A finite group action on $\mathbb{RP}^2$ by diffeomorphisms is conjugate to a subgroup of PO(3) $\cong$ SO(3). Since G_0 is abelian, its action on the standard representation of SO(3) can be simultaneously diagonalised, so G_0 is isomorphic to a subgroup of $\{\mathrm{diag}(\epsilon_1, \epsilon_2, \epsilon_3) \in \mathrm{SO}(3)\} \cong \mathbb{Z}_2^2$, which is impossible since $|G_0| = 8$. So G does not lift to Diff(X). □

7 Boundary Dehn twists

Let $X^{(n)}$ be obtained from X by removing n disjoint open balls. So $X^{(n)}$ is a compact 4-manifold with boundary consisting of n copies of S^3. Let Diff$(X^{(n)}, \partial X^{(n)})$ denote the group of diffeomorphisms of $X^{(n)}$ which are the identity in a neighbourhood of the boundary. Let $M_n(X) = \pi_0(\mathrm{Diff}(X^{(n)}, \partial X^{(n)})$ denote the group of components of Diff$(X^{(n)}, \partial X^{(n)})$. It is known that the map $M_n(X) \to M(X)$ is surjective and that the kernel is generated by Dehn twists on the boundary components [11]. More precisely, if $S^3 \to X^{(n)}$ is a boundary component, then $X^{(n)}$ has a tubular neighbourhood $[0,1] \times S^3 \to X$. The Dehn twist on this boundary component is defined by taking a nontrivial loop $\alpha_t : [0,1] \to \mathrm{SO}(4)$ and defining $\phi : [0,1] \times S^3 \to [0,1] \times S^3$ by $\phi(t,x) = (t, \alpha_t(x))$, where SO(4) acts on S^3 in the standard way. We assume that α_t is smooth and equals the identity in a neighbourhood of $\{0,1\}$, hence ϕ can be extended to an element of Diff$(X^{(n)}, \partial X^{(n)})$ by taking it to be the identity outside of the tubular neighbourhood.

Let $K_n(X)$ denote the kernel of $M_n(X) \to M(X)$, so we have an short exact sequence

$$1 \to K_n(X) \to M_n(X) \to M(X) \to 1.$$

Furthermore, we have a surjection $\mathbb{Z}_2^n \to K_n(X)$ given by Dehn twists on the boundary components [11, Proposition 3.1].

Proposition 7.1 *Let X be a compact, smooth, simply connected* 4*-manifold.*

(1) *If X is spin, then $K_n(X)$ is either $\mathbb{Z}_2^n$ or $\mathbb{Z}_2^n/\Delta\mathbb{Z}_2$, for all n, where $\Delta\mathbb{Z}_2$ is the diagonal copy of $\mathbb{Z}_2$.*

(2) *If X is not spin, then $K_n(X) = 0$ for all n, hence $M_n(X) \cong M(X)$.*

Proof Part (1) is given by [11, Corollary 2.5] and part (2) by [22, Corollary A.5]. □

In light of Proposition 7.1, boundary Dehn twists are only interesting when X is spin. In this case, we either have $K_n(X) \cong \mathbb{Z}_2^n$ or $K_n(X) \cong \mathbb{Z}_2^n/\Delta\mathbb{Z}_2$. Which of these two cases occurs is completely determined by the $n = 1$ case. We consider this case in more detail. There is a Serre fibration

$$\mathrm{Diff}(X^{(1)}, \partial X^{(1)}) \to \mathrm{Diff}(X) \to \mathrm{Emb}(D^4, X), \tag{7-1}$$

where $\mathrm{Emb}(D^4, X)$ is the space of embeddings of a disc in X which can be extended to a diffeomorphism. Furthermore, there is a homotopy equivalence $\mathrm{Emb}(D^4, X) \cong F(X)$, where $F(X)$ is the oriented frame bundle of X [11]. Since X is simply connected and spin, $\pi_1(F(X)) \cong \mathbb{Z}_2$. Then the fibration (7-1) induces an exact sequence

$$\pi_1(\mathrm{Diff}(X)) \xrightarrow{\phi} \mathbb{Z}_2 \to M_1(X) \to M(X) \to 1.$$

In the absence of a metric we can define the spin bundle of X to be the universal cover $\widetilde{F}(X) \to F(X)$ of $F(X)$. Since $\pi_1(F(X)) \cong \mathbb{Z}_2$, we have $\widetilde{F}(X) \to F(X)$ is a double cover. Since $\mathrm{Emb}(D^4, X) \cong F(X)$, it follows that ϕ is the map that measures whether or not a loop of diffeomorphisms of X lifts to a loop in the spin bundle of X. This leads to an alternative description of the group $M_1(X)$ when X is spin. Let $\mathrm{SpinDiff}(X)$ be the group whose elements consist of a diffeomorphism $f \in \mathrm{Diff}(X)$ and a choice of lift of $f_* : F(X) \to F(X)$ to $\widetilde{F}(X)$. We have a short exact sequence $1 \to \mathbb{Z}_2 \to \mathrm{SpinDiff}(X) \to \mathrm{Diff}(X) \to 1$ and the connecting homomorphism $\pi_1(\mathrm{Diff}(X)) \to \mathbb{Z}_2$ is precisely ϕ. The map $\mathrm{Diff}(X^{(1)}, \partial X^{(1)}) \to \mathrm{Diff}(X)$ admits a lift $\mathrm{Diff}(X^{(1)}, \partial X^{(1)}) \to \mathrm{SpinDiff}(X)$ by taking the unique lift which is the identity over $\partial X^{(1)}$. We then have a commutative diagram

$$\begin{tikzcd}
\pi_1(\mathrm{Diff}(X)) \arrow[r, "\phi"] \arrow[dr, "\phi"'] & \mathbb{Z}_2 \arrow[r] \arrow[d] & M_1(X) \arrow[r] \arrow[d] & M(X) \\
& \mathbb{Z}_2 \arrow[r] & \pi_0(\mathrm{SpinDiff}(X)) \arrow[ur] &
\end{tikzcd}$$

from which it follows that $M_1(X) \to \pi_0(\mathrm{SpinDiff}(X))$ is an isomorphism. If ϕ is nontrivial, then $K_1(X) = 0$ and $M_1(X) \to M(X)$ is an isomorphism. This happens for $S^2 \times S^2$, as seen by taking a loop of diffeomorphisms given by a circle action which rotates one of the spheres. Similarly, ϕ is nontrivial for $X = S^4$ or for a connected sum of copies of $S^2 \times S^2$. If ϕ is trivial, then $K_1(X) \cong \mathbb{Z}_2$ and $M_1(X) \to M(X)$ is an extension of $M(X)$ by $\mathbb{Z}_2$, hence corresponds to a class $\xi_X \in H^2(M(X); \mathbb{Z}_2)$. It is natural to ask what this class is and in particular, whether or not it is trivial. First, we need some examples of spin 4-manifolds where $\phi = 0$.

Theorem 7.2 *Let X be a compact, smooth, simply connected 4-manifold. If X is homeomorphic to $K3$ then $\phi = 0$. Similarly, if $X = X' \# (S^2 \times S^2)$, where X' is homeomorphic to $K3$, then $\phi = 0$.*

Proof In [4], it is proven that if $E \to S^2$ is a smooth family of $K3$ surfaces over S^2, then $w_2(TE) = 0$. As explained in [18], this implies that the homomorphism ϕ is zero. The same argument works for any X that is homeomorphic to $K3$, since by [20], the Seiberg–Witten invariant of the spin structure of X is odd.

Next, suppose $X = X' \# (S^2 \times S^2)$, where X' is homeomorphic to $K3$. Suppose that ϕ is nonzero. This means that the boundary Dehn twist $\tau \in M_1(X)$ is trivial. But this would imply that the Dehn twist on the neck of $K3 \# X'$ becomes trivial upon connected sum with $S^2 \times S^2$. However this contradicts [19] (in [19] the theorem is stated only for $X' = K3$, but the exact same proof works for any smooth 4-manifold homeomorphic to $K3$). □

Recall that an involution f on a simply connected spin 4-manifold X is called even or odd according to whether or not f lifts to an involution on the spin bundle of X.

Proposition 7.3 *Suppose X is spin and $\phi = 0$, so that the extension class $\xi_X \in H^2(M(X); \mathbb{Z}_2)$ is defined. Suppose that f is an odd involution. Then $\xi_X(f) \neq 0$. In particular, the extension $\mathbb{Z}_2 \to M_1(X) \to M(X)$ is nontrivial.*

Proof As explained above, the extension $1 \to \mathbb{Z}_2 \to M_1(X) \to M(X) \to 1$ is isomorphic to the extension $1 \to \mathbb{Z}_2 \to \pi_0(\mathrm{SpinDiff}(X)) \to M(X) \to 1$. But f defines a class $[f] \in M(X)$ such that $[f]^2 = 1$. But any lift of f to the spin bundle is not an involution. So there is no splitting $M(X) \to M_1(X)$ and more precisely, $\xi_X(f) \neq 0$. □

Corollary 7.4 *If $X = K3$ or $K3 \# (S^2 \times S^2)$, then $\xi_X \in H^2(M(X); \mathbb{Z}_2)$ is nontrivial.*

Proof This is immediate from Theorem 7.2 and Proposition 7.3, since both $K3$ and $K3 \# (S^2 \times S^2)$ admit odd involutions. □

In what follows we will completely determine the class $\xi_X \in H^2(M(X); \mathbb{Z}_2)$ when X is homeomorphic to $K3$.

Proposition 7.5 *Let $\pi : E \to B$ be a smooth fibre bundle, where B is a compact surface and the fibres of E are diffeomorphic to a compact, simply connected, smooth spin 4-manifold X. Then:*

1. *There exists a spin^c-structure $\mathfrak{s}_{E/B}$ on the vertical tangent bundle $TE/B = \mathrm{Ker}(\pi_*)$ whose restriction to each fibre is spin.*
2. *Let $\mathrm{ind}(D) \in K^0(B)$ denote the families index of the Dirac operator D with respect to the spin^c-structure $\mathfrak{s}_{E/B}$. Then*

$$c_1(\mathrm{ind}(D)) = \tfrac{1}{16}\sigma(X) w_2(TE/B) \pmod 2.$$

Proof (1) follows from [2, Proposition 2.1]. The Dirac operator D for the spin^c-structure $\mathfrak{s}_{E/B}$ defines a family of elliptic operators parametrised by B and $\mathrm{ind}(D)$ is the families index. Then $c_1(\mathrm{ind}(D))$ equals $c_1(\mathcal{L})$, where $\mathcal{L} = \det(\mathrm{ind}(D))$ is the determinant line bundle of D. Suppose that the family E is determined by transition function ψ_{ij} valued in $\mathrm{Diff}(X)$. Let $\widetilde{\psi}_{ij}$ be lifts of ψ_{ij} to $\mathrm{SpinDiff}(X)$. Then $\widetilde{\psi}_{ij}\widetilde{\psi}_{jk}\widetilde{\psi}_{ki} = g_{ijk}$, where g_{ijk} is a $\mathbb{Z}_2$-valued cocycle, defining a class $[g_{ijk}] \in H^2(B;\mathbb{Z}_2)$. Clearly $w_2(TE/B) = [g_{ijk}]$. Observe that $c(\mathfrak{s}_{E/B}) \in H^2(B;\mathbb{Z})$ is a lift of $[g_{ijk}]$ to integer coefficients. Therefore we can represent $c(\mathfrak{s}_{E/B})$ as an integer-valued 2-cocycle c_{ijk} such that $c_{ijk} = g_{ijk} \pmod 2$. Choose real-valued smooth functions u_{ij} such that $c_{ijk} = u_{ij} + u_{jk} + u_{ki}$. Set $f_{ij} = e^{2\pi i u_{ij}}$. Then f_{ij} define transition functions for a complex line bundle whose first Chern class is $[c_{ijk}]$. Note that $f_{ij} = h_{ij}^2$, where $h_{ij} = e^{\pi i u_{ij}}$. Then $h_{ij}h_{jk}h_{ki} = (-1)^{g_{ijk}}$. Define $\mathrm{Spin^cDiff}(X) = U(1) \times_{\mathbb{Z}_2} \mathrm{SpinDiff}(X)$. Then $\varphi_{ij} = h_{ij}\widetilde{\psi}_{ij}$ is a 2-cocycle valued in $\mathrm{Spin^cDiff}(X)$.

Consider now the transition functions for the determinant line bundle $\mathcal{L}$. Since $\mathfrak{s}_{E/B}$ restricts to a spin structure on the fibres, the spinor bundles have a quaternionic structure on each fibre. It follows that $\widetilde{\psi}_{ij}$ induces a trivial action on the determinant line. However, the $U(1)$-factor h_{ij} in $\varphi_{ij} = h_{ij}\widetilde{\psi}_{ij}$ acts on the spinor bundles as scalar multiplication which then acts on the determinant line by h_{ij}^d, where d is the virtual rank of $\mathrm{ind}(D)$, which is $d = -\sigma(X)/8$. Therefore $\mathcal{L}$ has transition functions $h_{ij}^{-\sigma(X)/8} = f_{ij}^{-\sigma(X)/16}$. Recalling that f_{ij} are transition functions for a line bundle with Chern class $c(\mathfrak{s}_{E/B})$, it follows that

$$c_1(\mathrm{ind}(D)) = c_1(\mathcal{L}) = -\tfrac{1}{16}\sigma(X)c(\mathfrak{s}_{E/B}) = \tfrac{1}{16}\sigma(X)w_2(TE/B) \pmod 2. \qquad \square$$

Proposition 7.6 *Let $\pi : E \to B$ be a smooth fibre bundle, where the fibres of E are homeomorphic to $K3$. Then $w_2(TE/B) = w_2(H^+)$, where $H^+ \to B$ denotes the bundle whose fibre over b is a maximal positive definite subspace of $H^2(E_b;\mathbb{R})$.*

Proof Since $H^2(B;\mathbb{Z}_2)$ is detected by maps of compact surfaces into B, it suffices to prove the result when B is a compact surface. Then by Proposition 7.5, $w_2(TE/B) = c_1(\mathrm{ind}(D)) \pmod 2$. On the other hand, since the fibres are homeomorphic to $K3$, their Seiberg–Witten with respect to the spin structure is odd [20]. Then by [4, Corollary 1.3], $c_1(\mathrm{ind}(D)) = w_2(H^+)$. □

Let L be a lattice and $A = \mathrm{Aut}(L)$ the group of automorphisms. Over the classifying space BA we have the tautological flat bundle $H = EA \times_A L$. Let $H^+ \to BA$ be a maximal positive subbundle. This defines a characteristic class $w_2(H^+) \in H^2(\mathrm{Aut}(L);\mathbb{Z}_2)$.

Theorem 7.7 *Let X be a smooth* 4*-manifold which is homeomorphic to $K3$. Let L_X be the intersection lattice of X. Then the extension class $\xi_X \in H^2(M(X);\mathbb{Z}_2)$ is the pullback of $w_2(H^+) \in H^2(\mathrm{Aut}(L_X);\mathbb{Z}_2)$ under the map $M(X) \to \mathrm{Aut}(L_X)$.*

Proof Let B be a compact surface and consider a map $\iota : B \to BM(X)$. This is equivalent to a homomorphism $\rho : \pi_1(B) \to M(X)$. We claim that ρ is the geometric monodromy of a family $E \to B$. We can take B to be given by attaching a 2-cell to a wedge of k circles. Each circle defines a generator $g_i \in \pi_1(B)$ and the 2-cell defines a relation $r = r(g_1, \dots, g_k)$, which is a word in the g_i. Choose a lift $f_i \in \mathrm{Diff}(X)$ of $\rho(g_i) \in M(X)$. Then we can construct a family E_1 over the 1-skeleton on B

as a wedge of mapping cylinders corresponding to the diffeomorphisms $f_1, \dots, f_k$. Since $g_1, \dots, g_k$ satisfy r, it follows that $r(f_1, \dots, f_k)$ is isotopic to the identity. Choosing such an isotopy, we can extend E_1 over the 2-cell, giving the desired family $E \to B$. As explained in [4, Remark 4.20], we can assume that the family $E \to B$ is smooth. Now consider the obstruction to lifting the structure group of E to SpinDiff(X). This is easily seen to coincide with the obstruction to lifting $\rho : \pi_1(B) \to M(X)$ to $M_1(X)$, which is $\iota^*(\xi_X) \in H^2(B; \mathbb{Z}_2)$. On the other hand, the obstruction to lifting the structure group of E to SpinDiff(X) is $w_2(TE/B)$, which by Proposition 7.6 equals $w_2(H^+)$. Since $H^2(M(X); \mathbb{Z}_2)$ is detected by maps of compact surfaces B into $BM(X)$, the result is proven. □

Acknowledgements

I thank Hokuto Konno for sending me his paper [15] and also for comments on a draft of this paper. I also thank Alexander Kupers for clarifying the finite generation of mapping class groups for simply connected 5-manifolds.

References

[1] **M Arabadji**, **R I Baykur**, *Nielsen realization in dimension four and projective twists*, Adv. Math. 463 (2025) art. id. 110112 MR

[2] **D Baraglia**, *Obstructions to smooth group actions on* 4*-manifolds from families Seiberg–Witten theory*, Adv. Math. 354 (2019) art. id. 106730 MR

[3] **D Baraglia**, **H Konno**, *A gluing formula for families Seiberg–Witten invariants*, Geom. Topol. 24:3 (2020) 1381–1456 MR

[4] **D Baraglia**, **H Konno**, *On the Bauer–Furuta and Seiberg–Witten invariants of families of* 4*-manifolds*, J. Topol. 15:2 (2022) 505–586 MR

[5] **D Baraglia**, **H Konno**, *A note on the Nielsen realization problem for K*3 *surfaces*, Proc. Amer. Math. Soc. 151:9 (2023) 4079–4087 MR

[6] **M Bustamante**, **M Krannich**, **A Kupers**, *Finiteness properties of automorphism spaces of manifolds with finite fundamental group*, Math. Ann. 388:4 (2024) 3321–3371 MR

[7] **M Dehn**, *Die Gruppe der Abbildungsklassen*: *Das arithmetische Feld auf Flächen*, Acta Math. 69:1 (1938) 135–206 MR

[8] **A L Edmonds**, *Aspects of group actions on four-manifolds*, Topology Appl. 31:2 (1989) 109–124 MR

[9] **B Farb**, **E Looijenga**, *The Nielsen realization problem for K3 surfaces*, J. Differential Geom. 127:2 (2024) 505–549 MR

[10] **M H Freedman**, *The topology of four-dimensional manifolds*, J. Differential Geometry 17:3 (1982) 357–453 MR

[11] **J Giansiracusa**, *The stable mapping class group of simply connected* 4*-manifolds*, J. Reine Angew. Math. 617 (2008) 215–235 MR

[12] **R E Gompf**, **A I Stipsicz**, 4*-manifolds and Kirby calculus*, Graduate Studies in Mathematics 20, Amer. Math. Soc., Providence, RI (1999) MR

[13] **S Hong**, **D McCullough**, *Mapping class groups of* 3*-manifolds, then and now*, from "Geometry and topology down under" (C D Hodgson, W H Jaco, M G Scharlemann, S Tillmann, editors), Contemp. Math. 597, Amer. Math. Soc., Providence, RI (2013) 53–63 MR

[14] **H Konno**, *Dehn twists and the Nielsen realization problem for spin* 4*-manifolds*, Algebr. Geom. Topol. 24:3 (2024) 1739–1753 MR

[15] **H Konno**, *The homology of moduli spaces of* 4*-manifolds may be infinitely generated*, Forum Math. Pi 12 (2024) art. id. e25 MR

[16] **H Konno**, **A Mallick**, **M Taniguchi**, *Exotic Dehn twists on* 4*-manifolds* (2023) arXiv 2306.08607

[17] **H Konno**, **J Miyazawa**, **M Taniguchi**, *Involutions, knots, and Floer K-theory*, Compos. Math. 161:11 (2025) 2852–2910 MR

[18] **P B Kronheimer**, **T S Mrowka**, *The Dehn twist on a sum of two $K3$ surfaces*, Math. Res. Lett. 27:6 (2020) 1767–1783 MR

[19] **J Lin**, *Isotopy of the Dehn twist on $K3 \# K3$ after a single stabilization*, Geom. Topol. 27:5 (2023) 1987–2012 MR

[20] **J W Morgan**, **Z Szabó**, *Homotopy $K3$ surfaces and mod 2 Seiberg–Witten invariants*, Math. Res. Lett. 4:1 (1997) 17–21 MR

[21] **L I Nicolaescu**, *Notes on Seiberg–Witten theory*, Graduate Studies in Mathematics 28, Amer. Math. Soc., Providence, RI (2000) MR

[22] **P Orson**, **M Powell**, *Mapping class groups of simply connected* 4*-manifolds with boundary*, J. Differential Geom. 131:1 (2025) 199–275 MR

[23] **F Quinn**, *Isotopy of* 4*-manifolds*, J. Differential Geom. 24:3 (1986) 343–372 MR

[24] **D Ruberman**, *An obstruction to smooth isotopy in dimension* 4, Math. Res. Lett. 5:6 (1998) 743–758 MR

[25] **D Ruberman**, *A polynomial invariant of diffeomorphisms of* 4*-manifolds*, from "Proceedings of the Kirbyfest" (Berkeley, CA, 1998) (J Hass, M Scharlemann, editors), Geom. Topol. Monogr. 2, Geom. Topol., Coventry (1999) 473–488 MR

[26] **D Ruberman**, *Positive scalar curvature, diffeomorphisms and the Seiberg–Witten invariants*, Geom. Topol. 5 (2001) 895–924 MR

[27] **P Seidel**, *Lectures on four-dimensional Dehn twists*, from "Symplectic 4-manifolds and algebraic surfaces" (F Catanese, G Tian, editors), Lecture Notes in Math. 1938, Springer (2008) 231–267 MR

[28] **A Seress**, *Permutation group algorithms*, Cambridge Tracts in Mathematics 152, Cambridge Univ. Press (2003) MR

[29] **C L Siegel**, *Einheiten quadratischer Formen*, Abh. Math. Sem. Hansischen Univ. 13 (1940) 209–239 MR

[30] **G Triantafillou**, *The arithmeticity of groups of automorphisms of spaces*, from "Tel Aviv Topology Conference: Rothenberg Festschrift" (Tel Aviv, 1998) (M Farber, W Lück, S Weinberger, editors), Contemp. Math. 231, Amer. Math. Soc., Providence, RI (1999) 283–306 MR

[31] **C T C Wall**, *Diffeomorphisms of* 4*-manifolds*, J. London Math. Soc. 39 (1964) 131–140 MR

DAVID BARAGLIA david.baraglia@adelaide.edu.au
School of Mathematical Sciences, The University of Adelaide, Adelaide, Australia

Received: November 6, 2023 Revised: March 16, 2025

Algebraic & Geometric Topology **26**:5 (2026) 1655–1680
DOI: 10.2140/agt.2026.26.1655
Published: May 29, 2026

Negative-definite spin filling and branched double covers

SOHEIL AZARPENDAR

We investigate the negative-definite spin fillings of branched double covers of alternating knots. We derive obstructions to the existence of such fillings and characterize special alternating knots using these results.

1 Introduction

Given a nonsplit link K, let $\Sigma(S^3, K)$ denote the branched double cover of S^3 along K. A filling of $\Sigma(S^3, K)$ is a 4-manifold X with $\partial X = \Sigma(S^3, K)$. A common approach to constructing fillings of $\Sigma(S^3, K)$ involves taking a spanning surface F of K and forming the branched double cover of D^4 over F^+, where F^+ is obtained by pushing the interior of F inside D^4. We use the term spanning filling to distinguish fillings that can be constructed through this method. One of the most important facts about spanning fillings of branched double covers of links is due to Gordon and Litherland [10]. They proved that the intersection form of $\Sigma(D^4, F^+)$ is equal to the Goeritz form of F. We call a spanning surface F positive- or negative-definite if the Goeritz form of F (or equivalently $\Sigma(D^4, F^+)$) is positive- or negative-definite, respectively.

A standard choice for a spanning surface of K comes from a checkerboard coloring of the regions in a knot diagram. Considering all of the white (resp. black) regions in S^2 and adding twisted bands between them around each crossing will result in a spanning surface of K. We refer to this surface as the white (resp. black) Tait surface and denote it by F_W (resp. F_B). Note that generally Tait surfaces are not invariants of the knot and depend on the choice of diagram.

A link diagram is alternating if its crossings alternate over and under around each link component, and a link is alternating if it admits an alternating diagram. One can prove that a diagram is alternating if and only if the associated Tait surfaces are definite and of opposite signs (see Proposition 4.1 of [14]). Greene [14] proved that the existence of definite spanning surfaces gives a topological characterization of alternating links as follows:

Theorem 1.1 [14, Theorem 1.1] *If F_P and F_N are positive- and negative-definite spanning surfaces for a nonsplit link K in S^3, then K is an alternating link, and it has an alternating diagram whose Tait surfaces are isotopic to F_P and F_N.*

In the rest of this paper, we only focus on nonsplit alternating links. Whenever we consider an alternating link K, we are in fact working with an arbitrary reduced alternating diagram of K. For convenience we do not differentiate between the link and its diagram. Furthermore, to fix a standard checkerboard coloring of an alternating diagram, we assume that the black Tait surface is a negative-definite spanning surface.

MSC2020: 57K10, 57K18, 57K40, 57K41.

We call an oriented alternating diagram *special* if the associated black Tait surface is orientable. Seifert's algorithm outputs this surface when applied to such a diagram. An oriented alternating link is special if it has a special alternating diagram. One can see that the orientability of F_B is equivalent to evenness of its Goeritz form (see Section 2 for more details). In this paper, whenever we consider a special alternating link K, we are working with a special alternating diagram of K.

Building on the preceding discussion, for an alternating link K, the 4-manifolds $\Sigma(D^4, F_W^+)$ and $\Sigma(D^4, F_B^+)$ are definite fillings of $\Sigma(S^3, K)$, which we refer to as the white and black Tait fillings, respectively. Furthermore, if K is a special alternating link, then $\Sigma(D^4, F_B^+)$ is also a spin filling of the branched double cover.

The existence and properties of definite fillings have been the subject of extensive research, including works by Ozsváth and Szabó [19], Scaduto [22], Choe and Park [3], Golla and Scaduto [8], and Aceto, McCoy, and Park [1]. These works also establish bounds on the Betti numbers using tools from Heegaard Floer and Seiberg–Witten theories. In this paper, we discuss how definite spin fillings can detect special alternating links among all alternating links. This is described in Theorem 1.2.

Before we state Theorem 1.2, note that, in the rest of the paper, we assume that all of the diagrams are decorated, i.e., they have a marked arc between two crossings. We refer to the two regions separated by the marked arc as marked (adjacent) regions. We use m to represent the number of unmarked white regions in a reduced alternating diagram.

Theorem 1.2 *Let K be a nonsplit alternating link and m be the number of unmarked white regions in a reduced alternating diagram. If X is a simply connected negative-definite spin filling of $\Sigma(S^3, K)$, then*

$$b_2(X) \leq m.$$

Furthermore, equality is achieved if and only if K is special alternating.

The existence of simply connected spin negative-definite fillings of the branched double cover of nonspecial alternating knots is not trivial. Using inequalities from Heegaard Floer and Seiberg–Witten theories, we develop several obstructions to the existence of such fillings. The main ones are in the form of Theorems 1.3 and 1.4. First, we need to explain some notation.

Consider a reduced alternating diagram of link K. The white (resp. black) Tait graph, denoted by W (resp. B), is constructed by considering a vertex for each white (resp. black) region in a diagram and drawing an edge between two regions if and only if they have a common crossing on their boundary. Note that W and B are planar duals. Let $\widetilde{W}$ be the reduced white Tait graph; i.e., the white Tait graph W with the vertex associated with the marked region deleted. Both reduced and unreduced Tait graphs are generally undirected multigraphs with no self-loops or degree-one vertices. This follows from the assumption that the diagram is reduced.

In this paper, we use V_G to denote the vertex set of a graph G and $E_G(\cdot,\cdot)$ to denote the set of edges between two disjoint subsets of V_G. A subgraph C of $\widetilde{W}$ is called *characteristic* if it satisfies

$$e_W(v, C) \equiv \deg_W(v) \mod 2 \qquad \text{for all } v \in V_{\widetilde{W}},$$

where $e_W(v, C)$ is defined by the formula

$$e_W(v,C) = \begin{cases} |E_W(\{v\}, V_C)| + \deg_W(v) & \text{if } v \in C, \\ |E_W(\{v\}, V_C)| & \text{if } v \notin C. \end{cases} \tag{1-1}$$

Using W, we will build a surgery diagram for the black Tait filling of $\Sigma(S^3, K)$ in Section 2. One can see that $e_W(\cdot,\cdot)$ is a reformulation of the intersection form of this filling.

Let $\mathcal{C}_{\widetilde{W}}$ denote the set of characteristic subgraphs of $\widetilde{W}$. We will see that these subgraphs classify spin structures on $\Sigma(S^3, K)$ (see Theorem 4.2). The empty subgraph is characteristic if and only if all of the vertices of W have even degrees. In this case, the dual plane graph B will be bipartite. The 2-coloring of B induces an orientation on F_B which means K is special alternating. As a result, if K is a nonspecial alternating knot, a characteristic subgraph can not be empty.

Theorem 1.3 *Let K be a nonspecial alternating knot. If*

$$\min_{C \in \mathcal{C}_{\widetilde{W}}} |E_W(V_C, V_W \setminus V_C)| \geq |V_W| - 1,$$

then $\Sigma(S^3, K)$ does not have a simply connected negative-definite spin filling.

Theorem 1.4 *Let K be a nonspecial alternating link. If*

$$\min_{C \in \mathcal{C}_{\widetilde{W}}} |E_W(V_C, V_W \setminus V_C)| \geq 9(|V_W| - 1),$$

then $\Sigma(S^3, K)$ does not have a simply connected negative-definite spin filling.

While investigating Theorem 1.4, we identify obstructions preventing 4-manifolds from having chainmail Kirby diagrams (see Section 4). This leads to Corollary 4.11 which states that any closed 4-manifold with a chainmail Kirby diagram is either spin or has a characteristic embedded sphere.

Theorems 1.3 and 1.4 turn out to be generalized versions of a known obstruction of negative-definite spin fillings given by the Neumann–Siebenmann invariant. See Remark 4.10.

We should point out that both of these theorems result in a *twisting phenomenon*. Consider a coherent twist region R in an alternating diagram of the link K. By coherent, we mean a twist region corresponding to a family of parallel edges between two vertices in the white Tait graph (see Figure 1). Let $K_{(R,i)}$ be the link obtained by adding i full twists to the twist region R. We call this operation *enlarging R in K*. In terms of the white Tait graph, the enlarging operation can be defined as adding $2i$ parallel edges.

The enlarging operation can increase the left-hand side of the inequalities of Theorems 1.3 and 1.4 while it does not change the right-hand side, i.e., $|V_W|$. Hence if you enlarge twist regions of the knot diagram, you will end up with links whose branched double cover does not have a negative-definite spin filling.

The main condition of Theorem 1.3 can be seen as an analogue of Elkies's condition [6, Theorem 1] (i.e., nonexistence of short characteristic vectors) for the *Goeritz lattice*. This is the content of Corollary 1.5. See Remark 3.2 for a detailed explanation.

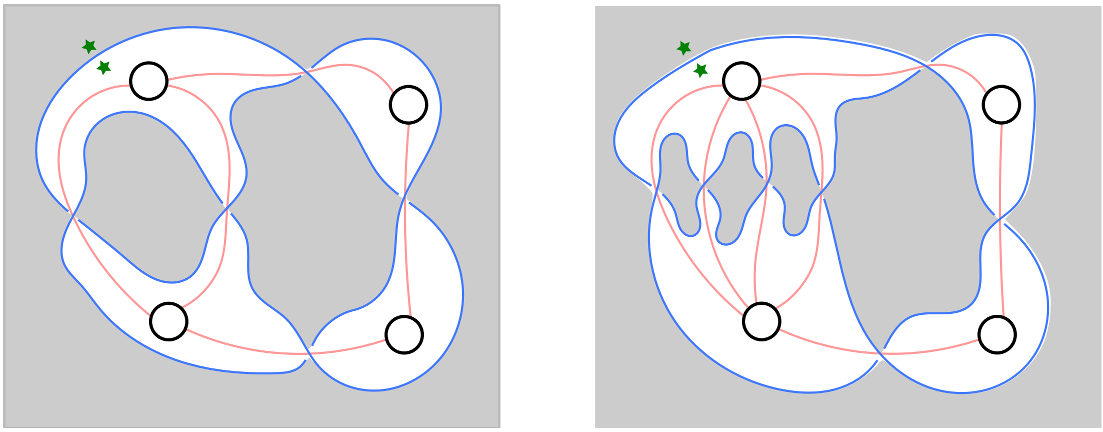


Figure 1: Enlarging a twist region.

Corollary 1.5 *Let K be a nonspecial alternating knot. Also let Λ be the lattice defined by the Goeritz form of F_B, and let* $\mathrm{Char}(\Lambda)$ *be the set of characteristic vectors of Λ. If*

$$\min_{w \in \mathrm{Char}(\Lambda)} |\langle w, w\rangle| \geq |V_W| - 1,$$

then $\Sigma(S^3, K)$ does not have a simply connected negative-definite spin filling.

Note that the only difference between Corollary 1.5 and the main result of Elkies is that the Goeritz lattice is not unimodular.

Generally, there are nonspecial alternating knots with branched double covers that admit simply connected negative-definite spin fillings. Examples of such knots are provided in Section 5. However, it turns out that by further restricting the search to the class of plumbed 4-manifolds, one can prove a nonexistence result explained in Theorem 1.7. First, we need to explain some notation.

Theorem 1.7 is about algebraic (or arborescent) links which are defined as follows.

Definition 1.6 The following algorithm associates a link to a planar weighted tree. Let T be a planar tree with weight $w(a_i) \in \mathbb{Z}$ associated with any vertex $a_i \in V_T$. For each $a_i \in V_T$ consider a twisted band F_i with $w(a_i)$ half twists. If two vertices are connected by an edge in T, plumb the two corresponding bands together. Let F be the resulting surface and consider the link ∂F. If a link can be constructed using this algorithm, we call it *algebraic* (*or arborescent*).

Following the work of Siebenmann [23], it is known that these are the only links whose branched double covers admit plumbed filling. We call an algebraic link *excessive* if it is constructed from a plumbing tree T with weight function w, satisfying

$$w(a_i) \leq \min\{-2, -\deg_T(a_i)\} \quad \text{for all } a_i \in V_T.$$

Excessiveness is a technical condition defined by Murasugi [16]. We use it to ensure that the algebraic link is alternating and relate the tree T to the Tait graph (see Lemma 5.4).

Theorem 1.7 *Let K be an excessive algebraic alternating knot. Then $\Sigma(S^3, K)$ admits a simply connected spin negative-definite plumbed filling if and only if K is special.*

This paper is organized as follows. In Section 2, we set our basic notation and recall some of the theorems from the literature. This section will include the construction of a Kirby diagram for the black Tait filling, some facts about the Heegaard–Floer homology of $\Sigma(S^3, K)$, and some inequalities about fillings of rational homology spheres and closed spin 4-manifolds. In Section 3, we discuss proofs of Theorems 1.2 and 1.3 which come from a formula for the correction term of the branched double cover. In Section 4, we discuss an algorithm of Kaplan which helps us to construct spin fillings and combine it with Furuta's $\frac{10}{8}$ theorem to prove Theorem 1.4. In Section 5, we will discuss Neumann's plumbing calculus and use it to prove Theorem 1.7.

2 Background and notation

Let $K \subset S^3$ be an alternating knot and let D be an alternating diagram of K in the plane. Since the diagram is alternating, one can construct a checkerboard coloring of the diagram such that all the crossings have $\mu = -1$, using the notation of Gordon and Litherland [10]. The coloring will look like Figure 2 around each crossing.

In this setting, one can define *white and black Tait graphs and Tait surfaces*. Tait graphs are constructed by considering regions with the same color as vertices and drawing an edge between two regions if and only if they have a common crossing on their boundary. We use the notation W and B for the graphs and F_W and F_B for the spanning surfaces. In this paper, we assume that diagrams are always decorated, i.e, they have a marked arc between two crossings. We refer to the two regions separated by the marked arc as marked (adjacent) regions. We refer to the graphs resulting from deleting the vertices associated with the marked regions as the reduced Tait graphs and denote them by $\widetilde{B}$ and $\widetilde{W}$. To define the Tait surfaces, let us first recall the definition of median construction.

Definition 2.1 Given a plane graph G, consider a thickening of $G \subset \mathbb{R}^2$. This thickening has a disc centered around each vertex of G, together with a band along each edge. Apply a right-handed (resp. left-handed) half-twist on each of the bands, which gives us a surface (in S^3). This construction is called the *positive* (*resp. negative*) median construction.

Now we define the white (resp. black) Tait surface, denoted by F_W (resp. F_B), as the result of negative (resp. positive) median construction on the Tait graph W (resp. B).

Figure 2: Standard coloring of a crossing in an alternating link.

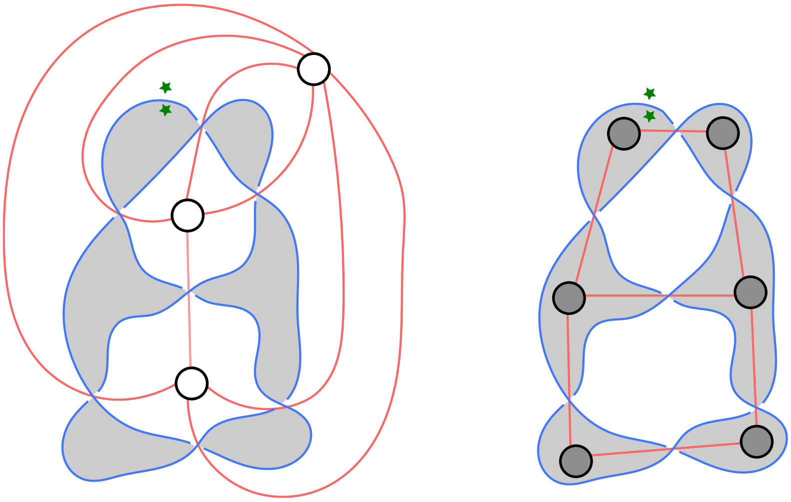


Figure 3: A special alternating knot and its Tait graphs.

An alternating knot is called *special* if the black Tait surface is orientable, i.e., a Seifert surface. This is equivalent to the black Tait graph being bipartite. Since black and white Tait graphs are dual planar graphs, this definition is also equivalent to the white Tait graph having no vertex with odd degree. An example of a special alternating knot and its Tait graphs can be seen in Figure 3.

As noted in the introduction and in [10], the branched double cover of D^4 over the black Tait surface F_B, which is denoted by $\Sigma(D^4, F_B)$, is a negative-definite filling of $\Sigma(S^3, K)$. The intersection form of $\Sigma(D^4, F_B)$ turns out to be the Goeritz form of F_B. There is a combinatorial description of the Goeritz form of the black Tait surface of an alternating knot in terms of the white Tait graph. Enumerate the vertices of $\widetilde{W}$ by $v_1, \dots, v_m$ and set

$$g_{ij} := |E_W(v_i, v_j)| \quad \text{for } i \neq j \text{ and } g_{ii} = -\deg_W(v_i).$$

Then the white Goeritz matrix $G_W := (g_{ij})$ represents the Goeritz form. Note that this is the definition of the Laplacian matrix of the graph W with the row and column associated with the marked vertex deleted. For the example presented in Figure 3, we have

$$G_W = \begin{pmatrix} -4 & 1 \\ 1 & -4 \end{pmatrix}.$$

Now we are going to describe a Kirby diagram of $\Sigma(D^4, F_B^+)$ which also acts as a surgery diagram for the $\Sigma(S^3, K)$. In the rest of the article, we refer to this construction as the *Tait surgery diagram*. This diagram originates from work of Ozsváth and Szabó [20].

We consider an unknot component with framing g_{ii} centered around each $v_i \in V_{\widetilde{W}}$ and then add a positive clasp between the unknot components corresponding to v_i and v_j for each edge $e \in E_{\widetilde{W}}$ between v_i and v_j. The intersection form of this Kirby diagram is clearly the same as G_W. Applying this to the example in Figure 3 gives us the surgery diagram in Figure 4.

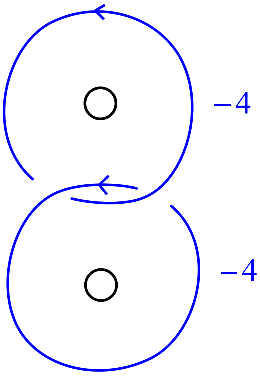


Figure 4: Surgery diagram of the branched double cover.

Greene [13] derives a Heegaard triple $\mathcal{H}_1$ subordinate to this surgery diagram and in combination with another Heegaard triple $\mathcal{H}_2$ originating from the Montesinos trick, he gives a combinatorial description of $\widehat{\mathrm{HF}}(\Sigma(S^3, K))$. We only need some of Greene's results about alternating links which we will summarize in the following.

Given a Kauffman state x for K, we induce an orientation on the white graph W in the following way. Given an edge $e \in E_W$, consider the crossing c to which it corresponds, as well as the white region which abuts c and lies on the same side of the over-strand as $x(c)$. We direct e to point towards the vertex corresponding to this white region. This process is illustrated in Figure 5.

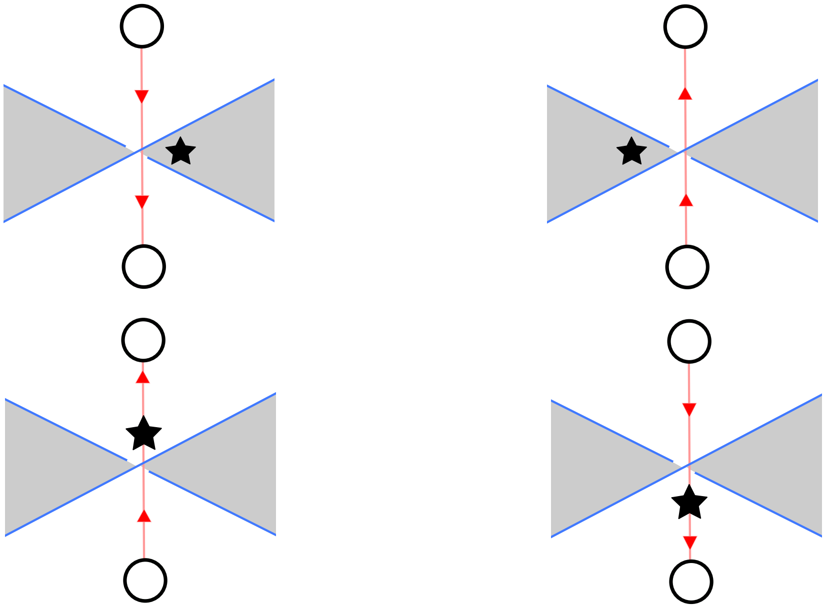


Figure 5: Orientation induced by a Kauffman state on the white graph. The Kauffman marker is depicted by the black star.

At a vertex $v \in V_{\widetilde{W}}$, we compute the signed degree $d_x(v)$ as

$$d_x(v) = d_x^+(v) - d_x^-(v),$$

where $d_x^+(v)$ (resp. $d_x^-(v)$) denotes the number of edges directed into (resp. out of) v with respect to this orientation on W.

Define $v_x^W := (d_x(v_1), \dots, d_x(v_m))^T$. Define the quadratic form $q(v)$ as

$$q(v) := v^T G_W^{-1} v \quad \text{for all } v \in \mathbb{Z}^m.$$

A *characteristic covector* $v \in \mathbb{Z}^m$ is defined by the condition that

$$v_i \equiv (G_W)_{ii} \mod 2 \qquad \text{for all } 1 \le i \le m.$$

The characteristic covectors are useful in studying the space of $\mathrm{Spin}^{\mathbb{C}}$ structures of $\Sigma(S^3, K)$. We know that $\mathrm{Spin}^{\mathbb{C}}(\Sigma(S^3, K))$ is an affine space over $H^2(\Sigma(S^3, K))$. Using Poincaré duality we have

$$H^2(\Sigma(S^3, K)) \cong H_1(\Sigma(S^3, K)).$$

We can compute this homology using the Tait surgery diagram and Mayer–Vietoris sequence as follows.

Let $L = \bigcup_{i=1}^m L_i$ be the underlying m-component link of the Tait surgery diagram. The Tait surgery diagram and the definition of Dehn surgery give us the decomposition

$$\Sigma(S^3, K) = (S^3 \setminus N(L)) \cup_\phi \bigsqcup_{i=1}^m H_i.$$

In this decomposition $H_i = D^2 \times S^1$ is the solid torus glued in a tubular neighborhood of L_i using the gluing map

$$\phi : \bigsqcup_{i=1}^m \partial H_i \to \partial(S^3 \setminus N(L)).$$

Let μ_i and λ_i respectively be the meridian and longitude of L_i in S^3. Also let μ_i' and λ_i' respectively be the meridian and longitude of ∂H_i. Then the gluing map ϕ induces a map ϕ_* on the first homology such that

$$\phi_*([\mu_i']) = g_{ii} \cdot [\mu_i] + [\lambda_i].$$

Now the Mayer–Vietoris sequence gives us

$$H_1\Bigl(\bigsqcup_{i=1}^m \partial H_i\Bigr) \to H_1(S^3 \setminus N(L)) \oplus H_1\Bigl(\bigsqcup_{i=1}^m H_i\Bigr) \to H_1(\Sigma(S^3, K)) \to 0.$$

We can rewrite this sequence as

$$\bigoplus_{i=1}^m \langle [\mu_i'], [\lambda_i'] \rangle \hookrightarrow \langle [\mu_1], \dots, [\mu_m] \rangle \oplus \bigoplus_{i=1}^m \langle [\lambda_i'] \rangle \twoheadrightarrow H_1(\Sigma(S^3, K)).$$

The first map in the sequence is defined, for all $i = 1, \dots, m$, as

$$[\mu_i'] \to g_{ii} \cdot [\mu_i] + [\lambda_i] = g_{ii} \cdot [\mu_i] + \sum_{\substack{1 \le j \le m \\ j \ne i}} \operatorname{lk}(L_i, L_j) \cdot [\mu_j] = \sum_{j=1}^{m} g_{ji} \cdot [\mu_j] \quad \text{and} \quad [\lambda_i'] \to [\lambda_i'].$$

As a result, we have

$$H_1(\Sigma(S^3, K)) \cong \frac{\mathbb{Z}^m}{\operatorname{im}(G_W)} \cong \operatorname{coker}(G_W).$$

The combination of these results gives us an identification of $H^2(\Sigma(S^3, K))$ with $\operatorname{coker}(G_W)$.

Now we go back to studying $\operatorname{Spin}^{\mathbb{C}}(\Sigma(S^3, K))$. Greene [13] proved that $\operatorname{Spin}^{\mathbb{C}}(\Sigma(S^3, K))$ can be identified with $2 \cdot \operatorname{im}(G_W)$-orbits of characteristic covectors. The identification comes from the first Chern class

$$c_1 : \operatorname{Spin}^{\mathbb{C}}(\Sigma(S^3, K)) \to H^2(\Sigma(S^3, K)) \cong \operatorname{coker}(G_W).$$

Now as before we assume that K is a nonsplit alternating link. Let $\mathcal{T}$ be the set of Kauffman states of K. Greene [13] also proved that there is a one-to-one correspondence

$$x \in \mathcal{T} \longleftrightarrow \mathfrak{t}(x) = v_x^W + 2 \cdot \operatorname{im}(G_W) \in \operatorname{Spin}^{\mathbb{C}}(\Sigma(S^3, K)).$$

Furthermore, when $\det(K)$ is odd, the first Chern class c_1 is a canonical identification of $\operatorname{Spin}^{\mathbb{C}}(\Sigma(S^3, K))$ and $\operatorname{coker}(G_W)$, and hence we have

$$c_1(\mathfrak{t}(x)) = [v_x^W] \in \operatorname{coker}(G_W).$$

We can now finally state Greene's computation of the Heegaard Floer homology of $\Sigma(S^3, K)$ in Theorem 2.2.

Theorem 2.2 [13] *Let K denote a nonsplit alternating link. Then*

$$\widehat{\mathrm{HF}}(\Sigma(S^3, K)) = \bigoplus_{x \in \mathcal{T}} \widehat{\mathrm{HF}}(\Sigma(S^3, K), \mathfrak{t}(x)) = \bigoplus_{x \in \mathcal{T}} \mathbb{Z}.$$

The correction term can be computed by the formula

$$d(\Sigma(S^3, K), \mathfrak{t}(x)) = \max_{v \in \mathfrak{t}(x)} \frac{q(v) + m}{4} = \frac{q(v_x^W) + m}{4},$$

where $m = |V_W| - 1$.

Remark 2.3 The notation v_x^W might be a bit confusing. Note that v_x^W is the degree vector of W restricted to the unmarked vertices $\{v_1, \dots, v_m\}$ which are also the vertices of $\widetilde{W}$.

We are going to use these formulas to obstruct branched double covers from having spin negative-definite fillings. To accomplish this, we are going to use known inequalities about fillings of rational homology spheres and closed spin 4-manifolds. We recall some of the theorems that we are going to use later.

Theorem 2.4 [19, Theorem 9.6] *Let Y be a rational homology three-sphere, and fix a $\mathrm{Spin}^{\mathbb{C}}$ structure $\mathfrak{t}$ over Y. Then, for each smooth, negative-definite filling X of Y, and for each $\mathfrak{s} \in \mathrm{Spin}^{\mathbb{C}}(X)$ with $\mathfrak{s}|_Y = \mathfrak{t}$, we have that*

$$c_1(\mathfrak{s})^2 + b_2(X) \leq 4d(Y, \mathfrak{t}).$$

Theorem 2.5 [7, Theorem 1] *If M is a closed spin manifold with indefinite intersection form, then*

$$b_2(M) \geq \tfrac{10}{8}|\sigma(M)| + 2.$$

3 Bounds from correction terms

Now we are ready to prove Theorems 1.2 and 1.3.

Proof of Theorem 1.2 Assume X is a simply connected, spin, negative-definite filling of $Y = \Sigma(S^3, K)$. Recall that $\mathrm{Spin}^{\mathbb{C}}$ structures of X correspond to integral lifts of the second Stiefel–Whitney class under the second map in the exact sequence

$$H^2(X;\mathbb{Z}) \xrightarrow{\times 2} H^2(X;\mathbb{Z}) \to H^2(X;\mathbb{Z}_2) \xrightarrow{\beta} \cdots.$$

The first Chern class of a $\mathrm{Spin}^{\mathbb{C}}$ structure is equal to this integral lift of the Stiefel–Whitney class. Since X is spin, the second Stiefel–Whitney class vanishes and, as a result, one can find a trivial lift $\mathfrak{s} \in \mathrm{Spin}^{\mathbb{C}}(X)$ with $c_1(\mathfrak{s}) = 0$. Due to Theorem 2.2, there is a Kauffman state x such that $[v_x^W] \in \operatorname{coker}(G_W)$ is identified with $\mathfrak{s}|_Y \in \mathrm{Spin}^{\mathbb{C}}(\Sigma(S^3, K))$. Using Theorem 2.4, we can write

$$b_2(X) \leq 4d(Y, \mathfrak{s}|_Y) = q(v_x^W) + m \leq m.$$

The last inequality follows from the fact that q is a negative-definite form as its defined using inverse of Goeritz matrix.

Now we are going to prove that the last inequality is sharp if K is not special. This again comes from Theorem 2.2. We only need to show that, for all Kauffman states x on a nonspecial alternating knot, $q(v_x^W) < 0$. Since q is negative-definite, we only need to prove that $v_x^W \neq 0$. This follows from the fact that, for nonspecial knots, W contains at least two vertices with odd degrees, since its dual can not be bipartite. As a result, there exist $v_i \in V_{\widetilde{W}}$ such that $\deg_W(v_i)$ is odd. On the other hand, v_x^W is a characteristic covector; i.e.,

$$v_x^W = (d_x(v_i)) = (d_x^+(v_i) - d_x^-(v_i)) \equiv (\deg_W(v_i)) = -g_{ii} \mod 2.$$

Finally, we need to show that, for special alternating knots, there exists a simply connected, spin, and negative-definite filling of $\Sigma(S^3, K)$ with $b_2 = m$. The 4-manifold $X = \Sigma(D^4, F_B^+)$ satisfies these conditions. Using the Tait surgery diagram (which is a Kirby diagram of X), one can see that X is simply connected with $b_2(X) = m$ and negative-definite with intersection form $Q_X = G_W$. Furthermore Q_X is even as $g_{ii} = -\deg_W(v_i)$ which is even since K is special. Hence, X is spin as well. □

Proof of Theorem 1.3 This proof is similar to the previous one. Due to Theorem 2.2, there is a Kauffman state x such that $[v_x^W] \in \operatorname{coker}(G_W)$ is identified with $\mathfrak{s}|_Y \in \operatorname{Spin}^{\mathbb{C}}(\Sigma(S^3, K))$. Using Theorem 2.4, we can write

$$b_2(X) \le 4d(Y, \mathfrak{s}|_Y) = q(v_x^W) + m.$$

Note that $\mathfrak{s}$ is induced by a Spin structure on X, and, as a result, $\mathfrak{s}|_Y$ is also the $\operatorname{Spin}^{\mathbb{C}}$ structure induced by the unique Spin structure on $\Sigma(S^3, K)$. The uniqueness of the Spin structure follows from the assumption that K is a knot. It is known that for a knot K, $\det(K) = |H_1(\Sigma(S^3, K); \mathbb{Z})|$ is odd. Consequently, $H_1(\Sigma(S^3, K); \mathbb{Z})$ has no 2-torsion, and $H^1(\Sigma(S^3, K); \mathbb{Z}_2)$ vanishes. Using this argument and the one made in the first lines of the proof of Theorem 1.2, we can deduce that $c_1(\mathfrak{s}|_Y) = 0$.

We will show that the inequality in the statement of Theorem 1.3 results in $d(Y, \mathfrak{s}|_Y)$ being negative. We know that $\operatorname{Spin}^{\mathbb{C}}(\Sigma(S^3, K))$ is identified with $\operatorname{coker}(G_W)$ through the first Chern class. As a result

$$[v_x^W] = [0] = [c_1(\mathfrak{s}|_Y)] \in \operatorname{coker}(G_W),$$

which means that there exists $y \in \mathbb{Z}^m$ such that $G_W y = v_x^W$. Note that we can rewrite

$$q(v_x^W) = (v_x^W)^T G_W^{-1} v_x^W = y^T G_W^T y = y^T G_W y.$$

We know that v_x^W is a characteristic covector and by definition we have

$$(v_x^W)_i \equiv -g_{ii} \mod 2. \tag{3-1}$$

We can rewrite (3-1) as

$$(v_x^W)_i = (G_W y)_i = \Sigma_j g_{ij} y_j \equiv \deg_W(v_i) = -g_{ii} \mod 2.$$

Let $y' \in \mathbb{Z}^m$ be the vector defined by $y'_i = (y_i \bmod 2) \in \{0, 1\}$. Let J be the support of y'. Let C be the subgraph of $\widetilde{W}$ induced by v_j for $j \in J$.

Now one can see that

$$\deg_W(v_i) \equiv (G_W y)_i \equiv (G_W y')_i \equiv e_W(v_i, C) \mod 2.$$

The last equality follows from the properties of the Laplacian matrix. Indeed, we have that

$$(G_W y')_i = \sum_{j \in J} g_{ij} = \sum_{j \in J - \{i\}} g_{ij} + \sum_{i \in J} g_{ii} = \sum_{j \in J} |E(\{v_i\}, \{v_j\})| - \sum_{i \in J} \deg_W(v_i),$$

which is equal to $e(v_i, C) \bmod 2$. As a result, C is a characteristic subgraph. Note that, since we assume the knot to be nonspecial, C can not be empty.

We can also use the interpretation of G_W as a submatrix of the Laplacian of W to reformulate $y^T G_W y$ as the following sum. Assume that L_W is the full Laplacian of W with the $(m+1)$-st (last) row and column associated to the distinguished vertex and set $y_{m+1} = 0$. Then we have

$$y^T G_W y = [y^T \ 0] L_W \begin{bmatrix} y \\ 0 \end{bmatrix} = - \sum_{\{v_i, v_j\} \in E_W} (y_i - y_j)^2. \tag{3-2}$$

Due to the definition of C, we will have $y_i \neq y_j$ for $v_i \in V_C$ and $v_j \in V_W \setminus V_C$. Combined with (3-2),

$$q(v_x^W) = y^T G_W y \leq -|E(V_C, V_W \setminus V_C)|.$$

Combining this with the statement of Theorem 1.3, we have that $q(v_x^W) \leq -m$, which gives us the result that we want. □

Remark 3.1 As we will mention in the next section, the definition of characteristic subgraph is an analogue of the definition of characteristic sublink in a surgery diagram (see Definition 4.1). Characteristic sublinks in a surgery diagram of Y are in one-to-one correspondence with $H^1(Y;\mathbb{Z}_2)$ and, as a result, in the setting of Theorem 1.3, there is only one characteristic sublink in the white Tait graph.

Remark 3.2 We can rephrase the main inequality of Theorems 1.3 and 1.4 in terms of the length of characteristic vectors in *Goeritz lattice* as follows.

Same as before let $m = |V_W| - 1$. The integral *Goeritz lattice* $\Lambda \subset \mathbb{R}^m$ is defined with the symmetric bilinear form

$$\langle x, y\rangle = x^T G_W y \quad \text{for all } x, y \in \mathbb{Z}^m.$$

A vector $w \in \mathbb{Z}^m$ is called *characteristic* if for all $x \in \mathbb{Z}^m$ we have

$$\langle w, x\rangle \equiv \langle x, x\rangle \mod 2.$$

Let $\mathrm{Char}(\Lambda)$ be the set of all characteristic vectors in Λ.

This definition is directly related to our definition of characteristic subgraph. First, let w' be the mod 2 reduction of w, i.e., for all $i \in \{1, \dots, m\}$ we have $w'_i = (w_i \bmod 2) \in \{0, 1\}$. Note that w is characteristic if and only if w' is characteristic. Let $J \subseteq \{1, \dots, m\}$ be the support of w', and let C be the subgraph of $\widetilde{W}$ induced by v_j for $j \in J$. One can see that w is a characteristic vector if and only if C is a characteristic subgraph.

Furthermore, similar to (3-2), for all $x \in \mathbb{Z}^m$ we have

$$\langle x, x\rangle = x^T G_W x = -\sum_{\{v_i, v_j\} \in E_W} (x_i - x_j)^2.$$

Now note that for any edge $\{v_i, v_j\} \in E_W(V_C, V_W \setminus V_C)$ we have

$$w_i \equiv w'_i = 1 \mod 2 \quad \text{and} \quad w_j \equiv w'_j = 0 \mod 2 \qquad \Longrightarrow \qquad (w_i - w_j)^2 \geq 1.$$

As a result,

$$\langle w, w\rangle = -\sum_{\{v_i, v_j\} \in E_W} (w_i - w_j)^2 \leq -|E_W(V_C, V_W \setminus V_C)| = \langle w', w'\rangle.$$

Finally we can conclude that

$$\min_{C \in \mathcal{C}_{\widetilde{W}}} |E_W(V_C, V_W \setminus V_C)| = \min_{w \in \mathrm{Char}(\Lambda)} |\langle w, w\rangle|.$$

As a result, we can conclude Corollary 1.5 from Theorem 1.4.

4 Bound from Furuta's $\frac{10}{8}$ theorem

It is well known that the third spin cobordism group vanishes. This means that any spin 3-manifold $(Y, \mathfrak{t})$ has a spin filling; i.e., there exists a spin 4-manifold $(W, \mathfrak{s})$ such that $\partial W = Y$ and $\mathfrak{s}|_Y = \mathfrak{t}$. In fact, Kaplan [15] built an algorithm that turns any surgery diagram of Y to a Kirby diagram of a spin filling through Kirby calculus. We will recall this algorithm from its exposition by Gompf and Stipsicz [9]. First, let us define the notion of a characteristic sublink of a framed link.

Definition 4.1 Let L be a framed link. A sublink $L' \subseteq L$ is called *characteristic* if and only if for any sublink $L_i \subseteq L$ we have

$$\mathrm{lk}(L', L_i) \equiv \mathrm{lk}(L_i, L_i) \mod 2.$$

Now assume that we have a surgery diagram L of a 3-manifold Y. Considering L as a Kirby diagram gives us a handlebody filling X_L of Y. For any component L_i of L, we get a class $[H_i] \in H_2(X_L; \mathbb{Z}_2)$ from capping off core of the 2-handle attached along L_i. These homology classes give a basis for $H_2(X_L; \mathbb{Z}_2)$. As a result, we have a bijection between sublinks of L and elements of $H_2(X_L; \mathbb{Z}_2)$ as

$$L' \subseteq L \longleftrightarrow \sum_{L_i \subseteq L'} [H_i] \in H_2(X_L; \mathbb{Z}_2).$$

Due to Poincaré duality we have $H_2(X_L; \mathbb{Z}_2) \cong H^2(X_L, Y; \mathbb{Z}_2)$. Combining these two facts, we end up with a bijection between sublinks of L and $H^2(X_L, Y; \mathbb{Z}_2)$.

Now we are ready to state the necessary tools from Gompf and Stipsicz [9] in Theorems 4.2 and 4.3.

Theorem 4.2 [9, Proposition 5.7.11] *For any* Spin *structure* $\mathfrak{t}$ *on* Y, *the nonvanishing of the relative Stiefel–Whitney class* $w_2(X_L, \mathfrak{t}) \in H^2(X_L, Y; \mathbb{Z})$ *serves as an obstruction for extending* $\mathfrak{t}$ *to* X. *The map*

$$\mathfrak{t} \in \mathrm{Spin}(Y) \to w_2(X_L, \mathfrak{t}) \in H^2(X_L, Y; \mathbb{Z})$$

gives a bijection between the set of Spin *structures on* Y *and characteristic sublinks of* L.

Proposition 4.3 [9, Theorem 5.7.14] *Assume that we have a surgery diagram* L *of a* 3*-manifold* Y. *Fix any* Spin *structure* $\mathfrak{t}$ *on* Y *and assume that* L' *is the corresponding characteristic sublink. The following steps will lead to a Kirby diagram of a spin filling of* $(Y, \mathfrak{t})$:

(1) *Slide one component* L'_0 *of* L' *over the rest of the components of* L'. *The characteristic sublink corresponding to* $\mathfrak{t}$ *in the new Kirby diagram will be the sublink consisting of one component* L'_0.

(2) *Unknot* L'_0 *using blow-ups. The blow-up circles can be imagined as connected sum of two small meridian circles* m_{i1}, m_{i2} *along a band* b_i *forming* D_i *as in Figure 6, left.*

The characteristic sublink will be the union of L'_0 *and all the blown-up circles.*

(3) *One can again use blow-ups to change the crossings between the bands and other components of* L, *without changing the characteristic sublink*; *see Figure 6, right. Use an isotopy to turn* L'_0 *into a circle in the plane and then use the operation of Figure 6, right*, *to turn the characteristic sublink into Figure 7, left.*

Figure 6: Left: Step 2 in Kaplan's algorithm [9, Figure 5.48]. Right: Step 3 in Kaplan's algorithm [9, Figure 5.49].

(4) *The operation shown in Figure 7, right, can be done using a blow-up. Use this operation to turn the characteristic sublink into an unlink.*

(5) *Consider each component of the characteristic sublink one by one. Blowing up its meridians turns the framing to* ± 1*. Then, by blowing down the characteristic sublink, one can turn it into the empty link.*

In the end, you will have a Kirby diagram with even framings. This 4*-manifold with its unique* Spin *structure is a spin filling of* $(Y, \mathfrak{t})$*.*

We are going to show that, in the setting of Theorem 1.4, Kaplan's algorithm simplifies and, as a result, one can compute the change in the signature and second Betti number and prove the obstruction of Theorem 1.4. We will use Lemma 4.5 in the proof.

Before we state the lemma we need to introduce some notation. We call a framed link a *chainmail link* if it is constructed in the following way. Let D be a weighted and signed plane multigraph. Assigned to each $v_i \in V_D$, there is an integer weight $w_i \in \mathbb{Z}$ and assigned to each $e_k \in E_D$, there is a sign $\mu_k \in \{+,-\}$. The framed link L_D is constructed by considering an unknot component L_i oriented counterclockwise and with framing w_i centered around each $v_i \in V_D$ and then add a left-handed (resp. right-handed) clasp between the unknot components corresponding to v_i and v_j for each edge $e_k \in E_D$ between v_i and v_j with $\mu_k = +$ (resp. $-$). This definition generalizes the construction of Tait surgery diagram explained in Section 2. Further details can be found in the work of Polyak [21].

Now we are going to explain a modification of the first step of Kaplan's algorithm. This procedure is called $MK1$ and is defined in Definition 4.4.

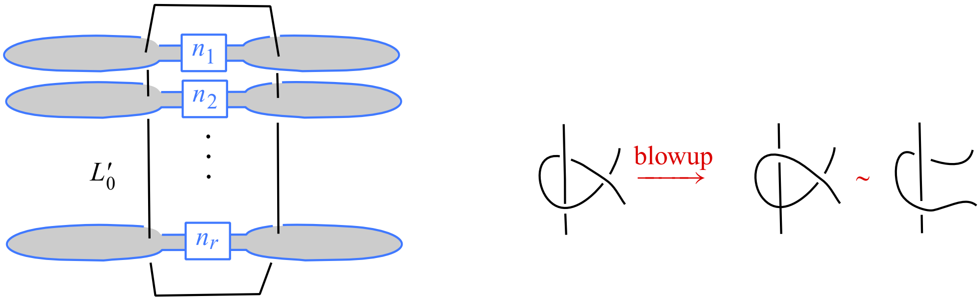


Figure 7: Left: Step 3 in Kaplan's algorithm [9, Figure 5.50]. Right: Step 4 in Kaplan's algorithm [9, Figure 5.51].

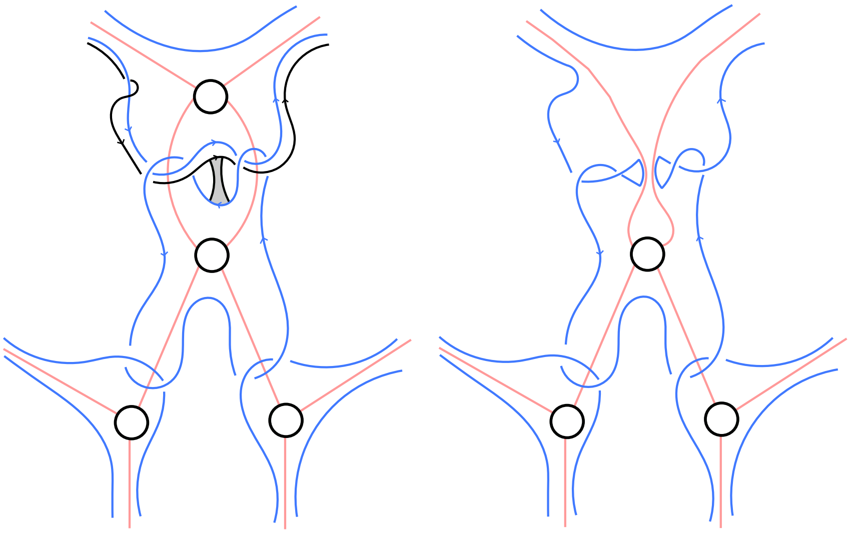


Figure 8: Handle slides in MK1.

Definition 4.4 Let L_D be the chainmail link based on the connected plane graph D and T be a spanning tree of D. Fix an arbitrary orientation (on each edge) and a total order $<$ on E_T. Consider the following procedure:

(1) Take the maximal edge e_m of T in the ordering and assume it is directed from v_p to v_t. Slide L_p over L_t with an orientation-preserving band.

(2) Contract e in T. The two vertices at two ends of e will form a new vertex which we denote by v_p (this labeling will be important in repeating step (1)).

(3) Repeat the process until $|V_T| = 1$.

The knot corresponding to the remaining vertex is denoted by $\mathrm{MK1}(L_D, T)$. See Figure 8.

Lemma 4.5 *Let L_D be the chainmail link based on the connected plane graph D. There exist a spanning rooted tree T with a total ordering and direction on edges such that* $\mathrm{MK1}(L_D, T)$ *is an unknot.*

Proof of Lemma 4.5 We construct T by induction on the number of vertices in D. For two adjacent vertices v_i, v_j in D, let $\mathrm{ER}(v_i, v_j)$ be the maximal bounded region in the plane enclosed by the edges between v_i and v_j with respect to inclusion. If only one edge connects v_i and v_j, let $\mathrm{ER}(v_i, v_j)$ be that single edge (as a subset of plane). Among all pairs of adjacent vertices, pick $\{v_r, v_s\}$ such that $\mathrm{ER}(v_r, v_s)$ is minimal, meaning it does not contain $\mathrm{ER}(v_i, v_j)$ for any other adjacent pair $\{v_i, v_j\}$. An example of such a minimal pair is a pair of adjacent vertices which only have one connecting edge.

Picking this minimal pair guarantees that all the other vertices lie in $\mathbb{R}^2 \setminus \mathrm{ER}(v_r, v_s)$ and this gives a standard model (see Figure 9) for the configuration of the clasps between L_r and L_s with respect to the other clasps involving L_r or L_s. We use this to control the result of sliding L_r over L_s.

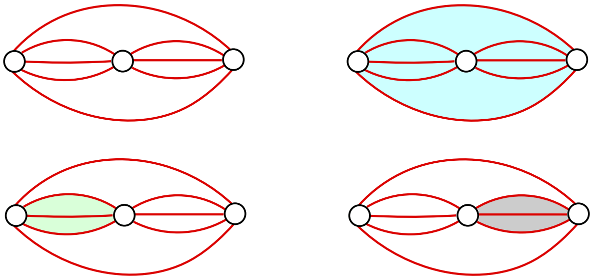


Figure 9: Examples of $\mathrm{ER}(v_i, v_j)$. The green and gray (shaded) regions in the bottom row are minimal.

Without loss of generality assume that $r = 1$ and $s = 2$. Figure 8 shows this sliding operation. It is clear that after a number of R1 moves, L_1 will be an unknot around v_1 and v_2 and there will be clasps associated with each edge between v_1 or v_2, and any v_j for $j \neq 1, 2$. This means that, after sliding L_1 over L_2 and deleting L_2, we will have a chainmail link on the plane graph $D/\{v_1, v_2\}$, which is the plane graph coming from contracting v_1 and v_2 to one vertex which we will denote by $v_{1,2}$. This decreases the size of the vertex set by one. Using induction, we know that there exists a spanning tree T' such that $\mathrm{MK1}(L_{D/\{v_1,v_2\}}, T')$ is an unknot. Spanning tree T can be constructed by replacing $v_{1,2}$ with v_1 and v_2 and the edge between them, directing the edge from v_1 to v_2 and putting it as the new maximal edge in the total ordering. □

Remark 4.6 As mentioned, the operation MK1 is designed to be a modification of the first step of Kaplan's algorithm for a chainmail link. We explain this in a bit more detail in the following. Note that we are using the notation of Theorems 4.2 and 4.3.

The main goal of the first step of the Kaplan algorithm is to construct a surgery diagram of Y such that this characteristic sublink only has one component. This can be achieved through handle slides. Note that the handle slides does not change the isomorphism type of the filling X_L. Sliding a component L'_i of L' over another component L'_j can be seen as a change of basis of $H^2(X_L, Y; \mathbb{Z}_2)$. As a result, for the purpose of the Kaplan's algorithm, we only need to find the expansion of $w_2(X_L, \mathfrak{t})$ in this new basis. This expansion shows that the characteristic sublink associated to $\mathfrak{t}$ in the new surgery diagram (after handle slide) is just $L' \setminus \{L'_j\}$.

Due to this fact, one can turn the characteristic sublink into a knot with handle slides in the first step of Kaplan's algorithm. The same reasoning shows that MK1 can replace the first step of Kaplan's algorithm when L' is a chainmail link.

Remark 4.7 The operation MK1 is only defined for chainmail links based on connected planar graphs. However, we can easily generalize MK1 and Lemma 4.5 to all chainmail links as follows.

If L_D is a chainmail link based on a disconnected graph $D = \bigsqcup_{i=1}^{k} D_i$, then L_D is also a split link which splits to chainmail sublinks as

$$L_D = \bigsqcup_{i=1}^{k} L_{D_i}.$$

We apply Lemma 4.5 to the each of L_{D_i}. For each $i = 1, \dots, k$, there exists a spanning rooted tree T_i in D_i such that $\mathrm{MK1}(L_{D_i}, T_i)$ is an unknot. Applying all of these operations turns L_D to the unlink $\bigsqcup_{i=1}^k \mathrm{MK1}(L_{D_i}, T_i)$. We can then slide one of the components of this unlink over the other components such that the final result is an unknot.

Due to Remark 4.6, this generalized operation can also replace the first step of Kaplan's algorithm.

Remark 4.8 The linking between the components of the link changes under the handle slides. We can update the linking matrix at each step with the rule

$$\mathrm{lk}(L_{1,2}, L_i) = \mathrm{lk}(L_1, L_i) + \mathrm{lk}(L_2, L_i).$$

Now we are ready to prove Theorem 1.4. In fact, we are going to prove a stronger result stated in Theorem 4.9.

Theorem 4.9 *Let K be a nonspecial alternating link, and $\mathfrak{t} \in \mathrm{Spin}^{\mathbb{C}}(\Sigma(S^3, K))$. Let C be the characteristic subgraph associated to $\mathfrak{t}$. If*

$$|E_W(V_C, V_W \setminus V_C)| \geq 9(|V_W| - 1),$$

then $(\Sigma(S^3, K), \mathfrak{t})$ does not have a simply connected negative-definite spin filling.

Proof of Theorem 4.9 Assume that $(X, \mathfrak{s})$ is a simply connected negative-definite spin filling of $Y = (\Sigma(S^3, K), \mathfrak{t})$. Let L be the Kirby diagram of $\Sigma(D^4, F_B)$ described in Section 2. Let L' be the characteristic sublink of L associated with $\mathfrak{t}$. We are going to apply the modified version of Kaplan's algorithm on L using L' which will result in a simply connected spin filling $(X', \mathfrak{s}')$ of $(\Sigma(S^3, K), \mathfrak{t})$. We can take $-X'$ and build the closed 4-manifold $W = X \cup_Y (-X')$ which is also spin since $\mathfrak{s}|_Y = \mathfrak{s}'|_Y = \mathfrak{t}$. Using Furuta's inequality (Theorem 2.5) on W gives

$$b_2(X) + b_2(-X') \geq \tfrac{10}{8}|\sigma(X) + \sigma(-X')| + 2. \tag{4-1}$$

The right-hand side comes from Novikov additivity. Now we need to compute $b_2(-X')$ and $\sigma(-X')$.

Using Lemma 4.5 and Remark 4.6, we know that Steps 2, 3, 4 of Kaplan's algorithm will not be needed in our setting, and we can easily compute the change of b_2 and σ. First note that in each of the described handle slides, the framings change in the following way. If we assume that framing of L_p, L_s are r_p, r_s, respectively, then the framing of the L_p after the sliding will be $r_p + r_s + 2\,\mathrm{lk}(L_p, L_r)$. Using induction and Remark 4.8, we can see that framing on the final component of the characteristic sublink (after finishing Step 1) is equal to

$$\sum_{v_i \in C} g_{ii} + 2 \sum_{\substack{i \neq j \\ v_i, v_j \in C}} \mathrm{lk}(L_i, L_j) = \sum_{v_i \in C} g_{ii} + 2 \sum_{\substack{i \neq j \\ v_i, v_j \in C}} g_{ij}.$$

Since $g_{ii} = -\deg_W(v_i)$ and $g_{ij} = |E(v_i, v_j)|$, we will have that

$$\begin{aligned}\sum_{v_i \in C} g_{ii} + 2 \sum_{\substack{i \neq j \\ v_i, v_j \in C}} g_{ij} &= -\sum_{v_i \in C} \deg_W(v_i) + 2 \sum_{\substack{i \neq j \\ v_i, v_j \in C}} |E(v_i, v_j)| \\ &= -|E_W(V_C, V_W \setminus V_C)| - 2|E_C| + 2|E_C| = -|E_W(V_C, V_W \setminus V_C)|.\end{aligned}$$

This is the right-hand side of the inequality stated in Theorem 1.4. We denote this value by $-f$. Note that in Step 1, we only use handle slides and isotopies which means that the filling will not change. To calculate the change in b_2 and signature, we only need to look at Step 5. In this step, we blow up $f-1$ meridians in order to turn the characteristic sublink into an unknot with framing -1 and then blow down this unknot. These increase b_2 and σ by $f-2$ and f, respectively. Now we only need to use this information in Furuta's inequality. Note that since X is negative-definite $\sigma(X)=-b_2(X)$. Assuming that $f \geq 9m$, which is the assumption of Theorem 1.4 (where $m=|V_W|-1$), we can rewrite (4-1),

$$\begin{aligned} b_2(X)+m+f-2 \geq \tfrac{10}{8}|-b_2(X)+m-f|+2 &\iff b_2(X)+m+f-2 \geq \tfrac{10}{8}(b_2(X)+f-m)+2 \\ &\iff \tfrac{18}{8}m \geq \tfrac{2}{8}f+\tfrac{2}{8}b_2(X)+4. \end{aligned}$$

The final inequality is a clear contradiction to $f \geq 9m$.

Note that the equivalence between the first and second inequality follows from the assumption that $f \geq 9m$ which means that $b_2(X)+f-m>0$. □

Remark 4.10 This procedure gives a generalization of a corollary of Ue [24, Theorem 1]. Recall that, for a plumbed 3-manifold Y, the Neumann–Siebenmann invariant $\bar{\mu}$ is defined as follows. Assume that Γ is the plumbing tree and $P(\Gamma)$ is the 4-manifold constructed from plumbing sphere bundles based on Γ. We know that $Y=\partial P(\Gamma)$. Let w_s be the indicator vector of the characteristic sublink associated with a Spin structure $\mathfrak{s}$ on Y. Then

$$\bar{\mu}(Y,\mathfrak{s}) := \tfrac{1}{8}\big(\sigma(P(\Gamma)) - \langle w_s, w_s\rangle\big),$$

where $\langle\cdot,\cdot\rangle$ represents the intersection pairing. Ue proves that a Seifert homology sphere Y with Spin structure $\mathfrak{s}$ bounding a negative-definite 4-manifold X with Spin structure $\mathfrak{s}_X$ must satisfy

$$-\tfrac{8}{9}\bar{\mu}(Y,\mathfrak{s}) \leq b_2(X) \leq -8\bar{\mu}(Y,\mathfrak{s}).$$

Now an obstruction to the existence of simply connected negative-definite spin fillings is $\bar{\mu}(Y,\mathfrak{s})>0$. In cases when $\widetilde{W}$ is a star-shaped tree (which is the plumbing tree of a Seifert homology S^3) one can apply this obstruction to our problem. Whenever $\widetilde{W}$ is a tree, any characteristic subgraph C will be a disjoint union of isolated vertices. As a result, if $w_C \in \mathbb{Z}^m$ is the indicator vector of V_C, then

$$\langle w_C, w_C\rangle = \sum_{v_i \in C} -g_{ii} = \sum_{v_i \in C} -\deg_W(v_i) = -|E_W(V_C, V_W \setminus V_C)|. \tag{4-2}$$

This comes from the fact that $E_C=\varnothing$. Combining (4-2) and definition of $\bar{\mu}(Y,\mathfrak{s})$ gives us

$$8\bar{\mu}(Y,\mathfrak{s}) = -m + |E_W(V_C, V_W \setminus V_C)|.$$

This means that Theorem 1.4 generalizes the obstruction $\bar{\mu}(Y,\mathfrak{s})>0$.

This simplification of Kaplan's algorithm and the fact that one can build a characteristic unknot without blowing up or down is of independent interest. The following corollary easily follows from this observation.

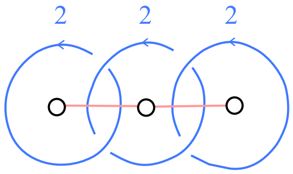


Figure 10: Kirby diagram of P_4.

Corollary 4.11 *If a closed* 4*-manifold* X *has a chainmail Kirby diagram, then it is either spin or has a characteristic sphere.*

This can act as an obstruction for a manifold to have a chainmail diagram. The following is an explicit example of this. This example was pointed out to us by Marco Golla.

Example 4.12 Let X be the Akhmedov–Park exotic $\mathbb{C}P^2 \# 2\overline{\mathbb{C}P^2}$ [2]. We prove that X does not have a characteristic sphere as follows.

Note that X is symplectic and minimal. Any embedded sphere $S \subset X$ satisfies $S \cdot S \leq -2$. Adjunction inequality gives us that $S \cdot S \leq 0$, and due to the minimality of X there is no embedded sphere with self-intersection -1. Let $S \cdot S = -m$. We use P_m to denote the 4-manifold constructed by negative linear plumbing of $m-1$ disk bundles with Euler number 2 over S^2. We can construct a Kirby diagram of P_m by considering the path graph on $m-1$ vertices and applying the chainmail construction (with all the weights equal to 2, and all the signs equal to $-$). We can see an example of this for P_4 in Figure 10.

There is an orientation-preserving diffeomorphism between ∂P_m and $\partial N(S)$. Using the mentioned Kirby diagram as a surgery diagram for ∂P_m, and the iterated slam-dunk move (see [18]), we see that $\partial P_m = L(m, m-1)$ since

$$\frac{m}{m-1} = [\underbrace{2, \dots, 2}_{m-1}]^- = 2 - \cfrac{1}{\cdots - \cfrac{1}{2 - \frac{1}{2}}}.$$

Finally, $\partial N(S)$ is the circle bundle over S^2 with Euler number $-m$. This means that $\partial N(S) \simeq -L(m, 1)$. We also have that $L(m, m-1) \simeq -L(m, 1)$. The composition of these two gives us the desired diffeomorphism.

Now one can form the closed 4-manifold $M = (X \setminus N(S)^{\mathrm{o}}) \cup P_m$. We are going to prove that M is spin. Note that $X \setminus N(S)$ is spin since S is characteristic. Let $\mathfrak{s}$ be the Spin structure induced on $\partial(X \setminus N(S)^{\mathrm{o}})$. We need to prove that P_m induces the same Spin structure on $\partial P_m = L(m, m-1)$. If m is odd, then $L(m, m-1)$ has a unique Spin structure and we are done. If m is even, then $L(m, m-1)$ has two different Spin structures $\mathfrak{s}, \mathfrak{s}'$. Furthermore when m is even, $N(S)$ is spin. However, since X is not spin, the induced Spin structure on $\partial N(S)$ must be $\mathfrak{s}'$. Now we only need to show that P_m and $N(S)$ induce different Spin structures on $L(m, m-1)$ which is equivalent to $P_m \cup -N(S)$ not being spin. The closed 4-manifold $P_m \cup -N(S)$ is positive-definite and $b_2(P_m \cup -N(S)) = m$, hence it can not be spin due to Donaldson's diagonalization theorem [4].

Now M is a closed, simply connected, spin 4-manifold. Using Novikov additivity we can deduce that $\sigma(M) = m-1$. Furthermore, $b_2^-(M) = 1$ since the negative-definite part of $H_2(M)$ lies inside $X \setminus N(S)$. Donaldson's theorem B [5] tell us that $\sigma(M) = 0$ which is a contradiction.

5 Spin negative-definite plumbed fillings

The previous results might lead one to ask if there are any nonspecial alternating knots K such that $\Sigma(S^3, K)$ has a simply connected, spin and negative-definite filling. A result such as the following theorem might further support this.

Theorem 5.1 *A nonspecial alternating link K does not have a spanning filling which is spin and negative-definite.*

Proof Assume $\Sigma(D^4, F^+)$ is a negative-definite spin filling. We know that the intersection form of $\Sigma(D^4, F^+)$ is the Goeritz form of the surface, which means that G_F is a negative-definite spanning surface. Using Theorem 1.1, we can conclude that F must be the black Tait surface in a diagram of K. Then we know, that for G_F to be even, the knot needs to be special, which contradicts the assumption. □

For general fillings, this is far from the truth. We present an example of a nonspecial alternating knot K with a spin negative-definite filling of $\Sigma(S^3, K)$. The main tool for the construction of this example is the lens space realization problem and one can generate a family of examples in the same way. We must mention that part of the inspiration for the example comes from Aceto, McCoy, and Park [1] which addresses negative-definite fillings of lens spaces with minimal b_2. The main difference is that we need to use the lens space fillings that emerge as the trace of a surgery on a knot instead of rational homology ball fillings. We use the notation K^m to denote the result of Dehn surgery on S^3 along K with slope m. We also use $\mathrm{Tr}(K^m)$ to denote the trace of this surgery. Using this notation we have

$$\partial \mathrm{Tr}(K^m) = K^m.$$

Example 5.2 The knot K will be the alternating knot in Figure 11. The white Tait graph is also drawn in the figure. The reduced white Tait graph will be a single path with framings $(-4, -2, -5, -2)$. Plumbing along this path with these framings will give us the standard negative-definite filling of $\Sigma(S^3, K)$. We use the notation $P(a_1, \dots, a_n)$ to denote the linear plumbing with framings $a_1, \dots, a_n$. For example,

$$\partial P(-4, -2, -5, -2) = \Sigma(S^3, K).$$

We have a standard embedding of $P(-2, -5, -2)$ in $P(-4, -2, -5, -2)$ which comes from the embedding of the plumbing graph of the first 4-manifold in the second. In other words, $P(-4, -2, -5, -2)$ can be constructed from $P(-2, -5, -2)$ by attaching a 2-handle with framing -4 to its boundary. The integers $(-2, -5, -2)$ also arise in the fractional expansion

$$\tfrac{16}{9} = 2 - \cfrac{1}{5 - \frac{1}{2}}.$$

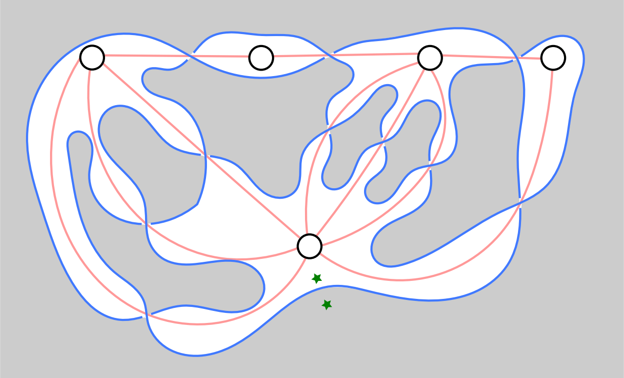


Figure 11: Alternating knot K of Example 5.2.

This means that $P(-2,-5,-2)$ bounds $L(16,9)$. We claim that $L(16,9)$ also has a filling in the form of the trace of a knot surgery, i.e.,

$$L(16,9) = \partial \mathrm{Tr}(K^{-16}).$$

This comes from the description of the Berge knots of type $I_{\pm}$ (see [12]). Picking $i,k \in \mathbb{Z}$ such that $\gcd(i,k) = 1$ and setting $p = ik \pm 1$ and $q \equiv -k^2 \bmod p$, leads to the lens space $L(p,q)$ which can be realized by a positive surgery on a knot. Setting $i = 3$ and $k = 5$, gives us $p = 16$ and $q \equiv -25 \bmod 16$, which means that there exists a knot K such that $L(16,7) = K^{+16}$, which in turn means that $L(16,7) = \partial \mathrm{Tr}(K^{+16})$. By reversing the orientation, we have $L(16,9) = -L(16,7) = \partial \mathrm{Tr}(K^{-16})$.

Let $i : P(-2,-5,-2) \hookrightarrow P(-4,-2,-5,-2) = X$ be the aforementioned embedding. Now we construct a 4-manifold X' by deleting the interior of $\mathrm{Im}(i)$ from X and gluing $\mathrm{Tr}(K^{-16})$ in its place. Then X' will be the result of attaching a -4-framed 2-handle to $\mathrm{Tr}(K^{-16})$, which means it is simply connected and has $b_2 = 2$, as it has a handle decomposition with two 2-handles. The intersection form is even as the framing of both 2-handles is even, and hence, X' is spin. Using Novikov additivity we can also prove that X' is negative-definite. This is again due to the fact that, while constructing X', we delete a submanifold of X with signature -3 and replace it with one with signature -1. This finally gives us the example we need.

Although there is no general nonexistence result for simply connected negative-definite spin fillings, by imposing suitable combinatorial conditions on the Kirby diagram of the filling we can prove such results. The first result of this type is Theorem 1.7. This result is directly rooted in Neumann's plumbing calculus. In the following theorem we recall the facts we need from [17]. Note that the branched double covers of links with nonzero determinant are rational homology spheres. As a result, they can be realized as plumbings of disk bundles over surfaces when the base surfaces are all spheres and the plumbing graph is a tree. We can describe these plumbings with a tree with integer weights on vertices and ± signs on edges. This means that we do not need Neumann's plumbing calculus in its full generality. In the following theorem, we only recall the facts we need.

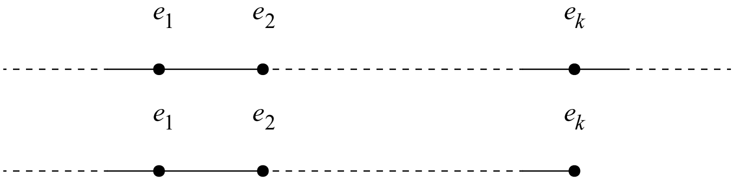


Figure 12: A chain in a plumbing graph [17].

In the rest of the paper, we use the term chain to refer to path subgraphs which contain at most one leaf of T (see Figure 12).

Theorem 5.3 [17, Theorem 4.1] *Any plumbing tree T can be reduced to a unique normal form using the following moves while keeping the boundary unchanged.*

(R0) *Reverse the sign of all the edges adjacent to a vertex v.*

(R1a) *Delete a component consisting of an isolated vertex with weight ± 1.*

(R1b–R3) *These are the moves which are described in Figure 13.*

The normal form is defined by the following properties:

(N1) *None of the operations can be applied*, *except that T might contain a component like Figure 14 with $k \geq 1$ and $e_i \leq -2$ for all i.*

(N2) *The weights e_i on all chains of T satisfy $e_i \leq -2$ for all i.*

(N3) *No portion of T has the form shown in Figure 15*, *top*, *unless it is in a component of the form shown in Figure 15*, *bottom*, *with $k \geq 1$ and $e_i \leq -2$ for all i.*

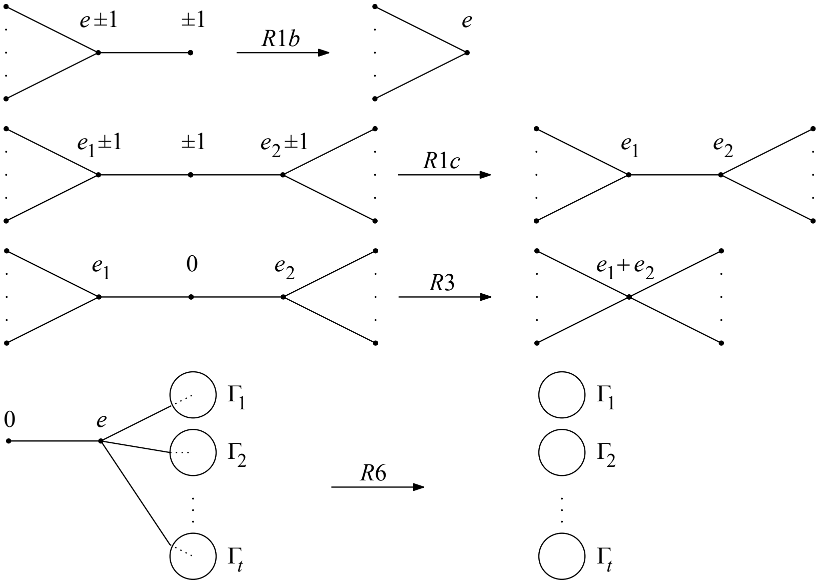


Figure 13: Neumann moves [17].

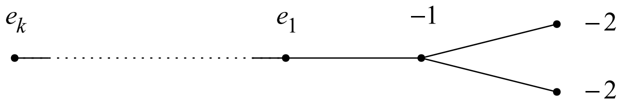


Figure 14: Property N1 of Neumann normal form [17].

You might notice that the moves described in Theorem 5.3 do not describe the change in the edge signs. When we are dealing with trees, the move R0 gives us that the edge signs do not matter.

Before we proceed with proving Theorem 1.7, we need to define the excessive property. This definition comes from Murasugi [16].

Recall that a link is called algebraic (or arborescent) if it can be constructed as the boundary of a plumbing of twisted bands according to a tree (see Definition 1.6). For a weighted tree T, we denote the algebraic link constructed from a plumbing based on T by $l(T)$. The tree T is called *negative excessive* if

$$w(a_i) \leq \min\{-2, -\deg_T(a_i)\} \quad \text{for all } a_i \in V_T.$$

The following lemma is proved by Murasugi.

Lemma 5.4 [16, Propositions 3.3 and 4.1] *For a negative excessive tree T, the link $l(T)$ is alternating. Furthermore, there is an alternating diagram of $l(T)$ such that T is isomorphic to the reduced white Tait graph $\widetilde{W}$. This isomorphism takes the weights of T to g_{ii} (diagonal of the Goeritz matrix).*

With this information, we can proceed with proving Theorem 1.7.

Proof of Theorem 1.7 We start by proving that any simply connected negative-definite spin plumbed filling is automatically in normal form. Due to the spin condition, we will not have framing ±1 on any vertex as all framings are even. Due to the negative-definite condition, we can not have any vertex with framing 0 as all framings are negative. This means that conditions N1 and N2 are satisfied. To show that N3 is satisfied, we use the assumption that K is a knot and hence has an odd determinant. We know that the determinant of the knot is equal to $|H_1(\Sigma(S^3,K);\mathbb{Z})|$. When the determinant is odd, the 2-torsion vanishes and, as a result, $H^1(\Sigma(S^3,K);\mathbb{Z}_2)=0$, which means that $\Sigma(S^3,K)$ has a unique Spin structure. We now use Proposition 4.3 to deduce that the number of characteristic sublinks of the Kirby diagram is equal to 1. This in turn means that the plumbing tree T has a unique characteristic subgraph.

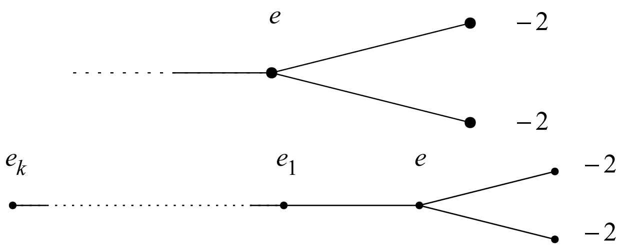


Figure 15: Property N3 of Neumann normal form [17].

We use proof by contradiction. Assume T violates condition N3, which means it contains the forbidden subgraph of Figure 15, top. A characteristic subgraph $C \subseteq T$ can not contain the parent vertex of the -2-framed leaves since the number of edges between a -2-framed leaf and C must be even (due to the definition of characteristic sublink). Let us use the names $L = \{l_1, l_2\}$ and p to denote the -2 framed leaves and their parent vertex. Also define $A := C \cap \{l_1, l_2\}$. Now we consider the subgraph C' defined as

$$C' = (C - A) \cup (L - A).$$

This subgraph is also characteristic. The only change happens with taking the complement of $C \cap L$ on L, which means that $E(v, C)$ and $E(v, C')$ are only different for $v \in \{p, l_1, l_2\}$. In all three cases, the parity of $|E(v, C)|$ and $|E(v, C')|$ are the same as $|E(p, C')| = |E(p, C)| - |A| + (2 - |A|)$ and $|E(l_i, C)| = |E(l_i, C')| \pm 2$. This construction builds a fixed-point-free bijection on the set of characteristic subgraphs, which means that the size of this set must be even. This gives us a contradiction with the argument in the previous paragraph. As a result, T must satisfy condition N3.

Let $\widetilde{W}$ be the reduced white Tait graph of K. By Lemma 5.4, We know that $\widetilde{W}$ is isomorphic to the plumbing tree associated to K. Using the Tait surgery diagram, we can see that $\Sigma(D^4, F_B^+)$ is a plumbed filling. We are going to prove that this plumbed filling is also in normal form. The excessive condition forces all weights to be ≤ -2 and as a result N1 and N2 are satisfied. Using the same argument as the previous paragraph, we can prove that condition N3 is also satisfied.

Now using the uniqueness of Neumann normal form, one can deduce that if a simply connected negative-definite spin plumbed filling exists, then its plumbing tree is exactly the reduced white Tait graph. This means that the framings in the white Tait graph; i.e., g_{ii}, must be all even, which is equivalent to the knot being special. □

The main idea behind Theorem 1.7 can be generalized to some other types of fillings.

Definition 5.5 We call a filling X of a 3-manifold Y a chainmail filling if and only if there exist a Kirby diagram of X which is a chainmail link

Following the discussion in Section 2, the 4-manifold $\Sigma(D^4, F_B^+)$ always gives a chainmail filling of the branched double cover. Unfortunately, there are no known normal forms for chainmail Kirby diagrams in the literature so the proof of Theorem 1.7 can not be replicated, but we can use the trick described here which is inspired by Murasugi [16].

Definition 5.6 A weighted planar graph is called *accessible* if it can be realized as the white Tait graph of an alternating link K such that the weights are equal to diagonal entries of the Goeritz matrix of K. We call a chainmail filling accessible if it has a chainmail Kirby diagram which is based on an accessible planar graph.

The main examples of accessible planar graphs come from the following example:

Example 5.7 Let G be a 2-connected plane graph such that all vertices are adjacent to the unbounded region, i.e., the boundary of the unbounded region includes all of the vertices of G. Furthermore, assume that G is negative excessive; i.e., weights satisfy

$$w(v_i) \leq \min\{-2, -\deg_G(v_i)\} \quad \text{for all } v_i \in V_G.$$

In this setting, one can add a vertex $\hat{v}$ in the unbounded region and connect it to all $v_i \in V_G$ such that

$$|E(\hat{v}, v_i)| + \deg_G(v_i) = |w(v_i)|.$$

The median construction on $G \cup \{\hat{v}\}$ gives an alternating link such that the reduced white Tait graph is isomorphic to G and the weights of G will become the diagonal entries of Goeritz matrix.

Theorem 5.8 *Let K be an alternating link. Then $\Sigma(S^3, K)$ admits a simply connected negative-definite spin accessible filling if and only if K is special alternating.*

Proof Assume such a filling X exists and it has a chainmail diagram based on an accessible plane graph like G. Let K' be an alternating link with $\widetilde{W}_{K'} = G$. This means that the chainmail Kirby diagram based on G is also a surgery diagram for $\Sigma(S^3, K')$, which means that branched double covers of K and K' are diffeomorphic. By a result of Greene [11, Theorem 1.1], we can deduce that K and K' are mutants. Planar mutation of alternating knots preserves the number of white regions of the diagram and as a result

$$b_2(X) = |V_G| = |V_{\widetilde{W}_{K'}}| = |V_{\widetilde{W}_K}|.$$

Using Theorem 1.2, we can deduce that K is special alternating. □

Acknowledgements

It is a pleasure to thank my advisor, Professor András Juhász, whose patience and guidance made this project possible. I am very grateful to Professor Marco Golla for our helpful discussions and for pointing out Example 4.12. I would also like to express my appreciation to the reviewer for valuable comments and constructive suggestions. The author was supported by the Clarendon Fund and the Soudavar Fund Scholarships during this research.

References

[1] **P Aceto**, **D McCoy**, **J Park**, *Definite fillings of lens spaces*, J. Differential Geom. 131:1 (2025) 1–41 MR

[2] **A Akhmedov**, **B D Park**, *Exotic smooth structures on small* 4*-manifolds with odd signatures*, Invent. Math. 181:3 (2010) 577–603 MR

[3] **D H Choe**, **K Park**, *On intersection forms of definite* 4*-manifolds bounded by a rational homology* 3*-sphere*, Topology Appl. 238 (2018) 59–75 MR

[4] **S K Donaldson**, *An application of gauge theory to four-dimensional topology*, J. Differential Geom. 18:2 (1983) 279–315 MR

[5] **S K Donaldson**, *Connections, cohomology and the intersection forms of* 4*-manifolds*, J. Differential Geom. 24:3 (1986) 275–341 MR

[6] **N D Elkies**, *A characterization of the $\mathbb{Z}^n$ lattice*, Math. Res. Lett. 2:3 (1995) 321–326 MR

[7] **M Furuta**, *Monopole equation and the $\frac{11}{8}$-conjecture*, Math. Res. Lett. 8:3 (2001) 279–291 MR

[8] **M Golla**, **C Scaduto**, *On definite lattices bounded by integer surgeries along knots with slice genus at most* 2, Trans. Amer. Math. Soc. 372:11 (2019) 7805–7829 MR

[9] **R E Gompf**, **A I Stipsicz**, *4-manifolds and Kirby calculus*, Graduate Studies in Mathematics 20, Amer. Math. Soc., Providence, RI (1999) MR

[10] **C M Gordon**, **R A Litherland**, *On the signature of a link*, Invent. Math. 47:1 (1978) 53–69 MR

[11] **J E Greene**, *Lattices, graphs, and Conway mutation*, Invent. Math. 192:3 (2013) 717–750 MR

[12] **J E Greene**, *The lens space realization problem*, Ann. of Math. (2) 177:2 (2013) 449–511 MR

[13] **J E Greene**, *A spanning tree model for the Heegaard Floer homology of a branched double-cover*, J. Topol. 6:2 (2013) 525–567 MR

[14] **J E Greene**, *Alternating links and definite surfaces*, Duke Math. J. 166:11 (2017) 2133–2151 MR

[15] **S J Kaplan**, *Constructing framed 4-manifolds with given almost framed boundaries*, Trans. Amer. Math. Soc. 254 (1979) 237–263 MR

[16] **K Murasugi**, *On the Alexander polynomial of alternating algebraic knots*, J. Austral. Math. Soc. Ser. A 39:3 (1985) 317–333 MR

[17] **W D Neumann**, *A calculus for plumbing applied to the topology of complex surface singularities and degenerating complex curves*, Trans. Amer. Math. Soc. 268:2 (1981) 299–344 MR

[18] **B Owens**, **S Strle**, *Dehn surgeries and negative-definite four-manifolds*, Selecta Math. (N.S.) 18:4 (2012) 839–854 MR

[19] **P Ozsváth**, **Z Szabó**, *Absolutely graded Floer homologies and intersection forms for four-manifolds with boundary*, Adv. Math. 173:2 (2003) 179–261 MR

[20] **P Ozsváth**, **Z Szabó**, *On the Heegaard Floer homology of branched double-covers*, Adv. Math. 194:1 (2005) 1–33 MR

[21] **M Polyak**, *From 3-manifolds to planar graphs and cycle-rooted trees*, lecture slides (2014) https://www.fields.utoronto.ca/programs/scientific/14-15/arnoldconf/slides/polyak.pdf

[22] **C Scaduto**, *On definite lattices bounded by a homology 3-sphere and Yang–Mills instanton Floer theory*, Geom. Topol. 28:4 (2024) 1587–1628 MR

[23] **L Siebenmann**, *On vanishing of the Rohlin invariant and nonfinitely amphicheiral homology 3-spheres*, from "Topology Symposium" (Siegen, 1979) (U Koschorke, W D Neumann, editors), Lecture Notes in Math. 788, Springer (1980) 172–222 MR

[24] **M Ue**, *Constraints on intersection forms of spin 4-manifolds bounded by Seifert rational homology 3-spheres in terms of $\bar{\mu}$ and κ invariants* (2022) arXiv 2206.05412

SOHEIL AZARPENDAR azarpendar@maths.ox.ac.uk
Mathematical Institute, University of Oxford, Oxford, United Kingdom

Received: November 27, 2023 Revised: May 26, 2025

Algebraic & Geometric Topology **26**:5 (2026) 1681–1747
DOI: 10.2140/agt.2026.26.1681
Published: May 29, 2026

On local fibrations of $(\infty, 2)$-categories

FERNANDO ABELLÁN

We provide a model independent notion of local fibrations of $(\infty, 2)$-categories which generalises the well-known theory of locally cocartesian fibrations of $(\infty, 1)$-categories. Based on previous work, we construct a model structure which serves as a specific combinatorial model for this type of fibration. Our main result is a generalisation of the locally cocartesian straightening and unstraightening construction of Lurie which yields for any scaled simplicial set S an equivalence of $(\infty, 2)$-categories between the $(\infty, 2)$-category $\mathbb{F}\mathrm{ib}_{0,1}(S)$ of $(0, 1)$-fibrations over S and the $(\infty, 2)$-category of functors $\mathrm{Fun}(S, \mathbb{B}\mathrm{icat}_\infty)$. Given an $(\infty, 2)$-category $\mathbb{B}$ our Grothendieck construction can be specialised to produce an equivalence between the $(\infty, 2)$-category of local fibrations over $\mathbb{B}$ and the $(\infty, 2)$-category of oplax normalised functors with values in $\mathbb{B}\mathrm{icat}_\infty$. Finally, as an application of our results we provide a version of the Yoneda lemma for $(\infty, 2)$-categories.

1 Introduction

The theory of $(\infty, 2)$-categories is enjoying in recent years a rapid and extensive development. Several fundamental constructions such as Gray tensor products [11], partially lax colimits [5; 8; 10] and a 2-dimensional theory of fibrations [3; 6; 13] are already available and ready to be used in the study of homotopy coherent structures. Even more remarkably, we are starting to see specific examples [9; 17] of how these techniques can be used to categorify existing areas of study and how they can put into perspective known constructions. In this paper, we continue our study of fibrations of $(\infty, 2)$-categories

---

MSC2020: 18N40, 18N65.

and provide an additional piece of technology which will be relevant for future applications: a theory of local fibrations.

In the homotopy-coherent world, fibrations are an essential tool that facilitates the construction of functors in a context where an infinite amount of coherent-data must be specified. More precisely, given an $(\infty,2)$-category $\mathbb{C}$ the so-called "Grothendieck construction" (also known in the higher-categorical world as the straightening-unstraightening equivalence due to Lurie [18]) states that there exists an equivalence between the $(\infty,2)$-category of $(0,1)$-fibrations[1] over $\mathbb{C}$ and the $(\infty,2)$-category of functors $F:\mathbb{C}\to\mathbb{C}\mathrm{at}_{(\infty,2)}$, with values in $(\infty,2)$-categories. This result has been realised in several models [3; 7; 19; 20; 21] and it is even available for general fibrations (and functors) of (∞,n)-categories.

However, in many situations we would like to have a version of the Grothendieck construction that is flexible enough to accommodate the notion of a (op)lax normalised functor, i.e., a functor which only preserves composition up to noninvertible coherent-data but preserves identity morphisms. This is achieved by means of the notion of a local fibration of $(\infty,2)$-categories as demonstrated by our main result.[2]

Theorem *Let $\mathbb{S}$ be an $(\infty,2)$-category. Then the straightening-unstraightening adjunction*

$$\mathbb{S}\mathrm{t}^{\mathrm{oplax}}_{\mathbb{S}}:\mathbb{L}\mathrm{Fib}(\mathbb{S})\rightleftarrows \mathrm{Fun}^{\mathrm{oplax}}(\mathbb{S},\mathbb{C}\mathrm{at}_{(\infty,2)}):\mathbb{U}\mathrm{n}^{\mathrm{oplax}}_{\mathbb{S}}$$

yields an equivalence of $(\infty,2)$-categories between the $(\infty,2)$-category of local $(0,1)$-fibrations over $\mathbb{S}$ and the $(\infty,2)$-category of oplax normalised functors with values in $(\infty,2)$-categories.

This result had already been proved for functors with values in $(\infty,1)$-categories by Lurie [19] and in full generality (in the model of 2-fold Segal spaces) by Ayala, Mazel-Gee and Rozenblyum [7]. Our theorem can be seen as a direct extension of Lurie's result. The main motivation for proving the general version of this result in the scaled simplicial model is due to the tractable description that the Gray tensor product admits in this setting [11] which can be used to provide further versions of the Grothendieck construction which are adapted to handle (op)lax natural transformations between functors, as seen in [2]. In the aforementioned work, we extensively exploit this result to produce a general calculus of mates, as well as providing formulas for computing partially lax (co)limits of functors with values in $\mathbb{C}\mathrm{at}_{(\infty,2)}$.

Another advantage of our construction in the scaled simplicial model is that we have full control over the "laxness" of our functors. This is materialised in the following result.

Theorem *Let $\mathbb{S}$ be an $(\infty,2)$-category and let $\mathcal{U}=\{(f_i,g_i)\}_{i\in I}$ be a collection of composable morphisms in $\mathbb{S}$. Then the straightening-unstraightening adjunction*

$$\mathbb{S}\mathrm{t}^{\mathcal{U}\mathrm{oplax}}_{\mathbb{S}}:\mathbb{L}\mathrm{Fib}^{\mathcal{U}}(\mathbb{S})\rightleftarrows \mathrm{Fun}^{\mathcal{U}\mathrm{oplax}}(\mathbb{S},\mathbb{C}\mathrm{at}_{(\infty,2)}):\mathbb{U}\mathrm{n}^{\mathcal{U}\mathrm{oplax}}_{\mathbb{S}}$$

yields an equivalence of $(\infty,2)$-categories between the $(\infty,2)$-category of $\mathcal{U}$-local $(0,1)$-fibrations over $\mathbb{S}$ and the $(\infty,2)$-category consisting in those oplax normalised functors which preserve the composites in $\mathcal{U}$.

---

[1] By an (i,j)-fibration we mean a 2-categorical version of the notion of a (co)cartesian fibration which we will define later in the introduction.

[2] For more detailed statements and references to the relevant theorems, see the following subsection in the introduction.

When we specialise to the case where $\mathcal{U}$ is precisely the collection of *all* pairs of composable morphisms the definition a $\mathcal{U}$-local $(0,1)$-fibration collapses to that of an ordinary $(0,1)$-fibration and we recover the usual Grothendieck construction for $(\infty,2)$-categories.

Local fibrations of $(\infty,2)$-categories

Let $p:\mathbb{X}\to\mathbb{S}$ be a functor of $(\infty,2)$-categories. We say that p is a (i,j)-fibration where $i,j\in\{0,1\}$ if:

(F1) For every $a,b\in\mathbb{X}$ the induced map $\mathbb{X}(a,b)\to\mathbb{S}(p(a),p(b))$ on mapping $(\infty,1)$-categories is a cocartesian fibration if $j=0$ or a cartesian fibration if $j=1$.

(F2) For every $a,b,c\in\mathbb{X}$ the composition functors

$$\mathbb{X}(a,b)\times\mathbb{X}(b,c)\to\mathbb{X}(a,c)$$

preserve cocartesian edges if $j=0$ (resp. cartesian edges if $j=1$).

(C1) Let $i=0$. Given an object $a\in\mathbb{X}$ and a morphism $e:p(a)\to y$ in $\mathbb{S}$, there exists an edge $\hat{e}:a\to\hat{y}$ over e with the following property: for every $z\in\mathbb{X}$ precomposition with $\hat{e}$ induces a pullback of $(\infty,1)$-categories

$$\begin{tikzcd}
\mathbb{X}(\hat{y},z) \arrow[r] \arrow[d] & \mathbb{X}(a,z) \arrow[d] \\
\mathbb{S}(y,p(z)) \arrow[r] & \mathbb{S}(p(a),p(z))
\end{tikzcd}$$

We say that $\hat{e}$ is a $(0,j)$-cartesian lift of e. If $i=1$ one defines a dual condition (which generalises the $(\infty,1)$-notion of cartesian edge) and obtains the definition of a $(1,j)$-cartesian edge.

Let us remind the reader that $(0,j)$-fibrations appear in the literature [6; 13] under the name of outer cocartesian when $j=0$ (resp. inner cocartesian if $j=1$) and similarly $(1,j)$-fibrations are called outer cartesian fibrations if $j=0$ and inner cartesian fibrations if $j=1$.

In [3; 6] we gave a systematic analysis of the theory of $(1,0)$-fibrations and provided the corresponding Grothendieck construction which identifies this kind of fibrations with contravariant functors $\mathcal{F}:\mathbb{S}^{\mathrm{op}}\to\mathbb{C}\mathrm{at}_{(\infty,2)}$. However, the models employed to study these fibrations in the aforementioned works do not generalise the theory of locally cocartesian fibrations developed in [18; 19]. In order to realise the local theory of fibrations in the scaled simplicial model [19], we will be working in this paper with the $(0,1)$-variance. Needless to say, in model independent terms these problems do not arise and we can give the following general definition.

Let $p:\mathbb{X}\to\mathbb{S}$ be a functor of $(\infty,2)$-categories. We say that p is a local (i,j)-fibration if these hold:

- Conditions (F1) and (F2) are satisfied. In this case, we say that p is *cocartesian*-enriched if $j=0$ and *cartesian*-enriched if $j=1$.
- For every morphism $e:\Delta^1\to\mathbb{S}$ the pullback $\mathbb{X}\times_{\mathbb{S}}\{e\}\to\Delta^1$ is an (i,j)-fibration. We say that an edge $\hat{e}:\Delta^1\to\mathbb{X}$ is a local $(0,1)$-cartesian edge if it is $(0,1)$-cartesian after taking the corresponding pullback.

In order to access the model-independent results mentioned above, we resort to the model of marked-biscaled simplicial sets constructed in [6], which provides us with a robust combinatorial framework to model local $(0,1)$-fibrations and to implement the desired Grothendieck construction. A marked-biscaled simplicial set denoted as $(X, E, T_X \subseteq C_X)$ consists of a simplicial set X, together with a subset of edges $E \subset X_1$ containing all degenerate 1-simplices, and two collections of 2-simplices (or triangles), $T_X \subseteq C_X \subseteq X_2$, such that T_X (and hence C_X) contains every degenerate 2-simplex. The collection E is used to model the cocartesian 1-morphisms of our fibrations, while C_X models the cartesian 2-morphisms. The subcollection T_X is used to represent the invertible 2-morphisms which are always cartesian. Equipped with this formalism we establish as the first result in this paper the existence a model structure whose fibrant objects we refer to as U-local $(0,1)$-fibrations over a scaled simplicial set (S, U) (see Theorems 3.25 and 3.34).

Theorem 1 *Let (S, U) be a scaled simplicial set. Then there exists a left proper, combinatorial, simplicial model structure*[3] *on $\left(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$, which is characterised uniquely by the following properties:*

(C) *A morphism $f : X \to Y$ is a cofibration if and only if f induces a monomorphism on the underlying simplicial sets.*

(F) *An object $p : X \to S$ is fibrant if and only if it is a U-local $(0,1)$-fibration.*

Moreover, if $S = \Delta^0$ then this model structure is Quillen equivalent to Lurie's model structure (see Theorem 4.27 in [19]) on $\mathrm{Set}_{\Delta}^{\mathbf{sc}}$, the category of scaled simplicial sets

In the case where $\mathbb{S} = (S, T_S)$ is a fibrant scaled simplicial set and thus models an $(\infty,2)$-category (here T_S consists in those 2-simplices representing commuting triangles in $\mathbb{S}$), we can consider a subcollection of triangles $M_S \subseteq T_S$ depicted visually as

$$\begin{tikzcd} & b \arrow[dr, "v"] & \\ a \arrow[ur, "u"] \arrow[rr] & {} \arrow[u, Rightarrow, "\simeq"'] & c \end{tikzcd}$$

where either u or v are equivalences. In Theorem 4.22, we specialise the previous result to (S, M_S) to obtain a model independent interpretation of the theory of M_S-local (or simply local) $(0,1)$-fibrations over (S, M_S). In order to achieve this, we construct (see Definition 4.4) from every M_S-local $(0,1)$-fibration $p : (X, E_X, T_X \subseteq C_X) \to (S, \sharp, M_S \subseteq \sharp)$ a fibration among fibrant scaled simplicial sets $\mathfrak{p} : \mathbb{X} \to \mathbb{S}$ satisfying (F1), (F2) and (C1) above which we call the *bicategorical interpretation* of p.

Theorem 2 *Let (S, T_S) be a fibrant scaled simplicial set and let $p : (X, E_X, T_X \subseteq C_X) \to (S, \sharp, M_S \subseteq \sharp)$ be an object of $\left(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\right)_{/(S,\sharp,M_S\subset\sharp)}$. Then p defines a fibrant object if and only if its bicategorical interpretation $\mathfrak{p} : \mathbb{X} \to \mathbb{S}$ is a local $(0,1)$-fibration of $(\infty,2)$-categories.*

---

[3] Here "$\sharp$" denotes the maximal collection consisting in *all* 1-simplices (or all 2-simplices).

It is well known (see Proposition 2.4.2.8 in [18]) that a locally cocartesian fibration whose locally cocartesian edges compose must be a cocartesian fibration. We extend this analysis to the $(\infty,2)$-categorical case by considering a fibrant scaled simplicial set (S,T_S) and a collection of triangles $M_S \subseteq U \subseteq T_S$ which allows us to make the following definition:

- A local $(0,1)$-fibration $p:\mathbb{X}\to\mathbb{S}$ is said to be $\mathcal{U}$-local if local $(0,1)$-cartesian edges compose along triangles lying over U.

It then follows from Theorem 4.29 that $\mathcal{U}$-local fibrations can also be characterised as fibrant objects in our model structure.

Theorem 3 *Let* (S,T_S) *be a fibrant scaled simplicial and let* U *be a collection of triangles such that* $M_S \subseteq U \subseteq T_S$*. Then an object* $p:X\to S$ *is fibrant in* $\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$ *if and only if its bicategorical interpretation* $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ *is a local* $(0,1)$*-fibration of* $(\infty,2)$*-categories which is in addition* $\mathcal{U}$*-local.*

Once the basics of the local theory of fibrations of $(\infty,2)$-categories are established we focus our attention into providing the expected Grothendieck construction which will allow us to interpret our fibrations as functors with values in $\mathbb{B}\mathrm{icat}_\infty$ the ∞-bicategory (i.e., fibrant scaled simplicial set) of ∞-bicategories.

Let $(S,U)=S_U$ be a scaled simplicial and denote by $\mathbb{F}\mathrm{ib}_{0,1}(S_U)$ the ∞-bicategory of U-local $(0,1)$-fibrations over S and by $\mathrm{Fun}(S_U,\mathbb{B}\mathrm{icat}_\infty)$ the functor ∞-bicategory. Our main construction is a generalisation of the straightening-unstraightening equivalence of Lurie [19] to the setting of $(0,1)$-fibrations whose fibres are ∞-bicategories. Combining Theorem 5.52 and Remark 5.54 we obtain:

Theorem 4 *Let* $S_U=(S,U)$ *be a scaled simplicial set. Then the straightening-unstraightening adjunction*

$$\mathbb{S}\mathrm{t}_{S_U}:\mathbb{F}\mathrm{ib}_{0,1}(S_U)\rightleftarrows \mathrm{Fun}(S_U,\mathbb{B}\mathrm{icat}_\infty):\mathbb{U}\mathrm{n}_{S_U}$$

yields an equivalence of ∞*-bicategories between the* ∞*-bicategory of* U*-local* $(0,1)$*-fibrations over* S *and the* ∞*-bicategory of covariant functors with values in* ∞*-bicategories.*

The general nature of the theorem above allows us to consider special cases which are of special interest. Namely, let $\mathbb{S}=(S,T_S)$ be an ∞-bicategory and let $S_{M_S}=(S,M_S)$. We can now consider $\mathbb{L}\mathrm{Fib}(\mathbb{S}):=\mathbb{F}\mathrm{ib}_{0,1}(S_{M_S})$ and apply our Grothendieck construction to obtain an equivalence between the ∞-bicategory of local $(0,1)$-fibrations over $\mathbb{S}$ and the category $\mathrm{Fun}(S_{M_S},\mathbb{B}\mathrm{icat}_\infty)$. The latter can be interpreted using the work of Gagna, Harpaz and Lanari [11] as a model for the ∞-bicategory of oplax normalised functors with values in ∞-bicategories. Our main theorem then specialises (Corollary 5.57) to the results presented at the beginning of the introduction.

The ∞-bicategorical Yoneda lemma

As an application of our results we give a fibrational proof of the Yoneda lemma for ∞-bicategories. We would like to stress that such a result was already present in the work of Hinich [16] where a general Yoneda lemma for enriched ∞-categories is established.

In this work we provide a proof of the Yoneda lemma as a direct application of the Grothendieck construction. Given an ∞-bicategory $\mathbb{C}$ we consider the $(0,1)$-fibration $\mathrm{ev}_1 : \mathrm{Fun}^{\mathrm{opgr}}(\Delta^1, \mathbb{C}) \to \mathbb{C}$ (see Proposition 2.2.6 in [13] for more details) which is sometimes referred as the oplax arrow category of $\mathbb{C}$. Carefully unwinding this construction reveals that the fibres of this map come equipped with maps

$$\mathbb{C}_{\nearrow c} = \mathrm{Fun}^{\mathrm{opgr}}(\Delta^1, \mathbb{C}) \times_{\mathbb{C}} \{c\} \to \mathbb{C},$$

which are in turn $(1,0)$-fibrations with ∞-categorical fibres. Applying the Grothendieck construction in two steps we obtain a functor $\mathcal{Y}_{\mathbb{C}} : \mathbb{C} \to \mathrm{Fun}(\mathbb{C}^{\mathrm{op}}, \mathbb{C}\mathrm{at}_\infty)$. Moreover, it follows from previous work (Theorem 3.17 in [4]) that $\mathbb{C}_{\nearrow c}$ corresponds under the Grothendieck construction to the representable functors $\mathbb{C}(-,c)$. We prove in Theorem 6.16 the final result of this paper.

Theorem 5 *For every ∞-bicategory $\mathbb{C}$ the Yoneda embedding*

$$\mathcal{Y}_{\mathbb{C}} : \mathbb{C} \to \mathrm{Fun}(\mathbb{C}^{\mathrm{op}}, \mathbb{C}\mathrm{at}_\infty), \quad c \mapsto \mathbb{C}(-,c),$$

is fully faithful. Moreover, given a functor $\mathcal{F} : \mathbb{C}^{\mathrm{op}} \to \mathbb{C}\mathrm{at}_\infty$ there is a equivalence of $\mathbb{C}\mathrm{at}_\infty$-valued functors

$$\mathrm{Nat}_{\mathbb{C}^{\mathrm{op}}}(\mathcal{Y}_{\mathbb{C}}(-), \mathcal{F}) \xRightarrow{\simeq} \mathcal{F},$$

(*where* $\mathrm{Nat}_{\mathbb{C}^{\mathrm{op}}}(-,-)$ *is the mapping ∞-category in* $\mathrm{Fun}(\mathbb{C}^{\mathrm{op}}, \mathbb{C}\mathrm{at}_\infty)$) *which is natural in $\mathcal{F}$.*

2 Preliminaries

In this section we will mainly gather the main definitions of the theory of scaled simplicial sets as presented by Lurie [19].

Definition 2.1 A scaled simplicial set (X, T_X) consists in a simplicial set X together with a collection of 2-simplices (also called triangles) T_X which contains *every degenerate* triangle. We call the elements of T_X the *thin triangles* of X. A morphism of scaled simplicial sets $f : (X, T_X) \to (Y, T_Y)$ is a map of simplicial sets $f : X \to Y$ such that $f(T_X) \subseteq T_Y$. We denote the corresponding category of scaled simplicial sets by $\mathrm{Set}_\Delta^{\mathbf{sc}}$.

Notation 2.2 Given a simplicial set A we have two canonical ways of viewing it as a scaled simplicial set:

- We define $A_\flat = (A, \flat)$ where $\flat$ is the collection consisting only in the degenerate triangles of A.
- We define $A_\sharp = (A, \sharp)$ where $\sharp$ is the collection consisting in every triangle of A.

Definition 2.3 The set of *generating scaled anodyne maps* $\mathbf{S}$ is the set of maps of scaled simplicial sets consisting of

(i) the inner horns inclusions

$$\big(\Lambda^n_i, \{\Delta^{\{i-1,i,i+1\}}\}\big) \to \big(\Delta^n, \{\Delta^{\{i-1,i,i+1\}}\}\big), \quad n \geq 2,\ 0 < i < n;$$

(ii) the map

$$(\Delta^4,T)\to(\Delta^4,T\cup\{\Delta^{\{0,3,4\}},\Delta^{\{0,1,4\}}\}),$$

where we define

$$T\overset{\mathrm{def}}{=}\{\Delta^{\{0,2,4\}},\Delta^{\{1,2,3\}},\Delta^{\{0,1,3\}},\Delta^{\{1,3,4\}},\Delta^{\{0,1,2\}}\};$$

(iii) the set of maps

$$(\Lambda^n_0\amalg_{\Delta^{\{0,1\}}}\Delta^0,\{\Delta^{\{0,1,n\}}\})\to(\Delta^n\amalg_{\Delta^{\{0,1\}}}\Delta^0,\{\Delta^{\{0,1,n\}}\}),\quad n\geqslant 3.$$

A general map of a scaled simplicial set is said to be *scaled anodyne* if it belongs to the weakly saturated closure of **S**.

Definition 2.4 A scaled simplicial set (X,T_X) is said to be an ∞-bicategory if it has the right lifting property against the class of scaled anodyne maps in Definition 2.3. In this case, we view the 2-simplices of T_X as the collection of commuting triangles.

Definition 2.5 Given an ∞-bicategory (X,T_X) we can construct an ∞-category $X^{\leqslant 1}$ by considering the subsimplicial set of X given by those n-simplices $\sigma:\Delta^n\to X$ such that each 2-dimensional face belongs to T_X. We call $X^{\leqslant 1}$ the underlying ∞-category of (X,T_X). We similarly define $X^{\simeq}$ as the underlying ∞-groupoid of $X^{\leqslant 1}$.

Definition 2.6 We denote by Cat^+_Δ the category of Set^+_Δ-enriched categories (i.e., categories enriched in marked simplicial sets). We note that we can view the category of (strict) 2-categories $2\,\mathrm{Cat}$ as a full subcategory of Cat^+_Δ by applying the nerve functor Hom-wise and marking the equivalences in each mapping category.

Proposition 2.7 *There is a left-proper, combinatorial model structure on* Cat^+_Δ *such that:*

(W) *The weak equivalences are those enriched functors which are essentially surjective on homotopy categories and induce weak equivalences on all mapping marked simplicial sets.*[4]

(C) *The cofibrations are the smallest weakly saturated class containing* $\varnothing\to[0]_{\mathrm{Set}^+_\Delta}$, *and each inclusion* $[1]_A\to[1]_B$ *where* $A\to B$ *is a generating cofibration for* Set^+_Δ.

Proof This is a special case of [18, A.3.2.4]. □

Definition 2.8 Let I be a linearly ordered finite set. We define a 2-category $\mathbb{O}^I$ as follows:

- The objects of $\mathbb{O}^I$ are the elements of I.
- The category $\mathbb{O}^I(i,j)$ of morphisms between objects $i,j\in I$ is defined as the poset of finite sets $S\subseteq I$ such that $\min(S)=i$ and $\max(S)=j$ ordered by inclusion.
- The composition functors are given, for $i,j,l\in I$, by

$$\mathbb{O}^I(i,j)\times\mathbb{O}^I(j,l)\to\mathbb{O}^I(i,l),\quad (S,T)\mapsto S\cup T.$$

---

[4] See [18, Section 3.1] for a model structure on marked simplicial sets.

When $I = [n]$, we denote $\mathbb{O}^I$ by $\mathbb{O}^n$. Note that the $\mathbb{O}^n$ form a cosimplicial object in 2Cat, which we denote by $\mathbb{O}^\bullet$.

Definition 2.9 The map

$$\Delta \xrightarrow{\mathfrak{C}} \mathrm{Cat}_\Delta^+,$$

which sends $[n]$ to $\mathbb{O}^n$ gives us a cosimplicial object in Cat_Δ^+. We can moreover send the thin 2-simplex $\Delta^2_\sharp$ to $\mathfrak{C}[\Delta^2]$ equipped with maximally marked mapping spaces. The usual machinery of nerve and realisation then gives us adjoint functors

$$\mathfrak{C}^{\mathrm{sc}} : \mathrm{Set}_\Delta^{\mathbf{sc}} \rightleftarrows \mathrm{Cat}_\Delta^+ : \mathrm{N}^{\mathbf{sc}},$$

which we will call the *scaled nerve* and *scaled rigidification*.

Theorem 2.10 *There is a left proper, combinatorial model structure on* $\mathrm{Set}_\Delta^{\mathbf{sc}}$ *in which:*

- (W) *The weak equivalences are the morphisms* $f : A \to B$ *such that* $\mathfrak{C}^{\mathrm{sc}}[f] : \mathfrak{C}^{\mathrm{sc}}[A] \to \mathfrak{C}^{\mathrm{sc}}[B]$ *is an equivalence in* Cat_Δ^+.
- (C) *The cofibrations are the monomorphisms.*

Moreover, the fibrant objects in this model structure are the ∞*-bicategories, and the adjunction*

$$\mathfrak{C}^{\mathrm{sc}} : \mathrm{Set}_\Delta^{\mathbf{sc}} \rightleftarrows \mathrm{Cat}_\Delta^+ : \mathrm{N}^{\mathbf{sc}}$$

is a Quillen equivalence.

Proof This is [19, Theorem A.3.2.4]. The characterisation of fibrant objects is [12, Theorem 5.1]. □

Definition 2.11 We say that a map of scaled simplicial sets is a *bicategorical equivalence* if it is a weak equivalence in the model structure given in Theorem 2.10. Similarly, call the fibrations in the model structure of scaled simplicial sets *bicategorical fibrations*.

Remark 2.12 In [12], the authors characterise the model structure on scaled simplicial sets as a Cisinski–Olschok model structure. This in turn implies that a map between fibrant scaled simplicial sets is a bicategorical fibration if and only if it is an isofibration and it has the right lifting property against the class of scaled anodyne maps.

Definition 2.13 Given a pair of scaled simplicial sets X, Y we denote by $\mathrm{Fun}(X, Y)$ the scaled simplicial set determined by the universal property

$$\mathrm{Hom}_{\mathrm{Set}_\Delta^{\mathbf{sc}}}(K, \mathrm{Fun}(X, Y)) \simeq \mathrm{Hom}_{\mathrm{Set}_\Delta^{\mathbf{sc}}}(K \times X, Y),$$

where $K \times X$ denotes the cartesian product of scaled simplicial sets.

Definition 2.14 Let $\mathbb{C}$ be an ∞-bicategory. Given an object $y \in \mathbb{C}$ we define a scaled simplicial set $\mathbb{C}_{\nearrow y}$ as follows. The data of an n-simplices $\Delta^n \to \mathbb{C}_{\nearrow y}$ is given by a map $\sigma : \Delta^{n+1} \to \mathbb{C}$ such that $\sigma(n+1) = y$. The inclusion $d_{n+1} : \Delta^n \to \Delta^{n+1}$ induces a map

$$\pi : \mathbb{C}_{\nearrow y} \to \mathbb{C},$$

which we use to declare a triangle in $\mathbb{C}_{\nearrow y}$ to be thin if and only if its image under π is thin in $\mathbb{C}$. It follows from [12, Proposition 2.33] that the fibre of π at an object $x \in \mathbb{C}$ is a model for $\mathbb{C}(x, y)$, the mapping ∞-category.

3 The model structure

Definition 3.1 A *marked biscaled* simplicial set (**MB** simplicial set) is given by

- a simplicial set X,
- a collection of edges $E_X \in X_1$ containing all degenerate edges,
- a collection of triangles $T_X \in X_2$ containing all degenerate triangles,
- a collection of triangles $C_X \in X_2$ such that $T_X \subseteq C_X$,

where we will refer to the elements of T_X as *thin triangles*, and we will refer to the elements of C_X as *lean triangles*. We will denote such objects as triples $(X, E_X, T_X \subseteq C_X)$. A map $(X, E_X, T_X \subseteq C_X) \to (Y, E_Y, T_Y \subseteq C_Y)$ is given by a map of simplicial sets $f : X \to Y$ compatible with the collections of edges and triangles above. We denote by $\mathrm{Set}_\Delta^{\mathbf{mb}}$ the category of **MB** simplicial sets.

Notation 3.2 Let $(X, E_X, T_X \subseteq C_X)$ be an **MB** simplicial set. Suppose that the collection E_X consists only of degenerate edges. Then we fix the notation $(X, E_X, T_X \subseteq C_X) = (X, \flat, T_X \subseteq E_X)$ and do similarly for T_X. If C_X consists only of degenerate triangles we fix the notation $(X, E_X, T_X \subseteq C_X) = (X, E_X, \flat)$. In an analogous fashion we will use the symbol "$\sharp$" to denote a collection containing all edges (or all triangles). Finally if $T_X = C_X$ then we will employ the notation (X, E_X, T_X).

Remark 3.3 We will often abuse notation when defining the collections E_X (resp. T_X or C_X) and just specify its nondegenerate edges (resp. triangles).

Definition 3.4 The set of *generating marked-biscaled-anodyne maps* **MB** is the set of maps of **MB** simplicial sets consisting of

(A1) the inner horn inclusions

$$(\Lambda^n_i, \flat, \flat \subset \{\Delta^{\{i-1,i,i+1\}}\}) \to (\Delta^n, \flat, \flat \subset \{\Delta^{\{i-1,i,i+1\}}\}), \quad n \geqslant 2,\ 0 < i < n;$$

(A2) the map

$$(\Delta^4, \flat, \flat \subset T) \to (\Delta^4, \flat, \flat \subset T \cup \{\Delta^{\{0,3,4\}}, \Delta^{\{0,1,4\}}\}),$$

where we define

$$T \overset{\text{def}}{=} \{\Delta^{\{0,2,4\}}, \Delta^{\{1,2,3\}}, \Delta^{\{0,1,3\}}, \Delta^{\{1,3,4\}}, \Delta^{\{0,1,2\}}\};$$

(A3) the set of maps

$$(\Lambda^n_0, \{\Delta^{\{0,1\}}\}, \{\Delta^{\{0,1,n\}}\}) \to (\Delta^n, \{\Delta^{\{0,1\}}\}, \{\Delta^{\{0,1,n\}}\}), \quad n \geqslant 2;$$

(A4) the inclusion of the initial vertex

$$(\Delta^0, \sharp, \sharp) \to (\Delta^1, \sharp, \sharp);$$

(S1) the map

$$(\Delta^2, \{\Delta^{\{0,1\}}, \Delta^{\{1,2\}}\}, \sharp) \to (\Delta^2, \sharp, \sharp);$$

(S2) the map

$$(\Delta^2, \flat, \flat \subset \sharp) \to (\Delta^2, \flat, \sharp);$$

(E) for every Kan complex K, the map

$$(K, \flat, \sharp) \to (K, \sharp, \sharp).$$

A map of **MB** simplicial sets is said to be **MB**-anodyne if it belongs to the weakly saturated closure of **MB**.

Remark 3.5 Let $p : (X, E_X, T_X \subseteq C_X) \to (S, E_S, T_S \subseteq C_S)$ be a morphism of **MB** simplicial sets. We will informally think of the collection of E_X as representing p-cocartesian edges. The collection C_X is understood to represent cartesian 2-morphisms, while the collection T_X is simply interpreted as representing commuting triangles. Equipped with this intuition let us clarify the meaning of our anodyne maps:

(A1) Having the right lifting property against this class guarantees the existence of enough cartesian 2-morphisms.

(A2) This class expresses a saturation property of cartesian 2-morphisms.

(A3) This class guarantees the marked morphisms to be p-cocartesian with respect to the given thin triangles.

(A4) This guarantees the existence of enough p-cocartesian lifts.

(S1) This class enforces that p-cocartesian morphisms compose across thin triangles.

(S2) This expresses that lean triangles lying over thin triangles are themselves thin. In other words, a cartesian 2-morphism lying over an invertible 2-morphism is itself invertible.

(E) This class simply expresses that equivalences must always be marked.

Remark 3.6 In [6] we also introduced the notion of a **MB**-anodyne map when dealing with the theory of $(1,0)$-fibrations. We would like to point out that both classes of maps are different but we are using the same name to avoid the overly cumbersome notation $\mathbf{MB}^{01}$-anodyne.

Definition 3.7 A map of **MB** simplicial sets is said to be an **MB**-*fibration* if it has the right lifting property against the class of **MB**-anodyne maps.

Lemma 3.8 *Let* $p : X \to S$ *be an* ***MB****-fibration. Then for every* $s \in S$ *the fibre, over* s,

$$\begin{tikzcd}
X_s \arrow[r] \arrow[d] & X \arrow[d, "p"] \\
\Delta^0 \arrow[r, "s"] & S
\end{tikzcd}$$

is of the form (X_s, E_s, T_s) *(see Notation 3.2), where* (X_s, T_s) *is an* ∞*-bicategory and where* E_s *is precisely given by the equivalences.*

Proof Observe that in X_s the lean and thin triangles coincide since in a **MB**-fibration a lean triangle lying over a thin triangle is itself thin. It follows that X_s has the right lifting property against the class of scaled anodyne maps and thus it is an ∞-bicategory. Note that by definition the equivalences must be marked in X_s. Moreover, since X_s lifts against the class of maps (A3) one checks easily that marked morphisms are equivalences. □

Lemma 3.9 *The morphism of* ***MB****-simplicial sets* $(\Delta^2, \{\Delta^{\{0,1\}}, \Delta^{\{0,2\}}\}, \sharp) \to (\Delta^2, \sharp, \sharp)$ *is* ***MB****-anodyne.*

Proof The proof is dual to [6, Lemma 3.11]. □

Lemma 3.10 *The morphism of* ***MB*** *simplicial sets*

$$\iota : (\Delta^3, \Delta^{\{0,1\}}, \{\Delta^{\{0,1,3\}}\} \subset U_0) \to (\Delta^3, \Delta^{\{0,1\}}, \{\Delta^{\{0,1,3\}}\} \subset \sharp)$$

is ***MB****-anodyne where* U_0 *is the collection of all* 2*-faces except* $\Delta^{\{1,2,3\}}$.

Proof Let $S = (S, E_S, T_S \subset \sharp)$ be an **MB** simplicial set and let $p : X \to S$ be an **MB**-fibration. We will show that p has the right lifting property against the map ι. Once this claim is established we will factor ι as an **MB**-anodyne morphism followed by an **MB**-fibration where B is of the form $(S, E_S, T_S \subset \sharp)$, that is,

$$\iota : (\Delta^3, \Delta^{\{0,1\}}, \{\Delta^{\{0,1,3\}}\} \subset U_0) := A \xrightarrow{\alpha} X \xrightarrow{p} (\Delta^3, \Delta^{\{0,1\}}, \{\Delta^{\{0,1,n\}}\} \subset \sharp) = B.$$

It will then follow that we can produce a solution to the lifting problem

$$\begin{tikzcd}
A \arrow[r] \arrow[d, "\iota"] & X \arrow[d] \\
B \arrow[r, "\mathrm{id}_B"] \arrow[ur, dotted] & B
\end{tikzcd}$$

which exhibits ι as a retract of the **MB**-anodyne map α and thus concluding the proof.

In order to complete the proof we must show the claim. Suppose we are given a lifting problem

$$\begin{tikzcd}
A \arrow[r, "\sigma"] \arrow[d, "\iota"] & X \arrow[d, "p"] \\
B \arrow[r] & S
\end{tikzcd}$$

Let $\sigma(1 \to 2) = u$, $\sigma(2 \to 3) = v$ and $\sigma(1 \to 3) = \omega$. Since p is an **MB**-fibration we can solve the lifting problem

$$\begin{tikzcd}
\Lambda^2_1 \arrow[r] \arrow[d] & X \arrow[d] \\
\Delta^2 \arrow[r, "d_0(p(\sigma))"'] \arrow[ur, dotted, "\varphi"] & S
\end{tikzcd}$$

and produce a lean triangle φ. We consider a subsimplicial set of $Q \subset \Delta^4$ consisting in the

- the face missing the vertex 2,
- the face missing the vertex 4,
- the 2-dimensional face $\Delta^{\{2,3,4\}}$.

We then produce a map $\theta : Q \to X$ as follows:

- We map the face missing the vertex 2 via σ.
- We map the face missing the vertex 4 via $s_1(d_3(\sigma))$.
- We map $\Delta^{\{2,3,4\}}$ via φ.

We equip Q with the induced decorations. It follows that we can extend Q to a map $\Xi : \Delta^4 \to X$ lying over $s_1(p(\sigma))$. Since p has the right lifting property against the morphism (A2) we see that $d_0(\sigma)$ is lean if and only if the image of $\Delta^{\{1,2,4\}}$ under Ξ is thin in X.

We observe that in $d_1(\Xi)$ every face is lean scaled except possibly the face missing the vertex 2. Again, as a consequence of (A2) it follows that this face must also be lean scaled. Moreover this face lies over a thin simplex of S so it must be itself thin. From now on we can focus our attention to $d_3(\Xi)$.

In $d_3(\Xi) = \rho$ every face is thin scaled except possibly the face missing the vertex 0 and the edge $0 \to 1$ is marked. One checks easily that the pullback of p along the simplex thin $d_2(p(\sigma))$ simplex yields a fibration of ∞-bicategories $X' \to (\Delta^2, \sharp)$ where we can identify the image of $d_0(\rho)$ with a morphism in the mapping ∞-category of X'. One easily shows that this morphism is an equivalence and thus it must be thin. □

Definition 3.11 We say that a map of **MB** simplicial sets is a cofibration if its underlying map of simplicial sets is a monomorphism. One can easily verify that the class of cofibrations is generated by

(C1) the boundary inclusions $(\partial\Delta^n, \flat, \flat) \to (\Delta^n, \flat, \flat)$ for $n \geqslant 0$,
(C2) the map $(\Delta^1, \flat, \flat) \to (\Delta^1, \sharp, \flat)$,
(C3) the map $(\Delta^2, \flat, \flat) \to (\Delta^2, \flat, \flat \subset \sharp)$,
(C4) the map $(\Delta^2, \flat, \flat \subset \sharp) \to (\Delta^2, \flat, \sharp)$.

Proposition 3.12 *Let $f : (X, E_X, T_X \subseteq C_X) \to (Y, E_Y, T_Y \subset C_Y)$ be a cofibration of **MB** simplicial sets and $g : (A, E_A, T_A \subseteq C_A) \to (B, E_B, C_B \subseteq T_B)$ be an **MB**-anodyne morphism. Then the pushout-product*

$$f \wedge g : X \times B \amalg_{X \times A} A \times Y \longrightarrow Y \times B$$

*is again **MB**-anodyne.*

Proof The proof is almost identical to the proof of [6, Proposition 3.14] and left as an exercise. $\square$

Remark 3.13 Observe that given a pair of **MB** simplicial sets X, Y we can produce a functor **MB** simplicial set $\mathrm{Fun}^{\mathbf{mb}}(X,Y)$ in an obvious way via the isomorphism

$$\mathrm{Hom}(A, \mathrm{Fun}^{\mathbf{mb}}(X,Y)) \simeq \mathrm{Hom}(A \times X, Y),$$

where Hom denotes the mapping set in the category $\mathrm{Set}_{\Delta}^{\mathbf{mb}}$.

Corollary 3.14 *Let* $p : Y \to S$ *be an* ***MB****-fibration. Then for every* ***MB*** *simplicial set* X *the induced map*

$$\mathrm{Fun}^{\mathbf{mb}}(X,Y) \longrightarrow \mathrm{Fun}^{\mathbf{mb}}(X,S)$$

is an ***MB****-fibration.*

Proof It follows from Proposition 3.12 after looking at the adjoint lifting problems. $\square$

Definition 3.15 Let $p : Y \to S$ be an **MB**-fibration and consider a map of **MB** simplicial sets $q : X \to S$. We define an ∞-bicategory of functors over S as the pullback

$$\begin{tikzcd}
\mathrm{Map}_S(X,Y) \arrow[r] \arrow[d] & \mathrm{Fun}^{\mathbf{mb}}(X,Y) \arrow[d] \\
\Delta^0 \arrow[r, "q"] & \mathrm{Fun}^{\mathbf{mb}}(X,S)
\end{tikzcd}$$

Definition 3.16 Given a scaled simplicial set (S,U) we define the category $\left(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$ of **MB** simplicial sets over $(S,\sharp,U \subset \sharp)$ as follows:

- The objects are maps $p : (X, E_X, T_X \subset C_X) \to (S, \sharp, U \subset \sharp)$.
- A morphism from $p : (X, E_X, T_X \subset C_X) \to (S, \sharp, U \subset \sharp)$ to $q : (Y, E_Y, T_Y \subset C_Y) \to (S, \sharp, U \subset \sharp)$ is given by a map $f : (X, E_X, T_X \subseteq C_X) \to (Y, E_Y, T_Y \subset C_Y)$ such that $q \circ f = p$.

An object of $\left(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$ is said to be a U-local $(0,1)$-fibration if the corresponding map of **MB** simplicial sets is an **MB**-fibration.

Definition 3.17 A morphism $f : A \to B$ in $\left(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$ is said to be

- a cofibration if its underlying map of **MB** simplicial sets is a cofibration;
- a weak equivalence if for every U-local $(0,1)$-fibration $p : X \to S$ the associated map of ∞-bicategories

$$f^* : \mathrm{Map}_S(B,X) \xrightarrow{\simeq} \mathrm{Map}_S(A,X)$$

is a bicategorical equivalence;

- a trivial fibration if it has the right lifting property against every cofibration in $\left(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$;
- a trivial cofibration if it is *both* a weak equivalence and a cofibration.

Lemma 3.18 *Let* $f : A \to B$ *be a morphism over* S *in* $\left(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$ *such that* f *is a trivial fibration. Then* f *is a weak equivalence.*

Proof Note that since f is a trivial fibration we can construct a section $g : B \to A$ such that $f \circ g = \mathrm{id}_B$. Moreover, we can further produce a marked homotopy $A \times (\Delta^1)^\sharp \to A$ between the identity morphism on A and the composite $g \circ f$ which is compatible with the projection map from A to S. This pair of homotopy inverse morphisms thus define the desired equivalence on mapping ∞-bicategories

$$f^* : \mathrm{Map}_S(B, X) \xrightarrow{\simeq} \mathrm{Map}_S(A, X)$$

and thus f is a weak equivalence. □

Proposition 3.19 *Let $f : A \to B$ be a morphism in $\big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)_{/(S,\sharp,U\subset\sharp)}$. Given a U-local $(0,1)$-fibration $p : X \to S$ let us consider the induced functor on mapping ∞-bicategories*

$$f^* : \mathrm{Map}_S(B, X) \longrightarrow \mathrm{Map}_S(A, X).$$

Then it follows that

(i) *if f is **MB**-anodyne then f^* is a trivial fibration of scaled simplicial sets*;

(ii) *if f is a cofibration then f^* is a fibration of ∞-bicategories*;

(iii) *if f is a trivial cofibration then f^* is a trivial fibration of scaled simplicial sets.*

Proof The first statement follows directly from Proposition 3.12. To see that (ii) holds we observe that again by Proposition 3.12 that f^* has the right lifting property against all scaled anodyne maps. Since the marked morphisms in the mapping ∞-bicategories are equivalences it follows that f^* is an isofibration. Therefore by Remark 2.12 it follows that f^* is a fibration of scaled simplicial sets. The final claim follows from (ii) together with the definition of the class of weak equivalences. □

Lemma 3.20 *Let us consider a pushout diagram, in $\big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)_{/(S,\sharp,U\subset\sharp)}$,*

$$\begin{tikzcd} A \arrow[r, "u"] \arrow[d, "v"] & B \arrow[d] \\ C \arrow[r, "i"] & P \end{tikzcd}$$

where u is a weak equivalence and v is a cofibration. Then $i : B \to C$ is also a weak equivalence.

Proof Given a U-local $(0,1)$-fibration $p : X \to S$ we observe that since v is a cofibration we obtain a pullback diagram of ∞-bicategories

$$\begin{tikzcd} \mathrm{Map}_S(P, X) \arrow[r, "i^*"] \arrow[d] & \mathrm{Map}_S(C, X) \arrow[d, "v^*"] \\ \mathrm{Map}_S(B, P) \arrow[r, "u^*"] & \mathrm{Map}_S(A, X) \end{tikzcd}$$

which shows that i^* is a bicategorical equivalence and consequently we see that $i : B \to C$ is a weak equivalence in $\big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)_{/(S,\sharp,U\subset\sharp)}$. □

Proposition 3.21 *An object $p : X \to S$ in $\big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)_{/(S,\sharp,U\subset\sharp)}$ has the right lifting property against the class of trivial cofibrations if and only if it is a U-local $(0,1)$-fibration.*

Proof Observe that due to (i) in Proposition 3.19 it follows that every **MB**-anodyne morphism is a trivial cofibration. Therefore, any object having the right lifting property against trivial cofibrations must be a U-local $(0,1)$-fibration. To check the converse we consider a U-local $(0,1)$-fibration $p : X \to S$ and a trivial cofibration $f : A \to B$. Then, in order to produce a solution to the lifting problem

$$\begin{tikzcd} A \arrow[r, "\alpha"] \arrow[d, "f"'] & X \arrow[d, "p"] \\ B \arrow[r] \arrow[ur, dotted] & S \end{tikzcd}$$

we observe that since the map $f^* : \operatorname{Map}_S(B, X) \to \operatorname{Map}_S(A, X)$ is a trivial fibration and in particular surjective, we can pick a preimage of $\alpha \in \operatorname{Map}_S(A, X)$ which yields the solution to our problem. $\square$

Definition 3.22 Let $p : X \to S$ be a U-local $(0,1)$-fibration. Consider an object $A \to S$ in $\big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)_{/(S,\sharp,U\subset\sharp)}$. We set the following notation:

(1) We denote by $\operatorname{Map}_S^{\mathrm{th}}(A, X)$ the underlying ∞-category (cf. Definition 2.5) of the mapping ∞-bicategory.

(2) We denote by $\operatorname{Map}_S^{\simeq}(A, X)$ the underlying groupoid of the mapping ∞-bicategory.

Proposition 3.23 *Let* $f : X \to Y$ *be a morphism in* $\big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)_{/(S,\sharp,U\subset\sharp)}$ *where both* X *and* Y *are* U*-local* $(0,1)$*-fibrations. Then the following are equivalent*:

(i) *For every* U*-local* $(0,1)$*-fibration* $Z \to S$ *the induced map* $f^* : \operatorname{Map}_S(Y, Z) \to \operatorname{Map}_S(X, Z)$ *is an equivalence of* ∞*-bicategories.*

(ii) *For every* U*-local* $(0,1)$*-fibration* $Z \to S$ *the induced map* $f^* : \operatorname{Map}_S^{\leq 1}(Y, Z) \to \operatorname{Map}_S^{\leq 1}(X, Z)$ *is an equivalence of* ∞*-categories.*

(iii) *For every* U*-local* $(0,1)$*-fibration* $Z \to S$ *the induced map* $f^* : \operatorname{Map}_S^{\simeq}(Y, Z) \to \operatorname{Map}_S^{\simeq}(X, Z)$ *is a homotopy equivalence of groupoids.*

(iv) *The exists a morphism* $g : Y \to X$ *over* S*, which is a homotopy inverse to* f*.*

(v) *For every* $s \in S$ *the induced morphism on fibres* $f_s : X_s \to Y_s$ *is a bicategorical equivalence.*

Proof The implications (i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii) are clear. We commence the proof by showing that (iii) $\Longrightarrow$ (iv). We consider the homotopy equivalence $\operatorname{Map}_S^{\simeq}(Y, X) \to \operatorname{Map}_S^{\simeq}(X, X)$ and pick an object $g \in \operatorname{Map}_S^{\simeq}(Y, X)$ such that $g \circ f \simeq \mathrm{id}_X$. To show that g is the desired homotopy inverse to f we need to show that $f \circ g \simeq \mathrm{id}_Y$. To see this we see that the map $\operatorname{Map}_S^{\simeq}(Y, Y) \to \operatorname{Map}_S^{\simeq}(X, Y)$ maps both $f \circ g$ and id_Y to morphisms which are equivalent to f. Consequently we see that $f \circ g \simeq \mathrm{id}_Y$.

Observe that (iv) $\Longrightarrow$ (v) follows from the fact that since our homotopies are fibrewise they descend to equivalences on the corresponding fibres.

In order to exhibit that (v) $\Longrightarrow$ (i) we use the small object argument to factor $f : X \to Y$ as a composite $X \to \widehat{X} \xrightarrow{\hat{f}} Y$ where the first morphism is **MB**-anodyne (and therefore a weak equivalence) and the second morphism has the right lifting property against the class of **MB**-anodyne morphisms.

In particular it follows that $\widehat{X} \to S$ is again a U-local $(0,1)$-fibration and that the induced maps on fibres $\hat{f}_s : \widehat{X}_s \to Y_s$ are bicategorical equivalences. These assumptions imply that $\hat{f}$ is itself a trivial fibration (see Proposition 4.24 for a proof) and so the claim follows. □

Lemma 3.24 *Given an object $A \to S$ in $\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$, the projection map $\pi_A : A \times (\Delta^n, \natural, \natural) \to A$ is a weak equivalence.*

Proof Let $\iota : \Delta^0 \to (\Delta^n, \natural, \natural)$ be the inclusion of the initial vertex. It is an easy exercise to show that ι is **MB**-anodyne. It then follows from Proposition 3.12 that $\iota_A : A \to A \times (\Delta^n, \natural, \natural)$ is also **MB**-anodyne. We conclude the proof by observing that $\pi_A \circ \iota_A = \mathrm{id}_A$. □

Theorem 3.25 *Let (S,U) be a scaled simplicial set. Then there exists a left proper combinatorial simplicial model structure on $\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$, which is characterised uniquely by the following properties:*

- (C) *A morphism $f : X \to Y$ in $\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$ is a cofibration if and only if f induces a monomorphism on the underlying simplicial sets.*
- (F) *An object $p : X \to S$ in $\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$ is fibrant if and only if it is a U-local $(0,1)$-fibration.*

Proof The proof is totally analogous to the proof of Theorem 3.42 in [6] where we verify that the conditions of [18, A.2.6.10] are satisfied. □

Remark 3.26 We will refer to the model structure in the previous theorem as model structure on U-local $(0,1)$-cartesian fibrations over (S,U).

Definition 3.27 Let $K = (K, E_K) \in \mathrm{Set}_\Delta^+$ and let $p : X \to S$ be an object in $\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$. We define the tensor $K \otimes X$ as $I(K) \times X \to X \to S$ where $I(K) = (K, E_K, \natural)$. Similarly, we define the cotensor X^K by declaring that a map **MB**-simplicial sets $\varphi : Y \to X^K$ over $\bar{\varphi} : Y \to S$ to be equivalent to the data of a commutative diagram

$$\begin{tikzcd}
(K, E_K, \natural) \times Y \arrow[r] \arrow[d] & X \arrow[d, "p"] \\
Y \arrow[r, "\bar{\varphi}"] & S
\end{tikzcd}$$

Theorem 3.28 *The model category $\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$ is a Set_Δ^+-enriched model category.*

Proof It is clear that the construction $\mathrm{Map}_S^{\leqslant 1}(-,-)$ equips $\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$ with the structure of a Set_Δ^+-enriched model category. Since the tensor preserves colimits in both variables separately it will be enough to show that given $i : L \to K$ a cofibration in Set_Δ^+ and a cofibration $f : X \to Y$ in $\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$ the corresponding pushout-product map

$$i \wedge f : L \otimes Y \amalg_{L \otimes X} K \otimes X \to K \otimes Y$$

is again cofibration which is a weak equivalence whenever i or f is. Note that $i \wedge f$ is clearly a cofibration so we can focus our attention in proving the weak equivalence part of the claim.

First, let us recall that given an anodyne morphism of marked simplicial sets $A \to B$ it follows that $I(A) \to I(B)$ is **MB**-anodyne. It then follows as a consequence Proposition 3.12 that we can assume that i and f are morphisms among fibrant objects in the corresponding model structures.

We note that given a pair of fibrant objects L and $p : X \to S$ it follows that $L \otimes X$ is again fibrant. To finish the proof we assume that $f : X \to Y$ is a weak equivalence (the case for i is totally analogous). Then it follows that for every $s \in S$ the map $(L \otimes X)_s \to (L \otimes Y)_s$ is identified with the map

$$L \times X_s \xrightarrow{\simeq} L \times Y_s,$$

which is a bicategorical equivalence by assumption. It follows that the map $L \otimes X \to L \otimes Y$ is a weak equivalence in $\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}$. We can now consider a pushout diagram

$$\begin{tikzcd}
L \otimes X \arrow[r, "\simeq"] \arrow[d] & L \otimes Y \arrow[d] \\
K \otimes X \arrow[r, "\simeq"] & L \otimes Y \amalg_{L\otimes X} K \otimes X
\end{tikzcd}$$

Moreover, using a similar argument as before we see that $K \otimes X \xrightarrow{\simeq} K \otimes Y$ is also a weak equivalence. The claim now follows from two-out-of-three. □

Proposition 3.29 *Let* $f : (S,U) \to (T,V)$ *be a map of scaled simplicial sets. Then postcomposition with* f *induces a left Quillen functor*

$$f_! : \left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)} \rightleftarrows \left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(T,\sharp,V\subset\sharp)} : f^*,$$

which is left adjoint to the pullback functor f^*.

Proof It is clear that $f_!$ preserves cofibrations. To finish the proof we only need to show that $f_!$ preserves weak equivalences. Given $\iota : A \to B$ and a fibrant object $p : Y \to S'$ we observe that we have a commutative diagram

$$\begin{tikzcd}
\mathrm{Map}_S(B, f^*Y) \arrow[r, "\simeq"] \arrow[d, "\simeq"] & \mathrm{Map}_S(A, f^*Y) \arrow[d, "\simeq"] \\
\mathrm{Map}_T(i_!B, Y) \arrow[r] & \mathrm{Map}_T(i_!A, Y)
\end{tikzcd}$$

so the conclusion holds by two-out-of-three. □

We finish this section by comparing the model structure in Theorem 3.25 with the model structure constructed in Theorem 3.2.6 in [19], and in the case where $(S,U) = (\Delta^0, \sharp)$ with the model structure on scaled simplicial sets given in Theorem 2.10.

Definition 3.30 Let (S,U) be a scaled simplicial set and consider the category $\left(\mathrm{Set}_\Delta^{+}\right)_{/(S,U)}$ of marked simplicial sets over S. We have a functor

$$R : \left(\mathrm{Set}_\Delta^{+}\right)_{/(S,U)} \to \left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(S,\sharp,U\subset\sharp)}, \quad (X, E_X) \mapsto (X, E_X, T_X \subset \sharp),$$

where T_X consists in those triangles in X lying over thin triangles in S.

Theorem 3.31 *Let (S,U) be a scaled simplicial set and consider an object (X,E_X) in $\left(\mathrm{Set}_\Delta^+\right)_{/(S,U)}$. Then $X=(X,E_X)$ is $\mathfrak{P}_S$-fibred (see Definition 3.2.1 and Example 3.2.9 in [19]) if and only if $R(X)$ is a U-local $(0,1)$-fibration. Moreover, if Y is a U-local $(0,1)$-fibration over S such that every triangle of Y is lean then there exists a $\mathfrak{P}_S$-fibred object T such that $R(T)=Y$.*

Proof Let us suppose that X is $\mathfrak{P}_S$-fibred and let us show that $R(X)$ defines a U-local $(0,1)$-fibration. We will show that $R(X)$ has the right lifting property against the class of **MB**-anodyne morphisms. To this end we recall that an object in $\left(\mathrm{Set}_\Delta^+\right)_{/S}$ is $\mathfrak{P}_S$-fibred if and only if it has the right lifting property against the class of $\mathfrak{P}_S$-anodyne morphisms described in Definition 3.2.10 in [19]. We first check that $R(X)$ has the right lifting property against the class of maps given in (E). Indeed given a map from a Kan complex $K\to R(X)$ we can use the morphisms of type (A_1) in [19, Definition 3.2.10] to see that every morphism of K maps to a marked edge in $R(X)$. Similarly, our construction of R guarantees the $R(X)$ has the right lifting property against morphisms of type (S2). The rest of the lifting problems follow immediately from the definition of the class of $\mathfrak{P}_S$-anodyne morphisms.

The converse follows by a similar argument. To finish the proof we suppose that we are given U-local $(0,1)$-fibration of the form $Y=(Y,E_Y,T_Y\subset\sharp)$. Then (S2) implies that T_Y is simply the collection of triangles lying over thin triangles in S. Therefore, it we can consider $\widehat{Y}=(Y,E_Y)$ and observe that $R(\widehat{Y})=Y$. The previous part of the proof shows that $\widehat{Y}$ must be $\mathfrak{P}_S$-fibred. □

Proposition 3.32 *Let $p:X\to S$ be a U-local $(0,1)$-fibration. Then for every $s\in S$ the fibre X_s is an ∞-category if and only if every triangle of X is lean.*

Proof It is clear that if every triangle of X is lean the fibres must be ∞-categories. For the converse, let us assume that for every $s\in S$ the fibre X_s is an ∞-category. Let $\sigma:\Delta^2\to X$ and assume that $p(\sigma)=s_1(p(f))$ for some edge $\Delta^1\to S$. We pick a marked edge lying over $p(f)$ and construct a 3-simplex $\theta:\Delta^3\to X$ lying over $s_1(p(\sigma))$ such that $\theta(0\to 1)$ is our chosen marked edge and such that every face is lean except possibly $d_1(\theta)=\sigma$. We conclude that σ is lean since p has the right lifting property against morphisms of type (A2).

For the general case we take a marked edge lying over $p(1\to 2)$ and construct another 3-simplex $\Xi:\Delta^3\to X$ such that $\Xi(1\to 2)$ is our chosen marked edge and such that $d_2(\Xi)=\sigma$. It follows that every face of Ξ is lean except possibly the face skipping the vertex 1 and the face missing the vertex 2. We conclude that the face missing the vertex 1 is lean since it falls in the previous case. It then follows that σ is lean. □

Definition 3.33 Let $\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(\Delta^0,\sharp,\sharp)}=\mathrm{Set}_\Delta^{\mathbf{mb}}$. We consider a functor $R:\mathrm{Set}_\Delta^{\mathbf{mb}}\to\mathrm{Set}_\Delta^{\mathbf{sc}}$ which sends an **MB** simplicial set $(X,E_X,T_X\subseteq C_X)$ to the scaled simplicial set $R(X,E_X,T_X\subseteq C_X)=(X,C_X)$. This functor has a left adjoint $L:\mathrm{Set}_\Delta^{\mathbf{sc}}\to\mathrm{Set}_\Delta^{\mathbf{mb}}$ which is given by $L(Y,T_Y)=(Y,\flat,\flat\subset T_Y)$.

Theorem 3.34 *The functor $L:\mathrm{Set}_\Delta^{\mathbf{sc}}\to\mathrm{Set}_\Delta^{\mathbf{mb}}$ is a left Quillen equivalence.*

Proof It is clear that L preserves cofibrations and colimits so in order to show that L is a left Quillen functor it will be enough to show that L preserves weak equivalences. Observe that given an object

$(X, E_X, T_X \subset C_X)$ we have an anodyne morphism $(X, E_X, T_X \subset C_X) \to (X, E_X, C_X)$ since every triangle in Δ^0 is thin. Using this observation it is easy to see that L maps scaled anodyne morphisms to trivial cofibrations in $\mathrm{Set}_\Delta^{\mathbf{mb}}$. This shows that it is enough to show that L preserves weak equivalences among fibrant scaled simplicial sets. However, this is clear since a weak equivalence between fibrant scaled simplicial sets has an inverse up to homotopy.

To conclude that L is a left Quillen equivalence we first observe that $R \circ L = \mathrm{id}$ by definition. This shows that it is only left to show that for any fibrant object $(Y, E_Y, T_Y \subseteq C_Y) \in \mathrm{Set}_\Delta^{\mathbf{mb}}$ the counit map $LR(Y) \to Y$ is a weak equivalence. Note that since our chosen **MB** simplicial set is fibrant $C_Y = T_Y$ and (Y, T_Y) is an ∞-bicategory. In particular, we see that $LR(Y) \to Y$ is the weakly saturated class of morphisms of type (E) and (S2) in Definition 3.4. □

3.1 Marked-scaled simplicial sets

We saw in Theorem 3.34 that the model structure of **MB** simplicial sets over Δ^0 is Quillen equivalent to the model structure on scaled simplicial sets given in [19]. While doing this, we observed that the collection of lean and thin triangles becomes redundant. To deal with this issue we introduce a third model structure on simplicial sets equipped with a collection of marked edges and *one* collection of triangles which we call marked-scaled simplicial sets.

Definition 3.35 A marked-scaled simplicial set denoted by (X, E_X, T_X) is given by

(1) a simplicial set X,

(2) a collection of edges $E_X \subseteq X_1$ which contains *all degenerate* edges,

(3) a collection of triangles $T_X \subseteq X_2$ which contains *all degenerate* triangles,

where we refer to the elements of E_X as *marked* edges, and we refer to the elements of T_X as *thin* triangles. A morphism of marked-scaled simplicial sets $f : (X, E_X, T_X) \to (Y, E_Y, T_Y)$ is given by a map of simplicial sets such that $f(E_X) \subseteq E_Y$ and $f(T_X) \subseteq T_Y$. We denote by $\mathrm{Set}_\Delta^{\mathbf{ms}}$ the category of marked-scaled simplicial sets.

Definition 3.36 The set of *generating marked-scaled anodyne maps* **MS** is the set of maps of marked-scaled simplicial sets consisting of

(M1) the inner horn inclusions

$$\left(\Lambda^n_i, \flat, \{\Delta^{\{i-1,i,i+1\}}\}\right) \to \left(\Delta^n, \flat, \{\Delta^{\{i-1,i,i+1\}}\}\right), \quad n \geqslant 2,\ 0 < i < n;$$

(M2) the map

$$(\Delta^4, \flat, T) \to (\Delta^4, \flat, T \cup \{\Delta^{\{0,3,4\}}, \Delta^{\{0,1,4\}}\}),$$

where we define

$$T \overset{\text{def}}{=} \{\Delta^{\{0,2,4\}}, \Delta^{\{1,2,3\}}, \Delta^{\{0,1,3\}}, \Delta^{\{1,3,4\}}, \Delta^{\{0,1,2\}}\};$$

(M3) the set of maps

$$(\Lambda^n_0, \{\Delta^{\{0,1\}}\}, \{\Delta^{\{0,1,n\}}\}) \to (\Delta^n, \{\Delta^{\{0,1\}}\}, \{\Delta^{\{0,1,n\}}\}), \quad n \geqslant 2;$$

(M4) the inclusion of the initial vertex

$$(\Delta^0,\sharp,\sharp)\to(\Delta^1,\sharp,\sharp);$$

(MS1) the map

$$(\Delta^2,\{\Delta^{\{0,1\}},\Delta^{\{1,2\}}\},\sharp)\to(\Delta^2,\sharp,\sharp);$$

(ME) for every Kan complex K, the map

$$(K,\flat,\sharp)\to(K,\sharp,\sharp),$$

which requires that every equivalence is a marked morphism.

A map of **MS** simplicial sets is said to be **MS**-anodyne if it belongs to the weakly saturated closure of **MS**.

Remark 3.37 Observe that (X,E_X,T_X) has the right lifting property against the class of **MS**-anodyne morphisms if and only if X is an ∞-bicategory, E_X is the collection of equivalences in X and T_X is the collection of thin triangles. Consequently, we might call such marked-scaled simplicial sets ∞-bicategories as well.

Definition 3.38 We say that a morphism of marked-scaled simplicial sets is a cofibration if its underlying map of simplicial sets is a cofibration.

Proposition 3.39 *Let $f:X\to Y$ be a cofibration in $\mathrm{Set}_\Delta^{\mathbf{ms}}$ and $g:A\to B$ be an **MS**-anodyne morphism. Then the pushout-product*

$$X\times B\amalg_{X\times A}Y\times A\to Y\times B$$

*is again **MS**-anodyne.*

Corollary 3.40 *Given a marked-biscaled simplicial set X which has the right lifting property against the class of **MS**-anodyne morphisms it follows that $\mathrm{Fun}^{\mathbf{ms}}(A,X)$ (compare to Remark 3.13) has the right lifting property against the class **MS**-anodyne morphisms for every $A\in\mathrm{Set}_\Delta^{\mathbf{ms}}$.*

Definition 3.41 A morphism of marked-scaled simplicial sets $i:A\to B$ is said to be a *weak equivalence* if for every ∞-bicategory X the induced map

$$i^*:\mathrm{Fun}^{\mathbf{ms}}(B,X)\xrightarrow{\simeq}\mathrm{Fun}^{\mathbf{ms}}(A,X)$$

is a bicategorical equivalence.

Remark 3.42 In a similar way as before given $K\in\mathrm{Set}_\Delta^{+}$ and $X\in\mathrm{Set}_\Delta^{\mathbf{ms}}$ we define the tensor $K\times X:=I(K)\times X$ where $I(K)=(K,E_K,\sharp)$. Similarly, we define the cotensor $X^K:=\mathrm{Fun}^{\mathbf{ms}}(I(K),X)$.

Theorem 3.43 *There exists a left proper combinatorial simplicial model category on $\mathrm{Set}_\Delta^{\mathbf{ms}}$, which is characterised uniquely by the following properties:*

(C) *A morphism $f:X\to Y$ in $\mathrm{Set}_\Delta^{\mathbf{ms}}$ is a cofibration if and only if f induces a monomorphism on the underlying simplicial set.*

(F) *An object* X *in* $\mathrm{Set}_\Delta^{\mathbf{ms}}$ *is fibrant if and only if it was the right lifting property against the class of marked-anodyne morphisms.*

Moreover the tensor and cotensor in Remark 3.42 equips $\mathrm{Set}_\Delta^{\mathbf{ms}}$ *with the structure of a* Set_Δ^{+}*-enriched model category.*

Theorem 3.44 *The functor* $L:\mathrm{Set}_\Delta^{\mathbf{sc}}\to\mathrm{Set}_\Delta^{\mathbf{ms}}$ *sending a scaled simplicial set* (X,T_X) *to the marked-scaled simplicial set* $(X,\flat,T_X)$ *is a left Quillen equivalence.*

Proof The proof is almost identical to the proof of Theorem 3.34 and thus omitted. □

4 Local fibrations

In this section we will show that the model independent definition of a locally cocartesian fibration of $(\infty,2)$-categories, or in our terminology a local $(0,1)$-fibration, given in the introduction can be realised within our framework of **MB** simplicial sets as follows:

Given a fibrant scaled simplicial set (S,T_S) we will construct a subcollection $M_S\subseteq T_S$ and show in Theorem 4.22 that a local $(0,1)$-fibration is precisely given by an M_S-local $(0,1)$-fibration. This perspective will allow us to give an easy proof of Proposition 4.24, as promised in Proposition 3.23.

Definition 4.1 Let (S,T_S) be an ∞-bicategory we define $M_S\subseteq T_S$ as the subcollection of triangles consisting in those *thin* triangles such that the edge $\Delta^{\{0,1\}}$ or the edge $\Delta^{\{1,2\}}$ is an equivalence in S. We call the elements of M_S the invertible 2-morphisms of S.

Remark 4.2 The collection M_S has been studied in [11] to give a definition of an oplax normalised functor of ∞-bicategories.

Notation 4.3 We will use boldface letters $\mathbb{S}:=(S,T_S)$ to describe fibrant scaled simplicial sets.

Definition 4.4 Let $\mathbb{S}=(S,T_S)$ be an ∞-bicategory and let $p:X\to S$ be an M_S-local (see Definition 4.1) $(0,1)$-fibration. We define a scaled simplicial set $\mathbb{X}$ whose underlying simplicial set is the underlying simplicial set of X and where a triangle is declared to be thin if it is lean in X and its image under p belongs to T_S. We denote the resulting map of scaled simplicial sets by $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ and call it the *bicategorical interpretation* of p.

Proposition 4.5 *Let* $\mathbb{S}=(S,T_S)$ *be an* ∞*-bicategory and let* $p:X\to S$ *be an* M_S*-local* $(0,1)$*-fibration. Then its bicategorical interpretation* $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ *(see Definition 4.4) is a bicategorical fibration.*

Proof It is clear that $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ has the right lifting property against the class of scaled anodyne maps and thus it follows that $\mathbb{X}$ is itself an ∞-bicategory. To finish the proof we will need to show that $\mathfrak{p}$ is an isofibration (see Remark 2.12). Given an equivalence $e:\Delta^1\to\mathbb{S}$ and a lift of the source $\Delta^0\to\mathbb{X}$ we can produce a marked edge $\hat{e}:\Delta^1\to X$ lying over e. We observe that the 2-simplices needed for exhibiting e as an equivalence in $\mathbb{S}$ are contained in M_S; in particular this can be used to show that $\hat{e}$ is an equivalence in $\mathbb{X}$. □

Proposition 4.6 *Let $\mathfrak{p} : \mathbb{X} \to \mathbb{S}$ be a bicategorical interpretation as in Definition 4.4. Then for every pair of objects $a, b \in \mathbb{X}$ the induced functor on mapping ∞-categories*

$$\mathbb{X}(a,b) \longrightarrow \mathbb{S}(\mathfrak{p}(a),\mathfrak{p}(b))$$

is a cartesian fibration. Moreover, the composition functors $\mathbb{X}(a,b) \times \mathbb{X}(b,c) \to \mathbb{X}(a,c)$ preserve cartesian edges.

Proof We pick a model for the mapping ∞-category discussed in Definition 2.14. Given a morphism $\alpha : \Delta^1 \to \mathbb{S}(\mathfrak{p}(a),\mathfrak{p}(b))$ and a lift of the target $g : \Delta^0 \to \mathbb{X}(a,b)$ we can consider a morphism of type (A1) to produce an edge in $\mathbb{X}(a,b)$ which is lean in the **MB** simplicial set X. We can similarly translate lifting problems of the form

$$\begin{tikzcd}
\Lambda^n_n \arrow[r] \arrow[d] & \mathbb{X}(a,b) \arrow[d] \\
\Delta^n \arrow[r] \arrow[ur, dotted] & \mathbb{S}(\mathfrak{p}(a),\mathfrak{p}(b))
\end{tikzcd}$$

where the last edge in the top horizontal morphism is mapped to a lean 2-simplex to lifting problems of the form

$$\begin{tikzcd}
\Lambda^{n+1}_n \arrow[r] \arrow[d] & X \arrow[d] \\
\Delta^{n+1} \arrow[r] \arrow[ur, dotted] & S
\end{tikzcd}$$

where the triangle $\Delta^{\{n-1,n,n+1\}}$ is mapped to a lean triangle in X and thus admits a solution. We conclude that $\mathbb{X}(a,b) \to \mathbb{S}(\mathfrak{p}(a),\mathfrak{p}(b))$ is a cartesian fibration. To finish the proof we consider the composition functor

$$\mathbb{X}(a,b) \times \mathbb{X}(b,c) \longrightarrow \mathbb{X}(a,c).$$

Given a pair of cartesian edges $\alpha : f \to g$ and $\beta : u \to v$ the composition map yields a commutative diagram in $\mathbb{X}(a,c)$ of the form

$$\begin{tikzcd}
f \circ u \arrow[r] \arrow[d] & f \circ v \arrow[d] \\
g \circ u \arrow[r] & g \circ v
\end{tikzcd}$$

which tells us that it will suffice to check that precomposition and postcomposition with a 1-morphism preserves cartesian 2-morphisms. This follows from the fact that $\mathfrak{p}$ has the right lifting property against morphisms of type (A2). □

Definition 4.7 A bicategorical fibration (see Notation 4.3) $\mathfrak{p} : \mathbb{X} \to \mathbb{S}$ of ∞-bicategories is said to be *cartesian-enriched* if:

- For every $a, b \in \mathbb{X}$ the morphisms $\mathbb{X}(a,b) \to \mathbb{S}(p(a), p(b))$ are cartesian fibrations.
- For every $a, b, c \in \mathbb{X}$ the composite maps $\mathbb{X}(a,b) \times \mathbb{X}(b,c) \to \mathbb{X}(a,c)$ preserve cartesian edges.

Given two cartesian-enriched bicategorical fibrations $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ and $\mathfrak{q}:\mathbb{Y}\to\mathbb{S}$ we say that a functor $\mathfrak{f}:\mathbb{X}\to\mathbb{Y}$ is cartesian-enriched if it preserves cartesian edges in the mapping categories.

Definition 4.8 Let $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ be a bicategorical fibration of ∞-bicategories. An edge $e:a\to b$ in $\mathbb{X}$ is said to be locally $(0,1)$-cartesian (or a local $(0,1)$-cartesian edge) if:

(i) Given $g:a\to c$ in $\mathbb{X}$ and a commutative diagram (represented by a thin simplex), in $\sigma:\Delta^2\to\mathbb{S}$,

$$\begin{tikzcd} & \mathfrak{p}(b) \arrow[dr, "\alpha"] & \\ \mathfrak{p}(a) \arrow[ur, "\mathfrak{p}(e)"] \arrow[rr, "\mathfrak{p}(g)"'] & & \mathfrak{p}(c) \end{tikzcd}$$

such that α is an equivalence, there exists a morphism $\hat{\alpha}:b\to c$ such that $\mathfrak{p}(\hat{\alpha})=\alpha$ and a thin 2-simplex $\hat{\sigma}$ exhibiting $e\circ\hat{\alpha}\simeq g$ such that $\mathfrak{p}(\hat{\sigma})=\sigma$.

(ii) Given any $\phi:b\to c$ such that $\phi\circ e\simeq g$ and such that $\mathfrak{p}(\phi)=\alpha$ as above, for any other $\varphi:b\to c$, precomposition along e induces a pullback diagram of spaces

$$\begin{tikzcd} \mathrm{Map}_{\mathbb{X}(b,c)}(\phi,\varphi) \arrow[r] \arrow[d] & \mathrm{Map}_{\mathbb{X}(a,c)}(\phi\circ e,\varphi\circ e) \arrow[d] \\ \mathrm{Map}_{\mathbb{S}(\mathfrak{p}(b),\mathfrak{p}(c))}(\mathfrak{p}(\phi),\mathfrak{p}(\varphi)) \arrow[r] & \mathrm{Map}_{\mathbb{S}(\mathfrak{p}(a),\mathfrak{p}(c))}(\mathfrak{p}(\phi\circ e),\mathfrak{p}(\varphi\circ e)) \end{tikzcd}$$

Remark 4.9 Given a bicategorical fibration $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ of ∞-bicategories, we observe that a local $(0,1)$-cartesian edge lying over an equivalence in $\mathbb{S}$ is necessarily an equivalence in $\mathbb{X}$. Moreover, the composition of a local $(0,1)$-cartesian edge with an equivalence is again locally $(0,1)$-cartesian.

Definition 4.10 Let $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ be a bicategorical fibration of ∞-bicategories. An edge $e:a\to b$ is $(0,1)$-cartesian if for every $c\in\mathbb{X}$ precomposition along e induces a pullback diagram of ∞-categories

$$\begin{tikzcd} \mathbb{X}(b,c) \arrow[r] \arrow[d] & \mathbb{X}(a,c) \arrow[d] \\ \mathbb{S}(\mathfrak{p}(b),\mathfrak{p}(c)) \arrow[r] & \mathbb{S}(\mathfrak{p}(a),\mathfrak{p}(c)) \end{tikzcd}$$

Proposition 4.11 *Let* $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ *be a bicategorical fibration of* ∞*-bicategories and suppose further that* $\mathfrak{p}$ *is cartesian-enriched. Given an edge* $e:a\to b$ *in* $\mathbb{X}$, *the following are equivalent*:

(i) *The edge* e *is locally* $(0,1)$*-cartesian.*

(ii) *For every* $\ell\in\mathbb{X}$ *such that* $\mathfrak{p}(\ell)=\mathfrak{p}(b)$ *we have a pullback diagram of* ∞*-categories*

$$\begin{tikzcd} \mathbb{X}_{\mathfrak{p}(b)}(b,\ell) \arrow[r, "(-)\circ e"] \arrow[d] & \mathbb{X}(a,\ell) \arrow[d] \\ \Delta^0 \arrow[r, "\mathfrak{p}(e)"] & \mathbb{S}(\mathfrak{p}(a),\mathfrak{p}(b)) \end{tikzcd}$$

where $X_{\mathfrak{p}(b)}$ *is the fibre of* $\mathfrak{p}$ *over* $\mathfrak{p}(b)$.

(iii) *The edge* e *is* $(0,1)$*-cartesian in the pullback along* $\mathfrak{p}(e)$, $\mathbb{X}_{\mathfrak{p}(e)}=\mathbb{X}\times_{\mathbb{S}}\Delta^1\to\Delta^1$.

Proof To show that (i) $\implies$ (ii) we observe that since

$$\mathbb{X}(a,\ell) \to \mathbb{S}(\mathfrak{p}(a),\mathfrak{p}(b))$$

is a cartesian fibration of ∞-categories it follows that the strict fibre over $\mathfrak{p}(e)$ is already a model for the ∞-categorical pullback. Therefore it suffices to show that the map

$$e^* : \mathbb{X}_{\mathfrak{p}(b)}(b,\ell) \to \mathbb{X}_{\mathfrak{p}(e)}(a,\ell)$$

is an equivalence of ∞-categories. Observe that the first condition in Definition 4.8 guarantees that e^* is essentially surjective. Given a pair of objects $\phi, \varphi \in \mathbb{X}_{\mathfrak{p}(b)}(b,\ell)$ it follows from condition (ii) in Definition 4.8 that we have a pullback diagram of spaces

$$\begin{tikzcd}
\operatorname{Map}_{\mathbb{X}(b,\ell)}(\phi,\varphi) \arrow[r] \arrow[d] & \operatorname{Map}_{\mathbb{X}(a,\ell)}(\phi\circ e,\varphi\circ e) \arrow[d] \\
\operatorname{Map}_{\mathbb{S}(\mathfrak{p}(b),\mathfrak{p}(b))}(\mathfrak{p}(\phi),\mathfrak{p}(\varphi)) \arrow[r] & \operatorname{Map}_{\mathbb{S}(\mathfrak{p}(a),\mathfrak{p}(b))}(\mathfrak{p}(\phi\circ e),\mathfrak{p}(\varphi\circ e))
\end{tikzcd}$$

so in particular after taking fibres we have a homotopy equivalence of spaces

$$\operatorname{Map}_{\mathbb{X}(b,\ell)}(\phi,\varphi)_{\mathrm{id}} \xrightarrow{\simeq} \operatorname{Map}_{\mathbb{X}(a,\ell)}(\phi\circ e,\varphi\circ e)_{\mathfrak{p}(e)},$$

which we identify with the action of e^* on mapping spaces which shows our claim.

Note that (ii) $\iff$ (iii) follows immediately from the definition so it will suffice to show that (ii) $\implies$ (i).

First let us remark that since $\mathfrak{p} : \mathbb{X} \to \mathbb{S}$ is a bicategorical fibration, it is in particular an isofibration. So, in order to show that e is locally $(0,1)$-cartesian, we specialise the conditions in Definition 4.8 to the case where $\mathfrak{p}(\alpha)$ is a degenerate edge. Let $\ell \in \mathbb{X}$ such that $\mathfrak{p}(\ell) = \mathfrak{p}(b)$ and consider an edge $u : a \to \ell$ such that $\mathfrak{p}(u) \simeq \mathfrak{p}(e)$ in $\mathbb{S}(\mathfrak{p}(a),\mathfrak{p}(b))$. Since $\mathfrak{p}$ is cartesian enriched we can pick an equivalence in $\mathbb{X}(a,\ell)$, $\hat{u} \simeq u$, such that $\mathfrak{p}(\hat{u}) = \mathfrak{p}(e)$. Then our assumptions guarantee the existence of an object $\phi \in \mathbb{X}(b,\ell)$ such that $\mathfrak{p}(\phi)$ is degenerate and such that

$$\phi\circ e \simeq \hat{u} \simeq u.$$

This shows that the first condition in Definition 4.8 holds. Let $\phi : b \to \ell$ in $\mathbb{X}$ such that $\mathfrak{p}(\phi)$ is degenerate on $\mathfrak{p}(b)$. To show that we have the necessary pullback square of spaces it will be enough to show that for every $\Xi \in \operatorname{Map}_{\mathbb{S}(\mathfrak{p}(b),\mathfrak{p}(b))}(\mathfrak{p}(\phi),\mathfrak{p}(\varphi))$ the associated morphism on fibres

$$\operatorname{Map}_{\mathbb{X}(b,\ell)}(\phi,\varphi)_{\Xi} \xrightarrow{\simeq} \operatorname{Map}_{\mathbb{X}(a,\ell)}(\phi\circ e,\varphi\circ e)_{\Xi\circ\mathfrak{p}(e)}$$

is a homotopy equivalence. Note that our assumptions guarantee that this holds whenever Ξ is the identity morphism. Since the map $\mathbb{X}(b,\ell) \to \mathbb{S}(\mathfrak{p}(b),\mathfrak{p}(b))$ is a cartesian fibration we pick a cartesian lift of Ξ in $\mathbb{X}(b,\ell)$ which we denote by $i : \hat{\varphi} \to \varphi$ and we note that the enrichment of $\mathfrak{p}$ implies that $i\circ e$ is again a cartesian morphism in $\mathbb{X}(a,\ell)$. The fact that i is cartesian allows us to construct a commutative diagram of spaces

$$\begin{tikzcd}
\operatorname{Map}_{\mathbb{X}(b,\ell)}(\phi,\widehat{\varphi})_{\mathrm{id}} \arrow[r, "\simeq"] \arrow[d, "\simeq"] & \operatorname{Map}_{\mathbb{X}(a,\ell)}(\phi\circ e,\widehat{\varphi}\circ e)_{\mathfrak{p}(e)} \arrow[d, "\simeq"] \\
\operatorname{Map}_{\mathbb{X}(b,\ell)}(\phi,\varphi)_{\Xi} \arrow[r] & \operatorname{Map}_{\mathbb{X}(a,\ell)}(\phi\circ e,\varphi\circ e)_{\Xi\circ\mathfrak{p}(e)}
\end{tikzcd}$$

so the conclusion follows by two-out-of-three. □

Definition 4.12 A bicategorical fibration $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ of ∞-bicategories is said to be a local $(0,1)$-fibration if the following conditions hold:

(1) The map $\mathfrak{p}$ is cartesian-enriched (see Definition 4.7).

(2) For every $a\in\mathbb{X}$ and every $e:\mathfrak{p}(a)\to b$ in $\mathbb{S}$ there exists a local $(0,1)$-cartesian edge $\hat{e}:a\to\hat{b}$ such that $\mathfrak{p}(\hat{e})=e$.

We say that a commutative diagram

$$\begin{tikzcd}
\mathbb{X} \arrow[rr, "\mathfrak{f}"] \arrow[dr, "\mathfrak{p}"'] & & \mathbb{Y} \arrow[dl, "\mathfrak{q}"] \\
& \mathbb{S} &
\end{tikzcd}$$

where $\mathfrak{p}$ and $\mathfrak{q}$ are local $(0,1)$-fibrations is a morphism of local $(0,1)$-fibrations if $\mathfrak{f}$ is a morphism of cartesian-enriched fibrations and it maps local $(0,1)$-cartesian edges in $\mathfrak{p}$ to local $(0,1)$-cartesian edges in $\mathfrak{q}$.

Lemma 4.13 *Let* $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ *be a local* $(0,1)$*-fibration of* ∞*-bicategories and consider a commutative diagram, in* $\mathbb{X}$,

$$\begin{tikzcd}
a \arrow[r, "\simeq"] \arrow[d, "f"'] & u \arrow[d, "g"] \\
b \arrow[r, "\simeq"] & v
\end{tikzcd}$$

such that the horizontal morphisms are equivalences in $\mathbb{X}$*. Then* f *is a local* $(0,1)$*-cartesian edge if and only if* g *is.*

Proof Left as an exercise. □

Theorem 4.14 *Let us consider a commutative diagram of* ∞*-bicategories*

$$\begin{tikzcd}
\mathbb{X} \arrow[r, "\simeq"] \arrow[d, "\mathfrak{p}"'] & \mathbb{C} \arrow[d, "\mathfrak{q}"] \\
\mathbb{S} \arrow[r, "\simeq"] & \mathbb{D}
\end{tikzcd}$$

where the vertical morphisms are bicategorical fibrations and the horizontal morphisms are bicategorical equivalences. Then $\mathfrak{p}$ *is a local* $(0,1)$*-fibration if and only if* $\mathfrak{q}$ *is.*

Proof Observe that we can view our diagram as an injectively fibrant-cofibrant diagram in the arrow category of $\mathrm{Set}_{\Delta}^{\mathbf{sc}}$. This guarantees the existence weak equivalences $\mathbb{C} \to \mathbb{X}$ and $\mathbb{D} \to \mathbb{S}$ making the obvious diagram commute. Therefore we might assume without loss of generality that $\mathfrak{q}$ is a local $(0,1)$-fibration. After inspecting the associated diagram in mapping categories we learn that $\mathfrak{p}$ must be cartesian enriched. To finish the proof we need to show that $\mathfrak{p}$ has a sufficient supply of local $(0,1)$-cartesian edges.

First, we observe that our commutative diagram is in fact a pullback diagram of ∞-bicategories. Therefore we have a weak equivalence $\varphi : \mathbb{X} \to \mathbb{P}$ where $\hat{\mathfrak{q}} : \mathbb{P} \to \mathbb{S}$ denotes the strict pullback of $\mathfrak{q}$ along the bottom horizontal morphism. It follows that $\hat{\mathfrak{q}}$ is a local $(0,1)$-fibration. Factoring φ as a cofibration (which is necessary a trivial cofibration) and a trivial fibration we might assume without loss of generality that φ is a trivial cofibration. We conclude that we have a section $\xi : \mathbb{P} \to \mathbb{X}$ such that $\xi \circ \varphi = \mathrm{id}$.

Given $a \in \mathbb{X}$ and an edge $e : \mathfrak{p}(a) \to b$ in $\mathbb{S}$ we pick a lift in $\mathbb{P}$ of e with source $\varphi(a)$ which we denote by $\hat{e}$. We finally consider $\xi(\hat{e}) = \tau$. We claim that τ is a local $(0,1)$-cartesian edge of $\mathfrak{p}$. To see this we note that due to Lemma 4.13 we have that $\varphi(\tau)$ is a local $(0,1)$-cartesian edge of $\mathbb{P}$. The existence of a section ξ and the fact that $\varphi(\tau)$ is locally $(0,1)$-cartesian shows that condition (i) in Definition 4.8 holds. The second condition follows immediately from the fact that φ is a bicategorical equivalence. □

Lemma 4.15 *Let $\mathfrak{p} : \mathbb{X} \to \mathbb{S}$ be a local $(0,1)$-fibration of ∞-bicategories. Let $\sigma : \Delta^2 \to \mathbb{X}$ be a 2-simplex whose associated 2-morphism in $\mathbb{X}(\sigma(0), \sigma(2))$ is cartesian. Given a lifting problem*

$$\begin{tikzcd}
\Lambda^n_i \arrow[r, "\varphi"] \arrow[d] & \mathbb{X} \arrow[d, "q"] \\
\Delta^n \arrow[r] \arrow[ur, dotted] & \mathbb{S}
\end{tikzcd}$$

such that restriction of φ to $\Delta^{\{i-1,i,i+1\}}$ equals σ, the dotted arrow exists.

Proof In virtue of Theorem 4.14 we might assume that our functor is of the form $\mathfrak{p} : \mathrm{N}^{\mathbf{sc}}(\mathbb{C}) \to \mathrm{N}^{\mathbf{sc}}(\mathbb{D})$. Then it follows that our lifting problem is equivalent to

$$\begin{tikzcd}
\mathfrak{C}^{\mathbf{sc}}[\Lambda^n_i](0,n) \arrow[r] \arrow[d] & \mathbb{C}(\varphi(0), \varphi(n)) \arrow[d] \\
\mathfrak{C}^{\mathbf{sc}}[\Delta^n](0,n) \arrow[r] \arrow[ur, dotted] & \mathbb{D}(\mathfrak{p}\varphi(0), \mathfrak{p}\varphi(n))
\end{tikzcd}$$

To show that the dotted arrow exists it will suffice to show that the left-hand side can be written as an iterated pushout of anodyne morphisms in the cartesian model structure. This follows from the (more general) argument given in [3, Lemma 3.41] by forgetting the scaling. □

Definition 4.16 Let $\pi : \mathcal{C} \to \mathcal{D}$ be a fibration of ∞-categories. We say that an object $x \in \mathcal{C}$ is *p-initial* if for every $y \in \mathcal{C}$ the functor π yields a homotopy equivalence

$$\mathrm{Map}_{\mathcal{C}}(x, y) \simeq \mathrm{Map}_{\mathcal{D}}(\pi(x), \pi(y)).$$

Remark 4.17 An object $x \in \mathcal{C}$ as above is p-initial if and only if for every $n \geqslant 1$ the lifting problems

$$\begin{tikzcd} \partial\Delta^n \arrow[r, "\varphi"] \arrow[d] & \mathcal{C} \arrow[d, "\pi"] \\ \Delta^n \arrow[r] \arrow[ur, dotted] & \mathcal{D} \end{tikzcd}$$

admit a solution provided $\varphi(0) = x$.

Lemma 4.18 *Let* $\mathfrak{p} : \mathbb{X} \to \mathbb{S}$ *be a local* $(0,1)$*-fibration of* ∞*-bicategories. Then an edge* $e : \Delta^1 \to \mathbb{X}$ *is locally* $(0,1)$*-cartesian if and only if for every* $n \geqslant 2$ *the lifting problems of the form*

$$\begin{tikzcd} \Lambda^n_0 \arrow[r, "\varphi"] \arrow[d] & \mathbb{X} \arrow[d, "\mathfrak{p}"] \\ \Delta^n \arrow[r, "\phi"] \arrow[ur, dotted] & \mathbb{S} \end{tikzcd}$$

admit a solution provided $\varphi(0 \to 1) = e$, $\varphi(0 \to 1 \to n)$ *is thin whenever* $n \geqslant 2$ *and* $\phi(1 \to n)$ *is an equivalence in* $\mathbb{S}$*. If* $n = 2$ *we require* $\phi(0 \to 1 \to n)$ *to be thin,* $\phi(1 \to n)$ *to be an equivalence in* $\mathbb{S}$ *and that our solution is a thin simplex in* $\mathbb{X}$*.*

Proof First let us remark that if we can produce solutions to those lifting problems, e must be a $(0,1)$-cartesian edge once restricted to Δ^1 and the claims follow from Proposition 4.11. We now prove the converse.

The case $n = 2$ is precisely the first condition in Definition 4.8. To tackle the cases $n \geqslant 3$ we will assume once more that $\mathfrak{p} : \mathrm{N}^{\mathbf{sc}}(\mathbb{C}) \to \mathrm{N}^{\mathbf{sc}}(\mathbb{D})$. Since $\varphi(0 \to 1 \to n)$ is thin we can solve the lifting problem

$$\begin{tikzcd} \mathfrak{C}^{\mathbf{sc}}[\Lambda^n_0](0,n) \arrow[r] \arrow[d] & \mathbb{C}(\varphi(0), \varphi(n)) \arrow[d] \\ \mathfrak{C}^{\mathbf{sc}}[\Delta^n](0,n) \arrow[r] \arrow[ur, dotted] & \mathbb{D}(\mathfrak{p}\varphi(0), \mathfrak{p}\varphi(n)) \end{tikzcd}$$

To conclude the proof we must show that we can produce the dotted arrow in

$$\begin{tikzcd} \mathfrak{C}^{\mathbf{sc}}[\Lambda^n_0](1,n) \arrow[r] \arrow[d] & \mathbb{C}(\varphi(1), \varphi(n)) \arrow[d, "\phi"] \\ \mathfrak{C}^{\mathbf{sc}}[\Delta^n](1,n) \arrow[r] \arrow[ur, dotted] & \mathbb{C}(\varphi(0), \varphi(n) \times_{\mathbb{D}(\mathfrak{p}\varphi(0), \mathfrak{p}\varphi(n))} \mathbb{D}(\mathfrak{p}\varphi(1), \mathfrak{p}\varphi(n)) \end{tikzcd}$$

However, by (ii) in Definition 4.8 the object $1n$ on the left-hand side gets mapped to a ϕ-initial object in $\mathbb{C}(\varphi(1), \varphi(n))$. Since the left-most vertical map can be obtained as an iterated pushout along boundary inclusions $\partial\Delta^n \to \Delta^n$, where the initial object is always $1n$, we conclude that the dotted arrow above can be constructed. □

Lemma 4.19 *Let* $\mathfrak{p} : \mathbb{X} \to \mathbb{S}$ *be a local* $(0,1)$*-fibration of* ∞*-bicategories. Suppose that we are given a simplex* $\sigma : \Delta^4 \to \mathbb{X}$ *such that the collection of triangles*

$$T = \{\Delta^{\{0,2,4\}}, \Delta^{\{1,2,3\}}, \Delta^{\{0,1,3\}}, \Delta^{\{1,3,4\}}, \Delta^{\{0,1,2\}}\}$$

gets mapped to 2*-simplices representing cartesian* 2*-morphisms in the corresponding mapping categories. Then the triangles* $\Delta^{\{0,1,4\}}$ *and* $\Delta^{\{0,3,4\}}$ *also represent cartesian* 2*-morphisms in* $\mathbb{X}(\sigma(0),\sigma(4))$*.*

Proof As usual, we will assume that our functor is of the form $\mathfrak{p} : \mathrm{N}^{\mathbf{sc}}(\mathbb{C}) \to \mathrm{N}^{\mathbf{sc}}(\mathbb{D})$. This allows us to reduce our problem to show that certain edges in $\mathcal{P} = \mathfrak{C}^{\mathbf{sc}}[\Delta^4](0,4)$ get mapped to cartesian edges in $\mathbb{C}(\sigma(0),\sigma(4))$. More specifically, we view $\mathcal{P}$ as the poset

$$\begin{tikzcd}
 & 0134 \arrow[rr, "\circ" marking] & & 01234 \\
034 \arrow[ur, "\circ" marking] \arrow[rr] & & 0234 \arrow[ur, "\circ" marking] & \\
 & 014 \arrow[uu, "\circ" marking] \arrow[rr] & & 0124 \arrow[uu] \\
04 \arrow[uu] \arrow[ur] \arrow[rr, "\circ" marking] & & 024 \arrow[uu] \arrow[ur, "\circ" marking] &
\end{tikzcd}$$

where the circled arrows are mapped by assumption to cartesian edges (note that to see this it is crucial to use that $\mathfrak{p}$ is cartesian-enriched) and we wish to show that $04 \to 014$ and $04 \to 034$ are mapped to cartesian edges in $\mathbb{C}(\sigma(0),\sigma(4))$.

Since $\pi : \mathbb{C}(\sigma(0),\sigma(4)) \to \mathbb{D}(\mathfrak{p}\sigma(0),\mathfrak{p}\sigma(4))$ is a cartesian fibration this is equivalent to require that certain morphisms are equivalences in the fibre over $\pi(04) = \alpha$. Using the functoriality of π we can move the diagram above to a diagram in the fibre over α where now the circled arrow are equivalences. It is easy to see that we can produce now an inverse as in Proposition 3.1.13 in [19]. □

Definition 4.20 Let $\mathfrak{p} : \mathbb{X} \to \mathbb{S}$ be a local $(0,1)$-fibration of ∞-bicategories and let $\mathbb{X} = (X, T_X)$ and $\mathbb{S} = (S, T_S)$. We define an **MB**-simplicial set $\mathrm{F}_{\mathfrak{p}}$ as follows:

- The underlying simplicial set of $\mathrm{F}_{\mathfrak{p}}$ is $\mathbb{X}$.
- An edge is declared to be marked if an only if it is a local $(0,1)$-cartesian edge in $\mathbb{X}$.
- A triangle is declared to be lean if its associated 2-morphism is a cartesian edge in $\mathbb{X}(a,b)$.
- A triangle is declared to be thin if it is lean and its image in $\mathbb{S}$ belongs to M_S.

This definition clearly yields a map $\pi_{\mathfrak{p}} : \mathrm{F}_{\mathfrak{p}} \to (S, \sharp, M_S \subseteq \sharp)$ which we call the **MB**-model of $\mathfrak{p}$.

Lemma 4.21 *Let* $\mathfrak{p} : \mathbb{X} \to \mathbb{S}$ *be a local* $(0,1)$*-fibration of* ∞*-bicategories. Then its associated* ***MB****-model (see Definition 4.20)* $\pi_{\mathfrak{p}} : \mathrm{F}_{\mathfrak{p}} \to (S, \sharp, M_S \subseteq \sharp)$ *defines a fibrant object in* $\left(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\right)_{/(S,\sharp,M_S\subset\sharp)}$.

Proof We need to show that $\pi_{\mathfrak{p}}$ has the right lifting property against the class of morphisms in Definition 3.4. It follows from Lemmas 4.15 and 4.18 that $\pi_{\mathfrak{p}}$ has the RLP property with respect to the morphisms of type (A1) and (A3) in Definition 3.4. Lemma 5.27 shows $\pi_{\mathfrak{p}}$ has the right lifting property against the class of morphisms of type (A2). The rest of the lifting problems follow immediately and thus our result is proved. □

Theorem 4.22 *Let* $\mathbb{S}=(S,T_S)$ *be a fibrant scaled simplicial set and let* $p:X\to S$ *be a fibrant object in* $\big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)_{/(S,\sharp,M_S\subset\sharp)}$*. Then its bicategorical interpretation (see Definition 4.4)* $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ *is a local* $(0,1)$*-fibration. Moreover, this assignment admits an inverse given by sending each local* $(0,1)$*-fibration of* ∞*-bicategories* $\mathfrak{q}:\mathbb{Y}\to\mathbb{S}$ *to its* ***MB****-model as in Definition 4.20.*

Proof Let $p:X\to S$ be a fibrant object in $\big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)_{/(S,\sharp,M_S\subset\sharp)}$. Then it follows from Propositions 4.5 and 4.6 that $p:\mathbb{X}\to\mathbb{S}$ is cartesian-enriched. To show that our map in question defines a local $(0,1)$-fibration it will be enough to show that the marked edges in X are local $(0,1)$-cartesian edges. However, one easily sees that a marked edge e is $(0,1)$-cartesian over Δ^1 and thus the claim holds.

Note that Lemma 4.21 shows that the **MB**-model of a local $(0,1)$-fibration of ∞-bicategories defines a fibrant object in $\big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)_{/(S,\sharp,M_S\subset\sharp)}$. To complete the proof, we must show that these two assignments are mutually inverse. Note that this simply amounts to verifying that the decorations of our simplicial sets remain unchanged after applying both procedures. This follows by virtue of the following observations:

- Let $p:X\to S$ be a fibrant object in $\big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)_{/(S,\sharp,M_S\subset\sharp)}$. Then a 2-simplex $\sigma:\Delta^2\to X$ is lean if and only if the associated 2-morphism in its bicategorical interpretation $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ is cartesian.
- Let $\mathfrak{q}:\mathbb{Y}\to\mathbb{S}$ be a local $(0,1)$-fibration of ∞-bicategories. Then a 2-simplex $\sigma:\Delta^2\to\mathbb{Y}$ is thin if and only if its associated 2-morphism is cartesian and its image in $\mathbb{S}$ is invertible. □

Proposition 4.23 *Let* $\mathfrak{f}:\mathbb{X}\to\mathbb{Y}$ *be a morphism of local* $(0,1)$*-fibrations. Then the following are equivalent:*

(i) *The map* $\mathfrak{f}$ *is a bicategorical equivalence.*

(ii) *For every* $s\in\mathbb{S}$ *the map* $\mathfrak{f}$ *induces an equivalence on fibres* $\mathfrak{f}_s:\mathbb{X}_s\xrightarrow{\simeq}\mathbb{Y}_s$*.*

Proof Let us suppose that $\mathfrak{f}$ is a bicategorical equivalence and let $s\in\mathbb{S}$. First we show that $\mathfrak{f}_s$ is essentially surjective. Given $y_s\in\mathbb{Y}_s$, we use that $\mathfrak{f}$ is essentially surjective to get some $x\in\mathbb{X}$ such that $\mathfrak{f}(x)\simeq y$. We denote by u the image of the equivalence $\mathfrak{f}(x)\simeq y$ under $\mathfrak{q}:\mathbb{Y}\to\mathbb{S}$ and use the fact that $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ is an isofibration to get an equivalence $v:x\to x_s$ where $\mathfrak{p}(x_s)=s$. It follows that $\mathfrak{f}(v)$ is again an equivalence and therefore defines a local $(0,1)$-cartesian edge in $\mathbb{Y}$ which allows to construct an equivalence $\mathfrak{f}(x_s)\simeq y_s$ lying over the identity on s. To show fully faithfulness of $\mathfrak{f}_s$ we consider $a,b\in\mathbb{X}_s$ and observe that we have a map of cartesian fibrations

$$\begin{tikzcd}
\mathbb{X}(a,b) \arrow[rr] \arrow[dr] & & \mathbb{Y}(\mathfrak{f}(a),\mathfrak{f}(b)) \arrow[dl] \\
& \mathbb{S}(s,s) &
\end{tikzcd}$$

which is an equivalence by assumption. It then follows that we have an equivalence after taking the fibre over the identity map on s. This morphism is then identified using Proposition 4.11 with the map

$$\mathbb{X}_s(a,b)\xrightarrow{\simeq}\mathbb{Y}_s(\mathfrak{f}(a),\mathfrak{f}(b)),$$

which shows that $\mathfrak{f}_s$ is fully faithful.

To show that the converse holds we note that by assumptions $\mathfrak{f}$ is already essentially surjective. It will then be enough to show that for every $a, b \in \mathbb{X}$ and every $\alpha : \mathfrak{p}(a) \to \mathfrak{p}(b)$ the induced morphism on fibres

$$\mathbb{X}(a,b)_\alpha \to \mathbb{Y}(\mathfrak{f}(a), \mathfrak{f}(b))_\alpha$$

is a categorical equivalence. By picking a cartesian lift $e : a \to \hat{b}$ such that $\mathfrak{p}(e) = \alpha$ we can again use Proposition 4.11 to produce a commutative diagram

$$\begin{tikzcd}
\mathbb{X}_{\mathfrak{p}(b)}(\hat{b}, b) \arrow[r, "\simeq"] \arrow[d, "\simeq"] & \mathbb{Y}_{\mathfrak{p}(b)}(\mathfrak{f}(\hat{b}), \mathfrak{f}(b)) \arrow[d, "\simeq"] \\
\mathbb{X}(a,b)_\alpha \arrow[r] & \mathbb{Y}(\mathfrak{f}(a), \mathfrak{f}(b))_\alpha
\end{tikzcd}$$

where we use two-out-of-three to conclude that the bottom horizontal morphism is a weak equivalence and thus our claim holds. □

Proposition 4.24 *Let (T, U) be a scaled simplicial set and consider a morphism $f : X \to Y$ between fibrant objects in $(\mathrm{Set}_\Delta^{\mathbf{mb}})_{/(T,\sharp,U\subset\sharp)}$. Suppose further that the following conditions hold*:

1. *The map f has the right lifting property against the class of* ***MB****-anodyne morphisms.*
2. *For every $t \in T$ the induced morphism $f_t : X_t \to Y_t$ is a bicategorical equivalence.*

Then f is a trivial fibration of ***MB****-simplicial sets.*

Proof We claim that f is a trivial fibration of **MB**-simplicial sets if and only if for every minimally scaled simplex $\Delta^n_\flat$ and every morphism $\sigma : \Delta^n_\flat \to T$ the restricted morphism

$$f_{|\sigma} : X \times_{\Delta^n_\flat} \Delta^n_\flat \to Y \times_{\Delta^n_\flat} \Delta^n_\flat$$

is a trivial fibration of **MB**-simplicial sets. One direction is obviously true. Let us assume that $f_{|\sigma}$ is always a trivial fibration. Then it is clear that f has the right-lifting property against the morphisms

- $(\partial\Delta^n, \flat, \flat) \to (\Delta^n, \flat, \flat)$,
- $(\Delta^2, \flat, \flat) \to (\Delta^2, \flat, \flat \subset \sharp)$,
- $(\Delta^1, \flat, \sharp) \to (\Delta^1, \sharp, \sharp)$.

Since a thin triangle in X is just a lean triangle lying over a thin triangle in T it follows that f also detects thin triangles and the claim holds.

Let us assume without loss of generality that $T = \Delta^n_\flat$ and consider the associated diagram (see Theorem 4.22) of local (0, 1)-fibrations ∞-bicategories

$$\begin{tikzcd}
\mathbb{X} \arrow[rr, "\mathfrak{f}"] \arrow[dr, "\mathfrak{p}"'] & & \mathbb{Y} \arrow[dl, "\mathfrak{q}"] \\
& \Delta^n &
\end{tikzcd}$$

First, let us show that $\mathfrak{f}$ is a trivial fibration of ∞-bicategories. Observe that our assumptions together with Proposition 4.23 imply that $\mathfrak{f}$ is a bicategorical equivalence. Moreover, $\mathfrak{f}$ has the right lifting property against the class of scaled anodyne maps given in Definition 2.3. It will then suffice by Remark 2.12 to show that $\mathfrak{f}$ is an isofibration. Given an equivalence in $e:\Delta^1\to\mathbb{Y}$ it follows that its image in Δ^n must be degenerate. Since this lifting problem is occurring in a fibre and by our assumptions the maps $\mathfrak{f}_t:\mathbb{X}_t\to\mathbb{Y}_t$ are trivial fibrations of scaled simplicial sets it follows that $\mathfrak{f}$ is an isofibration.

To finish the proof, we must show that $\mathfrak{f}$ detects local $(0,1)$-cartesian edges. Given an edge $e:\Delta^1\to\mathbb{X}$ such that $\mathfrak{f}(e)$ is a local $(0,1)$-cartesian edge we consider a local $(0,1)$-cartesian edge $u:\Delta^1\to\mathbb{X}$ such that $u(0)=e(0)$ and such that $\mathfrak{f}(u)=\mathfrak{f}(e)$. It follows that we have an edge $\alpha:u(1)\to e(1)$ such that $\alpha\circ u\simeq e$. Moreover, $\mathfrak{f}(\alpha)$ is an equivalence and lies over a degenerate morphism in Δ^n. We see then that α must be an equivalence in $\mathbb{X}$ and consequently e is a local $(0,1)$-cartesian edge. □

Definition 4.25 Let $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ be bicategorical fibration of ∞-bicategories and let $\sigma:\Delta^2\to\mathbb{S}$ be a thin triangle. We say that an edge $e:a\to b$ in $\mathbb{X}$ lying over $\sigma(0\to 1)$ is σ-local if the following conditions hold:

(i) For every $g:a\to c$ in $\mathbb{X}$ lying over $\sigma(0\to 2)$ there exists some $\widehat{\alpha}:b\to c$ and a thin simplex θ exhibiting $e\circ\widehat{\alpha}\simeq g$ such that $\mathfrak{p}(\theta)=\sigma$.

(ii) For any $\phi:b\to c$ such that $e\circ\phi\simeq g$ with associated simplex τ such that $\mathfrak{p}(\tau)=\sigma$ and for any $\varphi:b\to c$, precomposition along e induces a pullback diagram of spaces

$$\begin{tikzcd}
\mathrm{Map}_{\mathbb{X}(b,c)}(\phi,\varphi) \arrow[r] \arrow[d] & \mathrm{Map}_{\mathbb{X}(a,c)}(\phi\circ e,\varphi\circ e) \arrow[d] \\
\mathrm{Map}_{\mathbb{S}(\mathfrak{p}(b),\mathfrak{p}(c))}(\mathfrak{p}(\phi),\mathfrak{p}(\varphi)) \arrow[r] & \mathrm{Map}_{\mathbb{S}(\mathfrak{p}(a),\mathfrak{p}(c))}(\mathfrak{p}(\phi\circ e),\mathfrak{p}(\varphi\circ e))
\end{tikzcd}$$

Remark 4.26 Definition 4.25 implies that an edge is $(0,1)$-cartesian if and only if it is σ-local for every thin simplex $\sigma:\Delta^2\to\mathbb{S}$. Similarly, an edge is locally $(0,1)$-cartesian if and only if it is σ-local for every invertible 2-morphism (see Definition 4.1).

Proposition 4.27 *Let $\mathfrak{p}:\mathbb{X}\to\mathbb{S}$ be a local $(0,1)$-fibration of ∞-bicategories and let $\sigma:\Delta^2\to\mathbb{S}$ be a thin triangle. Then a local $(0,1)$-cartesian edge $e:a\to b$ is σ-local if and only if for every local $(0,1)$-cartesian edge $u:b\to c$ such that the composite $v\simeq u\circ e$ lies over σ then v is also a local $(0,1)$-cartesian edge.*

Proof Let us assume that e is σ-local and suppose that we have u and v as above. Let us suppose that we have $h:a\to d$ and let us show that condition (i) in Definition 4.8 is satisfied.

Since $\mathfrak{p}$ is a local $(0,1)$-fibration and, in particular, cartesian-enriched, we only need to show that v is $(0,1)$-cartesian once after pulling back along $\mathfrak{p}(v)$. Therefore, we can assume without loss of generality that $\mathfrak{p}(h)=\mathfrak{p}(v)$. We observe that since e is σ-local we can obtain a morphism $\alpha:b\to d$ such that $\alpha\circ e\simeq h$. Furthermore, we can use that u is a local $(0,1)$-cartesian edge to get a morphism $\phi:c\to d$ such that $\alpha\simeq\phi\circ u$. It follows that $h\simeq v\circ\phi$ and so the first condition holds.

Given $\phi : c \to d$ as above and another $\varphi : c \to d$ such that $\mathfrak{p}(\varphi) = \mathrm{id}$ we construct the commutative diagram

$$\begin{tikzcd}
\mathrm{Map}_{\mathbb{X}(c,d)}(\phi,\varphi) \arrow[r] \arrow[d] & \mathrm{Map}_{\mathbb{X}(b,d)}(\phi\circ u,\varphi\circ u) \arrow[r] \arrow[d] & \mathrm{Map}_{\mathbb{X}(a,d)}(\phi\circ v,\varphi\circ v) \arrow[d] \\
\mathrm{Map}_{\mathbb{S}(\mathfrak{p}(c),\mathfrak{p}(d))}(\mathrm{id},\mathrm{id}) \arrow[r] & \mathrm{Map}_{\mathbb{S}(\mathfrak{p}(b),\mathfrak{p}(d))}(\mathfrak{p}(u),\mathfrak{p}(u)) \arrow[r] & \mathrm{Map}_{\mathbb{S}(\mathfrak{p}(a),\mathfrak{p}(d))}(\mathfrak{p}(v),\mathfrak{p}(v))
\end{tikzcd}$$

We observe that the outer commutative diagram is obtained by pasting two pullback diagrams so it must be itself a pullback diagram. It follows that v is a local $(0,1)$-cartesian edge.

We wish now to show that the converse holds. Let $h : a \to d$ be an edge over $\sigma(0 \to 2)$. We take a local $(0,1)$-cartesian lift $u : b \to c$ of $\sigma(1 \to 2)$. Since by assumption $v \simeq u \circ e$ is again local $(0,1)$-cartesian we obtain a certain $\Xi : c \to d$ such that $h \simeq \Xi \circ v$. We can then set $\phi = \Xi \circ u$. It is then clear that $\phi \circ e = \Xi \circ u \circ e \simeq \Xi \circ v \simeq h$ and thus condition (i) in Definition 4.25 holds. Let $\phi : b \to d$ as above and assume we are given any other $\varphi : b \to d$. We wish to show that the associated commutative diagram (see (ii) in Definition 4.25) of spaces is cartesian. We note that a totally analogous argument as in Proposition 4.11 shows that it is enough to show that the associated map of fibres

$$\mathrm{Map}_{\mathbb{X}(b,d)}(\phi,\varphi)_{\mathfrak{p}(u)} \xrightarrow{\simeq} \mathrm{Map}_{\mathbb{X}(a,e)}(\phi\circ e,\varphi\circ e)_{\mathfrak{p}(v)}$$

is an equivalence whenever $\mathfrak{p}(\phi) = \mathfrak{p}(\varphi)$. Since u is a local $(0,1)$-cartesian edge we can find morphisms $\widetilde{\phi}, \widetilde{\varphi} : c \to d$ such that $\widetilde{\phi} \circ u \simeq \phi$ and $\widetilde{\varphi} \circ u = \varphi$. We can then produce the morphisms

$$\mathrm{Map}_{\mathbb{X}(c,d)}(\widetilde{\phi},\widetilde{\varphi})_{\mathrm{id}} \xrightarrow{\simeq} \mathrm{Map}_{\mathbb{X}(b,d)}(\phi,\varphi)_{\mathfrak{p}(u)} \to \mathrm{Map}_{\mathbb{X}(a,e)}(\phi\circ e,\varphi\circ e)_{\mathfrak{p}(v)}.$$

We conclude the proof by noting that the composite map must also be a weak equivalence since v is by assumption a local $(0,1)$-cartesian edge. □

Definition 4.28 Let $\mathbb{S}$ be an ∞-bicategory and let U be a subcollection of the thin triangles in $\mathbb{S}$ which contains the invertible 2-morphisms of $\mathbb{S}$ (see Definition 4.1). We say that a local $(0,1)$-fibration $\mathfrak{p} : \mathbb{X} \to \mathbb{S}$ is $\mathcal{U}$-local if given a pair of local $(0,1)$-cartesian edge $u, v : \Delta^1 \to \mathbb{X}$ and a thin 2-simplex σ pictured as

$$\begin{tikzcd}
& b \arrow[dr, "v"] & \\
a \arrow[ur, "u"] \arrow[rr, "w"] & & c
\end{tikzcd}$$

such that $p(\sigma) \in \mathcal{U}$ then we have that w is also locally $(0,1)$-cartesian. If $\mathcal{U}$ consists in all thin triangles we say that $\mathfrak{p} : \mathbb{X} \to \mathbb{S}$ is a $(0,1)$-cartesian fibration.

Theorem 4.29 *Let (S, T_S) be a fibrant scaled simplicial set and let $U \subset T_S$ be a subset containing all invertible 2-morphisms. Given a fibrant object $(\mathrm{Set}_{\Delta}^{\mathbf{mb}})_{/(S,\sharp,U\subset\sharp)}$, its bicategorical interpretation $\mathfrak{p} : \mathbb{X} \to \mathbb{S}$ defines a $\mathcal{U}$-local fibration. Conversely any $\mathcal{U}$-local fibration defines canonically a fibrant object in $(\mathrm{Set}_{\Delta}^{\mathbf{mb}})_{/(S,\sharp,U\subset\sharp)}$.*

Proof Let us assume that we are given a fibrant object in $(\mathrm{Set}_{\Delta}^{\mathbf{mb}})_{/(S,\sharp,U\subset\sharp)}$. In particular, we can use Theorem 4.22 to obtain a local $(0,1)$-fibration $\mathfrak{p} : \mathbb{X} \to \mathbb{S}$ of ∞-bicategories. Since our original object has the right lifting property against the class (S1) it follows that our local $(0,1)$-cartesian edges compose across triangles which lie over triangles in $\mathcal{U}$ and the claim follows.

We now show the converse. Note that due to Proposition 4.27 our local $(0,1)$-cartesian edges are σ-local with respect to the elements of $\mathcal{U}$. The only thing that we need to prove is that given a $\mathcal{U}$-local fibration we can produce the dotted arrow in

$$\begin{tikzcd}
\Lambda_0^n \arrow[r, "f"] \arrow[d] & \mathbb{X} \arrow[d, "\mathfrak{p}"] \\
\Delta^n \arrow[r] \arrow[ur, dotted] & \mathbb{S}
\end{tikzcd}$$

where $f(0 \to 1)$ is locally $(0,1)$-cartesian and $f(0 \to 1 \to n)$ lands in $\mathcal{U}$. The proof of this fact is essentially the same as the proof in Lemma 4.18 and therefore left as an exercise. □

5 The Grothendieck construction

5.1 The Quillen adjunction

Let $S = (S, T_S)$ be a scaled simplicial set and let $\mathfrak{C}^{\mathrm{sc}}[S]$ denote the scaled rigidification (Definition 2.9) of (S, T_S). The goal of this section is to prove the following theorem.

Theorem 6 *Let S be a scaled simplicial set. Then there exists a Quillen equivalence*

$$\mathbb{S}\mathrm{t}_S : (\mathrm{Set}_{\Delta}^{\mathbf{mb}})_{/(S,\sharp,T_S\subset\sharp)} \rightleftarrows \mathrm{Fun}(\mathfrak{C}^{\mathrm{sc}}[S], \mathrm{Set}_{\Delta}^{\mathbf{ms}}) : \mathbb{U}\mathrm{n}_S$$

between the model structure on $(0,1)$*-cartesian fibrations over S and the projective model structure of* $\mathrm{Set}_{\Delta}^{+}$*-enriched functors with values in marked-scaled simplicial sets.*

Our first order of business will be to define the left adjoint $\mathbb{S}\mathrm{t}_S$ which will be given by a 2-categorical enhancement of the straightening functor constructed in Section 3 of [19]. Before we present our main construction, we need to give some preliminary definitions.

Definition 5.1 Let $X, Y \in \mathrm{Set}_{\Delta}^{\mathbf{ms}}$. We define the Gray tensor product $X \otimes Y \in \mathrm{Set}_{\Delta}^{\mathrm{sc}}$ (see Definition 4.1.1 in [10]) as follows:

(1) The underlying simplicial set of $X \otimes Y$ is given by $X \times Y$, the cartesian product of the underlying simplicial sets.

(2) Given a simplex $\sigma : \Delta^2 \to X \otimes Y$ let us denote by σ_X and σ_Y the projections to the corresponding factors in the cartesian product. We say that σ is scaled in $X \otimes Y$ if and only if the following conditions hold:

(i) The projection of the simplex σ is both scaled in X and in Y.

(ii) The restriction $\sigma_X(1 \to 2)$ is marked in X *or* the restriction $\sigma_Y(0 \to 1)$ is marked in Y.

Given marked scaled simplicial sets X, C we define marked scaled simplicial sets $\operatorname{Fun}^{\mathrm{gr}}(X,C)$ and $\operatorname{Fun}^{\mathrm{opgr}}(X,C)$ by means of the universal properties,

$$\operatorname{Hom}_{\mathrm{Set}^{\mathrm{sc}}_{\Delta}}(A,\operatorname{Fun}^{\mathrm{gr}}(X,C))\cong \operatorname{Hom}_{\mathrm{Set}^{\mathrm{sc}}_{\Delta}}(A\otimes X,C),$$
$$\operatorname{Hom}_{\mathrm{Set}^{\mathrm{sc}}_{\Delta}}(A,\operatorname{Fun}^{\mathrm{opgr}}(X,C))\cong \operatorname{Hom}_{\mathrm{Set}^{\mathrm{sc}}_{\Delta}}(X\otimes A,C).$$

Definition 5.2 Let $n\geqslant 0$. We define a poset P_n as follows:

- The objects are given by subsets $S\subseteq[n]$ such $S\neq\varnothing$ and $\max(S)=n$.
- We define a partial order on P_n by declaring $S\leqslant T$ whenever $\min(S)\leqslant\min(T)$ and there exists some U such that $\min(U)=\min(S)$ and $\max(U)=\min(T)$ and such that $S\subseteq U\cup T$.

Remark 5.3 In the definition above, $U\leqslant V$ if and only if $\min(U)\leqslant\min(V)$ and for every $x\in U$ such that $x\geqslant\min(V)$ then $x\in V$. Moreover, we can identify those inequalities $U<V$ in P_n which cannot be decomposed as $U<W<V$ as:

(O1) We have $U<V$ with $\min(U)=\min(V)$ and $V=U\cup\{s\}$.

(O2) We have $U<V$ with $V=U\setminus\min(U)$.

Remark 5.4 Given $S,T\in P_n$ such that $S\leqslant T$ we can have several subsets U as above such that $S\subseteq U\cup T$. Moreover, we can order such subsets by inclusion and define $U_{S,T}$ to be the minimal subset such that $S\subseteq U_{S,T}\cup T$. Let $\min S=s$ and let $\min T=t$. We can then write $U_{S,T}=\{s,t\}\cup\{s<i<t\mid i\in S\}$.

Definition 5.5 Let $\mathcal{P}_n=\mathrm{N}(P_n)$. We promote $\mathcal{P}_n$ to a scaled simplicial set as follows. Given a 2-simplex σ represented by $S\leqslant T\leqslant W$ we declare σ to be thin if $U_{S,W}=U_{S,T}\cup U_{T,W}$.

Remark 5.6 Let $\Delta^n_\flat=(\Delta^n,\flat,\flat)$ and let $\Delta^1\otimes_\flat\Delta^n=\Delta^1_\flat\otimes\Delta^n_\flat$. We consider $\mathfrak{C}^{\mathbf{sc}}[\Delta^1\otimes_\flat\Delta^n]$. Recall that given $(i,j)\leqslant(k,\ell)$ in $\Delta^1\times\Delta^n$ we have that $\mathfrak{C}^{\mathbf{sc}}[\Delta^1\otimes_\flat\Delta^n]((i,j),(k,\ell))$ is given by the nerve of the poset of chains C,

$$(i,j)=(i_0,j_0)<(i_1,j_1)<\cdots<(i_{\alpha-1},j_{\alpha-1})<(i_\alpha,j_\alpha)=(k,\ell),$$

ordered by refinement. Let us suppose that $i=0$ and that $k=1$. Then, given a chain $C=\{(i_\alpha,j_\alpha)\}_{\alpha\in A}$ we can define m_C to be the biggest index in A such that $i_{m_C}=0$. This allows us to define a map

$$\pi_{j,\ell}:\mathfrak{C}^{\mathbf{sc}}[\Delta^1\otimes_\flat\Delta^n]\,((0,j),(1,\ell))\to\mathcal{P}_{[j,\ell]},\quad C\mapsto\bigcup_{\alpha\geqslant m_C} j_\alpha.$$

This assignment is clearly a map of posets which sends marked edges in our mapping simplicial set to identities in $\mathcal{P}_{[j,l]}$. We use the map $\pi_{j,\ell}$ to equip the left-hand side with the scaling induced by $\mathcal{P}_{[j,l]}$.

Definition 5.7 We define a colimit preserving functor $\Pi:\mathrm{Set}^{\mathbf{mb}}_{\Delta}\to\mathrm{Cat}^{\mathbf{ms}}_{\Delta}$ with values in the category of $\mathrm{Set}^{\mathbf{ms}}_{\Delta}$-enriched categories by specifying its values on the generators under colimits of $\mathrm{Set}^{\mathbf{mb}}_{\Delta}$ as follows:

(1) Given a minimally marked and biscaled simplex $\Delta^n_\flat=(\Delta^n,\flat,\flat)$ we define $\Pi(\Delta^n)$ to have as underlying Set^+_Δ-category the scaled rigidification of the Gray tensor product $\Delta^1\otimes_\flat\Delta^n$ studied in Remark 5.6. Given $(i,j)<(k,\ell)$ in $\Pi(\Delta^n_\flat)$ we equip the mapping simplicial sets with a scaling by

declaring every triangle to be scaled if $i \neq 0$ and $k \neq 1$. If $i = 0$ and $k = 1$ we scale $\Pi(\Delta^n_\flat)((0,j),(1,\ell))$ according to Remark 5.6.

(2) Given a lean scaled 2-simplex, i.e., $\Delta^2_\dagger = (\Delta^2, \flat, \flat \subset \sharp)$ we define $\Pi(\Delta^2_\dagger)$ from $\Pi(\Delta^2_\flat)$ by scaling every triangle in the mapping simplicial sets.

(3) Given a thin scaled 2-simplex, i.e., $\Delta^2_\sharp = (\Delta^2, \flat, \sharp)$ we define $\Pi(\Delta^2_\sharp)$ from $\Pi(\Delta^2_\dagger)$ by additionally marking every morphism in $\Pi(\Delta^2_\sharp)((0,0),(1,2))$ which gets maps under the map in Remark 5.6 to the morphism $02 \to 012$ in $\mathcal{P}_2$.

(4) Given a marked edge $(\Delta^1)^\sharp = (\Delta^1, \sharp, \sharp)$ we can identify $\Pi((\Delta^1)^\sharp) = \mathfrak{C}^{\mathbf{sc}}[\Delta^1 \times \Delta^1]$.

One easily checks that our choice of decorations is compatible with composition and thus our definition yields well-defined $\mathrm{Set}^{\mathbf{ms}}_\Delta$-enriched categories and that our definition is functorial on the set of generators of $\mathrm{Set}^{\mathbf{mb}}_\Delta$. Since $\mathrm{Cat}^{\mathbf{ms}}_\Delta$ is cocomplete our functor can be extended by colimits and the definition is complete.

Definition 5.8 Let $j : \mathrm{Cat}^{+}_\Delta \to \mathrm{Cat}^{\mathbf{ms}}_\Delta$ be the functor that scales every 2-simplex in the mapping simplicial sets. Given a scaled simplicial set S we define a functor

$$\Pi_S : \left(\mathrm{Set}^{\mathbf{mb}}_\Delta\right)_{/(S,\sharp,T_S\subset\sharp)} \to \mathrm{Cat}^{\mathbf{ms}}_\Delta, \quad X \mapsto \Pi(X) \amalg_{j\circ\mathfrak{C}^{\mathrm{sc}}[X]} j \circ \mathfrak{C}^{\mathbf{sc}}[S],$$

where $\mathfrak{C}^{\mathbf{sc}}[X]$ denotes the scaled rigidification of the *underlying* scaled simplicial set of X and where the morphism $j \circ \mathfrak{C}^{\mathbf{sc}}[X] \to \Pi(X)$ is given by the inclusion of $\Delta^{\{1\}} \times X$.

We define a further functor

$$C_S : \left(\mathrm{Set}^{\mathbf{mb}}_\Delta\right)_{/(S,\sharp,T_S\subset\sharp)} \to \mathrm{Cat}^{\mathbf{ms}}_\Delta, \quad X \mapsto \Pi_S(X) \amalg_{j\circ\mathfrak{C}^{\mathrm{sc}}[X]} \Delta^0,$$

where the morphism $j \circ \mathfrak{C}^{\mathbf{sc}}[X] \to \Pi_S(X)$ is induced by the inclusion $\Delta^0 \times X \to \Delta^1 \otimes X$.

Remark 5.9 From this point on we will drop the notation $j \circ \mathfrak{C}^{\mathbf{sc}}$ and we will view Set^{+}_Δ-enriched categories as a full subcategory of $\mathrm{Set}^{\mathbf{ms}}_\Delta$-enriched categories consisting in those enriched categories whose mapping simplicial sets are fully scaled.

Remark 5.10 Let $f : S \to S'$ be a map of scaled simplicial sets. Given $p : X \to S$ in $\left(\mathrm{Set}^{\mathbf{mb}}_\Delta\right)_{/(S,\sharp,T_S\subset\sharp)}$ we claim that we have an isomorphism of $\mathrm{Set}^{\mathbf{ms}}_\Delta$-enriched categories

$$C_S(X) \amalg_{\mathfrak{C}^{\mathrm{sc}}[S]} \mathfrak{C}^{\mathbf{sc}}[S'] \xrightarrow{\simeq} C_{S'}(f_! X),$$

where $f_! X$ denotes the value of the functor C'_S at the object $f \circ p : X \to S'$. The isomorphism on the underlying Set^{+}_Δ-categories is clear. The only thing to show is that the scaling on mapping simplicial sets of the form $C_{S'}(f_! X)((0,x),(1,s'))$ is the same for both $\mathrm{Set}^{\mathbf{ms}}_\Delta$-enriched categories. This follows after a direct inspection since the scaling on those simplicial sets (which does not factor through some mapping simplicial set between objects $(1,s)$ and $(1,s')$) is independent of the base.

Definition 5.11 Let us denote by v the collapsed point in the definition of $C_S(X)$. Then for every $p : X \to S$ and every morphism of Set^{+}_Δenriched categories $\phi : \mathfrak{C}^{\mathbf{sc}}[S] \to \mathcal{C}$ we can define a functor

$$\mathbb{S}\mathrm{t}_\phi(X) : \mathcal{C} \to \mathrm{Set}^{\mathbf{ms}}_\Delta, \qquad c \mapsto \mathrm{Map}_{C_\phi(X)}(v, c), \quad \text{where } C_\phi(X) := C_S(X) \amalg_{\mathfrak{C}^{\mathrm{sc}}[S]} \mathcal{C},$$

which we call the *straightening* of $p : X \to S$. This definition extends to a functor

$$\mathbb{S}\mathrm{t}_{\phi} : \big(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\big)_{/(S,\sharp,T_S \subset \sharp)} \to \mathrm{Fun}(\mathcal{C}, \mathrm{Set}_{\Delta}^{\mathbf{ms}})$$

with values in the category of $\mathrm{Set}_{\Delta}^{+}$-enriched functors. If ϕ is an isomorphism we will use the notation $\mathbb{S}\mathrm{t}_S$.

Definition 5.12 Let $\mathcal{F} : \mathcal{C} \to \mathrm{Set}_{\Delta}^{\mathbf{mb}}$ where $\mathcal{C}$ is a $\mathrm{Set}_{\Delta}^{+}$-enriched category. We define $T_{\mathcal{F}} \in \mathrm{Cat}_{\Delta}^{\mathbf{ms}}$ as follows:

- The objects of $T_{\mathcal{F}}$ are those of $\mathcal{C}$ in addition to an object v which we call the cone point.
- The mapping marked-scaled simplicial sets are as follows: We declare $T_{\mathcal{F}}(x, v) = \varnothing$ unless $x = v$ in which case $T_{\mathcal{F}}(v, v) = \Delta^0$. Let $c, c' \in \mathcal{C}$. We declare $T_{\mathcal{F}}(v, c) = \mathcal{F}(c)$ and $T_{\mathcal{F}}(c, c') = \mathcal{C}(c, c')$ (cf. Remark 5.9).
- Given $a, b, c \in T_{\mathcal{F}}$ such that $a, b, c \neq v$, the composition rule is that of $\mathcal{C}$. If $a = v$, then the composition rule is given by functoriality of $\mathcal{F}$, i.e., $\mathcal{F}(b) \times \mathcal{C}(b, c) \to \mathcal{F}(c)$.

Proposition 5.13 *The straightening functor given in Definition 5.11 admits a right adjoint*

$$\mathbb{U}\mathrm{n}_{\phi} : \mathrm{Fun}(\mathcal{C}, \mathrm{Set}_{\Delta}^{\mathbf{ms}}) \to \big(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\big)_{/(S,\sharp,T_S \subset \sharp)},$$

which we call the ***unstraightening*** *functor.*

Proof By the adjoint functor theorem it suffices to show $\mathbb{S}\mathrm{t}_{\phi}$ preserves all colimits in $\big(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\big)_{/(S,\sharp,T_S \subset \sharp)}$. To see this, we observe that by construction the functor

$$C_{\phi} : \big(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\big)_{/(S,\sharp,T_S \subset \sharp)} \to \mathrm{Cat}_{\Delta}^{\mathbf{ms}}$$

preserves all colimits since it is built out of colimit-preserving functors. We now observe that a functor of $\mathrm{Set}_{\Delta}^{\mathbf{ms}}$-categories $C_{\phi}(X) \to T_{\mathcal{F}}$ (see Definition 5.12) which preserves cone points and restricts to the identity on $\mathcal{C}$ is precisely the data of a natural transformation $\mathbb{S}\mathrm{t}_{\phi}(X) \Rightarrow \mathcal{F}$. The claim now follows since C_{ϕ} preserves colimits. □

Proposition 5.14 *Let $f : S \to S'$ be a map of scaled simplicial sets and consider the commutative diagram of $\mathrm{Set}_{\Delta}^{+}$-enriched categories*

$$\begin{tikzcd}
\mathfrak{C}^{\mathbf{sc}}[S] \arrow[r, "\mathfrak{C}^{\mathbf{sc}}[f]"] \arrow[d, "\phi"] & \mathfrak{C}^{\mathbf{sc}}[S'] \arrow[d, "\phi'"] \\
\mathcal{C} \arrow[r, "\psi"] & \mathcal{C}'
\end{tikzcd}$$

Then

$$\begin{tikzcd}
\big(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\big)_{/(S,\sharp,T_S \subset \sharp)} \arrow[r, "\mathbb{S}\mathrm{t}_{\phi}"] \arrow[d, "f_!"] & \mathrm{Fun}(\mathcal{C}, \mathrm{Set}_{\Delta}^{\mathbf{ms}}) \arrow[d, "\psi_!"] \\
\big(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\big)_{/(S',\sharp,T_{S'} \subset \sharp)} \arrow[r, "\mathbb{S}\mathrm{t}_{\phi'}"] & \mathrm{Fun}(\mathcal{C}', \mathrm{Set}_{\Delta}^{\mathbf{ms}})
\end{tikzcd}$$

commutes up to invertible natural transformation where $\psi_!$ is the left adjoint to the restriction functor ψ^*.

Proof Let $\theta = \psi \circ \phi$. We will show that $\mathbb{S}\mathrm{t}'_{\phi} \circ f_! \simeq \mathbb{S}\mathrm{t}_{\theta}$ and that $\psi_! \circ \mathbb{S}\mathrm{t}_{\phi} \simeq \mathbb{S}\mathrm{t}_{\theta}$. We proceed case by case:

- To show that $\mathbb{S}\mathrm{t}'_{\phi} \circ f_! \simeq \mathbb{S}\mathrm{t}_{\theta}$, we consider the diagram of $\mathrm{Set}_{\Delta}^{\mathbf{ms}}$-categories

$$\begin{tikzcd}
\mathfrak{C}^{\mathbf{sc}}[S] \arrow[r] \arrow[d] & \mathfrak{C}^{\mathbf{sc}}[S'] \arrow[r] \arrow[d] & \mathcal{C}' \arrow[d] \\
C_S(X) \arrow[r] & C_{S'}(f_! X) \arrow[r] & C_{\phi'}(f_! X)
\end{tikzcd}$$

The left-most square is a pushout by Remark 5.10 and right-most is also a pushout square by definition. We conclude that $C_{\phi'}(f_! X) \cong C_{\theta}(X)$. Since our constructions are natural the claim holds.

- We now show that $\psi_! \circ \mathbb{S}\mathrm{t}_{\phi} \simeq \mathbb{S}\mathrm{t}_{\theta}$. We will show that for every $\varphi : \mathfrak{C}^{\mathbf{sc}}[S] \to \mathcal{C}$, we have that $\mathbb{S}\mathrm{t}_{\varphi} \cong \varphi_! \circ \mathbb{S}\mathrm{t}_S$ which immediately implies the claim. Let $\mathcal{F} : \mathcal{C} \to \mathrm{Set}_{\Delta}^{\mathbf{ms}}$ and recall the construction $T_{\mathcal{F}}$ given in Definition 5.12. Then it follows that commutative diagrams of the form

$$\begin{tikzcd}
\mathfrak{C}^{\mathbf{sc}}[S] \arrow[r] \arrow[d] & \mathcal{C} \arrow[d] \arrow[ddr, "\iota"] & \\
C_S(X) \arrow[r] \arrow[drr] & C_{\varphi}(X) \arrow[dr, dashed] & \\
 & & T_{\mathcal{F}}
\end{tikzcd}$$

where $\iota : \mathcal{C} \to T_{\mathcal{F}}$ is the obvious fully faithful functor, correspond to natural transformations $\mathbb{S}\mathrm{t}_S(X) \Rightarrow \varphi^* \mathcal{F}$. Since $C_{\varphi}(X)$ is a pushout we conclude that this data is equivalent to natural transformations $\mathbb{S}\mathrm{t}_{\varphi}(X) \Rightarrow \mathcal{F}$ which provides the desired isomorphism $\mathbb{S}\mathrm{t}_{\varphi}(X) \cong \varphi_! \, \mathbb{S}\mathrm{t}_S(X)$. Our constructions are natural and so the claim holds. □

Remark 5.15 In the situation above passing to right adjoints we obtain an isomorphism of functors $\mathbb{U}\mathrm{n}_{\phi} \circ \psi^* \cong f^* \circ \mathbb{U}\mathrm{n}_{\phi'}$.

Lemma 5.16 *For any simplicial set* (S, T_S) *the functor* $\mathbb{S}\mathrm{t}_S$ *preserves cofibrations.*

Proof Since $\mathbb{S}\mathrm{t}_S$ preserves colimits it will be enough to prove the claim on the generating class of cofibrations given in Definition 3.11. Moreover, given a cofibration $\alpha : A \to B$ we can use Proposition 5.14 to reduce to the case where $S = B$. The case $\varnothing \to \Delta^0$ is obviously true. For the rest of the generators we have that $B = (\Delta^n, \flat)$ for $n \geqslant 1$ or $B = (\Delta^2, \sharp)$ and that the map $\mathbb{S}\mathrm{t}_B \, A(i) \to \mathbb{S}\mathrm{t}_B \, B(i)$ is the identity except when $i = n$ in which case the map is a cofibration. It is immediate to see that the map in this situation $\mathbb{S}\mathrm{t}_B \, A \to \mathbb{S}\mathrm{t}_B \, B$ has the left lifting property against the class of trivial fibrations. □

Remark 5.17 Let $\iota : \mathrm{Set}_{\Delta}^{+} \to \mathrm{Set}_{\Delta}^{\mathbf{ms}}$ be the functor defined by $\iota(X, E_X) = (X, E_X, \sharp)$. Given a scaled simplicial set (S, T_S) let $\iota_* \, \mathrm{St}_S : (\mathrm{Set}_{\Delta}^{+})_{/S} \to \mathrm{Fun}(\mathfrak{C}^{\mathbf{sc}}[S], \mathrm{Set}_{\Delta}^{\mathbf{ms}})$ be the straightening functor given in Definition 3.5.4 in [19] postcomposed with the enriched functor ι. Recalling the definition of the functor R

(see Definition 3.30), we can then define a functor

$$\mathbb{S}\mathrm{t}_S \circ R : \big(\mathrm{Set}_\Delta^+\big)_{/S} \to \mathrm{Fun}(\mathfrak{C}^{\mathbf{sc}}[S], \mathrm{Set}_\Delta^{\mathbf{ms}}).$$

It follows that $\iota_* \mathrm{St}_S$ and $\mathbb{S}\mathrm{t}_S \circ R$ differ only in the scaling and thus induce a natural transformation $\eta_S : \mathbb{S}\mathrm{t}_S \circ R \Rightarrow \iota_* \mathrm{St}_S$.

Proposition 5.18 *Let (S, T_S) be a scaled simplicial set. Then $\eta_S : \mathbb{S}\mathrm{t}_S \circ R \Rightarrow \iota_* \mathrm{St}_S$ is an object-wise weak equivalence.*

Proof It is clear that both functors preserve colimits. Moreover, a totally analogous proof to that of Lemma 5.16 shows that $\iota_* \mathrm{St}_S$ preserves cofibrations. It is also not hard to verify that $\iota_* \mathrm{St}_S$ satisfies similar base change properties as those in Proposition 5.14. We conclude that it will be enough to show that $\eta_k = \eta_{\Delta^k_\flat}(\Delta^k)$ is an equivalence for $k \geqslant 0$ and similarly for $\eta_1^\sharp = \eta_{(\Delta^1)^\sharp}((\Delta^1)^\sharp)$. It is easy to check that η_0, η_1, η_2 and $\eta_1^\sharp$ are all isomorphisms.

For $k \geqslant 3$ let us define $\mathbb{L}^k = \Pi_{\Delta^k_\flat}((\Delta^k, \flat, \flat \subset \sharp))$ and $\mathbb{L}^k_\sharp$ by scaling every triangle in the mapping simplicial categories of $\mathbb{L}^k$. It is not hard to see that our problem can be reduced to showing that the map

$$\varphi_k : \mathbb{L}^k \to \mathbb{L}^k_\sharp$$

is an equivalence of $\mathrm{Set}_\Delta^{\mathbf{ms}}$-enriched categories. Using induction on k we can assume that the mapping simplicial sets $\mathbb{L}^k((0,i),(1,j))$ are maximally scaled except if $i = 0$ and $j = k$. Let $\mathbb{L}^k((0,0),(1,k)) = A^k$ and similarly $\mathbb{L}^k_\sharp((0,0),(1,k)) = A^k_\sharp$. We observe that for every face $d_i : \Delta^{k-1} \to \Delta^k$ we have a commutative diagram

$$\begin{tikzcd}
A^{k-1} \arrow[r, "\simeq"] \arrow[d, "\alpha_i"] & A^{k-1}_\sharp \arrow[d] \\
A^k \arrow[r] & A^k_\sharp
\end{tikzcd}$$

where the top horizontal morphism is a weak equivalence. We can therefore assume inductively that the triangles of A^k which are in the image of the maps α_i for $i = 0, \dots, k$ are all scaled.

Let $\sigma : C_0 \subset C_1 \subset C_2$ be a triangle in A^k and let us fix the notation $C_j = \{(\varepsilon_i^j, a_i^j\}_{i=0}^{\ell}$. Let $e = C_0 \subset C_1$ be an edge in A^k. We define $S(e)$ as the set of nondegenerate simplices with initial vertex C_0 and final vertex C_1. We finally set $|e| = \max\{\dim(\phi) \mid \phi \in S(e)\}$. Note that this a well-defined number. Given a 2-simplex σ as above we define $|\sigma| = |d_1(\sigma)|$. We will show that we can scale σ using a pushout along a **MS**-anodyne morphism using induction on $|\sigma| = \ell$. The case $\ell = 2$ follows easily after direct inspection. So we will assume from this point on that $\ell > 1$.

Let us suppose that the claim holds for those 2-simplices τ such that $|d_2(\tau)| = 1$ and for those 2-simplices σ such that $|\sigma| < \ell - 1$. Then given $\sigma : C_0 \subset C_1 \subset C_2$ such that $|d_2(\sigma)| > 1$ and such that $|\sigma| = \ell$ we can construct a 3-simplex $\rho : C_0 \subset D \subset C_1 \subset C_2$ such that the following holds:

(1) We have $|d_2 d_3(\rho)| = 1$; in particular, $d_i(\rho)$ is thin scaled for $i = 2, 3$.

(2) We have $|d_0(\rho)| = \ell - 1$ and is therefore scaled.

It follows that we can fully scale ρ using a pushout along a morphism of type (M2) in Definition 3.36. Therefore, we have reduced our problem to proving the claim above.

Let $\sigma : C_0 \subset C_1 \subset C_2$ such that $|d_2(\sigma)| = 1$. We observe that unless $C_0 = (0,0) < (1,k)$ then σ is scaled. Otherwise we could express σ as a certain composition in the category $\mathbb{L}^k$ and it would follow from the induction hypothesis that σ is thin.

We can now see that $C_1 = (0,0) < (\varepsilon, a) < (1,k)$ which leads us to consider cases depending on the parameter $\varepsilon \in \{0,1\}$.

($\boldsymbol{\varepsilon = 1}$) Then we can assume without loss of generality that C_2 contains an element of the form $(0,x)$ in C_2 with $x \neq 0$. This is true since otherwise the maps $\pi_{0,k}$ in Remark 5.6 would show that σ is already scaled. Moreover, we can further assume that there is only element of the form $(0,x)$ in C_2. Indeed, if we had some $(0,y) < (0,x)$ then we can produce a 3-simplex $\rho : C_0 \to C_1 \to \widetilde{C}_2 \to C_2$ where $\widetilde{C}_2 = C_2 \setminus \{(0,y)\}$. Since $d_0(\rho)$ and $d_1(\rho)$ must be scaled by definition it follows that we can scale $d_2(\rho)$ if and only if we can scale $d_3(\rho)$ and thus the claims follows. Additionally, we note that if $x \neq 1$ then σ factors through one of the morphisms $\alpha_j : A^{k-1} \to A^k$ above. We finally see that in this case $\pi_{0,k}(\sigma)$ is given by a simplex of the form $0n \to 0an \to S$ with $\min(S) = 1$ and it is consequently scaled in $\mathcal{P}_k$.

($\boldsymbol{\varepsilon = 0}$) Observe that if C_2 contains an element of the form $(0,x)$ with $x < a$ we can define $\widetilde{C}_2$ as above an produce a 3-simplex $\rho : C_0 \to C_1 \to \widetilde{C}_2 \to C_2$ which shows that we can scale $\sigma = d_2(\rho)$ if and only if we can scale $d_3(\rho)$. In a totally analogous way as in the case $\varepsilon = 1$ we can assume that $a = 1$. If C_2 does not contain any element of the form $(0,z)$ with $z > 1$ then σ must be already scaled. Moreover, we can assume without loss of generality that C_2 only contains one element of the form $(0,z)$ using a similar argument as before by constructing a certain $\widetilde{C}_2$. We can assume that in this case $z = 2$ since otherwise, the simplex factors through a certain morphism $\alpha_j : A^{k-1} \to A^k$. If C_2 does not contain an element of the form $(1,s)$ with $s \neq k$ it follows by direct inspection that σ is already scaled. If this is not the case we consider D which is obtained from C_2 by discarding every element of the form $(1,s)$ with $s \neq k$. Then we get a 3-simplex $\Xi : C_0 \to C_1 \to D \to C_2$. One easily checks that every face of Ξ is scaled except possibly $d_2(\Xi) = \sigma$ and thus our result follows. □

Corollary 5.19 *Let* (S, T_S) *be a scaled simplicial set. Consider a weak equivalence* $u : (A, E_A) \to (B, E_B)$ *in* $(\mathrm{Set}_\Delta^+)_{/S}$*. Then the functor* $\mathbb{S}\mathrm{t}_S$ *sends the morphism* $R(u) : (A, E_A, T_A \subset \sharp) \to (B, E_B, T_B \subset \sharp)$ *(see Definition 3.30) to a weak equivalence in* $\mathrm{Fun}(\mathfrak{C}^{\mathrm{sc}}[S], \mathrm{Set}_\Delta^{\mathbf{ms}})$*.*

Remark 5.20 Let $i : A \to B$ be a cofibration of **MB** simplicial sets which we view as a morphism in $(\mathrm{Set}_\Delta^{\mathbf{mb}})_{/B}$ and recall the definition of Π_B in Definition 5.8. We define $\Pi_B(A)^{\dagger}$ as the pushout

$$\begin{tikzcd}
j \circ \mathfrak{C}^{\mathrm{sc}}[A] \times \Delta^{\{0\}} \arrow[r] \arrow[d] & j \circ \mathfrak{C}^{\mathrm{sc}}[B] \times \Delta^{\{0\}} \arrow[d] \\
\Pi_B(A) \arrow[r] & \Pi_B(A)^{\dagger}
\end{tikzcd}$$

It follows from our definition that in order to check that $\mathbb{S}\mathrm{t}_B(A) \Rightarrow \mathbb{S}\mathrm{t}_B(B)$ is a pointwise weak equivalence it suffices to check that the induced map $\Pi_B(A)^{\uparrow} \to \Pi_B(B)$ is a weak equivalence of $\mathrm{Set}_\Delta^{\mathbf{ms}}$-categories.

Remark 5.21 Let $\theta : \Delta^{n+1} \to \Delta^1 \times \Delta^n$ be a nondegenerate simplex and let $i \in \Delta^{n+1}$ be the biggest element such that $\theta(i) = (i, 0)$. We will use the notation $\theta = \sigma_i$ and give an order in the set of nondegenerate simplices in $\Delta^1 \times \Delta^n$ of maximal dimension by declaring $\sigma_i < \sigma_j$ if $i < j$.

Definition 5.22 Let $\mathbb{K}_n = \Pi_{\Delta^n_\flat}((\Delta^n, \flat, \flat))$ and let $K_n = \mathbb{K}_n((0,0),(1,n))$. For every σ_i as in Remark 5.21, define K_n^i as the subposet (with the inherited decorations) of K_n consisting in those chains whose elements are in the image of σ_i. Observe that K_n^i is isomorphic as a simplicial set to $C_n = (\Delta^1)^n$.

Given $0 < s < n+1$ we define $d_s(K_n^j)$ as the simplicial set of K_n^j consisting in those chains whose elements are in the image of $d_s(\sigma_j)$.

Lemma 5.23 *Let $0 < j < n$ and let $d_{j+1}(K_n^j) \subset K_n^j$ be the simplicial subset (with the induced decorations) consisting in those chains whose elements factor through $d_{j+1}(\sigma_j)$. We view $\Delta^1 \times d_{j+1}(K_n^j)$ as marked-scaled simplicial set as follows:*

- *The marking consists in those marked edges in the cartesian product together with the edges contained in $\Delta^1 \times \{e\}$ where $e : C_0 \to C_1$ is a marked edge in K_n^j such that C_0 contains the element $(0, j)$.*
- *The scaling is given by the usual scaling on the cartesian product.*

Then we have an isomorphism of marked-scaled simplicial sets $\Delta^1 \times d_{j+1}(K_n^j) \simeq K_n^j$

Proof We only need to check that the decorations agree on each side. To show the claim regarding the marking we note that the marked edges of K_n^j always factor through a face of the cube C_n (see Definition 5.22) and thus the conclusion follows after direct inspection.

To see that the scaling of K_n^j is given by the product scaling we consider a triangle $\sigma : C_0 \to C_1 \to C_2$ and we define $D_i = C_i \setminus \{(1, j)\}$ which yields another triangle $\varphi : D_0 \to D_1 \to D_2$. We claim that σ is scaled in K_n^j if and only if φ is. Observe that since φ lies always in $d_{j+1}(K_n^j)$ this will be enough to show the claim regarding the scaling.

We set the notation $\pi_{0,n}(C_i) = S_i$ and $\pi_{0,n}(D_i) = T_i$. Then we have the following:

(1) We have that $S_i = T_i$ if $(0, j) \in C_i$ and $T_i = S_i \setminus \{j\}$ otherwise.

(2) We have $\min(S_i) = \min(T_i)$.

(3) Given $C \in d_{j+1}(K_n^j)$ and setting $V = \pi_{0,n}(C)$, it follows that $\min(V) \leq j$.

Our final claim is that for $i < j$ we have that (see Remark 5.4) $U_{S_i,S_j} = U_{T_i,T_j}$. Observe that $j \in U_{S_i,S_j}$ if and only if $j \in \{\min(S_i), \min(S_j)\}$ or if $j \in S_i$ and $\min(S_i) < j < \min(S_j)$. However by (3) above this cannot be the case. □

Lemma 5.24 *Let $(A, E_A, T_A) \subset (B, E_B, T_B)$ be an inclusion of marked-scaled simplicial sets such that $E_A = E_B$. Suppose that there exists some vertex $v \in A$ with the following property:*

- *For every simplex $\sigma : \Delta^n \to B$ which does not factor through A, v is the final vertex of σ.*

Let M_A *(resp.* M_B*) be the collection of marked edges in the cartesian product (of marked-scaled simplicial sets)* $\Delta^1 \times A$ *(resp.* $\Delta^1 \times B$*) together with the edge* $\Delta^1 \times \{v\}$*. Then the induced morphism*

$$j : \Delta^1 \times A \amalg_{\Delta^{\{1\}} \times A} B \to \Delta^1 \times B,$$

where both simplicial sets are equipped with the product scaling and the marking given by M_A *and* M_B*, respectively, is a trivial cofibration in* $\mathrm{Set}_\Delta^{\mathbf{ms}}$*.*

Proof First let us assume that the claim holds for $E_A = E_B = \flat$. Then it follows that for a general marking the map j is obtained as a pushout of the map $j_\flat$ where the latter map is the inclusion associated to the minimal marking. This shows that the general result will follow.

Working simplex by simplex we can reduce the problem to the cases

(i) $(\partial\Delta^n, \flat, \flat) \to (\Delta^n, \flat, \flat)$ for $n \geqslant 1$,

(ii) $(\Delta^2, \flat, \flat) \to (\Delta^2, \flat, \sharp)$,

where v is given by the final vertex.

To check (i) we use an argument analogous to Lemma 3.5.12 in [19] which tells us that in this case the map j above is in the weakly saturated class of morphisms of type (M1) in Definition 3.36 and of type

$(*)$ $(\Lambda^n_n, \Delta^{\{n-1,n\}}, \Delta^{\{0,n-1,n\}}) \to (\Delta^n, \Delta^{\{n-1,n\}}, \Delta^{\{0,n-1,n\}})$

and thus the claim holds. To prove (ii) we note that we can scale the remaining simplices using pushouts along morphisms of type (M2) in Definition 3.36 together with the morphism

$(\diamond)$ $(\Delta^3, \Delta^{\{n-1,n\}}, U_3) \to (\Delta^3, \Delta^{\{n-1,n\}}, \sharp)$,

where U_3 is the collection of all triangles except $\Delta^{\{0,1,2\}}$. □

Lemma 5.25 *Let* $i : A \to B$ *be a morphism of type* (A1) *in Definition 3.4. Then the induced natural transformation* $\mathbb{S}\mathrm{t}_B(A) \Rightarrow \mathbb{S}\mathrm{t}_B(B)$ *is a pointwise weak equivalence.*

Proof Since $\mathbb{S}\mathrm{t}_B$ preserves colimits and cofibrations it will be enough to prove the claim in the specific case where $A = (\Lambda^n_i, \flat, \flat \subset \Delta^{\{i-1,i,i+1\}})$ and $B = (\Delta^n, \flat, \flat \subset \Delta^{\{i-1,i,i+1\}})$. We will show according to Remark 5.20 that the map $\Pi_B(A)^\uparrow \to \Pi_B(B)$ is an equivalence of $\mathrm{Set}_\Delta^{\mathbf{ms}}$-categories.

We observe that the induced morphism of marked-scaled simplicial sets $\Pi_B(A)^\uparrow(x, y) \to \Pi_B(B)(x, y)$ is an isomorphism except when $x = (0, 0)$ and $y = (1, n)$. Recall the definition of K_n in Definition 5.22 and equip this marked-scaled simplicial set with the decorations induced from $\Pi_B((0, 0), (1, n))$. We similarly define $\Lambda^n_i K = \Pi_B(A)^\uparrow((0, 0), (1, n))$. We define a filtration

$$\Lambda^n_i K = A_{-1} \to A_0 \to \cdots \to A_{n-1} \to A_n = K_n,$$

where A_s is the subsimplicial set of K_n containing every simplex which factors through K_n^j for $j \leqslant s$ (see Definition 5.22 for a definition of K_n^j). We will show that each of the steps in this filtration is a weak equivalence. The case $n = 2$ follows easily by a direct computation. For the rest of the proof we will assume that $n \geqslant 3$.

For $0 \leqslant j \leqslant n$ we consider the pullback-pushout diagram

$$\begin{tikzcd}
Q_n^j \arrow[r] \arrow[d] & K_n^j \arrow[d] \\
A_{j-1} \arrow[r] & A_j
\end{tikzcd}$$

We will show that the top horizontal morphism is a trivial cofibration. First we consider the case $0 \leqslant j < n$. We describe Q_n^j as a simplicial subset of K_n^j which contains every face of the cube C_n, except those that factor through $d_\alpha(K_n^j)$ for $\alpha \notin \Phi(i)$ where

$$\Phi(i) = \begin{cases} \{j+1, i+1\} & \text{if } j < i, \\ \{j+1\} & \text{if } j = i, \\ \{j+1, i\} & \text{if } j > i. \end{cases}$$

We produce a 2-step filtration $Q_n^j \to Z_n^j \to K_n^j$ where Z_n^j is obtained from Q_n^j by attaching the simplices in $d_\beta(K_n^j)$ where $\beta \neq j+1$ and $\beta \in \Phi(i)$.

We observe that for $n \geqslant 3$ we have that every marked edge in $d_{j+1}(K_n^j)$ factors through Z_n^j. Therefore we can use Lemma 5.24 with $B = d_{j+1}(K_n^j)$ and $A = (Z_n^j)_{|d_{j+1}(K_n^j)}$ to obtain a trivial cofibration of marked-scaled simplicial sets

$$\varphi : \Delta^1 \times (Z_n^j)_{|d_{j+1}(K_n^j)} \to \Delta^1 \times d_{j+1}(K_n^j).$$

As a consequence of Lemma 5.23 we see that the scaling of $\Delta^1 \times A$ and the scaling of Z_n^j coincide except possibly in those triangles coming from $\Delta^{\{i-1,i,i+1\}}$ and similarly for $\Delta^1 \times B$ and K_n^j. We further note that every marked edge of K_n^j factors through Z_n^j. After direct inspection we observe that we can produce a pushout diagram

$$\begin{tikzcd}
\Delta^1 \times (Z_n^j)_{|d_{j+1}(K_n^j)} \arrow[r] \arrow[d] & \Delta^1 \times d_{j+1}(K_n^j) \arrow[d] \\
Z_n^j \arrow[r] & K_n^j
\end{tikzcd}$$

Therefore we can add the remaining decorations via a pushout of φ along a cofibration which shows that the last step in the filtration is a trivial cofibration. To show that $Q_n^j \to Z_n^j$ is a trivial cofibration we consider a pushout-pullback diagram

$$\begin{tikzcd}
Z_{n-1}^u \arrow[r] \arrow[d] & d_\beta(K_n^j) = K_{n-1}^u \arrow[d] \\
Q_n^j \arrow[r] & Z_n^j
\end{tikzcd}$$

and conclude by the previous argument or by a direct computation if $n = 3$.

To finish the proof we will show that $Q_n^n \to K_n^n$ is a trivial cofibration. In this case Q_n^n contains every simplex that factors through a face of C_n except those factoring through $d_i(K_n^n)$. For every $k \in [n]$ we define a chain

$$D_k = (0,0) < (0,1) < \cdots < (0,k) < (1,n).$$

We observe that using morphisms of type (M2) in Definition 3.36 and morphisms of type ($\diamond$) as in the proof of Lemma 5.24 we can scale the triangle $D_{i-1} \subset D_i \subset D_{i+1}$ in A_{n-1} and in A_n. Recall that for every chain $C = \{(i_\alpha, j_\alpha)\}_{\alpha \in A}$ we defined m_C to be the biggest element in A such that $i_{m_C} = 0$. We can use this parameter to define a map

$$r_n : K_n^n \to (\Delta^n, \flat, \Delta^{\{i-1,i,i+1\}}), \quad C \mapsto j_{m_C}.$$

Moreover r_n admits a section s_n which sends j to D_j as defined above. It follows that there exists a marked homotopy between $s_n \circ r_n$ and the identity map on K_n^n. Furthermore, r_n restricts to a map $\hat{r}_n : Q_n^n \to \Lambda_i^n$. One checks that since Q_n^n and Λ_i^n can be expressed as iterated pushouts along cofibrations indexed by $(n-1)$-dimensional faces of Q_n^n and the map r_n restricts to an equivalence in each of its faces, that $\hat{r}_n$ is also a weak equivalence. We conclude that we have a commutative diagram

$$\begin{tikzcd}
Q_n^n \arrow[r] \arrow[d, "\simeq"] & K_n^n \arrow[d, "\simeq"] \\
(\Lambda_i^n, \flat, \Delta^{\{i-1,i,i+1\}}) \arrow[r, "\simeq"] & (\Delta^n, \flat, \Delta^{\{i-1,i,i+1\}})
\end{tikzcd}$$

and so the result follows by two-out-of-three. □

Lemma 5.26 *Let* $i : A \to B$ *be a morphism of type* (A3) *in Definition 3.4. Then the induced natural transformation* $\mathbb{S}\mathrm{t}_B(A) \Rightarrow \mathbb{S}\mathrm{t}_B(B)$ *is a pointwise weak equivalence.*

Proof The proof will mirror the strategy of the previous lemma. Again, we observe that the induced morphism of mapping simplicial sets $\Pi_B(A)^{\uparrow}(x,y) \to \Pi_B(B)(x,y)$ is an isomorphism except when $x = (0,0)$ and $y = (1,n)$ or when $x = (0,1)$ and $y = (1,n)$. However we observe that since the edge $(0,0) \to (0,1)$ will be collapsed in order to define the value of the functor $\mathbb{S}\mathrm{t}_B$ it will suffice to construct the analogous filtration

$$\Lambda_0^n K = A_{-1} \to A_0 \to \cdots \to A_{n-1} \to A_n = K_n$$

and show that each step is a trivial cofibration. As before, we will leave the case $n = 2$ as an easy exercise and focus our attention to the cases $n \geqslant 3$. Note that in this case, we decorations coming from the marked edge $0 \to 1$ and the thin triangle $0 \to 1 \to n$ are already contained in $\Lambda_0^n K$.

For $0 \leqslant j < n$ we consider pullback-pushout diagrams

$$\begin{tikzcd}
Q_n^j \arrow[r] \arrow[d] & K_n^j \arrow[d] \\
A_{j-1} \arrow[r] & A_j
\end{tikzcd}$$

where Q_n^j is the simplicial subset of K_n^j which contains every face of C_n except the $d_{j+1}(K_n^j)$ and if $j > 0$ the face C_n^0 consisting in those chains that have the element $(0,1)$. The proof at this point is totally analogous to the proof of Lemma 5.25. We construct Z_n^j by adding to Q_n^j the face C_n^0 and conclude by Lemma 5.24 that each step in the filtration $Q_n^j \to Z_n^j \to K_n^j$ is given by a trivial cofibration.

To show that $A_{n-1} \to A_n$ is a trivial cofibration we need to work a little bit harder. First we consider a commutative diagram where we are using circled arrows to represent marked morphisms

$$\begin{tikzcd}
(0,0) < (1,0) < (1,n) \arrow[r, "\circ" marking] & (0,0) < (1,0) < (1,1) < (1,n) \\
(0,0) < (1,n) \arrow[r] \arrow[u, "\circ" marking] \arrow[d] & (0,0) < (1,1) < (1,n) \arrow[u, "\circ" marking] \arrow[d, "\circ" marking] \\
(0,0) < (0,1) < (1,n) \arrow[r, "\circ" marking] & (0,0) < (0,1) < (1,1) < (1,n)
\end{tikzcd}$$

We note that every 2-simplex in this diagram is scaled and therefore we can mark every morphism via a pushout along a trivial cofibration. Let us remark that we can mark this morphisms in both A_{n-1} and A_n since $n \geqslant 3$. Recall the definition of D_i in Lemma 5.25 and consider a 3-simplex

$$\rho_W : D_0 \subset D_1 \subset (0,0) < (0,1) < (0,n) < (1,n) \subset (0,0) < (0,1) < (0,n) < (1,n) \cup W,$$

where W is any chain starting at $(0,1)$ and ending at $(0,n)$. It follows that every face of ρ_W is scaled except possibly $d_2(\rho_W)$. Therefore we might scale that face using a pushout along a morphism of type (M2) in Definition 3.36. Note that any possible ρ_W factors through A_{n-1} except in the case where W is the maximal chain. A similar argument as in the previous lemma shows that we have a commutative diagram

$$\begin{tikzcd}
A_{n-1} \arrow[r] \arrow[d, "\simeq"] & A_n \arrow[d, "\simeq"] \\
(\Lambda_0^n, \{\Delta^{\{0,1\}}\}, \{\Delta^{\{0,1,n\}}\}) \arrow[r, "\simeq"] & (\Delta^n, \{\Delta^{\{0,1\}}\}, \{\Delta^{\{0,1,n\}}\})
\end{tikzcd}$$

and the claim follows from two-out-of-three. □

Lemma 5.27 *Let $i : A \to B$ be a morphism of type* (M2) *in* (M2). *Then the induced natural transformation $\mathbb{S}\mathrm{t}_B(A) \Rightarrow \mathbb{S}\mathrm{t}_B(B)$ is a pointwise weak equivalence.*

Proof Let $L = \Pi_B(A)^{\uparrow}((0,0),(1,4))$ and let $K_4 = \Pi_B(B)((0,0),(1,4))$. The only thing we need to show is that the map $L \to K_4$ is a weak equivalence. Using the notation from the previous proofs it follows that the missing scaled simplices in L are all contained in K_4^4. Therefore if we denote by L_4^4 the restriction of K_4^4 to L we see that it will be enough to show that the induced map $L_4^4 \to K_4^4$ is a weak equivalence. However, one can easily construct a commutative diagram

$$\begin{tikzcd}
L_4^4 \arrow[r] \arrow[d, "\simeq"] & K_4^4 \arrow[d, "\simeq"] \\
(\Delta^4, \flat, T) \arrow[r, "\simeq"] & (\Delta^4, \flat, T')
\end{tikzcd}$$

where vertical maps are weak equivalences and the scaling in the bottom horizontal map is that of (A2) in Definition 3.4. □

Proposition 5.28 *Let* (S,T_S) *be a scaled simplicial set and let* $i:A\to B$ *be an* **MB***-anodyne morphism in* $\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(S,\sharp,T_S\subset\sharp)}$*. Then* $\mathbb{S}\mathrm{t}_S(i)$ *is a weak equivalence in* $\mathrm{Fun}(\mathfrak{C}^{\mathbf{sc}}[S],\mathrm{Set}_\Delta^{\mathbf{ms}})$*.*

Proof Using Proposition 5.14 we can assume that $S=B$. The rest of the proof will consist in verifying that the claim holds for each of the generators given in Definition 3.4. We proceed case by case:

(A1) This follows from Lemma 5.25.

(A2) This follows from Lemma 5.27.

(A3) This follows from Lemma 5.26.

(A4) This follows from Corollary 5.19.

(S1) This follows from Corollary 5.19.

(S2) This follows by explicit verification.

(E) This follows from Corollary 5.19. □

Proposition 5.29 *Let* (S,T_S) *be a scaled simplicial set and let* $\phi:\mathfrak{C}^{\mathbf{sc}}[S]\to\mathcal{C}$ *be a functor of* Set_Δ^+*-enriched categories. Then the straightening functor*

$$\mathbb{S}\mathrm{t}_\phi:\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/(S,\sharp,T_S\subset\sharp)}\to\mathrm{Fun}(\mathcal{C},\mathrm{Set}_\Delta^{\mathbf{ms}})$$

is a left Quillen functor.

Proof We will show that $\mathbb{S}\mathrm{t}_\phi$ preserves cofibrations and weak equivalences. First, we point out that due to Proposition 5.14 it will be enough to consider the case $\phi=\mathrm{id}$. In this case, we saw in Lemma 5.16 that our functor preserves cofibrations.

To address the claim regarding weak equivalences, we see that Proposition 5.28 implies $\mathbb{S}\mathrm{t}_S$ preserves **MB**-anodyne morphisms. We can therefore restrict our attention to showing that $\mathbb{S}\mathrm{t}_S$ preserves weak equivalences between fibrant objects. To this end it will be enough to show the following:

($*$) Let $f,g:X\to Y$ be morphisms between fibrant objects such that $\mathbb{S}\mathrm{t}_S(f)$ is a weak equivalence. Then given a homotopy $H:X\times(\Delta^1)^\sharp\to Y$ between f and g it follows that $\mathbb{S}\mathrm{t}_S(g)$ is also a weak equivalence.

The claim follows after noting that we have an anodyne morphism $X\times\Delta^{\{0\}}\to X\times(\Delta^1)^\sharp$ due to Proposition 3.12 which implies that $\mathbb{S}\mathrm{t}_S(H)$ is a weak equivalence as well as the map induced by the projection onto X, $\mathbb{S}\mathrm{t}_S(p):\mathbb{S}\mathrm{t}_S(X\times(\Delta^1)^\sharp)\to\mathbb{S}\mathrm{t}_S(X)$. □

5.2 Straightening over a point

The goal of this section is to prove the following result.

Proposition 5.30 *The straightening-unstraightening adjunction over the point*

$$\mathbb{St}_* : \big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)_{/\Delta^0} \rightleftarrows \mathrm{Set}_\Delta^{\mathbf{ms}} : \mathbb{U}\mathrm{n}_*$$

is a Quillen equivalence.

To do this we will construct a left Quillen equivalence

$$L_* : \big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)_{/\Delta^0} \to \mathrm{Set}_\Delta^{\mathbf{ms}}$$

and a natural transformation $\alpha : \mathbb{St}_* \Rightarrow L$ which is pointwise a weak equivalence of marked-scaled simplicial sets.

Proposition 5.31 *Let* $L_* : \big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)_{/\Delta^0} \to \mathrm{Set}_\Delta^{\mathbf{ms}}$ *be the functor that assigns to an* **MB** *simplicial set* $(X, E_X, T_X \subseteq C_X)$ *the marked scaled simplicial set* (X, E_X, C_X)*. Then* L_* *is a left Quillen equivalence.*

Proof The functor L_* admits a right adjoint R_* which is given by $R_*(X, E_X, T_X) = (X, E_X, T_X)$. We observe that $L_* \circ R_* = \mathrm{id}$ and that the unit $\mathrm{id} \Rightarrow R_* \circ L_*$ is given by $(X, E_X, T_X \subset C_X) \to (X, E_X, C_X)$ which is in the weakly saturated class of morphisms of type (S2) in Definition 3.4.

To finish the proof we observe that L_* preserves cofibrations and maps **MB**-anodyne morphisms to **MS**-anodyne morphisms. It is then easy to see that L_* preserves weak equivalences between fibrant objects and the result follows. □

Recall the definition of the maps $\pi_{0,n}$ in Remark 5.6. Then postcomposing this map with the morphism $m_n : \mathcal{P}_n \to (\Delta^n, \flat)$ which assigns to every $S \in \mathcal{P}_n$ the value $m_n(S) = \min(S)$ we obtain a map of marked scaled simplicial sets

$$\alpha_n^\flat : \mathbb{St}_*((\Delta^n, \flat, \flat)) \to (\Delta^n, \flat, \flat).$$

One can easily produce marked variants of this maps $\alpha_2^\sharp$ and $\alpha_1^\sharp$ associated to the **MB** simplicial sets $(\Delta^2, \flat, \sharp)$ and $(\Delta^1, \sharp, \sharp)$. We would like to remark that $\mathbb{St}_*(\Delta^2, \flat, \flat \subset \sharp) = \mathbb{St}_*(\Delta^2, \flat, \sharp)$ which justifies why we are defining only one $\alpha_2^\sharp$. Since our definitions are functorial with respect to monotone morphisms $[k] \to [n]$ this collection of maps assemble into a natural transformation $\alpha : \mathbb{St}_* \Rightarrow L_*$.

Proposition 5.32 *The natural transformation* $\alpha : \mathbb{St}_* \Rightarrow L_*$ *is a pointwise weak equivalence.*

Proof Since both functors are left Quillen and both model categories are left proper, a simplex by simplex argument shows that it will be enough to show that the components $\alpha_n^\flat$, $\alpha_2^\sharp$ and $\alpha_1^\sharp$ are weak equivalences. We further observe $\alpha_2^\sharp$ and $\alpha_1^\sharp$ can be obtained from their undecorated countermarks via pushouts along cofibrations. This shows that we can restrict our attention to $\alpha_n^\flat$ for $n \geqslant 0$.

Let $S = (\Delta^n, \flat)$ and consider $\Pi_S(\Delta^n, \flat, \flat)$ (see Definition 5.8). Denote by Φ_n the $\mathrm{Set}_\Delta^{\mathbf{ms}}$-category obtained from $\Pi_S(\Delta^n, \flat, \flat)$ by marking every edge in the mapping simplicial sets of the form

$$\Pi_S(\Delta^n, \flat, \flat)((1, a), (1, b)).$$

We will show that the map $\widehat{\alpha}_n^\flat : K_n = \Phi_n((0, n), (1, n)) \to (\Delta^n, \flat)$ is a weak equivalence. Note that $\alpha_n^\flat$ is obtained from $\widehat{\alpha}_n^\flat$ after identifying certain simplices. It is easy to see that we can mark every edge in K_n

whose image under $\widehat{\alpha}^{\flat}_n$ becomes degenerate using pushouts along **MB**-anodyne morphisms. We consider a filtration

$$A_{-1} = K^0_n \to A_1 \to A_2 \to \cdots \to A_{n-1} \to A_n = K_n,$$

where A_i is obtained from A_{i-1} by attaching those simplices contained in K^i_n (see Definition 5.22) where K^i_n has the decorations induced from Φ_n. We further denote by $\overline{A}_i$ the image of A_i under the collapse map in the definition of $\mathbb{S}\mathrm{t}_*((\Delta^n, \flat, \flat))(*)$ and similarly denote $\overline{K}^i_n$.

We will show that the restriction of $\alpha^{\flat}_n$ to each $\overline{A}_i$ defines a weak equivalence

$$\alpha^{\flat}_{n,i} : \overline{A}_i \to (\Delta^{[0,i]}, \flat, \flat).$$

Since weak equivalences are stable under filtered colimits this will imply the result. Assume that $\alpha^{\flat}_{n,j}$ is a weak equivalence for $j \leqslant i-1$ and consider the pullback-pushout square

$$\begin{tikzcd} Q^i_n \arrow[r] \arrow[d] & K^i_n \arrow[d] \\ A_{i-1} \arrow[r] & A_i \end{tikzcd}$$

Observe that $\widehat{\alpha}^{\flat}_n$ induces a commutative diagram

$$\begin{tikzcd} Q^i_n \arrow[r] \arrow[d] & K^i_n \arrow[d, "r_i"] \\ \Delta^{[0,i-1]} \arrow[r] & \Delta^{[0,i]} \end{tikzcd}$$

We claim that the vertical morphisms are weak equivalences. Note that we can define, for $0 \leqslant j \leqslant i$,

$$C_j = (0,0) < (0,1) \cdots < (0,j) < (1,i) < \cdots < (1,n),$$

which provides us with a section $s_i : \Delta^{[0,i]} \to K^i_n$ sending j to C_j. One checks that $r_i \circ s_i = \mathrm{id}$ and that there is a marked homotopy between the identity of K^i_n and $s_i \circ r_i$. Moreover, the section s_i and the homotopy restrict to Q^i_n. It is immediate to see that both the section an the homotopy can be factored through the quotient simplicial sets $\overline{K}^i_n$ and $\overline{Q}^i_n$ which shows that $\alpha^{\flat}_{n,i}$ is again a weak equivalence. □

Corollary 5.33 *Let S be a scaled simplicial set. Let $\phi : \mathfrak{C}^{\mathrm{sc}}[S] \to \mathcal{C}$ be a Set^+_Δ-enriched functor. Assume that ϕ is essentially surjective, and let $\alpha : \mathcal{F} \to \mathcal{F}'$ be a map between fibrant objects of $\mathrm{Fun}(\mathcal{C}, \mathrm{Set}^{\mathbf{ms}}_\Delta)$. Then the following conditions are equivalent:*

(1) *The map α is a weak equivalence in $(\mathrm{Set}^{\mathbf{ms}}_\Delta)^{\mathcal{C}}$.*

(2) *For every $C \in \mathcal{C}$ the induced map $\alpha_C : \mathcal{F}(C) \to \mathcal{F}'(C)$ is a weak equivalence.*

(3) *For every vertex $s \in S$ the induced map on fibres*

$$\mathbb{U}\mathrm{n}_\phi(\mathcal{F})_s \to \mathbb{U}\mathrm{n}_\phi(\mathcal{F}')_s$$

is a bicategorical equivalence.

(4) *The map $\mathbb{U}\mathrm{n}_\phi(\alpha)$ is a weak equivalence in $(\mathrm{Set}^{\mathbf{mb}}_\Delta)_{/(S,\sharp,T_S \subset \sharp)}$.*

Proof The equivalence (1) $\iff$ (2) is immediate from the definition. Since both $\mathcal{F}$ and $\mathcal{F}'$ are fibrant it follows from Proposition 5.29 that $\mathbb{U}\mathrm{n}_\phi(\alpha)$ is a map between fibrant objects in $(\mathrm{Set}_\Delta^{\mathbf{mb}})_{/(S,\sharp,T_S\subset\sharp)}$. Then the equivalence (3) $\iff$ (4) follows from (v) in Proposition 3.23. To finish the proof we need to show that (2) $\iff$ (3). However, the fact that ϕ is surjective allows us to reduce to the case where $S=\Delta^0$ and conclude by Remark 5.15 and Proposition 5.30. □

5.3 Straightening over a simplex

In this section we establish the key element in the proof of our main theorem.

Proposition 5.34 *Denote by $\Delta^n_\flat=(\Delta^n,\flat)$ the minimally scaled n-simplex. Then the straightening–unstraightening adjunction, over $\Delta^n_\flat$,*

$$\mathbb{S}\mathrm{t}_{\Delta^n_\flat} : (\mathrm{Set}_\Delta^{\mathbf{mb}})_{(\Delta^n,\sharp,\flat\subset\sharp)} \rightleftarrows \mathrm{Fun}(\mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat],\mathrm{Set}_\Delta^{\mathbf{ms}}) : \mathbb{U}\mathrm{n}_{\Delta^n_\flat}$$

is a Quillen equivalence.

Let us comment on the general structure of the proof before we dive into the details. First, let us observe that Corollary 5.33 shows that $\mathbb{U}\mathrm{n}_{\Delta^n_\flat}$ detects weak equivalences between fibrant objects and thus it descends to a conservative right adjoint on homotopy categories. Consequently, in order to prove the result we only need to show that given an object $X\in(\mathrm{Set}_\Delta^{\mathbf{mb}})_{(\Delta^n,\sharp,\flat\subset\sharp)}$ and an equivalence of Set_Δ^+-enriched functors

$$\mathbb{S}\mathrm{t}_{\Delta^n}(X)\xrightarrow{\simeq}\mathcal{F},$$

where $\mathcal{F}$ is a fibrant functor, the adjoint map $X\to\mathbb{U}\mathrm{n}_{\Delta^n}(\mathcal{F})$ is a weak equivalence in $(\mathrm{Set}_\Delta^{\mathbf{mb}})_{(\Delta^n,\sharp,\flat\subset\sharp)}$. Using two-out-of-three we can assume without loss of generality that $p:X\to\Delta^n$ is a fibrant object. Now, in order to check that our adjoint map is a weak equivalence it suffices by Proposition 3.23 to check that the induced morphisms on fibres

$$\varphi_i^n : X_i\xrightarrow{\simeq}\mathbb{U}\mathrm{n}_{\Delta^n}(\mathcal{F})_i$$

are bicategorical equivalences for $0\leqslant i\leqslant n$. We will use induction on n, the dimension of our simplex $\Delta^n_\flat$. Let us assume that we have proved Proposition 5.34 for $0\leqslant k\leqslant n-1$ and note the base case was already shown in Proposition 5.30. We claim that for every $0\leqslant i\leqslant n-1$ the map φ_i^n is a bicategorical equivalence.

We consider the morphism $\alpha:\Delta^{[0,n-1]}_\flat\to\Delta^n_\flat$ and denote by $\overline{X}$ the pullback of X along α. Similarly we denote by $\overline{\mathcal{F}}=j^*\mathcal{F}$ the restriction of $\mathcal{F}$ along $\mathfrak{C}^{\mathbf{sc}}[\alpha]=j$. We observe the following:

(1) As a direct consequence of the point-wise formula for left (enriched) Kan extensions in terms of weighted colimits (see [18, Proposition A.3.3.7]) and fully faithfulness of j we see that for every functor $\mathcal{H}:\mathfrak{C}^{\mathbf{sc}}[\Delta^{n-1}_\flat]\to\mathrm{Set}_\Delta^{\mathbf{ms}}$ we have an isomorphism $j^*j_!\mathcal{H}\simeq\mathcal{H}$. This implies that for every $\mathcal{G}:\mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat]\to\mathrm{Set}_\Delta^{\mathbf{ms}}$ we have that $j_!j^*\mathcal{G}(i)\to\mathcal{G}(i)$ is an isomorphism for $0\leqslant i<n$.

(2) For every $p:X\to\Delta^n$ we have that the map $\mathbb{S}\mathrm{t}_{\Delta^n_\flat}(\overline{X})(i)\to\mathbb{S}\mathrm{t}_{\Delta^n_\flat}(X)(i)$ is an isomorphism for $0\leqslant i<n$. This follows from the previous point and Proposition 5.14 since $\mathbb{S}\mathrm{t}_{\Delta^n_\flat}\overline{X}\simeq j_!\,\mathbb{S}\mathrm{t}_{\Delta^{n-1}_\flat}\overline{X}$.

(3) Using Proposition 5.14 and Remark 5.15 we obtain a commutative diagram

$$\begin{tikzcd} j_! \, \mathbb{S}\mathrm{t}_{\Delta^{n-1}_\flat} \overline{X} \arrow[r] \arrow[d] & j_! j^*(\mathcal{F}) \arrow[d] \\ \mathbb{S}\mathrm{t}_{\Delta^{n}_\flat} X \arrow[r] & \mathcal{F} \end{tikzcd}$$

We conclude from (1)–(3) above that $\mathbb{S}\mathrm{t}_{\Delta^{n-1}_\flat} \overline{X} \to \overline{\mathcal{F}}$ is a weak equivalence. It follows from our induction hypothesis that $\overline{X} \to \mathbb{U}\mathrm{n}_{\Delta^{n-1}_\flat}(\overline{\mathcal{F}})$ is a weak equivalence and thus the maps φ_n^i are equivalences for $0 \leqslant i < n$.

We have reduced our problem to showing that φ_n^n is a weak equivalence. We claim that is enough to show the following.

($*$) The map $\mathbb{S}\mathrm{t}_{\Delta^n}(X_n)(n) \to \mathbb{S}\mathrm{t}_{\Delta^n}(X)(n)$ is a weak equivalence.

Indeed, let $r^n : \Delta^0 \to \Delta^n_\flat$ denote the inclusion of the terminal vertex. Given $K \in \mathrm{Set}_\Delta^{\mathbf{ms}}$ it follows that we have an isomorphism $r^n_!(K)(n) \simeq K$ which in turn shows that the adjoint morphism to

$$r^n_!(\mathbb{S}\mathrm{t}_* X_n) \simeq \mathbb{S}\mathrm{t}_{\Delta^n}(X_n) \to \mathbb{S}\mathrm{t}_{\Delta^n}(X) \to \mathcal{F},$$

which is given by $\mathbb{S}\mathrm{t}_* X_n \to \mathcal{F}(n)$ is a weak equivalence. We can now use Proposition 5.30 to conclude that we have a bicategorical equivalence

$$X_n \xrightarrow{\simeq} \mathbb{U}\mathrm{n}_{\Delta^n}(\mathcal{F})_n.$$

Therefore, we will devote the rest of this section to the proof of the claim ($*$) above.

Definition 5.35 Let I be a finite linearly ordered set and let $i \in I$. We define simplicial set as the nerve of a poset $\mathcal{O}^I_{i\nearrow}$ whose elements are given by subsets $S \subseteq I$ such that $S \neq \varnothing$ and such that $\min(S) = i$. We declare $S \leqslant T$ if $S \subseteq T$. We observe that we have a map

$$\pi_I : \mathcal{O}^I_{i\nearrow} \to \Delta^I, \quad S \mapsto \max(S).$$

We upgrade $\pi_I : \mathcal{O}^I_{i\nearrow} \to \Delta^I$ to an object of $\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{/\Delta^I_\flat}$ as follows:

- We declare an edge $S \to T$ to be marked if $T = S \cup \max(T)$.
- We declare every triangle of $\mathcal{O}^I_{i\nearrow}$ to be lean.
- We declare a triangle to be thin if its image in Δ^I is degenerate.

Definition 5.36 To ease the notation we set $\mathbb{O}^n = \mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat]$ (see Definition 2.8).

Remark 5.37 Let $I = [n]$. For every $i \leqslant j$ we view $\mathbb{O}^n(i, j)$ as an **MB** simplicial set by declaring every triangle to be thin scaled and only degenerate edges to be marked. We further note that we have functors

$$\mathbb{O}^n(i, j) \times \mathcal{O}^n_{j\nearrow} \to \mathcal{O}^n_{i\nearrow}, \quad (S, T) \mapsto S \cup T,$$

which preserve the decorations.

Definition 5.38 Let $\mathcal{F} : \mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat] \to \mathrm{Set}^{\mathbf{ms}}_\Delta$ be a Set^+_Δ-enriched functor. We define an **MB** simplicial set $\mathcal{M}(\mathcal{F})$ over $\Delta^n_\flat$ as follows:

(1) Let $i \in [n]$. We upgrade the marked-scaled simplicial set $\mathcal{F}(i)$ to an **MB** simplicial set by declaring the collection of thin triangles and lean triangles to coincide.

(2) We define $\mathcal{M}(\mathcal{F})$ as the coequaliser of the diagram in the category of **MB** simplicial sets

$$\coprod_{i<j} \mathcal{F}(i) \times \mathbb{O}^n(i,j) \times \mathcal{O}^n_{j\nearrow} \rightrightarrows \coprod_{i} \mathcal{F}(i) \times \mathcal{O}^n_{i\nearrow}, \quad i,j \in [n],$$

where the maps in the diagram are given by

$$\mathcal{F}(i) \times \mathbb{O}^n(i,j) \times \mathcal{O}^n_{j\nearrow} \to \mathcal{F}(j) \times \mathcal{O}^n_{j\nearrow}, \quad \mathcal{F}(i) \times \mathbb{O}^n(i,j) \times \mathcal{O}^n_{j\nearrow} \to \mathcal{F}(i) \times \mathcal{O}^n_{i\nearrow}.$$

(3) The maps $\mathcal{F}(i) \times \mathcal{O}^n_{i\nearrow} \to \mathcal{O}^n_{i\nearrow} \to \Delta^n$ assemble into a functor $\mathcal{M}(\mathcal{F}) \to \Delta^n$.

Remark 5.39 Given $0 \leqslant i < n$ we have a map (actually an isomorphism) of simplicial sets

$$\Xi^i : \Delta^1 \times \mathcal{O}^{n-1}_{i\nearrow} \to \mathcal{O}^n_{i\nearrow}, \qquad \Xi^i(\varepsilon, S) = \begin{cases} S & \text{if } \varepsilon = 0, \\ S \cup \{n\} & \text{if } \varepsilon = 1, \end{cases}$$

which we use to equip $\Delta^1 \times \mathcal{O}^{n-1}_{i\nearrow}$ with the induced decorations from $\mathcal{O}^n_{i\nearrow}$.

Given a functor $\mathcal{F} : \mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat] \to \mathrm{Set}^{\mathbf{ms}}_\Delta$ let us denote by $\overline{\mathcal{F}}$ the restriction of $\mathcal{F}$ to $\mathfrak{C}^{\mathbf{sc}}[\Delta^{n-1}_\flat]$ along the map $i : \Delta^{[0,n-1]}_\flat \to \Delta^n_\flat$. We can use the maps Ξ^i to construct a map of simplicial sets

$$\Xi_{\mathcal{F}} : \Delta^1 \times \mathcal{M}(\overline{\mathcal{F}}) \to \mathcal{M}(\mathcal{F}).$$

Finally, we equip $\Delta^1 \times \mathcal{M}(\overline{\mathcal{F}})$ with the decorations induced (via $\Xi_{\mathcal{F}}$) from $\mathcal{M}(\mathcal{F})$ and denote the resulting **MB** simplicial set by $\Delta^1 \mathbin{\overline{\otimes}} \mathcal{M}(\overline{\mathcal{F}})$.

Remark 5.40 The construction of $\mathcal{M}(\mathcal{F})$ defines a colimit-preserving functor

$$\mathcal{M}(-) : \mathrm{Fun}(\mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat], \mathrm{Set}^{\mathbf{ms}}_\Delta) \to \left(\mathrm{Set}^{\mathbf{mb}}_\Delta\right)_{/\Delta^n_\flat},$$

which enjoys several properties:

(i) For every $j \in [n]$ we have an isomorphism $\mathcal{M}(\mathcal{F})_j \simeq \mathcal{F}(j)$.

(ii) Given a functor $\mathcal{F} : \mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat] \to \mathrm{Set}^{\mathbf{ms}}_\Delta$ let us denote by $\overline{\mathcal{F}}$ as in Remark 5.39. Then we have a pushout diagram

$$\begin{tikzcd}
\Delta^{\{1\}} \times \mathcal{M}(\overline{\mathcal{F}}) \arrow[r] \arrow[d] & \mathcal{F}(n) \arrow[d] \\
\Delta^1 \mathbin{\overline{\otimes}} \mathcal{M}(\overline{\mathcal{F}}) \arrow[r] & \mathcal{M}(\mathcal{F})
\end{tikzcd}$$

of **MB** simplicial sets.

(iii) The functor $\mathcal{M}(-)$ preserves cofibrations.

Lemma 5.41 *Let* $A \to \Delta^n_\flat$ *be an object of* $\left(\mathrm{Set}^{\mathbf{mb}}_\Delta\right)_{(\Delta^n,\sharp,\flat\subset\sharp)}$. *Define an* ***MB*** *simplicial set* $\Delta^1 \overline{\otimes} A \to \Delta^{n+1}_\flat$ *as follows*:

- *The underlying simplicial set is given by the cartesian product.*
- *The projection map* $\Delta^1 \times A \to \Delta^{n+1}_\flat$ *is induced by the map*

$$r : \Delta^1 \times \Delta^n \to \Delta^{n+1}, \qquad r(\varepsilon, i) = \begin{cases} i & \text{if } i = 0, \\ n & \text{if } i = 1. \end{cases}$$

- *Let* $(e_1, e_A) : \Delta^1 \to \Delta^1 \times A$ *be marked if* e_A *is marked in* A *and* $e_1 = 0 \to 0$ ***or*** *if* e_A *is degenerate.*
- *A triangle is lean if and only if it is lean in* A.
- *A triangle is thin if and only if it is lean and its image in* Δ^{n+1} *is degenerate.*

Then the map $\Delta^{\{0\}} \times A \to \Delta^1 \overline{\otimes} A$ *is* ***MB****-anodyne.*

Proof The claim follows from a standard simplex by simplex argument and is left as an exercise to the reader. □

Remark 5.42 The scaling of $\Delta^1 \overline{\otimes} \mathcal{M}(\overline{\mathcal{F}})$ given in Remark 5.39 is precisely that of Lemma 5.41 so in particular we obtain an anodyne morphism $\mathcal{M}(\overline{\mathcal{F}}) \to \Delta^1 \overline{\otimes} \mathcal{M}(\overline{\mathcal{F}})$. Moreover, in the particular case where $\mathcal{F}$ is the corepresentable functor on the object 0 it follows $\mathcal{M}(\mathcal{F}) = \mathcal{O}^n_{0\nearrow}$ so applying Lemma 5.41 n times we obtain an anodyne morphism

$$\Delta^0 \xrightarrow{\simeq} \mathcal{O}^n_{0\nearrow},$$

where the map above selects the subset $\{0\}$.

Definition 5.43 Let $\mathcal{F} : \mathfrak{C}^{\mathrm{sc}}[\Delta^n_\flat] \to \mathrm{Set}^{\mathbf{ms}}_\Delta$ be a Set^+_Δ-enriched functor. We define a scaled simplicial set $\mathbb{M}(\mathcal{F})$ whose underlying simplicial set is given by that of $\mathcal{M}(\mathcal{F})$ and whose thin triangles are precisely the lean triangles of $\mathcal{M}(\mathcal{F})$.

Lemma 5.44 *Let* $\mathcal{F} : \mathfrak{C}^{\mathrm{sc}}[\Delta^n_\flat] \to \mathrm{Set}^{\mathbf{ms}}_\Delta$ *be a fibrant* Set^+_Δ*-enriched functor. Given a* $\flat$*-local* $(0,1)$*-fibration* $p : X \to \Delta^n_\flat$ *and a morphism, over* $\Delta^n_\flat$,

$$f : \mathcal{M}(\mathcal{F}) \to X$$

such that for every $j \in [n]$ *the map* f *induces bicategorical equivalences* $\mathcal{F}(j) \simeq X_j$, *it follows that* f *is a weak equivalence in* $\left(\mathrm{Set}^{\mathbf{mb}}_\Delta\right)_{(\Delta^n,\sharp,\flat\subset\sharp)}$.

Proof Since $p : X \to \Delta^n_\flat$ is a $\flat$-local $(0,1)$-fibration it follows that we can construct its associated bicategorical interpretation (see Definition 4.4) $\mathfrak{p} : \mathbb{X} \to \Delta^n$ by declaring lean triangles to be scaled. We claim that is enough to show the following:

($\diamond$) The associated map $\mathfrak{f} : \mathbb{M}(\mathcal{F}) \to \mathbb{X}$ is a bicategorical equivalence.

Indeed, given a $\flat$-local $(0,1)$-fibration $q : Z \to \Delta^n_\flat$ with associated bicategorical interpretation $\mathfrak{q} : \mathbb{Z} \to \Delta^n$ it follows from our claim that we have a bicategorical equivalence

$$\phi : \operatorname{Fun}^{\mathbf{sc}}(\mathbb{X}, \mathbb{Z}) \xrightarrow{\simeq} \operatorname{Fun}^{\mathbf{sc}}(\mathbb{M}(\mathcal{F}), \mathbb{Z}).$$

Moreover, we obtain a commutative diagram

$$\begin{tikzcd}
\operatorname{Map}^{\mathbf{sc}}_{\Delta^n}(\mathbb{X}, \mathbb{Z}) \arrow[r, "\psi"] \arrow[d] & \operatorname{Map}^{\mathbf{sc}}_{\Delta^n}(\mathbb{M}(\mathcal{F}), \mathbb{Z}) \arrow[d] \\
\operatorname{Fun}^{\mathbf{sc}}(\mathbb{X}, \mathbb{Z}) \arrow[r, "\phi"] & \operatorname{Fun}^{\mathbf{sc}}(\mathbb{M}(\mathcal{F}), \mathbb{Z})
\end{tikzcd}$$

where $\operatorname{Map}^{\mathbf{sc}}_{\Delta^n}(-,-)$ denotes the full subcategory on maps which preserve the marked edges and commute with the projection maps. We observe that higher simplices in the aforementioned scaled simplicial set are also compatible with the projection maps. Since a simplex in $\mathcal{M}(\mathcal{F})$ is thin if and only if it is lean and its image is thin in $\Delta^n_\flat$ we see that if ψ is a bicategorical equivalence it will follow that f is a weak equivalence $\big(\mathrm{Set}^{\mathbf{mb}}_\Delta\big)_{(\Delta^n,\sharp,\flat\subset\sharp)}$. It is clear by construction that ψ is fully faithful so it will suffice to show that it is essentially surjective.

Given $u \in \operatorname{Map}^{\mathbf{sc}}_{\Delta^n}(\mathbb{M}(\mathcal{F}), \mathbb{Z})$ we can find some $v \in \operatorname{Fun}^{\mathbf{sc}}(\mathbb{X}, \mathbb{Z})$ such that $\phi(v) \simeq u$. Consequently, it will be enough to show that v factors through $\operatorname{Map}^{\mathbf{sc}}_{\Delta^n}(\mathbb{X}, \mathbb{Z})$. Let $x \in X$ such that $p(x) = i$ and pick an equivalence $f(y) \xrightarrow{\simeq} x$. We then see that

$$v(f(y)) \simeq u(y), \quad v(f(y)) \simeq v(x) \Longrightarrow q(v(x)) = i.$$

Therefore, we can focus our attention into proving the statement $(\diamond)$ above. Let $x_i \in \mathbb{M}(\mathcal{F})$ be an object represented by a pair $(a_i, \{i\})$ in $\mathcal{F}(i) \times \mathcal{O}^n_{i\nearrow}$. We consider a marked morphism $f : x_i \to \hat{x}_i$ given by $(a_i, \{i\}) \to (a_i, \{in\})$. Given an object x_n, lying over n, we claim the following:

($\star$) Restriction along f induces a weak equivalence of marked simplicial sets

$$\mathfrak{C}^{\mathbf{sc}}[\mathbb{M}(\mathcal{F})](\hat{x}_i, x_n) \xrightarrow{\simeq} \mathfrak{C}^{\mathbf{sc}}[\mathbb{M}(\mathcal{F})](x_i, x_n).$$

It is easy to see that $(\diamond)$ follows from $(\star)$ together with a routine inductive argument.

We observe that we have cofibrations of Set^+_Δ-enriched categories

$$\mathfrak{C}^{\mathbf{sc}}[\mathbb{M}(\overline{\mathcal{F}})] \to \mathfrak{C}^{\mathbf{sc}}[\Delta^1 \times \mathbb{M}(\overline{\mathcal{F}})], \quad \mathfrak{C}^{\mathbf{sc}}[\mathbb{M}(\overline{\mathcal{F}})] \to \mathfrak{C}^{\mathbf{sc}}[\Delta^1] \times \mathfrak{C}^{\mathbf{sc}}[\mathbb{M}(\overline{\mathcal{F}})]$$

and a diagram

$$\begin{tikzcd}
\mathfrak{C}^{\mathbf{sc}}[\mathbb{M}(\overline{\mathcal{F}})] \arrow[r] \arrow[d] & \mathfrak{C}^{\mathbf{sc}}[\Delta^1 \times \mathbb{M}(\overline{\mathcal{F}})] \arrow[r] \arrow[d] & \mathfrak{C}^{\mathbf{sc}}[\Delta^1] \times \mathfrak{C}^{\mathbf{sc}}[\mathbb{M}(\overline{\mathcal{F}})] \arrow[d] \\
\mathfrak{C}^{\mathbf{sc}}[\mathcal{F}(n)] \arrow[r] & \mathfrak{C}^{\mathbf{sc}}[\mathbb{M}(\mathcal{F})] \arrow[r] & \mathbb{P}(\mathcal{F})
\end{tikzcd}$$

where both squares are pushouts. It is not hard to see using the fact $\mathfrak{C}^{\mathbf{sc}}[-]$ is a left Quillen equivalence that the top right-most horizontal map is a weak equivalence. We conclude that the bottom right-most horizontal morphism is also a weak equivalence. It is easy to see by direct inspection that the analogous claim to $(\star)$ holds for $\mathbb{P}(\mathcal{F})$. □

Proposition 5.45 *Let* $p : X \to \Delta^n_\flat$ *be a* $\flat$*-local* $(0,1)$*-fibration. Then there exists a projectively fibrant-cofibrant functor* $\mathcal{F} : \mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat] \to \mathrm{Set}_\Delta^{\mathbf{ms}}$ *and a weak equivalence* $\mathcal{M}(\mathcal{F}) \to X$ *in* $\left(\mathrm{Set}_\Delta^{\mathbf{mb}}\right)_{(\Delta^n,\sharp,\flat\subset\sharp)}$.

Proof Using Lemma 5.44 it will be enough to construct a map $\mathcal{M}(\mathcal{F}) \to X$ inducing categorical equivalences on fibres. We proceed using induction on n, the case $n = 0$ being clear. Let us suppose that the claim holds for $n-1$ and let $i : \Delta^{[0,n-1]}_\flat \to \Delta^n_\flat$. We denote by $\overline{X}$ the restriction of X along i. Using our induction hypothesis we obtain a projectively fibrant-cofibrant functor $\overline{\mathcal{F}} : \mathfrak{C}^{\mathbf{sc}}[\Delta^{n-1}_\flat] \to \mathrm{Set}_\Delta^{\mathbf{ms}}$ and fibrewise equivalence $\mathcal{M}(\overline{\mathcal{F}}) \to \overline{X}$. We use Remark 5.42 to provide a solution to the lifting problem

$$\begin{tikzcd}
\mathcal{M}(\overline{\mathcal{F}}) \arrow[r] \arrow[d] & X \arrow[d, "p"] \\
\Delta^1 \mathbin{\overline{\otimes}} \mathcal{M}(\overline{\mathcal{F}}) \arrow[r] \arrow[ur, dotted] & \Delta^n
\end{tikzcd}$$

which provides us with a map $\mathcal{M}(\overline{\mathcal{F}}) \to X_n$. We factor this map as

$$\mathcal{M}(\overline{\mathcal{F}}) \xrightarrow{u} \widehat{X}_n \xrightarrow{v} X_n,$$

where u is a cofibration and v is a trivial fibration. Note that it follows that $\widehat{X}_n$ is also an ∞-bicategory. We can use the map u to extend $\overline{\mathcal{F}}$ to a fibrant-cofibrant functor $\mathcal{F} : \mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat] \to \mathrm{Set}_\Delta^{\mathbf{ms}}$ such that $\mathcal{F}(n) = \widehat{X}_n$. □

Proposition 5.46 *Let* $p : X \to \Delta^n_\flat$ *be a* $\flat$*-local* $(0,1)$*-fibration. Then there exists an equivalence of marked-scaled simplicial sets*

$$\mathbb{S}\mathrm{t}_{\Delta^n}(X_n)(n) \xrightarrow{\simeq} \mathbb{S}\mathrm{t}_{\Delta^n}(X)(n),$$

where X_n *denotes the fibre over* n *of* X.

Proof We note that due to Proposition 5.45 we have a fibrant-cofibrant functor $\mathcal{F} : \mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat] \to \mathrm{Set}_\Delta^{\mathbf{ms}}$ and a weak equivalence $\mathcal{M}(\mathcal{F}) \to X$. We will show that for every projectively cofibrant functor $\mathcal{G}$ the map

$$\eta_{\mathcal{G}} : \mathbb{S}\mathrm{t}_{\Delta^n}(\mathcal{G}(n))(n) \xrightarrow{\simeq} \mathbb{S}\mathrm{t}_{\Delta^n}(\mathcal{M}(\mathcal{G}))(n)$$

is a weak equivalence of marked-scaled simplicial sets. We observe that we have a pair of functors

$$L_i : \mathrm{Fun}(\mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat], \mathrm{Set}_\Delta^{\mathbf{ms}}) \to \mathrm{Set}_\Delta^{\mathbf{ms}}, \quad i = 1,2,$$

given by $L_1(\mathcal{G}) = \mathbb{S}\mathrm{t}_{\Delta^n}(\mathcal{G}(n))(n)$ and $L_2 = \mathbb{S}\mathrm{t}_{\Delta^n}(\mathcal{M}(\mathcal{G}))(n)$ which preserve colimits and cofibrations together with a natural transformation $\eta : L_1 \Rightarrow L_2$. We say that a functor $\mathcal{G}$ is *good* if $\eta_{\mathcal{G}}$ is a weak equivalence. To finish the proof we need to show that every cofibrant functor is good.

For every $-1 \leqslant j \leqslant n$ let $i_j : \Delta^{[0,j]}_\flat \to \Delta^n_\flat$ be the obvious inclusion with the convention $\Delta^{[0,-1]} = \varnothing$. Given a functor $\mathcal{G}$ we define $\mathcal{G}_j$ as the result of first restricting $\mathcal{G}$ along $\mathfrak{C}^{\mathbf{sc}}[i_j] = f^j$ and then applying the left Kan extension along f^j. We set $\mathcal{G}_{-1}$ to be the initial functor. We further denote by $r^j : \mathfrak{C}^{\mathbf{sc}}[\Delta^0] \to \mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat]$ the functor that picks the object j for $0 \leqslant j \leqslant n$.

Note that given a projectively cofibrant functor $\mathcal{G}$, it follows that the canonical map $\mathcal{G}_{j-1}(j) \to \mathcal{G}(j)$ is a cofibration for $0 \leqslant j \leqslant n$. We further note that we have a pushout diagram

$$\begin{tikzcd}
r^j_! \mathcal{G}_{j-1}(j) \arrow[r] \arrow[d] & r^j_! \mathcal{G}(j) \arrow[d] \\
\mathcal{G}_{j-1} \arrow[r] & \mathcal{G}_j
\end{tikzcd}$$

where $r^j_!$ denotes the left Kan extension functor along r^j. Since the top horizontal map is a cofibration it follows that in order to show that $\mathcal{G}_j$ is good it is enough to show that $r^j_! \mathcal{G}_{j-1}(j)$, $r^j_! \mathcal{G}(j)$ and $\mathcal{G}_{j-1}$ are good. We note the following:

- If $j > 0$ it follows that $r^j_! A(i) = \varnothing$ for $i < j$. In particular, it follows that $\mathcal{M}(r^j_! A)$ factors through $\Delta^{[j,n]}_\flat$. We can now induct on n to see that $r^j_! A$ is good for $j > 0$.

Finally, we can use induction on j to reduce our problem to show that for every $K \in \mathrm{Set}^{\mathbf{ms}}_\Delta$ the functor

$$\underline{K} : \mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat] \to \mathrm{Set}^{\mathbf{ms}}_\Delta, \quad i \mapsto K \times \mathbb{O}^n(0,i),$$

is good. Note that we can use further simplify our computation to the cases where $K = \Delta^k_\flat$ for $k \geqslant 0$, $\Delta^2_\sharp$ and $(\Delta^1)^\sharp$. We will only show the case $K = \Delta^k_\flat$ for $k \geqslant 0$ the other cases will follow by a totally analogous argument.

We can identify $\mathbb{S}\mathrm{t}_\Delta(\Delta^k \times \mathcal{O}^n_{0\nearrow})(n)$ as a quotient of the (decorated) poset of chains of $\Delta^1 \times \Delta^k \times \mathcal{O}^n_{0\nearrow}$ starting at $(0,0,\{0\})$ at ending at some element $(1,\ell,S)$ with $\max(S) = n$. We define a map

$$\psi : \mathbb{S}\mathrm{t}_\Delta(\Delta^k_\flat \times \mathcal{O}^n_{0\nearrow})(n) \to \Delta^k_\flat \times \mathbb{O}^n(0,n)$$

by sending a chain $C = \{(\varepsilon_i, k_i, S_i)\}_{i=0}^{i=\ell}$ to $(k_{i_{m_C}}, S_{m_C} \cup \{\max(S_j)\}_{j \geqslant m_C})$ where m_C is the biggest index such that $\varepsilon_i = 0$. We consider the map $p_0 : \Delta^k \to \Delta^n_\flat$ where Δ^k has the minimal decorations and where p_0 is constant on the vertex 0. We can now look at the commutative diagram

$$\begin{tikzcd}
\mathbb{S}\mathrm{t}_{\Delta^n}(\Delta^k)(n) \arrow[r, "\varphi"] \arrow[dr, "u"'] & \mathbb{S}\mathrm{t}_\Delta(\Delta^k_\flat \times \mathcal{O}^n_{0\nearrow})(n) \arrow[d, "\psi"] & \mathbb{S}\mathrm{t}_{\Delta^n}(\Delta^k \times \mathbb{O}^n(0,n))(n) \arrow[l, "\phi"'] \arrow[dl, "v"] \\
& \Delta^k_\flat \times \mathbb{O}^n(0,n) &
\end{tikzcd}$$

and make the following observations:

(1) It follows from Lemma 5.41, Remark 5.42 and Proposition 3.12 that φ is a weak equivalence.

(2) The map $u : \mathbb{S}\mathrm{t}_{\Delta^n}(\Delta^k)(n) \simeq \mathbb{S}\mathrm{t}_*(\Delta^k) \times \mathbb{O}^n(0,n) \to \Delta^k_\flat \times \mathbb{O}^n(0,n)$ can be identified with a product of the natural transformation α at Δ^k considered in the proof of Proposition 5.30 and the identity map on $\mathbb{O}^n(0,n)$ and its a consequently a weak equivalence. It follows from 1 that ψ is also a weak equivalence.

(3) The map $v : \mathbb{S}\mathrm{t}_{\Delta^n}(\Delta^k \times \mathbb{O}^n(0,n))(n) \simeq \mathbb{S}\mathrm{t}_*(\Delta^k \times \mathbb{O}^n(0,n)) \to \Delta^k_\flat \times \mathbb{O}^n(0,n)$ can also be identified with the component of the natural transformation α and thus it is a weak equivalence. We conclude that ϕ is a weak equivalence. □

Our final claim is established. Therefore, Proposition 5.34 is proved.

5.4 The main theorem

Let $\mathrm{Set}^{\mathbf{sc}}_\Delta$ be the category of scaled simplicial sets and observe that we have a pair of functors

$$F_1, F_2 : (\mathrm{Set}^{\mathbf{sc}})^{\mathrm{op}}_\Delta \to \mathrm{Set}^+_\Delta\text{–}\mathrm{Cat}, \quad F_1(S) = \mathrm{Fun}^{\mathrm{o}}(\mathfrak{C}^{\mathbf{sc}}[S], \mathrm{Set}^{\mathbf{ms}}_\Delta), \quad F_2(S) = \big(\mathrm{Set}^{\mathbf{mb}}_\Delta\big)^{\mathrm{o}}_{/(S,\sharp,T_S\subset\sharp)},$$

which values in the category of Set^+_Δ-enriched categories and where the superscript "o" denotes the full (enriched) subcategory on fibrant-cofibrant objects. Let us remind the reader that it then follows that for every $S \in \mathrm{Set}^{\mathbf{sc}}_\Delta$ we have fibrant Set^+_Δ-enriched categories $F_i(S)$ for $i = 1, 2$.

We claim that the unstraightening construction $\mathbb{U}\mathrm{n}_{(-)}$ defines a natural transformation. In virtue of Remark 5.15 it will be enough to show that for every scaled simplicial set S we have that $\mathbb{U}\mathrm{n}_S$ defines a Set^+_Δ-enriched functor. Given a fibrant-cofibrant functors $\mathcal{F}, \mathcal{G} : \mathfrak{C}^{\mathbf{sc}}[S] \to \mathrm{Set}^{\mathbf{ms}}_\Delta$ let $\mathrm{Nat}(\mathcal{F}, \mathcal{G})$ denote the corresponding mapping marked simplicial set. Then a map $K \to \mathrm{Nat}(\mathcal{F}, \mathcal{G})$ is precisely the data of an enriched natural transformation $K \otimes \mathcal{F} \Rightarrow \mathcal{G}$ where

$$K \otimes \mathcal{F}(s) = K \times \mathcal{F}(s).$$

We can consequently define a map $K \to \mathrm{Map}_S(\mathbb{U}\mathrm{n}_S(\mathcal{F}), \mathcal{G})$ as the composite

$$K \times \mathbb{U}\mathrm{n}_S(\mathcal{F}) \to \mathrm{Un}_*(K) \times \mathbb{U}\mathrm{n}_S(\mathcal{F}) \simeq \mathbb{U}\mathrm{n}_S(K \times \mathcal{F}) \to \mathbb{U}\mathrm{n}_S(\mathcal{G}),$$

which shows that $\mathbb{U}\mathrm{n}_S$ defines a Set^+_Δ-enriched functor. The main goal of this section is to show that for every scaled simplicial set S, it follows that $\mathbb{U}\mathrm{n}_S$ is an equivalence of Set^+_Δ-enriched categories.

Proposition 5.47 *Let S be a scaled simplicial set. Then the following are equivalent*:

(i) *The functor $\mathbb{U}\mathrm{n}_S$ is a right Quillen equivalence.*

(ii) *The functor $\mathbb{U}\mathrm{n}_S$ defines an equivalence of fibrant Set^+_Δ-enriched categories after restriction to the full subcategories of fibrant-cofibrant objects.*

Proof It follows from Proposition 3.1.10 in [18] and our previous discussion that (ii) $\Longrightarrow$ (i). To show that (i) $\Longrightarrow$ (ii) we show that given a marked simplicial set K, $\mathbb{U}\mathrm{n}_S$ induces an isomorphism in the homotopy category of Set^+_Δ between $[K, \mathrm{Nat}(\mathcal{F}, \mathcal{G})] \simeq [K, \mathrm{Map}_S(\mathbb{U}\mathrm{n}_S(\mathcal{F}, \mathbb{U}\mathrm{n}_S(\mathcal{G}))]$. This follows from the chain of isomorphisms

$$[K, \mathrm{Nat}(\mathcal{F}, \mathcal{G})] \simeq [K \otimes \mathcal{F}, \mathcal{G}]^{\mathrm{Fun}} \simeq [\mathbb{U}\mathrm{n}_*(K) \times \mathbb{U}\mathrm{n}_S(\mathcal{F}), \mathbb{U}\mathrm{n}_S(\mathcal{G})]^{\mathrm{Fib}} \simeq [K, \mathrm{Map}_S(\mathbb{U}\mathrm{n}_S(\mathcal{F}), \mathbb{U}\mathrm{n}_S(\mathcal{G}))],$$

where the second isomorphism is a consequence of (i). □

Remark 5.48 Let S be a scaled simplicial set and $\phi : \mathfrak{C}^{\mathbf{sc}}[S] \to \mathcal{C}$ an equivalence of $\mathrm{Set}_{\Delta}^{+}$-enriched categories. Observe that it follows from Proposition 5.14 that $\mathbb{U}\mathrm{n}_{\phi}$ is a right Quillen equivalence if and only if $\mathbb{U}\mathrm{n}_S$ is a right Quillen equivalence. Therefore for the rest of the section we will let $\phi = \mathrm{id}$.

Corollary 5.49 *Let $\Delta^2_{\sharp} = (\Delta^2, \sharp)$ denote a maximally scaled 2-simplex. Then the straightening-unstraightening adjunction*

$$\mathbb{S}\mathrm{t}_{\Delta^2_{\sharp}} : \left(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\right)_{(\Delta^2,\sharp,\sharp)} \rightleftarrows \mathrm{Fun}(\mathfrak{C}^{\mathbf{sc}}[\Delta^2_{\sharp}], \mathrm{Set}_{\Delta}^{\mathbf{ms}}) : \mathbb{U}\mathrm{n}_{\Delta^2_{\sharp}}$$

is a Quillen equivalence.

Proof Observe that we have a commutative diagram

$$\begin{tikzcd}
\mathrm{Fun}^{\mathrm{o}}(\mathfrak{C}^{\mathbf{sc}}[\Delta^2_{\sharp}], \mathrm{Set}_{\Delta}^{\mathbf{ms}}) \arrow[r, "\mathbb{U}\mathrm{n}_{\Delta^2_{\sharp}}"] \arrow[d] & \left(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\right)^{\mathrm{o}}_{(\Delta^2,\sharp,\sharp)} \arrow[d] \\
\mathrm{Fun}^{\mathrm{o}}(\mathfrak{C}^{\mathbf{sc}}[\Delta^2_{\flat}], \mathrm{Set}_{\Delta}^{\mathbf{ms}}) \arrow[r, "\mathbb{U}\mathrm{n}_{\Delta^2}"] & \left(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\right)^{\mathrm{o}}_{/(\Delta^2,\sharp,\flat\subset\sharp)}
\end{tikzcd}$$

where the vertical maps are fully faithful functors. We conclude that $\mathbb{U}\mathrm{n}_{\Delta^2_{\sharp}}$ is fully faithful. It follows from Proposition 5.47 that it will be enough to show that $\mathbb{U}\mathrm{n}_{\Delta^2_{\sharp}}$ is essentially surjective.

Let $p : X \to \Delta^2_{\sharp}$ be a fibrant object and pick a fibrant-cofibrant functor $\mathcal{F} : \mathfrak{C}^{\mathbf{sc}}[\Delta^2_{\flat}] \to \mathrm{Set}_{\Delta}^{\mathbf{ms}}$ such that $\mathbb{U}\mathrm{n}_{\Delta^2}(\mathcal{F}) \simeq X$. To finish the proof we need to show that $\mathcal{F}$ factors through $\mathfrak{C}^{\mathbf{sc}}[\Delta^2_{\sharp}]$. Since $\mathbb{U}\mathrm{n}_{\Delta^2}(\mathcal{F})$ is equivalent to X it follows that the composition of local $(0,1)$-cartesian edges in this fibration remains $(0,1)$-cartesian. Direct inspection reveals that our functor must factor through $\mathfrak{C}^{\mathbf{sc}}[\Delta^2_{\sharp}]$. □

Remark 5.50 It follows from Lemma A.3.6.17 and Corollary A.3.6.18 in [18] that F_1 sends homotopy colimits of scaled simplicial sets to homotopy limits of $\mathrm{Set}_{\Delta}^{+}$-enriched categories.

Lemma 5.51 *Let $f : S_0 \to S$ be a cofibration of scaled simplicial sets. Then the functor*

$$f^* : \left(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\right)^{\mathrm{o}}_{S} \to \left(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\right)^{\mathrm{o}}_{S_0}$$

is a fibration of $\mathrm{Set}_{\Delta}^{+}$*-enriched categories.*

Proof Since both $\mathrm{Set}_{\Delta}^{+}$-enriched categories are fibrant it follows from Theorem A.3.2.24 in [18] that it will be enough to show the following:

$(*)$ Given a pair of fibrant objects $X, Y \in \left(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\right)^{\mathrm{o}}_{S}$, the induced morphism on mapping ∞-categories

$$\mathrm{Map}_S^{\leq 1}(X, Y) \to \mathrm{Map}_{S_0}^{\leq 1}(f^*X, f^*Y)$$

is a fibration of marked simplicial sets.

More generally we consider a pair of adjoint lifting problems

$$\begin{tikzcd} A \arrow[r] \arrow[d] & \mathrm{Map}_S(X,Y) \arrow[d] \\ B \arrow[r] \arrow[ur, dotted] & \mathrm{Map}_{S_0}(f^*X, f^*Y) \end{tikzcd} \quad \leftrightsquigarrow \quad \begin{tikzcd} A \times X \coprod_{A \times f^*X} B \times f^*X \arrow[r] \arrow[d] & Y \arrow[d] \\ B \times X \arrow[r] \arrow[ur, dashed] & S \end{tikzcd}$$

where $A \to B$ is **MB**-anodyne. Since $f : S_0 \to S$ is a cofibration it follows that the canonical map $f^*X \to X$ is a cofibration so we can use Proposition 3.12 to conclude that we can produce the desired solution to the lifting problem. □

Theorem 5.52 *Let* S *be a scaled simplicial set. Then the functor* $\mathbb{U}\mathrm{n}_S$ *induces an equivalence of* Set_Δ^+*-enriched categories*

$$\mathbb{U}\mathrm{n}_S : \mathrm{Fun}^{\mathrm{o}}(\mathfrak{C}^{\mathbf{sc}}[S], \mathrm{Set}_\Delta^{\mathbf{ms}}) \to \big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)^{\mathrm{o}}_{/(S,\sharp,T_S \subset \sharp)}.$$

Proof We say that a scaled simplicial set S is good if the conclusion of the theorem holds. In virtue of Proposition 5.47 we know that Proposition 5.34 shows that the scaled simplicial sets $(\Delta^n, \flat)$ are good for $n \geqslant 0$. Moreover, it follows from Corollary 5.49 that $(\Delta^2, \sharp)$ is also good.

Recall that every scaled simplicial set S can be expressed as a filtered colimit over the natural numbers

$$S_0 \to S_1 \to \cdots \to S_k \to \cdots$$

such that each map $S_i \to S_{i+1}$ is a cofibration and such that S_0 is a disjoint union of points. Moreover, a simplex by simplex argument shows that each step in this filtration can be obtained via pushouts along the cofibrations

- $(\partial\Delta^n, \flat) \to (\Delta^n, \flat)$ for $n \geqslant 0$,
- $(\Delta^2, \flat) \to (\Delta^2, \sharp)$.

We saw in Remark 5.50 that F_1 maps homotopy colimits to homotopy limits. We see that in order to finish the proof it will be enough to show that F_2 maps the colimits appearing in our filtration to homotopy limits. Using Lemma 5.51 we reduce the problem to verifying that F_2 maps those colimits to ordinary limits which follows from direct inspection. □

Corollary 5.53 *Let* S *be a scaled simplicial set and let* $\phi : \mathfrak{C}^{\mathbf{sc}}[S] \to \mathcal{C}$ *be an equivalence of* Set_Δ^+*-enriched categories. Then the straightening-unstraightening adjunction*

$$\mathbb{S}\mathrm{t}_\phi : \big(\mathrm{Set}_\Delta^{\mathbf{mb}}\big)_{/(S,\sharp,T_S \subset \sharp)} \rightleftarrows \mathrm{Fun}(\mathcal{C}, \mathrm{Set}_\Delta^{\mathbf{ms}}) : \mathbb{U}\mathrm{n}_\phi$$

is a Quillen equivalence.

Remark 5.54 Let N^{sc} be as in Definition 2.9. Given a scaled simplicial set S we define $\mathbb{F}\mathrm{ib}_{0,1}(S) = \mathrm{N}^{\mathrm{sc}}((\mathrm{Set}_\Delta^{\mathbf{mb}})^{\mathrm{o}}_{/(S,\sharp,T_S \subset \sharp)})$. We also define $\mathbb{B}\mathrm{icat}_\infty = \mathrm{N}^{\mathrm{sc}}((\mathrm{Set}_\Delta^{\mathbf{ms}})^{\mathrm{o}})$ and observe that an analogous discussion to that of [3] shows that Theorem 5.52 shows that we have an equivalence of ∞-bicategories

$$\mathbb{U}\mathrm{n}_S : \mathrm{Fun}(S, \mathbb{B}\mathrm{icat}_\infty) \to \mathbb{F}\mathrm{ib}_{0,1}(S).$$

Definition 5.55 Let $\mathbb{S} = (S, T_S)$ be an ∞-bicategory and let $S_{M_S} = (S, M_S)$ (see Definition 4.1). We denote by $\mathbb{L}\mathrm{Fib}(\mathbb{S}) = \mathbb{F}\mathrm{ib}_{0,1}(S_{M_S})$ the ∞-bicategory of local $(0,1)$-fibrations over $\mathbb{S}$. Similarly, given another ∞-bicategory $\mathbb{D}$ we define $\mathrm{Fun}^{\mathrm{oplax}}(\mathbb{S}, \mathbb{D}) = \mathrm{Fun}(S_{M_S}, \mathbb{D})$.

Remark 5.56 In [11], the definition of $\mathrm{Fun}^{\mathrm{oplax}}(\mathbb{S}, \mathbb{D})$ is proposed as a model for oplax normalised functors. In [1], we establish an equivalence between the aforementioned notion of an oplax normalised functor and that given in [14; 15].

Corollary 5.57 *Let $\mathbb{S}$ be an ∞-bicategory. Then the straightening-unstraightening adjunction, associated to the scaled simplicial set (S, M_S),*

$$\mathbb{S}\mathrm{t}_{\mathbb{S}}^{\mathrm{oplax}} : \mathbb{L}\mathrm{Fib}(\mathbb{S}) \rightleftarrows \mathrm{Fun}^{\mathrm{oplax}}(\mathbb{S}, \mathbb{B}\mathrm{icat}_\infty) : \mathbb{U}\mathrm{n}_{\mathbb{S}}^{\mathrm{oplax}}$$

yields an equivalence of ∞-bicategories between the ∞-bicategory of local $(0,1)$-fibrations over $\mathbb{S}$ and the ∞-bicategory of oplax normalised functors with values in ∞-bicategories.

Proof This follows immediately from Theorem 5.52 and Remark 5.54. ☐

Definition 5.58 Let $\mathbb{S} = (S, T_S)$ be an ∞-bicategory and let $S_{\mathcal{U}} = (S, \mathcal{U})$ where $M_S \subset \mathcal{U} \subset T_S$. We denote by $\mathbb{L}\mathrm{Fib}^{\mathcal{U}}(\mathbb{S}) = \mathbb{F}\mathrm{ib}_{0,1}(S_{\mathcal{U}})$ the ∞-bicategory of $\mathcal{U}$-local $(0,1)$-fibrations over $\mathbb{S}$ (see Definition 4.28 and Theorem 4.29). Similarly, given another ∞-bicategory $\mathbb{D}$ we define $\mathrm{Fun}^{\mathcal{U}\mathrm{oplax}}(\mathbb{S}, \mathbb{D}) = \mathrm{Fun}(S_{\mathcal{U}}, \mathbb{D})$.

Corollary 5.59 *Let $\mathbb{S}$ be an ∞-bicategory. Then the straightening-unstraightening adjunction associated to the scaled simplicial set $(S, \mathcal{U})$ (see Definition 5.58)*

$$\mathbb{S}\mathrm{t}_{\mathbb{S}}^{\mathcal{U}\mathrm{oplax}} : \mathbb{L}\mathrm{Fib}^{\mathcal{U}}(\mathbb{S}) \rightleftarrows \mathrm{Fun}^{\mathcal{U}\mathrm{oplax}}(\mathbb{S}, \mathbb{B}\mathrm{icat}_\infty) : \mathbb{U}\mathrm{n}_{\mathbb{S}}^{\mathcal{U}\mathrm{oplax}}$$

yields an equivalence of ∞-bicategories between the ∞-bicategory of $\mathcal{U}$-local $(0,1)$-fibrations over $\mathbb{S}$ and the ∞-bicategory consisting in those oplax normalised functors which preserve the composites in $\mathcal{U}$.

6 The ∞-bicategorical Yoneda embedding

In this section, we present a proof of the Yoneda lemma for $(\infty, 2)$-categories. We would like to point out that this result has already appeared in the work of Hinich [16] in the context of enriched ∞-category theory. Throughout this section we fix an ∞-bicategory $\mathbb{C}$.

Definition 6.1 Let $\mathbb{C}$ be an ∞-bicategory and let $\mathbb{F}(\mathbb{C}) = \mathrm{Fun}^{\mathrm{opgr}}(\Delta^1, \mathbb{C})$ (see Definition 5.1). We observe that evaluation at 0 and evaluation at 1 induce maps $\mathrm{ev}_i : \mathbb{F}(\mathbb{C}) \to \mathbb{C}$ for $i = 0, 1$. It follows from Proposition 2.2.6 in [13] that evaluation at 0 yields a $(1,0)$-fibration. One can similarly show that the map ev_1 defines a $(0,1)$-fibration.

Remark 6.2 Let us recall that a morphism in $\mathbb{F}(\mathbb{C})$ is $(0,1)$-cartesian if and only if it is sent to an equivalence under ev_0. Similarly, a 2-morphism is cartesian in $\mathbb{F}(\mathbb{C})(x, y)$ if and only if its image in $\mathbb{C}$ under ev_0 is an invertible. The same description holds for the $(1,0)$-cartesian morphisms by replacing ev_0 with ev_1.

Definition 6.3 Let $\mathbb{C}$ be an ∞-bicategory. We define an ∞-bicategory $\mathbb{F}\mathrm{ib}_{1,0}(\mathbb{C})$ in a manner entirely analogous to Remark 5.54, that is, as the nerve (see Definition 2.9) of the fibrant $\mathrm{Set}_{\Delta}^{+}$-enriched category $\big(\mathrm{Set}_{\Delta}^{\mathbf{mb}}\big)^{\circ}_{/(\mathbb{C},\sharp,T_{\mathbb{C}}\subset\sharp)}$ (see Section 5.4) associated to the model structure on $(1,0)$-fibrations (see [6]).

Remark 6.4 For the rest of this section we will work simultaneously with $(0,1)$- and $(1,0)$-fibrations and we will consequently introduce some notation to avoid confusion. Given an (i,j)-fibration, incarnated as a fibrant **MB** simplicial set $p:(X,E_X,T_X\subseteq C_X)\to(S,\sharp,\sharp)$ and any other **MB** simplicial set $(A,E_A,T_A\subseteq E_A)$, we will denote by $\mathrm{Map}_S^{i,j}(A,X)$ the ∞-bicategory obtained (see Definition 3.15) as the pullback

$$\begin{tikzcd}
\mathrm{Map}_S^{i,j}(A,X) \arrow[r] \arrow[d] & \mathrm{Fun}^{\mathbf{mb}}(A,X) \arrow[d] \\
\Delta^0 \arrow[r, "q"] & \mathrm{Fun}^{\mathbf{mb}}(A,S)
\end{tikzcd}$$

Note that the underlying marked simplicial set of $\mathrm{Map}_{\mathbb{C}}^{i,j}(-,-)$ is precisely the mapping marked simplicial set in $\mathbb{F}\mathrm{ib}_{i,j}(\mathbb{C})$.

Remark 6.5 Let $f:\mathbb{A}\to\mathbb{C}$ be a functor of ∞-bicategories and consider a pullback diagram

$$\begin{tikzcd}
\mathbb{F}(\mathbb{C})\times_{\mathbb{C}}\mathbb{A} \arrow[r] \arrow[d] & \mathbb{F}(\mathbb{C}) \arrow[d, "\mathrm{ev}_1"] \\
\mathbb{A} \arrow[r, "f"] & \mathbb{C}
\end{tikzcd}$$

It follows that evaluation at 0 induces a map $\mathbb{F}(\mathbb{C})\times_{\mathbb{C}}\mathbb{A}\to\mathbb{C}$ which is a $(1,0)$-fibration by Proposition 3.8 in [4]. We observe that we have a commutative diagram

$$\begin{tikzcd}
\mathbb{A} \arrow[rr] \arrow[dr, "f"'] & & \mathbb{F}(\mathbb{C})\times_{\mathbb{C}}\mathbb{A} \arrow[dl, "\mathrm{ev}_0"] \\
& \mathbb{C} &
\end{tikzcd}$$

where the horizontal morphism is induced by the map $\Delta^1\otimes\mathbb{A}\to\Delta^0\otimes\mathbb{A}\simeq\mathbb{A}$. Moreover, we see as a consequence of Theorem 3.17 in [4] that for every $(1,0)$-fibration $\pi:\mathbb{X}\to\mathbb{C}$ we have a trivial fibration of ∞-bicategories

$$\mathrm{Map}_{\mathbb{C}}^{1,0}(\mathbb{F}(\mathbb{C})\times_{\mathbb{C}}\mathbb{A},\mathbb{X})\to\mathrm{Fun}_{\mathbb{C}}(\mathbb{A},\mathbb{X}),$$

where we are denoting by $\mathrm{Map}_{\mathbb{C}}^{1,0}(-,-)$ the mapping ∞-bicategory (as in Remark 6.4) between $(1,0)$-fibrations and by $\mathrm{Fun}_{\mathbb{C}}(\mathbb{A},\mathbb{X})$ the ∞-bicategory obtained via the pullback

$$\begin{tikzcd}
\mathrm{Fun}_{\mathbb{C}}(\mathbb{A},\mathbb{X}) \arrow[r] \arrow[d] & \mathrm{Fun}(\mathbb{A},\mathbb{X}) \arrow[d, "\pi_*"] \\
\Delta^0 \arrow[r, "f"] & \mathrm{Fun}(\mathbb{A},\mathbb{C})
\end{tikzcd}$$

We are now ready to define the Yoneda embedding. Let us consider the $(0,1)$-fibration $\mathrm{ev}_1 : \mathbb{F}(\mathbb{C}) \to \mathbb{C}$ and observe that we have a commutative diagram, over $\mathbb{C}$,

$$\begin{tikzcd}
\mathbb{F}(\mathbb{C}) \arrow[rr, "\mathrm{ev}_0 \times \mathrm{ev}_1"] \arrow[dr, "\mathrm{ev}_1"'] & & \mathbb{C} \times \mathbb{C} \arrow[dl, "\pi_2"] \\
& \mathbb{C} &
\end{tikzcd}$$

where π_2 is the projection onto the second factor. It follows from our definitions that the map $\mathrm{ev}_0 \times \mathrm{ev}_1$ can be seen as a map of $(0,1)$-fibrations where $\mathbb{C} \times \mathbb{C}$ is classified by the constant functor with value $\mathbb{C}$. Let $\mathcal{F}$ be the functor classified by $\mathbb{F}(\mathbb{C}) \to \mathbb{C}$. We make the following observations:

(1) For every $c \in \mathbb{C}$ we have a functor $p_c : \mathcal{F}(c) \to \mathbb{C}$ and for every morphism $u : c \to c'$ we have a commutative diagram $p_{c'} \circ \mathcal{F}(u) = p_c$.

(2) For every $c \in \mathbb{C}$ it follows that we have a morphism

$$\mathbb{F}(\mathbb{C}) \times_{\mathbb{C}} * = \mathbb{C}_{\nearrow c} \to \mathbb{C},$$

which is a $(1,0)$-fibration by Remark 6.5. Furthermore, we note that for every $c \in \mathbb{C}$ it follows that we have a morphism

$$\mathbb{F}(\mathbb{C}) \times_{\mathbb{C}} * = \mathbb{C}_{\nearrow c} \to \mathbb{C},$$

which is a $(1,0)$-fibration by Remark 6.5 whose fibres are ∞-categories. Careful inspection reveals that for every $u : c \to c'$ the induced morphism $u_* : \mathbb{C}_{\nearrow c} \to \mathbb{C}_{\nearrow c'}$ preserves $(1,0)$-cartesian edges.

(3) Combining (1) and (2) we see that $p_c : \mathcal{F}(c) \to \mathbb{C}$ is a $(1,0)$-fibration and that $\mathcal{F}(u)$ is a functor of $(1,0)$-fibrations.

We conclude that $\mathcal{F}$ can be expressed as a composite

$$\mathcal{F} : \mathbb{C} \to \mathbb{F}\mathrm{ib}_{1,0}(\mathbb{C})^{\text{1-fib}} \to \mathbb{B}\mathrm{icat}_\infty,$$

where the second functor is the obvious projection and the superscript "1-fib" denotes the ∞-bicategory spanned by $(1,0)$-fibrations whose fibres are ∞-categories. Using a dual version of our main result or equivalently Corollary 3.90 in [3] we obtain a functor

$$\mathcal{Y}_{\mathbb{C}} : \mathbb{C} \to \mathrm{Fun}(\mathbb{C}^{\mathrm{op}}, \mathbb{C}\mathrm{at}_\infty),$$

which we call the bicategorical Yoneda embedding. Dually, using the map $\mathrm{ev}_0 : \mathbb{F}(\mathbb{C}) \to \mathbb{C}$ we obtain the co-Yoneda embedding

$$\mathcal{Y}^{\mathrm{co}}_{\mathbb{C}} : \mathbb{C}^{\mathrm{op}} \to \mathrm{Fun}(\mathbb{C}, \mathbb{C}\mathrm{at}_\infty).$$

Definition 6.6 Let $f : (A, E_A, T_A \subset C_A) \to (\mathbb{C}, \sharp, T_{\mathbb{C}} \subset \sharp)$ be a map of **MB** simplicial sets. We define an **MB** simplicial set $(Q(f), E_f, T_f \subset C_f)$ over $(\mathbb{C}, \sharp, T_{\mathbb{C}} \subset \sharp)$ as follows:

- The underlying map of simplicial sets is the composite

$$\mathbb{F}(\mathbb{C}) \times_{\mathbb{C}} A \to \mathbb{F}(\mathbb{C}) \xrightarrow{\mathrm{ev}_0} \mathbb{C},$$

where the pullback is taken along to the map $\mathrm{ev}_1 : \mathbb{F}(\mathbb{C}) \to \mathbb{C}$.

- An edge is marked if its associated map $\Delta^1 \otimes \Delta^1 \to \mathbb{C}$ factors through $\Delta^1 \times \Delta^1$ and the restriction to $\Delta^1 \times \Delta^{\{1\}}$ is marked in A.
- A triangle is lean if its restriction $\Delta^{\{1\}} \otimes \Delta^2$ is lean in A.
- A triangle is thin if it is lean and its image in $\mathbb{C}$ is thin.

The proofs in [4, Theorem 3.17, Corollary 3.20] are of an entirely combinatorial nature and under close inspection one sees we always have an **MB** anodyne morphism $(A, E_A, T_A \subset C_A) \to (Q(f), E_f, T_f \subset C_f)$ (in the model structure for $(1,0)$-fibrations) whose definition is induced by $\Delta^1 \otimes A \to A$ as in Remark 6.5.

Proposition 6.7 *There exists a functor* $\mathbb{I} : \mathbb{F}\mathrm{ib}_{1,0}(\mathbb{C})^{\text{1-fib}} \to \mathbb{F}\mathrm{ib}_{1,0}(\mathbb{C})^{\text{1-fib}}$ *which sends a* $(1,0)$*-fibration* $p : \mathcal{G} \to \mathbb{C}$ *with* ∞*-categorical fibres to the* $(1,0)$*-fibration* $\mathbb{I}(\mathcal{G})$ *which is defined as the* ***MB*** *simplicial set characterised uniquely by the isomorphism*

$$\mathrm{Map}^{1,0}_{\mathbb{C}}(A, \mathbb{I}(\mathcal{G})) \cong \mathrm{Map}^{1,0}_{\mathbb{C}}(Q(f), \mathcal{G}),$$

where $f : (A, E_A, T_A \subseteq C_A) \to (\mathbb{C}, \sharp, T_{\mathbb{C}} \subset \sharp)$ *is a map of* ***MB*** *simplicial sets and* $Q(f)$ *is defined as in Definition 6.6.*

Proof Note that since $Q(-)$ is functorial on $\left(\mathrm{Set}^{\mathbf{mb}}_{\Delta}\right)_{/(\mathbb{C},\sharp,T_{\mathbb{C}}\subset\sharp)}$ it follows that our definition yields an **MB** simplicial set over $(\mathbb{C}, \sharp, T_{\mathbb{C}} \subset \sharp)$.

There are two main things to prove: we need to show that $\mathbb{I}(\mathcal{G}) \to \mathbb{C}$ is a $(1,0)$-fibration with ∞-categorical fibres and that the construction $\mathbb{I}(-)$ is functorial. We will start by first proving the second assertion. Note that by Definition 6.3, it suffices to verify that $\mathbb{I}(-)$ yields a functor of Set^+_{Δ}-enriched categories.

Let $K \in \mathrm{Set}^+_{\Delta}$ (which we view as having the maximal scaling) and consider a map $K \to \mathrm{Map}^{1,0}_{\mathbb{C}}(\mathcal{G}, \mathcal{H})$. We will construct a morphism $K \to \mathrm{Map}^{1,0}_{\mathbb{C}}(\mathbb{I}(\mathcal{G}), \mathbb{I}(\mathcal{H}))$ as follows:

Given a simplex $\Delta^n \xrightarrow{\sigma \times \bar{\varphi}} \mathbb{I}(\mathcal{G}) \times K$ we consider a morphism

$$\mathbb{F}(\mathbb{C}) \times_{\mathbb{C}} \Delta^n \xrightarrow{f_\sigma \times \varphi} \mathcal{G} \times K \to \mathcal{H},$$

where φ is given by the composite $\mathbb{F}(\mathbb{C}) \times_{\mathbb{C}} \Delta^n \to \Delta^n \xrightarrow{\bar{\varphi}} K$. This map is clearly compatible with the projection and functorial a thus yields a map $\mathbb{I}(\mathcal{G}) \times K \to \mathbb{I}(\mathcal{H})$.

To finish the proof we will show that $\mathbb{I}(\mathcal{G}) \in \mathbb{F}\mathrm{ib}_{1,0}(\mathbb{C})^{\text{1-fib}}$. First let us observe that given $c \in \mathbb{C}$ we have a canonical isomorphism

$$\mathbb{I}(\mathcal{G}) \times_{\mathbb{C}} \{c\} \simeq \mathrm{Map}^{1,0}_{\mathbb{C}}(\mathbb{C}_{\nearrow c}, \mathcal{G}),$$

which shows that the fibres of $\mathbb{I}(\mathcal{G})$ are in fact ∞-categories. Let us consider a commutative diagram of **MB** simplicial sets

$$\begin{tikzcd}
A \arrow[rr, "i"] \arrow[dr, "u"'] & & B \arrow[dl, "v"] \\
& \mathbb{C} &
\end{tikzcd}$$

where i is an anodyne morphism in the model structure for $(1,0)$-fibrations developed in [6]. It follows from Definition 6.6 that we have a commutative diagram

$$\begin{tikzcd}
A \arrow[r, "\simeq"] \arrow[d, "\simeq"] & Q(u) \arrow[d] \\
B \arrow[r, "\simeq"] & Q(v)
\end{tikzcd}$$

where the horizontal morphisms are weak equivalences. It then follows by two-out-of-three that the right-most vertical morphism is also a weak equivalence. We conclude that $\mathbb{I}(\mathcal{G})$ is a $(1,0)$-fibration. □

Remark 6.8 By the previous proposition the fibration $\mathbb{I}(\mathcal{G})$ corresponds under the straightening equivalence to a functor $\mathbb{C}^{\mathrm{op}} \to \mathbb{C}\mathrm{at}_\infty$ mapping an object c to

$$\mathrm{Map}^{1,0}_{\mathbb{C}}(\mathbb{C}_{\nearrow c}, \mathcal{G}) \simeq \mathrm{Nat}_{\mathbb{C}^{\mathrm{op}}}(\mathbb{C}(-,c), \mathbb{S}\mathrm{t}_{\mathbb{C}}(\mathcal{G})),$$

where $\mathrm{Nat}_{\mathbb{C}^{\mathrm{op}}}(-,-)$ is the mapping ∞-category in $\mathrm{Fun}(\mathbb{C}^{\mathrm{op}}, \mathbb{C}\mathrm{at}_\infty)$. We will omit the explicit verification of the fact that there exists an equivalence of contravariant functors

$$\mathbb{S}\mathrm{t}_{\mathbb{C}}(\mathbb{I}(\mathcal{G})) \simeq \mathrm{Nat}_{\mathbb{C}^{\mathrm{op}}}(\mathcal{Y}_{\mathbb{C}}(-), \mathbb{S}\mathrm{t}_{\mathbb{C}}(\mathcal{G})).$$

Remark 6.9 The coming proofs will involve using the straightening-unstraightening equivalence for $(0,1)$- and $(1,0)$-fibrations. We will employ the notation $\mathbb{S}\mathrm{t}^{i,j}_S$ to distinguish between both variances.

Definition 6.10 Let $\mathbb{C}$ be an ∞-bicategory and pick a fibrant replacement $s : \mathfrak{C}^{\mathrm{sc}}[\mathbb{C}] \xrightarrow{\simeq} \mathbb{C}$. We consider a functor of fibrant Set^+_Δ-enriched categories

$$\mathbb{Y}_{\mathbb{C}} : \mathfrak{C}^{\mathrm{sc}}[\mathbb{C}] \xrightarrow{J} \mathrm{Fun}_{\mathrm{Set}_{\Delta^+}}(\mathfrak{C}^{\mathrm{sc}}[\mathbb{C}]^{\mathrm{op}}, \mathbb{C}\mathrm{at}_\infty) \xrightarrow{\mathbb{U}\mathrm{n}^{1,0}_{\mathbb{C}}} \mathbb{F}\mathrm{ib}_{1,0}(\mathbb{C}) \to \mathbb{B}\mathrm{icat}_\infty,$$

where the first functor sends each $c \in \mathfrak{C}^{\mathrm{sc}}[\mathbb{C}]$ to the functor $\mathbb{C}(s(-), s(c))$. Now we can consider the Grothendieck construction and obtain a $(0,1)$-fibration $p : \mathfrak{Y}_{\mathbb{C}} \to \mathbb{C}$.

The definition above yields another reasonable candidate for the Yoneda embedding. In the next theorem, we will identify our purely fibrational approach to the Yoneda embedding with the previous definition, which relies heavily on the model of Set^+_Δ-enriched categories. Before proving the main results of the section we will need some preliminary work.

Construction 6.11 Let $\mathbb{C}$ be an ∞-bicategory and let $\mathfrak{C}^{\mathrm{sc}}[\Delta^n_\flat] = \mathbb{O}^n$. Let us denote by $\iota_n : \Delta^n_\flat \to \mathrm{N}^{\mathrm{sc}}(\mathbb{O}^n)$ (see Definition 2.9) the canonical trivial cofibration adjoint to the identity map. Then for every $\sigma : \Delta^n_\flat \to \mathbb{C}$ we can pick an extension $E_\sigma : \mathrm{N}^{\mathrm{sc}}(\mathbb{O}^n) \to \mathbb{C}$ such that $E_\sigma \circ \iota_n = \sigma$. We claim that we can make a choice $\{E_\sigma\}_\sigma$ where σ runs over the nondegenerate simplices of $\mathbb{C}$ which is in addition functorial. This means that if we are given a map $\tau : \Delta^\ell_\flat \to \Delta^n_\flat$ then $E_\sigma \circ \mathfrak{C}^{\mathrm{sc}}[\tau] = E_{\sigma\circ\tau}$.

Since $\mathrm{N}^{\mathrm{sc}}(\mathbb{O}^n) = \Delta^n_\flat$ for $n = 0, 1$, the choices are already fixed for dimensions $n \leqslant 1$. Suppose we have made a compatible choice for simplices up to dimension $k-1$ and consider $\Delta^k_\flat \to \mathbb{C}$. Then we can consider the homotopy pushout diagram

$$\begin{tikzcd}
\partial\Delta^k_\flat \arrow[r, "\simeq"] \arrow[d] & \operatorname{colim}_{\partial\Delta^k} \mathrm{N}^{\mathrm{sc}}(\mathbb{O}^{k-1}) \arrow[d] \\
\Delta^n_\flat \arrow[r, "\simeq"] & P
\end{tikzcd}$$

where top horizontal (and hence the bottom) map is a weak equivalence the colimits involved are homotopy colimits. We obtain by the universal property of the pushout a map $P \to \mathbb{C}$. Moreover, we have a factorisation of ι_n as $\Delta^n_\flat \to P \xrightarrow{j} \mathrm{N}^{\mathrm{sc}}(\mathbb{O}^n)$, so we conclude by two-out-of-three that j is also a weak equivalence. The solution of the lifting problem

$$\begin{tikzcd}
P \arrow[r] \arrow[d, "j"] & \mathbb{C} \\
\mathrm{N}^{\mathrm{sc}}(\mathbb{O}^n) \arrow[ur, dotted] &
\end{tikzcd}$$

provides the desired functorial extension.

Definition 6.12 We define a functor $\mathbb{P}^n : \mathfrak{C}^{\mathrm{sc}}[\Delta^n_\flat] \to \mathrm{Set}^{\mathbf{ms}}_\Delta$ as follows:

- For every $i \in [n]$ we declare the value $\mathbb{P}^n(i)$ to be the scaled poset $\mathcal{P}_{[0,i]}$ given in Definitions 5.2 and 5.5 (equipped with the minimal marking on 1-simplices).
- For every $S \subset [n]$ with $\min(S) = i$ and $\max(S) = j$ we consider the functor $\mathcal{P}_{[0,i]} \to \mathcal{P}_{[0,j]}$ sending $U \in \mathcal{P}_{[0,i]}$ to $S \cup U$. Similarly given an inclusion $S \subset T$ we consider the natural transformation

$$\mathcal{P}_{[0,i]} \times \Delta^1 \to \mathcal{P}_{[0,j]},$$

with components given by $S \cup U \subset T \cup V$ (see Remark 5.3).

Remark 6.13 We have functors $\mathcal{P}_{[0,j]} \to \mathrm{N}^{\mathrm{sc}}(\mathbb{O}^n)$ which are uniquely determined by the assignment given by $S \mapsto \min(S)$, $S \leqslant T \mapsto U_{S,T}$ (see Definition 5.5 and Remark 5.4 to see why this definition is compatible with the scaling). These functors allow us to produce a lift $\mathbb{P}^n : \mathfrak{C}^{\mathrm{sc}}[\Delta^n_\flat] \to \left(\mathrm{Set}^{\mathbf{mb}}_\Delta\right)_{/\mathrm{N}^{\mathrm{sc}}(\mathbb{O}^n)}$ by picking the maximal lean scaling.

Lemma 6.14 *Let* $\phi : \mathfrak{C}^{\mathrm{sc}}[\mathrm{N}^{\mathrm{sc}}(\mathbb{O}^n)] \to \mathbb{O}^n$ *be the counit map associated to the Quillen equivalence* $\mathfrak{C}^{\mathrm{sc}} \dashv \mathrm{N}^{\mathrm{sc}}$ *(see Definition 2.9) where* $\mathbb{O}^n = \mathfrak{C}^{\mathrm{sc}}[\Delta^n_\flat]$*. Then there exists a natural transformation* $\alpha_n : \mathbb{P}^n \Rightarrow \mathbb{U}\mathrm{n}^{1,0}_\phi \circ \mathbb{T}$ *where* $\mathbb{T}^n(i) = \mathfrak{C}^{\mathrm{sc}}[\Delta^n_\flat](-,i)$*.*

Proof By adjunction, it will enough to produce natural transformation $\mathbb{S}\mathrm{t}^{1,0}_\phi \circ \mathbb{P}^n \Rightarrow \mathbb{T}$. Let $i \in \mathbb{O}^n$ and let $\mathcal{P}^{\triangleright}_{[0,i]}$ the scaled simplicial set obtained from the join $P^{\triangleright}_{[0,i]} = P_{[0,i]} * \Delta^0$, by scaling those simplices of the form $u(\sigma)$ where $u : P_{[0,i]} \to P^{\triangleright}_{[0,i]} = P_{[0,i]} * \Delta^0$ is the canonical map, and σ is scaled in $\mathcal{P}_{[0,i]}$. We consider a pushout diagram

$$\begin{tikzcd}
\mathfrak{C}^{\mathrm{sc}}[\mathcal{P}_{[0,i]}] \arrow[r] \arrow[d, "\Phi"] & \mathfrak{C}^{\mathrm{sc}}[\mathcal{P}^{\triangleright}_{[0,i]}] \arrow[d] \\
\mathbb{O}^n \arrow[r, "r"] & C_\phi(\mathbb{P}(i))
\end{tikzcd}$$

where Φ is the adjoint to the map given in Remark 6.13. Let $*$ be the cone point of $\mathfrak{C}^{\mathrm{sc}}[\mathcal{P}^{\triangleright}_{[0,i]}]$. Then it follows that $C_\phi(\mathbb{P}(i))(r(-),*) = \mathbb{S}\mathrm{t}^{1,0}_{\phi}(\mathbb{P}(i))$ as constructed in [3, Section 3]. Let $S \in \mathcal{P}_{[0,i]}$ with $\min(S) = s$. We consider the map

$$\bar{\alpha}^n_i(S) : \mathfrak{C}^{\mathrm{sc}}[\mathcal{P}^{\triangleright}_{[0,i]}](S,*) \to \mathbb{O}^n(s,i), \quad \{S = S_0 < S_1 < \cdots < S_n < *\} \mapsto S_n \cup \bigcup_{i=0}^{n} U_{S_i,S_{i-1}},$$

which is compatible with the marking of $\mathfrak{C}^{\mathrm{sc}}[\mathcal{P}_{[0,i]}](S,*)$ and descends to a map

$$\alpha^n_i(s) : \mathbb{S}\mathrm{t}^{1,0}_{\phi}(\mathbb{P}(i))(s) \to \mathbb{O}^n(s,i).$$

Note that if we are given $S < T$ with $\min(S) = s$ and $\min(T) = t$, then we have a commutative diagram

$$\begin{tikzcd}
\mathfrak{C}^{\mathrm{sc}}[\mathcal{P}^{\triangleright}_{[0,i]}](T,*) \arrow[r, "\bar{\alpha}^n_i(T)"] \arrow[d] & \mathbb{O}^n(t,i) \arrow[d, "\cup U_{S,T}"] \\
\mathfrak{C}^{\mathrm{sc}}[\mathcal{P}^{\triangleright}_{[0,i]}](S,*) \arrow[r, "\bar{\alpha}^n_i(S)"] & \mathbb{O}^n(s,i)
\end{tikzcd}$$

which guarantees that the maps $\{\alpha^n_i(s)\}_{s \in \mathbb{O}^n}$ assemble into a natural transformation $\alpha^n_i : \mathbb{S}\mathrm{t}^{1,0}_{\phi}(\mathbb{P}^n(i)) \Rightarrow \mathbb{O}^n(-,i)$. It is straightforward to verify that the maps $\{\alpha^n_i\}_{i \in \mathbb{O}^n}$ also assemble into the desired natural transformation $\alpha^n : \mathbb{S}\mathrm{t}^{1,0}_{\phi} \circ \mathbb{P}^n \Rightarrow \mathbb{T}$. □

Proposition 6.15 *Let $\mathbb{C}$ be an ∞-bicategory. Then there exists an equivalence of $(0,1)$-fibrations, over $\mathbb{C}$,*

$$\begin{tikzcd}
\mathbb{F}(\mathbb{C}) \arrow[rr, "\simeq"] \arrow[dr, "\mathrm{ev}_1"'] & & \mathfrak{Y}_{\mathbb{C}} \arrow[dl, "p"] \\
& \mathbb{C} &
\end{tikzcd}$$

Proof We observe that by construction there exists a map of $(0,1)$-fibrations $\mathfrak{Y}_{\mathbb{C}} \to \mathbb{C} \times \mathbb{C}$ such that for every $c \in \mathbb{C}$ the induced map on fibres

$$\mathfrak{Y}_{\mathbb{C}} \times_{\{c\}} \mathbb{C} \to \mathbb{C}$$

defines a $(1,0)$-fibration whose fibres are ∞-categories. Since the map $\mathbb{C} \to \mathbb{F}(\mathbb{C})$ is a trivial cofibration (this is the dual to Theorem 3.17 in [4]) it follows that to produce a map $\mathbb{F}(\mathbb{C}) \to \mathfrak{Y}_{\mathbb{C}}$ it suffices to construct a section of p.

Let $\sigma : \Delta^n_\flat \to \mathbb{C}$. Our goal is to produce an n-simplex in the pullback

$$\begin{tikzcd}
\mathfrak{Y}_n \arrow[r] \arrow[d] & \mathfrak{Y}_{\mathbb{C}} \arrow[d] \\
\Delta^n_\flat \times \Delta^n_\flat \arrow[r, "\sigma\times\sigma"] & \mathbb{C} \times \mathbb{C}
\end{tikzcd}$$

in a functorial way. Let $\mathbb{S}\mathrm{t}^{0,1}_{\Delta^n_\flat}(\Delta^n_\flat) : \mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat] \to \mathrm{Set}^{\mathbf{ms}}_{\Delta}$ be the straightening of the identity functor $\Delta^n_\flat \to \Delta^n_\flat$. The maps defined in Remark 5.6 provide us with a natural transformation $\mathbb{S}\mathrm{t}^{1,0}_{\Delta^n_\flat}(\Delta^n_\flat) \Rightarrow \mathbb{P}^n$ with $\mathbb{P}^n$ as in Definition 6.12. It then follows from Remark 6.13 that $\mathbb{S}\mathrm{t}^{(0,1)}_{\Delta^n_\flat}(\Delta^n_\flat)$ admits a lift to $(\mathrm{Set}^{\mathbf{mb}}_{\Delta})_{/\mathrm{N}^{\mathrm{sc}}(\mathbb{O}^n)}$.

Note that since we have a map $\mathfrak{C}^{\mathbf{sc}}[\sigma] : \mathfrak{C}^{\mathbf{sc}}[\Delta^n_\flat] \to \mathfrak{C}^{\mathbf{sc}}[\mathbb{C}] \xrightarrow{\simeq} \mathbb{C}$, it follows that we have a commutative diagram, for every $i \in \mathbb{O}^n$,

$$\begin{tikzcd}
\mathbb{U}\mathrm{n}^{1,0}_{\phi}(\mathbb{T}(i)) \arrow[r] \arrow[d] & \mathbb{U}\mathrm{n}^{1,0}_{\mathbb{C}}\big(J(\mathfrak{C}^{\mathbf{sc}}[\sigma](i))\big) \arrow[d] \\
\mathrm{N}^{\mathbf{sc}}(\mathbb{O}^n) \arrow[r, "E_\sigma"] & \mathbb{C}
\end{tikzcd}$$

where E_σ was given in Construction 6.11 and $\mathbb{T}$ in Lemma 6.14. This family of maps is natural in $i \in \mathbb{O}$, thus yielding a natural transformation $\mathbb{U}\mathrm{n}^{(1,0)}_{\phi} \circ \mathbb{T} \Rightarrow \mathbb{U}\mathrm{n}^{(1,0)}_{\mathbb{C}} \circ J \circ \mathfrak{C}^{\mathbf{sc}}[\sigma]$. Finally, let us observe that

$$\mathbb{S}\mathrm{t}^{0,1}_{\Delta^n_\flat}(\Delta^n_\flat) \Rightarrow \mathbb{P}^n \Rightarrow \mathbb{U}\mathrm{n}^{(1,0)}_{\phi} \circ \mathbb{T} \Rightarrow \mathbb{U}\mathrm{n}^{(1,0)}_{\mathbb{C}} \circ J \circ \mathfrak{C}^{\mathbf{sc}}[\sigma]$$

yields the desired simplex in $T(\sigma) : \Delta^n_\flat \to \mathfrak{Y}_n$. Note that the functoriality of our lifts E_σ (Construction 6.11) guarantees that the assignment $\sigma \mapsto T(\sigma)$ is functorial. It is also straightforward to verify that our assignment is compatible with the decorations thus yielding the desired map $\mathbb{F}(\mathbb{C}) \to \mathfrak{Y}_{\mathbb{C}}$ of $(0,1)$-fibrations over $\mathbb{C}$.

To finish the proof we need to show that for every $c \in \mathbb{C}$ the induced map on fibres

$$\begin{tikzcd}
\mathbb{C}_{\nearrow c} \arrow[rr, "f"] \arrow[dr, "\mathrm{ev}_0"'] & & \mathfrak{Y}_{\mathbb{C}} \times_{\{c\}} \mathbb{C} \arrow[dl] \\
& \mathbb{C} &
\end{tikzcd}$$

is a weak equivalence. We note that by construction the horizontal map above sends the object $i : \Delta^0 \to \mathbb{C}_{\nearrow c}$, represented to the identity morphism to an object in $\mathfrak{Y}_{\mathbb{C}} \times_{\{c\}} \mathbb{C}$ also represented by the identity morphism. We claim that we have a sequence of maps

$$\Delta^0 \xrightarrow{i} \mathbb{C}_{\nearrow c} \xrightarrow{f} \mathfrak{Y}_{\mathbb{C}} \times_{\{c\}} \mathbb{C}$$

such that i and $f \circ i$ are weak equivalences of in the model structure of $(1,0)$-fibrations. The first map is a weak equivalence by [4, Theorem 3.17]. Since $\mathfrak{Y}_{\mathbb{C}} \times_{\{c\}} \mathbb{C} \simeq \mathbb{U}\mathrm{n}^{1,0}_{\mathbb{C}}\big(\mathbb{C}(s(-), s(c))\big)$ (see Definition 6.10), we see that $f \circ i$ is adjoint to the natural transformation $\mathfrak{C}^{\mathbf{sc}}[\mathbb{C}](-, c) \Rightarrow \mathbb{C}(s(-), s(c))$ which is a pointwise weak equivalence. The result follows by two-out-of-three. □

Theorem 6.16 *For every ∞-bicategory $\mathbb{C}$ the Yoneda embedding*

$$\mathcal{Y}_{\mathbb{C}} : \mathbb{C} \to \mathrm{Fun}(\mathbb{C}^{\mathrm{op}}, \mathbb{C}\mathrm{at}_\infty), \quad c \mapsto \mathbb{C}(-, c),$$

is fully faithful. Moreover, given a functor $\mathcal{F} : \mathbb{C}^{\mathrm{op}} \to \mathbb{C}\mathrm{at}_\infty$ there is a equivalence of $\mathbb{C}\mathrm{at}_\infty$-valued functors

$$\mathrm{Nat}_{\mathbb{C}^{\mathrm{op}}}(\mathcal{Y}_{\mathbb{C}}(-), \mathcal{F}) \overset{\simeq}{\Longrightarrow} \mathcal{F},$$

which is natural in $\mathcal{F}$.

Proof It follows from Proposition 6.15 that the embedding $\mathcal{Y}_{\mathbb{C}}$ is fully faithful if and only if the functor J in Definition 6.10 induces a weak equivalence of the corresponding mapping marked simplicial sets. This is equivalent to showing that the composite

$$\mathfrak{C}^{\mathrm{sc}}[\mathbb{C}] \xrightarrow{s} \mathbb{C} \xrightarrow{j} \mathrm{Fun}_{\mathrm{Set}_\Delta^+}(\mathbb{C}^{\mathrm{op}}, \mathrm{Set}_\Delta^+) \xrightarrow{s^*} \mathrm{Fun}_{\mathrm{Set}_\Delta^+}(\mathfrak{C}^{\mathrm{sc}}[\mathbb{C}]^{\mathrm{op}}, \mathrm{Set}_\Delta^+)$$

induces a weak equivalence of the corresponding mapping marked simplicial sets. This follows easily as the functor s is an equivalence of Set_Δ^+-enriched categories, j is the enriched Yoneda embedding (which is fully faithful in the enriched sense) and s^* is also an equivalence of Set_Δ^+-enriched categories by [18, A.3.3.8]. We conclude that $\mathcal{Y}_{\mathbb{C}}$ is fully faithful.

To prove the final claim we show that the functor $\mathbb{I}$ in Proposition 6.7 is naturally equivalent to the identity. We construct a natural transformation $\mathbb{I} \Rightarrow \mathbb{1}$ to the identity functor which is induced by the canonical map $A \to Q(f)$ (see Definition 6.6). We observe that the induced map on fibres

$$\mathrm{Map}_{\mathbb{C}}^{1,0}(\mathbb{C}_{\nearrow c}, \mathcal{F}) \xrightarrow{\simeq} \mathcal{F}(c)$$

is given by restriction along the anodyne map $\Delta^0 \to \mathbb{C}_{\nearrow c}$ selecting the identity morphism and thus is a weak equivalence. □

Acknowledgements

The author would like to thank Rune Haugseng for the helpful discussions regarding the model independent definition of local fibrations. The author is also thankful to the referee for providing suggestions that improved the quality of this article.

References

[1] **F Abellán**, *Comparing lax functors of $(\infty, 2)$-categories* (2023) arXiv 2311.12746

[2] **F Abellán**, **A Gagna**, **R Haugseng**, *Lax transformations, fibrations and mates for $(\infty, 2)$-categories* In preparation

[3] **F Abellán**, **W H Stern**, *2-Cartesian fibrations, II: A Grothendieck construction for ∞-bicategories* (2022) arXiv 2201.09589

[4] **F Abellán**, **W H Stern**, *On cofinal functors of ∞-bicategories* (2023) arXiv 2304.07028

[5] **F Abellán García**, *Marked colimits and higher cofinality*, J. Homotopy Relat. Struct. 17:1 (2022) 1–22 MR

[6] **F Abellán García**, **W H Stern**, *2-Cartesian fibrations, I: A model for ∞-bicategories fibred in ∞-bicategories*, Appl. Categ. Structures 30:6 (2022) 1341–1392 MR

[7] **D Ayala**, **A Mazel-Gee**, **N Rozenblyum**, *Stratified noncommutative geometry* (2019) arXiv 1910.14602

[8] **J D Berman**, *On lax limits in infinity categories* (2020) arXiv 2006.10851

[9] **M Christ**, **T Dyckerhoff**, **T Walde**, *Complexes of stable* ∞*-categories* (2023) arXiv 2301.02606

[10] **A Gagna**, **Y Harpaz**, **E Lanari**, *Fibrations and lax limits of* $(\infty, 2)$*-categories* (2020) arXiv 2012.04537

[11] **A Gagna**, **Y Harpaz**, **E Lanari**, *Gray tensor products and Lax functors of* $(\infty, 2)$*-categories*, Adv. Math. 391 (2021) art. id. 107986 MR

[12] **A Gagna**, **Y Harpaz**, **E Lanari**, *On the equivalence of all models for* $(\infty, 2)$*-categories*, J. Lond. Math. Soc. (2) 106:3 (2022) 1920–1982 MR

[13] **A Gagna**, **Y Harpaz**, **E Lanari**, *Cartesian fibrations of* $(\infty, 2)$*-categories*, Algebr. Geom. Topol. 24:9 (2024) 4731–4778 MR

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

[15] **D Gaitsgory**, **N Rozenblyum**, *A study in derived algebraic geometry, I*: *Correspondences and duality*, Mathematical Surveys and Monographs 221, Amer. Math. Soc., Providence, RI (2017) MR

[16] **V Hinich**, *Yoneda lemma for enriched* ∞*-categories*, Adv. Math. 367 (2020) art. id. 107129 MR

[17] **S Linskens**, **D Nardin**, **L Pol**, *Global homotopy theory via partially lax limits* (2022) arXiv 2206.01556

[18] **J Lurie**, *Higher topos theory*, Annals of Mathematics Studies 170, Princeton Univ. Press (2009) MR

[19] **J Lurie**, $(\infty, 2)$*-categories and the Goodwillie calculus, I* (2009) arXiv 0905.0462

[20] **L Moser**, **N Rasekh**, **M Rovelli**, *An* (∞, n)*-categorical straightening-unstraightening construction* (2023) arXiv 2307.07259

[21] **J Nuiten**, *On straightening for Segal spaces* (2021) arXiv 2108.11431

FERNANDO ABELLÁN fernando.abellan.garcia@uni-hamburg.de
Department of Mathematics, Universität Hamburg, Hamburg, Germany

Received: January 9, 2024 Revised: March 14, 2025

msp

Algebraic & Geometric Topology **26**:5 (2026) 1749–1780
DOI: 10.2140/agt.2026.26.1749
Published: May 29, 2026

Brauer–Wall groups and truncated Picard spectra of K-theory

JONATHAN BEARDSLEY, KIRAN LUECKE AND JACK MORAVA

We compute the first two k-invariants of the Picard spectra of KU and KO by analyzing their Picard groupoids and constructing their unit spectra as global sections of sheaves on the category of manifolds. This allows us to determine the $\mathbb{E}_\infty$-structures of their truncations $\mathrm{Pic}(KU)[0,3]$ and $\mathrm{Pic}(KO)[0,2]$. It follows that these truncated Picard spaces represent the Brauer groups of $\mathbb{Z}/2$-graded algebra bundles of Donovan, Karoubi, Moutuou and Maycock; the Brauer groups of super 2-lines; and the K-theory twists of Freed, Hopkins and Teleman. Our results also imply that these spaces represent twists of **String**- and **Spin**-structures on manifolds and can be used to twist tmf-cohomology. Finally, we are able to identify $\mathrm{pic}(KU)[0,3]$ with a cotruncation of the Anderson dual of the sphere spectrum.

1 Introduction

We study the first two k-invariants of the Picard spectra of KO and KU, or equivalently, the infinite loop space structures on the 2-truncation of $\mathrm{Pic}(KO)$ and the 3-truncation of $\mathrm{Pic}(KU)$, which we denote by $\mathrm{Pic}_0^2(KO)$ and $\mathrm{Pic}_0^3(KU)$, respectively. We will also work with their connected covers which we denote by $\mathrm{Pic}_1^2(KO)$ and $\mathrm{Pic}_1^3(KU)$, respectively. These last two spectra are of course equivalent to truncations of $B\mathrm{GL}_1(KO)$ and $B\mathrm{GL}_1(KU)$, respectively. The homotopy types of $\mathrm{Pic}_0^2(KO)$ and $\mathrm{Pic}_0^3(KU)$ are not particularly interesting. Indeed, they both split as products of Eilenberg–Mac Lane spaces:

$$\mathrm{Pic}_0^3(KU) \simeq \mathbb{Z}/2 \times K(\mathbb{Z}/2,1) \times K(\mathbb{Z},3),$$
$$\mathrm{Pic}_0^2(KO) \simeq \mathbb{Z}/8 \times K(\mathbb{Z}/2,1) \times K(\mathbb{Z}/2,2).$$

What we show below however is that neither the first nor second k-invariants of the associated Picard *spectra* $\mathrm{pic}_0^3(KU)$ and $\mathrm{pic}_0^2(KO)$ are trivial. In other words, none of the above splittings are splittings of infinite loop spaces. The first theorems we prove are the computations of these k-invariants:

Theorem *The first k-invariant of $\mathrm{pic}_0^3(KU)$ is $\mathrm{Sq}^2 : H\mathbb{Z}/2 \to \Sigma^2 H\mathbb{Z}/2$. The second k-invariant can be taken to be either of the generators of $H^4(\mathrm{pic}_0^1(KU);\mathbb{Z}) \cong \mathbb{Z}/4$, both of which restrict to the map $\beta \circ \mathrm{Sq}^2 : \Sigma H\mathbb{Z}/2 \to \Sigma^4 H\mathbb{Z}$ upon taking a connected cover.*

MSC2020: 19L50, 55N15, 55P42.

Theorem *The first k-invariant of* $\mathrm{pic}_0^2(KO)$ *is* $\mathrm{Sq}^2\circ\rho : H\mathbb{Z}/8\to\Sigma^2 H\mathbb{Z}/2$, *where* $\rho : H\mathbb{Z}/8\to H\mathbb{Z}/2$ *is the reduction mod-2 map. The second k-invariant is one of two classes in* $H^3(\mathrm{pic}_0^1(KO);\mathbb{Z}/2)\cong\mathbb{Z}/2\times\mathbb{Z}/2$, *both of which restrict to* Sq^2 *upon taking connected covers.*

The authors have not been able to find these results in the literature although they do seem to be at least partially known to experts (see, e.g., the proof of [23, Proposition 7.14]). From these theorems we can deduce the group structures of $\mathrm{pic}_0^2(KO)^0(X)$ and $\mathrm{pic}_0^3(KU)^0(X)$, which are nontrivial extensions of ordinary singular cohomology groups:

Corollary *There are bijections of sets*

$$\mathrm{pic}_0^2(KO)^0(X)\cong H^0(X;\mathbb{Z}/8)\times H^1(X;\mathbb{Z}/2)\times H^2(X;\mathbb{Z}/2),$$
$$\mathrm{pic}_0^3(KU)^0(X)\cong H^0(X;\mathbb{Z}/2)\times H^1(X;\mathbb{Z}/2)\times H^3(X;\mathbb{Z}).$$

The group laws on these sets, in the same order, are

$$(a,b,c)+(a',b',c')\mapsto(a+a',b+b',c+c'+b\cup b'),$$
$$(a,b,c)+(a',b',c')\mapsto(a+a',b+b',c+c'+\beta(b\cup b')),$$

where β denotes the Bockstein homomorphism.

Knowing these group structures allows us to identify other roles that these spectra play in algebraic topology, mathematical physics, and the theory of C^*-algebras. Almost all of these manifestations are related to twisted K-theory but in the literature they are not often directly related to Picard spectra and spaces, which are the universal receptacles for twists of any highly structured cohomology theory. Indeed, because of the infinite loop space splittings $B\mathrm{GL}_1(KO)\simeq B\mathrm{GL}_1(KO)[0,2]\times B\mathrm{GL}_1(KO)[3,\infty]$ and $B\mathrm{GL}_1(KU)\simeq B\mathrm{GL}_1(KU)[0,3]\times B\mathrm{GL}_1(KU)[4,\infty]$, any map from a connected space into $\mathrm{Pic}_0^2(KO)$ or $\mathrm{Pic}_0^3(KU)$ gives a purely homotopy-theoretic twist of K-theory by either composing with the inclusion or multiplying by -1 and then including. We now detail the various interpretations of $\mathrm{pic}_0^2(KO)$ and $\mathrm{pic}_0^3(KU)$.

The group laws written above imply that the $\mathrm{pic}_0^2(KO)$ and $\mathrm{pic}_0^3(KU)$ cohomology groups of a space are isomorphic to well-known Brauer groups of C^*-algebras over that space:

Theorem *If* $\mathrm{Br}_{\mathbf{O}}^{\infty}(X)$ *and* $\mathrm{Br}_{\mathrm{U}}^{\infty}(X)$ *are the Brauer groups of (possibly infinite-dimensional) real and complex continuous trace graded C^*-algebras with spectrum X (see [43; 44]) then there are isomorphisms*

$$\mathrm{Br}_{\mathbf{O}}^{\infty}(X)\cong\mathrm{pic}_0^2(KO)^0(X),$$
$$\mathrm{Br}_{\mathrm{U}}^{\infty}(X)\cong\mathrm{pic}_0^3(KU)^0(X).$$

It is a corollary of this fact that $\mathrm{pic}_0^2(KO)^0(X)$ and $\mathrm{Tors}(\mathrm{pic}_0^3(KU)^0(X))$ are also isomorphic to the graded Brauer groups of Donovan and Karoubi [10], where Tors denotes taking the torsion subgroup. Elements of both the C^*-algebra Brauer groups and the Brauer groups of Donovan and Karoubi are known to produce twists of K-theory but our isomorphisms allow one to produce such twists as actual maps of spaces $X\to B\mathrm{GL}_1(KO)$ and $X\to B\mathrm{GL}_1(KU)$. In the case that we replace $\mathrm{pic}_0^3(KU)$ with $\Sigma^3 H\mathbb{Z}$, and use *ungraded* C^*-algebras, the agreement of the two notions of twisted K-theory is shown in [26]. Presumably the same is true in our more general case, but we do not prove it in this paper.

Our computations also imply that $B\mathrm{GL}_1(KO)[0,2]$ and $B\mathrm{GL}_1(KU)[0,3]$ are equivalent, as infinite loop spaces, to certain fibers in the Postnikov tower for $\mathbf{BO}$.

Theorem *Let* $\mathbf{BString} \to \mathbf{BSO}$ *and* $\mathbf{BSpin} \to \mathbf{BO}$ *be the connective covers in the Postnikov tower of* $\mathbf{BO}$. *Then there are equivalences of infinite loop spaces*

$$\mathrm{Pic}_1^2(KO) \simeq \mathrm{fib}(\mathbf{BSpin} \to \mathbf{BO}),$$
$$\mathrm{Pic}_1^3(KU) \simeq \mathrm{fib}(\mathbf{BString} \to \mathbf{BSO}).$$

From this it follows that if M is a manifold (resp. oriented manifold) then the set of **Spin**-structures on M (resp. **String**-structures) is a torsor for $[M, \mathrm{Pic}_1^2(KO)] \cong H^1(M;\mathbb{Z}/2) \ltimes H^2(M;\mathbb{Z}/2)$ (resp. $[M, \mathrm{Pic}_1^3(KU)] \cong H^1(M;\mathbb{Z}/2) \ltimes H^3(M;\mathbb{Z})$). This should be compared with the fact that the set of **Spin**-structures on an oriented manifold is a torsor for $H^2(M;\mathbb{Z}/2)$ and the set of **String**-structures on a **Spin**-manifold is a torsor for $H^3(M;\mathbb{Z})$ (see [45, 2.11, 2.16]). It is claimed in [9, §2.1] that **Spin**-structures are a torsor for $H^1(M;\mathbb{Z}/2) \times H^2(M;\mathbb{Z}/2)$ but a proof is not included. Moreover, because there are equivalences of infinite loop spaces $\mathrm{Pic}_1^3(KU) \simeq B\mathrm{GL}_1(KO[0,1])$ and $\mathrm{Pic}_1^3(KU) \simeq B\mathrm{GL}_1(KU[0,2])$ we can twist **Spin**-structures (resp. **String**-structures) by bundles of $KO[0,1]$ (resp. $KU[0,2]$) modules, which we can interpret as real and complex super 2-line bundles, respectively.

Not surprisingly, our work here is also related to the work on twisted K-theory and mathematical physics by Freed, Hopkins, Teleman and others [9; 16; 17; 20; 21].

Theorem *Let* $\mathrm{cAlg}_{\mathbb{R}}^{\times}$ *and* $\mathrm{cAlg}_{\mathbb{C}}^{\times}$ *be the spectra associated to the Picard* 2*-groupoids of invertible topological* $\mathbb{R}$*- and* $\mathbb{C}$*-superalgebras*, *respectively. Then there are equivalences of spectra*

$$\mathrm{cAlg}_{\mathbb{R}}^{\times} \simeq \mathrm{pic}_0^2(KO),$$
$$\mathrm{cAlg}_{\mathbb{C}}^{\times} \simeq \mathrm{pic}_0^3(KU).$$

It follows that $\mathrm{gl}_1(KO)[0,1]$ and $\mathrm{gl}_1(KU)[0,2]$ are the Picard spectra of real and complex superlines, respectively. Again using the fact that $\mathrm{gl}_1(KO)[0,2] \simeq \mathrm{gl}_1(KO[0,2])$ and $\mathrm{gl}_1(KU)[0,3] \simeq \mathrm{gl}_1(KU[0,3])$, we have that a bundle of real (resp. complex) superlines on a space X is the same data as a bundle of $KO[0,2]$-lines (resp. $KU[0,3]$-lines). This also identifies $\mathrm{pic}_0^3(KU)$ with the spectrum R_{-1} of [9] whose associated cohomology theory is proposed as the container for the flux of the oriented superstring B-field.

The above theorem can be restated in the context of the K-theory twists of [16, 1.80; 21, Corollary 2.25] assuming we take X to be only a space rather than a topological groupoid.

Theorem *Let* $\pi_0\mathfrak{Twist}_{KU}(X)$ *and* $\pi_0\mathfrak{Twist}_{KO}(X)$ *be the groups of isomorphism classes of KU and KO twists on* X *in the sense of ibid. Then there are group isomorphisms*

$$\pi_0\mathfrak{Twist}_{KO}(X) \cong \mathrm{pic}_0^2(KO)^0(X),$$
$$\pi_0\mathfrak{Twist}_{KU}(X) \cong \mathrm{pic}_0^3(KU)^0(X).$$

In related work by the same authors, the Anderson dual of the sphere spectrum, $I_{\mathbb{Z}}$, often arises (see, for instance, [18, Hypothesis 5.17; 19, Theorem 5.27]). We show that, at least in the complex case, the truncated Picard spectrum is closely related:

Theorem *There is an equivalence of spectra*

$$\Sigma^3(I_{\mathbb{Z}}[-3,\infty)) \simeq \mathrm{pic}^3_0(KU).$$

This suggests a connection between invertible topological field theories (and deformations thereof) and bundles of truncated $KU[0,2]$-lines.

1.1 Conventions and notation

In this paper we will often work with ∞-categories (in the sense of [36]) of *spectra* and ∞-*groupoids*, which we denote by $\mathcal{S}\mathrm{p}$ and $\mathcal{S}$, respectively. There has been some disagreement lately about an efficacious term for the objects of the ∞-category $\mathcal{S}$. We find the term "∞-groupoid" to be too long, the term "anima" to be unpleasant to pluralize, and the term "space" to be far too ambiguous. Therefore, going forward, we will use "h-type" to refer to these structures and propose this as an alternative to the terms listed above. We will still occasionally call them ∞-groupoids when we want to emphasize their use as ∞-categories in which all morphisms are invertible. When we use the term "space" we will specifically be referring to a compactly generated, weakly Hausdorff topological space, the category of which we will denote by Top. Recall that if $E \in \mathcal{S}\mathrm{p}$ is a spectrum then it has an associated cohomology theory E^*. This cohomology theory is defined equally well on $\mathcal{S}$ and Top, so we will apply it to both sorts of objects without comment.

The term "infinite loop space" frequently denotes what, in the language of [38], one might call "grouplike $\mathbb{E}_\infty$-monoids in $\mathcal{S}$." Another commonly used term for such structures, and the one we will employ, is "abelian ∞-group," recognizing that these objects are the higher algebraic analogues of abelian groups. We emphasize however that being an abelian ∞-group is a *structure* rather than a property. If we wish to refer to an h-type which has an $\mathbb{E}_\infty$-structure without concerning ourselves with whether or not it is grouplike, we will often say "$\mathbb{E}_\infty$-type" instead of the unwieldy "$\mathbb{E}_\infty$-h-type."

Many of our constructions will involve a field $\mathbb{F}$. We will always assume that $\mathbb{F}$ has a topology which is accounted for by these constructions, e.g., vector bundles. We will occasionally wish to consider a field with its discrete topology, in which case we will say so.

We will often be interested in truncations and connective covers of spectra and h-types. For integers $n, m \geq 0$ and a spectrum or h-type X we will write $X[n,m]$ for the n-connective cover of the m-truncation of X. Note the use of connective here, as opposed to *connected*, which differs by 1. In the special, and ubiquitous, case that X is $\mathrm{Pic}(R)$ or $\mathrm{pic}(R)$ for a commutative ring spectrum R, we will use the nonstandard notation of Definition 2.2. We do this to keep the names of these frequently used objects compact.

2 Background

In this section we recall, for an $\mathbb{E}_\infty$-ring spectrum R, the construction of the so-called *Picard space* of R along with various spectra and h-types which can be built from it. More detailed constructions of these can be found in [3; 4].

Definition 2.1 Let R be an $\mathbb{E}_\infty$-ring spectrum with symmetric monoidal ∞-category of left modules LMod_R. Then we make the following definitions.

(1) We write $\mathrm{Pic}(R)$ for the maximal ∞-groupoid in LMod_R spanned by modules which are invertible with respect to the tensor product over R. Recall that $\mathrm{Pic}(R)$ is an abelian ∞-group with base point $R \in \mathrm{Pic}(R)$. We denote its infinite delooping by $\mathrm{pic}(R)$, i.e., $\Omega^\infty \mathrm{pic}(R) \simeq \mathrm{Pic}(R)$.

(2) We write $\mathrm{GL}_1(R)$ for the pullback of the projection $\Omega^\infty R \to \pi_0(R)$ along the inclusion $\pi_0(R)^\times \hookrightarrow \pi_0(R)$. Recall that $\mathrm{GL}_1(R)$ is equivalent, as an abelian ∞-group, to $\Omega\,\mathrm{Pic}(R)$. Equivalently, we could take $B\mathrm{GL}_1(R)$ to be the base point component of $\mathrm{Pic}(R)$. We denote the infinite deloopings of $\mathrm{GL}_1(R)$ and $B\mathrm{GL}_1(R)$ by $\mathrm{gl}_1(R)$ and $\mathrm{bgl}_1(R)$, respectively.

Definition 2.2 Given an $\mathbb{E}_\infty$-ring R, we write $\mathrm{Pic}_n^m(R)$ and $\mathrm{pic}_n^m(R)$ for the m-truncated, n-connective covers of $\mathrm{Pic}(R)$ and $\mathrm{pic}(R)$, respectively. In other words, $\mathrm{Pic}_n^m(R)$ and $\mathrm{pic}_n^m(R)$ are equivalent to $\mathrm{Pic}(R)$ and $\mathrm{pic}(R)$, respectively, in homotopy degrees n through m, and have trivial homotopy groups elsewhere. We will use $\mathrm{bgl}_1(R)[0,n]$ and $B\mathrm{GL}_1(R)[0,n]$ interchangeably with $\mathrm{pic}_1^n(R)$ and $\mathrm{Pic}_1^n(R)$. This notation should not be confused with the notation Pic^0 of algebraic geometry, which denotes the identity component of the Picard scheme.

Remark 2.3 Both m-truncating and taking n-connective covers determine symmetric monoidal functors $\mathcal{S} \to \mathcal{S}$ for all n and m, so $\mathrm{Pic}_n^m(R)$ has an abelian ∞-group structure induced by that of $\mathrm{Pic}(R)$.

It will also be useful to have the following definition recorded here, though we will not make frequent use of it. This definition appears, e.g., in [29].

Definition 2.4 Let R be a commutative ring spectrum. Then $\mathrm{LMod}(R)$ is an $\mathbb{E}_\infty$-algebra in the symmetric monoidal ∞-category of ∞-categories, Cat_∞, which in turn has its own category of modules. We write $\mathrm{Br}(R)$ for the abelian ∞-group of invertible $\mathrm{LMod}(R)$-modules and equivalences between them.

3 Nontriviality of k-invariants

In this section we show that the k-invariants of $\mathrm{pic}_0^1(KO)$, $\mathrm{pic}_0^1(KU)$, $\mathrm{pic}_1^2(KO)$ and $\mathrm{pic}_1^3(KU)$ are nontrivial. This will follow from showing that these spectra can be modeled by Picard groupoids, or sheaves thereof, that have nontrivial symmetries.

3.1 Picard groupoids and symmetries

Recall that there is an equivalence between spectral 1-types, i.e., spectra with homotopy groups only in degrees n and $n+1$ for $n \in \mathbb{Z}$, and symmetric monoidal categories in which every morphism is invertible and every object has a tensor inverse, i.e., Picard groupoids (see [30, Proposition B.12; 32]). In particular, Picard groupoids are equivalent to spectra with nontrivial homotopy groups only in degrees 0 and 1 by [32, Theorem 1.5].

Note that the proof of ibid. proceeds by lifting the usual looping/delooping equivalence between groupoids and homotopy 1-types to their categories of grouplike commutative monoids: Picard groupoids and 1-truncated grouplike infinite loop spaces, respectively. It follows immediately that a Picard groupoid, up to equivalence, is entirely determined by the data of two homotopy groups, π_0 and π_1, and a single k-invariant $\alpha \in \mathrm{Map}(H\pi_0, \Sigma^2 H\pi_1)$. However, Picard groupoids are an especially useful model of stable 1-types since, up to symmetric monoidal equivalence, they can always be rigidified to categories that are *permutative* and *skeletal* [32, Theorem 2.2]. In other words we can always assume, up to symmetric monoidal equivalence, that a Picard groupoid is strictly associative and strictly unital and that its only morphisms are automorphisms.

Let us be more explicit about the data by which a Picard groupoid $\mathcal{P}$ is classified. What follows is mostly a restatement of [23, Remark 2.12] but we also direct the reader to [15, Lecture 17, Section 27]. We assume that $\mathcal{P}$ is permutative and skeletal. It is clear from the preceding paragraph that the group of objects of $\mathcal{P}$ must be π_0 of the associated spectrum. Moreover, one can show that for any $x, y \in \mathcal{P}$ there must be an isomorphism $\mathrm{Aut}_{\mathcal{P}}(x) \cong \mathrm{Aut}_{\mathcal{P}}(y)$. By again invoking the description of stable 1-types as grouplike 1-truncated infinite loop spaces, we see that the associated spectrum must have $\pi_1 \cong \mathrm{Aut}_{\mathcal{P}}(x)$ for any choice of basepoint.

A standard diagram chase shows that the symmetry natural isomorphism of $\mathcal{P}$ must be entirely determined by a choice, for each $x \in \mathcal{P}$, of an element $\varepsilon \in \mathrm{Aut}_{\mathcal{P}}(x)$ such that $\varepsilon_x^2 = \mathrm{id}_x$ (since $\mathcal{P}$ can be assumed to have no isomorphic objects which are not identical). This is equivalent to an element of $\varepsilon \in \mathrm{Hom}_{\mathrm{Ab}}(\pi_0 \otimes \mathbb{Z}/2, \pi_1)$. By assuming permutativity, we have no need to define an associativity natural transformation. Therefore, the Picard groupoid is entirely determined by the data of π_0, π_1 and ε. Using the isomorphism

$$\mathrm{Hom}_{\mathrm{Ab}}(\pi_0 \otimes \mathbb{Z}/2, \pi_1) \cong H^2(\pi_0; \pi_1) \cong [H\pi_0, \Sigma^2 H\pi_1] \tag{1}$$

of [12, Section 27], we see that the symmetry datum ε is exactly the k-invariant of the spectrum. Note that if the k-invariant, equivalently the function ε, is trivial, i.e., if the associated spectrum is equivalent to $H\pi_0 \vee \Sigma H\pi_1$, then the symmetry natural isomorphism of $\mathcal{P}$ is the identity transformation.

The above discussion has a somewhat more modern interpretation. A 1-truncated connective spectrum X is equivalent to the data of a connective module over the 1-truncation of the sphere $\mathbb{S}$, which in turn is completely determined by $\pi_0 X$, $\pi_1 X$, and the action of the Hopf element $\eta \in \pi_1 \mathbb{S}$. This is a map

$$\eta \cdot : \pi_0 X \otimes \mathbb{Z}/2 \to \pi_1 X.$$

On the other hand, X is completely determined by its k-invariant in $H^2(\pi_0; \pi_1)$. Therefore the data of two homotopy groups and a k-invariant can be identified with the data of two homotopy groups and an action of η, which is precisely an element of $\mathrm{Hom}_{\mathrm{Ab}}(\pi_0 \otimes \mathbb{Z}/2, \pi_1)$.

In what follows, we show that the first k-invariants of $\mathrm{pic}_0^1(KO)$ and $\mathrm{pic}_0^1(KU)$ are nontrivial. These are both stable 1-types so it suffices to show that the symmetry isomorphisms of the associated Picard groupoids are nontrivial. In fact, it turns out that we only need to consider the symmetry isomorphism of the sphere spectrum. Note that this data is determined at the level of homotopy categories.

Lemma 3.1 *The k-invariant of the Picard spectrum* $\mathrm{pic}_0^1(\mathbb{S})$ *is nontrivial.*

Proof Let $\sigma : \mathbb{S} \otimes \mathbb{S} \to \mathbb{S} \otimes \mathbb{S}$ be the symmetry map of $\mathbb{S}$ in $\mathcal{S}\mathrm{p}$. This map is the stabilization of the "swap" equivalence $S^1 \wedge S^1 \to S^1 \wedge S^1$ in $\pi_2(S^2)^\times$, which is not homotopic to the identity. □

Lemma 3.2 *The functors* $\mathrm{pic}(\mathbb{S}) \to \mathrm{pic}(KO)$ *and* $\mathrm{pic}(\mathbb{S}) \to \mathrm{pic}(KU)$ *given by tensoring with KO and KU, respectively, equivalently the maps induced by applying the Picard spectrum functor to the units $\mathbb{S} \to KO$ and $\mathbb{S} \to KU$, are surjective on π_0 and isomorphisms on π_1.*

Proof We only prove the statement for KO as the proof for KU is identical. Recall from [23] that the objects of $\mathrm{pic}(KO)$ are precisely the shifts $\Sigma^i KO$ for $0 \le i \le 7$, using the eightfold periodicity of KO. Each of these is covered by $\Sigma^i \mathbb{S}$ for $0 \le i \le 7$ under the given functor $\mathrm{pic}(\mathbb{S}) \to \mathrm{pic}(KO)$, so the functor induces a surjection on π_0.

On π_1 this is precisely π_0 of the induced map $\mathrm{gl}_1(\mathbb{S}) \to \mathrm{gl}_1(KO)$. Recall that the unit map $\mathbb{S} \to KO$ is an isomorphism on π_0, for instance because it is a ring map $\mathbb{Z} \to \mathbb{Z}$. Now using the fact that $\mathrm{gl}_1(R)$ is the infinite delooping of the pullback of $\pi_0(R)^\times \to \pi_0 R \leftarrow \Omega^\infty R$ for any commutative ring spectrum R, we have that the induced map $\mathrm{gl}_1(\mathbb{S}) \to \mathrm{gl}_1(KO)$ is also an isomorphism π_0. □

Corollary 3.3 *The first k-invariants of* $\mathrm{pic}(KO)$ *and* $\mathrm{pic}(KU)$ *are nontrivial.*

3.2 Chain bundle models of topological K-theory

In [8, Section 6] it is shown that the commutative ring spectrum *ku* can be recovered as the homotopification of the sheaf of groupoids of complex vector bundles, suitably group completed. The constructions below are similar, but replace vector bundles with $\mathbb{Z}/2$-*graded chain complexes* of vector bundles. This sheaf is still not concordance invariant, so we must localize it with respect to $\mathbb{R}$ (i.e., make it *concordance invariant*). The resulting localized sheaf represents complex topological K-theory by [48, Appendix I]. Moreover, since it is concordance invariant, the results of [1] imply that it is a constant sheaf determined by its value at the point, which we show to be *ku*. This sheaf then has a subsheaf of *invertible* $\mathbb{Z}/2$-graded chain complexes (with respect to the usual tensor product) which evaluates to $\mathrm{gl}_1(ku) \simeq \mathrm{gl}_1(KU)$ on the point.

These sheaves are slightly more complicated than those of [8] but give us better access to the k-invariants of $\mathrm{gl}_1(KU)$ and therefore the k-invariants of $\mathrm{pic}(KU)$. Of course all of our arguments apply equally well to case of real K-theory. It is worth noting that the sheaf constructed in [8] returns $K(\mathbb{C})$ when evaluated at the point before localizing, whereas ours most certainly does not. We suspect that $K(\mathbb{C})$ however, or $K(\mathbb{F})$ in general for a discrete field $\mathbb{F}$, can be recovered by a variant of our construction using principal $\mathrm{GL}(\mathbb{F})$-bundles instead of vector bundles, i.e., by always equipping $\mathbb{F}$ with the discrete topology (see Conjecture 3.38).

Throughout this section, $\mathrm{Ch}_{\mathbb{F}}^{\mathrm{perf}} : \mathrm{Top}^{\mathrm{op}} \to \mathrm{Gpd}$ will denote the sheaf whose value at a space is the groupoid of bounded chain complexes of finite-dimensional $\mathbb{F}$-vector bundles on X with homotopy classes of homotopy equivalences between them. We assume all the standard structures of this sheaf, e.g., tensor products and direct sums. For our purposes it will be convenient to work with a slightly different sheaf, which we define below.

Definition 3.4 Let $\mathbb{F}$ be a field. We begin by defining chain complexes of $\mathbb{F}$ vector spaces which are graded by $\mathbb{Z}/2$. We call these *differential super* $\mathbb{F}$*-vector spaces* or DSVs for short.

(1) A differential super $\mathbb{F}$-vector space is a $\mathbb{Z}/2$-graded $\mathbb{F}$-vector space $V = V_0 \oplus V_1$ equipped with two maps $d_0 : V_0 \to V_1$ and $d_1 : V_1 \to V_0$ such that $d_0 d_1 = d_1 d_0 = 0$. A morphism of differential super $\mathbb{F}$-vector spaces is a morphism of $\mathbb{Z}/2$-graded vector spaces which commutes with the differential. These data form a category which we denote by $\mathrm{DSV}_{\mathbb{F}}$.

(2) If $V = (V_0 \oplus V_1, d_0, d_1)$ and $W = (W_0 \oplus W_1, d'_0, d'_1)$ are DSVs, we will write $V \otimes W$ for the DSV which is $(V \otimes W)_0 = (V_0 \otimes W_0) \oplus (V_1 \otimes W_1)$ in degree zero and $(V \otimes W)_1 = (V_1 \otimes W_0) \oplus (V_0 \otimes W_1)$ in degree one. The differentials are given in the usual way after reducing all indices modulo 2.

(3) For $V = (V_0 \oplus V_1, d_0, d_1)$ a DSV, we define $H_0(V) = \ker(d_0)/\mathrm{im}(d_1)$ and $H_1(V) = \ker(d_1)/\mathrm{im}(d_0)$. We say that a morphism of DSVs is a quasi-isomorphism if it induces isomorphisms of these two homology groups.

(4) Let $V = (V_0 \oplus V_1, d_0, d_1)$ and $W = (W_0 \oplus W_1, d'_0, d'_1)$ be DSVs. Given two maps $f, g : V \to W$ of DSVs we say that a chain homotopy between f and g is a pair of maps $h_0 : V_0 \to W_1$ and $h_1 : V_1 \to W_0$ such that $f_0 - g_0 = d'_1 h_0 + h_1 d_0$ and $f_1 - g_1 = d'_0 h_1 + h_0 d_1$. If there is a chain homotopy between f and g we write $f \sim g$.

(5) We say that a map $f : V \to W$ of DSVs is a homotopy equivalence if there exists $g : W \to V$ and chain homotopies $f \circ g \sim \mathrm{id}_W$ and $g \circ f \sim \mathrm{id}_V$.

(6) We write $\varepsilon : \mathrm{Ch}_{\mathbb{F}}^{\mathrm{perf}} \to \mathrm{DSV}_{\mathbb{F}}$ for the functor which takes a bounded and finite-dimensional $\mathbb{F}$-chain complex $(E_\bullet, \partial)$ to the DSV whose graded vector space is $E = \left(\bigoplus_i E_{2i}\right) \oplus \left(\bigoplus_i E_{2i+1}\right)$ and whose differential is the obvious restriction of ∂.

We leave it to the reader to check the following lemmas which are standard arguments in homological algebra.

Lemma 3.5 *A map $f : V \to W$ is a quasi-isomorphism of DSVs if and only if it is a homotopy equivalence.*

Lemma 3.6 *The tensor product of DSVs defines a symmetric monoidal structure on* $\mathrm{DSV}_{\mathbb{F}}$ *with monoidal unit* $(\mathbb{F} \oplus 0, 0, 0)$. *This tensor product distributes over the direct sum of DSVs.*

Lemma 3.7 *The functor ε is symmetric monoidal. Moreover, it takes quasi-isomorphisms to quasi-isomorphisms and chain homotopies to chain homotopies.*

Remark 3.8 As with the classical case, one direction of Lemma 3.5 depends on $\mathbb{F}$ being a field. For a general ring, homotopy equivalences are strictly stronger than quasi-isomorphisms.

Definition 3.9 Let X be a space and $\mathbb{F}$ a field. We write $\mathrm{DSV}_{\mathbb{F}}(X)$ for the groupoid of bundles of DSVs on X whose morphisms are chain homotopy classes of homotopy equivalences. Because bundles can be pulled back along maps of spaces this defines a functor $\mathrm{DSV}_{\mathbb{F}} : \mathrm{Top}^{\mathrm{op}} \to \mathrm{Gpd}$. The tensor product and direct sum bundles are defined in the usual way.

Going forward, many of our results will apply equally well to the functors $\mathrm{DSV}_{\mathbb{F}}$ and $\mathrm{Ch}_{\mathbb{F}}^{\mathrm{perf}}$. We will therefore write $\mathscr{G}_{\mathbb{F}}$ to denote either.

The following lemma says that $\mathscr{G}_{\mathbb{F}}$ is a sheaf of *ring groupoids* in the sense of [11], which are a special case of the ring categories of [35]. These are essentially commutative monoids in the category of Picard groupoids. Without loss of generality ring groupoids can always be assumed to have underlying *strict* Picard groupoid.

Lemma 3.10 *The functor $\mathscr{G}_{\mathbb{F}}$ is a presheaf of symmetric monoidal groupoids with respect to direct sum of chain bundles, a presheaf of symmetric monoidal groupoids with respect to tensor product of chain bundles, and the latter structure distributes over the former. Moreover, the natural transformation $\varepsilon(X):\mathrm{Ch}_{\mathbb{F}}^{\mathrm{perf}}(X)\to\mathrm{DSV}_{\mathbb{F}}(X)$ preserves this structure.*

Proof The lemma follows from standard arguments for bundles extended suitably to chain complexes. □

Remark 3.11 It will be convenient to extend the codomain and restrict the domain of $\mathscr{G}_{\mathbb{F}}$. By taking nerves there is an inclusion $\mathrm{Gpd}\subset\mathcal{S}$ under which the symmetric monoidal structure of Lemma 3.10 makes $L_{\mathbb{R}}\mathscr{G}_{\mathbb{F}}$ into a presheaf of $\mathbb{E}_\infty$-types. For convenience, we will also restrict $\mathscr{G}_{\mathbb{F}}$ to the subcategory of smooth manifolds and smooth maps $\mathrm{Mfd}^{\mathrm{op}}\subset\mathrm{Top}^{\mathrm{op}}$.

Lemma 3.12 *The functor $\mathscr{G}_{\mathbb{F}}$ is a sheaf of symmetric monoidal groupoids on* Mfd *with respect to coverings by families of open embeddings.*

Proof To prove that $\mathscr{G}_{\mathbb{F}}$ is a sheaf, it suffices to show that, for a fixed manifold M, it satisfies descent on the "little" site of open submanifolds of M (indeed by [1, Lemma 3.5.3] it suffices to check only on Euclidean spaces). This follows immediately from the definitions, as bundles themselves are defined locally. Further, $\mathscr{G}_{\mathbb{F}}$ is valued in *symmetric monoidal* ∞-groupoids (i.e., $\mathbb{E}_\infty$-types) as a result of the fact that the inclusion $\mathrm{Gpd}\subset\mathcal{S}$ is symmetric monoidal (with respect to the Cartesian product). □

We recall the $\mathbb{R}$-invariantization functor from Definition 4.2.5 and Proposition 5.1.2 of [1].

Definition 3.13 Let Δ^n_{sm} denote the hyperplane in $\mathbb{R}^{n+1}$ spanned by points $(x_1,x_2,\dots,x_{n+1})$ such that $\sum_{i=1}^{n+1}x_i=1$, also known as the *smooth n-simplex.* Let $\Delta^\bullet_{\mathrm{sm}}$ denote the cosimplicial manifold which is Δ^n_{sm} is degree n. Its coface maps $\Delta^n_{\mathrm{sm}}\to\Delta^{n+1}_{\mathrm{sm}}$ are the $n+2$ inclusions $\Delta^n_{\mathrm{sm}}\to\Delta^{n+1}_{\mathrm{sm}}$ given by the $n+2$ possible intersections of $\Delta^{n+1}_{\mathrm{sm}}$ with the coordinate hyperplanes. The codegeneracies $\Delta^{n+1}_{\mathrm{sm}}\to\Delta^n_{\mathrm{sm}}$ are given by the n possible ways of adding adjacent coordinates.

Definition 3.14 Let $\mathscr{E}:\mathrm{Mfd}^{\mathrm{op}}\to\mathrm{Top}$ be a presheaf of h-types. Then define $L_{\mathbb{R}}\mathscr{E}:\mathrm{Mfd}^{\mathrm{op}}\to\mathrm{Top}$ to be the presheaf of h-types whose value at X is given by $L_{\mathbb{R}}\mathscr{E}(X)=|\mathscr{E}(X\times\Delta^\bullet_{\mathrm{alg}})|$.

Remark 3.15 If a and a' are objects of $\mathscr{G}_{\mathbb{F}}(X)$ (thought of as a Picard groupoid), b is an object of $\mathscr{G}_{\mathbb{F}}(X\times\Delta^1)$, and we have isomorphisms $d_0(b)\cong a$ and $d_1(b)\cong a'$, where d_0 and d_1 are the face maps, then a and a' are equivalent in $L_{\mathbb{R}}\mathscr{G}_{\mathbb{F}}(X)$. Similar statements hold for the $\mathscr{G}_{\mathbb{F}}(X\times\Delta^n)$ and the relevant higher face maps. This has the effect of making $\mathscr{G}_{\mathbb{F}}$ insensitive to the difference between a space X and its "stabilizations" $X\times\mathbb{R}\simeq X\times\Delta^1$, $X\times\mathbb{R}^2\simeq X\times\Delta^2$, and so forth.

Remark 3.16 In the Appendix of [48], working with $\mathrm{Ch}_{\mathbb{C}}^{\mathrm{perf}}$, Segal uses an equivalence relation to identify bundles which can be homotoped into one another along $X \times \mathbb{R} \cong X \times \Delta^1$. The above construction can be thought of actually inserting the homotopies (and homotopies between homotopies etc.) between such bundles, forcing the sheaf to be insensitive to deformations of X. However, after taking connected components the resulting group is the same as Segal's (see Proposition 3.18 below).

Lemma 3.17 *The presheaf of* $\mathbb{E}_\infty$*-types* $L_{\mathbb{R}}\mathscr{G}_{\mathbb{F}}$ *is a sheaf of grouplike* $\mathbb{E}_\infty$*-types, i.e., connective spectra.*

Proof It is a general fact that $L_{\mathbb{R}}$ preserves sheaves (see [1, Remark 4.2.6; 8, Proposition 2.6]). The symmetric monoidal structure is now *grouplike* because for any DSVs or chain complex $E_\bullet \to X$ we can always find $E'_\bullet \to X$ such that $E_\bullet \oplus E'_\bullet$ is concordant to an acyclic: that is, there exists an $E''_\bullet \to X \times [0,1]$ whose restriction to $X \times \{0\}$ is isomorphic to $E_\bullet \oplus E'_\bullet$ and whose restriction to $X \times \{1\}$ is acyclic. This is achieved by using the $[0,1]$ coordinate to turn on differentials killing any nontrivial cohomology. □

Proposition 3.18 *There is a group isomorphism* $\pi_0 L_{\mathbb{R}}\mathrm{DSV}_{\mathbb{C}}(X) \cong KU^0(X)$ *which is natural in* X.

Proof By 3.2.3.1 and 1.4.3.9 of [38] the simplicial colimits defining $L_{\mathbb{R}}\mathrm{DSV}_{\mathbb{C}}(X)$ and $L_{\mathbb{R}}\mathrm{Ch}_{\mathbb{F}}^{\mathrm{perf}}$ can be taken in spaces. Now, π_0 of a simplicial colimit depends only on the subdiagram (which is a coequalizer) involving the 0- and 1-simplices. From the definition of $L_{\mathbb{R}}\mathrm{Ch}_{\mathbb{C}}^{\mathrm{perf}}(X)$, π_0 of that coequalizer is the set of chain-bundles on X up to concordance. Remark 3.16 therefore implies that [48, Proposition A.I] (in which "concordant" is called "homotopic") gives an isomorphism $\pi_0 L_{\mathbb{R}}\mathrm{Ch}_{\mathbb{C}}^{\mathrm{perf}}(X) \xrightarrow{\cong} KU^0(X)$.

This isomorphism is given by taking a bundle to its Euler characteristic. Specifically, a bundle of chain complexes $E_\bullet$ is taken to the alternating sum of the K-theory classes of each grade, $\Sigma_{i\in\mathbb{Z}}(-1)^i[E^i]$. This clearly factors through $\pi_0(\varepsilon(X))$ so that we have a composite isomorphism

$$\pi_0 L_{\mathbb{R}}\mathrm{Ch}_{\mathbb{C}}^{\mathrm{perf}}(X) \xrightarrow{\varepsilon} \pi_0 L_{\mathbb{R}}\mathrm{DSV}_{\mathbb{C}}(X) \to KU^0(X)$$

in which the last morphism forgets the differential. This implies that ε is injective. We have already seen that it is surjective, so it is an isomorphism. □

Theorem 3.19 *There is an equivalence of commutative ring spectra* $L_{\mathbb{R}}\mathrm{DSV}_{\mathbb{C}}(*) \simeq ku$.

Proof This follows from [1, Propositions 4.3.1 and I.2] and Proposition 3.18. Specifically, we know that evaluation at a point is an equivalence between concordance invariant sheaves of spectra on Mfd and spectra. The inverse of this equivalence is given by taking the constant sheaf. Moreover, given a spectrum E, the value of $\mathrm{const}(E)$ at a manifold X is $\mathrm{Map}(\Sigma^\infty_+ X, E)$. Hence $L_{\mathbb{R}}\mathrm{DSV}_{\mathbb{C}}(X) \simeq \mathrm{Map}(\Sigma^\infty_+ X, L_{\mathbb{R}}\mathrm{DSV}_{\mathbb{C}}(*))$ and thus $\pi_0(X, L_{\mathbb{R}}\mathrm{DSV}_{\mathbb{C}}(*)) \cong KU^0(X)$ for every manifold X. So $L_{\mathbb{R}}\mathrm{DSV}_{\mathbb{C}}(*) \simeq ku$ by Brown representability. □

The above arguments, along with those of [48], can be repeated mutatis mutandis with chain complexes of $\mathbb{R}$-vector spaces rather than $\mathbb{C}$-vector spaces. This leads to:

Theorem 3.20 *There is an equivalence of commutative ring spectra* $L_{\mathbb{R}}\mathrm{DSV}_{\mathbb{R}}(*) \simeq ko$.

Corollary 3.21 *Consider the full symmetric monoidal subgroupoid* $\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}(X)$ *of* $\mathrm{DSV}_{\mathbb{C}}(X)$ *on the objects which are invertible in the tensor product monoidal structure. Then* $\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}$ *is a sheaf of connective spectra and* $L_{\mathbb{R}}\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}(*) \simeq \mathrm{gl}_1(ku) \simeq \mathrm{gl}_1(KU)$.

Proof The subgroupoid $\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}(X)$ is the full subgroupoid of $\mathrm{DSV}_{\mathbb{C}}(X)$ on bundles $E_\bullet$ such that $\dim(E_0) - \dim(E_1) = \pm 1$. The same argument for Lemma 3.12 shows that $\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}$ is a sheaf of connective spectra (but not ring spectra). Let $\mathbb{Z}(-)$ denote the sheaf whose value at X is the discrete groupoid of continuous $\mathbb{Z}$-valued functions on X. Similarly let $\mathbb{Z}/2(X)$ be the discrete groupoid of continuous ± 1-valued functions on X. The inclusion $i_X : \mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}(X) \hookrightarrow \mathrm{DSV}_{\mathbb{C}}(X)$ fits (essentially by definition) into a pullback square of h-types

$$\begin{tikzcd}
\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}(X) \arrow[r, "i_X"] \arrow[d] & \mathrm{DSV}_{\mathbb{C}}(X) \arrow[d] \\
\mathbb{Z}/2(X) \arrow[r, hook] & \mathbb{Z}(X)
\end{tikzcd}$$

where the vertical maps send a DSV to its graded dimension. This extends to a levelwise pullback square of simplicial h-types

$$\begin{tikzcd}
\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}(X \times \Delta^\bullet_{\mathrm{alg}}) \arrow[r, "i_X^\bullet"] \arrow[d] & \mathrm{DSV}_{\mathbb{C}}(X \times \Delta^\bullet_{\mathrm{alg}}) \arrow[d] \\
\mathbb{Z}/2(X \times \Delta^\bullet_{\mathrm{alg}}) \arrow[r, hook] & \mathbb{Z}(X \times \Delta^\bullet_{\mathrm{alg}})
\end{tikzcd}$$

Forgetting the monoidal structure and considering this as just a diagram of h-types, the diagram is still a pullback after applying $L_{\mathbb{R}}$ (see [46, Definition 1.1, Proposition 5.4]). When $X = *$, this gives the pullback diagram of h-types

$$\begin{tikzcd}
\Omega^\infty L_{\mathbb{R}}\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}(*) \arrow[r] \arrow[d] & \Omega^\infty L_{\mathbb{R}}\mathrm{DSV}_{\mathbb{C}}(*) \simeq \mathbb{Z} \times BU \arrow[d] \\
\mathbb{Z}/2 \arrow[r, hook] & \mathbb{Z}
\end{tikzcd}$$

exhibiting $\Omega^\infty L_{\mathbb{R}}\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}(*)$ as $\mathrm{GL}_1\,ku$. Since i_X is symmetric monoidal, there is a natural-in-Y isomorphism of abelian groups

$$L_{\mathbb{R}}\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}(*)^0(Y) \cong \mathrm{Map}(\Sigma^\infty_+ Y, L_{\mathbb{R}}\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}(*)) \cong ku^0(Y)^\times$$

for any space Y (where the right side has the tensor product abelian group structure). Hence by Brown representability there is an equivalence of spectra $L_{\mathbb{R}}\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}(*) \simeq \mathrm{gl}_1(ku)$. □

The same arguments apply to the real case, which gives Corollary 3.22. Our arguments also seem likely to apply in the case that $\mathbb{F}$ is a discrete field which we codify in Conjecture 3.38.

Corollary 3.22 *Consider the full symmetric monoidal subgroupoid* $\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{R}}(X)$ *of* $\mathrm{DSV}_{\mathbb{R}}(X)$ *on the objects which are invertible in the tensor product monoidal structure. Then* $\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{R}}$ *is a sheaf of connective spectra and* $L_{\mathbb{R}}\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{R}}(*) \simeq \mathrm{gl}_1(ko) \simeq \mathrm{gl}_1(KO)$.

3.3 The groupoid of $\mathbb{Z}/2$-graded line bundles

We now define a simpler sheaf that will map to $\mathrm{DSV}_{\mathbb{F}}$ and help us to understand its structure. The sheaves $L_{\mathbb{R}}\mathscr{L}^{\mathbb{C}}$ and $L_{\mathbb{R}}\mathscr{L}^{\mathbb{R}}$ that we describe here will end up being truncations of $L_{\mathbb{R}}\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}} \simeq \mathrm{gl}_1(KU)$ and $L_{\mathbb{R}}\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{R}} \simeq \mathrm{gl}_1(KO)$ after evaluating at the point. Some of the ideas of this section exist in [16] but we go a step further in relating these structures to the unit spectra of K-theory. This is in contrast to [16, 1.45] in which they are described as truncations of *ko* itself. Our results are arguably more conceptually satisfying given that these structures are in fact used for twisting both real and complex K-theory. Indeed, in [16], Freed remarks that he does not have a conceptual reason for the appearance of *ko* in his constructions.

Definition 3.23 Let X be a space, and $\mathbb{F}$ a (topological) field. Let $\mathscr{L}^{\mathbb{F}}(X)$ be the groupoid of pairs (ξ, n) where $\xi : E \to X$ is an $\mathbb{F}$-line bundle and $n : X \to \mathbb{Z}/2 = \{0, 1\}$ is a continuous function. The morphisms between (ξ, m) and (ξ', m') in $\mathscr{L}^{\mathbb{F}}(X)$ will be the empty set if $n \neq m$ and the set of bundle isomorphisms otherwise. We will refer to $\mathscr{L}^{\mathbb{F}}(X)$ as the groupoid of $\mathbb{F}$-superline bundles on X.

Definition 3.24 Given two pairs $(\xi, n), (\xi', m) \in \mathscr{L}^{\mathbb{F}}(X)$, we define a symmetric monoidal structure on $\mathscr{L}^{\mathbb{F}}(X)$ by declaring that:

(1) The tensor product $(\xi, n) \otimes (\xi', m)$ is the object $(\xi \otimes \xi', n + m)$, where the tensor product of the left-hand coordinate is the standard tensor product of principal bundles.

(2) The symmetry isomorphism $(\xi, n) \otimes (\xi', m) \to (\xi', m) \otimes (\xi, n)$ is given on the fiber over x by $(v, w) \mapsto (w, (-1)^{n(x)m(x)}v)$.

Remark 3.25 The above definition is almost identical to the sheaf of graded $\mathbb{T}$-bundles $\mathcal{B}\mathbb{T}^{\pm}$ in [21, Definition 2.1]. Their definition differs from ours only in their definition of the symmetry isomorphism which would be written in our notation as $(v, w) \mapsto (w, vn(x)m(x))$. This makes sense if we take $\mathbb{Z}/2 = \{-1, 1\}$, but then the second coordinate in the tensor product formula, i.e., $n + m$, does not make sense.

Lemma 3.26 *The groupoid* $\mathscr{L}^{\mathbb{F}}(X)$ *does not decompose as a product of symmetric monoidal groupoids.*

Proof Let $\mathbb{Z}/2(X)$ be the discrete symmetric monoidal groupoid of $\mathbb{Z}/2$-valued functions with the pointwise group structure. The sign in the second component of the symmetry isomorphism of $\mathscr{L}^{\mathbb{F}}(X)$ prevents the natural projection $\mathscr{L}^{\mathbb{F}}(X) \to \mathbb{Z}/2(X)$ from having a symmetric monoidal section. □

Remark 3.27 The argument in the proof of Lemma 3.26, along with [32], also shows that the connective spectrum associated to $\mathscr{L}^{\mathbb{F}}(X)$ has nontrivial k-invariant, though we will not need this fact.

Remark 3.28 Definition 3.23 extends to a presheaf of grouplike symmetric monoidal groupoids on Top whose domain, following Remark 3.11, we restrict to Mfd and whose codomain we extend to $\mathcal{S}\mathrm{p}_{\geq 0}$.

Lemma 3.29 *The presheaf $\mathscr{L}^{\mathbb{F}}$ of connective spectra on* Mfd *is a sheaf.*

Proof Because the inclusion $\mathrm{Gpd} \subset \mathcal{S}$ is symmetric monoidal (and the category of connective spectra is equivalent to the category of grouplike $\mathbb{E}_\infty$ h-types), it suffices to show that $\mathscr{L}^{\mathbb{F}}$ is a sheaf of grouplike symmetric monoidal groupoids. The fact that it is a sheaf of groupoids, without symmetric monoidal structure, follows immediately from the fact that it decomposes as a sum of presheaves of groupoids which are clearly sheaves. The symmetric monoidal structure glues as well since limits of symmetric monoidal groupoids are computed in Cat. The grouplike condition is certainly satisfied for each $X \in \mathrm{Mfd}$. □

Proposition 3.30 *With respect to the symmetric monoidal structures given by bundle tensor product, there is a natural-in-X symmetric monoidal functor $\ell(X) : \mathscr{L}^{\mathbb{F}}(X) \to \mathrm{DSV}_{\mathbb{F}}(X)$ defined by taking (ξ, n) to the chain bundle which has line bundle ξ concentrated in degree n.*

Proof It suffices to check for trivial bundles on a path connected space. One checks readily that $\ell(X)((\xi, n) \otimes (\xi', m)) \cong \ell(X)(\xi, n) \otimes \ell(X)(\xi', m)$. If $n = m$, so $n + m = 0$, then we obtain the tensor product line bundle $\xi \otimes \xi'$ in degree 0 and the 0-line bundle in degree 1, with the zero differential between them. If $n \neq m$ then we have the reverse situation. The definition of the morphisms in $\mathscr{L}^{\mathbb{F}}(X)$ makes it clear that their tensor product is similarly preserved by ℓ. □

Theorem 3.31 *When $\mathbb{F}$ equals $\mathbb{C}$ the connective spectrum $L_{\mathbb{R}}\mathscr{L}^{\mathbb{C}}(*)$ fits into a cofiber sequence $\Sigma^2 H\mathbb{Z} \to L_{\mathbb{R}}\mathscr{L}^{\mathbb{C}}(*) \to H\mathbb{Z}/2$ and its k-invariant is nontrivial.*

Proof Consider the simplicial colimit of Definition 3.14 used to define $L_{\mathbb{R}}\mathscr{L}^{\mathbb{C}}(*)$. The connective spectra $\mathscr{L}^{\mathbb{C}}(\Delta^k)$ appearing in this colimit have exactly two nonzero homotopy groups, i.e., they are stable 2-types in the sense of [32] (see the discussion preceding Lemma 3.1). These homotopy groups are determined by the group of isomorphism classes of objects of $\mathscr{L}^{\mathbb{C}}(\Delta^k)$ and by the automorphisms of any one of those objects. In this case we have that $\pi_0(\mathscr{L}^{\mathbb{C}}(\Delta^k)) \cong \mathbb{Z}/2$ since Δ^k is connected and $\pi_1(\mathscr{L}^{\mathbb{C}}(\Delta^k)) = \mathrm{Top}(\Delta^k, \mathbb{C}^\times)$ (since that is the group of automorphisms of the trivial bundle on Δ^k), where $\mathbb{C}^\times$ is considered as a topological group with the usual subspace topology inherited from $\mathbb{C}$.

Again as a result of the analysis of [32], we get that the k-invariant is determined by the element of $H^2(H\mathbb{Z}/2; \mathrm{Top}(\Delta^k, \mathbb{C}^\times)) \simeq \mathrm{Hom}_{\mathrm{Ab}}(\mathbb{Z}/2, \mathrm{Top}(\Delta^k, \mathbb{C}^\times))$ (see (1) and the discussion that follows) corresponding to the function that assigns to an element in each isomorphism class the symmetry isomorphism of its tensor square. In this case this is the function that sends $n \in \mathbb{Z}/2$ to the constant function on Δ^k with value $(-1)^{n^2}$.

The ring $\mathbb{Z}/2$ can be thought of as a Picard groupoid with $\pi_0 \cong \mathbb{Z}/2$ and $\pi_1 = 0$. Therefore there is a forgetful functor of Picard groupoids $(\xi, n) \mapsto n$, corresponding to a map of spectra $\mathscr{L}^{\mathbb{C}}(\Delta^k) \to H\mathbb{Z}/2$ whose fiber is the Eilenberg–Mac Lane spectrum of $\pi_1(\mathscr{L}^{\mathbb{C}}(\Delta^k))$. Therefore there is a cofiber sequence of simplicial spectra

$$\Sigma H\mathrm{Top}(\Delta^\bullet, \mathbb{C}^\times) \to \mathscr{L}^{\mathbb{C}}(\Delta^\bullet) \to H\mathbb{Z}/2 \to \Sigma^2 H\mathrm{Top}(\Delta^\bullet, \mathbb{C}^\times)$$

in which the third term is a constant simplicial diagram and the last map is (the amalgam of) the k-invariants just described.

Note that the leftmost term is equivalent to $\Sigma H\mathrm{Sing}_\bullet(\mathbb{C}^\times)$, the levelwise suspension of the Eilenberg–Mac Lane spectrum of the singular simplicial group of $\mathbb{C}^\times$. Because Ω^∞ preserves geometric realizations we can compute the homotopy groups of this colimit in $\mathcal{S}$, which are trivial except in π_1, where they are $\mathbb{C}^\times$ (with the discrete topology). This implies that the colimit of the above cofiber sequence is the cofiber sequence $\Sigma H\mathbb{C}^\times \to L_{\mathbb{R}}\,\mathscr{L}^{\mathbb{F}}(*) \to H\mathbb{Z}/2 \to \Sigma^2 H\mathbb{C}^\times$ with the last map being the element of $H^2(H\mathbb{Z}/2;\mathbb{C}^\times) = \mathrm{Hom}_{\mathrm{Ab}}(\mathbb{Z}/2,\mathbb{C}^\times)$ sending n to $(-1)^{n^2} \in \mathbb{C}^\times$. Now $\mathbb{C}^\times \simeq B\mathbb{Z}$ so $H\mathbb{C}^\times \simeq \Sigma H\mathbb{Z}$. Thus the preceding cofiber sequence can be rewritten as

$$\Sigma^2 H\mathbb{Z} \to L_{\mathbb{R}}\,\mathscr{L}^{\mathbb{C}}(*) \to H\mathbb{Z}/2 \to \Sigma^3 H\mathbb{Z}$$

and the k-invariant is still nontrivial of course (and therefore represented by $\beta\,\mathrm{Sq}^2$). □

Remark 3.32 Although $\mathscr{L}^{\mathbb{C}}(\Delta^k)$ has homotopy groups concentrated in degrees 0 and 1 for all k, its localization $L_{\mathbb{R}}\,\mathscr{L}^{\mathbb{C}}(*)$ has homotopy groups concentrated in degrees 0 and 2.

Corollary 3.33 *When $\mathbb{F}$ equals $\mathbb{C}$ the spectrum $\mathrm{gl}_1(KU)$ splits as*

$$L_{\mathbb{R}}\,\mathscr{L}^{\mathbb{C}}(*) \oplus \mathrm{gl}_1(KU)[3,\infty).$$

Proof After applying $L_{\mathbb{R}}$, the functor of Proposition 3.30 induces a morphism of sheaves of commutative connective ring spectra. By Corollary 3.21, we have a map of spectra

$$L_{\mathbb{R}}\mathscr{L}^{\mathbb{C}}(*) \to L_{\mathbb{R}}\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}(*) \simeq \mathrm{gl}_1(KU)$$

after evaluating at the point. This map is an isomorphism on π_0 because π_0 of these spectra can be computed by looking only at the bottom two levels of the simplicial diagram defining $L_{\mathbb{R}}$ (the coequalizer diagrams). It is of course an isomorphism on π_1 because both spectra have trivial homotopy in degree 1. Even further, it is an isomorphism on π_2. To see this, first note that whenever $\mathcal{F}$ is a concordance invariant sheaf there is an equivalence (as in the proof of Theorem 3.19) $\mathcal{F}(X) \simeq \mathrm{Map}(\Sigma^\infty_+ X, \mathcal{F}(*))$. Therefore the claimed isomorphism on π_2 would follow from an isomorphism $\pi_0 L_{\mathbb{R}}\,\mathscr{L}^{\mathbb{C}}(S^2) \simeq \pi_0 L_{\mathbb{R}}\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}(S^2)$. Since the inclusion of spaces induces an isomorphism $\pi_0 L_{\mathbb{R}}\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{C}}(S^2) \to \pi_0 L_{\mathbb{R}}\mathrm{DSV}_{\mathbb{C}}(S^2)$ and the latter group is $ku^0(S^2)$, the desired claim follows from the fact that both generators of $ku^0(S^2)$ are represented by line bundles and are therefore in the image of $L_{\mathbb{R}}\,\mathscr{L}^{\mathbb{C}}$. Therefore the composite

$$L_{\mathbb{R}}\,\mathscr{L}^{\mathbb{C}}(*) \to \mathrm{gl}_1(KU) \to \mathrm{gl}_1(KU)[0,2]$$

is an equivalence, and the lemma follows. □

Corollary 3.34 *The first k-invariant of $\mathrm{gl}_1(KU)$ is nontrivial.*

Similar arguments prove the analogous statements for $L_{\mathbb{R}}\mathscr{L}^{\mathbb{R}}$:

Theorem 3.35 *When $\mathbb{F}$ equals $\mathbb{R}$ the connective spectrum $L_{\mathbb{R}}\,\mathscr{L}^{\mathbb{R}}(*)$ fits into a cofiber sequence $\Sigma H\mathbb{Z}/2 \to L_{\mathbb{R}}\,\mathscr{L}^{\mathbb{R}}(*) \to H\mathbb{Z}/2$ and its k-invariant is nontrivial.*

Corollary 3.36 *When $\mathbb{F}$ equals $\mathbb{R}$ the spectrum $\mathrm{gl}_1(KO)$ splits as*

$$L_{\mathbb{R}}\mathscr{L}^{\mathbb{R}}(*)\oplus\mathrm{gl}_1(KO)[2,\infty).$$

Corollary 3.37 *The first k-invariant of $\mathrm{gl}_1(KO)$ is nontrivial.*

Conjecture 3.38 *Let $\mathbb{F}$ be a discrete field. Then there are equivalences $L_{\mathbb{R}}\mathrm{DSV}_{\mathbb{F}}(*)\simeq K(\mathbb{F})$ and $L_{\mathbb{R}}\mathrm{gl}_1\,\mathrm{DSV}_{\mathbb{F}}(*)\simeq\mathrm{gl}_1(K(\mathbb{F}))$. Moreover, there is a splitting $L_{\mathbb{R}}\mathscr{L}^{\mathbb{F}}(*)$ is a split summand of $\mathrm{gl}_1\,K(\mathbb{F})$ and the first and second k-invariants of $\mathrm{pic}(K(\mathbb{F}))$ are nontrivial.*

4 Computations of k-invariants

In this section we compute the possible $\mathbb{E}_\infty$-structures on h-types with the same homotopy groups as $\mathrm{Pic}_0^1(KO)$, $\mathrm{Pic}_0^1(KU)$, $\mathrm{Pic}_1^2(KO)$, and $\mathrm{Pic}_1^3(KU)$. We do this by computing the possible k-invariants of spectra with the same homotopy groups. It will be useful to recall that $\mathrm{Pic}_1^2(KO)$ and $\mathrm{Pic}_1^3(KU)$ are equivalent to $B\mathrm{GL}_1(KO)[0,2]$ and $B\mathrm{GL}_1(KU)[0,3]$, respectively. In most cases we can explicitly name these k-invariants in terms of cohomology operations.

Proposition 4.1 *There are equivalences of h-types*

$$\mathrm{Pic}_0^2(KO)\simeq\mathbb{Z}/8\times K(\mathbb{Z}/2,1)\times K(\mathbb{Z}/2,2)$$

and

$$\mathrm{Pic}_0^3(KU)\simeq\mathbb{Z}/2\times K(\mathbb{Z}/2,1)\times K(\mathbb{Z},3).$$

Proof The k-invariant connecting π_0 to π_1 must be zero, since on each connected component the relevant cohomology group is clearly zero. From [41, Lemma 3.1] we have that the 1-component of $\mathrm{Pic}(KO)$ splits as $K(\mathbb{Z}/2,1)\times\mathbf{BSO}$. The fact that $\mathbf{BSO}[0,2]\simeq K(\mathbb{Z}/2,2)$ completes the proof. The case of $\mathrm{Pic}(KU)$ is essentially identical (the result of [41] applies to both KO and KU). □

4.1 First k-invariants

Now we determine the possible first k-invariants of the associated spectra $\mathrm{pic}_0^2(KO)$ and $\mathrm{pic}_0^3(KU)$.

Lemma 4.2 *The first k-invariant of $\mathrm{pic}(KO)$ is either trivial or $\mathrm{Sq}^2\circ\rho: H\mathbb{Z}/8\to\Sigma^2H\mathbb{Z}/2$, where $\rho: H\mathbb{Z}/8\to H\mathbb{Z}/2$ is the reduction mod-2 map, and the first k-invariant of $\mathrm{pic}(KU)$ is either trivial or $\mathrm{Sq}^2: H\mathbb{Z}/2\to\Sigma^2\mathbb{Z}/2$.*

Proof The case for KU is immediate because $H^2(H\mathbb{Z}/2;\mathbb{Z}/2)\cong\mathbb{Z}/2$. The case of KO follows from the fact that $H^2(H\mathbb{Z}/8;\mathbb{Z}/2)\cong\mathbb{Z}/2$, which follows, e.g., from the argument given after (1). □

Corollary 4.3 *The k-invariants of $\mathrm{pic}_0^1(KO)$ and $\mathrm{pic}_0^1(KU)$ are $\mathrm{Sq}^2\circ\rho$ and Sq^2, respectively, where $\rho:\mathbb{Z}/8\to\mathbb{Z}/2$ is the reduction mod-2 map. Therefore $\mathrm{Sq}^2\circ\rho$ and Sq^2 are also the first k-invariants of $\mathrm{pic}_0^2(KO)$ and $\mathrm{pic}_0^3(KU)$.*

Proof This follows from Lemmas 3.1, 3.2 and 4.2. □

Next we wish to determine the *second* k-invariants of $\mathrm{pic}_0^3(KU)$ and $\mathrm{pic}_0^2(KO)$. We begin by determining the k-invariants of their connected covers $\mathrm{bgl}_1(KO)[0,2]$ and $\mathrm{bgl}_1(KU)[0,3]$. This is of course equivalent to determining the k-invariants of $\mathrm{gl}_1(KO)[0,1]$ and $\mathrm{gl}_1(KU)[0,2]$. The first is almost trivial, and we prove it in Proposition 4.4. For the second case, more work is required, and we first prove several lemmas.

Proposition 4.4 *The k-invariant of $\mathrm{pic}_1^2(KO)$ is $\mathrm{Sq}^2 : \Sigma H\mathbb{Z}/2 \to \Sigma^3 H\mathbb{Z}/2$. Equivalently, there are two $\mathbb{E}_\infty$-structures on $K(\mathbb{Z}/2,1) \times K(\mathbb{Z}/2,2)$ and the one on $B\mathrm{GL}_1(KO)[0,2]$ is the one which is not the product structure.*

Proof By considering the Postnikov tower and knowing that $H^1(H\mathbb{Z}/2;\mathbb{Z}/2) \cong \mathbb{Z}/2$ is generated by Sq^2 we see that the only two possible k-invariants are 0 and Sq^2. The result then follows from Theorem 3.35 and Corollary 3.36. □

Next we determine the k-invariant of $\mathrm{pic}_1^3(KU)$. We begin by determining all possible k-invariants of a spectrum with the same homotopy groups, or equivalently, all possible infinite loop space structures on the space $\Omega^\infty \mathrm{pic}_1^3(KU) \simeq B\mathrm{GL}_1(KU)[0,3]$.

Lemma 4.5 *There are exactly two h-types whose only nontrivial homotopy groups are $\pi_1 \cong \mathbb{Z}/2$ and $\pi_3 \cong \mathbb{Z}$.*

Proof The result follows immediately from the computation $H^4(B\mathbb{Z}/2;\mathbb{Z}) \cong \mathbb{Z}/2$. □

Lemma 4.6 *Let X be the h-type with $\pi_1(X) \cong \mathbb{Z}/2$ and $\pi_3(X) \cong \mathbb{Z}$ and nontrivial k-invariant. Then X does not admit an $\mathbb{E}_\infty$-structure.*

Proof If X admitted an $\mathbb{E}_\infty$-structure then its k-invariant $K(\mathbb{Z}/2,1) \to K(\mathbb{Z},4)$ would be Ω^∞ of a stable k-invariant $\Sigma H\mathbb{Z}/2 \to \Sigma^4 H\mathbb{Z}$ and would therefore induce a map of abelian groups on cohomology. Let $\gamma \in H^4(K(\mathbb{Z}/2;1);\mathbb{Z}) \cong \mathbb{Z}/2$ be the k-invariant of X and let $\alpha \in H^2(K(\mathbb{Z}/2;1);\mathbb{Z})$ be the generator given by the inclusion $B\mathbb{Z}/2 \to BU(1) = B^2\mathbb{Z}$. Then γ must be the cup-square of α.

Now for any h-type Y the cohomology operation induced by γ is the map

$$H^1(Y;\mathbb{Z}/2) \to H^4(Y;\mathbb{Z}), \quad a \mapsto \beta(a)^2,$$

where β is the Bockstein map. Let $Y = K(\mathbb{Z}/2,1) \times K(\mathbb{Z}/2,1)$. If we take a and b to be the generators of $H^1(Y;\mathbb{Z}/2)$ then we see that the cross term in $\beta(a+b)^2 = \beta(a)^2 + 2\beta(a)\beta(b) + \beta(b)^2$ is nonzero, and therefore the above map is not an abelian group homomorphism. In other words, the k-invariant cannot be Ω^∞ of a stable k-invariant. □

Proposition 4.7 *For any n, $H^{n+3}(\Sigma^n H\mathbb{Z}/2;\mathbb{Z})$ is isomorphic to $\mathbb{Z}/2$, generated by $\Sigma^n(\beta \circ \mathrm{Sq}^2) : \Sigma^n H\mathbb{Z}/2 \to \Sigma^{n+1} H\mathbb{Z}/2 \to \Sigma^{n+3} H\mathbb{Z}$, the appropriate suspension of the composite of the Bockstein and Sq^2.*

Proof It suffices to calculate $H^4(\Sigma H\mathbb{Z}/2;\mathbb{Z}) \simeq H^3(H\mathbb{Z}/2;\mathbb{Z})$. Note that a map $f : X \to \Sigma^3 H\mathbb{Z}$ factors through the Bockstein $\Sigma^3\beta : \Sigma^2 H\mathbb{Z}/2 \to \Sigma^3 H\mathbb{Z}$ if and only if the composite of f with the map $\Sigma^3 2 : \Sigma^3 H\mathbb{Z} \to \Sigma^3 H\mathbb{Z}$ is null, since β is the fiber of multiplication by 2. But the composite of any map $\alpha : H\mathbb{Z}/2 \to \Sigma^3 H\mathbb{Z}$ with $\Sigma^3 2 : \Sigma^3 H\mathbb{Z} \to \Sigma^3 H\mathbb{Z}$ is null because $H^3(H\mathbb{Z}/2;\mathbb{Z})$ is 2-torsion. So every element $\alpha \in H^3(H\mathbb{Z}/2;\mathbb{Z})$ factors as $\beta\alpha'$ for some $\alpha' \in H^2(H\mathbb{Z}/2;\mathbb{Z}/2)$. Therefore $\beta_* : \mathbb{Z}/2 \cong H^2(H\mathbb{Z}/2;\mathbb{Z}/2) \to H^3(H\mathbb{Z}/2;\mathbb{Z})$ is a surjection and $H^3(H\mathbb{Z}/2;\mathbb{Z})$ has only two elements, 0 and $\beta \circ \mathrm{Sq}^2$. Therefore it only remains to check that $\beta \circ \mathrm{Sq}^2$ is nonzero.

Suppose that $\beta \circ \mathrm{Sq}^2$ were null. Then Sq^2 would lift to a map $H\mathbb{Z}/2 \to \Sigma^2 H\mathbb{Z}$, i.e., there would be a factorization, through the fiber of $\Sigma^3\beta$,

$$\begin{tikzcd}
 & \Sigma^2 H\mathbb{Z}/2 \arrow[dl, "\omega"'] \arrow[d, "\mathrm{Sq}^2"] & \\
\Sigma^2 H\mathbb{Z} \arrow[r, "\Sigma^2\rho"] & \Sigma^2 H\mathbb{Z} \arrow[r, "\Sigma^3\beta"] & \Sigma^3 H\mathbb{Z}
\end{tikzcd}$$

where ρ is reduction mod 2. But, similarly to above, every class in $H^2(H\mathbb{Z}/2;H\mathbb{Z})$ is 2-torsion and therefore there is another factorization

$$\begin{tikzcd}
H\mathbb{Z}/2 \arrow[r, "\omega"] \arrow[d, "\mathrm{Sq}^1"'] & \Sigma^2 H\mathbb{Z} \\
\Sigma H\mathbb{Z}/2 \arrow[ur, "\beta"'] &
\end{tikzcd}$$

Therefore we have a commutative diagram

$$\begin{tikzcd}
H\mathbb{Z}/2 \arrow[r, "\mathrm{Sq}^2"] \arrow[d, "\beta\circ\mathrm{Sq}^1"'] & \Sigma^2 H\mathbb{Z}/2 \\
\Sigma^2 H\mathbb{Z} \arrow[ur, "\Sigma^2\rho"'] &
\end{tikzcd}$$

This however is a contradiction because the composite $\Sigma^2\rho \circ \beta \circ \mathrm{Sq}^1$ must be trivial on cohomology classes in degree greater than 1, whereas Sq^2 is not. Therefore $H^3(H\mathbb{Z}/2;\mathbb{Z})$ is isomorphic to $\mathbb{Z}/2$ and is generated by $\beta \circ \mathrm{Sq}^2$. □

Proposition 4.8 *Any h-type equivalent to $K(\mathbb{Z}/2,1) \times K(\mathbb{Z},3)$ admits exactly two $\mathbb{E}_\infty$-structures. One of them is the product structure and the other is the one associated to the stable k-invariant $\Sigma(\beta \circ \mathrm{Sq}^2) : \Sigma H\mathbb{Z}/2 \to \Sigma^4 H\mathbb{Z}$.*

Proof By Proposition 4.7 there are at most two $\mathbb{E}_\infty$-structures on $K(\mathbb{Z}/2,1) \times K(\mathbb{Z},3)$, one associated to the coproduct spectrum $\Sigma H\mathbb{Z}/2 \vee \Sigma^3 H\mathbb{Z}$ and the other associated to the spectrum with the same homotopy groups but nontrivial k-invariant $\Sigma(\beta \circ \mathrm{Sq}^2) : \Sigma H\mathbb{Z}/2 \to \Sigma^4 H\mathbb{Z}$. By taking Ω^∞, the latter yields an infinite loop space, distinct from the product infinite loop space, with homotopy groups $\pi_1 \cong \mathbb{Z}/2$ and $\pi_3 \cong \mathbb{Z}$. By Lemma 4.5 this space has either the trivial or nontrivial k-invariant, and by Lemma 4.6 it cannot have the nontrivial k-invariant. Thus each element of $H^4(\Sigma H\mathbb{Z}/2;\mathbb{Z})$ induces a distinct infinite loop space structure on the product $K(\mathbb{Z}/2,1) \times K(\mathbb{Z},3)$. □

Corollary 4.9 *The k-invariant of $\mathrm{pic}_1^3(KU)$ is $\beta \circ \mathrm{Sq}^2$ and therefore the infinite loop space structure on $\mathrm{Pic}_1^3(KU) \simeq K(\mathbb{Z}/2, 1) \times K(\mathbb{Z}, 3)$ is not the product structure.*

Proof This follows from Theorem 3.31, Corollary 3.33 and Proposition 4.8. □

The following proposition is not immediately relevant but will be used later and follows naturally from Corollary 4.9. We do not know if the splitting also exists at the level of $\mathbb{E}_\infty$-types.

Proposition 4.10 *There is a splitting of $\mathbb{E}_1$-types $\mathrm{Pic}_0^3(KU) \simeq \mathbb{Z}/2 \times \mathrm{Pic}_1^3(KU)$.*

Proof Consider the cofiber sequence of spectra

$$\mathrm{pic}_1^3(KU) \to \mathrm{pic}_0^3(KU) \to H\mathbb{Z}/2.$$

We will calculate the set of possible k-invariants $[H\mathbb{Z}/2, \Sigma\,\mathrm{pic}_1^3(KU)]$. Consider the second cofiber sequence

$$\Sigma H\mathbb{Z}/2 \xrightarrow{\beta\,\mathrm{Sq}^2} \Sigma^4 H\mathbb{Z} \to \Sigma\,\mathrm{pic}_1^3(KU) \to \Sigma^2 H\mathbb{Z}/2 \xrightarrow{\beta\,\mathrm{Sq}^2} \Sigma^5 H\mathbb{Z}.$$

Applying $[H\mathbb{Z}/2, -]$ produces an exact sequence

$$\begin{aligned} H^1(H\mathbb{Z}/2; \mathbb{Z}/2) \xrightarrow{\beta\,\mathrm{Sq}^2} H^4(H\mathbb{Z}/2; \mathbb{Z}) &\to [\Sigma H\mathbb{Z}/2, \Sigma^2\,\mathrm{pic}_1^3(KU)] \\ &\to H^2(H\mathbb{Z}/2; \mathbb{Z}/2) \xrightarrow{\beta\,\mathrm{Sq}^2} H^5(H\mathbb{Z}/2; \mathbb{Z}). \end{aligned}$$

In order to compute the two relevant integral cohomology groups of $H\mathbb{Z}/2$, we use the exact sequence $H\mathbb{Z} \to H\mathbb{Z}/2 \xrightarrow{\beta} \Sigma H\mathbb{Z}$ to identify $H^k(H\mathbb{Z}/2; \mathbb{Z})$ as the image of the multiplication-by-Sq^1 map

$$H^{k-1}(H\mathbb{Z}/2; \mathbb{Z}/2) \to H^k(H\mathbb{Z}/2; \mathbb{Z}/2),$$

which in turn can be made explicit via the standard generators of the Steenrod algebra and the Adem relations. We find that our sequence of interest can be rewritten as

$$\begin{aligned} \mathbb{Z}/2\{\mathrm{Sq}^1\} \xrightarrow{\beta\,\mathrm{Sq}^2} \mathbb{Z}/2\{\beta\,\mathrm{Sq}^2\,\mathrm{Sq}^1\} &\to [H\mathbb{Z}/2, \Sigma\,\mathrm{pic}_1^3(KU)] \\ &\to \mathbb{Z}/2\{\mathrm{Sq}^2\} \xrightarrow{\beta\,\mathrm{Sq}^2} \mathbb{Z}/2\{\beta\,\mathrm{Sq}^4\}. \end{aligned}$$

From this it immediately follows that the first map in the sequence is surjective, and the final map is zero, so the third map gives an isomorphism

$$[H\mathbb{Z}/2, \Sigma\,\mathrm{pic}_1^3(KU)] \cong [H\mathbb{Z}/2, \Sigma^2 H\mathbb{Z}/2] \cong \mathbb{Z}/2\{\mathrm{Sq}^2\}.$$

Now recall that $\mathrm{Sq}^2 : \Sigma H\mathbb{Z}/2 \to \Sigma^3 H\mathbb{Z}/2$ induces the null map on underlying spaces $B\mathbb{Z}/2 \to B^3\mathbb{Z}/2$. Indeed, by the instability relation Sq^2 kills the (degree-1) generator of $H\mathbb{Z}/2^* B\mathbb{Z}/2$. Therefore both infinite-loop maps $B\mathbb{Z}/2 \to B^2\mathrm{Pic}_1^3(KU)$ are null as maps of spaces, and the claim follows. □

4.2 Extending to $\mathrm{pic}_0^2(KO)$ and $\mathrm{pic}_0^3(KU)$

So far we have computed the k-invariants of $\mathrm{pic}_0^1(KO)$, $\mathrm{pic}_0^1(KU)$, $\mathrm{pic}_1^2(KO)$ and $\mathrm{pic}_1^3(KU)$. Now we wish to determine how this data can be glued together to understand the k-invariants of $\mathrm{pic}_0^2(KO)$ and $\mathrm{pic}_0^3(KU)$. Our computations only determine the second k-invariants of these spectra up to isomorphism. In other words, for KO we determine that the second k-invariant of $\mathrm{pic}_0^2(KO)$ is one of two elements of $H^3(\mathrm{pic}_0^1(KO);\mathbb{Z}/2)\cong\mathbb{Z}/2\times\mathbb{Z}/2$, and for KU it is one of the two generators of $H^4(\mathrm{pic}_0^1(KU))\cong\mathbb{Z}/4$. In the latter case there is an equivalence of Postnikov towers which interchanges the two generators, but in the former case it is less clear which generator one should choose. Ultimately, however, the choice will not matter because both k-invariants work for the applications of Section 5. In particular, both choices satisfy the conclusions of Propositions 4.18 and 4.19.

The proof of Lemma 4.12 was sketched for us by Tyler Lawson. Any mistakes are of course due to our own misunderstanding.

Lemma 4.11 *There is an exact sequence*

$$0\to H^4(H\mathbb{Z}/2;\mathbb{Z})\to H^4(\mathrm{pic}_0^1(KU);\mathbb{Z})\to H^3(H\mathbb{Z}/2;\mathbb{Z})\to 0$$

in which the first group is generated by $\beta\,\mathrm{Sq}^2\,\mathrm{Sq}^1$ *and the last group is generated by* $\beta\,\mathrm{Sq}^2$.

Proof The generators of the first and last groups are standard computations. The fact that

$$H^4(H\mathbb{Z}/2;\mathbb{Z})\to H^4(\mathrm{pic}_0^1(KU);\mathbb{Z})\to H^3(H\mathbb{Z}/2;\mathbb{Z})$$

is exact follows immediately from applying $H^4(-;\mathbb{Z})$ to the cofiber sequence

$$\Sigma H\mathbb{Z}/2\to\mathrm{pic}_0^1(KU)\to H\mathbb{Z}/2$$

from the Postnikov tower of $\mathrm{pic}(KU)$. The fact that the entire sequence is short exact follows on the left from the fact that $H^2(H\mathbb{Z}/2;\mathbb{Z})=0$. For the right-hand side, first recall that the first k-invariant of $\mathrm{pic}(KU)$, i.e., the only k-invariant of $\mathrm{pic}_0^1(KU)$, is $\mathrm{Sq}^2: H\mathbb{Z}/2\to\Sigma^2H\mathbb{Z}/2$. Therefore there is an exact sequence

$$H^4(\mathrm{pic}_0^1(KU);\mathbb{Z})\to H^3(H\mathbb{Z}/2;\mathbb{Z})\to H^5(H\mathbb{Z}/2;\mathbb{Z})$$

in which the last map is Sq^2. Since $H^3(H\mathbb{Z}/2;\mathbb{Z})\cong\mathbb{Z}/2$ is generated by $\beta\,\mathrm{Sq}^2$ the image of this map is $\beta\,\mathrm{Sq}^2\,\mathrm{Sq}^2$. Taking the quotient by 2, which is an injection $H^*(H\mathbb{Z}/2;\mathbb{Z})\to H^*(H\mathbb{Z}/2;\mathbb{Z}/2)$ (its kernel is the image of multiplication by 2 between $\mathbb{Z}/2$-modules), we get $\mathrm{Sq}^1\,\mathrm{Sq}^2\,\mathrm{Sq}^2$, which is zero by the Adem relations. Therefore the image of $\mathrm{Sq}^2: H^3(H\mathbb{Z}/2;\mathbb{Z})\to H^5(H\mathbb{Z}/2;\mathbb{Z})$ is zero. □

Lemma 4.12 *There is an isomorphism* $H^4(\mathrm{pic}_0^1(KU);\mathbb{Z})\cong\mathbb{Z}/4$.

Proof Consider the cofiber sequence $\mathbb{S}\xrightarrow{2}\mathbb{S}\xrightarrow{q}\mathbb{S}/2$, where $\mathbb{S}/2$ is the mod-2 Moore spectrum. Let $\bar{f}:\mathbb{S}\to\mathrm{pic}_0^1(KU)$ be the generator of $\pi_0(\mathrm{pic}_0^1(KU))\cong\mathbb{Z}/2$. Then since $\pi_0(\mathrm{pic}_0^1(KU))$ is 2-torsion, $\bar{f}$ lifts to a nonzero map $f:\mathbb{S}/2\to\mathrm{pic}_0^1(KU)$. The long exact sequence in homotopy applied to the above cofiber sequence shows that $\pi_0(\mathbb{S}/2) \cong \mathbb{Z}/2$, generated by q and $\pi_1(\mathbb{S}/2) \cong \mathbb{Z}/2$, generated by $q \circ \eta$. By considering the commutative diagram

$$\begin{tikzcd}
\Sigma^d\mathbb{S} \arrow[d, "\phi"'] \arrow[dr, "q\circ\phi"] & \\
\mathbb{S} \arrow[r, "q"] \arrow[d, "\bar{f}"'] & \mathbb{S}/2 \arrow[dl, "f"] \\
\mathrm{pic}_0^1(KU) &
\end{tikzcd}$$

in which $d = 0, 1$ and $\phi = \mathrm{id}, \eta$, respectively, we have that f must be nonzero on π_0 and π_1 and therefore an isomorphism on those two homotopy groups. As a result, the cofiber of f, $\mathrm{cofib}(f)$, is 2-connected. Applying the Hurewicz theorem to $\mathrm{cofib}(f)$, we find that $\pi_3(\mathrm{cofib}(f)) \cong H^3(\mathrm{cofib}(f); \mathbb{Z})$. Applying the long exact sequence in homotopy to the cofiber sequence $\mathbb{S}/2 \to \mathrm{pic}_0^1(KU) \to \mathrm{cofib}(f)$ we see that $\pi_3(\mathrm{cofib}(f)) \cong \pi_2(\mathbb{S}/2)$. Now applying the long exact sequence for homology to the same cofiber sequence, we get $\pi_2(\mathbb{S}/2) \cong H^3(\mathrm{pic}_0^1(KU); \mathbb{Z})$. The universal coefficient theorem tells us that $H^4(\mathrm{pic}_0^1(KU); \mathbb{Z}) \cong \mathrm{Ext}(\pi_2(\mathbb{S}/2), \mathbb{Z}) \cong \pi_2(\mathbb{S}/2)^\vee \cong \pi_2(\mathbb{S}/2)$. We conclude by pointing out that $\pi_2(\mathbb{S}/2) \cong \mathbb{Z}/4$. The authors cannot find this final fact in the published literature, but several sketch proofs of it are provided in [40]. □

Proposition 4.13 *The second k-invariant of $\mathrm{pic}_0^3(KU)$ generates $H^4(\mathrm{pic}_0^1(KU); \mathbb{Z}) \cong \mathbb{Z}/4$.*

Proof By Lemma 4.12 and Theorem 3.31 there are three options for the second k-invariant of $\mathrm{pic}_0^3(KU)$: any of the nontrivial maps $\mathrm{pic}_0^1(KU) \to \Sigma^4 H\mathbb{Z}$. However, by Corollary 4.9, the actual second k-invariant must give $\beta\,\mathrm{Sq}^2$ on $\mathrm{pic}_1^3(KU)$ after taking a connected cover. This corresponds to asking for maps $\mathrm{pic}_0^1(KU) \to \Sigma^4 H\mathbb{Z}$ which restrict to $\beta\,\mathrm{Sq}^2$ when precomposing with the map $\Sigma H\mathbb{Z}/2 \to \mathrm{pic}_0^1(KU)$ in the Postnikov tower of $\mathrm{pic}(KU)$. But by Lemma 4.11, $\mathbb{Z}/4 \cong H^4(\mathrm{pic}_0^1(KU); \mathbb{Z}) \to H^3(\mathbb{Z}/2; \mathbb{Z}) \cong \mathbb{Z}/2$ is a surjection, and the codomain is generated by $\beta\,\mathrm{Sq}^2$. Therefore only the two generators of $H^4(\mathrm{pic}_0^1(KU); \mathbb{Z})$ satisfy the necessary property, so one of them of must be the k-invariant of $\mathrm{pic}_0^3(KU)$. □

Theorem 4.14 *The two generators of $\mathbb{Z}/4 \cong H^4(\mathrm{pic}_0^1(KU); \mathbb{Z})$ yield equivalent Postnikov sections, and hence both present $\mathrm{pic}_0^3(KU)$.*

Proof Let a and b be the two generators of $\mathbb{Z}/4 \cong H^4(\mathrm{pic}_0^1(KU); \mathbb{Z})$. Since $a = -b$ there is a commutative diagram

$$\begin{tikzcd}
\mathrm{pic}_0^1(KU) \arrow[r, "a"'] \arrow[d, "\mathrm{id}"] & \Sigma^4 H\mathbb{Z} \arrow[d, "-1"] \\
\mathrm{pic}_0^1(KU) \arrow[r, "b"] & \Sigma^4 H\mathbb{Z}
\end{tikzcd}$$

The induced map between the fibers of the horizontal maps is an equivalence between the Postnikov sections corresponding to the k-invariants a and b. □

Remark 4.15 While we do not have a geometric argument at hand, it seems almost certain that the automorphism used in the proof of Theorem 4.14 corresponds to the complex conjugation automorphism on KU.

Proposition 4.16 *There is an isomorphism* $H^3(\mathrm{pic}_0^1(KO);\mathbb{Z}/2) \cong \mathbb{Z}/2 \times \mathbb{Z}/2$. *Moreover, pulling back along the fiber in the Postnikov tower for* $\mathrm{pic}_0^1(KO)$, $\Sigma H\mathbb{Z}/2 \to \mathrm{pic}_0^1(KO)$, *induces a surjection* $\mathbb{Z}/2 \times \mathbb{Z}/2 \to \mathbb{Z}/2\{\mathrm{Sq}^2\}$.

Proof For readability we do not include the entire proof. We only note that it arises from considering the long exact sequence in mod-2 cohomology applied to the Postnikov tower $\Sigma H\mathbb{Z}/2 \to \mathrm{pic}_0^1(KO) \to H\mathbb{Z}/8$ and a large number of low-degree cohomology computations for the Eilenberg–Mac Lane spectra $H\mathbb{Z}/8$ and $H\mathbb{Z}/2$. □

Remark 4.17 For the time being, we do not know how to specify the "correct" k-invariant for $\mathrm{pic}_0^2(KO)$, as it could be one of two elements in the preimage of Sq^2. In the case of $\mathrm{pic}(KU)$ the ambiguity is irrelevant up to equivalence (see Theorem 4.14), but a similar approach will not work here. It may be possible to resolve the ambiguity by taking homotopy fixed points of $\mathrm{pic}(KU)$ and comparing the second k-invariant of the resulting fixed point spectrum, via the homotopy fixed points spectral sequence, to the two possibilities given in Proposition 4.16. Luckily, this uncertainty does not effect the group structure on the $\mathrm{pic}_0^2(KO)$-cohomology of a space.

4.3 Group structures

Now that we know the k-invariants of $\mathrm{pic}_0^2(KO)$ and $\mathrm{pic}_0^3(KU)$, we can determine the group laws for $\mathrm{pic}_0^2(KO)^0(X)$ and $\mathrm{pic}_0^3(KU)^0(X)$, which we will use in the next section.

Proposition 4.18 *For a space* X, *the group law on the set*

$$\mathrm{pic}_0^3(KU)^0(X) \cong H^0(X;\mathbb{Z}/2) \times H^1(X;\mathbb{Z}/2) \times H^3(X;\mathbb{Z})$$

is given by $(a,b,c) \boxplus (a',b',c') = (a+a', b+b', c+c'+\beta(b \cup b'))$, *where we abuse notation and use the symbol* $+$ *to denote the usual addition in* $H^0(X;\mathbb{Z}/2)$, $H^1(X;\mathbb{Z}/2)$ *and* $H^3(X;\mathbb{Z})$.

Proof Proposition 4.10 implies that the first coordinate of the group splits off. Therefore it suffices to prove that the group structure on $\mathrm{pic}_1^3(KU)^0(X)$ is $(b,c)+(b',c') = (b+b', c+c'+\beta(b \cup b'))$. Consider the natural short exact sequence of abelian groups

$$H^3(X;\mathbb{Z}) \to \mathrm{pic}_1^3(KU)^0(X) \to H^1(X;\mathbb{Z}/2).$$

The cocycle for that group extension (in the sense of [49, Section 6.6]) is a natural map

$$H^1(X;\mathbb{Z}/2) \times H^1(X;\mathbb{Z}/2) \to H^3(X;\mathbb{Z}),$$

which (since it is natural in X) is represented by some map $K(\mathbb{Z}/2,1) \times K(\mathbb{Z}/2,1) \to K(\mathbb{Z},3)$. There are only two such maps, the trivial one and $\beta(- \cup -)$. The cocycle cannot be trivial, otherwise $\mathrm{pic}_1^3(KU)$ would be a direct sum of Eilenberg–Mac lane spectra, which contradicts Corollary 4.9. □

The proof of the following proposition is similar.

Proposition 4.19 *For an h-type X, the group law on the set*

$$\mathrm{pic}_0^2(KO)^0(X) \cong H^0(X;\mathbb{Z}/2)\times H^1(X;\mathbb{Z}/2)\times H^2(X;\mathbb{Z}/2)$$

is given by $(a,b,c)\boxplus(a',b',c')=(a+a',b+b',c+c'+b\cup b')$, where we abuse notation and use the symbol $+$ to denote the usual addition in $H^0(X;\mathbb{Z}/8)$, $H^1(X;\mathbb{Z}/2)$ and $H^2(X;\mathbb{Z}/2)$.

5 What $\mathrm{pic}_0^2(KO)$ and $\mathrm{pic}_0^3(KU)$ represent

In the following section we describe several algebraic and geometric interpretations of the cohomology theories associated to $\mathrm{pic}_0^2(KO)$ and $\mathrm{pic}_0^3(KU)$.

5.1 Brauer groups

By determining the group structures of $\mathrm{pic}_0^3(KU)^0(X)$ and $\mathrm{pic}_0^2(KO)^0(X)$ for a space X we now have that $\mathrm{Pic}_0^3(KU)$ and $\mathrm{Pic}_0^2(KO)$ represent well-known classical Brauer groups whose elements are Morita classes of bundles of $\mathbb{Z}/2$-graded central simple algebras [10, Theorem 6, Theorem 11]; bundles of $\mathbb{Z}/2$-graded continuous trace C^*-algebras (as described below); and bundles of super (i.e., $\mathbb{Z}/2$-graded) 2-lines (see Corollary 5.8). All of these data were previously known to be isomorphic, at least at the level of folks theorems, but interpreting them in terms of $\mathrm{Pic}_0^3(KU)$ and $\mathrm{Pic}_0^2(KO)$ is new. However, this is consistent with the fact that bundles of graded central simple algebras; bundles of graded C^*-algebras; and bundles of super 2-lines; can all be used to twist K-theory.

Definition 5.1 We let $\mathrm{GBr}\mathbf{O}(X)$ denote the Brauer group of (possibly infinite-dimensional) graded, continuous trace, complex C^*-algebras with spectrum X, as in [44]. We let $\mathrm{GBr}\mathbf{U}(X)$ denote the Brauer group of (possibly infinite-dimensional) graded, continuous trace, real C^*-algebras with spectrum X, as in [43].

The following theorems are proven in [43; 44]:

Theorem 5.2 (Maycock) *If X is a space homotopy equivalent to a CW-complex then*

$$\mathrm{GBr}\mathbf{U}(X)\cong H^0(X;\mathbb{Z}/2)\times H^1(X;\mathbb{Z}/2)\times H^3(X;\mathbb{Z})$$

with group law $(a,b,c)+(a',b',c')=(a+a',b+b',c+c'+\beta(b\cup b'))$ where β is the Bockstein homomorphism.

Remark 5.3 In [44] the H^0 term is mostly ignored, since Maycock requires that her bundles have isomorphic fibers over every connected component. This assumption is unnecessary, as shown in [43].

Theorem 5.4 (Moutuou) *If X is a space homotopy equivalent to a CW-complex then*

$$\mathrm{GBr}\mathbf{O}(X)\cong H^0(X;\mathbb{Z}/8)\times H^1(X;\mathbb{Z}/2)\times H^2(X;\mathbb{Z}/2)$$

with group law $(a,b,c)+(a',b',c')=(a+a',b+b',c+c'+(b\cup b'))$.

Propositions 4.18 and 4.19 now imply the following corollaries:

Corollary 5.5 *If X is a space homotopy equivalent to a CW-complex then there is an isomorphism*

$$\mathrm{GBr}\mathbf{U}(X) \cong \mathrm{pic}_0^3(KU)^0(X).$$

Corollary 5.6 *If X is a space homotopy equivalent to a CW-complex then there is an isomorphism*

$$\mathrm{GBr}\mathbf{O}(X) \cong \mathrm{pic}_0^2(KO)^0(X).$$

Remark 5.7 The Brauer group of $\mathbb{Z}/2$-graded continuous trace C^*-algebras with spectrum X is equivalent to the Brauer group of $\mathbb{Z}/2$-equivariant C^*-algebras with spectrum X with the property that the induced $\mathbb{Z}/2$-action on X is trivial. Our constructions of $\mathrm{pic}_1^3(KU)$ and $\mathrm{pic}_1^2(KO)$, and the fact that they represent these Brauer groups, should be compared to the construction of the cohomology theory $E_{\mathbb{C},\mathbb{T}}$ in [14]. It is shown therein that $E^0_{\mathbb{C},\mathbb{T}}(X)$ is the group of $\mathbb{T}$-equivariant line bundles on X and that $E^1_{\mathbb{C},\mathbb{T}}(X)$ is the Brauer group of $\mathbb{T}$-equivariant C^*-algebras with spectrum X where X has trivial induced $\mathbb{T}$-action. Our constructions on the other hand show that $\mathrm{pic}_1^3(KU)^{-1}(X)$ is the group of super line bundles on X and that $\mathrm{pic}_1^3(KU)^0(X)$ is the Brauer group of $\mathbb{Z}/2$-equivariant C^*-algebras with spectrum X having trivial induced $\mathbb{Z}/2$-action. Moreover, $\Omega^\infty E_{\mathbb{C},\mathbb{T}} \simeq \mathbb{Z} \times K(\mathbb{Z},2)$ and $\Omega^\infty \Sigma^{-1} \mathrm{pic}_1^3(KU) \simeq \mathbb{Z}/2 \times K(\mathbb{Z},2)$. This suggests that Evans and Pennig's $E_{\mathbb{C},\mathbb{T}}$ spectrum is equivalent to $\mathrm{gl}_1(KU_{\mathbb{T}})$, the space of units of $\mathbb{T}$-equivariant complex K-theory.

Propositions 4.18 and 4.19 also give $\mathrm{pic}_0^3(KU)$ and $\mathrm{pic}_0^2(KO)$ the following interpretations in terms of graded 2-line bundles.

Corollary 5.8 *Let X be a smooth manifold. Then*

$$\mathrm{pic}_0^3(KU)^0(X) \cong \mathrm{sLBdl}_{\mathbb{C}}(X) \quad \textit{and} \quad \mathrm{pic}_0^2(KO)^0(X) \cong \mathrm{sLBdl}_{\mathbb{R}}(X),$$

where $\mathrm{sLBdl}_{\mathbb{C}}$ and $\mathrm{sLBdl}_{\mathbb{R}}$ are the Brauer groups of complex and real super 2-line bundles on X as defined in [34].

Proof This follows immediately from either [34, Theorem 4.4] or [42, Theorem 2.2.6]. □

Corollaries 5.5 and 5.8 have the following interpretation in terms of parameterized stable homotopy theory:

Corollary 5.9 *If X is connected then there is an isomorphism between the Brauer group of graded, continuous trace, complex C^*-algebras with spectrum X, equivalently the Brauer group of complex super 2-line bundles on X, and $[X, B(\mathrm{GL}_1(KU[0,2]))]$, the group of $KU[0,2]$-line bundles on X (where the group structure on the latter arises from the abelian ∞-group structure of the target). The same holds in the real case with $KU[0,2]$ replaced by $KO[0,1]$.*

Proof We have a string of equivalences of infinite loop spaces

$$\mathrm{Pic}_1^3(KU) \simeq B\Omega\,\mathrm{Pic}_0^3(KU) \simeq B\mathrm{GL}_1(KU)[0,3] \simeq B(\mathrm{GL}_1(KU)[0,2]).$$

Therefore it suffices to show that $\mathrm{GL}_1(KU)[0,2] \simeq \mathrm{GL}_1(KU[0,2])$, which follows from Lemma 5.12 below. The argument for the real case is identical. □

Remark 5.10 The group structure on $KU[0,2]$-line bundles over X also arises by interpreting it as the set of connected components of the symmetric monoidal slice category ∞-category $\mathrm{Top}_{/B(\mathrm{GL}_1(KU[0,2]))}$. Here, the symmetric monoidal structure is given by [3, Proposition 6.12]. This can be interpreted as applying Lurie's straightening construction to the slice category and equipping the resulting presheaf category with the Day convolution monoidal structure.

Remark 5.11 Corollary 5.9 cannot be stated in terms of all of $\mathrm{Pic}_0^3(KU)$ for the following reason: the spectrum $KU[0,2]$ is no longer 2-periodic, so $\pi_0(\mathrm{Pic}(KU[0,2]))$ will be at least $\mathbb{Z}$. As a result we cannot think of maps $X \to \mathrm{Pic}_0^3(KU)$ as bundles of invertible $KU[0,2]$-modules on X.

Lemma 5.12 *Let R be a connective commutative ring spectrum and $n \in \mathbb{Z}$ a positive integer. Then there is an equivalence of infinite loop spaces*

$$\mathrm{GL}_1(R[0,n]) \simeq \mathrm{GL}_1(R)[0,n].$$

Proof Consider the zigzag of infinite loop maps

$$\mathrm{GL}_1(R[0,n]) \to \mathrm{GL}_1(R[0,n])[0,n] \leftarrow \mathrm{GL}_1(R)[0,n],$$

where the arrow from left to right is the usual map from a space to its truncation and the arrow from right to left is obtained by truncating after applying the functor GL_1 to the ring map $R \to R[0,n]$. Recall that the homotopy groups of $\mathrm{GL}_1(R)$ are isomorphic to those of R except in degree zero where $\pi_0(\mathrm{GL}_1(R)) \cong \pi_0(R)^\times$. Therefore $\mathrm{GL}_1(R[0,n])$ is an n-truncated space and the left to right map from itself to its truncation is an equivalence. Because $R \to R[0,n]$ is an equivalence through homotopy degree n, so is $\mathrm{GL}_1(R) \to \mathrm{GL}_1(R[0,n])$ and therefore the left to right map is also an equivalence. □

Remark 5.13 There is an equivalence of commutative ring spectra $KU[0,2] \simeq ku[0,2]$ which is the truncation of the equivalence $ku \to KU[0,\infty)$. This is essential to the use of Lemma 5.12 in the proof of Corollary 5.9. In particular we must construct $KU[0,2]$ by first taking the connective cover of KU and then truncating, as $KU(-\infty,2]$ is not even a ring spectrum.

Remark 5.14 Recall that there is a C_2-action on KU by complex conjugation whose fixed point spectrum is KO. The C_2-equivariant complex K-theory spectrum is often denoted by $K\mathbb{R}$, following Atiyah. One can show that the second C_2-equivariant Postnikov slice of $K\mathbb{R}$, denoted by $P^2 K\mathbb{R}$ (in the sense of [28]) has underlying spectrum $KU[0,2]$ and fixed point spectrum $KO[0,1]$ [27]. Therefore both perspectives are encompassed by the space of *equivariant* units of $K\mathbb{R}$, i.e., $\mathrm{GL}_1(K\mathbb{R})$. In other words, given the correct equivariant generalization of Lemma 5.12, both the real and complex Brauer groups described above are determined by a delooping of $P^2\mathrm{GL}_1(K\mathbb{R})$. This is closely related the operator-theoretic perspective on real K-theory described in [43]. We will return to this question in later work.

5.2 Twisting Spin- and String-structures

Recall from [45] that the set of **String**-structures on a **Spin**-manifold M is a torsor for $H^3(M;\mathbb{Z})$. Our results admit a similar interpretation, except that we are considering **String**-structures on a manifold with a fixed orientation (as opposed to a fixed **Spin**-structure). Our results imply that the set of **String**-structures on M relative to a fixed orientation is a torsor for $\mathrm{pic}_1^3(KU)^0(M) \cong H^1(M;\mathbb{Z}/2) \ltimes H^3(M;\mathbb{Z})$. Similarly, the **Spin**-structures on a real manifold M are a torsor for $\mathrm{pic}_1^2(KO)^0(M) \cong H^1(M;\mathbb{Z}/2) \ltimes H^2(M;\mathbb{Z}/2)$.

Definition 5.15 Let $\mathbf{SO}/\!\!/\mathbf{String}$ denote the abelian ∞-group arising as the fiber of the connective cover $\mathbf{BString} \to \mathbf{BSO}$.

Proposition 5.16 *There is an equivalence of abelian ∞-groups* $\mathbf{SO}/\!\!/\mathbf{String} \simeq \mathrm{Pic}_1^3(KU)$.

Proof Let F be the fiber in $\mathcal{S}\mathrm{p}$ of the connective cover $\mathbf{bstring} \to \mathbf{bso}$. Therefore F has $\pi_1(F) \cong \mathbb{Z}/2$ and $\pi_3(F) \cong \mathbb{Z}$ as its only nontrivial homotopy groups. By Proposition 4.8 and Corollary 4.9 it suffices to show that the k-invariant of F is nontrivial. We will show that the k-invariant of ΣF is nontrivial, which is equivalent.

Note that there is a fiber sequence $\mathbf{bstring} \to \mathbf{bso} \to \mathbf{bso}[0,4] \simeq \Sigma F$. By Lemma 4.6, there is an equivalence of h-types

$$\mathbf{BSO}[0,4] \simeq B(B\mathrm{GL}_1(KU)[0,3]) \simeq K(\mathbb{Z}/2,2) \times K(\mathbb{Z},4)$$

so there is a natural isomorphism of sets $\mathbf{bso}[0,4]^0(X) \cong H^2(X;\mathbb{Z}/2) \oplus H^4(X;\mathbb{Z})$. If the k-invariant of F were trivial, this would be an isomorphism of abelian groups. We will show that is not the case.

Because of the h-type splitting $\mathbf{BSO}[0,4] \simeq K(\mathbb{Z}/2,2) \times K(\mathbb{Z},4)$ described above, there is a projection $\mathbf{BSO}[0,4] \to K(\mathbb{Z}/2,2)$ and the composite of that projection with the truncation $\mathbf{BSO} \to \mathbf{BSO}[0,4]$ must be nontrivial. Therefore the second Stiefel Whitney class $w_2 : \mathbf{BSO} \to K(\mathbb{Z}/2,2)$ can be factored as $\mathbf{BSO} \to \mathbf{BSO}[0,4] \to K(\mathbb{Z}/2,2)$. As a result, the composite $\mathbf{bso}^0(X) \to \mathbf{bso}[0,4]^0(X) \to H^2(X,\mathbb{Z}/2)$ must take an oriented vector bundle V on X to $w_2(V)$.

The composite $p : \mathbf{BSO} \to \mathbf{BSO}[0,4] \to K(\mathbb{Z},4)$ determines *some* integral characteristic class (and is therefore some multiple of the first Pontryagin class p_1). We argue that it must be either p_1 or $-p_1$. First note that by the computations of [12] and the Künneth formula, $H_4(\mathbf{BSO}[0,4];\mathbb{Z})$ splits as $H_4(K(\mathbb{Z}/2,2);\mathbb{Z}) \oplus H_4(K(\mathbb{Z},4);\mathbb{Z}) \cong \mathbb{Z}/4 \oplus \mathbb{Z}$. This implies that $H^4(K(\mathbb{Z}/2,2) \times K(\mathbb{Z},4);\mathbb{Z}) \cong \mathrm{Hom}_{\mathrm{Ab}}(\mathbb{Z}/4 \oplus \mathbb{Z},\mathbb{Z}) \cong \mathbb{Z}$. Thus the projection $\mathbf{BSO}[0,4] \to K(\mathbb{Z},4)$ gives an isomorphism in H^4. The restriction along $\mathbf{BSO} \to \mathbf{BSO}[0,4]$ is surjective on H^4 for connectivity reasons, and hence an isomorphism, so the composite p must be a generator of $H^4(\mathbf{BSO},\mathbb{Z})$, which proves the claim.

Since the natural map $\mathbf{bso}^0(X) \to \mathbf{bso}[0,4]^0(X)$ is a map of abelian groups, if the isomorphism $\mathbf{bso}[0,4]^0(X) \cong H^2(X;\mathbb{Z}/2) \oplus H^4(X;\mathbb{Z})$ were also one of abelian groups then the map which sends an oriented vector bundle V to $(w_2(V), \pm p_1(V))$ would be a map of abelian groups. But the Whitney sum formulas that dictate the behavior of w_2 and p_1 under the direct sum of bundles make this impossible. □

Corollary 5.17 *Let X be a space with an oriented real vector bundle $\xi : X \to$ **BSO** which lifts to a string bundle. Then the set of string bundles which lift ξ is a torsor for* $\mathrm{pic}_1^3(KU)^0(X) \simeq \mathrm{bgl}_1(KU[0,2])^0(X)$.

Proof This follows from applying the limit preserving functor $\mathrm{Map}(X,-)$ to the pullback of h-types

$$\begin{tikzcd}
\mathrm{Pic}_1^3(KU) \arrow[r] \arrow[d] & \mathbf{BString} \arrow[d] \\
\{*\} \arrow[r, "\xi"] & \mathbf{BSO}
\end{tikzcd}$$

□

Remark 5.18 Corollary 5.17 implies that if X is a connected and oriented manifold which admits a string structure then those string structures can be twisted by elements of $\mathrm{bgl}_1^0(KU[0,2])(X)$. These are $KU[0,2]$-line bundles on X and therefore, in light of Corollary 5.8, complex super 2-line bundles.

Remark 5.19 We suspect it is also true that $\mathrm{Pic}_0^3(KU) \simeq \mathrm{fib}(\mathbf{BString} \to \mathbf{BO})$, but we don't have an interesting interpretation of this fact, so we do not investigate it here.

The next proposition can be proven by methods similar to those used in the proof of Proposition 5.16:

Proposition 5.20 *If* $\mathbf{O}/\!\!/\mathbf{Spin}$ *denotes the fiber of the map* $\mathbf{BSpin} \to \mathbf{BO}$ *then there is an equivalence of abelian* ∞*-groups* $\mathbf{O}/\!\!/\mathbf{Spin} \simeq \mathrm{Pic}_1^2(KO) \simeq B\mathrm{GL}_1(KO[0,1])$.

Corollary 5.21 *For an* h*-type* X *with a real vector bundle* $\xi : X \to \mathbf{BO}$ *the set of lifts of* ξ *to* $\mathbf{BSpin}$ *is a torsor for* $\mathrm{bgl}_1^0(KO[0,1])$.

Remark 5.22 In light of the results of [7; 13], Propositions 5.16 and 5.20 imply that $M\mathbf{String} \to M\mathbf{SO}$ and $M\mathbf{Spin} \to M\mathbf{O}$ are Hopf–Galois extensions (or co-Galois extensions) in the sense of [47] with Galois algebras $\mathbb{S}[B\mathrm{GL}_1(KU[0,2])]$ and $\mathbb{S}[B\mathrm{GL}_1(KO[0,1])$, respectively. In other words, there are morphisms of affine spectral schemes $\mathrm{Spec}(M\mathbf{SO}) \to \mathrm{Spec}(M\mathbf{String})$ and $\mathrm{Spec}(M\mathbf{O}) \to \mathrm{Spec}(M\mathbf{Spin})$ which are torsors for the affine commutative group schemes $\mathrm{Spec}(\mathbb{S}[B\mathrm{GL}_1(KU[0,2])])$ and $\mathrm{Spec}(\mathbb{S}[B\mathrm{GL}_1(KO[0,1])]$, respectively. We note however that this in no way implies that either $M\mathbf{String} \to M\mathbf{SO}$ or $M\mathbf{Spin} \to M\mathbf{O}$ is an actual *Galois* extension. These statements could be made more precise with the language of [39] but we leave that for another day.

5.3 Twisting cohomology theories

There are maps of h-types $\mathrm{Pic}_0^3(KU) \to \mathrm{Pic}(KU)$ and $\mathrm{Pic}_0^3(KU) \to \mathrm{Pic}(KO)$, but we do not know if these are maps of abelian ∞-groups because we do not know if there are $\mathbb{E}_\infty$-splittings $\mathrm{Pic}(KU) \simeq \mathrm{Pic}_0^3(KU) \times \mathrm{Pic}_4^\infty(KU)$ or $\mathrm{Pic}(KO) \simeq \mathrm{Pic}_0^2(KO) \times \mathrm{Pic}_3^\infty(KO)$. However, there *are* $\mathbb{E}_\infty$-splittings

$$B\mathrm{GL}_1(KU) \simeq \mathrm{Pic}_1^3(KU) \times B\mathrm{GL}_1(KU)[4,\infty) \quad \text{and} \quad B\mathrm{GL}_1(KO) \simeq \mathrm{Pic}_1^2(KO) \times B\mathrm{GL}_1(KO)[3,\infty)$$

(this is well known, but also follows from our Corollaries 3.33 and 3.36). Therefore for a connected h-type X there are twists of real and complex K-theory by $\mathrm{pic}_1^3(KU)^0(X)$ and $\mathrm{pic}_1^2(KO)^0(X)$.

Question 5.23 *Given a class* $\alpha \in \mathrm{pic}_1^3(KU)^0(X)$ *there is a twisted* K*-theory group* $K^\alpha(X)$. *On the other hand, via the isomorphism of Corollary 5.5,* α *is also a class in* $\mathrm{GBr}\mathbf{U}(X)$, *i.e., a Morita class of graded continuous class* C^**-algebras with spectrum* X. *Then, by* [44, Section 4.1] *there is a twisted* *operator-theoretic KK-theory group $KK^{\alpha}(X)$. Are these two groups always isomorphic? This question, for the comparison between ungraded C^*-algebras and $H^3(X;\mathbb{Z})$ is answered in the affirmative in* [26].

More generally, using the results of Section 5.2, we may repeat the constructions of [2] to obtain twists of other cohomology theories that now can be interpreted as coming from $KU[0,2]$ and $KO[0,1]$ line bundles.

Example 5.24 By taking Thom spectra of the fiber sequence $B\mathrm{GL}_1(KO[0,1]) \to \mathbf{BSpin} \to \mathbf{BO}$ of Proposition 5.20 and composing with the $\mathbb{E}_\infty$ Atiyah–Bott–Shapiro orientation of [31], we obtain a composite of maps of $\mathbb{E}_\infty$-ring spectra

$$\Sigma^\infty_+ B\mathrm{GL}_1(KO[0,1]) \to M\mathbf{Spin} \to KO.$$

The $\mathrm{gl}_1 \vdash \Sigma^\infty_+\Omega^\infty$ adjunction then induces a map of abelian ∞-groups $B\mathrm{GL}_1(KO[0,1]) \to \mathrm{GL}_1(KO)$ which deloops to

$$B^2\mathrm{GL}_1(KO[0,1]) \to B\mathrm{GL}_1(KO).$$

Recall that $B^2\mathrm{GL}_1(KO[0,1])$ is the base space component of $\mathrm{Br}(KO[0,1])$ and $B\mathrm{GL}_1(KO[0,1]) \simeq \Omega B\mathrm{GL}_1(KO[0,1])$ classifies super 2-line bundles. We think it reasonable to interpret this map as giving twists of KO-theory by real super 3-line bundles (though there does not appear to be an agreed upon definition of 3-line bundles in the literature).

Remark 5.25 Note that the fiber of the composite $\mathbf{BSpin}^c \to \mathbf{BSO} \to \mathbf{BO}$ is, as an h-type, equivalent to $\mathbb{Z}/2 \times K(\mathbb{Z},2)$. Under the assumption that this fiber has nontrivial $\mathbb{E}_\infty$-structure, Theorem 3.35 implies that it is equivalent as an abelian ∞-group to the h-type classifying complex superline bundles. In other words, $\mathbf{Spin}^c$-structures on manifolds can be twisted by complex superline bundles. We believe that this fact has an interpretation in terms of Clifford algebra bundles which is the subject of joint work of the first two authors and Pacheco-Tallaj.

Example 5.26 Similarly to Example 5.24, we can use the fiber sequence in the proof of Proposition 5.16, $B\mathrm{GL}_1(KU[0,2]) \to \mathbf{BString} \to \mathbf{BSO}$, along with the $\mathbb{E}_\infty$-orientation $M\mathbf{String} \to KU$ of Ando, Hopkins and Rezk [5] to obtain twists

$$B^2\mathrm{GL}_1(KU[0,2]) \to B\mathrm{GL}_1(\mathrm{tmf}).$$

Again one might interpret such twists as twists of tmf-theory by complex super 3-line bundles, or maps to the connected component of $\mathrm{Br}(KU[0,2])$.

Remark 5.27 Recall that, for a commutative ring spectrum R, the base point component of $\mathrm{Br}(R)$ is the space of LMod_R-modules (in Cat_∞) which are equivalent to LMod_R and equivalences between them. Moreover, there is a canonical map of spectra $\mathrm{bgl}_1(R) \to \mathrm{gl}_1(K(R))$, and hence a map of abelian ∞-groups $B^2\mathrm{GL}_1(R) \to B\mathrm{GL}_1(K(R))$, by which we may think of maps $X \to B^2\mathrm{GL}_1(R)$ as $K(R)$-line bundles. Example 5.26 then suggests that our twists $B^2\mathrm{GL}_1(KU[0,2]) \to B\mathrm{GL}_1(\mathrm{tmf})$ are related to $K(KU)$ being a form of elliptic cohomology (see, for instance, [6]).

5.4 Mathematical physics

In this section we describe how our work is connected to work in mathematical physics of Freed, Hopkins and others. In [16, 1.34, 1.38], Freed describes four spectra: $\mathrm{cAlg}_{\mathbb{R}}^{\times}$, $\mathrm{cAlg}_{\mathbb{C}}^{\times}$, $\mathrm{Alg}_{\mathbb{R}}^{\times}$ and $\mathrm{Alg}_{\mathbb{C}}^{\times}$. These are each Picard spectra of Morita 2-categories of invertible algebras, bimodules between them, and intertwiners between bimodules. In the first two cases Freed requires that the bimodule structures and the intertwiners are all continuous with respect to the topologies of $\mathbb{R}$ and $\mathbb{C}$, respectively. In the second two cases, everything is with respect to the discrete topologies on $\mathbb{R}$ and $\mathbb{C}$. Freed computes the homotopy groups and k-invariants of each of these spectra (implicitly using results which are made concrete in [24; 32]). Each of these have a finite number of nonzero homotopy groups, all of which we exhibit below:

$$\begin{aligned} &\pi_{\{0,1,2,3\}}\mathrm{cAlg}_{\mathbb{C}}^{\times} \cong \{\mathbb{Z}/2, \mathbb{Z}/2, 0, \mathbb{Z}\}, && \pi_{\{0,1,2\}}\mathrm{cAlg}_{\mathbb{R}}^{\times} \cong \{\mathbb{Z}/8, \mathbb{Z}/2, \mathbb{Z}/2\}, \\ &\pi_{\{0,1,2\}}\mathrm{Alg}_{\mathbb{C}}^{\times} \cong \{\mathbb{Z}/2, \mathbb{Z}/2, \mathbb{C}^{\times}\}, && \pi_{\{0,1,2\}}\mathrm{Alg}_{\mathbb{R}}^{\times} \cong \{\mathbb{Z}/8, \mathbb{Z}/2, \mathbb{R}^{\times}\}. \end{aligned}$$

In the last two cases, $\mathbb{C}^{\times}$ and $\mathbb{R}^{\times}$ both have the discrete topology. Freed also computes the k-invariants of these spectra to all be nontrivial. Freed's computations when combined with ours (as well as Conjecture 3.38) yield the following:

Theorem 5.28 *There are equivalences of spectra*

$$\mathrm{cAlg}_{\mathbb{C}}^{\times} \simeq \mathrm{pic}_0^3(KU), \quad \mathrm{cAlg}_{\mathbb{R}}^{\times} \simeq \mathrm{pic}_0^2(KO), \quad \mathrm{Alg}_{\mathbb{C}}^{\times}[1,2] \simeq \mathrm{pic}_1^2(K(\mathbb{C})),^1 \quad \mathrm{Alg}_{\mathbb{R}}^{\times}[1,2] \simeq \mathrm{pic}_1^2(K(\mathbb{R})).^1$$

Note that the second and last spectra above are essentially the same because there is no nondiscrete topology to put on $\mathbb{Z}/2$.

In [17, 4.3, 4.4] Freed again introduces these spectra, though with different names (and of course ignoring the difference between $\mathrm{cAlg}_{\mathbb{R}}^{\times}$ and $\mathrm{Alg}_{\mathbb{R}}^{\times}$). Specifically, he writes R^{-1} for $\mathrm{cAlg}_{\mathbb{C}}^{\times}$, $R_{\mathbb{R}/\mathbb{Z}}^{-2}$ for $\mathrm{Alg}_{\mathbb{C}}^{\times}$, and E for $\mathrm{cAlg}_{\mathbb{R}}^{\times}$. However, Freed is also interested in the connective covers of the one-fold desuspensions of these spectra, which he denotes by R^{-2}, $R_{\mathbb{R}/\mathbb{Z}}^{-3}$ and E^{-1}. This immediately implies the following.

Corollary 5.29 *There are equivalences of spectra*

$$R^{-2} \simeq \mathrm{gl}_1(KU)[0,2], \quad R_{\mathbb{R}/\mathbb{Z}}^{-3} \simeq \mathrm{gl}_1(K(\mathbb{C}))[0,1],^2 \quad E^{-1} \simeq \mathrm{gl}_1(KO)[0,1].$$

The spectra R^{-2}, $R_{\mathbb{R}/\mathbb{Z}}^{-3}$ and E^{-1} are explained by Freed to be the Picard spectra of the groupoids of complex superlines, *flat* complex superlines, and real superlines. This should not be surprising in light of the identifications made in Corollaries 3.33 and 3.36.

In [17; 16], Freed identifies $\mathrm{cAlg}_{\mathbb{C}}^{\times}$ with $\Sigma^{-1}ko[1,4]$, which is abstractly equivalent to $\mathrm{pic}_0^3(KU)$, but remarks that he does not have a conceptual reason for this identification. Similarly, for $\mathrm{cAlg}_{\mathbb{R}}^{\times} \simeq \mathrm{Alg}_{\mathbb{R}}^{\times}$, Freed mentions that there is not an "off the shelf" spectrum representing it. We claim that the constructions of this paper, and the identifications made in this section, provide spectra which naturally arise in stable

---

[1] While the nontriviality of the k-invariants of these spectra relies on Conjecture 3.38, they do have the correct homotopy groups in the correct degrees (see [33]).

[2] Again, conjecturally.

homotopy theory (they are "off the shelf") which are the Picard spectra of these Morita categories. Moreover, the constructions of Sections 3.2 and 3.3 essentially prove that the Picard and unit spectra of interest in this paper have the desired geometric interpretations.

5.5 The Anderson dual of $\mathbb{S}$

Recall that there is a spectrum called the *Anderson dual of* $\mathbb{S}$, denoted by $I_{\mathbb{Z}}$, which defines a functor on $\mathcal{S}\mathrm{p}$, $\mathrm{Map}(-, I_{\mathbb{Z}}) : \mathcal{S}\mathrm{p}^{\mathrm{op}} \to \mathcal{S}\mathrm{p}$ (see, e.g., [25, §2; 37, 4.3.9]). Given a spectrum E we will write $I_{\mathbb{Z}}E$ for $\mathrm{Map}(E, I_{\mathbb{Z}})$ and refer to this as the *Anderson dual of* E.

We now show that there is a close relationship between the truncated Picard spectrum $\mathrm{pic}_0^3(KU)$ and the Anderson dual of the sphere spectrum. This is consistent with the results of the prior section as the Anderson dual often appears in the mathematical physics work of Freed, Hopkins and others (see, for instance, [18, Hypothesis 5.17; 19, Ansatz 5.26, Theorem 5.27]).

We will be particularly interested in $\Sigma^n(I_{\mathbb{Z}}[-n,\infty))$, the n-fold suspension of the $-n$-connective cover of $I_{\mathbb{Z}}$. We will simplify notation by writing this as ${}_nI_{\mathbb{Z}}$.

Lemma 5.30 *The spectrum* ${}_3I_{\mathbb{Z}}$ *has homotopy groups* $\pi_0 = \mathbb{Z}/2$, $\pi_1 = \mathbb{Z}/2$, $\pi_2 = 0$, $\pi_3 = \mathbb{Z}$.

Proof The uppermost homotopy groups of $I_{\mathbb{Z}}$ (which is coconnective) are well known, see again [37, 4.3.9], and the result follows by suspending. □

Lemma 5.31 *The unique* k*-invariant of the spectrum* ${}_2I_{\mathbb{Z}}$ *is* $\beta\,\mathrm{Sq}^2$ *and the bottom* k*-invariant of* ${}_3I_{\mathbb{Z}}$ *is* Sq^2.

Proof By [37, 4.2.7(1)] the functor $\mathrm{Map}(-, I_{\mathbb{Z}})$ is a contravariant equivalence on the full subcategory of spectra with finitely many homotopy groups all of which are finitely generated. The first k-invariant of $\mathbb{S}$ is nontrivial and therefore so is the map $\mathrm{Sq}^2 \circ \rho : H\mathbb{Z} \to \Sigma^2 H\mathbb{Z}/2$. This is also of course the k-invariant of $\mathbb{S}[0,1]$. It follows that the uppermost k-invariant of $I_{\mathbb{Z}}$, hence the k-invariant of $I_{\mathbb{Z}}[-2,0]$, or equivalently of ${}_2I_{\mathbb{Z}}$, is a nontrivial map $H\mathbb{Z}/2 \to \Sigma^3 H\mathbb{Z}$, and therefore must be $\beta\,\mathrm{Sq}^2$, the generator of $H^3(H\mathbb{Z}/2;\mathbb{Z}) \cong \mathbb{Z}/2$. By a similar argument the bottom k-invariant of ${}_3I_{\mathbb{Z}}$ is nontrivial and therefore Sq^2. □

Theorem 5.32 *There is an equivalence of spectra* $\mathrm{pic}_0^3(KU) \simeq {}_3I_{\mathbb{Z}}$.

Proof The spectra have the same homotopy groups and both have Sq^2 as their bottom k-invariant. The calculations of Lemmas 4.11 and 4.12 and Propositions 4.13 and 4.14 proceed identically for ${}_3I_{\mathbb{Z}}$ and show that although it has two possible second k-invariants they yield equivalent spectra, both of which must be equivalent to $\mathrm{pic}_0^3(KU)$. □

Corollary 5.33 *There is an equivalence of spectra* $\mathrm{gl}_1(KU[0,2]) \simeq {}_2I_{\mathbb{Z}}$.

We recall the hypothetical characterization of topological field theories given by Freed and Hopkins [16, Ansatz 5.26]:

Ansatz 5.34 *A continuous, invertible, n-dimensional extended topological field theory with symmetry group H_n is a map*

$$\phi : \Sigma^n \mathrm{MTH}_n \to \Sigma^{n+1}(I_{\mathbb{Z}}[-n, 0]),$$

where MTH_n *is the Madsen–Tillman spectrum for H_n of* [22].

If we believe Ansatz 5.34 then a topological field theory determines a map, for a manifold X, $\phi^* : \mathrm{MTH}^n_n(X) \to {}_n I^0_{\mathbb{Z}}(X)$. The domain of this map is, by [22], a set of submersions $E \to X$ with n-dimensional fibers, up to cobordism, and therefore essentially a bundle of cobordism classes of n-dimensional manifolds on X. If we let $n = 1, 2$ then Theorem 5.32 and Corollary 5.8 imply that we have a map whose input is bundles of 1- or 2-dimensional manifolds over X and whose output is either super lines bundles on X or super 2-line bundles on X, which is the behavior one would expect of a fully extended, invertible topological field theory (especially in the case that $X = *$). This also suggests that super n-lines in general should be classified by maps into ${}_n I_{\mathbb{Z}}$ (or rather, should be defined as such, as the authors are not aware of any generally accepted definition of super n-line).

The following conjecture is a 2-local and real version of Theorem 5.32, but the indeterminacy of the k-invariants of $\mathrm{pic}^2_0(KO)$ in our calculations prevents us from proving it. Recall that $\pi_{-4}(I_{\mathbb{Z}}) \cong \mathbb{Z}/24$ which becomes $\mathbb{Z}/8$ after 2-localizing. Therefore the left-hand spectrum below has homotopy groups $\{\mathbb{Z}/8, \mathbb{Z}/2, \mathbb{Z}/2, 0, \mathbb{Z}_{(2)}\}$, which agree with the homotopy groups of the right-hand side.

Conjecture 5.35 *There is an equivalence of spectra* ${}_4(I_{\mathbb{Z}})_{(2)} \simeq \mathrm{pic}^4_0(KO_{(2)})$.

We conclude with an attempt to collect the ideas of Sections 5.4 and 5.5, at least in the complex case, in the following table:

$\mathrm{gl}_1(KU)[0,2] \simeq \mathrm{gl}_1(KU[0,2])$	$\mathrm{pic}^3_0(KU)$	$\mathrm{br}(KU)[0,4]$
n/a	central simple $\mathbb{C}$-superalgebras	central simple $KU[0,2]$-algebras?
invertible (super) $\mathbb{C}$-modules	invertible $\mathrm{sVect}_{\mathbb{C}}$-modules ($\cong$ invertible $KU[0,2]$-modules?)	invertible $\mathrm{LMod}(KU[0,2])$-modules?
complex superlines	complex super 2-lines	complex super 3-lines?
$\Sigma^2(I_{\mathbb{Z}}[-2,\infty))$	$\Sigma^3(I_{\mathbb{Z}}[-3,\infty))$	$\Sigma^4(I_{\mathbb{Z}}[-4,\infty))$?

Certain entries are conjectural, and therefore ended with question marks. Note that the passage from the first row to the second is obtained by applying the functor $\mathrm{LMod}(-)$ and that the third row is essentially just a renaming of the second row.

Acknowledgements

The authors thank Tyler Lawson, Leon Liu, Natalia Pacheco-Tallaj, Eric Peterson (especially Eric Peterson) and Charles Rezk for helpful conversations regarding this material. They are also grateful to the referee for several very helpful recommendations.

References

[1] **A Amabel**, **A Debray**, **P J Haine**, *Differential cohomology*: *categories*, *characteristic classes*, *and connections* (2021) arXiv 2109.12250

[2] **M Ando**, **A J Blumberg**, **D Gepner**, *Twists of K-theory and TMF*, from "Superstrings, geometry, topology, and C^*-algebras" (R S Doran, G Friedman, J Rosenberg, editors), Proc. Sympos. Pure Math. 81, Amer. Math. Soc., Providence, RI (2010) 27–63 MR

[3] **M Ando**, **A J Blumberg**, **D Gepner**, *Parametrized spectra, multiplicative Thom spectra and the twisted Umkehr map*, Geom. Topol. 22:7 (2018) 3761–3825 MR

[4] **M Ando**, **A J Blumberg**, **D Gepner**, **M J Hopkins**, **C Rezk**, *An ∞-categorical approach to R-line bundles, R-module Thom spectra, and twisted R-homology*, J. Topol. 7:3 (2014) 869–893 MR

[5] **M Ando**, **M J Hopkins**, **C Rezk**, *Multiplicative orientations of KO-theory and of the spectrum of topological modular forms*, preprint (2010) https://rezk.web.illinois.edu/koandtmf.pdf

[6] **N A Baas**, **B I Dundas**, **B Richter**, **J Rognes**, *Stable bundles over rig categories*, J. Topol. 4:3 (2011) 623–640 MR

[7] **J Beardsley**, *On bialgebras, comodules, descent data and Thom spectra in ∞-categories*, Homology Homotopy Appl. 25:2 (2023) 219–242 MR

[8] **U Bunke**, **T Nikolaus**, **M Völkl**, *Differential cohomology theories as sheaves of spectra*, J. Homotopy Relat. Struct. 11:1 (2016) 1–66 MR

[9] **J Distler**, **D S Freed**, **G W Moore**, *Spin structures and superstrings*, from "Surveys in differential geometry, XV: Perspectives in mathematics and physics" (T Mrowka, S-T Yau, editors), Surv. Differ. Geom. 15, International Press, Somerville, MA (2011) 99–130 MR

[10] **P Donovan**, **M Karoubi**, *Graded Brauer groups and K-theory with local coefficients*, Inst. Hautes Études Sci. Publ. Math. 38 (1970) 5–25 MR

[11] **V Drinfeld**, *On a notion of ring groupoid* (2021) arXiv 2104.07090

[12] **S Eilenberg**, **S Mac Lane**, *On the groups $H(\Pi, n)$, II*: *Methods of computation*, Ann. of Math. (2) 60 (1954) 49–139 MR

[13] **A Ergus**, *Hopf-algebras and Hopf–Galois extensions in ∞-categories*, PhD thesis, École polytechnique fédérale de Lausanne (2022) https://infoscience.epfl.ch/entities/publication/d5d0d090-2457-4dc1-a6af-56fd0ed1dc4e

[14] **D E Evans**, **U Pennig**, *Equivariant higher Dixmier–Douady theory for circle actions on UHF-algebras*, Adv. Math. 410 (2022) art. id. 108745 MR

[15] **D S Freed**, *Bordism*: *old and new*, lecture notes (2012) https://people.math.harvard.edu/~dafr/M392C-2012/

[16] **D S Freed**, *Lecture notes on k-theory and orientifolds*, lecture notes (2012) https://web.ma.utexas.edu/users/dafr/vienna.pdf

[17] **D S Freed**, *Anomalies and invertible field theories*, from "String-Math 2013" (R Donagi, M R Douglas, L Kamenova, M Roček, editors), Proc. Sympos. Pure Math. 88, Amer. Math. Soc., Providence, RI (2014) 25–45 MR

[18] **D S Freed**, *Short-range entanglement and invertible field theories* (2014) arXiv 1406.7278

[19] **D S Freed**, **M J Hopkins**, *Reflection positivity and invertible topological phases*, Geom. Topol. 25:3 (2021) 1165–1330 MR

[20] **D S Freed**, **M J Hopkins**, **C Teleman**, *Consistent orientation of moduli spaces*, from "The many facets of geometry" (O García-Prada, J P Bourguignon, S Salamon, editors), Oxford Univ. Press (2010) 395–419 MR

[21] **D S Freed**, **M J Hopkins**, **C Teleman**, *Loop groups and twisted K-theory, I*, J. Topol. 4:4 (2011) 737–798 MR

[22] **S Galatius**, **U Tillmann**, **I Madsen**, **M Weiss**, *The homotopy type of the cobordism category*, Acta Math. 202:2 (2009) 195–239 MR

[23] **D Gepner**, **T Lawson**, *Brauer groups and Galois cohomology of commutative ring spectra*, Compos. Math. 157:6 (2021) 1211–1264 MR

[24] **N Gurski**, **N Johnson**, **A M Osorno**, **M Stephan**, *Stable Postnikov data of Picard 2-categories*, Algebr. Geom. Topol. 17:5 (2017) 2763–2806 MR

[25] **D Heard**, **V Stojanoska**, *K-theory, reality, and duality*, J. K-Theory 14:3 (2014) 526–555 MR

[26] **F Hebestreit**, **S Sagave**, *Homotopical and operator algebraic twisted K-theory*, Math. Ann. 378:3-4 (2020) 1021–1059 MR

[27] **M A Hill**, *personal communication* (2022)

[28] **M A Hill**, **M J Hopkins**, **D C Ravenel**, *Equivariant stable homotopy theory and the Kervaire invariant problem*, New Mathematical Monographs 40, Cambridge Univ. Press (2021) MR

[29] **M Hopkins**, **J Lurie**, *On Brauer groups of Lubin–Tate spectra, I*, preprint (2017) https://people.math.harvard.edu/∼lurie/papers/Brauer.pdf

[30] **M J Hopkins**, **I M Singer**, *Quadratic functions in geometry, topology, and M-theory*, J. Differential Geom. 70:3 (2005) 329–452 MR

[31] **M Joachim**, *Higher coherences for equivariant K-theory*, from "Structured ring spectra" (A Baker, B Richter, editors), London Math. Soc. Lecture Note Ser. 315, Cambridge Univ. Press (2004) 87–114 MR

[32] **N Johnson**, **A M Osorno**, *Modeling stable one-types*, Theory Appl. Categ. 26 (2012) art. id. 20 MR

[33] **M Karoubi**, *Bott periodicity in topological, algebraic and Hermitian K-theory*, from "Handbook of K-theory" (E M Friedlander, D R Grayson, editors), Springer (2005) 111–137 MR

[34] **P Kristel**, **M Ludewig**, **K Waldorf**, *2-vector bundles*, High. Struct. 9:1 (2025) 36–87 MR

[35] **M L Laplaza**, *Coherence for distributivity*, from "Coherence in categories" (S MacLane, editor), Lecture Notes in Math. 281, Springer (1972) 29–65 MR

[36] **J Lurie**, *Higher topos theory*, Annals of Mathematics Studies 170, Princeton Univ. Press (2009) MR

[37] **J Lurie**, *Derived algebraic geometry, XIV: Representability theorems*, preprint (2012) https://people.math.harvard.edu/∼lurie/papers/DAG-XIV.pdf

[38] **J Lurie**, *Higher algebra*, preprint (2017) https://url.msp.org/Lurie-HA

[39] **J Lurie**, *Spectral algebraic geometry*, preprint (2018) https://www.math.ias.edu/∼lurie/papers/SAG-rootfile.pdf

[40] **A Mathew**, *Second homotopy group of the mod 2 Moore spectrum* (2012) http://mathoverflow.net/questions/100272/

[41] **J P May**, *E_∞ ring spaces and E_∞ ring spectra*, Lecture Notes in Math. 577, Springer (1977) MR

[42] **D Mertsch**, *Geometric models of twisted K-theory based on bundle gerbes and algebra bundles*, PhD thesis, Universität Greifswald (2020) https://epub.ub.uni-greifswald.de/frontdoor/index/index/docId/4097

[43] **E-k M Moutuou**, *Graded Brauer groups of a groupoid with involution*, J. Funct. Anal. 266:5 (2014) 2689–2739 MR

[44] **E M Parker**, *The Brauer group of graded continuous trace C^*-algebras*, Trans. Amer. Math. Soc. 308:1 (1988) 115–132 MR

[45] **C Redden**, *String structures and canonical 3-forms*, Pacific J. Math. 249:2 (2011) 447–484 MR

[46] **C Rezk**, *When are homotopy colimits compatible with homotopy base change?*, preprint (2014) https://rezk.web.illinois.edu/i-hate-the-pi-star-kan-condition.pdf

[47] **J Rognes**, *Galois extensions of structured ring spectra: stably dualizable groups*, Mem. Amer. Math. Soc. 898, Amer. Math. Soc., Providence, RI (2008) MR

[48] **G Segal**, *Equivariant K-theory*, Inst. Hautes Études Sci. Publ. Math. 34 (1968) 129–151 MR

[49] **C A Weibel**, *An introduction to homological algebra*, Cambridge Studies in Advanced Mathematics 38, Cambridge Univ. Press (1994) MR

JONATHAN BEARDSLEY jbeardsley@unr.edu
Department of Mathematics and Statistics, University of Nevada, Reno, Reno, NV, United States

KIRAN LUECKE kiranl@illinois.edu
Department of Mathematics, University of Illinois Urbana-Champaign, Urbana, IL, United States

JACK MORAVA
Department of Mathematics, The Johns Hopkins University, Baltimore, MD, United States

Received: January 31, 2024 Revised: April 23, 2025

Algebraic & Geometric Topology **26**:5 (2026) 1781–1800
DOI: 10.2140/agt.2026.26.1781
Published: May 29, 2026

Polyhedral coproducts

STEVEN AMELOTTE, WILLIAM HORNSLIEN AND LEWIS STANTON

Dualising the construction of a polyhedral product, we introduce the notion of a *polyhedral coproduct* as a certain homotopy limit over the face poset of a simplicial complex. We begin a study of the basic properties of polyhedral coproducts, surveying the Eckmann–Hilton duals of various familiar examples and properties of polyhedral products. In particular, we show that polyhedral coproducts give a functorial interpolation between the wedge and cartesian product of spaces which differs from the one given by polyhedral products, and we establish a general loop space decomposition for these spaces which is dual to the suspension splitting of a polyhedral product due to Bahri, Bendersky, Cohen and Gitler.

1 Introduction

Polyhedral products are natural subspaces of cartesian products defined as certain colimits over the face poset of a finite simplicial complex K. This construction generalises and unifies into a common combinatorial framework many familiar methods of constructing new topological spaces from given ones — for example, products, wedge sums, joins, half-smash products and the fat wedge construction are all special cases. Since their introduction by Bahri, Bendersky, Cohen and Gitler [2], the topology of polyhedral products has become a growing topic of investigation within homotopy theory and has made fruitful contact with many other areas of mathematics. Notable examples include toric topology, following Buchstaber and Panov's [4] formulation of moment-angle complexes and Davis–Januszkiewicz spaces as polyhedral products; commutative algebra, where polyhedral products give geometric realisations of Stanley–Reisner rings and their Tor algebras; and geometric group theory, where polyhedral products model the classifying spaces of right-angled Artin and Coxeter groups. For more on the history and far-reaching applications of polyhedral products, we recommend the excellent survey [1].

Motivated by the ubiquity and utility of polyhedral products, the purpose of this paper is to propose a definition for the dual notion of a *polyhedral coproduct* and begin a study of its basic properties. Before describing the main results, we first review the construction of polyhedral products more precisely.

Let $(\underline{X}, \underline{A}) = \{(X_i, A_i)\}_{i=1}^m$ be an m-tuple of pointed CW-pairs, $\mathbf{SCpx}_m$ be the category of simplicial complexes on the vertex set $[m] = \{1, \dots, m\}$ with morphisms given by simplicial inclusions, and $\mathbf{Top}_*$ be the category of pointed topological spaces. For a simplicial complex K, let $\operatorname{cat}(K)$ denote the face poset of K, regarded as a small category with objects given by faces $\sigma \in K$ and morphisms given by face inclusions $\tau \subset \sigma$. We denote the initial object of $\operatorname{cat}(K)$ by $\varnothing$, which corresponds to the empty face of K.

MSC2020: primary 55P10, 55P35; secondary 18A30, 55P30.

The *polyhedral product* associated to $(\underline{X}, \underline{A})$ is the functor

$$(\underline{X}, \underline{A})^{(-)} : \mathbf{SCpx}_m \to \mathbf{Top}_*,$$

which associates to each simplicial complex K the (homotopy) colimit

$$(\underline{X}, \underline{A})^K = \operatorname*{hocolim}_{\mathrm{cat}(K)} \prod_{i=1}^{m} Y_i(\sigma),$$

where $Y_i : \mathrm{cat}(K) \to \mathbf{Top}_*$ is defined for each $i \in [m]$ by

$$Y_i(\sigma) = \begin{cases} X_i & \text{if } i \in \sigma, \\ A_i & \text{if } i \notin \sigma. \end{cases}$$

As has been pointed out in [13; 16; 21], for example, the homotopy colimit above agrees up to homotopy with the usual colimit $\bigcup_{\sigma \in K} \prod_{i=1}^{m} Y_i(\sigma)$ since each (X_i, A_i) is an NDR-pair. In particular, each map in the diagram defining the polyhedral product is a cofibration, and the polyhedral product $(\underline{X}, \underline{A})^K$ is a cellular subcomplex of $\prod_{i=1}^{m} X_i$ for all K. In the case that $A_i = *$ for all $i \in [m]$, this subcomplex $(\underline{X}, \underline{*})^K$ naturally interpolates between the wedge $\bigvee_{i=1}^{m} X_i$ (when K consists of m disjoint vertices) and the product $\prod_{i=1}^{m} X_i$ (when $K = \Delta^{m-1}$ is the simplex on m vertices).

Dualising the definition of a polyhedral product as a homotopy colimit of products, we define a polyhedral coproduct as a homotopy limit of coproducts, as follows.[1]

Definition 1.1 Let $\underline{f} = (f_1, \dots, f_m)$ be an m-tuple of maps $f_i : X_i \to A_i$ of pointed spaces. Define the *polyhedral coproduct* associated to $\underline{f}$ as the functor

$$\underline{f}_{\mathrm{co}}^{(-)} : \mathbf{SCpx}_m \to \mathbf{Top}_*,$$

which associates to each simplicial complex K with m vertices the homotopy limit

$$\underline{f}_{\mathrm{co}}^{K} = \operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}} D(\sigma)$$

of a diagram $D : \mathrm{cat}(K)^{\mathrm{op}} \to \mathbf{Top}_*$, where $D(\sigma) = \bigvee_{i=1}^{m} Y_i(\sigma)$ and

$$Y_i(\sigma) = \begin{cases} X_i & \text{if } i \in \sigma, \\ A_i & \text{if } i \notin \sigma. \end{cases}$$

Note that for a face inclusion $\tau \subset \sigma \in K$, there are maps $Y_i(\sigma) \to Y_i(\tau)$ defined for each $i \in [m]$ by f_i if $i \in \sigma \backslash \tau$ and by the identity map otherwise, and hence there is an induced map

$$\bigvee_{i=1}^{m} Y_i(\sigma) \to \bigvee_{i=1}^{m} Y_i(\tau).$$

For a family $(\underline{X}, \underline{A})$ of pairs of spaces, if the maps $f_i : X_i \to A_i$ are clear from context, we will sometimes denote $\underline{f}_{\mathrm{co}}^{K}$ by $(\underline{X}, \underline{A})_{\mathrm{co}}^{K}$. One example is the case that $A_i = *$ is a point, and f_i is the constant

---

[1] Since we work only with homotopy limits, we define polyhedral coproducts with respect to any set of maps f_i, rather than insisting on fibrations.

map for all $i \in [m]$. In this case, as we show in Section 2, the polyhedral coproduct $(\underline{X}, \underline{*})^K_{\mathrm{co}}$ naturally interpolates between $\bigvee_{i=1}^m X_i$ (when $K = \Delta^{m-1}$) and $\prod_{i=1}^m X_i$ (when K is m disjoint vertices).

For polyhedral products, the relationship between the combinatorics of K and the homotopy type of the space $(\underline{X}, \underline{*})^K$ interpolating between the m-fold wedge and m-fold product is made clear after suspending. By [3, Theorem 2.15], there is a natural homotopy equivalence

$$\Sigma(\underline{X}, \underline{*})^K \simeq \bigvee_{\sigma \in K} \Sigma X^{\wedge\sigma}, \tag{1}$$

where $X^{\wedge\sigma} = X_{i_1} \wedge \cdots \wedge X_{i_k}$ for each face $\sigma = \{i_1, \dots, i_k\} \in K$. Notice that this generalises the well-known splitting of $\Sigma\bigl(\prod_{i=1}^m X_i\bigr)$ when $K = \Delta^{m-1}$, in which case the wedge above is indexed over all subsets of the vertex set $[m]$. For polyhedral coproducts, we dualise the suspension splitting (1) by establishing a loop space decomposition for $(\underline{X}, \underline{*})^K_{\mathrm{co}}$, which similarly generalises a product decomposition due to Porter [18] for $\Omega\bigl(\bigvee_{i=1}^m X_i\bigr)$ when $K = \Delta^{m-1}$.

Theorem 1.2 (Theorem 4.3) *Let K be a simplicial complex on $[m]$ and let $X_1, \dots, X_m$ be pointed, simply connected CW-complexes. Then there is a homotopy equivalence*

$$\Omega(\underline{X}, \underline{*})^K_{\mathrm{co}} \simeq \prod_{i=1}^m \Omega X_i \times \prod_{b \in I} \Omega\Sigma\Bigl(\bigwedge_{\tau \in \mathcal{F}} ((\Omega X)^{\wedge\tau})^{\wedge b(\tau)}\Bigr).$$

As with the splitting (1), the indexing set I above is defined in terms of the faces of the simplicial complex K, and made explicit in Section 4. The equivalence (1) is a special case of the more general Bahri–Bendersky–Cohen–Gitler splitting (henceforth, BBCG splitting) which identifies the homotopy type of any polyhedral product $(\underline{X}, \underline{A})^K$ as a certain wedge after suspending once. In the case that each X_i is contractible, the authors of [3] obtain their splitting using a lemma regarding homotopy colimits of certain diagrams due to Welker, Ziegler and Živaljević [21, Proposition 3.5]. We first dualise the Welker–Ziegler–Živaljević lemma (see Lemma 3.7), and then use this to dualise the BBCG splitting to obtain the following result. This is a simplified version of the full statement of Theorem 4.5 where the indexing sets are made explicit in terms of the underlying simplicial complex.

Theorem 1.3 (Theorem 4.5) *Let K be a simplicial complex on $[m]$ and let $f_i : X_i \to A_i$ be a map of pointed, simply connected CW-complexes where X_i is contractible for $1 \le i \le m$. Then there is a homotopy equivalence*

$$\Omega \underline{f}^K_{\mathrm{co}} \simeq \prod_{b \in I} \Omega\mathrm{Map}_*(\Sigma|K_{I_b}|, \Sigma\Omega A_1^{\wedge l_1(b)} \wedge \cdots \wedge \Omega A_m^{\wedge l_m(b)}).$$

Both Theorems 1.2 and 1.3 are special cases of a general loop space decomposition of an arbitrary polyhedral coproduct (see Theorem 4.2), dual to the suspension splitting of a polyhedral product.

Definition 1.1 is alternate to Theriault's definition of a *dual polyhedral product*, which was introduced in [20] and used to identify the Lusternik–Schnirelmann cocategory of a simply connected space X with the homotopy nilpotency of its loop space ΩX. Although the two notions coincide in some special cases (see Remark 2.2), the diagrams defining polyhedral coproducts and dual polyhedral products are very different in general, and our definition is more suitable for dualising the BBCG splitting of $\Sigma(\underline{X}, \underline{A})^K$ (see Section 4).

Although we restrict our attention to constructions in $\mathbf{Top}_*$ in this paper, note that polyhedral (co)products could be defined more generally in any model category $\mathcal{C}$, for example, by replacing the category of pointed spaces with $\mathcal{C}$ in the definitions above. Since any (closed) model category has an initial object and a terminal object, the polyhedral products and coproducts of the form $(\underline{X}, \underline{*})^K$ and $(\underline{X}, \underline{*})^K_{\mathrm{co}}$ can be defined in this setting to yield functorial interpolations between the categorical product and coproduct in $\mathcal{C}$.

2 Basic properties

2.1 Basic examples

We begin by computing some basic examples of polyhedral coproducts, in each case illustrating the Eckmann–Hilton duality between these constructions and their corresponding polyhedral products.

Example 2.1 (the $\underline{A} = \underline{*}$ case) (1) Let K be m disjoint vertices. In this case, the polyhedral product associated to the m-tuple of pairs $(\underline{X}, \underline{*}) = \{(X_i, *)\}_{i=1}^m$ is the wedge

$$(\underline{X}, \underline{*})^K \simeq X_1 \vee \cdots \vee X_m.$$

Dually, if $f_i : X_i \to *$ is the constant map for each $i = 1, \dots, m$, then by definition the corresponding polyhedral coproduct is given by

$$(\underline{X}, \underline{*})^K_{\mathrm{co}} \simeq X_1 \times \cdots \times X_m.$$

(2) On the other extreme, let $K = \Delta^{m-1}$. The polyhedral product associated to $(\underline{X}, \underline{*})$ in this case is

$$(\underline{X}, \underline{*})^K \simeq X_1 \times \cdots \times X_m.$$

Since the diagram defining $(\underline{X}, \underline{*})^K_{\mathrm{co}}$ has an initial object corresponding to the maximal face of the simplex Δ^{m-1},

$$(\underline{X}, \underline{*})^K_{\mathrm{co}} \simeq X_1 \vee \cdots \vee X_m.$$

(3) Let $K = \partial\Delta^{m-1}$. The polyhedral product $(\underline{X}, \underline{*})^K$ in this case is precisely the *fat wedge* of the spaces $X_1, \dots, X_m$, which is defined as

$$\mathrm{FW}(X_1, \dots, X_m) = \{(x_1, \dots, x_m) \mid x_i = * \text{ for at least one } i\}.$$

Dual to the fat wedge is the *thin product* of $X_1, \dots, X_m$, as defined by Hovey in [10, Definition 1]. This construction is realised by the polyhedral coproduct $(\underline{X}, \underline{*})^K_{\mathrm{co}}$.

Remark 2.2 The dual polyhedral product, denoted by $(\underline{X}, \underline{A})^K_D$, defined by Theriault [20] also models some of the spaces in Example 2.1. In particular, when K is m disjoint points, $(\underline{X}, \underline{*})^K_D$ is equal to the thin product of $X_1, \dots, X_m$. When $K = \partial\Delta^{m-1}$, we have $(\underline{X}, \underline{*})^K_D \simeq X_1 \times \cdots \times X_m$. Outside of these cases, it is not clear whether there is any correspondence between the dual polyhedral product, and the polyhedral coproduct. Theriault also used the dual polyhedral product to give a loop space decomposition of the thin product. An alternate loop space decomposition of the thin product can be recovered in the context of polyhedral coproducts by Theorem 4.3.

Just like the polyhedral product $(\underline{X}, \underline{*})^K$, the polyhedral coproduct $(\underline{X}, \underline{*})^K_{\mathrm{co}}$ interpolates between the categorical product $X_1 \times \cdots \times X_m$ and coproduct $X_1 \vee \cdots \vee X_m$ as K interpolates between a discrete set of vertices and a full simplex. Next, we compute two further examples of $\underline{f}^K_{\mathrm{co}}$ where the m-tuple $\underline{f}$ involves maps other than the constant map $X_i \to *$. An important class of polyhedral products (which includes generalised moment-angle complexes $(D^n, S^{n-1})^K$) is given by those associated to CW-pairs $(\underline{CX}, \underline{X}) = \{(CX_i, X_i)\}_{i=1}^m$ consisting of cones and their bases. The first example below dualises this case by replacing the cofibrations $X_i \hookrightarrow CX_i$ with path space fibrations $PX_i \to X_i$.

Example 2.3 (dual of the join) Let $K = \partial\Delta^1$ be two disjoint vertices so that the only faces of K are $\varnothing$, $\{1\}$ and $\{2\}$, and its face poset is given by $\{1\} \leftarrow \varnothing \rightarrow \{2\}$. In this case the polyhedral product $(\underline{CX}, \underline{X})^K$ recovers the join of X_1 and X_2 as a pushout:

$$(\underline{CX}, \underline{X})^K = CX_1 \times X_2 \cup_{X_1 \times X_2} X_1 \times CX_2 \simeq X_1 \star X_2.$$

For $i \in \{1, 2\}$, let $f_i : PX_i \to X_i$ be the path space fibration over X_i. Since each PX_i is contractible, the polyhedral coproduct $\underline{f}^K_{\mathrm{co}} = (\underline{PX}, \underline{X})^K_{\mathrm{co}} = \mathrm{holim}(PX_1 \vee X_2 \to X_1 \vee X_2 \leftarrow X_1 \vee PX_2)$ agrees with the homotopy limit of the middle column of the commutative diagram

$$\begin{tikzcd}
* \arrow[r] \arrow[d] & X_1 \arrow[r] \arrow[d] & X_1 \arrow[d] \\
F \arrow[r] & X_1 \vee X_2 \arrow[r] & X_1 \times X_2 \\
* \arrow[r] \arrow[u] & X_2 \arrow[r] \arrow[u] & X_2 \arrow[u]
\end{tikzcd}$$

where the vertical maps are inclusions and the rows are homotopy fibrations. The homotopy limit of the right column is contractible, so by taking homotopy limits of the columns we obtain a homotopy equivalence $(\underline{PX}, \underline{X})^K_{\mathrm{co}} \simeq \Omega F$. By [6, p. 302], $F \simeq \Sigma(\Omega X_1 \wedge \Omega X_2)$, and so there is a homotopy equivalence $(\underline{PX}, \underline{X})^K_{\mathrm{co}} \simeq \Omega\Sigma(\Omega X_1 \wedge \Omega X_2)$. This space is known as the *cojoin* of X_1 and X_2.

Example 2.4 (dual of the half-smash) Let $K = \partial\Delta^1$ be two disjoint vertices and let $(\underline{X}, \underline{A})$ consist of the CW-pairs $\{(X_1, A_1), (X_2, A_2)\} = \{(CX, X), (Y, *)\}$. As in the previous example, the polyhedral product is a pushout $(\underline{X}, \underline{A})^K = CX \times * \cup_{X \times *} X \times Y$. Since $CX \times *$ is contractible, this is simply the cofibre of the inclusion $X \times * \hookrightarrow X \times Y$, which by definition is the *half-smash product*

$$(\underline{X}, \underline{A})^K \simeq X \ltimes Y.$$

To dualise this example, let $\underline{f} = (f_1, f_2)$ where $f_1 : PX \to X$ is the path space fibration and $f_2 : Y \to *$ is the constant map. Then by definition, the polyhedral coproduct is given by

$$\underline{f}^K_{\mathrm{co}} = \operatorname{holim}(PX \vee * \to X \vee * \leftarrow X \vee Y) \simeq \operatorname{hofib}(X \vee Y \xrightarrow{\pi_X} X),$$

the expected Eckmann–Hilton dual of the cofibre $(\underline{X}, \underline{A})^K = \operatorname{hocofib}(X \xrightarrow{i_X} X \times Y)$ above. The homotopy fibre of the projection onto a wedge summand can be identified using Mather's cube lemma [14], and we therefore obtain that the dual of the half-smash is given by

$$\underline{f}^K_{\mathrm{co}} \simeq \Omega X \ltimes Y.$$

Moreover, by Mather's cube lemma or [6, Theorem 1.1], there is a homotopy fibration

$$\Sigma\Omega X \wedge \Omega Y \to \Omega X \ltimes Y \to Y,$$

where the right map is the projection map. The projection has a right homotopy inverse, implying there is a homotopy equivalence

$$\Omega(\Omega X \ltimes Y) \simeq \Omega Y \times \Omega(\Sigma\Omega X \wedge \Omega Y).$$

This result can be recovered in the context of polyhedral coproducts by Theorem 4.2.

2.2 Functorial properties

The polyhedral product is a bifunctor (see [3, Remark 2.3]). Namely, it defines a functor from the category of (m-tuples of) CW-pairs to the category of CW-complexes, and it also defines a functor from the category of simplicial complexes to the category of CW-complexes. In this section, we prove that the polyhedral coproduct enjoys similar functorial properties. First, we show naturality with respect to maps of spaces.

Theorem 2.5 *Let K be a simplicial complex on $[m]$. For $1 \le i \le m$, let $f_i : X_i \to A_i$ and $f_i' : X_i' \to A_i'$ be maps. If there are maps $g_i : X_i \to X_i'$ and $h_i : A_i \to A_i'$ such that the diagram*

$$\begin{tikzcd} X_i \arrow[r, "g_i"] \arrow[d, "f_i"] & X_i' \arrow[d, "f_i'"] \\ A_i \arrow[r, "h_i"] & A_i' \end{tikzcd} \tag{2}$$

homotopy commutes, then there is an induced map $\underline{f}^K_{\mathrm{co}} \to \underline{f}'^K_{\mathrm{co}}$.

Proof Let D_K and D_K' be the diagrams defining $\underline{f}^K_{\mathrm{co}}$ and $\underline{f}'^K_{\mathrm{co}}$, respectively. For a face $\sigma \in K$, define a map $F_\sigma : D_K(\sigma) \to D_K'(\sigma)$, defined by

$$F_\sigma : D_K(\sigma) = \bigvee_{i=1}^m Y_i \xrightarrow{\bigvee_{i=1}^m \phi_i} \bigvee_{i=1}^m Y_i' = D_K'(\sigma),$$

where $\phi_i = g_i$ if $i \in \sigma$, and $\phi_i = h_i$ if $i \notin \sigma$. By (2), F_σ induces a natural transformation $D_K \to D_K'$, which in turn induces a map $\underline{f}^K_{\mathrm{co}} \to \underline{f}'^K_{\mathrm{co}}$. □

The definition of $\underline{f}^K_{\mathrm{co}}$ is also natural with respect to simplicial inclusions.

Theorem 2.6 *Let K be a simplicial complex on $[m]$, and let L be a subcomplex of K on $[n]$ with $n \leq m$. Then the simplicial inclusion $L \to K$ induces a map $\underline{f}^K_{\mathrm{co}} \to \underline{f}^L_{\mathrm{co}}$.*

Proof Let D_K and D_L be the diagrams defining $\underline{f}^K_{\mathrm{co}}$ and $\underline{f}^L_{\mathrm{co}}$, respectively. Let $D^K_L : \mathrm{cat}(L)^{\mathrm{op}} \to \mathbf{Top}_*$ be the diagram with $D^K_L(\sigma) = \bigvee_{i=1}^m Y_i(\sigma)$, where $Y_i(\sigma) = X_i$ if $i \in \sigma$, and $Y_i(\sigma) = A_i$ if $i \notin \sigma$. By definition of $\underline{f}^K_{\mathrm{co}}$ as a homotopy limit, there are canonical maps $\underline{f}^K_{\mathrm{co}} \to D^K_L(\sigma)$ for all $\sigma \in L$, and so the inclusion $\mathrm{cat}(L)^{\mathrm{op}} \to \mathrm{cat}(K)^{\mathrm{op}}$ induces a map $\underline{f}^K_{\mathrm{co}} \to \mathrm{holim}_{\mathrm{cat}(L)^{\mathrm{op}}} D^K_L$.

Now define a natural transformation of diagrams $D^K_L \to D_L$ by the pinch map

$$D^K_L(\sigma) = \bigvee_{i=1}^m Y_i \to \bigvee_{i=1}^n Y_i = D_L(\sigma).$$

This induces a map $\mathrm{holim}_{\mathrm{cat}(L)^{\mathrm{op}}} D^K_L \to \underline{f}^L_{\mathrm{co}}$. Therefore, the simplicial inclusion induces the composite

$$\underline{f}^K_{\mathrm{co}} \to \operatorname*{holim}_{\mathrm{cat}(L)^{\mathrm{op}}} D^K_L \to \underline{f}^L_{\mathrm{co}}. \qquad \square$$

Remark 2.7 The map $\underline{f}^K_{\mathrm{co}} \to \underline{f}^L_{\mathrm{co}}$ can be represented as the homotopy limit of a map of diagrams $D_K \to D_L$. For each $\sigma \in L$, we have a pinch map $D_K(\sigma) = \bigvee_{i=1}^m Y_i \to \bigvee_{i=1}^n Y_i = D_L(\sigma)$. By computing $\mathrm{holim}\, D_K$, one can see that the maps $\underline{f}^K_{\mathrm{co}} \to D_L(\sigma)$ for $\sigma \in L$ are the maps described in the proof of Theorem 2.6.

2.3 Retractions

Let K be a simplicial complex and L a full subcomplex of K. For polyhedral products, by [5, Lemma 2.2.3], there is a map $(\underline{X}, \underline{A})^K \to (\underline{X}, \underline{A})^L$ which is a left inverse for the map $(\underline{X}, \underline{A})^L \to (\underline{X}, \underline{A})^K$. In the case of polyhedral coproducts, there is an analogous statement.

Theorem 2.8 *Let K be a simplicial complex on $[m]$ and L be a full subcomplex of K on $[n]$, with $n < m$. Then there is a right homotopy inverse for the map $\underline{f}^K_{\mathrm{co}} \to \underline{f}^L_{\mathrm{co}}$ induced by the simplicial inclusion $L \to K$.*

Proof Let D_K and D_L be the diagrams defining $\underline{f}^K_{\mathrm{co}}$ and $\underline{f}^L_{\mathrm{co}}$, respectively. Recall from the proof of Theorem 2.6 the diagram D^K_L indexed by $\mathrm{cat}(L)^{\mathrm{op}}$, which is defined by $D^K_L(\sigma) = \bigvee_{i=1}^m Y_i(\sigma)$, where $Y_i(\sigma) = X_i$ if $i \in \sigma$, and $Y_i(\sigma) = A_i$ if $i \notin \sigma$. Define a natural transformation $D_L \to D^K_L$ by the inclusion

$$D_L(\sigma) = \bigvee_{i=1}^n Y_i \hookrightarrow \bigvee_{i=1}^m Y_i = D^K_L(\sigma).$$

This induces a map $\underline{f}^L_{\mathrm{co}} \xrightarrow{f} \mathrm{holim}_{\mathrm{cat}(L)^{\mathrm{op}}} D^K_L$. Define a functor $F : \mathrm{cat}(K)^{\mathrm{op}} \to \mathrm{cat}(L)^{\mathrm{op}}$ by sending $\sigma \in K$ to the face $\tau \in L$, where τ is obtained from σ by removing any instances of the vertices $\{n+1, \dots, m\}$. Since L is a full subcomplex, F is well defined. The functor F induces a map $\mathrm{holim}_{\mathrm{cat}(L)^{\mathrm{op}}} D^K_L \xrightarrow{g} \underline{f}^K_{\mathrm{co}}$. Therefore, we obtain a composite

$$\underline{f}^L_{\mathrm{co}} \xrightarrow{f} \operatorname*{holim}_{\mathrm{cat}(L)^{\mathrm{op}}} D^K_L \xrightarrow{g} \underline{f}^K_{\mathrm{co}}.$$

Now consider the composite

$$\phi : \underline{f}^L \xrightarrow{f} \operatorname*{holim}_{\mathrm{cat}(L)^{\mathrm{op}}} D_L^K \xrightarrow{g} \underline{f}^K_{\mathrm{co}} \xrightarrow{h} \operatorname*{holim}_{\mathrm{cat}(L)^{\mathrm{op}}} D_L^K \xrightarrow{k} \underline{f}^L_{\mathrm{co}},$$

where the composite $k \circ h : \underline{f}^K_{\mathrm{co}} \to \underline{f}^L_{\mathrm{co}}$ is defined as in Theorem 2.6. By definition of the functor F, the composite $\mathrm{cat}(L)^{\mathrm{op}} \hookrightarrow \mathrm{cat}(K)^{\mathrm{op}} \xrightarrow{F} \mathrm{cat}(L)^{\mathrm{op}}$ is the identity, and so the composite $h \circ g$ is the identity. For a face σ, the natural transformation inducing the composite $k \circ f$ is the identity on $D_L^K(\sigma)$, and so $k \circ f$ is the identity. Hence, ϕ is the identity map, and so the composite $g \circ f$ is a right homotopy inverse for the map induced by $L \to K$. □

2.4 Homotopy cofibrations

For polyhedral products, if K is a simplicial complex on $[m]$, it was shown in [5, Lemma 2.3.1] that there exists a homotopy fibration

$$(\underline{C\Omega X}, \underline{\Omega X})^K \to (\underline{X}, \underline{*})^K \to \prod_{i=1}^m X_i,$$

which splits after looping. More generally, it was shown in [8, Theorem 2.1] that there is a homotopy fibration

$$(\underline{CY}, \underline{Y})^K \to (\underline{X}, \underline{A})^K \to \prod_{i=1}^m X_i,$$

where Y_i is the homotopy fibre of the inclusion $A_i \to X_i$. Note that the map $(\underline{X}, \underline{A})^K \to \prod_{i=1}^m X_i$ is induced by the inclusion $K \to \Delta^{m-1}$. Moreover, the second homotopy fibration above also splits after looping, giving a homotopy equivalence

$$\Omega(\underline{X}, \underline{A})^K \simeq \prod_{i=1}^m \Omega X_i \times \Omega(\underline{CY}, \underline{Y})^K.$$

This implies that to understand the loop spaces of polyhedral products, and therefore their homotopy groups, it suffices to study polyhedral products of the form $\Omega(\underline{CY}, \underline{Y})^K$. Loop space decompositions of certain polyhedral products of this form have been studied, for example, in [17; 19].

For polyhedral coproducts, by considering the map induced by the inclusion $K \to \Delta^{m-1}$, one might hope there is a homotopy cofibration

$$\bigvee_{i=1}^m X_i \to (\underline{X}, \underline{*})^K_{\mathrm{co}} \to (\underline{P\Sigma X}, \underline{\Sigma X})^K_{\mathrm{co}},$$

or more generally,

$$\bigvee_{i=1}^m X_i \to (\underline{X}, \underline{A})^K_{\mathrm{co}} \to (\underline{PY}, \underline{Y})^K_{\mathrm{co}},$$

where Y_i is the homotopy cofibre of $f_i : X_i \to A_i$. This would allow us to understand the suspension of polyhedral coproducts, and therefore their homology. However, if K is two disjoint points, by part (1) of

Example 2.1, the map $X_1 \vee X_2 \to (\underline{X}, \underline{*})^K_{\mathrm{co}}$ is the inclusion $X_1 \vee X_2 \to X_1 \times X_2$. The homotopy cofibre of this map is $X_1 \wedge X_2$, but Example 2.3 implies that $(\underline{P\Sigma X}, \underline{\Sigma X})^K_{\mathrm{co}} \simeq \Omega\Sigma(\Omega\Sigma X_1 \wedge \Omega\Sigma X_2)$. This is reminiscent of how Ganea's theorem [6, Theorem 1.1] does not dualise canonically; see [6, Remark 3.5]. This gives rise to the following problem.

Problem 2.9 For certain classes of polyhedral coproducts, find a decomposition for their suspensions.

3 Preliminary results

3.1 Preliminary decompositions

To decompose the loop space of a polyhedral coproduct, we will use a result known as the Porter decomposition. Let K be m disjoint points. By [5, Lemma 2.3.1], there is a homotopy fibration

$$(\underline{C\Omega X}, \underline{\Omega X})^K \to \bigvee_{i=1}^m X_i \xrightarrow{i} \prod_{i=1}^m X_i. \tag{3}$$

A result of Porter [18, Theorem 1] identifies the homotopy type of $(\underline{C\Omega X}, \underline{\Omega X})^K$ in the case that each X_i is simply connected. For a pointed space X and $k \geq 1$, let $X^{\vee k}$ be the k-fold wedge of X.

Theorem 3.1 *Let $X_1, \dots, X_m$ be pointed, simply connected CW-complexes, and let K be m disjoint points. There is a homotopy equivalence*

$$(\underline{C\Omega X}, \underline{\Omega X})^K \simeq \bigvee_{k=2}^m \bigvee_{1 \leq i_1 < \cdots < i_k \leq m} (\Sigma\Omega X_{i_1} \wedge \cdots \wedge \Omega X_{i_k})^{\vee(k-1)}.$$

Moreover, this homotopy equivalence is natural for maps $X_i \to Y_i$. □

There is a special case of the naturality in Theorem 3.1 which will be important. Let $n < m$ and let $Y_i = X_i$ for $1 \leq i \leq n$, and let $Y_i = CX_i$ for $n+1 \leq i \leq m$. In this case, we obtain the following.

Proposition 3.2 *Let $n < m$, and let $X_1, \dots, X_m$ be pointed, simply connected CW-complexes. There is a homotopy commutative diagram*

$$\begin{tikzcd}
\bigvee_{k=2}^m \bigvee_{1 \leq i_1 < \cdots < i_k \leq m} \Sigma(\Omega X_{i_1} \wedge \cdots \wedge \Omega X_{i_k})^{\vee(k-1)} \arrow[r] \arrow[d, "p'"] & \bigvee_{i=1}^m X_i \arrow[r, hook] \arrow[d, "p"] & \prod_{i=1}^m X_i \arrow[d, "\pi"] \\
\bigvee_{k=2}^n \bigvee_{1 \leq i_1 < \cdots < i_k \leq n} \Sigma(\Omega X_{i_1} \wedge \cdots \wedge \Omega X_{i_k})^{\vee(k-1)} \arrow[r] & \bigvee_{i=1}^n X_i \arrow[r, hook] & \prod_{i=1}^n X_i
\end{tikzcd}$$

where p and p' are pinch maps and π is the projection. □

After looping the homotopy fibration (3), there is a natural right homotopy inverse s of the canonical inclusion i, given by multiplying the inclusions $X_i \to \prod_{i=1}^m X_i$. The naturality of s and the homotopy fibration in Theorem 3.1 imply the following.

Theorem 3.3 *Let $X_1, \dots, X_m$ be pointed, simply connected spaces. There is a homotopy equivalence*

$$\Omega\Big(\bigvee_{i=1}^{m} X_i\Big) \simeq \prod_{i=1}^{m} \Omega X_i \times \Omega\Big(\bigvee_{k=2}^{m} \bigvee_{1\le i_1<\cdots<i_k\le m} (\Sigma\Omega X_{i_1} \wedge \cdots \wedge \Omega X_{i_k})^{\vee(k-1)}\Big).$$

Moreover, this homotopy equivalence is natural for maps $X_i \to Y_i$. □

For a subset $I = \{i_1, \dots, i_k\} \subseteq [m]$ and pointed spaces $X_1, \dots, X_m$, define $X^{\wedge I} = X_{i_1} \wedge \cdots \wedge X_{i_k}$.

Remark 3.4 In Theorem 3.3, the wedge summand in the right-hand product term can be indexed as

$$\bigvee_{\substack{I\subseteq[m]\\ |I|\ge 2}} (\Sigma(\Omega X)^{\wedge I})^{\vee(|I|-1)}.$$

Now we recall the Hilton–Milnor theorem. Let L be the free (ungraded) Lie algebra over $\mathbb{Z}$ on the elements $x_1, \dots, x_m$, and let B be a Hall basis of L. For a bracket $b \in B$, let $k_i(b)$ be the number of instances of x_i in b. For a space X and $k \ge 0$, denote by $X^{\wedge k}$ the k-fold smash of X. The following is from [15, Theorem 4], which is a generalisation of [9, Theorem A]. We will define the 0-fold smash of X to be omission of the corresponding term, rather than a trivial space.

Theorem 3.5 *Let $X_1, \dots, X_m$ be connected topological spaces. Then there is a homotopy equivalence*

$$\Omega\Big(\bigvee_{i=1}^{m} \Sigma X_i\Big) \simeq \prod_{b\in B} \Omega\Sigma(X_1^{\wedge k_1(b)} \wedge \cdots \wedge X_m^{\wedge k_m(b)}).$$

Moreover, this homotopy equivalence is natural for maps $X_i \to Y_i$. □

As in the case of the Porter decomposition, there is a special case which will be important. Let $n < m$ and let $Y_i = X_i$ for $1 \le i \le n$, and let $Y_i = CX_i$ for $n+1 \le i \le m$. By contracting out the CX_i terms, we obtain the following.

Corollary 3.6 *Let $n < m$, let B_n be a Hall basis on the free Lie algebra generated by $x_1, \dots, x_n$, and let B_m be a Hall basis on the free Lie algebra generated by $x_1, \dots, x_m$. Then the diagram*

$$\begin{tikzcd}
\Omega\Sigma(\bigvee_{i=1}^{m} X_i) \arrow[r, "\simeq"] \arrow[d, "\Omega p"] & \prod_{b\in B_m} \Omega\Sigma(X_1^{\wedge k_1(b)} \wedge \cdots \wedge X_m^{\wedge k_m(b)}) \arrow[d, "\pi"] \\
\Omega\Sigma(\bigvee_{i=1}^{n} X_i) \arrow[r, "\simeq"'] & \prod_{b\in B_n} \Omega\Sigma(X_1^{\wedge k_1(b)} \wedge \cdots \wedge X_n^{\wedge k_n(b)})
\end{tikzcd}$$

homotopy commutes. □

3.2 Preliminary homotopy limit decompositions

In this section, we prove some decompositions of certain homotopy limits indexed by the opposite of the face category of a simplicial complex. The first lemma is the dual statement of the "wedge lemma" from [21, Proposition 3.5].

Lemma 3.7 *Let K be a simplicial complex. Let X be a space and let $\mathcal{D}:\mathrm{cat}(K)^{\mathrm{op}}\to\mathbf{Top}_*$ be a diagram with $\mathcal{D}(\varnothing)=X$ and $\mathcal{D}(\sigma)=*$ for all $\sigma\neq\varnothing$. Then*

$$\operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}}\mathcal{D}\simeq\mathrm{Map}_*(\Sigma|K|,X).$$

Proof Let $\mathrm{cat}(K)^{\mathrm{op}}_{>\varnothing}$ denote the subposet category consisting of all $\sigma\in K$ where $\sigma\neq\varnothing$. For a topological space A, let $\underline{A}:\mathrm{cat}(K)^{\mathrm{op}}\to\mathbf{Top}_*$ be the diagram with constant value A. The opposite category $(\mathrm{cat}(K)^{\mathrm{op}}_{>\varnothing})^{\mathrm{op}}$ is the category $\mathrm{cat}(K)_{<\varnothing}$ and its geometric realisation coincides with the realisation of K as a topological space

$$|K|\simeq\operatorname*{hocolim}_{\mathrm{cat}(K)_{<\varnothing}}\underline{*}.$$

Thus there are homotopy equivalences

$$\mathrm{Map}(|K|,X)\simeq\mathrm{Map}(\operatorname*{hocolim}_{\mathrm{cat}(K)_{<\varnothing}}\underline{*},X)\simeq\operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}_{>\varnothing}}\mathrm{Map}(*,X)\simeq\operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}_{>\varnothing}}\underline{X}.$$

Let $\mathcal{X}:\mathrm{cat}(K)^{\mathrm{op}}\to\mathbf{Top}_*$ be the diagram with $\mathcal{X}(\varnothing)=*$ and $\mathcal{X}(\sigma)=X$ for all $\sigma\neq\varnothing$. The diagram $\mathcal{D}$ can be written as the (homotopy) pullback

$$\underline{*}\to\mathcal{X}\leftarrow\underline{X},\tag{4}$$

where the right-hand map is the constant map to the basepoint for $\sigma=\varnothing$, and the identity on X for $\sigma\neq\varnothing$. The left-hand map is the inclusion of the basepoint for each $\sigma\in K$.

Let $\mathcal{Z}$ be the category with two objects, 1 and 2, and a morphism $f:2\to1$ in addition to the identity morphisms. Consider the diagram $\mathcal{Y}:\mathrm{cat}(K)^{\mathrm{op}}_{>\varnothing}\times\mathcal{Z}\to\mathbf{Top}_*$, with $\mathcal{Y}(\sigma,1)=*$ and $\mathcal{Y}(\sigma,2)=X$ for all $\sigma\neq\varnothing$. Let $F_1:\mathrm{cat}(K)^{\mathrm{op}}_{>\varnothing}\times\mathcal{Z}\to\mathrm{cat}(K)^{\mathrm{op}}$ be the functor sending $(\sigma,1)\mapsto\varnothing$ and $(\sigma,2)\mapsto\sigma$. Let $F_2:\mathrm{cat}(K)^{\mathrm{op}}_{>\varnothing}\times\mathcal{Z}\to\mathrm{cat}(\Delta^0)^{\mathrm{op}}$ be the functor sending $(\sigma,1)\mapsto1$ and $(\sigma,2)\mapsto2$ for all $\sigma\neq\varnothing$. The right Kan extension of $\mathcal{Y}$ along F_1 is $\mathcal{X}$ and the right Kan extension of $\mathcal{Y}$ along F_2 is the diagram

$$\mathrm{Map}(|K|,X)\to\mathrm{Map}(|K|,*).$$

Since homotopy limits are preserved by right Kan extensions, we obtain the equivalences

$$\operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}}\mathcal{X}\simeq\operatorname{holim}(\mathrm{Map}(|K|,X)\to\mathrm{Map}(|K|,*))\simeq\mathrm{Map}(|K|,\operatorname{holim}(X\to*))\simeq\mathrm{Map}(|K|,X).$$

Recall that the diagram $\mathcal{D}$ was equivalent to diagram (4). Using that $\operatorname{holim}_{\mathrm{cat}(K)^{\mathrm{op}}}\underline{*}$ is contractible and the previous observations about $\mathcal{X}$ yields the homotopy equivalence

$$\operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}}\mathcal{D}\simeq\operatorname{holim}(*\to\mathrm{Map}(|K|,X)\leftarrow X).\tag{5}$$

The map $X\to\mathrm{Map}(|K|,X)$ is a section for the evaluation map

$$\mathrm{ev}_k:\mathrm{Map}(|K|,X)\to X,$$

where $k \in |K|$. In particular, there is a homotopy fibration diagram

$$\begin{tikzcd}
\Omega\mathrm{Map}_*(|K|, X) \arrow[r] \arrow[d, "\simeq"] & * \arrow[r] \arrow[d] & \mathrm{Map}_*(|K|, X) \arrow[d] \\
F \arrow[r] & X \arrow[r] \arrow[d, equal] & \mathrm{Map}(|K|, X) \arrow[d, "\mathrm{ev}_k"] \\
 & X \arrow[r, equal] & X
\end{tikzcd}$$

where the top right square is a homotopy pullback and F is the homotopy limit of (5). Hence, there are homotopy equivalences

$$\operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}} \mathcal{D} \simeq F \simeq \Omega\mathrm{Map}_*(|K|, X) \simeq \mathrm{Map}_*(|K|, \Omega X) \simeq \mathrm{Map}_*(\Sigma|K|, X). \qquad \square$$

Lemma 3.8 *Let K be a simplicial complex on $[m]$. Let $I \subseteq [m]$, and let $\mathcal{D} : \mathrm{cat}(K)^{\mathrm{op}} \to \mathbf{Top}_*$ be a diagram. Suppose that all maps induced by $\sigma \subset \tau$, where σ is obtained from τ by removing a single vertex not contained in I, are identity maps. Then the homotopy limit of $\mathcal{D}$ is equivalent to the homotopy limit of a diagram $\mathcal{D}' : \mathrm{cat}(K_I) \to \mathbf{Top}_*$, where $\mathcal{D}'(\sigma_I) = \mathcal{D}(\sigma)$.*

Proof The inclusion of $K_I \subset K$ induces a map of face categories $i : \mathrm{cat}(K_I)^{\mathrm{op}} \to \mathrm{cat}(K)^{\mathrm{op}}$. The diagram $\mathcal{D}$ is the right Kan extension of $\mathcal{D}'$ along i. Hence, $\mathcal{D}$ and $\mathcal{D}'$ have the same homotopy limit since right Kan extensions preserve homotopy limits. $\square$

4 Loop spaces of polyhedral coproducts

4.1 A general loop space decomposition

In [3, Definition 2.2], for a simplicial complex K, a construction known as the *polyhedral smash product* is defined and denoted by $\widehat{(\underline{X}, \underline{A})}^K$. By [3, Theorem 2.10], there is a homotopy equivalence

$$\Sigma(\underline{X}, \underline{A})^K \simeq \bigvee_{I \subseteq [m]} \Sigma\widehat{(\underline{X}, \underline{A})}^{K_I}.$$

In this subsection, we show a dual statement for polyhedral coproducts. Recall that for spaces X and Y there is a (homotopy) cofibration $X \vee Y \to X \times Y \to X \wedge Y$. To dualise this, by [6, p. 302], there is a homotopy fibration $\Sigma(\Omega X \wedge \Omega Y) \to X \vee Y \to X \times Y$. This motivates the following definition.

Definition 4.1 The *polyhedral smash coproduct* is defined as the homotopy limit

$$\underline{\hat{f}}^K_{\mathrm{co}} = \operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}} \Sigma\widehat{D}, \qquad \text{where } \widehat{D} = \bigwedge_{i=1}^{m} \Omega Y_i \quad \text{and} \quad Y_i(\sigma) = \begin{cases} X_i & \text{if } i \in \sigma, \\ A_i & \text{if } i \notin \sigma. \end{cases}$$

For a set of positive integers $N = \{k_1(N), \dots, k_m(N)\}$, we define *the weighted polyhedral smash coproduct* as

$$\underline{\hat{f}}^K_{N,\mathrm{co}} = \operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}} \Sigma \widehat{D}^N, \qquad \text{where } \widehat{D}^N = \bigwedge_{i=1}^m (\Omega Y_i)^{\wedge k_i(N)} \quad \text{and} \quad Y_i(\sigma) = \begin{cases} X_i & \text{if } i \in \sigma, \\ A_i & \text{if } i \notin \sigma. \end{cases}$$

Before stating the result, we set up some notation which will be used throughout the rest of Section 4. For a subset $I = \{i_1, \dots, i_k\} \subseteq [m]$, let S_I be the set

$$\{a_{J,i} \mid J \subseteq I,\ |J| \ge 2,\ 1 \le i \le |J|-1\}.$$

Denote by B_I a Hall basis of the free ungraded Lie algebra on the set S_I. For a bracket $b \in B_I$ and $J \subseteq I$, let $b(J)$ be the sum of the number of instances of $a_{J,i}$ in b for each $1 \le i \le |J|-1$. For $1 \le i \le m$ and a bracket $b \in B_{[m]}$, define

$$l_i(b) := \sum_{\substack{I \subseteq [m] \\ i \in I}} b(I),$$

which counts the number of instances of each vertex i in the faces in b. Let $L_b = (l_1(b), \dots, l_m(b))$. For any $I \subseteq [m]$ and $b \in B_{[m]}$, define

$$I_b := I \cap \{j \mid 1 \le j \le m,\ l_j(b) \ne 0\}.$$

This set contains the elements of I which appear in the subsets in b. To ensure that ΩX_i is connected in order to apply Theorem 3.5, we need the hypothesis that each X_i is simply connected.

Theorem 4.2 *Let $f_i : X_i \to A_i$ be a map of pointed, simply connected CW-complexes for all $1 \le i \le m$. There is a homotopy equivalence*

$$\Omega \underline{f}^K_{\mathrm{co}} \simeq \prod_{i=1}^m \Omega X_i \times \prod_{b \in B_{[m]}} \Omega \underline{\hat{f}}^{K_{I_b}}_{L_b,\mathrm{co}}.$$

Proof Since loops commute with homotopy limits, $\Omega \underline{f}^K_{\mathrm{co}} \simeq \operatorname{holim}_{\mathrm{cat}(K)^{\mathrm{op}}} \Omega D$. By Theorem 3.3 and Remark 3.4, ΩD decomposes as a product

$$\Omega D \simeq \prod_{i=1}^m \Omega Y_i \times \Omega\Sigma\Bigl(\bigvee_{\substack{I \subseteq [m] \\ |I| \ge 2}} ((\Omega Y)^{\wedge I})^{\vee(|I|-1)} \Bigr).$$

We can apply the Hilton–Milnor theorem (Theorem 3.5) to the right-hand product term to obtain the equivalence

$$\Omega\Bigl(\bigvee_{\substack{I \subseteq [m] \\ |I| \ge 2}} \Sigma((\Omega Y)^{\wedge I})^{\vee(|I|-1)} \Bigr) \simeq \prod_{b \in B_{[m]}} \Omega\Sigma\Bigl(\bigwedge_{\substack{I \subseteq [m] \\ |I| \ge 2}} ((\Omega Y)^{\wedge I})^{\wedge b(I)} \Bigr).$$

Note that for any $b \in B_{[m]}$, by definition

$$\Sigma\Bigl(\bigwedge_{\substack{I\subseteq[m]\\|I|\geq 2}}((\Omega Y)^{\wedge I})^{\wedge b(I)}\Bigr)=\Sigma\widehat{D}^{L_b}.$$

Using the loop space decomposition of D we obtain a new description of $\Omega\underline{f}^K_{\mathrm{co}}$:

$$\begin{aligned}\Omega\underline{f}^K_{\mathrm{co}}\simeq\operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}}\Omega D&\simeq\operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}}\Bigl(\prod_{i=1}^{m}\Omega Y_i\times\prod_{b\in B_{[m]}}\Omega\Sigma\widehat{D}^{L_b}\Bigr)\\&\simeq\prod_{i=1}^{m}(\operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}}\Omega Y_i)\times\prod_{b\in B_{[m]}}(\operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}}\Omega\Sigma\widehat{D}^{L_b}).\end{aligned}$$

Fix $i\in[m]$ and consider the term $\operatorname{holim}_{\mathrm{cat}(K)^{\mathrm{op}}}\Omega Y_i$. By Lemma 3.8 in the case $I=\{i\}$, there is an equivalence

$$\operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}}\Omega Y_i\simeq\operatorname{holim}(\Omega X_i\to\Omega A_i)\simeq\Omega X_i.$$

For any $b\in B_{[m]}$ and any pair of simplices $\sigma\subset\tau$ where σ is obtained from τ by removing a vertex not in I_b, the induced maps $\Omega\Sigma\widehat{D}^{L_b}(\sigma)\to\Omega\Sigma\widehat{D}^{L_b}(\tau)$ are identity maps. Therefore, Lemma 3.8 implies

$$\operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}}\Omega\Sigma\widehat{D}^{L_b}\simeq\operatorname*{holim}_{\mathrm{cat}(K_{I_b})^{\mathrm{op}}}\Omega\Sigma\widehat{D}^{L_b}\simeq\Omega\underline{\hat{f}}^{K_{I_b}}_{L_b,\mathrm{co}}.\qquad\square$$

4.2 Loop space decompositions of $(\underline{X},\underline{*})^K_{\mathrm{co}}$

For polyhedral products of the form $(\underline{X},\underline{*})^K$, by [3, Theorem 2.15], there is a homotopy equivalence

$$\Sigma(\underline{X},\underline{*})^K\simeq\bigvee_{\sigma\in K}\Sigma X^{\wedge\sigma}.$$

In this subsection, we prove a dual statement for polyhedral coproducts of the form $(\underline{X},\underline{*})^K_{\mathrm{co}}$. Let $\mathcal{F}$ and $\mathcal{M}$ be the set of faces and maximal faces of K on two or more vertices, respectively. The following result could be shown using Theorem 4.2 by showing that certain polyhedral smash coproducts are contractible in this case. However, this would require a technical argument involving choices of vector space bases for free Lie algebras. To avoid these technicalities, and make clearer the connection to Hall bases, we provide a proof using Corollary 3.6.

Theorem 4.3 *Let $X_1,\dots,X_m$ be pointed, simply connected CW-complexes. There is a homotopy equivalence*

$$\Omega(\underline{X},\underline{*})^K_{\mathrm{co}}\simeq\prod_{i=1}^{m}\Omega X_i\times\prod_{b\in\bigcup_{\sigma\in\mathcal{M}}B_\sigma}\Omega\Sigma\Bigl(\bigwedge_{\tau\in\mathcal{F}}((\Omega X)^{\wedge\tau})^{\wedge b(\tau)}\Bigr).$$

Proof By definition of the polyhedral coproduct,

$$(\underline{X},\underline{*})^K_{\mathrm{co}}=\operatorname*{holim}_{\mathrm{cat}(K)^{\mathrm{op}}}D,$$

where, if $\sigma = \{i_1, \dots, i_k\}$, $D(\sigma) = \bigvee_{j=1}^{k} X_{i_j}$, and for each $\sigma' \subset \sigma$, the map $D(\sigma) \to D(\sigma')$ is given by the pinch map $p : \bigvee_{i \in \sigma} X_i \to \bigvee_{j \in \sigma'} X_j$. Since looping commutes with homotopy limits, we obtain a homotopy equivalence $\Omega(\underline{X}, \underline{*})^K_{\mathrm{co}} \simeq \operatorname{holim}_{\mathrm{cat}(K)^{\mathrm{op}}} \Omega D$. By Theorem 3.3 and Remark 3.4, there is a natural homotopy equivalence

$$\Omega\Big(\bigvee_{j=1}^{k} X_{i_j}\Big) \simeq \prod_{j=1}^{k} \Omega X_{i_j} \times \Omega\Big(\bigvee_{\substack{\tau\subseteq\sigma\\ |\tau|\geq 2}} (\Sigma(\Omega X)^{\wedge\tau})^{\vee|\tau|-1}\Big). \tag{6}$$

Under this equivalence, it follows from Proposition 3.2 that for $\sigma' \subseteq \sigma$, the maps $\Omega D(\sigma) \to \Omega D(\sigma')$ are given by $\pi \times \Omega p'$ up to homotopy, where π is the projection and p' is the pinch map.

Applying the Hilton–Milnor theorem to the second factor on the right-hand side in (6), we obtain a natural homotopy equivalence

$$\Omega\Big(\bigvee_{\substack{\tau\subseteq\sigma\\ |\tau|\geq 2}} (\Sigma(\Omega X)^{\wedge\tau})^{\vee|\tau|-1}\Big) \simeq \prod_{b\in B_\sigma} \Omega\Sigma\Big(\bigwedge_{\substack{\tau\subseteq\sigma\\ |\tau|\geq 2}} ((\Omega X)^{\tau})^{\wedge b(\tau)}\Big).$$

By Corollary 3.6, for $\sigma' \subseteq \sigma$, there is a homotopy commutative diagram

$$\begin{tikzcd}
\Omega\big(\bigvee_{\tau\subseteq\sigma,|\tau|\geq 2}(\Sigma(\Omega X)^{\wedge\tau})^{\vee|\tau|-1}\big) \arrow[r, "\simeq"] \arrow[d, "\Omega p'"] & \prod_{b\in B_\sigma} \Omega\Sigma\big(\bigwedge_{\tau\subseteq\sigma,|\tau|\geq 2}((\Omega X)^{\tau})^{\wedge b(\tau)}\big) \arrow[d, "\pi'"] \\
\Omega\big(\bigvee_{\tau'\subseteq\sigma',|\tau'|\geq 2}(\Sigma(\Omega X)^{\wedge\tau'})^{\vee|\tau'|-1}\big) \arrow[r, "\simeq"] & \prod_{b\in B_{\sigma'}} \Omega\Sigma\big(\bigwedge_{\tau'\subseteq\sigma',|\tau'|\geq 2}((\Omega X)^{\tau'})^{\wedge b(\tau')}\big)
\end{tikzcd}$$

where π' is the projection. Summarising, for $\sigma' \subseteq \sigma$, there is a homotopy commutative diagram

$$\begin{tikzcd}
\Omega\big(\bigvee_{i\in\sigma} X_i\big) \arrow[r, "\simeq"] \arrow[d, "\Omega p"] & \prod_{i\in\sigma} X_i \times \prod_{b\in B_\sigma} \Omega\Sigma\big(\bigwedge_{\tau\subseteq\sigma,|\tau|\geq 2}((\Omega X)^{\tau})^{\wedge b(\tau)}\big) \arrow[d, "\pi''"] \\
\Omega\big(\bigvee_{j\in\sigma'} X_j\big) \arrow[r, "\simeq"] & \prod_{j\in\sigma'} X_j \times \prod_{b\in B_{\sigma'}} \Omega\Sigma\big(\bigwedge_{\tau'\subseteq\sigma',|\tau'|\geq 2}((\Omega X)^{\tau'})^{\wedge b(\tau')}\big)
\end{tikzcd}$$

Hence $\Omega(\underline{X}, \underline{*})^K_{\mathrm{co}} \simeq \operatorname{holim}_{\mathrm{cat}(K)^{\mathrm{op}}} \Omega D$ is the product of each of the distinct factors that appear in the diagram. For $\sigma' \subseteq \sigma$, the product terms appearing in the decomposition for $\Omega D(\sigma)$ strictly contain the product terms in the decomposition for $\Omega D(\sigma')$. Therefore, enumerating the distinct factors that appear for the maximal faces, we obtain the desired equivalence. □

Example 4.4 Let K be a 1-dimensional simplicial complex on $[m]$. In this case, the set $\mathcal{M}$ consists of all the 1-simplices in K. For each $\sigma = \{i, j\} \in \mathcal{M}$, we have $B_\sigma = \{\sigma\}$. Therefore, Theorem 4.3 implies there is a homotopy equivalence

$$\Omega(\underline{X}, \underline{*})^K_{co} \simeq \prod_{i=1}^{m} \Omega X_i \times \prod_{\sigma\in\mathcal{M}} \Omega\Sigma(\Omega X_i \wedge \Omega X_j).$$

4.3 Loop space decompositions when the domain is contractible

For a simplicial complex K, let $|K|$ be the geometric realisation of K as a topological space. For polyhedral products of the form $(\underline{CX}, \underline{X})^K$, by [3, Theorem 2.21], there is a homotopy equivalence

$$\Sigma(\underline{CX}, \underline{X})^K \simeq \bigvee_{I \notin K} \Sigma(|K_I| \wedge X^{\wedge I}). \tag{7}$$

In this subsection, we prove a dual statement for polyhedral coproducts of the form $\underline{f}^K_{\mathrm{co}}$ where the domain of each f_i is contractible.

Theorem 4.5 *Let K be a simplicial complex on $[m]$ and $f_i : X_i \to A_i$ where X_i is contractible and A_i is a pointed, simply connected CW-complex for $1 \le i \le m$. Then there is a homotopy equivalence*

$$\Omega \underline{f}^K_{\mathrm{co}} \simeq \prod_{\substack{b \in B_{[m]} \\ I_b \notin K}} \Omega \mathrm{Map}_*(\Sigma|K_{I_b}|, \Sigma\Omega A_1^{\wedge l_1(b)} \wedge \cdots \wedge \Omega A_m^{\wedge l_m(b)}).$$

To prove Theorem 4.5, we will use the following consequence of Theorem 4.2.

Lemma 4.6 *Assume that X_i is contractible and A_i is a pointed, simply connected CW-complex for all i and $N \in \mathbb{N}^m$. There is a homotopy equivalence*

$$\hat{\underline{f}}^K_{N,\mathrm{co}} \simeq \mathrm{Map}_*(\Sigma|K|, \Sigma\Omega A_1^{\wedge k_1(N)} \wedge \cdots \wedge \Omega A_m^{\wedge k_m(N)}).$$

Proof By definition, $\hat{D}^N(\varnothing) = \Sigma\Omega A_1^{\wedge k_1(N)} \wedge \cdots \wedge \Omega A_m^{\wedge k_m(N)}$ and $\hat{D}^N(\sigma) \simeq *$ for all $\sigma \neq \varnothing$ since all the X_i are contractible. Thus, we may apply Lemma 3.7 to the diagram defining $\hat{\underline{f}}^K_{N,\mathrm{co}}$ which yields the claimed result. □

With the lemma above, it is straightforward to prove Theorem 4.5

Proof of Theorem 4.5 By Lemma 4.6, if $I_b \in K$, then $\hat{\underline{f}}^{K_{I_b}}_{L_b,\mathrm{co}}$ is contractible. One can then apply Lemma 4.6 to the decomposition in Theorem 4.2 to prove the statement. □

Example 4.7 Let $K = \partial\Delta^{m-1}$. In this case, the only missing face of K is $\{1, \dots, m\}$. By Theorem 4.5, there is a homotopy equivalence

$$\Omega \underline{f}^K_{\mathrm{co}} \simeq \prod_{\substack{b \in B_{[m]} \\ I_b = \{1,\dots,m\}}} \Omega \mathrm{Map}_*(\Sigma|K_{I_b}|, \Sigma\Omega A_1^{\wedge l_1(b)} \wedge \cdots \wedge \Omega A_m^{\wedge l_m(b)}),$$

where the indexing set of the product consists of brackets b such that for each $i \in [m]$, there is a face $\sigma \in K$ in b which contains i.

In the case of polyhedral products, it is known that the decomposition in (7) desuspends in certain cases. For example, when K is a shifted complex [7; 11], a flag complex with chordal 1-skeleton [17, Theorem 6.4], or more generally, a totally fillable simplicial complex [12, Corollary 7.3]. Specialising, polyhedral products of the form $(D^2, S^1)^K$ are known as *moment-angle complexes*, which are denoted by $\mathcal{Z}_K$. In the aforementioned cases, $\mathcal{Z}_K$ is homotopy equivalent to a wedge of spheres.

Consider the case where K is a simplicial complex on $[m]$, and is either a shifted complex, or a flag complex with chordal 1-skeleton. The dual of the polyhedral product $(\underline{CX}, \underline{X})^K$ is the polyhedral coproduct $(\underline{PX}, \underline{X})^K_{\mathrm{co}}$. In the first case, $|K_I|$ is homotopy equivalent to a wedge of spheres for all $I \subseteq [m]$, and in the second case, $|K_I|$ is homotopy equivalent to a set of disjoint points for all $I \subseteq [m]$. Therefore, in the case where each X_i is a simply connected sphere, Theorem 4.5 implies $\Omega(\underline{PX}, \underline{X})^K_{\mathrm{co}}$ is homotopy equivalent to a product of iterated loop spaces of spheres. Dual to the polyhedral product case, we give the next conjecture.

Conjecture 4.8 Let K be a shifted complex or a flag complex with chordal 1-skeleton. Then the decomposition in Theorem 4.5 deloops.

5 Polyhedral coproducts under operations on simplicial complexes

5.1 Joins of simplicial complexes

For any polyhedral product, if $K = K_1 \star K_2$ is the join of K_1 and K_2, then $(\underline{X}, \underline{A})^K \cong (\underline{X}, \underline{A})^{K_1} \times (\underline{X}, \underline{A})^{K_2}$. Therefore, we may expect a homotopy equivalence $(\underline{X}, \underline{A})^K_{\mathrm{co}} \simeq (\underline{X}, \underline{A})^{K_1}_{\mathrm{co}} \vee (\underline{X}, \underline{A})^{K_2}_{\mathrm{co}}$. However, this does not hold in general for polyhedral coproducts.

For $1 \le i \le 4$, let $X_i = \mathbb{C}P^\infty$, and let $K = \{1,2\} \star \{3,4\}$ be the boundary of a square. Since $(\mathbb{C}P^\infty, *)^{\{1,2\}}_{\mathrm{co}}$ and $(\mathbb{C}P^\infty, *)^{\{3,4\}}_{\mathrm{co}}$ are homotopy equivalent to $\mathbb{C}P^\infty \times \mathbb{C}P^\infty$ by Example 2.1, suppose that $(\mathbb{C}P^\infty, *)^K_{\mathrm{co}} \simeq (\mathbb{C}P^\infty \times \mathbb{C}P^\infty) \vee (\mathbb{C}P^\infty \times \mathbb{C}P^\infty)$. Since the space $\mathbb{C}P^\infty$ is simply connected and $\Omega\mathbb{C}P^\infty \simeq S^1$, by Theorem 4.3, there is a homotopy equivalence

$$\Omega(\mathbb{C}P^\infty, *)^K_{\mathrm{co}} \simeq \prod_{i=1}^{4} (S^1 \times \Omega S^3).$$

Now by Theorem 3.3 applied to $(\mathbb{C}P^\infty \times \mathbb{C}P^\infty) \vee (\mathbb{C}P^\infty \times \mathbb{C}P^\infty)$, there is a homotopy equivalence

$$\Omega((\mathbb{C}P^\infty \times \mathbb{C}P^\infty) \vee (\mathbb{C}P^\infty \times \mathbb{C}P^\infty)) \simeq \prod_{i=1}^{4} S^1 \times \Omega\Sigma((S^1 \times S^1) \wedge (S^1 \times S^1)).$$

For spaces X and Y, there is a well-known homotopy equivalence $\Sigma(X \times Y) \simeq \Sigma X \vee \Sigma Y \vee \Sigma(X \wedge Y)$. By shifting the suspension coordinate, we obtain homotopy equivalences

$$\prod_{i=1}^{4} S^1 \times \Omega\Sigma((S^1 \vee S^1 \vee S^2) \wedge (S^1 \vee S^1 \vee S^2)) \simeq \prod_{i=1}^{4} S^1 \times \Omega\Sigma\Big(\bigvee_{i=1}^{4} S^2 \vee \bigvee_{i=1}^{4} S^3 \vee S^4\Big).$$

By Theorem 3.5, $\Omega\Sigma\big(\bigvee_{i=1}^{4} S^2 \vee \bigvee_{i=1}^{4} S^3 \vee S^4\big)$ decomposes as an infinite, finite-type product of spheres and loops on spheres. However, since $\Omega(\mathbb{C}P^\infty, *)^K_{\mathrm{co}}$ is homotopy equivalent to a finite product of spheres and loops on spheres,

$$\Omega(\mathbb{C}P^\infty, *)^K_{\mathrm{co}} \not\simeq \Omega((\mathbb{C}P^\infty \times \mathbb{C}P^\infty) \vee (\mathbb{C}P^\infty \times \mathbb{C}P^\infty)),$$

which implies that

$$(\mathbb{C}P^\infty, *)^K_{\mathrm{co}} \not\simeq (\mathbb{C}P^\infty \times \mathbb{C}P^\infty) \vee (\mathbb{C}P^\infty \times \mathbb{C}P^\infty).$$

However, it is possible to say something about certain joins. Let K be a simplicial complex on $[m]$ and for $1 \leq i \leq m$, let (X_i, A_i) be CW-pairs. If (X_{m+1}, A_{m+1}) is a CW-pair where X_{m+1} is contractible, then there are homotopy equivalences $(\underline{X}, \underline{A})^{K*\{m+1\}} \cong (\underline{X}, \underline{A})^{K} \times X_{m+1} \simeq (\underline{X}, \underline{A})^{K}$. The following dualises this case.

Proposition 5.1 *Let K be a simplicial complex on the vertex set $[m]$ and let $\underline{f}^{K}_{\mathrm{co}}$ be any polyhedral coproduct. Let $K' = K \star \{m+1\}$ where $f_{m+1} : * \to Y$ for some space Y. Then $\underline{f}^{K'}_{\mathrm{co}} \simeq \underline{f}^{K}_{\mathrm{co}}$.*

Proof There is an equivalence of categories $\mathrm{cat}(K')^{\mathrm{op}} \cong \mathrm{cat}(K)^{\mathrm{op}} \times \mathrm{cat}(\{m+1\})^{\mathrm{op}}$ and a projection map $p : \mathrm{cat}(K')^{\mathrm{op}} \to \mathrm{cat}(K)^{\mathrm{op}}$. Geometrically, the map p deletes vertex $m+1$ from any simplex in K'. Let $D' : \mathrm{cat}(K')^{\mathrm{op}} \to \mathbf{Top}_*$ (resp. $D : \mathrm{cat}(K)^{\mathrm{op}} \to \mathbf{Top}_*$) be the diagram defining $\underline{f}^{K'}_{\mathrm{co}}$ (resp. $\underline{f}^{K}_{\mathrm{co}}$). By right Kan extending D' along p, we recover D. Note that had we not placed the assumption on the domain of f_{m+1} then the Kan extension would not recover D. Since right Kan extensions preserve homotopy limits, we get the equivalence $\underline{f}^{K'}_{\mathrm{co}} \simeq \underline{f}^{K}_{\mathrm{co}}$. □

5.2 Pullbacks of polyhedral coproducts

Let K_1 be a simplicial complex on $\{1, \dots, n\}$ and K_2 be a simplicial complex on $\{l, \dots, m\}$ with $n < m$ and $l \leq m$, and let L be a subcomplex (possibly empty) of K_1 and K_2 on $\{l, \dots, n\}$. Define $K = K_1 \cup_L K_2$, and for M one of K_1, K_2 or L, let $\overline{M}$ be the simplicial complex considered on the vertex set $\{1, \dots, m\}$. For polyhedral products, by [7, Proposition 3.1], there is a pushout

$$\begin{tikzcd}
(\underline{X}, \underline{A})^{\overline{L}} \arrow[r] \arrow[d] & (\underline{X}, \underline{A})^{\overline{K_1}} \arrow[d] \\
(\underline{X}, \underline{A})^{\overline{K_2}} \arrow[r] & (\underline{X}, \underline{A})^{K}
\end{tikzcd}$$

For polyhedral coproducts, we can prove a dual statement.

Proposition 5.2 *Let K_1 be a simplicial complex on $\{1, \dots, n\}$ and K_2 be a simplicial complex on $\{l, \dots, m\}$ with $n < m$ and $l \leq m$, and let L be a subcomplex (possibly empty) of K_1 and K_2 on $\{l, \dots, n\}$. Define $K = K_1 \cup_L K_2$. Then there is a homotopy pullback of polyhedral coproducts*

$$\begin{tikzcd}
\underline{f}^{K}_{\mathrm{co}} \arrow[r] \arrow[d] & \underline{f}^{\overline{K_2}}_{\mathrm{co}} \arrow[d] \\
\underline{f}^{\overline{K_1}}_{\mathrm{co}} \arrow[r] & \underline{f}^{\overline{L}}_{\mathrm{co}}
\end{tikzcd}$$

where the maps $\underline{f}^{\overline{K_1}}_{\mathrm{co}} \to \underline{f}^{\overline{L}}_{\mathrm{co}}$ and $\underline{f}^{\overline{K_2}}_{\mathrm{co}} \to \underline{f}^{\overline{L}}_{\mathrm{co}}$ are induced by the simplicial inclusions.

Proof Let $\mathcal{C}$ be the category associated to the diagram

$$\mathrm{cat}(K_1)^{\mathrm{op}} \to \mathrm{cat}(L)^{\mathrm{op}} \leftarrow \mathrm{cat}(K_2)^{\mathrm{op}},$$

where the maps are induced by the inclusions of L into K_1 and K_2. By Remark 2.7, one may write the elements of the pullback

$$\begin{tikzcd} & \underline{f}^{\overline{K_2}}_{\mathrm{co}} \arrow[d] \\ \underline{f}^{\overline{K_1}}_{\mathrm{co}} \arrow[r] & \underline{f}^{\overline{L}}_{\mathrm{co}} \end{tikzcd}$$

as diagrams and we are left with a diagram $D' : \mathcal{C} \to \mathbf{Top}_*$. For each $\sigma \in L$, let $\sigma_{K_1}, \sigma_{K_2}$ denote the copies in $\mathcal{C}$. The objects of $\mathcal{C}$ are the same as $\mathrm{cat}(K)^{\mathrm{op}}$, but with three copies of each $\sigma \in L$. For each $\sigma \in L$, the maps $D'(\sigma_{K_1}) \to D'(\sigma)$ and $D(\sigma_{K_2}) \to D(\sigma)$ are the identity map. Let $D : \mathrm{cat}(K)^{\mathrm{op}} \to \mathbf{Top}_*$ be the diagram defining $\underline{f}^{K}_{\mathrm{co}}$. There's a projection map $p : \mathcal{C} \to \mathrm{cat}(K)^{\mathrm{op}}$ collapsing the tripled simplices. The right Kan extension of D' along p recovers the diagram D. As right Kan extensions preserve homotopy limits, the homotopy pullback diagram has the desired limit. □

Let K_1 and K_2 be simplicial complexes and let $K = K_1 \sqcup K_2$. By definition of the polyhedral product, $(\underline{X}, \underline{*})^K = (\underline{X}, \underline{*})^{K_1} \vee (\underline{X}, \underline{*})^{K_2}$. In the case of a polyhedral coproduct $(\underline{X}, \underline{*})^K_{\mathrm{co}}$, using Proposition 5.2, we show that the dual holds.

Theorem 5.3 *Let K_1 and K_2 be simplicial complexes, and let $K = K_1 \sqcup K_2$. There is a homotopy equivalence*

$$(\underline{X}, \underline{*})^K_{\mathrm{co}} \simeq (\underline{X}, \underline{*})^{K_1}_{\mathrm{co}} \times (\underline{X}, \underline{*})^{K_2}_{\mathrm{co}}.$$

Proof By definition, since each $A_i = *$, we have $(\underline{X}, \underline{*})^{\varnothing} = *$ and $(\underline{X}, \underline{*})^{\overline{K_i}} = (\underline{X}, \underline{*})^{K_i}$ for $i \in \{1, 2\}$. Therefore, Proposition 5.2 implies there is a homotopy pullback

$$\begin{tikzcd} (\underline{X}, \underline{*})^K_{\mathrm{co}} \arrow[r] \arrow[d] & (\underline{X}, \underline{*})^{K_2}_{\mathrm{co}} \arrow[d] \\ (\underline{X}, \underline{*})^{K_1}_{\mathrm{co}} \arrow[r] & * \end{tikzcd}$$

Hence, there is a homotopy equivalence

$$(\underline{X}, \underline{*})^K_{\mathrm{co}} \simeq (\underline{X}, \underline{*})^{K_1}_{\mathrm{co}} \times (\underline{X}, \underline{*})^{K_2}_{\mathrm{co}}.$$ □

Acknowledgements

The authors would like to thank the International Centre for Mathematical Sciences (ICMS), Edinburgh, for support and hospitality during the workshop “Polyhedral Products: a Path Between Homotopy Theory and Geometric Group Theory”, where work on this paper began. We are grateful to the workshop organisers and especially to Martin Bendersky, Mark Grant and Sarah Whitehouse for helpful and encouraging conversations at the beginning of this project. Hornslien would like to thank Louis Martini for helpful discussions regarding Lemma 3.7 and the wedge lemma of Welker, Ziegler and Živaljević [21]. The authors would also like to thank Stephen Theriault for reading a draft of this work, and the referee for valuable comments that improved the quality of this paper.

References

[1] **A Bahri**, **M Bendersky**, **F R Cohen**, *Polyhedral products and features of their homotopy theory*, from "Handbook of homotopy theory" (H Miller, editor), CRC Press, Boca Raton, FL (2020) 103–144 MR

[2] **A Bahri**, **M Bendersky**, **F R Cohen**, **S Gitler**, *Decompositions of the polyhedral product functor with applications to moment-angle complexes and related spaces*, Proc. Natl. Acad. Sci. USA 106:30 (2009) 12241–12244 MR

[3] **A Bahri**, **M Bendersky**, **F R Cohen**, **S Gitler**, *The polyhedral product functor*: *a method of decomposition for moment-angle complexes, arrangements and related spaces*, Adv. Math. 225:3 (2010) 1634–1668 MR

[4] **V M Buchstaber**, **T E Panov**, *Actions of tori, combinatorial topology and homological algebra*, Uspekhi Mat. Nauk 55:5(335) (2000) 3–106 MR In Russian; translated in Russian Math. Surv. 55 (2000) 825–921

[5] **G Denham**, **A I Suciu**, *Moment-angle complexes, monomial ideals and Massey products*, Pure Appl. Math. Q. 3:1 (2007) 25–60 MR

[6] **T Ganea**, *A generalization of the homology and homotopy suspension*, Comment. Math. Helv. 39 (1965) 295–322 MR

[7] **J Grbić**, **S Theriault**, *The homotopy type of the polyhedral product for shifted complexes*, Adv. Math. 245 (2013) 690–715 MR

[8] **Y Hao**, **Q Sun**, **S Theriault**, *Moore's conjecture for polyhedral products*, Math. Proc. Cambridge Philos. Soc. 167:1 (2019) 23–33 MR

[9] **P J Hilton**, *On the homotopy groups of the union of spheres*, J. London Math. Soc. 30 (1955) 154–172 MR

[10] **M Hovey**, *Lusternik–Schnirelmann cocategory*, Illinois J. Math. 37:2 (1993) 224–239 MR

[11] **K Iriye**, **D Kishimoto**, *Decompositions of polyhedral products for shifted complexes*, Adv. Math. 245 (2013) 716–736 MR

[12] **K Iriye**, **D Kishimoto**, *Fat-wedge filtration and decomposition of polyhedral products*, Kyoto J. Math. 59:1 (2019) 1–51 MR

[13] **D Kishimoto**, **R Levi**, *Polyhedral products over finite posets*, Kyoto J. Math. 62:3 (2022) 615–654 MR

[14] **M Mather**, *Pull-backs in homotopy theory*, Canadian J. Math. 28:2 (1976) 225–263 MR

[15] **J W Milnor**, *On the construction FK*, from "Algebraic topology: a student's guide" (J F Adams, editor), Lond. Math. Soc. Lect. Note Ser. 4, Cambridge Univ. Press (1972) 118–136

[16] **D Notbohm**, **N Ray**, *On Davis–Januszkiewicz homotopy types, I*: *Formality and rationalisation*, Algebr. Geom. Topol. 5 (2005) 31–51 MR

[17] **T Panov**, **S Theriault**, *The homotopy theory of polyhedral products associated with flag complexes*, Compos. Math. 155:1 (2019) 206–228 MR

[18] **G J Porter**, *The homotopy groups of wedges of suspensions*, Amer. J. Math. 88 (1966) 655–663 MR

[19] **L Stanton**, *Loop space decompositions of moment-angle complexes associated to flag complexes*, Q. J. Math. 75:2 (2024) 457–477 MR

[20] **S Theriault**, *The dual polyhedral product, cocategory and nilpotence*, Adv. Math. 340 (2018) 138–192 MR

[21] **V Welker**, **G M Ziegler**, **R T Živaljević**, *Homotopy colimits—comparison lemmas for combinatorial applications*, J. Reine Angew. Math. 509 (1999) 117–149 MR

STEVEN AMELOTTE steven.amelotte@carleton.ca
School of Mathematics and Statistics, Carleton University, Ottawa, ON, Canada

WILLIAM HORNSLIEN w.hornslien.math@proton.me
Institut Fourier, Université Grenoble-Alpes, Grenoble, France

LEWIS STANTON L.R.Stanton@soton.ac.uk
School of Mathematical Sciences, University of Southampton, Southampton, United Kingdom

Received: June 6, 2024 Revised: April 24, 2025

Algebraic & Geometric Topology **26**:5 (2026) 1801–1866
DOI: 10.2140/agt.2026.26.1801
Published: May 29, 2026

Homoclinic leaves, Hausdorff limits and homeomorphisms

IAN BIRINGER AND CYRIL LECUIRE

We show that except for one exceptional case, a lamination on the boundary of a handlebody H is commensurable to a Hausdorff limit of meridians if and only if it is commensurable to a lamination with a "homoclinic leaf". This is an "if and only if" version of a theorem called Casson's criterion. Applications of our techniques include a characterisation of when a nonminimal lamination is a Hausdorff limit of meridians, in terms of properties of its minimal components, and a related characterisation of which reducible self-homeomorphisms of ∂H have powers that extend to subcompression bodies of H.

1 Introduction

Let H be a 3-dimensional handlebody[1] with genus $g \geq 2$ and let $S := \partial H$, which we usually equip with a reference hyperbolic metric. A simple closed curve m on S is called a *meridian* if it bounds an embedded disk in H but not in S. These curves play a fundamental role when studying the topology and geometry of handlebodies and more general 3-manifolds with compressible boundary. Our work here is centred around geodesic laminations $\lambda \subset S$ that are *Hausdorff limits of meridians*, i.e., λ where there is a sequence of geodesic meridians (m_i) on S such that $m_i \to \lambda$ in the Hausdorff topology on the set of closed subsets of S. See Section 2.8 for definitions.

Limits of meridians are important in the geometry and topology of hyperbolic 3-manifolds [2; 13; 42], convergence of sequences of Kleinian groups [12; 35; 38] and the action of $\mathrm{Out}(F_n)$ on the character variety $\chi(F_n, \mathrm{PSL}_2(\mathbb{C}))$; see [31; 36]. In many of the above references, the focus is on a closely related set called the "Masur domain", defined in [46], but to work with that one usually has to understand Hausdorff limits of meridians too.

We have three main goals. First, we investigate the relationship between Hausdorff limits of meridians and "homoclinic leaves", defined by Casson in unpublished notes and later exploited by Otal [56] in his thesis. Using this analysis, we study Hausdorff limits of meridians that are *nonminimal* laminations, in essence reducing their classification to the minimal case. Finally, using similar techniques we characterise when a reducible homeomorphism $f : S \to S$ extends (partially, up to a power) into H, in terms of extension properties of its Nielsen–Thurston components.

Another motivation for writing this paper is that it provides all the technical topological machinery necessary for our forthcoming follow-up paper, called *Iterations in Schottky space*. In that paper, we

---

[1] The body of this paper is written in greater generality, with (H, S) replaced by a compact, orientable, 3-manifold M with hyperbolisable interior, together with an essential connected subsurface $S \subset \partial M$ such that the multicurve ∂S is incompressible in M. However, everything we do is just as interesting in the handlebody case.

MSC2020: 57K30, 57K32.

fix a point $[X] \in \mathcal{T}(S)$ and describe the geometric limits of sequences of convex cocompact hyperbolic structures on H whose conformal boundaries are the iterates $f^n([X]) \in \mathcal{T}(S)$. This is a compressible version of Jeff Brock's thesis [8]. Brock's result was a first step toward work of Minsky [52] and Brock, Canary and Minsky [9], in which they developed bilipschitz models for hyperbolic structures on $S \times \mathbb{R}$ in terms of their end invariants. Our work on iterations in Schottky space will likewise be a step toward a bilipschitz model theory for hyperbolic structures on a handlebody.

1.1 Homoclinic leaves

Let $\widetilde{H}$ be the universal cover of H, which is homeomorphic to a thickened infinite tree. We equip H with any Riemannian metric, and equip $\widetilde{H}$ with the lift of this metric. A geodesic $\ell \subset S$ is called *homoclinic*[2] if it has a (possibly periodic) parametrisation $h : \mathbb{R} \to \lambda \subset S = \partial H$ that lifts to a path $\tilde{h} : \mathbb{R} \to \partial \widetilde{H}$ where there are sequences $s_i, t_i \in \mathbb{R}$ such that

$$|s_i - t_i| \to \infty \quad \text{and} \quad \sup_i d_{\widetilde{H}}(\tilde{\ell}(s_i), \tilde{\ell}(t_i)) < \infty.$$

Intuitively, ℓ is homoclinic if it travels very inefficiently in H, even though on the boundary of H it is a geodesic. As an example, one can check that a simple closed geodesic ℓ on S is homoclinic if and only if ℓ is a meridian. Similarly, a biinfinite geodesic that spirals onto a meridian will also be homoclinic.

The following was first written by Otal [56], building on work in an unpublished manuscript of Casson.

Theorem 1.1 ("Casson's criterion") *Let λ be a geodesic lamination on S. If λ is a Hausdorff limit of meridians, then λ contains a homoclinic leaf.*

The converse of Theorem 1.1 is not true. Of course, any Hausdorff limit of meridians is connected, and there are disconnected laminations on S that have homoclinic leaves, e.g., the union of two disjoint simple closed curves, one of which is a meridian. There are also connected laminations containing a meridian as a leaf (say) that are not even Hausdorff limits of simple closed curves. And there are also some more interesting examples that are related more intimately to the structure of H; see the beginning of Section 7.

Nevertheless, in certain particular cases homoclinic leaves have been used to construct limits of sequences of meridians (or annuli) in H; see [41; 43; 56] and particularly [38; 42]. And in Otal's thesis, Casson's criterion was even stated (incorrectly) in passing as an "if and only if". Our goal is to state precisely a result that is closest to a converse of Theorem 1.1. This will tie together, extend and explain the partial results in the papers above.

We say that μ_1, μ_2 are *strongly commensurable* if they contain a common sublamination ν such that for both i, the difference $\mu_i \setminus \nu$ is the union of finitely many isolated leaves, none of which are simple closed curves. Additionally, let's say that a lamination λ on S is *exceptional* if S has genus 2, there is a

---

[2] This is not quite what Casson and Otal call "homoclinic" (in French, "homoclinique"), but rather what Otal calls "faiblement homoclinique", or "weakly homoclinic". However, the definition we give has been adopted in almost all subsequent papers. See Section 5.1 for an explanation of the difference between the two definitions.

separating meridian m on S that does not intersect λ transversely, and λ intersects transversely the two nonseparating meridians disjoint from m.

Theorem 1.2 (a bidirectional Casson's criterion; see Theorem 7.2) *If λ is a geodesic lamination on S that is not exceptional, then λ is strongly commensurable to a Hausdorff limit of meridians if and only if it is strongly commensurable to a lamination with a homoclinic leaf.*

This is the first part of Theorem 7.2. In our view, it is the strongest converse for Casson's criterion that is likely to be true for arbitrary laminations.

The main tool in the proof of Theorem 1.2 is a complete characterisation of the minimal laminations onto which the two ends of a homoclinic simple geodesic on S can accumulate.

Theorem 1.3 (limits of homoclinic geodesics; see Corollary 6.2) *Suppose that h is a homoclinic simple biinfinite geodesic on S and that the two ends of h limit onto minimal laminations $\lambda_-, \lambda_+ \subset S$. Then either*

(1) *the two ends of h are asymptotic on S,*

(2) *one of λ_-, λ_+ is an intrinsic limit of meridians, or*

(3) *λ_-, λ_+ are contained in incompressible subsurfaces $S_-, S_+ \subset S$ that bound an essential interval bundle $B \subset H$ through which λ_- and λ_+ are homotopic.*

Here, a minimal lamination $\lambda \subset S$ is *an intrinsic limit of meridians* if it is strongly commensurable to the Hausdorff limit of a sequence of meridians that are contained in the smallest essential subsurface $S(\lambda) \subset S$ containing λ; see Proposition 5.12 for a number of equivalent definitions. We refer the reader to Theorem 6.1 and Corollary 6.2 for more precise and more general versions of the above that apply both to homoclinic biinfinite geodesics, and also to pairs of "mutually homoclinic" geodesic rays on S, as well as many examples.

Interval bundles as in (3) are essential to the study of meridians on handlebodies, and it is no surprise that they appear in Theorem 1.3. For example, subsurfaces bounding such interval bundles are the "incompressible holes" studied by Masur and Schleimer [49], and interval bundles appear frequently in Hamenstädt's work on the disk set; see, e.g., [25; 26; 27]. We note that the interval bundles B appearing in Theorem 1.3 may be twisted interval bundles over nonorientable surfaces, in which case $\lambda_- = \lambda_+$ and $S_- = S_+$. See Proposition 4.5 for background on interval bundles, and Section 6 for examples of (3).

1.2 Hausdorff limits via their minimal sublaminations

The previous two theorems suggest that if a lamination λ that is a Hausdorff limits of meridians, one might expect to see minimal sublaminations of λ that are intrinsic limits of meridians, or pairs of components that are homotopic through essential interval bundles in H. In fact, we show the following.

Theorem 1.4 (the second part of Theorem 7.2) *Suppose that $\lambda \subset S$ is a nonexceptional geodesic lamination that is a finite union of minimal components. Then λ is strongly commensurable to a Hausdorff limit of meridians if and only if either*

(1) *λ is disjoint from a meridian on S*,

(2) *some component of λ is an intrinsic limit of meridians*, or

(3) *there are components $\lambda_\pm \subset \lambda$ that fill incompressible subsurfaces $S_\pm \subset S$, such that $S_\pm$ bound an essential interval bundle $B \subset H$, the laminations $\lambda_\pm$ are essentially homotopic through B, and there is a compression arc α for B that is disjoint from λ.*

In (3), a *compression arc* for B is an arc from ∂S_- to ∂S_+ that is homotopic in H, keeping its endpoints in $\partial S_\pm$, to a fibre of the interval bundle B. See Section 2.7 and Figure 2 for more explanation.

Note that Theorem 1.4 does not say anything interesting about which minimal filling laminations λ on S are strongly commensurable to Hausdorff limits of meridians. Indeed, for minimal filling laminations, it is not clear that there should be an easy way to "identify" Hausdorff limits of meridians. The point of Theorem 1.4, though, is that it reduces the characterisation of Hausdorff limits of meridians to the minimal filling case. We note that one should also be able to replace the part of the proof of Theorem 1.4 that references homoclinic geodesics with arguments inspired by Masur and Schleimer's paper [49].

1.3 Extension of reducible maps to compression bodies

Lackenby [39] studied a generalisation of Dehn filling in which one starts with a compact 3-manifold M with (say, connected) genus-g boundary Σ, and glues a genus-g handlebody H via a homeomorphism $\phi : \Sigma \to S := \partial H$. Among other results, he showed that when an arbitrary ϕ is precomposed by a high power of a sufficiently "generic" homeomorphism of S, the gluing is hyperbolic:

Theorem 1.5 (Lackenby [39]) *Suppose that M is hyperbolisable and acylindrical with incompressible boundary, and $f : S \to S$ is a homeomorphism such that no nonzero power of f extends to a nontrivial subcompression body of H. Then for infinitely many n, the gluing $M \cup_{\phi \circ f^n} H$ is hyperbolic.*[3]

See also the work of Namazi and Souto [55] and Brock, Minsky, Namazi and Souto [10] for more general hyperbolisation theorems inspired by Lackenby's result.

Above, a *subcompression body* of H is a 3-dimensional submanifold $C \subset H$ with $S \subset \partial C$ that is obtained by choosing a finite collection Γ of disjoint meridians on S, taking a regular neighbourhood of S and a collection of discs in H with boundary Γ, and adding in any complementary components that are topological 3-balls. We say C is obtained by *compressing* Γ. We usually consider subcompression bodies only up to isotopy, and we allow the case that $\Gamma = \varnothing$, in which case we recover the *trivial subcompression body*, which is just a regular neighbourhood of S. See Section 2.4 for details.

In light of Lackenby's theorem, it is natural to investigate which homeomorphisms $f : S \to S$ have powers that extend to a nontrivial subcompression body of H. Biringer, Johnson and Minsky [2] showed that for pseudo-Anosov f, this condition is equivalent to the attracting lamination λ_+ being a (projective) limit of meridians. In fact, Lackenby also proved a version of his theorem for pseudo-Anosovs with the latter condition on λ_+, not knowing the equivalence of the two properties at that time.

---

[3] Lackenby's paper was written pregeometrisation, so he actually proves that the hypotheses of the hyperbolisation theorem are satisfied, not that the gluing is hyperbolic.

Here, we show that extension of powers of a homeomorphism $f : S \to S$ to subcompression bodies can be detected by looking at extension of powers of its components in the Nielsen–Thurston decomposition. More precisely, recall that f is *pure* if there are disjoint, nonisotopic,[4] essential subsurfaces $S_i \subset S$, such that $f = \mathrm{id}$ on $S_{\mathrm{id}} := S \setminus \bigcup_i S_i$, and where for each i, if we set $f_i := f|_{S_i}$, then either

(1) S_i is an annulus and f_i is a nonzero power of a Dehn twist, or

(2) f_i is a pseudo-Anosov map on S_i.

It follows from the Nielsen–Thurston classification [21] that every homeomorphism of S has a power that is isotopic to a pure homeomorphism.

Theorem 1.6 (partial extension of reducible maps; see Theorem 9.2) *Let $f : S \to S$ be a pure homeomorphism. Then f has a power that extends to a nontrivial subcompression body of H if and only if either*

(1) *there is a meridian in S_{id},*

(2) *for some i, the map $f_i : S_i \to S_i$ has a power that extends to a nontrivial subcompression body of H that is obtained by compressing a set of meridians in S_i, or*

(3) *there are (possibly equal) indices i, j such that S_i, S_j bound an essential interval bundle B in H, such that some power of $f|_{S_i \cup S_j}$ extends to B, and there is a compression arc α for B whose interior lies in S_{id}.*

Moreover, Theorem 8.1 says that when f_i is a pseudo-Anosov map on S_i, then (2) above is equivalent to the condition that the attracting lamination of f_i is a (projective) limit of meridians in S_i. This is a relative version of Biringer, Johnson and Minsky's article [2], that generalises their theorem from pseudo-Anosov maps on S to partial pseudo-Anosovs.

In (3), note that if (for simplicity) B is a trivial interval bundle, then $f|_{S_i \cup S_j}$ extends to B exactly when f_i, f_j become isotopic maps when S_i, S_j are identified through B. More generally, a power of $f|_{S_i \cup S_j}$ extends to B when f_j is obtained from f_i by multiplying by a periodic map that commutes with f_i.

1.4 Other results of interest

There are two other theorems in this paper that we should mention in the introduction.

In Section 3 we study the *disk set* $\mathcal{D}(S, M)$ of all isotopy classes of meridians in an essential subsurface $S \subset \partial M$ with ∂S incompressible, where here M is a compact, irreducible 3-manifold with boundary. We show in Proposition 3.1 that either $\mathcal{D}(S, M)$ is *small*, meaning that it is either empty, has a single element, or has a single nonseparating element and infinitely many separating elements that one can explicitly describe, or $\mathcal{D}(S, M)$ is *large*, meaning that it has infinite diameter in the curve complex $\mathcal{C}(S)$. This result will probably not surprise any experts, but we have never seen it in the literature.

In Section 4 we show how essential interval bundles in a compact 3-manifold with boundary M can be seen in the limit sets in $\partial\mathbb{H}^3$ associated to hyperbolic metrics on $\mathrm{int}(M)$. This picture was originally

---

[4] This is only important for annuli.

known to Thurston [59], and was studied previously under more restrictive assumptions by Walsh [60] and Lecuire [40]. We need a more general theorem in the proof of Theorem 6.1: in particular, we need a version that allows accidental parabolics. This is Theorem 4.1. Our proof is also more direct and more elementary than those of [40; 60]. See Section 4 for more context and details.

1.5 Outline of the paper

Section 2 contains all the necessary background for the rest of the paper. We discuss the curve complex, the disc set, compression bodies, interval bundles, the Jaco–Shalen and Johannson characteristic submanifold theory, compression arcs, and geodesic laminations. Sections 3 and 4 are described in the previous subsection. Section 5 contains a discussion of homoclinic geodesics, intrinsic limits of meridians, and some of their basic properties. In Section 6 we explain how to build essential discs and annulus from a homotopic leaf. Section 7 is devoted to Theorem 7.2, which characterises Hausdorff limits of meridians and their structure; see Theorem 1.2 above. Section 8 contains our extension of [2] to partial pseudo-Anosovs, and Section 9 contains the proof of Theorem 9.2, which generalises Theorem 1.6 above.

2 Preliminaries

2.1 Subsurfaces with geodesic boundary

Suppose S is a finite-type hyperbolic surface with geodesic boundary. A *connected subsurface with geodesic boundary* in S is by definition either

(1) a simple closed geodesic X on S, which is the degenerate case, or

(2) an immersed surface $X \to S$ such that the restriction to $\operatorname{int}(X)$ and to each component of ∂X is an embedding, and where each component of ∂X maps to a simple closed geodesic on S.

In (2), the point is that our surface is basically an embedding, except that we allow two boundary components of X to map to the same geodesic in S. We will usually suppress the immersion and write $X \subset S$, abusing notation. We consider $X, Y \subset S$ to be *equal* if they are either the same simple closed geodesic, or if they are both immersions as in (2) and the interiors of their domains have the same images. We say X, Y are *essentially disjoint* if either

- X, Y are disjoint simple closed geodesics,
- X is a simple closed geodesic, Y is not, and X is disjoint from $\operatorname{int}(Y)$, or vice versa with X, Y exchanged, or
- X, Y have nonempty disjoint interiors.

More generally, we define a (possibly disconnected) *subsurface with geodesic boundary* in S to be a finite union of essentially disjoint connected subsurfaces with geodesic boundary.

Any connected essential subsurface $T \subset S$ that is not an annulus homotopic into a cusp of S determines a unique connected subsurface with geodesic boundary X such that the images of $\pi_1 T$ and $\pi_1 X$ in $\pi_1 S$

are conjugate. Here, we say that X is obtained by *tightening* T. More generally, we can tighten a disconnected T to a disconnected X by tightening all its components.

Tightening is performed as follows. If T is an annulus, then we let X be the unique simple closed geodesic homotopic to the core curve of T. Otherwise, we obtain X by homotoping T so that every component of ∂T is either geodesic or bounds a cusp in $S \setminus T$, and then adding in any components of $S \setminus T$ that are cusp neighbourhoods. Alternatively, let $\widetilde{T}$ be a component of the preimage of T in the universal cover $\widetilde{S}$, which is isometric to a convex subset of $\mathbb{H}^2$, let $\Lambda_T \subset \partial \mathbb{H}^3$ be the set of limit points of $\widetilde{T}$ on $\partial_\infty \mathbb{H}^2$, and let $\widetilde{X}$ be the convex hull of Λ_T within $\widetilde{S}$. Then $\widetilde{X}$ projects to an X as desired.

Conversely, suppose X is a subsurface with geodesic boundary in S. Then there is a compact essential subsurface $T \hookrightarrow S$, unique up to isotopy and called a *resolution* of X, that tightens to X. When X is a simple closed geodesic, we take T to be a regular neighbourhood of X. Otherwise, construct T by deleting half-open collar neighbourhoods of all boundary components of X, and deleting open neighbourhoods of all cusps of T.

Note that subsurfaces with geodesic boundary X, Y are essentially disjoint if and only if they admit disjoint resolutions.

2.2 The curve complex

Let S be a compact orientable surface, possibly with boundary, and assume that S is not an annulus.

Definition 2.1 The *curve complex* of S, written $\mathcal{C}(S)$, is the graph whose vertices are homotopy classes of nonperipheral, essential simple closed curves on S and whose edges connect homotopy classes that intersect minimally.

When S is a 4-holed sphere, minimally intersecting simple closed curves intersect twice, while on a punctured torus they intersect once. Otherwise, edges in $\mathcal{C}(S)$ connect homotopy classes that admit disjoint representatives.

Masur and Minsky [47] have shown that the curve complex is Gromov hyperbolic, when considered with the path metric in which all edges have unit length. Klarreich [37] (see also [24]) showed that the Gromov boundary $\partial_\infty \mathcal{C}(S)$ is homeomorphic to the space of *ending laminations* of S: i.e., filling, measurable geodesic laminations on S with the topology of Hausdorff superconvergence.

2.3 The disc set

Suppose $S \subset \partial M$ is an essential subsurface of the boundary of a compact, irreducible 3-manifold M, and that ∂S is incompressible in M. An essential simple closed curve γ on M is called a *meridian* if it bounds an embedded disc in M. By the loop theorem, γ is a meridian if and only if it is homotopically trivial in M.

Definition 2.2 The *disc set* of S in M, written $\mathcal{D}(S, M)$, is the (full) subgraph of $\mathcal{C}(S)$ whose vertices are the meridians of S in M.

When convenient, we will sometimes regard $\mathcal{D}(S, M)$ as a subset of the space of projective measured laminations $\mathcal{PML}(S)$, instead of as a graph.

The following is an extension of a theorem of Masur and Minsky [48, Theorem 1.1], which they prove in the case that S is an entire component of ∂M.

Theorem 2.3 (Masur–Minsky) *The subset $\mathcal{D}(S, M)$ of $\mathcal{C}(S)$ is quasiconvex.*

To prove Theorem 2.3 as stated above, one follows the outline of [48]: given $a, b \in \mathcal{D}(S, M)$, the goal is to construct a *well-nested curve replacement sequence* from $a = a_1, \dots, a_n = b$ consisting of meridians, which must be a quasigeodesic by their Theorem 1.2. The sequence (a_i) is created by successive surgeries along innermost discs, and the only difference here is that one needs to ensure that none of the surgeries create peripheral curves. However, the surgeries create meridians and S has incompressible boundary.

2.4 Compression bodies

We refer the reader to Section 2 of [3] for a more detailed discussion of compression bodies, and state here only a few definitions that will be used later on.

A *compression body* is a compact, orientable, irreducible 3-manifold C with a π_1-surjective boundary component $\partial_+ C$, called the *exterior boundary* of C. The complement $\partial C \setminus \partial_+ C$ is called the *interior boundary*, and is written $\partial_- C$. Note that the interior boundary is incompressible. For if an essential simple closed curve on $\partial_- C$ bounds a disk $D \subset C$, then $C \setminus D$ has either one or two components, and in both cases, Van Kampen's theorem implies that $\partial_+ C$, which is disjoint from D, cannot π_1-surject.

Suppose M is a compact irreducible 3-manifold with boundary, let Σ be a component of ∂M. A *subcompression body of* (M, Σ) is a compression body that is embedded as a submanifold $C \subset M$ with exterior boundary Σ. Up to isotopy, any such subcompression body can be constructed as follows. Choose a set Γ of disjoint, pairwise nonhomotopic simple closed curves on Σ that are all meridians in M. Let $C' \subset M$ be the union of Σ with a set of disjoint disks in M whose boundaries are the components of Γ, and define $C \subset M$ to be the union of a regular neighbourhood of $C' \subset M$ together with any components of the complement of this neighbourhood that are topological 3-balls. Here, we say that $C \subset M$ is obtained *by compressing* Γ. Note that the irreducibility of M implies that no component of ∂C is a 2-sphere, and hence that C is irreducible, and therefore a compression body. See [3, Section 2] for details about this construction. More generally, if $S \subset \Sigma$ is an essential subsurface, a *subcompression body of* (M, S) is a subcompression body of (M, Σ) obtained by compressing a multicurve $\Gamma \subset S$.

Two subcompression bodies of a handlebody H are illustrated in Figure 1. On the left, we compress a separating meridian $\Gamma = \{m_1\}$ and obtain a subcompression body of H that has two "interior" boundary components contained in $\operatorname{int}(H)$; these are the tori drawn in grey. On the right, we compress a nonseparating meridian $\Gamma = \{m_2\}$ and obtain a subcompression body with a single torus interior boundary component. Note that compressing $\Gamma = \{m_1, m_2\}$ gives the same subcompression body as compressing m_2, because we fill in complementary components that are balls.

When the compressing set Γ is empty, we obtain the *trivial subcompression body* of (M, S), which is just a regular neighbourhood of $\Sigma \subset \partial M$. (In other words, the trivial compression body $\Sigma \times [0, 1]$ can be considered as a subcompression body of (M, S) for any subsurface $S \subset \Sigma$.) At the other extreme, we

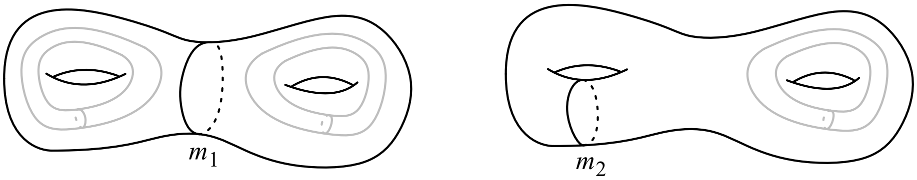


Figure 1: Compression bodies inside a genus-2 handlebody.

can compress a maximal $\Gamma \subset S$, which gives the "characteristic compression body" of (M, S), defined via the following fact.

Fact 2.4 (the characteristic compression body) *Suppose M is an irreducible compact 3-manifold, that Σ is a component of ∂M and that $S \subset \Sigma$ is an essential subsurface such that the multicurve ∂S is incompressible in M. Then there is a unique (up to isotopy) subcompression body*

$$C := C(M, S) \subset M$$

of (M, S), called the **characteristic compression body** *of (M, S), such that a curve γ in S is a meridian in C if and only if it is a meridian in M.*

Moreover, C can be constructed by compressing any maximal set of disjoint, pairwise nonhomotopic meridians in S.

This is a version of a construction of Bonahon [5], except that he only defines the characteristic compression body when S is an entire boundary component of M. In that case, the interior boundary components of M are incompressible in M, so Bonahon's construction can be used to reduce problems about 3-manifolds to problems about compression bodies and about 3-manifolds with incompressible boundary.

The reader can also compare Fact 2.4 to Lemma 2.1 in [3], which is the special case of the fact where M is a compression body and S is its exterior boundary, so that $C = M$ is obtained by compressing any maximal set of disjoint, nonhomotopic meridians in M.

Proof Let Γ be a maximal set of disjoint, pairwise nonhomotopic M-meridians on S, and define C by compressing Γ. We have to check that any curve in S that is an M-meridian is also a C-meridian. Suppose not, and take an M-meridian $m \subset S$ that is not a C-meridian, and that intersects Γ minimally. Since Γ is maximal, m intersects some component $\gamma \subset \Gamma$. Then there is an arc $\alpha \subset \gamma$ with endpoints on m and interior disjoint from m, that is homotopic rel endpoints in M to the arcs $\beta', \beta'' \subset m$ with the same endpoints. (Here α is an "outermost" arc of intersection on a disk bounded by γ, where the intersection is with the disk bounded by m; see, e.g., Lemma 2.8 in [3].) Since m is in minimal position with respect to Γ, the curves $m' = \alpha \cup \beta'$ and $m'' = \alpha \cup \beta''$ are both essential, and are M-meridians in S that intersects Γ fewer times than m. So by minimality of m, both m', m'' are C-meridians, implying that α is homotopic rel endpoints to β' and β'' in C. This implies m is a C-meridian, contrary to assumption.

For uniqueness, suppose we have two subcompression bodies C_1, C_2 of (M, S) in which all curves in S that are meridians in M are also meridians in C_1, C_2. Since C_1, C_2 are subcompression bodies of (M, S), the kernels of the maps

$$\pi_1 \Sigma \to \pi_1 C_i$$

induced by inclusion are both normally generated by the set of all elements of $\pi_1 \Sigma$ that represent simple closed curves in S that are meridians in M. Hence, the disk sets $\mathcal{D}(\Sigma, C_i)$ are the same for $i = 1, 2$. It follows that C_1, C_2 are isotopic in M, say by Corollary 2.2 of [3]. ☐

2.5 Interval bundles

In this paper, an *interval bundle* always means a fibre bundle $B \to Y$, where Y is a compact surface with boundary, and where all fibres are closed intervals I. Regarding the fibres as "vertical", we call the associated ∂I-bundle over Y the *horizontal boundary* of B, written $\partial_H B$. An interval bundle that is isomorphic to $Y \times [-1, 1]$ is called *trivial*, and we often call nontrivial interval bundles *twisted*.

All 3-manifolds in this paper are assumed to be orientable, but even when the total space B of an interval bundle is orientable, the base surface Y may not be. Indeed, let Y be a compact nonorientable surface and let $\pi : \widehat{Y} \to Y$ be its orientation cover. Then the mapping cylinder

$$B := \widehat{Y} \times [0, 1]/\sim, \quad (x, 1) \sim (x', 1) \iff \pi(x) = \pi(x'),$$

is orientable, and is a twisted interval bundle over Y, where the fibre over $y \in Y$ is obtained by gluing together the two intervals $\{x\} \times [0, 1]$ and $\{x'\} \times [0, 1]$ along $(x, 1)$ and $(x', 1)$, where $\pi^{-1}(y) = \{x, x'\}$. The horizontal boundary $\partial_H B$ here is $\widehat{Y} \times \{0\}$, which is homeomorphic to the orientable surface $\widehat{Y}$. Note that B is double covered by the trivial interval bundle $\widehat{Y} \times [-1, 1]$.

Fact 2.5 *Suppose that $B \to Y$ is an interval bundle and B is orientable. If Y is orientable, then B is a trivial interval bundle. If Y is nonorientable, then B is isomorphic to the mapping cylinder of the orientation cover of Y.*

Proof If Y and B are orientable, so is the line bundle, so the bundle is trivial. If Y is nonorientable, the horizontal boundary $\partial_H B \subset \partial B$ is an orientable surface that double covers Y, and from there it's easy to construct the desired isomorphism to the mapping cylinder of the projection $\partial_H B \to Y$. ☐

An interval bundle $B \to Y$ comes with a *canonical involution* σ, which is well defined up to isotopy, and which is defined as follows. If $B \cong Y \times [0, 1]$ is a trivial interval bundle, we define

$$\sigma : Y \times [-1, 1] \to Y \times [-1, 1], \quad \sigma(y, t) = (y, -t).$$

And if B is the twisted interval bundle $B \cong \widehat{Y} \times [0, 1]/\sim$ above, we define

$$\sigma : \widehat{Y} \times [0, 1]/\sim \to \widehat{Y} \times [0, 1]/\sim, \quad \sigma(\hat{y}, t) = (\iota(\hat{y}), t),$$

where ι is the nontrivial deck transformation of the orientation cover. Note that σ is always an orientation reversing involution of B, so in particular, when we give the surface $\partial_H B$ its boundary orientation, the restriction $\sigma|_{\partial_H B}$ is also orientation reversing.

We also recall the following well-known fact.

Fact 2.6 *Suppose $B \to Y$ is an interval bundle* (*as always, over a compact surface with boundary*) *and B is orientable. Then B is homeomorphic to a handlebody.*

It's a nice topology exercise to visualise the homeomorphism. Regard Y as the union of a polygon and a collection of bands (long, skinny rectangles), each of which is glued along its short sides to two sides of the polygon. Thickening, the picture becomes a ball with 1-handles attached, so since B is orientable, it is a handlebody.

Note that if $S = S_{g,b}$ has genus g and b boundary components, then the handlebody $S \times [-1,1]$ has genus $2g+b-1$, since that is the rank of the free group $\pi_1(S \times [-1,1]) \cong \pi_1 S$.

Finally, suppose $\pi : B \to Y$ is an interval bundle and $f : \partial_H B \to \partial_H B$ is a homeomorphism. We say that f *extends to B* if there is a homeomorphism $F : B \to B$ such that $F|_{\partial_H B} = f$. We leave the following to the reader.

Fact 2.7 *The following are equivalent*:

(1) *f extends to B.*

(2) *$f \circ \sigma$ is isotopic to f on $\partial_H B$.*

(3) *After isotoping f, there is a homeomorphism $\bar{f} : Y \to Y$ such that $\pi \circ f = \bar{f} \circ \pi$.*

(4) *There is a homeomorphism from B to either*

$$Y \times [-1,1] \quad \text{or} \quad \widehat{Y} \times [0,1]/\sim,$$

taking horizontal boundary to horizontal boundary, such that $f = F|_{\partial_H B}$, and where either

$$F : Y \times [-1,1] \to Y \times [-1,1], \quad F(y,t) = (\bar{f}(y), t),$$

for some homeomorphism $\bar{f} : Y \to Y$, or

$$F : \widehat{Y} \times [0,1]/\sim \to \widehat{Y} \times [0,1]/\sim, \quad F(y,t) = (\bar{f}(y), t),$$

for some homeomorphism $\bar{f} : \widehat{Y} \to \widehat{Y}$ commuting with the deck group of $\widehat{Y} \to Y$, and hence covering a homeomorphism of Y.

2.6 The characteristic submanifold of a pair

Suppose that M is a compact, orientable 3-manifold and that $S \subset \partial M$ is an incompressible subsurface. In the late 1970s, Jaco and Shalen [30] and Johannson [32] described a "characteristic" submanifold of (M,S) that contains the images of all nondegenerate maps from interval bundles and Seifert fibred spaces.

Theorem 2.8 (Jaco–Shalen [29, page 138]) *There is a perfectly embedded Seifert pair* $(X,\Sigma)\subset(M,S)$*, unique up to isotopy and called the* **characteristic submanifold** *of* (M,S)*, such that any nondegenerate map* $(B,F)\to(M,S)$ *from a Seifert pair* (B,F) *is homotopic as a map of pairs into* (X,Σ)*.*

A *Seifert pair* is 3-manifold pair that is a finite disjoint union of interval bundle pairs $(B,\partial_H B)$ and S^1-bundle pairs. Here, an S^1*-bundle pair* (B,F) is a 3-manifold B fibred by circles, where $F\subset\partial B$ is a compact subsurface saturated by fibres. A Seifert pair $(X,\Sigma)\subset(M,S)$ is *well embedded* if $X\cap\partial M=\Sigma\subset S$ and the frontier of X in M is a π_1-injective surface, and is *perfectly embedded* if it is well embedded, no component of the frontier of X in M is homotopic into S, and no component of X is homotopic into another component.

When (B,F) is a connected Seifert pair, a map $f:(B,F)\to(M,S)$ is *essential* if it is not homotopic as a map of pairs into S. Notice that this only depends on the image of f and not on f itself. One says f is *nondegenerate* if it is essential, its π_1-image is nontrivial, its π_1-image is noncyclic when $F=\varnothing$, and no fibre of B is nullhomotopic in (M,S). For disconnected (B,F), one says f is nondegenerate if its restriction to every component is nondegenerate.

The following is very well known.

Fact 2.9 *If* $\mathrm{int}(M)$ *is hyperbolisable and* (B,F) *is an* S^1*-bundle pair that is perfectly embedded in* (M,S)*, then either*

(1) (B,F) *is a "fibred solid torus", i.e.,* B *is an* S^1*-bundle over a disk, and* $F\subset\partial B\cong T^2$ *is a collection of fibred parallel annuli, or*

(2) (B,F) *is a "thickened torus", i.e.,* B *is an* S^1*-bundle over an annulus, so is homeomorphic to* $T^2\times[0,1]$*, and each component of* F *is either a torus or a fibred annulus.*

So in particular, the components of the characteristic submanifold of (M,S) are either interval bundles, solid tori, or thickened tori.

Proof Suppose that (B,F) is a perfectly embedded S^1-bundle pair in M. Then $B\to Y$ is an S^1-bundle, where Y is a compact 2-orbifold, and the cyclic subgroup $Z\subset\pi_1 B$ corresponding to a regular fibre is normal in $\pi_1 B$. In a hyperbolic 3-manifold, any subgroup of π_1 that has a cyclic normal subgroup is elementary, say by a fixed point analysis on $\partial_\infty\mathbb{H}^3$. So, $\pi_1 B$ is either cyclic or isomorphic to $\mathbb{Z}^2$. It follows that Y is a disc, in which case B is a fibred solid torus, or Y is an annulus, in which case B is a thickened torus. □

In this paper we will mostly be interested in interval bundles. For brevity, we'll use the following terminology, which differs slightly from the terminology above used by Jaco and Shalen.

Definition 2.10 An *essential interval bundle* in (M,S) is an essential, well-embedded interval bundle pair $(B,\partial_H B)\hookrightarrow(M,S)$.

Note that the horizontal boundary of any essential interval bundle is an incompressible subsurface of S.

The definition above differs from a well-embedded interval bundle pair in that we are excluding boundary-parallel interval bundles over annuli, and differs from a perfectly embedded interval bundle pair in that we are allowing components of the frontier of an interval bundle over a surface that is not an annulus to be boundary parallel. For instance, if Y is a surface with boundary and $Y' \subset Y$ is obtained by deleting collar neighbourhoods of the boundary components, and we set $M = Y \times [-1, 1]$, which is a handlebody, then $(Y' \times [-1, 1], Y' \times \{-1, 1\})$ is an essential interval bundle in $(M, \partial M)$, but is not perfectly embedded. However, note that any essential interval bundle $(B, \partial_H B) \hookrightarrow (M, S)$ is perfectly embedded in $(M, \partial_H B)$.

2.7 Compression arcs

Suppose $(B, \partial_H B) \subset (M, S)$ is an essential interval bundle. An arc $\alpha \subset S$ with endpoints on $\partial(\partial_H B)$ and interior disjoint from $\partial_H B$ is called a *compression arc* if it is homotopic in M to a fibre of B, while keeping its endpoints on $\partial(\partial_H B)$. See Figure 2. To link this definition with more classical ones, it is easy to see that there is a compression arc for B if and only if $\overline{\mathrm{Fr}(B)}$ is boundary compressible; see [29, pages 36–37] for a definition.

Write our interval bundle as $\pi : B \to Y$. Let α be a compression arc for B. After isotoping the bundle map π, we can assume that α is homotopic rel endpoints to a fibre $\pi^{-1}(y)$, where $y \in Y$. Suppose c is an oriented, two-sided, essential, simple closed loop Y based at y, and suppose that either c is nonperipheral in Y, or that Y is an annulus or Möbius band. Write $\pi^{-1}(c) = c_- \cup c_+$, where $c_\pm$ are disjoint simple closed oriented loops in X based at $y_\pm$, and where the orientations of $c_\pm$ project to that of c.

Claim 2.11 *The concatenation* $m(c) := c_- \cdot \alpha \cdot c_+^{-1} \cdot \alpha^{-1}$ *is homotopic to a meridian on* S.

So, a compression arc α allows one to make compressible curves on S from essential curves on Y. See Figure 2.

Proof Since α is homotopic rel endpoints to the fibre $\pi^{-1}(y)$, the curve $m(c)$ is homotopic in M to a curve in B that projects under π to $c \cdot c^{-1}$, and hence $m(c)$ is nullhomotopic in M. Checking orientations, one can see that $m(c)$ is homotopic to a simple closed curve on S. So, we only have to prove that $m(c)$ is homotopically essential on S.

Suppose that c_-, c_+ are freely homotopic on $\partial_H B$ as oriented curves. (This happens exactly when the curve $c \subset Y$ bounds a Möbius band in Y.) Then $m(c)$ is homotopic to the commutator of two essential simple closed curves on S that intersect once, and hence is essential since S is not a torus.

We can now assume that $c_\pm$ are not freely homotopic in $\partial_H B$ as oriented curves. If $m(c)$ is inessential, then $c_\pm$ *are* freely homotopic on S, so $c_\pm$ are homotopic in $\partial_H B$ to boundary components $c'_\pm \subset \partial_H B$ that bound an annulus in $S \setminus \partial_H B$. In this case $c_\pm$ are peripheral, so we may assume that Y is either an annulus or a Möbius band. If Y is a Möbius band, we are in the situation of the previous paragraph and are done. So, Y is an annulus, and $\partial_H B$ is a pair of disjoint annuli on S, where $c'_\pm$ lie in different components of $\partial_H B$. Since $c'_\pm$ bound an annulus in $S \setminus \partial_H B$, the interval bundle B is inessential, contrary to our assumption. □

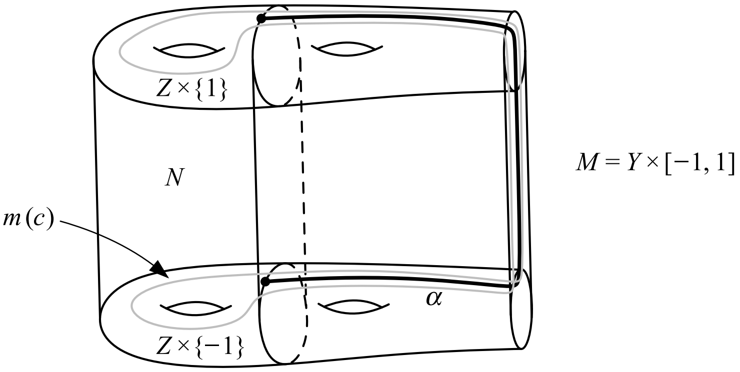


Figure 2: $Z\subset Y$ is a compact surface, the interval bundle $B=Z\times[-1,1]$ embeds in the manifold $M=Y\times[-1,1]$, and α above is a compression arc. Also pictured in light grey is a meridian as described in Claim 2.11.

In fact, more is true.

Fact 2.12 (arcs that produce meridians) *Suppose $(B,\partial_H B)\subset(M,S)$ is an essential interval bundle and let $\alpha\subset S$ be an arc with endpoints on $\partial_H B$ and interior disjoint from $\partial_H B$. Let $X\subset S$ be a regular neighbourhood of $\alpha\cup\partial_H B$ within S. Then there is a meridian in X if and only if we have either*

(1) *the endpoints of α lie on the same component c of $\partial(\partial_H B)$, and there is an arc $\beta\subset c$ such that $\alpha\cup\beta$ is a meridian, or*

(2) *α is a compression arc.*

Note that in the second case the endpoints of α lie on distinct components of $\partial(\partial_H B)$, so in particular the two cases are mutually exclusive.

The reason we say X "contains a meridian" instead of "is compressible" is that X may not be an essential subsurface of S, and we want to emphasise that the essential curve in X that is compressible in M is actually essential in S. For example, let Y be a compact surface with boundary, $Y'\subset Y$ be obtained by deleting a collar neighbourhood of ∂Y, set $B=Y'\times[-1,1]$ and $M=Y\times[-1,1]$, and let α be a spanning arc of B in ∂M.

Proof The "if" direction is immediate: in case (1) we are essentially given a meridian in X, and in case (2) we can appeal to Claim 2.11.

We now work on the "only if" direction. Write our regular neighbourhood of $\partial_H B\cup\alpha$ as $X=\partial_H B\cup R$ where R is a rectangle with two opposite "short" sides on the boundary of $\partial_H B$. Let $D\subset M$ be an essential disc whose boundary is contained in X, and where D intersects the frontier $\mathrm{Fr}(B)\subset M$ in a minimal number of components. Let $a\subset D\cap\mathrm{Fr}(B)$ be an arc that is "outermost" in D, i.e., there is some arc $a'\subset\partial D$ with the same endpoints as a such that a,a' bound an open disk in D that does not intersect $\mathrm{Fr}(B)$.

We claim that $a'\subset R$. If not, then $a'\subset\partial_H B$, and bounds a disk in B with the arc $a\subset\partial B$. Writing the interval bundle as $\pi:B\to Y$, the projection $\pi(a\cup a')$ in Y is then also nullhomotopic, so $\pi(a')$ is homotopic rel endpoints into ∂Y. Lifting this homotopy through the covering map $\partial_H B\to Y$ we get that

a' is inessential in $\partial_H B$, i.e., is homotopic in $\partial_H B$ rel endpoints into $\partial(\partial_H B)$. Lifting this homotopy through the covering map $\partial_H Y \to Y$ $\pi(a) \subset \partial Y$, it follows that $\pi(a')$ is an inessential arc in Y. We can then decrease the number of components of $D \cap \mathrm{Fr}(B)$, contradicting that this number is minimal.

So, $a' \subset R$. Again by minimality of the intersection, the endpoints of a' lie on opposite short sides of R, so α is homotopic to a' through arcs in R with endpoints on $\mathrm{Fr}(B)$. Since a' is homotopic rel endpoints to $a \subset \mathrm{Fr}(B)$, it follows that α is homotopic rel endpoints into $\mathrm{Fr}(B)$. If the two endpoints of α lie on the same component of $\partial(\partial_H B)$, we are in case (1), and otherwise we are in case (2). □

2.8 Laminations

We assume the reader is familiar with geodesic and measured laminations on finite-type hyperbolic surfaces. See, e.g., [16; 33].

Suppose λ is a connected geodesic lamination on a finite-type hyperbolic surface S with geodesic boundary. We say that λ *fills* an essential subsurface $T \subset S$ if $\lambda \subset T$ and λ intersects every essential, nonperipheral simple closed curve in T.

Fact 2.13 *For every connected λ, there is a unique subsurface with geodesic boundary (as in Section 2.1) that is filled by λ, which we denote by $S(\lambda)$. It is the minimal subsurface with geodesic boundary in S that contains λ.*

Here, $S(\lambda)$ can be constructed by taking a component $\tilde{\lambda} \subset \widetilde{S} \subset \mathbb{H}^2$ of the preimage of λ, letting $C \subset \mathbb{H}^2$ be the convex hull of the set of endpoints of leaves of $\tilde{\lambda}$ in $\partial\mathbb{H}^2$, and projecting C into S.

Suppose that M is a compact, orientable irreducible 3-manifold let $S \subset \partial M$ be an essential subsurface. The *limit set* of (S, M) is the closure

$$\Lambda(S, M) = \overline{\{\text{meridians } \gamma \subset S\}} \subset \mathcal{PML}(S),$$

where $\mathcal{PML}(S)$ is the space of projective measured laminations on S. The limits set was first studied by Masur [46] in the case that M is a handlebody, with S its entire boundary. In this case, Kerckhoff [34] later proved that the limit set has measure zero in $\mathcal{PML}(S)$, although a mistake in his argument was later found and fixed by Gadre [23].

In some ways, $\Lambda(S, M)$ acts as a dynamical limit set. For instance, let $\mathrm{Map}(S)$ be the mapping class group of S, and let $\mathrm{Map}(S, M) \subset \mathrm{Map}(S)$ be the subgroup consisting of mapping classes represented by restrictions of homeomorphisms of M. Then we have:

Fact 2.14 (1) *If $\Lambda(S, M)$ is nonempty, it is the smallest nonempty closed subset of $\mathcal{PML}(S)$ that is invariant under* $\mathrm{Map}(S, H)$.

(2) *If* $\mathrm{Map}(S, M)$ *contains a pseudo-Anosov map on S, then $\Lambda(S, M)$ is the closure of the set of the attracting and repelling laminations of pseudo-Anosov elements of* $\mathrm{Map}(S, M)$.

Note that $\mathrm{Map}(S, M)$ contains a pseudo-Anosov map on S if and only if the disk set $\mathcal{D}(S, M)$ has infinite diameter in the curve complex $\mathcal{C}(S)$, where the latter condition was discussed earlier in Proposition 3.1. See also [2; 41].

Proof For the first part just note that Dehn twist T_m around meridians $m \subset S$ are in $\mathrm{Map}(S, M)$, so if $A \subset \mathcal{PML}(S)$ is nonempty and invariant, $\lambda \in A$ and m is a meridian, then $m = \lim_i T_m^i(\lambda)$ is also in A, implying $\Lambda(S, M) \subset A$.

For the second part, take a pseudo-Anosov $f \in \mathrm{Map}(S, M)$ with attracting lamination λ_+, say. If m is a meridian in S, then $T_m^i \circ f \circ T_m^{-i}$ are pseudo-Anosov maps on S and their attracting laminations converge to m, and then the argument finishes as before. □

2.9 Laminations on interval bundles

Suppose that Y is a compact hyperbolisable surface with boundary, and that $B \to Y$ is an interval bundle over Y. Endow Y and the horizontal boundary $\partial_H B$ with arbitrary hyperbolic metrics such that the boundary components are all geodesic.

Suppose we have two geodesic laminations $\lambda_\pm$ on $\partial_H B$.

Definition 2.15 We say that $\lambda_\pm$ are *essentially homotopic through B* if there is a lamination λ and a homotopy $h_t : \lambda \to B$, $t \in [-1, 1]$, such that $h_{\pm 1}$ is a homeomorphism onto $\lambda_\pm$, and where (h_t) is not homotopic into $\partial_H B$.

When B is a trivial interval bundle, $\lambda_\pm$ are essentially homotopic through B if and only if we can write $B \cong Y \times [0, 1]$ in such a way that $\lambda_\pm = \lambda \times \{\pm 1\}$ for some geodesic lamination on Y. This is an easy consequence of the fact that on a surface, homotopic laminations are isotopic. In general:

Fact 2.16 *Suppose that $\lambda_\pm$ are disjoint or equal geodesic laminations on $\partial_H B$. Then these are equivalent*:

(1) *$\lambda_\pm$ are essentially homotopic through B.*

(2) *$\lambda_\pm$ is isotopic on $\partial_H B$ to $\sigma(\lambda_\mp)$, where σ is the canonical involution of B discussed in Section 2.5.*

Moreover, (1) *and* (2) *imply*

(3) *there is a geodesic lamination $\bar{\lambda}$ on Y such that $\lambda_- \cup \lambda_+$ is isotopic on $\partial_H B$ to $(\pi|_{\partial_H B})^{-1}(\bar{\lambda})$.*

Here, (3) does not always imply (1) and (2), since it could be that $\bar{\lambda}$ has two components, $(\pi|_{\partial_H B})^{-1}(\bar{\lambda})$ has four, and these components are incorrectly partitioned into the two laminations $\lambda_\pm$. However, that's the only problem, so, for instance, if $\lambda_\pm$ are minimal then (1)–(3) are equivalent.

While we have phrased things more generally in the section, we can always assume in proofs that our hyperbolic metrics have been chosen so that the covering map $\pi|_{\partial_H B} : \partial_H B \to Y$ is locally isometric. Here, we're using the fact that given two hyperbolic metrics with geodesic boundary on a compact surface, a geodesic lamination with respect to one metric is isotopic to a unique geodesic lamination with respect to the other hyperbolic metric. In this case, we can remove the word "isotopic" from (2) and (3).

Proof The fact is trivial when B is a trivial interval bundle. When B is nontrivial, lift the homotopy to the trivial interval bundle $B' \to B$ that double covers B, giving homotopic laminations $\lambda'_\pm \subset \partial_H B'$. The statement (1) $\Longleftrightarrow$ (2) follows since the canonical involution on B' covers that of B. For (2) $\Longrightarrow$ (3), note that since $\lambda_\pm$ are disjoint or equal and differ by σ, their projections $\pi(\lambda_\pm) \subset Y$ are the same, and are a geodesic lamination $\bar{\lambda}$ on Y. □

3 Large and small disk sets and compression bodies

Suppose that $S \subset \partial M$ is an essential subsurface of the boundary of a compact, irreducible 3-manifold M, and that ∂S is incompressible in M. The following is probably known to some experts, but we don't think it appears anywhere in the literature, so we give a complete proof.

Proposition 3.1 (diameters of disk sets) *With M, S as above, either*

(1) *$\mathcal{D}(S, M)$ has infinite diameter in $\mathcal{C}(S)$,*

(2) *S has one nonseparating meridian δ, and every other meridian is a band sum of δ,*

(3) *S has a single meridian, which is separating, or*

(4) *$\mathcal{D}(S, M) = \varnothing$.*

In case (1) we will say that $\mathcal{D}(S, M)$ is *large*, and in cases (2)–(4), we will say that $\mathcal{D}(S, M)$ is *small*. Similarly, if $C(S, M)$ is the characteristic compression body defined in Fact 2.4, then $C(S, M)$ is said to be large or small depending on whether $\mathcal{D}(S, M)$ is large or small. See also the discussion of small compression bodies in Section 3 of [3].

Here, recall that a *band sum* of a meridian δ is the boundary of a regular neighbourhood of $\delta \cup \beta$, where β is a simple closed curve on S that intersects δ once. Any such band sum must be a meridian: for instance, as an element of $\pi_1 M$ it is a commutator with a trivial element. Also, (3) includes the case when M is a solid torus and $S = \partial M$, in which case there is only one (nonseparating) meridian. When M is not a solid torus, though, every nonseparating curve has infinitely many band sums. Similarly, a *band sum* of two disjoint meridians δ and γ over an arc k joining δ to γ is the boundary of a regular neighbourhood of $\delta \cup k \cup \gamma$.

Before beginning the proof, we first establish the following:

Claim 3.2 *Suppose S is not a torus, $\gamma \subset S$ is a meridian on S and δ is a meridian that lies in a component $T \subset S \setminus \gamma$. If γ is not a band sum of δ, there is a pseudo-Anosov $f : T \to T$ that extends to a homeomorphism of M.*

The condition that γ is not a band sum is necessary. For if M is a handlebody with $S = \partial M$, and γ is a separating meridian that bounds a compressible punctured torus $T \subset S$, then T has only a single meridian δ. This δ is nonseparating and γ is a band sum of δ. Any map $T \to T$ that extends to a homeomorphism of M must then fix δ, so cannot be pseudo-Anosov.

Similarly, if S is a torus and γ is a meridian, the complement of γ is an annulus, which does not admit any pseudo-Anosov map.

Proof of Claim 3.2 Suppose first that γ is not separating. Any simple closed curve that intersects γ once can be used to create a band sum. Now S is not a torus, and cannot be a punctured torus either, since then its boundary would be compressible. So, there are a pair α, β of band sums of γ that fill $S \setminus \gamma$. By a theorem of Thurston [22, III.3 in 13], the composition of twists $T_\alpha \circ T_\beta^{-1}$ is pseudo-Anosov. Each twist extends to M, because twist about meridians can be extended to twists along the disks they bound.

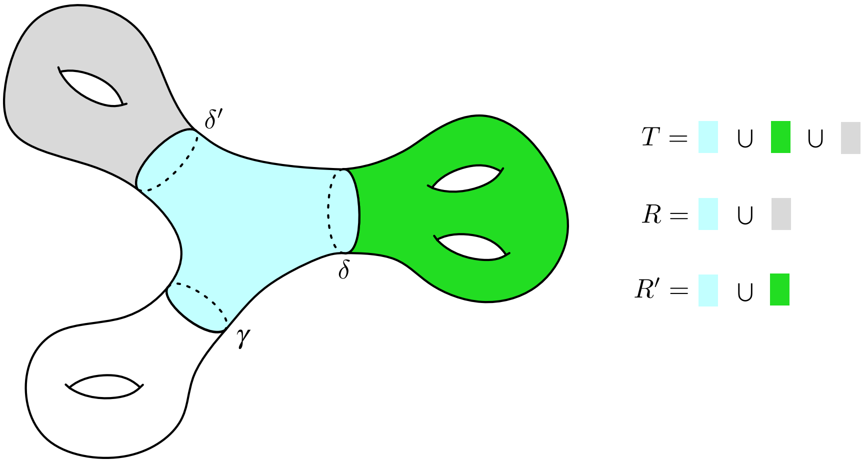


Figure 3: The surfaces R and R' fill T.

Now suppose γ separates S, and suppose that R is the component of $T \setminus (\gamma \cup \delta)$ adjacent to γ and δ. Any curve in R that bounds a pair of pants with γ and δ is also a meridian. Such curves are constructed as the boundary of a neighbourhood of the union of γ, δ and any arc in R joining the two. Therefore, there is a pair α, β of such curves that fills R. As before, $f = T_\alpha \circ T_\beta^{-1}$ is a pseudo-Anosov on R that extends to M.

However, there was nothing special about δ in the above construction. So if there is some (nonperipheral) meridian $\delta' \subset T$ with $\delta \neq \delta'$, there is also a pseudo-Anosov f' on the corresponding surface R', such that f' extends to M. Since R and R' fill T, [19, Theorem 6.1] says that for large i the composition $f^i (f')^i$ is a pseudo-Anosov on T. See Figure 3.

The only case left to consider is when δ is the only (nonperipheral) meridian in T. Since new meridians can be created by taking a band sum of δ and γ over any arc joining them, the only possibility here is that T is a punctured torus, in which case this construction always just produces γ again. But then γ is a band sum of δ. □

Proof of Proposition 3.1 When S is a torus, distinct curves have nonzero algebraic intersection number, so either there are no meridians or there is a single meridian. So, we assume $S \neq T^2$ below.

It follows easily from Claim 3.2 that if S contains two disjoint meridians, neither of which is a band sum of the other, then $\mathcal{D}(S, M)$ has infinite diameter in the curve complex. We first claim that if there are two disjoint meridians in S, neither of which is a band sum of the other, then $\mathcal{D}(S, M)$ has infinite diameter in the curve complex. To see this, suppose γ_1, γ_2 are such meridians. Claim 3.2 gives two pseudo-Anosov maps f_1, f_2, each defined on the component of $S \setminus \gamma_i$ that contains γ_j, where $i \neq j$. Since the component of $S \setminus \gamma_1$ containing γ_2 and the component of $S \setminus \gamma_2$ containing γ_1 together fill S, for large k the composition $f_1^k f_2^k$ is a pseudo-Anosov map on the entire surface S, by [19, Theorem 6.1]. Any such composition extends to M, so maps meridians to meridians. As pseudo-Anosovs act with unbounded orbits on the curve complex [47], this implies that the set of meridians has infinite diameter in $\mathcal{C}(S)$.

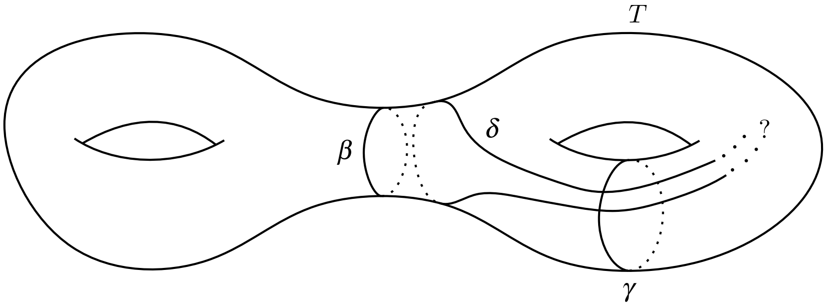


Figure 4: A surgery of a curve δ along a nonseparating γ cannot produce a curve β that is a band sum of γ.

Starting now with the proof of the proposition, suppose there are *no nonseparating meridians* in S. If γ, δ are distinct (separating) meridians, then an innermost disk surgery produces another separating meridian γ_2 disjoint from γ_1; see [3, Lemma 2.8]. By the previous paragraph, $\mathcal{D}(S,M)$ has infinite diameter in the curve complex. So, the only other options are if $\mathcal{D}(S,M)=\varnothing$, or if the only meridian is a single separating curve.

Suppose now that there is a nonseparating meridian γ in S. By Claim 3.2, unless the disc set has infinite diameter in the curve complex, any meridian disjoint from γ must be a band sum of γ. So, either we are in case (2) of the proposition, or there is some meridian δ that intersects γ. Any innermost disk surgery of δ along γ must produce a band sum β of γ. However, this β must then bound a punctured torus T containing γ, and δ is then forced to lie inside T, which gives a contradiction; see Figure 4. □

4 Windows from limit sets

Let $N=\Gamma\backslash\mathbb{H}^3$ be an orientable, geometrically finite hyperbolic 3-manifold, let $\Lambda\subset\partial\mathbb{H}^3$ be the limit set of Γ, and let

$$\mathrm{CC}(N):=\Gamma\backslash\mathrm{CH}(\Lambda)\subset N$$

be the convex core of N. Equip $\partial\mathrm{CC}(N)$ with its intrinsic length metric, which is hyperbolic; see, for instance, [58, Proposition 8.5.1].

Let $S_\pm$ be (possibly degenerate) incompressible subsurfaces with geodesic boundary in $\partial\mathrm{CC}(N)$ that are either equal or are essentially disjoint, as in Section 2.1. Let

$$\widetilde{S}_\pm\subset\partial\mathrm{CH}(\Lambda)\subset\mathbb{H}^3$$

be lifts of $S_\pm$, where if $S_-=S_+$, we require that $\widetilde{S}_-\neq\widetilde{S}_+$. Let $\Gamma_\pm\subset\Gamma$ be the stabilisers of $\widetilde{S}_\pm$, let $\Lambda_\pm\subset\partial\mathbb{H}^3$ be their limit sets and $\Delta=\Gamma_+\cap\Gamma_-$.

The lift $\widetilde{S}_\pm$ is isometric to a convex subset of $\mathbb{H}^2$. Let $\partial_\infty S_\pm\subset\partial\mathbb{H}^2$ be the boundary of $\widetilde{S}_\pm$. By [53, Theorem 5.6], say, the inclusion $\widetilde{S}_\pm\hookrightarrow\mathbb{H}^3$ extends continuously to a $\Gamma_\pm$-equivariant quotient map

$$\iota_\pm:\partial_\infty\widetilde{S}_\pm\to\Lambda_\pm\subset\partial\mathbb{H}^3.$$

Theorem 4.1 (windows from limit sets) *We have* $\Lambda_- \cap \Lambda_+ = \Lambda_\Delta$. *Next, suppose* Δ *is nonempty and is not a cyclic group acting parabolically on either* $\widetilde{S}_-$ *or* $\widetilde{S}_+$, *and let* $\widetilde{C}_\pm \subset \widetilde{S}_\pm$ *be the convex hulls of the subsets* $\iota_\pm^{-1}(\Lambda_\Delta) \subset \partial_\infty \widetilde{S}_\pm$. *Then* $\widetilde{C}_\pm$ *are* Δ*-invariant, the quotients* $C_\pm := \Delta \backslash \widetilde{C}_\pm$ *are (possibly degenerate) subsurfaces with geodesic boundary in* $S_\pm$, *and there is an essential homotopy from* C_- *to* C_+ *in* $\mathrm{CC}(N)$ *that is the projection of a homotopy from* $\widetilde{C}_-$ *to* $\widetilde{C}_+$.

Above, $C_\pm$ are (possibly degenerate) subsurfaces with geodesic boundary in $S_\pm$, as defined in Section 2.1, but it follows from the above and Theorem 2.8 that there are "resolutions" (see Section 2.1) $C'_\pm \subset S_\pm$ such that $C'_\pm$ bound an interval bundle in $\mathrm{CC}(M)$. So informally, the theorem says that the intersection $\Lambda_- \cap \Lambda_+$ is exactly the limit set of the fundamental group of some essential interval bundle in $(\mathrm{CC}(M), S'_- \cup S'_+)$. The term "window" comes from Thurston [59] and refers to interval bundles; for example, one can "see through" a trivial interval bundle from one horizontal boundary component to the other.

The assumption that Δ is not cyclic and acting parabolically on either $\widetilde{S}_\pm$ is just for convenience in the statement of the theorem. (Just to be clear, note that an element $\gamma \in \Delta$ can act parabolically as an isometry of $\mathbb{H}^3$, but hyperbolically on the convex subsets $\widetilde{S}_\pm \subset \mathbb{H}^2$.) If Δ is cyclic and acts parabolically on $\widetilde{S}_+$ the subset $\widetilde{C}_+$ in the statement of the theorem will be empty. However, using the same proof one can construct a homotopy from a simple closed curve on S_+ bounding a cusp of S_+ to some simple closed curve on S_-.

As mentioned in the introduction, a version of Theorem 4.1 was known to Thurston; see his discussion of the only windows break theorem in [59]. Precise statements for geometrically finite N without accidental parabolics were worked out in Lecuire's thesis [40] and by Walsh [60]; note that Walsh uses the conformal boundary instead of the convex core boundary, but the two points of view are equivalent. However, for our applications in this paper, we need to allow accidental parabolics in $\widetilde{S}_\pm$, which are not allowed in those theorems. Also, our proof is more direct and natural[5] than those in [40; 60], despite the extra complication coming from parabolics.

Finally, the assumption that N is geometrically finite is not really essential for the theorem statement. With a bit more work dealing with degenerate ends, one can prove the theorem for all finitely generated Γ. Essentially, the point is to use Canary's covering theorem [14] to show that degenerate NP-ends in the covers $N_\pm := \Gamma_\pm \backslash \mathbb{H}^3$ have neighbourhoods that embed in N, and then to use this to prove that geodesic rays in $\mathbb{H}^3$ that converge to points in $\Lambda_- \cap \Lambda_+$ cannot exit degenerate ends in $N_\pm$. After showing this, the proof of Claim 4.3 extends to the general case. However, we don't have an application for that theorem in mind, so we'll spare the reader the details.

Proof of Theorem 4.1

We first focus on proving that $\Lambda_- \cap \Lambda_+ = \Lambda_\Delta$. For each $\xi \in \partial\mathbb{H}^3$, let $\Gamma_\pm(\xi) \subset \Gamma_\pm$ be the stabiliser of ξ.

---

[5] In both [40; 60], the authors focus on proving that the boundary components of $\widetilde{C}_\pm$ project to simple closed curves in $S_\pm$, but that isn't sufficient to say that $\widetilde{C}_\pm$ projects to a subsurface with geodesic boundary in $S_\pm$, which is what they then claim. E.g., in [60] it is stated that under a covering map, the boundary of a subset goes to the boundary of the image, but this isn't true.

Claim 4.2 *Let $\xi \in \partial\mathbb{H}^3$ and suppose that $\Gamma_-(\xi)$ and $\Gamma_+(\xi)$ are both nontrivial. Then they are equal.*

Proof By the tameness theorem [1; 11], we can identify $\mathrm{CC}(N)$ topologically with a subset of a 3-compact manifold with boundary M, where

$$\mathrm{CC}(N) \supset \mathrm{int}(M), \quad \mathrm{CC}(N) \cap \partial M = \partial\mathrm{CC}(N), \tag{1}$$

and where $\partial\mathrm{CC}(N)$ is a collection of essential subsurfaces of $\partial\overline{N}$. Let $\partial_{\chi=0}M$ be the union of all torus boundary components of M, and let (X, Σ) be the characteristic submanifold of the pair $(M, S_- \cup S_+ \cup \partial_{\chi=0}M)$, as in Section 2.6.

Since $\Gamma_\pm$ are both contained in a discrete group Γ, both $\Gamma_\pm(\xi)$ are contained in the stabiliser $\Gamma(\xi)$, which is either infinite cyclic, or rank-2 parabolic.

Suppose first that $\Gamma(\xi)$ is rank-2 parabolic. The groups $\Gamma_\pm(\xi)$ are both cyclic, since $S_\pm$ are incompressible hyperbolic surfaces, so their fundamental groups do not contain $\mathbb{Z}^2$ subgroups. So, we can write $\Gamma_\pm(\xi) = \langle\gamma_\pm\rangle$ for closed curves $\gamma_\pm$ on $S_\pm$. Both $\gamma_\pm$ are homotopic into some fixed component $T \subset \partial_{\chi=0}M$, the component whose fundamental group can be conjugated to stabilise ξ. So, there is a component $(X_0, \Sigma_0) \subset (X, \Sigma)$ of the characteristic submanifold such that Σ_0 intersects T and both $\gamma_\pm$ are homotopic on $S_\pm$ into Σ_0. Since $M \ncong T^2 \times [0, x]$, the component (X_0, Σ_0) is either an interval bundle over an annulus (so, a fibred solid torus), or an S^1-bundle pair, so by Fact 2.9, X_0 is either a fibred solid torus or a thickened torus. In either case, Σ_0 intersects each of $S_\pm$ in a fibred annulus, and these annuli are disjoint, so they are parallel on a torus boundary component of X_0, implying that $\gamma_\pm$ are homotopic in M, and hence $\Gamma_\pm(\xi)$ are conjugate in Γ. But since $\Gamma_\pm(\xi)$ have the same fixed point at infinity, the conjugating element must fix ξ, and therefore commute with the two groups, implying $\Gamma_-(\xi) = \Gamma_+(\xi)$.

Now assume $\Gamma(\xi)$ is cyclic. Pick a basepoint $p \in S_-$, say, and let $\gamma_- \subset S_-$ be a loop based at p representing a generator of $\Gamma_-(\xi)$. Represent a generator of $\Gamma_+(\xi)$ as $\alpha \cdot \gamma_+ \cdot \alpha^{-1}$, where α is an arc from $p \in S_-$ to a point in S_+, and γ_+ is a loop in S_+. Since $\Gamma_\pm$ stabilise distinct components $\widetilde{S}_\pm$, the arc α is not homotopic into $S_- \cup S_+$. So, α is a spanning arc of an essential map from an annulus, where the boundary components of the annulus map to powers of $\gamma_\pm$. It follows that the loops $\gamma_\pm$ are homotopic on $S_\pm$ into Σ_0 for some component $(X_0, \Sigma_0) \subset (X, \Sigma)$.

If X_0 is an I-bundle with horizontal boundary Σ_0, then as $\gamma_\pm$ are not proper powers in $\pi_1 S_\pm$, they are both primitive in $\pi_1 X_0$, and hence $\gamma_\pm$ (rather than their powers) are homotopic in $X_0 \subset M$. Similarly, if (X_0, Σ_0) is a fibred solid torus, Σ_0 is a collection of parallel annuli on ∂X_0, so since $\gamma_\pm$ are primitive in $\pi_1 S_\pm$, they are homotopic on $S_\pm$ to *simple* closed curves in these annuli, and hence are homotopic to each other in X_0.

It follows that there are generators for $\Gamma_\pm(\xi)$ that are conjugate in Γ, but since these generators both fix ξ, they are equal. □

Claim 4.3 *For all $\xi \in \Lambda_- \cap \Lambda_+$, we have $\Gamma_-(\xi) = \Gamma_+(\xi)$. Moreover,*

$$\Lambda_\Delta = \Lambda_- \cap \Lambda_+.$$

Proof Let $N_\pm \subset \mathbb{H}^3$ be the 1-neighbourhood of the convex hull of $\Lambda_\pm$, and for small $\epsilon > 0$, let $T_\pm(\epsilon) \subset \mathbb{H}^3$ be the set of all points that are translated less than ϵ by some parabolic element of $\Gamma_\pm$. If ϵ is at most the Margulis constant ϵ_0, then $T_\pm(\epsilon)$ is a disjoint union of horoballs in $\mathbb{H}^3$.

The sets $N_\pm$ and $T_\pm(\epsilon)$ are $\Gamma_\pm$ invariant. Since $\Gamma_\pm$ is a finitely generated subgroup of Γ, which is geometrically finite, $\Gamma_\pm$ is geometrically finite as well by [14]. So, the action of $\Gamma_\pm$ on $N_\pm \setminus T_\pm(\epsilon)$ is cocompact, see, e.g., Theorem 3.7 in [50], implying that either the function

$$D_+ : \mathbb{H}^3 \to \mathbb{R}_{>0}, \quad D_+(x) = \min\{d(x, \gamma(x)) \mid \gamma \in \Gamma_+ \text{ loxodromic}\},$$

is bounded above on $N_+ \setminus T_+(\epsilon)$ by some $B(\epsilon) > 0$, or Γ_+ is elementary parabolic. A similar statement holds for $-$ instead of $+$. With ϵ_0 the Margulis constant, the Margulis lemma then implies that *if* $\epsilon > 0$ *is sufficiently small with respect to* $B(\epsilon_0)$, *and* Γ_+ *is not elementary parabolic*, *then*

$$T_-(\epsilon) \cap N_+ \subset T_+(\epsilon_0), \tag{2}$$

and similarly with $-$, $+$ exchanged. Indeed, if not then we have (say) a point $p \in \mathbb{H}^3$ that is translated by less than ϵ by some parabolic $\gamma_- \in \Gamma_-$ and by at most B by some loxodromic $\gamma_+ \in \Gamma_+$. If ϵ is small with respect to B, then both γ_- and $[\gamma_+, \gamma_-]$ translates p by at most ϵ_0, so they generate an elementary discrete group by the Margulis lemma applied to Γ, implying that γ_+ fixes the fixed point of γ_-, which contradicts that they generate a discrete group.

Fix $\xi \in \Lambda_+ \cap \Lambda_-$. We claim that $\Gamma_-(\xi) = \Gamma_+(\xi)$. By Claim 4.2 it suffices to show that whenever $\Gamma_-(\xi)$ is nontrivial, say, so is $\Gamma_+(\xi)$.

First, assume that $\Gamma_-(\xi)$ is elementary parabolic. We claim that $\Gamma_+(\xi)$ is elementary parabolic as well. Assume not, and let α be a geodesic ray in $\mathbb{H}^3$ converging to ξ. Then $\alpha(t)$ lies in $T_-(\epsilon) \cap N_+$ for large t, and therefore in $T_+(\epsilon_0)$ for large t by (2), which implies ξ is a parabolic fixed point of Γ_+ as well, a contradiction.

Next, suppose that $\Gamma_-(\xi)$ is elementary loxodromic. If ξ is a parabolic fixed point of Γ_+, we are done, so let's assume this isn't the case. Let α be the axis of $\Gamma_-(\xi)$, parametrised so $\alpha(t) \to \xi$ as $t \to \infty$. Since ξ is not a Γ_+ parabolic fixed point, there are $t_i \to \infty$ such that $\alpha(t_i) \notin T_+(\epsilon)$ for all i. Since the action of Γ_+ on $N_+ \setminus T_+(\epsilon)$ is cocompact, if $p \in \mathbb{H}^3$ is a fixed basepoint, there are elements $\gamma_i^+ \in \Gamma_+$ such that $\sup_i d(\gamma_i^+(p), \alpha(t_i)) < \infty$. Since the action of $\Gamma_-(\xi)$ on α is cocompact, there are then elements $\gamma_i^- \in \Gamma_-(\xi)$ with

$$\sup_i d(\gamma_i^+(p), \gamma_i^-(p)) < \infty.$$

By discreteness of Γ, after passing to a subsequence we can assume $\gamma_i^+ = \gamma_i^- \circ g$ for some fixed $g \in \Gamma$. Hence, for all i we have

$$\gamma_i^+ \circ (\gamma_1^+)^{-1} = \gamma_i^- \circ (\gamma_1^-)^{-1} \in \Gamma_+ \cap (\Gamma_-(\xi)) \subset \Gamma_+(\xi),$$

so we are done.

Finally, we want to show that $\Lambda_- \cap \Lambda_+ = \Lambda_\Delta$. The inclusion $\Lambda_\Delta \subset \Lambda_- \cap \Lambda_+$ is clear. So, take $\xi \in \Lambda_- \cap \Lambda_+$. We can assume that $\Gamma_\pm(\xi) = 1$, since otherwise we're in the cases handled above. Let α be a geodesic ray in $\mathbb{H}^3$ converging to ξ. As in the previous case, since ξ is not a parabolic fixed point of Γ_+, there are $t_i \to \infty$ such that $\alpha(t_i) \notin T_+(\epsilon_0)$ for all i. Discarding finitely many i, we have $\alpha(t_i) \in N_+$, so it follows from (2) that $\alpha(t_i) \notin T_-(\epsilon)$. Fixing a base point $p \in \mathbb{H}^3$, as Γ_- acts cocompactly on $N_- \setminus T_-(\epsilon)$ and Γ_+ acts cocompactly on $N_+ \setminus T_+(\epsilon_0)$, there are elements $\gamma_i^\pm \in \Gamma_\pm$ such that

$$\sup_i d(\gamma_i^\pm(p), \alpha(t_i)) < \infty.$$

So passing to a subsequence, $\gamma_i^+ = \gamma_i^- \circ g$ for some fixed $g \in \Gamma$, and then

$$\gamma_i^+ \circ (\gamma_1^+)^{-1} = \gamma_i^- \circ (\gamma_1^-)^{-1} \in \Gamma_+ \cap \Gamma_- = \Delta$$

for all i. But applying this sequence to p and letting $i \to \infty$ gives a sequence of points in the orbit $\Delta(p)$ that converge to ξ, so $\xi \in \Lambda_\Delta$. ☐

Now assume that $\Delta \neq 1$. We want to construct the interval bundle W mentioned in the statement of the theorem. After an isotopy on $\partial\mathrm{CC}(N)$, let's assume that $S_\pm$ is a subsurface of $\partial\mathrm{CC}(N)$ with geodesic boundary. Consequently, we allow degenerate subsurfaces, where $S_\pm$ is a simple closed geodesic, as well as subsurfaces where only the interior is embedded and two boundary components can coincide. As $\partial\mathrm{CC}(N)$ may have cusps, we also must allow $S_\pm$ to be noncompact with finite volume, rather than compact.

Recall that $\widetilde{S}_\pm$ is isometric to a convex subset of $\mathbb{H}^2$, and that if $\partial_\infty S_\pm \subset \partial\mathbb{H}^2$ is the boundary of $\widetilde{S}_\pm$ the inclusion $\widetilde{S}_\pm \hookrightarrow \mathbb{H}^3$ extends continuously to a $\Gamma_\pm$-equivariant quotient map

$$\iota_\pm : \partial_\infty \widetilde{S}_\pm \to \Lambda_\pm \subset \partial\mathbb{H}^3.$$

Moreover, if $\xi, \xi' \in \partial_\infty \widetilde{S}_+$, say, we have $\iota_+(\xi) = \iota_+(\xi')$ if and only if there is an element $\gamma \in \Gamma_+$ that acts hyperbolically on $\widetilde{S}_+ \cup \partial_\infty \widetilde{S}_+$ with fixed points $\xi, \xi' \in \partial_\infty \widetilde{S}_+$, but acts parabolically on $\mathbb{H}^3$. By discreteness of the action $\Gamma_+ \curvearrowright \widetilde{S}_+$, each $\xi \in \partial_\infty \widetilde{S}_+$ has the same image under ι_+ as *at most one* other ξ'. Similar statements holds with $-$ instead of $+$. All this is a consequence (for instance) of Bowditch's theory of the boundary of a relatively hyperbolic group [6]: since the action $\Gamma_\pm \curvearrowright \mathbb{H}^3$ is geometrically finite, $\Lambda_\pm$ is a model for the Bowditch boundary of the group $\Gamma_\pm$ relatively to its maximal parabolic subgroups, so the statement above follows from Theorem 1.3 of [44], say.[6]

Let $\iota_\pm^{-1}(\Lambda_\Delta) \subset \partial_\infty \widetilde{S}_\pm$. Since $\Delta \neq 1$ and is not cyclic parabolic, $\iota_\pm^{-1}(\Lambda_\Delta)$ has at least two points, so it has a well-defined convex hull $\widetilde{C}_\pm \subset S_\pm$.

Claim 4.4 (convex hulls) *One of the following holds.*

1. *Δ is cyclic and acts hyperbolically on $\widetilde{S}_+$. The convex hull $\widetilde{C}_+$ is its geodesic axis, which is precisely invariant under $\Delta \subset \Gamma$, so that the quotient $C_+ := \Delta \backslash \widetilde{C}_+$ embeds as a simple closed geodesic in S_+.*
2. *$\widetilde{C}_+$ is a subsurface of $\widetilde{S}_+$ with geodesic boundary,* $\mathrm{int}(\widetilde{C}_+)$ *is precisely invariant under $\Delta \subset \Gamma$, and the quotient $C_+ := \Delta \backslash \widetilde{C}_+$ is a generalised subsurface of S_+ with compact geodesic boundary.*

A similar statement holds with $-$ instead of $+$.

---

[6] See also Theorem 5.6 of [53], which says that there is a continuous equivariant extension $\iota_\pm$ of the inclusion $\widetilde{S}_\pm \hookrightarrow \mathbb{H}^3$ as above. This theorem is stated in a much more general setting, though, and our statement is a trivial case.

Proof Let's work with $+$ for concreteness. If $g \in \Delta$, then $g(\Lambda_\Delta) = \Lambda_\Delta$, so g leaves $\iota_+^{-1}(\Lambda_\Delta)$ invariant by equivariance of ι_+. Hence g leaves $\widetilde{C}_+$ invariant.

Let's suppose first that $\widetilde{C}_+$ has nonempty interior, since that is the more interesting case. We'll address the case that $\widetilde{C}_+$ is a biinfinite geodesic at the end of the proof. Let $g \in \Gamma_+ \setminus \Delta$. We want to show

$$g(\operatorname{int}(\widetilde{C}_+)) \cap \operatorname{int}(\widetilde{C}_+) = \varnothing.$$

Assume this is not the case. By Claim 4.3, the fixed points of g in $\partial_\infty S_+$ lie outside $\iota_+^{-1}(\Lambda(\Delta))$. So, we cannot have $g(\widetilde{C}_+) \subset \widetilde{C}_+$, as then we'd have $g^n(\widetilde{C}_+) \subset \widetilde{C}_+$ for all n, contradicting that points of $\widetilde{S}_+$ converge to the fixed points of g under iteration. Considering backwards iterates, we also cannot have $\widetilde{C}_+ \subset g(\widetilde{C}_+)$. Therefore, $\partial\widetilde{C}_+$ and $\partial g(\widetilde{C}_+)$ intersect transversely.

Since $\widetilde{C}_+$ has nonempty interior, Δ is nonelementary, and therefore the fixed points of loxodromic isometries of Δ are dense in Λ_Δ. Loxodromic fixed points of Δ are in particular *not* parabolic fixed points in $\Gamma_\pm$, so any biinfinite geodesic in $\widetilde{C}_+$ is a limit of biinfinite geodesics in $\widetilde{C}_+$ whose endpoints are *not* fixed points of parabolic isometries of $\Gamma_\pm$. By the previous paragraph, there are then biinfinite geodesics α_+, β_+ in $\widetilde{C}_+$ such that $g(\alpha_+)$ and β_+ intersect transversely, and where the endpoints of α_+, β_+ project under ι_+ to points $\xi_\alpha, \xi'_\alpha, \xi_\beta, \xi'_\beta \in \Lambda_\Delta$ that are not parabolic fixed points in $\Gamma_\pm$.

Let α_- be the geodesic in $\widetilde{S}_-$ whose endpoints in $\partial_\infty\widetilde{S}_-$ map to the points ξ_α, ξ'_α under ι_-. Define β_- similarly. Then

$$\alpha := \alpha_+ \cup \{\xi_\alpha, \xi'_\alpha\} \cup \alpha_-, \quad \beta := \beta_+ \cup \{\xi_\beta, \xi'_\beta\} \cup \beta_-$$

are two *simple* closed curves on the closure $\operatorname{cl}(\partial \mathrm{CH}(\Lambda_\Gamma)) \subset \mathbb{H}^3 \cup \partial\mathbb{H}^3$, which is homeomorphic to a sphere. For instance, the arcs $\alpha_\pm$ are disjoint and $\xi_\alpha \neq \xi'_\alpha$, since the endpoints of α_+ are not parabolic fixed points.

Now consider how the two simple closed curves $g(\alpha)$, β intersect. The arcs β_- and $g(\alpha_+)$ are disjoint since $\widetilde{S}_- \neq \widetilde{S}_+$. The arcs $g(\alpha_-)$, β_- are disjoint since $g \notin \Gamma_-$ and hence $g(\alpha_-)$ lies on a different translate of $\widetilde{S}_-$ than $\beta_- \subset \widetilde{S}_-$. Moreover, since $g(\alpha_+)$, β_+ intersect transversely in $\widetilde{S}_+$, the endpoints of $g(\alpha_+)$ and β_+ are distinct in $\partial_\infty\widetilde{S}_+$, and since none of them are parabolic fixed points, the points $g(\xi_\alpha)$, $g(\xi'_\alpha)$, ξ_β, ξ'_β are all distinct. But by assumption, $g(\alpha_+)$ intersects β_+ transversely in a single point! This shows that $g(\alpha)$ and β intersect exactly once, transversely, which is a contradiction.

By precise invariance of the action on the interior, the quotient $\operatorname{int}(C_+) = \Delta \setminus \operatorname{int}(\widetilde{C}_+)$ embeds in the finite volume surface S_+, so C_+ has finite volume itself. So if ∂C_+ is noncompact, it must have two noncompact boundary components that are asymptotic. Lifting, we get two boundary components β_1, β_2 of $\widetilde{C}_+$ that are asymptotic. Since $\widetilde{C}_+$ is convex, it is contained in the subset of $\mathbb{H}^2$ bounded by β_1, β_2, and hence the common endpoint of β_1, β_2 is an isolated point of Λ_Δ, which is a contradiction since Δ is not elementary.

The case when $\widetilde{C}_+$ is a biinfinite geodesic is similar. Here, Δ must be cyclic, acting on $\widetilde{S}_+$ with axis $\widetilde{C}_+$, and acting either parabolically or loxodromically on $\mathbb{H}^3$. In the parabolic case, $\widetilde{C}_+$ compactifies to a simple closed curve on the sphere $\operatorname{cl}(\mathrm{CH}(\Lambda_\Gamma)) \subset \mathbb{H}^3 \cup \partial\mathbb{H}^3$, so no translate $g(\widetilde{C}_+)$, $g \in \Gamma_+$, can intersect $\widetilde{C}_+$ transversely, since if it did we'd get two simple closed curves on the sphere that intersect once. In the loxodromic case, we get a similar contradiction by looking at the simple closed curve $\mathrm{cl}(\widetilde{C}_+ \cup \widetilde{C}_-) \subset \mathrm{cl}(\partial \widetilde{M})$ and its g-image. So, $\widetilde{C}_+$ is precisely invariant under $\Delta \subset \Gamma$. The quotient $C_+ := \Delta \backslash \widetilde{C}_+$ is obviously compact, and is therefore a simple closed geodesic in S_+. □

We claim that C_- and C_+ are homeomorphic. If $C_\pm$ are isotopic in $\partial \mathrm{CC}(M)$ this is clear, and otherwise we argue as follows. The subgroups $\pi_1 C_\pm$ are both represented by Δ, so are conjugate in $\pi_1 M$. The fact that every curve in C_- is homotopic to a curve in C_+ (and vice versa) implies that $C_\pm$ are isotopic to subsurfaces $C'_\pm \subset \Sigma$ in the boundary Σ of a component (X, Σ) of the characteristic submanifold[7] of $(\mathrm{CC}(M), S_- \cup S_+)$, see Section 2.6, and that even within X every closed curve in C'_- is homotopic to a closed curve in C'_+, and vice versa. When X is a solid torus or thickened torus, $C'_\pm$ are annuli, while if X is an interval bundle, $C'_\pm$ bound a vertical interval bundle in X, and are homeomorphic.

So, let $f : C_- \to C_+$ be a homeomorphism, lift f to a Δ-equivariant homeomorphism $\tilde{f} : \widetilde{C}_- \to \widetilde{C}_+$ and let

$$F : \widetilde{C}_- \times [0,1] \to \mathrm{CH}(\Lambda),$$

where $F(x, \cdot)$ parametrises the geodesic from x to $f(x)$. Then F is Δ-equivariant, and projects to an essential homotopy from C_- to C_+, as desired.

4.1 An annulus theorem for laminations

Suppose M is a compact, orientable, hyperbolisable 3-manifold with nonempty boundary and let $S = \partial_{\chi<0} M$ be the union of all nontorus boundary components of M. When $\alpha, \beta \subset S$ are disjoint simple closed curves that are essential and homotopic in M, but not homotopic in S, the annulus theorem says that there is an essential embedded annulus $A \subset M$ with $\partial A = \alpha \cup \beta$; see Scott's paper [57].

More generally, equip S with an arbitrary hyperbolic metric. An *essential homotopy* between two geodesic laminations $\lambda_\pm$ on S is a map

$$H : (\lambda \times [-1,1], \lambda \times \{-1,1\}) \to (M, S),$$

where λ is a lamination, such that H maps $\lambda \times \{\pm 1\}$ homeomorphically onto $\lambda_\pm$, and where H is not homotopic rel $\lambda \times \{-1,1\}$ into ∂M.

Here is an "annulus theorem" for minimal laminations.

Proposition 4.5 (an annulus theorem for laminations) *Let λ_-, λ_+ be two minimal geodesic laminations on S that are either disjoint or equal, and assume that $S(\lambda_\pm)$ are incompressible in M. If $\lambda_\pm$ are essentially homotopic in (M, S), there is an essential interval bundle $(B, \partial_H B) \subset (M, S)$ such that $\lambda_\pm$ fill $\partial_H B$, and where $\lambda_\pm$ are essentially homotopic through B, as in Section 2.9.*

Here, $S(\lambda_\pm)$ are the subsurfaces with geodesic boundary filled by $\lambda_\pm$, as in Section 2.8. The assumption that they are incompressible generalises the assumption that α, β are homotopically essential in M in the annulus theorem.

---

[7] Really, we need to be using resolutions of our subsurfaces with geodesic boundary, as discussed in Section 2.1.

Proof Identify $M \setminus \partial_{\chi=0} M$ with the convex core of a geometrically finite hyperbolic 3-manifold. Set $S_\pm := S(\lambda_\pm)$. Lift the essential homotopy from λ_- to λ_+ to a homotopy from lifts $\tilde{\lambda}_- \subset \widetilde{S}_-$ to $\tilde{\lambda}_+ \subset \widetilde{S}_+$ in $\mathbb{H}^3$. Under the homotopy, which has bounded tracks, corresponding leaves of $\tilde{\lambda}_\pm$ have the same endpoints in $\partial \mathbb{H}^3$. The endpoints of $\tilde{\lambda}_\pm$ are dense in $\partial_\infty \widetilde{S}_\pm$, so this means that the subsurfaces $C_\pm \subset S_\pm$ constructed in Theorem 4.1 are just $C_\pm = S_\pm$. Passing to disjoint or equal resolutions $S'_\pm$ of $S_\pm$ and applying Theorem 2.8 gives an interval bundle B where $\lambda_\pm$ fill $\partial_H B = S'_- \cup S'_+$.

We claim that $\lambda_\pm$ are essentially homotopic through B. By Fact 2.16, it suffices to show that if σ is the canonical involution of B, as described in Proposition 4.5, then $\sigma(\lambda_\pm)$ is isotopic to $\lambda_\mp$ on $S'_\mp$. Using the notation of Theorem 4.1, σ lifts to a Δ-equivariant involution $\tilde{\sigma}$ of $\widetilde{B}$ that exchanges $\widetilde{S}'_-$ and $\widetilde{S}'_+$, where here $\Delta = \Gamma_- \cap \Gamma_+$. By equivariance, $\tilde{\sigma}$ extends continuously to the identity on Λ_Δ, so $\tilde{\sigma}(\tilde{\lambda}_-)$ is a lamination on $\widetilde{S}_+$ with all the same endpoints at infinity as $\tilde{\lambda}_+$, and hence equals $\tilde{\lambda}_+$. □

5 Laminations on the boundary

Suppose M is a compact, orientable 3-manifold with hyperbolisable interior and nonempty boundary ∂M. Equip M with an arbitrary Riemannian metric and lift it to a Riemannian metric on the universal cover $\widetilde{M}$. As in the introduction, a biinfinite path or ray h on $\partial \widetilde{M}$ is called *homoclinic* if there are points s^i, t^i with $|s^i - t^i| \to \infty$ such that

$$\sup_i d_{\widetilde{M}}(h(s^i), h(t^i)) < \infty.$$

Two rays h_+, h_- on $\partial \widetilde{M}$ are called *mutually homoclinic* if there are parameters $s^i_\pm \to \infty$ such that

$$\sup_i d_{\widetilde{M}}(h_+(s^i_+), h_-(s^i_-)) < \infty.$$

Here, a *ray* is a continuous map from an interval $[a, \infty)$, and a *biinfinite path* is a continuous map from $\mathbb{R}$. We will also call rays and paths on ∂M (mutually) homoclinic if they have lifts that are (mutually) homoclinic paths on $\partial \widetilde{M}$. We refer the reader to Section 5.1 for some comments on alternate definitions of homoclinic that exist in the literature.

Note that if we divide a biinfinite homoclinic path into two rays, then either one of the two rays is itself homoclinic, or the two rays are mutually homoclinic. Also, these definitions are metric independent: since M is compact, any two Riemannian metrics on M lift to quasi-isometric metrics on $\widetilde{M}$, and a path is homoclinic or mutually homoclinic with respect to one metric if and only if it is with respect to the other metric.

Here are some examples.

Example 5.1 (1) Suppose that D is a properly embedded disc in M, and $h : \mathbb{R} \to \partial M$ is a path that covers $\partial D \subset \partial M$. Then h is homoclinic: indeed, D lifts homeomorphically to $\widetilde{M}$, so h lifts to a path in $\widetilde{M}$ with compact image.

(2) Suppose that $\phi : (S^1 \times [0,1], S^1 \times \{0,1\}) \to (M, \partial M)$ is an essential embedded annulus. Then rays covering the two boundary components of the annulus are mutually homoclinic: indeed, ϕ lifts to

$$\tilde{\phi} : \mathbb{R} \times [0,1] \to \widetilde{M},$$

and we have $\sup_{t \in \mathbb{R}} d(\tilde{\phi}(t,0), \tilde{\phi}(t,1)) < \infty$, so restricting to $t \in [0,\infty)$ we get two mutually homoclinic rays in $\widetilde{M}$.

It will be convenient below to work with a particular choice of metric on M.

Definition 5.2 (an explicit metric on M) Let $\partial_{\chi<0} M$ be the union of all components of ∂M that have negative Euler characteristic, i.e., are not tori. Thurston's Haken hyperbolisation theorem, see [33], implies that there is a hyperbolic 3-manifold $N = \mathbb{H}^3 / \Gamma$ homeomorphic to the interior of M, where every component of $\partial_{\chi<0} M$ corresponds to a convex cocompact end of N. A torus $T \subset \partial M$, on the other hand, determines a cusp of N. So, in other words, N is "minimally parabolic": the only parabolics come from torus boundary components of M. For each T, pick an open neighbourhood $N_T \subset N$ of the associated cusp that is the quotient of a horoball in $\mathbb{H}^3$ by a $\mathbb{Z}^2$-action. Then

$$M \cong \mathrm{CC}(N) \setminus \bigcup_{\text{tori } T \subset \partial M} N_T, \tag{3}$$

and we will identify M with the right-hand side everywhere below. Then

- $\widetilde{M} \subset \mathbb{H}^3$ is obtained from the convex hull $\mathrm{CH}(\Gamma) \subset \mathbb{H}^3$ of the limit set of Γ by deleting an equivariant collection of horoballs, and
- the path metric induced on $\partial_{\chi<0} M$ is hyperbolic [58, Proposition 8.5.1], and the path metric induced on every torus $T \subset \partial M$ is Euclidean.

We now specialise to the case of paths that are *geodesics* on ∂M. Recall from Example 5.1(1) above that one can make homoclinic paths by running around the boundaries of disks in ∂M. The following shows that discs are essential in such constructions.

Fact 5.3 *Suppose that $S \subset \partial M$ is an essential subsurface. Then the inclusion of any lift $\widetilde{S} \subset \partial \widetilde{M}$ is a quasi-isometric embedding into $\widetilde{M}$. Moreover if S is incompressible then any pair of mutually homoclinic infinite rays on S are asymptotic and no biinfinite geodesic γ in S is homoclinic.*

Proof Think of M as embedded in a complete hyperbolic 3-manifold N as in (3), write $N = \Gamma \backslash \mathbb{H}^3$, and let $\widetilde{M} \subset \mathbb{H}^3$ be the preimage of M, so that $\widetilde{M}$ is obtained from the convex hull $\mathrm{CH}(\Gamma)$ be deleting an equivariant collection of horoballs. Fix a subgroup $\Delta < \Gamma$ that represents the conjugacy class associated to the image of the fundamental group of $S \subset M$. To show that

$$\widetilde{S} \hookrightarrow \partial \widetilde{M}$$

is a quasi-isometric embedding, it suffices to show that Δ is undistorted in Γ. But since M is geometrically finite and Δ is finitely generated, it follows from a result of Thurston (see Proposition 7.1 in [54]) that the group Δ is geometrically finite, and geometrically finite subgroups of (say, geometrically finite) hyperbolic 3-manifold groups are undistorted; see Corollary 1.6 in [28].

For the "moreover" statement, assume S is incompressible, so that $\widetilde{S}$ is simply connected, and consider a pair of infinite rays

$$h^{\pm} : \mathbb{R}^+ \to \widetilde{S}$$

that are geodesic for the induced hyperbolic metric and $t_n^{\pm} \to +\infty$ such that $d_{\widetilde{M}}(h^+(t_n^+), h^-(t_n^-))$ is bounded. Since $\widetilde{S} \subset \partial \widetilde{M}$ is a quasi-isometric embedding, $d_{\widetilde{S}}(h^+(t_n^+), h^-(t_n^-))$ is also bounded. Since $\widetilde{S}$ is simply connected and hyperbolic, this is possible only if h^+ and h^- are asymptotic on S. Taking $h^+ = h^-$ we get that a geodesic ray on $\widetilde{S}$ cannot be homoclinic. Taking $h^+ \neq h^-$, we get that any pair of mutually homoclinic infinite rays on S are asymptotic. In particular two disjoint geodesic rays in a homoclinic geodesic should be asymptotic. This is impossible for a geodesic in a simply connected hyperbolic surface. □

Example 5.1(2) above shows how embedded annuli in M can be used to create mutually homoclinic rays. In analogy to Fact 5.3, one can show that annuli are essential in such a construction. For instance, suppose M is acylindrical. Then work of Thurston, see [33] and more generally [42], says that we can choose the hyperbolic manifold N so that $\partial \mathrm{CC}(N) \cong \partial_{\chi<0} M$ is totally geodesic. Hence, the preimage of $\partial_{\chi<0} M$ in $\widetilde{M} \subset \mathbb{H}^3$ is a collection of hyperbolic planes. Any geodesic ray on $\partial_{\chi<0} M$ then lifts to a geodesic in $\mathbb{H}^3$, and two geodesic rays on $\partial_{\chi<0} M$ are mutually homoclinic if and only if their geodesic lifts are asymptotic in $\mathbb{H}^3$, which implies that they were asymptotic on $\partial_{\chi<0} M$.

5.1 Alternate definitions of homoclinic

Above, we defined a path

$$h : I \to \partial \widetilde{M}$$

to be homoclinic if there is are $s^i, t^i \in I$ with $|s^i - t^i| \to \infty$ such that

$$\sup_i d_{\widetilde{M}}(h(s^i), h(t^i)) < \infty.$$

Some other papers use slight variants of this definition. For example, the definition of (*faiblement*) *homoclinique* in Otal's thesis [56] is almost the same as what is written above, except that distances are computed *in the intrinsic metric on* $\partial \widetilde{M}$ instead of in $\widetilde{M}$. This is equivalent to our definition, though: the nonobvious direction follows from Fact 5.3, which says that boundary components of $\widetilde{M}$ quasi-isometrically embed in $\widetilde{M}$. And in the definition of *homoclinique* in Lecuire's earlier work [42], distances are computed not in $\widetilde{M}$, but within $\mathbb{H}^3$, with respect to a given identification of $\widetilde{M}$ with the convex core of some minimally parabolic hyperbolic 3-manifold, as discussed in Definition 5.2. When M has tori in its boundary, the inclusion $\widetilde{M} \hookrightarrow \mathbb{H}^3$ is not a quasi-isometric embedding, but the following lemma says that $d_{\mathbb{H}^3}$ is bounded if and only if $d_{\widetilde{M}}$ is bounded, so Lecuire's earlier definition is equivalent to ours.

Lemma 5.4 *Whenever* $x, y \in \widetilde{M}$, *we have*

$$d_{\mathbb{H}^3}(x, y) \le d_{\widetilde{M}}(x, y) \le e^{d_{\mathbb{H}^3}(x,y)/2} d_{\mathbb{H}^3}(x, y).$$

Proof Set $N := \Gamma \backslash \mathbb{H}^3$, so that $\widetilde{M}$ is obtained from the convex hull $\mathrm{CH} := \mathrm{CH}(\Lambda(\Gamma))$ of the limit set of Γ by deleting horoball neighbourhoods around all rank-two cusps. Take a $\mathbb{H}^3$-geodesic γ from x to y. Then γ lies inside CH, and it can only penetrate the deleted horoball neighbourhoods to a depth of $d(x, y)/2$. Now, whenever $B \supset B'$ are horoballs in $\mathbb{H}^3$ such that $d_{\mathbb{H}^3}(\partial B, B') \le d(x, y)/2$, the closest point projection

$$\pi : B \setminus B' \to \partial B$$

is well defined and $e^{d(x,y)/2}$-lipschitz. (Indeed, it suffices to take B as the height 1 horoball in the upper half space model and B' as the height $e^{d(x,y)/2}$ horoball, and then the claim is obvious.) So, the parts of γ above that penetrate the deleted horoballs can be projected back into $\partial \widetilde{M}$, and if we do this the resulting path has length at most $e^{d(x,y)/2} d(x, y)$. □

We should mention the version of homoclinic defined in Casson's original unpublished notes. There, M is a handlebody, and if we regard $\partial \widetilde{M} \hookrightarrow \mathbb{H}^3$ as above, then a simple geodesic $h : I \to \partial \widetilde{M}$ is called *homoclinic* if when we subdivide h into two rays $h_\pm$, these rays limit onto subsets $A_\pm \subset \mathbb{H}^3 \cap \partial \mathbb{H}^3$ such that $A_+ \cap A_- \neq \varnothing$. This definition is stronger than all the ones mentioned above: if $A_+ \cap A_-$ contains a point on $\partial \widetilde{M}$, rather than at infinity, then the definition of homoclinic above is obviously satisfied. Otherwise, $h_\pm$ have to have a common accumulation point in $\partial \mathbb{H}^3$, which corresponds to an end ξ of $\widetilde{M}$, and one can use the treelike structure of the universal cover $\widetilde{M}$ of the handlebody M to say that $h_\pm$ have to both intersect a sequence of meridians (m_i) on $\widetilde{M}$ that cut off smaller and smaller neighbourhoods of ξ. The times $t_p^i m$ when $h_\pm$ intersects m_i then work in the definition of homoclinic above. In fact, Casson's definition is strictly stronger. For instance, if $h_\pm$ both spiral around disjoint meridians $\gamma_\pm \subset \partial \widetilde{M}$, then h is homoclinic by our definition but not by Casson's. However, Theorem 1.1 still fails using Casson's original definition, due to the examples in Figure 13 on page 1849.

5.2 Waves, tight position, and intrinsic limits

As in the previous section, let M be a compact, orientable hyperbolisable 3-manifold with nonempty boundary ∂M, which we think of as the convex core of a hyperbolic 3-manifold with horoball neighbourhoods of its rank-2 cusps deleted.

Definition 5.5 (waves and tight position) Suppose that m is a meridian multicurve on ∂M, and let $\gamma \subset \partial M$ be a simple closed geodesic, a simple geodesic ray, or a simple biinfinite geodesic. An *m-wave* is a segment $\beta \subset \gamma$ that has endpoints on m, and is homotopic rel endpoints in M to an arc of m. If γ has no m-waves, and every infinite length segment of γ intersects m, then we say that γ is in *tight position* with respect to m.

Waves and tight position were discussed previously in [38; 42], for instance. Note that in our definition, an m-wave β can intersect m in its interior. More generally, an m-wave of a lamination is an m-wave of one of its leaves, and a lamination is in *tight position* with respect to m if all of its leaves are. Note that from this perspective, if a geodesic γ is in tight position with respect to some multicurve m (regarded as a lamination), then it is in tight position with respect to some component of m.

As an example, a meridian γ can never be in tight position with respect to another meridian m: taking discs with boundaries γ and m that are transverse and intersect minimally, any arc of intersection of these disks terminates in a pair of intersection points of γ and m that bound a m-wave of γ.

More generally, we have the following fact.

Fact 5.6 (tight position $\Longrightarrow \mathbb{H}^3$ quasigeodesic) *Let γ be a simple geodesic ray or biinfinite geodesic on ∂M. If γ is in tight position with respect to some meridian m then any lift $\tilde{\gamma} \subset \partial \widetilde{M}$ of γ is a quasigeodesic in $\mathbb{H}^3$. In particular, $\tilde{\gamma}$ is an $\widetilde{M}$-quasigeodesic, and is not homoclinic.*

Proof Intersecting with m breaks γ into a union of finite arcs. By simplicity of γ, these arcs fall into only finitely many homotopy classes rel m, and there is a universal upper bound $L = L(\gamma, m)$ on their lengths. Let D be a disc with boundary m and let $\widetilde{D}$ be the entire preimage of D in $\widetilde{M}$. Tightness means that the path $\tilde{\gamma}$ intersects infinitely many components of $\widetilde{D}$, and intersects no single component more than once.

In the notation of Definition 5.2, we have that $\widetilde{M} \subset \mathbb{H}^3$ is obtained from $\mathrm{CH}(\Gamma)$ by deleting an equivariant collection of horoballs. Each component of $\widetilde{D}$ separates $\mathrm{CH}(\Gamma)$, so if γ' is a segment of γ, any geodesic in $\mathbb{H}^3$ joining the endpoints of γ' must intersect each of the discs that γ' intersects. Hence, if $\epsilon > 0$ is the minimum distance between any two components of $\widetilde{D}$, then $\tilde{\gamma}$ is a $(L/\epsilon, L)$-quasigeodesic. □

Given a lamination, we now describe how to create a system of meridians with respect to which the lamination is in tight position.

Definition 5.7 (surgery) Suppose that λ is a geodesic lamination on ∂M and $m = \bigsqcup_{i=1}^{n} m_i$ is a geodesic meridian multicurve on ∂M, and β is an m-wave in λ whose interior is disjoint from m. Then the pair of points $\partial\beta$ separates some component m_i of m into two arcs m_i^1, m_i^2, both of which are homotopic to β rel endpoints in M. We perform a *λ-surgery* on m by replacing m_i^1 (say) with β, thus constructing a new multicurve $m' := (\beta \cup m_i^2) \sqcup \bigsqcup_{j \neq i} m_j$.

This notion of surgery appears in many other references, e.g., [3; 18; 38; 42]. We summarise its elementary properties here:

Fact 5.8 *Suppose that λ is a geodesic lamination.*

1. *If m is a meridian and λ has an m-wave, it also has an m-wave whose interior is disjoint from m, so a λ-surgery can be performed.*
2. *Any curve m' obtained by λ-surgery on a meridian m as above is a meridian.*
3. *If m is a **cut system** for M, i.e., a multicurve of meridians bounding discs that cut M into balls and 3-manifolds with incompressible boundary, then some λ-surgery on a component of m is another cut system.*

Proof For (1), suppose that λ has an m-wave β. Let $\tilde{m}$ be the entire preimage of m in the cover $\partial\widetilde{M}$, and lift β to an arc $\tilde{\beta}$ starting and ending on some fixed component $m_0 \subset \tilde{m}$. Since each component of $\tilde{m}$ separates $\partial\widetilde{M}$, there is some "outermost" subarc $\tilde{\beta}'$ that has endpoints on the same component of $\tilde{m}$, and that has interior disjoint from $\tilde{m}$. This $\tilde{\beta}'$ projects to an m-wave of λ whose interior is disjoint from m.

For (2), note that if $m' := \beta \cup m_2$ is obtained by λ-surgery on m, as above, then m, m' are homotopic in M, and hence m' is nullhomotopic. Also, if m' is inessential in ∂M, then β is homotopic on ∂M to m_2, implying that λ and m were not in minimal position, a contradiction since they are both geodesic. Hence m' is a meridian.

For (3), consider an m-wave in λ whose interior is disjoint from m and say that $\partial\beta \subset m_1$. Then $\partial\beta$ separates m_1 into two arcs m_1^1, m_1^2. It is not difficult to see that either $(\beta \cup m_1^1) \sqcup \bigsqcup_{j\neq 1} m_j$ or $(\beta \cup m_1^2) \sqcup \bigsqcup_{j\neq 1} m_j$ is a cut system. □

The following lemma is a modification of a result of Kleineidam and Souto [38, Lemmas 7 and 8] that is essential for everything below.

Lemma 5.9 (no waves, or a sequence of meridians) *Suppose λ is a geodesic lamination on $S = \partial M$ and m is a meridian multicurve. Then either*

(1) *there exists a finite sequence of λ-surgeries on m that terminates in some meridian multicurve m' where λ has no m'-waves,*

(2) *$S(\lambda)$ contains a sequence of meridians (γ_i) such that $i(\lambda, \gamma_i) \to 0$, with respect to every transverse measure on λ.*

Here, (2) makes sense even when λ admits no transverse measure of full support. Note that if λ is a minimal lamination and $\partial S(\lambda)$ is incompressible, then (2) implies that λ is an intrinsic limit of meridians.

Proof of Lemma 5.9 The two cases depend on whether λ contains infinitely many homotopy classes of m-waves, or not. Here, our homotopies are through arcs on S, keeping their endpoints on m.

If there are only finitely many classes of m-waves in λ, then a finite sequence of λ-surgeries converts m into a multicurve m' such that λ has no m'-waves, as each surgery decreases the number of waves by at least one. If there are infinitely many homotopy classes of m-waves in λ, then we can choose a sequence of parameterised m-waves $\alpha_i : [0,1] \to \mathbb{R}$ such that

(1) the two sequences of endpoints $(\alpha_i(0))$ and $(\alpha_i(1))$ both converge, and if either sequence converges into a simple closed curve $\gamma \subset \lambda$, then it approaches γ from only one side,

(2) no α_i and α_j are homotopic keeping their endpoints on m, for $i \neq j$.

To construct the desired sequence of meridians, let β_i^0 be the shortest geodesic on S from $\alpha_i(0)$ to $\alpha_{i+1}(0)$, and define β_i^1 similarly. Since $(\alpha_i(0)) \subset m$ and $(\alpha_i(1)) \subset m$ converge, we have $\beta_i^0, \beta_i^1 \subset m$ for i large enough. For large i, the union $\beta_i^0 \cup \alpha_i \cup \beta_i^1 \cup \alpha_{i+1}$ is an essential closed curve in $S(\lambda)$ that is nullhomotopic in M. It may not be simple, since β_i^0 and β_i^1 may overlap, but it has at most one self intersection. So by the loop theorem [29], one of the three simple closed curves obtained by surgery on it is a meridian γ_i.

Now, the fact that the endpoints can approach a simple closed curve in λ only from one side implies that for large i, the curves γ_i do not intersect any simple closed curve contained in λ. Since γ_i only intersects λ along the arcs β_i^0 and β_i^1, whose hyperbolic lengths converge to zero, it follows that $i(\gamma_i, \lambda) \to 0$ for any transverse measure on λ. □

Here is an important application of Lemma 5.9.

Lemma 5.10 (quasigeodesic or a sequence of meridians) *Suppose $\lambda \subset \partial_{\chi<0} M$ is a minimal geodesic lamination and that $\partial S(\lambda)$ is incompressible in M. Let $h \subset S(\lambda)$ be a simple geodesic ray or biinfinite geodesic that is disjoint from λ or contained in λ. Then either*

(1) *any lift $\tilde{h} \subset \partial \widetilde{M}$ of h is a quasigeodesic in $\widetilde{M}$*, *or*

(2) *$S(\lambda)$ contains a sequence of meridians (γ_i) such that $i(\lambda, \gamma_i) \to 0$, with respect to every transverse measure on λ.*

In particular, *if h is homoclinic*, *then λ satisfies* (2).

Proof Assume that (2) does not hold. If h is a geodesic ray, it is asymptotic to a geodesic ray $l^+ \subset \lambda$ and any lift of h is a quasigeodesic if and only if any lift of l^+ is quasigeodesic. Let $\mu = \lambda \cup h$ if h is biinfinite and $\mu = \lambda$ otherwise. Given a cut system m for M, Lemma 5.9 and Fact 5.8(3) say that we can perform μ-surgeries until we obtain a new cut system m such that μ has no m-waves. If m intersects μ, then μ is in tight position with respect to m, so (1) follows from Fact 5.6. Therefore, we can assume m does not intersect μ λ. Up to isotopy, we can also assume that $S(\lambda)$ does not intersect m. Since $\partial S(\lambda)$ is assumed to be a collection of incompressible curves, it follows that $S(\lambda)$ is itself incompressible, so (1) follows from Fact 5.3. □

We now come to the central definition of the section.

Definition 5.11 A minimal geodesic lamination $\lambda \subset \partial_{\chi<0} M$ is an *intrinsic limit of meridians* if there is a transverse measure[8] on λ and a sequence of meridians (γ_i) contained in $S(\lambda)$ such that $\gamma_i \to \lambda$ in $\mathcal{PML}(S(\lambda))$.

Using Lemma 5.10, we can prove the following proposition, which gives several equivalent characterisations of intrinsic limits.

Proposition 5.12 (intrinsic limits) *Suppose $\lambda \subset S = \partial M$ is a minimal geodesic lamination and $\partial S(\lambda)$ is incompressible. The following are equivalent*:

(1) *λ is an intrinsic limit of meridians.*

(2) *Given* (*some/any*) *transverse measure on λ*, *there is a sequence of meridians (γ_i) in $S(\lambda)$ such that $i(\gamma_i, \lambda) \to 0$.*

---

[8] It is currently unknown whether the particular transverse measure matters: we might suspect that a measured lamination is a projective limit of meridians if and only if the same is true for any other measured lamination with the same support, but there could also very well be a counterexample.

(3) *There is a homoclinic geodesic in* $S(\lambda)$ *that is either a leaf of* λ, *or is disjoint from* λ.

(4) *Given any transverse measure on* λ, *there is a sequence of essential (possibly nonsimple) closed curves* (γ_i) *in* $S(\lambda)$ *such that each* γ_i *is nullhomotopic in* M, *and* $i(\gamma_i,\lambda) \to 0$.

Note that when we say $\partial S(\lambda)$ is incompressible, we mean that no closed curve that is a boundary component of $S(\lambda)$ is nullhomotopic in M. This condition is mainly here to make statements and proofs easier. For instance, without this assumption our proof of (4) $\Longrightarrow$ (2) may produce peripheral meridians, but peripheral meridians can't be used in (2) $\Longrightarrow$ (1).

Proof **(2) $\Longrightarrow$ (1)** Fix some transverse measure on λ. By (2),

$$i(\gamma_i,\lambda)/\mathrm{length}(\gamma_i) \to 0,$$

so after passing to a subsequence we can assume that (γ_i) converges to a measured lamination μ in $S(\lambda)$ that does not intersect λ transversely. As λ fills $S(\lambda)$, μ is supported on λ.

(1) $\Longrightarrow$ (3) After passing to a subsequence, we can assume that (γ_i) converges in the Hausdorff topology to some lamination, which must then be an extension of λ by finitely many leaves. The statement (3) follows from an unpublished criterion of Casson, see Lecuire [42, Théorème B.1] for a proof, that states that any Hausdorff limit of meridians has a homoclinic leaf.

(3) $\Longrightarrow$ (2) This is an immediate corollary of Lemma 5.10.

(4) $\Longleftrightarrow$ (2) The direction $\Leftarrow$ is immediate, so suppose (γ_i) is a sequence of essential closed curves in $S(\lambda)$ that are nullhomotopic in H and $i(\lambda,\gamma_i) \to 0$. By Stallings' version of the loop theorem, for each i there is a meridian γ_i' that is obtained from γ_i by surgery at the self intersection points. Such surgeries can only decrease the intersection number with λ, so (2) follows. □

We will also need the following criterion in the next section.

Lemma 5.13 (intrinsic limits of annuli) *Suppose* $\lambda \subset \partial_{\chi<0} M$ *is a minimal lamination such that* $S(\lambda)$ *is compressible but* $\partial S(\lambda)$ *is incompressible, and that there is a sequence* (A_i) *of essential embedded annuli in* $(M, S(\lambda))$ *with* $i(\partial A_i,\lambda) \to 0$. *Then* λ *is an intrinsic limit of meridians.*

Proof Pick a meridian $m \subset S(\lambda)$. For each i, let $T_i : M \to M$ be the Dehn twist along the annulus A_i. Then for any sequence $n_i \in \mathbb{Z}$, the curves $T_i^{n_i}(m)$ are meridians, and if n_i grows sufficiently fast, then

$$i(T_i^{n_i}(m),\lambda)/\mathrm{length}(T_i^{n_i}(m)) \to 0.$$

Hence, after passing to a subsequence $T_i^{n_i}(m)$ converges to a lamination λ' supported in $S(\lambda)$ with zero intersection number with λ, implying λ' and λ have the same support, so λ is an intrinsic limit of meridians. □

6 Limits of homoclinic rays

In this section we characterise the laminations onto which pairs of disjoint mutually homoclinic rays can accumulate.

Theorem 6.1 (mutually homoclinic rays) *Let M be a compact orientable hyperbolisable 3-manifold and equip $\partial_{\chi<0} M$ with an arbitrary hyperbolic metric. Let $h_\pm$ be two disjoint, mutually homoclinic simple geodesic rays on $\partial_{\chi<0} M$ that accumulate onto (possibly equal) minimal laminations $\lambda_\pm$, and where the multicurve $\partial S(\lambda_\pm)$ is incompressible in M. Then one of the following holds:*

(1) *one of λ_+ or λ_- is an intrinsic limit of meridians,*

(2) *h_+ and h_- are asymptotic on $\partial_{\chi<0} M$, and either*

  (a) *any two mutually homoclinic lifts $\tilde{h}_\pm$ to $\partial \widetilde{M}$ are asymptotic on $\partial \widetilde{M}$, or*

  (b) *$\lambda := \lambda_- = \lambda_+$ is a simple closed curve that is homotopic in M to a nontrivial power γ^n, $n > 1$ of some closed curve γ in M,*

(3) *$h_\pm$ are not asymptotic on $\partial_{\chi<0} M$, and there is an essential (possibly nontrivial) interval bundle $B \subset M$ such that $\lambda_\pm$ each fill a component of $\partial_H B$, and $\lambda_\pm$ are essentially homotopic through B, as in Section 2.9.*

The proof of Theorem 6.1 is given in Section 6.1. One can construct examples of mutually homoclinic rays in (1)–(3), as follows.

For (1), pick two meridians λ_-, λ_+ on M and let $h_\pm$ spiral onto $\lambda_\pm$. One can also produce similar examples by letting $\lambda_\pm$ be arbitrary laminations in disjoint subsurfaces $S(\lambda_\pm)$ that are spheres with at least 4 boundary components, all of which are compressible in M, and letting $h_\pm$ accumulate onto $\lambda_\pm$. In case (1), we expect it is possible that $S(\lambda_-)$ is incompressible, say, while λ_+ is an intrinsic limit of meridians. For instance, suppose C is a compression body with connected, genus-at-least-two interior boundary $\partial_- C$, and exterior boundary $\partial_+ C$. Let $f : C \to C$ be a homeomorphism such that $f|_{\partial_+ C}$ and $f|_{\partial_- C}$ are both pseudo-Anosov, with attracting laminations λ_+ and λ_-, respectively. We expect that there are rays $\ell_\pm \subset \lambda_\pm$ that are mutually homoclinic. But λ_+ is an intrinsic limit of meridians, while $S(\lambda_-)$ is incompressible.

For (2)(a), take M to be a handlebody, let λ be any simple closed curve on ∂M that is essential in M but has no nontrivial roots in $\pi_1 M$, and let $h_\pm$ spiral around λ in the same direction. For (2)(b), take $h_\pm$ to be the two ends of the homoclinic geodesic h on the left in Figure 5. In the picture, we have drawn a solid torus that is a boundary-connect-summand of M, which (say) is a handlebody. The rays $h_\pm$ both spiral onto a simple closed curve λ, the $(2,1)$-curve on the solid torus. This λ is homotopic to the square of the core curve of the solid torus. Although $h_\pm$ are asymptotic on ∂M, any lift $\tilde{h}$ in $\partial \widetilde{M}$ will have ends that are mutually homoclinic, but nonasymptotic. On the right, we have drawn the preimage $\tilde{\lambda}$ of λ, and two lifts $\tilde{h}_1, \tilde{h}_2$ of h. Note that, since $h_\pm$ are asymptotic on ∂M, one end of $\tilde{h}_1$ is asymptotic to an end of $\tilde{h}_2$.

When M is a compression body, Casson and Gordon [17, Theorem 4.1] proved that any simple closed curve $\lambda \subset \partial M$ that has a nontrivial root in $\pi_1 M$ lies on the boundary of a solid torus that is a boundary

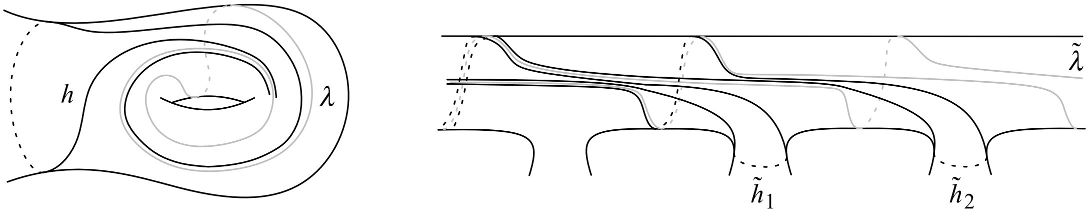


Figure 5: An example of (2)(b) in Theorem 6.1.

connect summand of M, exactly as in Figure 5. When M has incompressible boundary, such λ come from components of the characteristic submanifold of M, see Section 2.6, that are either solid tori or twisted interval bundles over nonorientable surfaces.

Examples of (3) are shown in Figure 6, with $h_\pm$ being the two ends of a homoclinic geodesic h. On the left, λ_-, λ_+ are simple closed curves that bound an embedded annulus A in M and B is a regular neighbourhood of A. The rays $h_\pm$ are mutually homoclinic since the annulus A lifts to an embedded infinite strip $\mathbb{R}\times[-1,1]\subset\widetilde{H}$ and the $h_\pm$ are asymptotic to $\mathbb{R}_+\times\{-1\}$ and $\mathbb{R}_+\times\{1\}$, respectively. On the right, we write $M=Y\times[-1,1]$ where Y is a genus-two surface with one boundary component. The laminations $\lambda_\pm$ are minimal (in the picture they are drawn as "train tracks") and fill $Z\times\{\pm1\}$, where $Z\subset Y$ is a torus with two boundary components. Here, $B=Z\times[-1,1]$. One can also construct similar examples of (3) where the interval bundle B is twisted.

The assumption that $\partial S(\lambda_\pm)$ is incompressible is necessary in Theorem 6.1. For instance, suppose M is a compression body with exterior boundary a genus-3 surface S, where the only meridian on S is a single separating curve γ. Let T be the component of $S\setminus\gamma$ that is a punctured genus-2 surface. Then there are distinct minimal geodesic laminations $\lambda,\lambda'\subset T$, each of which fills T, that are properly homotopic in M: just take distinct laminations that are identified when we cap off the puncture of T to get a closed genus-2 surface. Corresponding ends of corresponding leaves of λ, λ' are mutually homoclinic rays that accumulate onto λ, λ', respectively, but none of (2)–(3) hold. One could write down a version of Theorem 6.1 that omits the assumption that $\partial S(\lambda_\pm)$ is incompressible, but the conclusion would be relative to capping off $S(\lambda_\pm)$, and the statement would be more complicated.

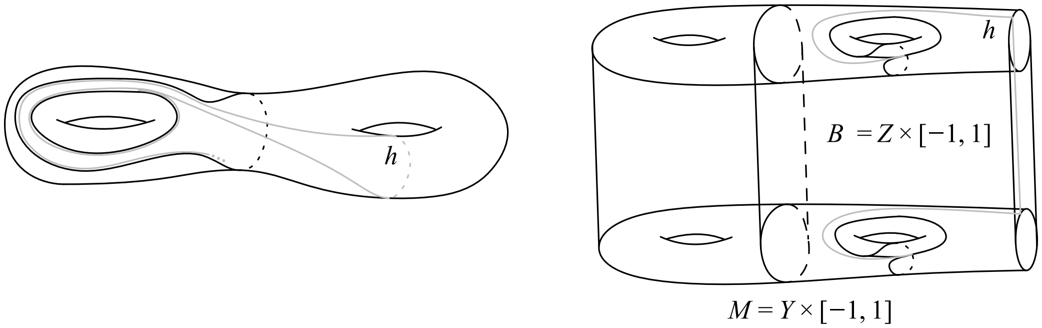


Figure 6: Examples of homoclinic geodesics (in grey) satisfying (3) in Theorem 6.1.

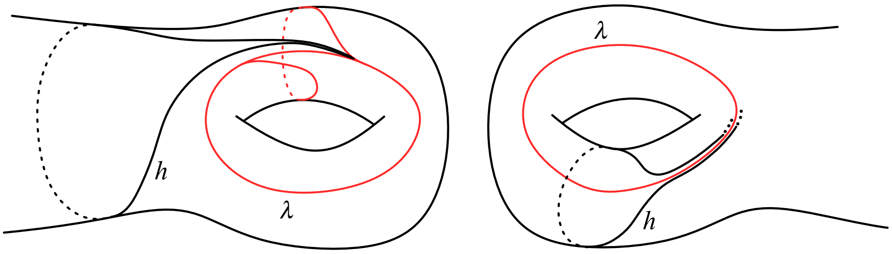


Figure 7: Homoclinic geodesics h as in cases (i) and (ii) in Corollary 6.2, respectively.

Here is a slightly more refined version of Theorem 6.1 that applies to homoclinic biinfinite geodesics on $\partial_{\chi<0}M$.

Corollary 6.2 (homoclinic biinfinite geodesics) *Suppose that M is as in Theorem 6.1, that h is a homoclinic biinfinite simple geodesic on some component $S \subset \partial_{\chi<0}M$, that $h_\pm$ are the two ends of h, that $h_\pm$ limit onto $\lambda_\pm$, and that $\partial S(\lambda_\pm)$ is incompressible in M.*

Suppose case (2) *of Theorem 6.1 holds, in which case the two limiting laminations are equal, and we can set $\lambda := \lambda_\pm$. If this λ is not an intrinsic limit of meridians then either*

- (i) *after reparametrising h, we have that for all large s, the points $h(-s)$ and $h(s)$ are joined by a geodesic segment c with $h \cap \operatorname{int}(c) = \varnothing$, such that c, $h|_{(-\infty,-s)}$ and $h|_{[s,\infty)}$ bound an embedded geodesic triangle $\Delta \subset S$ with one ideal vertex, and $c \cup h([-s,s])$ is a meridian in M, or*
- (ii) *λ is a simple closed curve on S, the two ends of h spiral around λ in the same direction, and any neighbourhood of the union $h \cup \lambda \subset S$ contains a meridian.*

See Figure 7. From this dichotomy, one can alternatively deduce that either

- (i′) *any neighbourhood of $h \cup \lambda$ contains a meridian disjoint from λ, or*
- (ii′) *λ is a simple closed curve and $h_\pm$ spiral around λ in the same direction but from opposite sides.*

Finally, if case (3) *of Theorem 6.1 holds, we can choose the interval bundle B such that h contains a subarc $\alpha \subset h$ that is a compression arc for B.*

Proof Let $\tilde{h}$ be a homoclinic lift of h on $\partial\widetilde{M}$. By Lemma 5.10, either one of $\lambda_\pm$ is an intrinsic limit of meridians in M, in which case we're in case (1) and are done, or both ends of $\tilde{h}$ are quasigeodesic in $\widetilde{M}$. Since $\tilde{h}$ is homoclinic, it follows that its two ends are mutually homoclinic, so we're in the setting of Theorem 6.1 and one of (2)–(3) holds.

Assume we're in case (2) of Theorem 6.1, and set $\lambda := \lambda_\pm$. Assume first that $\lambda := \lambda_\pm$ is a simple closed curve. Since the two ends of h are asymptotic, they spiral around λ in the same direction. Let U be a neighbourhood of $h \cup \lambda$ on $\partial_{\chi<0}M$. Then h is a homoclinic geodesic contained in U, so Fact 5.3 implies that U is compressible as desired in (ii).

Now suppose that λ is *not* a simple closed curve, in which case we're in case (2)(a) of Theorem 6.1. We show (i) holds. Let's start by constructing the desired geodesic triangle. Parametrise h, pick a universal covering map $\mathbb{H}^2 \to S$, lift h to a parametrised geodesic $\hat{h}$ in $\mathbb{H}^2$, and let

$$\xi = \lim_{t \to +\infty} \hat{h}(t) \in \partial_\infty \mathbb{H}^2.$$

Note that since λ is not a simple closed curve, ξ is not a fixed point of any deck transformation of $\mathbb{H}^2 \to S$. The two ends of h are asymptotic on S, so there is then a unique deck transformation $\gamma : \mathbb{H}^2 \to \mathbb{H}^2$ such that

$$\xi = \lim_{t \to -\infty} \gamma \circ \hat{h}(t).$$

It follows that if we use a particular arc-length parametrisation of h, we may assume that for each $t \in \mathbb{R}$, the points $\hat{h}(t), \gamma \circ \hat{h}(-t)$ lie on a common horocycle tangent to ξ. Fix a parameter $t = s$ large enough such that the geodesic segment $\hat{c}$ joining $\hat{h}(s)$ and $\gamma \circ \hat{h}(-s)$ is shorter than the injectivity radius of S, and therefore projects to a simple geodesic segment c in S.

Let $\hat{\Delta} \subset \mathbb{H}^2$ be the triangle bounded by $\hat{c}$ and the two rays $\hat{h}([s, \infty))$ and $\gamma \circ \hat{h}((-\infty, s])$. Let $g : \mathbb{H}^2 \to \mathbb{H}^2$ be a deck transformation. We claim that $g \circ \hat{h}(\mathbb{R}) \cap \operatorname{int}(\hat{\Delta}) = \varnothing$. If not, then since $\hat{\Delta}$ has geodesic sides, two of which are disjoint from $g \circ \hat{h}$, it follows that one of the two endpoints of $g \circ \hat{h}$ is ξ. If it's the positive endpoint, then g fixes ξ, and the axis of g projects to a (simple) closed curve on S, around which the two ends of h spiral, contradicting that λ isn't a simple closed curve. If the negative endpoint of $g \circ \hat{h}$ is ξ, then $g \circ \gamma^{-1}$ fixes ξ and we get a similar contradiction.

Next, we claim that we have $g(\operatorname{int}(\hat{\Delta})) \cap \operatorname{int}(\hat{\Delta}) = \varnothing$ as long as $g \neq \mathrm{id}$. Suppose that for some $g \neq \mathrm{id}$ the intersection is nonempty. Then $g(\xi) \neq \xi$, since otherwise we have a contradiction as in the previous paragraph. The previous paragraph implies that the sides of the triangles $g(\hat{\Delta}), \hat{\Delta}$ that are lifts of rays of h do not intersect the interior of the other triangle. So, the only way the interiors of $g(\hat{\Delta}), \hat{\Delta}$ can intersect is if $\hat{c}$ and $g(\hat{c})$ intersect. However, this does not happen since we chose s large enough so that $\hat{c}$ projects to a simple geodesic segment in S.

The previous two paragraphs imply that $\hat{\Delta}$ projects to an embedded geodesic triangle Δ in S whose interior is disjoint from h, as desired in (i). By construction, c and $h([-s, s])$ are simple geodesic segments, and since $g \circ \hat{h}(\mathbb{R}) \cap \operatorname{int}(\hat{\Delta}) = \varnothing$ for any $g \neq \mathrm{id}$, they are disjoint. It follows that $c \cup h([-s, s])$ is an essential simple closed curve on S.

We need to show that $c \cup h([-s, s])$ is nullhomotopic in M. For this, remember that since λ is not a simple closed curve, it cannot be that two distinct lifts of $h([s, \infty))$ to $\mathbb{H}^2$ are asymptotic, for if so they would differ by a deck transformation fixing the endpoint $\xi \in \partial \mathbb{H}^2$. Therefore, no two distinct lifts of $h([s, \infty))$ to $\partial \widetilde{M}$ can be asymptotic, and similarly for $h((-\infty, s])$. But the two ends of any lift of $\tilde{h} \subset \partial \widetilde{M}$ of h are mutually homoclinic, and hence asymptotic by the assertion in case (2) of Theorem 6.1. So, the projection $\hat{\Delta} \to \Delta$ factors through a geodesic triangle $\tilde{\Delta} \subset \partial \widetilde{M}$ bounded by $\tilde{h}([s, \infty))$, $\tilde{h}((-\infty, -s])$ and a geodesic segment $\tilde{c}$. The curve $c \cup h([-s, s])$ is the projection of the closed curve $\tilde{c} \cup \tilde{h}([-s, s]) \subset \widetilde{M}$, and so is nullhomotopic in M.

For the (i′) versus (ii′) dichotomy, note that if we're in case (i) then by taking s large, we can ensure that $c \cup h([-s, s])$ is inside a given neighbourhood of $h \cup \lambda$. Moreover, as h is disjoint from the interior of the geodesic triangle Δ mentioned in (i), so is λ, and hence after increasing s slightly we can assume $c \cup h([-s,s])$ is disjoint from λ. If we're in case (ii), and $h_\pm$ spiral onto λ *from the same side*, then λ is a peripheral curve in the specified regular neighbourhood of $h \cup \lambda$, so the meridian given in that neighbourhood can be taken disjoint from λ. Otherwise, we're in case (ii′).

Now assume we are in case (3). Let $S_\pm$ be the components of $\partial_H B$ containing $\lambda_\pm$. We may assume that h is in minimal position with respect to $\partial S_\pm$. Since h is simple and the ends of h limit onto minimal laminations that fill $S_\pm$, we have that h intersects $\partial S_- \cup \partial S_+$ at most twice. Furthermore, in the case that $S_- = S_+$, the homoclinic geodesic h cannot be contained entirely in the incompressible surface $S_\pm$, by Fact 5.3. So, h is the concatenation of two rays in $S_\pm$ and an arc α such that $\operatorname{int}(\alpha)$ lies outside $S_\pm$.

Let $X \subset S$ be the union of $S_\pm$ and a regular neighbourhood of α. Since h is homoclinic, there is a meridian on X by Fact 5.3. If the two endpoints of α lie on different boundary components of $\partial_H B$, then α is a compression arc for B by Fact 2.12. So, we may assume that the two endpoints of α lie on the same boundary component c of $\partial_H B$. Fact 2.12 then says that α is homotopic rel endpoints in M to an arc of c. So, if we make a new path $h' \subset \partial_H B$ from h by replacing α with that arc of c, then h' is still homoclinic, so it cannot be boundedly homotopic to a geodesic in $\partial_H B$ by Fact 5.3, which implies that its ends $h_\pm$ are asymptotic, a contradiction to the assumption in (3). □

If h is a parametrised biinfinite geodesic, let's denote by $h_\pm$ the associated *positive and negative rays*, namely $h_+ := h|_{[0,\infty)}$ and $h_- := h|_{(-\infty,0]}$.

Definition 6.3 (mutually bihomoclinic) We say that two biinfinite geodesics h, h' on some component $S \subset \partial_{\chi<0} M$ are *mutually bihomoclinic* if they have distinct lifts $\tilde{h}$, $\tilde{h}'$ on $\partial \widetilde{M}$ such that the associated rays $\tilde{h}_+$, $\tilde{h}'_+$ are mutually homoclinic, as are the rays $\tilde{h}_-$, $\tilde{h}'_-$.

Above, we allow $h_0 = h_1$, but we require the two lifts to be distinct. Here is a variant of Corollary 6.2 for pairs of mutually bihomoclinic rays that we will use in a sequel to this paper. For simplicity, we'll state only the analogue of (i′) versus (ii′) in Corollary 6.2.

Corollary 6.4 *Suppose that M is as in Theorem 6.1, that h, h' are simple biinfinite geodesics on a component $S \subset \partial_{\chi<0} M$ that are either disjoint or equal, and that h, h' are mutually bihomoclinic.*

Suppose case (2) *of Theorem 6.1 holds both for the positive rays h_+, h'_+, and for the negative rays h_-, h'_-. Then either*

(i′) *in any neighbourhood of $h \cup h' \cup \lambda_- \cup \lambda_+$, there is a meridian m disjoint from λ_- and λ_+, or*

(ii′) *for either $+$ or $-$, say $+$, we have that λ_+ is a simple closed curve and h_+, h'_+ spiral around λ_+ in the same direction but from different sides.*

The proof is similar to the proof of Corollary 6.4.

6.1 Proof of Theorem 6.1

The proof proceeds in a few cases. As in Definition 5.2, we identify M with the convex core $\mathrm{CC}(N)$ of a geometrically finite hyperbolic 3-manifold $N = \mathbb{H}^3/\Gamma$, and we identify the universal cover $\widetilde{M}$ with the preimage of $\mathrm{CC}(N)$ in $\mathbb{H}^3$, which is the convex hull of the limit set of Γ. Note that the closure of $\widetilde{M}$ in $\mathbb{H}^3 \cup \partial\mathbb{H}^3$ is a ball.

There are four cases to consider:

(A) Both $\lambda_\pm$ are simple closed curves. We show that either (2) or (3) holds.

(B) Both $\lambda_\pm$ are distinct, in which case the surfaces $S(\lambda_\pm)$ are disjoint, but one of these surfaces is compressible, say $S(\lambda_+)$. We show (1).

(C) At least one of $\lambda_\pm$ is not a simple closed curve, and both $S(\lambda_\pm)$ are incompressible. We show (2)(a) or (3) holds.

(D) $\lambda_- = \lambda_+$, which is not a simple closed curve, and $S(\lambda_\pm)$ is compressible. We show either (1) or (2)(a) holds.

Cases (A) and (B) above are the easiest. Our proof in case (C) involves a hyperbolic geometric interpretation of the characteristic submanifold of a pair, as discussed in Section 3 of [40] and by Walsh [60]; our argument is a bit more complicated than theirs, since we have to deal with accidental parabolics. In case (D), our argument adapts and fills some gaps in a surgery argument of Kleineidam and Souto [38] and Lecuire [42, Appendix C].

Proof of (A) Assume that both $\lambda_\pm$ are simple closed curves. Since $S(\lambda_\pm)$ is assumed to be incompressible, both $\lambda_\pm$ are incompressible in M. If $\lambda_- \neq \lambda_+$, they have mutually homoclinic lifts and hence are homotopic. Then we are in case (3) by the annulus theorem. So we may assume the two curves are the same, and write $\lambda := \lambda_\pm$.

We claim that $h_\pm$ spiral around λ in the same direction, so that they are asymptotic on ∂M. Suppose not, and pick mutually homoclinic lifts $\tilde{h}_\pm$ in $\widetilde{M}$. Then $\tilde{h}_-$ and $\tilde{h}_+$ are asymptotic to lifts $\tilde{\lambda}$ and $\alpha(\tilde{\lambda})$ of λ, where $\alpha \in \Gamma$ is a deck transformation. Any lift of λ is a quasigeodesic in $\widetilde{M}$, and hence in $\mathbb{H}^3$, so $\tilde{h}_\pm$ are quasigeodesic rays, and therefore have well-defined endpoints in $\partial\mathbb{H}^3$, which must be the same since the two rays are mutually homoclinic. Since $h_\pm$ spiral around λ in opposite directions, this means that $\alpha \in \Gamma$ takes one endpoint of $\tilde{\lambda}$ in $\partial\mathbb{H}^3$ to the other endpoint of $\tilde{\lambda}$. Since $\tilde{\lambda}$ is stabilised by a loxodromic isometry in Γ, and Γ is torsion-free and discrete, this is impossible.

Suppose we are not in case (2)(a), so there are mutually homoclinic lifts $\tilde{h}_\pm$ that are not asymptotic on $\partial\widetilde{M}$. As in the previous paragraph, we may assume that $\tilde{h}_-$ and $\tilde{h}_+$ are asymptotic to lifts $\tilde{\lambda}$ and $\alpha(\tilde{\lambda})$ for some deck transformation $\alpha \in \Gamma$. Since $\tilde{h}_\pm$ are not asymptotic, $\tilde{\lambda} \neq \alpha(\tilde{\lambda})$. As before, α fixes the common endpoint of $\tilde{h}_\pm$ in $\partial\mathbb{H}^3$, which is a fixed point of the cyclic group $\langle\beta\rangle \subset \Gamma$ of loxodromic isometries fixing $\tilde{\lambda}$. As Γ is discrete and torsion-free, and $\alpha \notin \langle\beta\rangle$, we have that α is a root of β or β^{-1} in Γ, and (2)(b) follows. □

Proof of (B) Suppose that $\lambda_\pm$ are distinct, in which case the surfaces $S(\lambda_\pm)$ are disjoint, but that one of these surfaces is compressible, say $S(\lambda_+)$. We claim that λ_+ is an intrinsic limit of meridians, in which case (1) holds and we are done. If not, take a meridian $m \subset S(\lambda_+)$ and apply Lemma 5.9. We obtain a new meridian $m' \subset S(\lambda_+)$ such that λ_+ has no m-waves. Since λ_+ fills $S(\lambda_+)$ and the boundary components $\partial S(\lambda_\pm)$ are incompressible, it follows that λ_+ is in tight position with respect to m. So after possibly restricting the domains, h_+ is in tight position with respect to m', while h_- never intersects m'. This contradicts the fact that $h_\pm$ are mutually homoclinic, since if $\tilde h_\pm$ are homoclinic lifts in $\widetilde M$, for large t the point $\tilde h_+(t)$ is separated from the image of $\tilde h_-$ by arbitrarily many lifts of m'. □

Proof of (C) Assume that at least one of $\lambda_\pm$ is not a simple closed curve, and that $S_\pm := S(\lambda_\pm)$ are incompressible. Note that $S_\pm$ are equal or have disjoint interiors. We want to prove that we're in case (2) or (3). Lift $h_\pm$ to mutually homoclinic rays $\tilde h_\pm \subset \partial \widetilde M$. Fact 5.3 implies that each inclusion $\widetilde S_\pm \hookrightarrow \widetilde M$ is a quasi-isometric embedding, so if $\widetilde S_- = \widetilde S_+$, then the two mutually homoclinic rays $\tilde h_\pm$ are actually asymptotic on $\partial \widetilde M$. If this is true for all lifts $\tilde h_\pm$, we are in case (2)(a) and are done. So, we can assume below that $\widetilde S_- \neq \widetilde S_+$. Note that it may still be that $\lambda_- = \lambda_+$ and $S_- = S_+$.

Let $\Gamma_\pm \subset \Gamma$ be the stabiliser of $\widetilde S_\pm$ and let $\Lambda_\pm \subset \partial \mathbb{H}^3$ be the limit set of $\Gamma_\pm$. Since $\Gamma_\pm$ acts cocompactly on $\widetilde S_\pm$, the inclusion $\widetilde S_\pm \hookrightarrow \mathbb{H}^3$ extends continuously to a map $\partial_\infty \widetilde S_\pm \to \Lambda_\pm \subset \partial \mathbb{H}^3$, by the main result of [53]. In particular, $\tilde h_\pm$ have well-defined endpoints in $\partial \mathbb{H}^3$, and since they're mutually homoclinic, they have the same endpoint $\xi \in \Lambda_- \cap \Lambda_+ \subset \partial \mathbb{H}^3$.

We now apply Theorem 4.1. Since $\xi \in \Lambda_- \cap \Lambda_+$, using the notation of Theorem 4.1, the rays $\tilde h_\pm$ are either eventually contained in the convex hulls $\widetilde C_\pm \subset \widetilde S_\pm$, or are asymptotic onto their boundaries. But $\widetilde C_\pm$ project to (possibly degenerate) generalised subsurfaces $C_\pm$ with geodesic boundary in $S_\pm$, and the rays $h_\pm$ limit onto filling laminations in $S_\pm$, so it follows that actually $C_\pm = S_\pm$, and that there is a homotopy from S_- to S_+ in M that is the projection of a homotopy from $\widetilde S_-$ to $\widetilde S_+$. Since one of $\lambda_\pm$ is not a simple closed curve, this means they are *both* not simple closed curves and the (a priori degenerate) subsurfaces with geodesic boundary $S_\pm$ are not simple closed geodesics.

Let $S'_\pm \subset \operatorname{int}(S_\pm)$ be obtained by deleting small collar neighbourhoods of $\partial S_\pm$, so that $S'_\pm$ are both actually embedded, still contain $\lambda_\pm$, and are either disjoint or equal. Since $S'_\pm$ are incompressible and homotopic in M, Theorem 2.8 implies that they bound an essential interval bundle $B \subset M$. Moreover, the fact that the homotopy from S_- to S_+ is the projection of a homotopy from $\widetilde S_-$ to $\widetilde S_+$ means that we can assume that there is a component $\widetilde B \subset \widetilde M$ of the preimage of B that intersects $\partial \widetilde M$ in $\widetilde S'_\pm$. Note that $\widetilde B$ is invariant under $\Delta = \Gamma_- \cap \Gamma_+$, since any element of Δ preserves $\widetilde S'_\pm$, and hence $\widetilde B$.

We claim that $\lambda_\pm$ are essentially homotopic through B. By Fact 2.16, it suffices to show that if σ is the canonical involution of B, as described in Proposition 4.5, then $\sigma(\lambda_\pm)$ is isotopic to $\lambda_\mp$ on $S'_\mp$. Well, σ lifts to a Δ-equivariant involution $\tilde\sigma$ of $\widetilde B$ that exchanges $\widetilde S'_-$ and $\widetilde S'_+$, where here $\Delta = \Gamma_- \cap \Gamma_+$. By equivariance, $\tilde\sigma$ extends continuously to the identity on Λ_Δ, so in particular its extension fixes ξ, and hence $\tilde\sigma(h_\pm)$ is properly homotopic to $h_\mp$ on $S_\mp$, which implies $\tilde\sigma(\lambda_\pm)$ is isotopic to $\lambda_\mp$ as desired.

If $h_\pm$ are not asymptotic on ∂M, then we are in case (3) and are done. So, assume $h_\pm$ are asymptotic. Then there is some $\gamma \in \Gamma$ such that $\gamma(\tilde h_-) \subset \widetilde S'_+$ and is asymptotic to $\tilde h_+$. This γ fixes the endpoint $\xi \in \partial \mathbb{H}^3$ of $\tilde h_\pm$. Moreover, $\gamma(\widetilde B)$ is a component of the preimage of B that contains $\widetilde S'_+$, and therefore equals $\widetilde B$. So, γ exchanges $\widetilde S'_\pm$, and therefore $\gamma^2 \in \Delta$. But then $\tilde h_\pm$ are asymptotic to the axes of $\gamma^2 \curvearrowright \widetilde S_\pm$, implying that $h_\pm$ accumulate onto simple closed curves in ∂M, contradicting our assumption in (C). □

Proof of (D) Assume that $\lambda_- = \lambda_+$, write $\lambda = \lambda_\pm$ for brevity, assume that λ is not a simple closed curve, and that $S(\lambda)$ is compressible. We want to prove that either λ is an intrinsic limit of meridians, or $h_\pm$ are asymptotic, as are any pair of mutually homoclinic lifts $\tilde{h}_\pm$.

If λ is an intrinsic limit of meridians, we are done, so since $S(\lambda)$ is compressible with incompressible boundary, by Lemma 5.9 we can choose a meridian $m \subset S(\lambda)$ with respect to which λ is in tight position. Let $\tilde{m}$ be its full preimage in $\partial \widetilde{M}$, and let $\tilde{h}_\pm$ be *any* pair of mutually homoclinic lifts in $\partial \widetilde{M}$. Truncating if necessary, we can assume that $h_\pm$ are in tight position with respect to m, and hence the lifts $\tilde{h}_\pm$ are quasigeodesic rays in $\mathbb{H}^3$, by Fact 5.6. Since they are mutually homoclinic, $\tilde{h}_-$ and $\tilde{h}_+$ converge to the same point $\xi \in \partial_\infty \mathbb{H}^3$, and tightness further implies that after restricting to appropriate subrays, $\tilde{h}_-$ and $\tilde{h}_+$ intersect exactly the same components of $\tilde{m}$, in the same order. Reparametrising, we have

$$\tilde{h}_\pm : [0,\infty) \to \partial \widetilde{M}, \qquad \tilde{h}_+(i), \tilde{h}_-(i) \in \tilde{m}_i \quad \text{for all } i \in \mathbb{N},$$

where each $\tilde{m}_i$ is a component of $\tilde{m}$, and where $\tilde{h}_\pm(t) \notin \tilde{m}$ when $t \notin \mathbb{N}$. Let

$$d_i := d_{\tilde{m}}(\tilde{h}_+(i), \tilde{h}_-(i))$$

be the distance along $\tilde{m}$ between $\tilde{h}_+(i)$ and $\tilde{h}_-(i)$. □

Claim 6.5 *There is some uniform $\epsilon > 0$, independent of the particular chosen lifts $\tilde{h}_\pm$, such that either*

(1) *$\tilde{h}_\pm$ are asymptotic on $\partial \widetilde{M}$, and hence $h_\pm$ are asymptotic on ∂M, or*

(2) *$\liminf_i d_i \geq \epsilon$.*

Proof Let's assume that $\tilde{h}_+$ and $\tilde{h}_-$ are not asymptotic on $\partial \widetilde{M}$, and write $d = \liminf_i d_i$. Fix some transverse measure on λ. If d is small, we will construct meridians $\gamma \subset S(\lambda)$ with very small intersection number with λ. Since λ is not an intrinsic limit of meridians, there is some fixed lower bound for such intersection numbers, which will give a contradiction for small d.

Suppose d is small and pick $0 \ll i < j$ such that

$$d_i, d_j < 2d,$$

let b_i be the (unique) shortest path on $\tilde{m}$ from $\tilde{h}_-(i)$ to $\tilde{h}_+(i)$, and define b_j similarly. Let $\tilde{\gamma}_{ij}$ be the loop on $\partial \widetilde{M}$ obtained by concatenating the four segments $\tilde{h}_+([i,j])$, $\tilde{h}_-([i,j])$, b_i and b_j in the obvious way.

We first claim that after fixing i, it is possible to choose j such that $\tilde{\gamma}_{ij}$ is homotopically essential on $\partial \widetilde{M}$. Assume not, let $\widetilde{S} \subset \partial \widetilde{M}$ be the component containing $\tilde{h}_\pm$, fix a universal covering map

$$\mathbb{H}^2 \to \widetilde{S},$$

and lift the rays $\tilde{h}_\pm|_{[i,\infty)}$ to rays

$$\mathfrak{h}_\pm : [i,\infty) \to \mathbb{H}^2$$

in such a way that b_i lifts to a segment connecting $\mathfrak{h}_-(i)$ to $\mathfrak{h}_+(i)$. Now, there are infinitely many $j > i$ with $d_j < 2d$. For each such j, we know that $\tilde{\gamma}_{ij}$ is homotopically inessential on $\partial \widetilde{M}$, so the points $\mathfrak{h}_-(j)$ to $\mathfrak{h}_+(j)$ are at most $2d$ away from each other in $\mathbb{H}^2$. This gives a sequences of points exiting the rays $\mathfrak{h}_\pm$

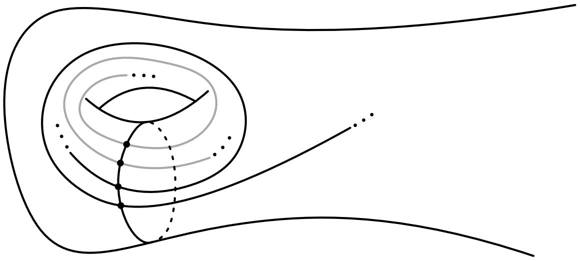


Figure 8: Two rays spiralling onto a simple closed curve (which is not allowed below), where the points in Claim 6.6 are linked.

that are always at most $2d$ apart, so $\mathfrak{h}_-$ is asymptotic to $\mathfrak{h}_+$. Hence, $\tilde{h}_-$ is asymptotic to $\tilde{h}_+$, contrary to our assumption.

We now fix large i, j such that $d_i, d_j < 2d$ and $\tilde{\gamma} := \tilde{\gamma}_{ij}$ is homotopically essential on $\partial\widetilde{M}$. Then $\tilde{\gamma}$ projects to a homotopially essential loop $\gamma \subset \partial M$ that is homotopically trivial in M ($\tilde{\gamma}$ is homotopically trivial in the simply connected space $\widetilde{M}$). Note that if i, j are chosen large enough and d is small, then $\gamma \subset S(\lambda)$. Furthermore, since the segments b_i, b_j are the only parts of $\tilde{\gamma}$ that intersect λ, and these segments have hyperbolic length less than $2d$, the intersection number $i(\gamma, \lambda)$ is small when d is small. (Recall that λ is a minimal lamination that is not a simple closed curve, so no leaves have positive weight, and hence hyperbolic length can be compared to intersection number.) But λ is not an intrinsic limit of meridians, so Proposition 5.12(4) says that there is some positive lower bound for the intersection numbers of λ with essential curves that are nullhomotopic in M. Hence, we get a contradiction if d is small. □

Suppose we have two pairs $\{a, b\}$ and $\{c, d\}$ of points in m, all four of which are distinct. We say the two pairs are *unlinked* in m if in the induced cyclic ordering on $\{a, b, c, d\} \subset m$, a is adjacent to b and c is adjacent to d, and we say that the two pairs are *linked* otherwise.

Claim 6.6 *If $i, j \in \mathbb{N}$, $i < j$, then the pairs $\{h_+(i), h_-(i)\}$ and $\{h_+(j), h_-(j)\}$ are unlinked in m.*

For an example where the pairs are *linked*, see Figure 8. The proof below works in general whenever $h_\pm$ are simple geodesic rays on ∂M in tight position with respect to m, where neither h_+ nor h_- spirals onto a simple closed curve.

Proof The essential observation used in the proof is that the closure

$$\mathrm{cl}(\partial\widetilde{M}) \subset \mathbb{H}^3 \cup \partial\mathbb{H}^3$$

is homeomorphic to a sphere: indeed, the closure of $\widetilde{M}$ in $\mathbb{H}^3 \cup \partial\mathbb{H}^3$ is a ball, since $\widetilde{M} \subset \mathbb{H}^3$ is convex with nonempty interior, and the closure of the boundary is the boundary of the closure. We obtain the unlinking property above by exploiting separation properties of arcs and curves on $\mathrm{cl}(\partial\widetilde{M})$.

Since $h_\pm$ are in tight position with respect to m, both lifts $\tilde{h}_\pm$ cross $\tilde{m}_i$ exactly once. Since $\tilde{h}_\pm$ limit to the same point in $\partial\mathbb{H}^3$, they must then cross $\tilde{m}_i$ in the same direction. In other words, the tangent vectors $h_+(i)'$, $h_-(i)'$ point to the same side of m. The same statement holds for j. This allows us to break into the following two cases:

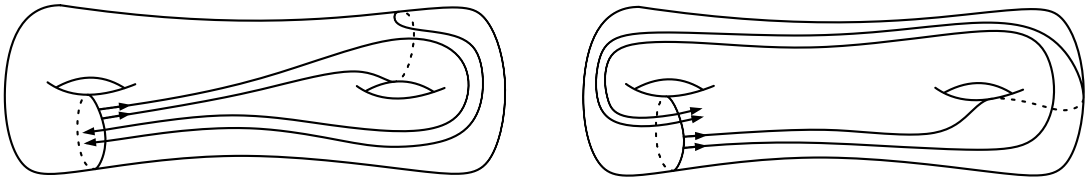


Figure 9: The cases (a) and (b) in the proof of Claim 6.6.

(a) the arcs $h_{\pm}|_{[i,j]}$ start and end on the same side of m, i.e., the vectors $h_{\pm}(i)'$ point to the opposite side of m as the vectors $h_{\pm}(j)'$, or

(b) the arcs $h_{\pm}|_{[i,j]}$ start and end on different sides of m, i.e., all four velocity vectors $h_{\pm}(i)', h_{\pm}(j)'$ point to the same side of m;

see Figure 9.

First, assume we're in case (a). Let

$$\alpha_{\pm} := \tilde{h}_{\pm}|_{[i,j]},$$

which we regard as oriented arcs in $\partial\widetilde{M}$ starting on $\widetilde{m}_i$ and ending on $\widetilde{m}_j$. Let $\gamma : \widetilde{M} \to \widetilde{M}$ be the deck transformation taking $\widetilde{m}_j$ to $\widetilde{m}_i$. Then the arcs

$$\beta_{\pm} := \gamma \circ \tilde{h}_{\pm}|_{[i,j]}$$

start on $\gamma(\widetilde{m}_i)$ and end on $\gamma(\widetilde{m}_j) = \widetilde{m}_i$, and since we're in case (a) they end on the *same side* of $\widetilde{m}_i$ as the arcs $\alpha_{\pm}$ start. Note that $\gamma(\widetilde{m}_i)$ is not $\widetilde{m}_i$ or $\widetilde{m}_j$. Indeed, if $\gamma(\widetilde{m}_i) = \widetilde{m}_i$ then we'd have $\gamma = \mathrm{id}$, contradicting that $\gamma(\widetilde{m}_j) = \widetilde{m}_i$. And if $\gamma(\widetilde{m}_i) = \widetilde{m}_j$, then γ^2 would leave $\widetilde{m}_i$ invariant, implying that $\gamma^2 = \mathrm{id}$, which is impossible since $\pi_1 M$ has no torsion.

We claim *the interiors of the arcs $\beta_{\pm}$ do not intersect $\widetilde{m}_i$ or $\widetilde{m}_j$, and the arcs $\alpha_{\pm}$ do not intersect $\gamma(\widetilde{m}_i)$.* Indeed, the interiors of $\beta_{\pm}$ don't intersect $\widetilde{m}_i$ because the arcs $\beta_{\pm}$ end on $\widetilde{m}_i$ and intersect each component of $\widetilde{m}$ at most once, by tight position. The interiors of $\beta_{\pm}$ don't intersect $\widetilde{m}_j$ because any arc from $\widetilde{m}_i$ to $\widetilde{m}_j$ intersect at least $j-i+1$ components of $\widetilde{m}$ (counting $\widetilde{m}_i$ and $\widetilde{m}_j$), while any proper subarc of $\beta_{\pm}$ intersects at most $j-i$ components of $\widetilde{m}$. Here, for the $j-i+1$ bound we are using tight position of $h_{\pm}$, the definitions of $\widetilde{m}_i, \widetilde{m}_j$, and the fact that each component of $\widetilde{m}$ separates $\partial\widetilde{M}$. The fact that the arcs $\alpha_{\pm}$ don't intersect $\gamma(\widetilde{m}_i)$ is similar: any arc from m_i to $\gamma(\widetilde{m}_i)$ must pass through at least $j-i+1$ components of $\widetilde{m}$, while any proper subarc of $\alpha_{\pm}$ intersects at most $j-i$ components, and $\alpha_{\pm}$ do not end on $\gamma(\widetilde{m}_i) \neq \widetilde{m}_j$.

Let $A \subset \mathrm{cl}(\partial\widetilde{M}) \cong S^2$ be the annulus that is the closure of the component of $\mathrm{cl}(\partial\widetilde{M}) \setminus (\widetilde{m}_i \cup \widetilde{m}_j)$ that contains the side of $\widetilde{m}_i$ on which the arcs $\alpha_{\pm}$ start and the arcs $\beta_{\pm}$ end. Then $\alpha_{\pm}$ are two disjoint arcs in A that join $\widetilde{m}_i$ to $\widetilde{m}_j$, and therefore $\alpha_{\pm}$ separate A into two rectangles. The component $\gamma(\widetilde{m}_i)$ on which the arcs $\beta_{\pm}$ start is contained in the interior of one of these two rectangles. Therefore, the two arcs $\beta_{\pm}$ must lie in the same component of $A \setminus (\alpha_+ \cup \alpha_-)$. Looking at endpoints, this means the pairs $\{\tilde{h}_+(i), \tilde{h}_-(i)\}$ and $\{\gamma \circ \tilde{h}_+(j), \gamma \circ \tilde{h}_-(j)\}$ are unlinked in $\widetilde{m}_i$, and the claim follows.

Now assume that we're in case (b). The curve $\widetilde{m}_i$ separates $\partial\widetilde{M}$, and we let $X \subset \partial\widetilde{M}$ be the closure of the component of $\partial\widetilde{M} \setminus \widetilde{m}_i$ into which the velocity vectors $\tilde{h}'_\pm(i)$ and $(\gamma \circ \tilde{h}_\pm)'(j)$ all point. The closure

$$\mathrm{cl}(X) \subset \mathbb{H}^3 \cup \partial\mathbb{H}^3$$

is homeomorphic to a disk, since $\mathrm{cl}(\partial\widetilde{M})$ is a sphere. As before, we let $\gamma : \widetilde{M} \to \widetilde{M}$ be the deck transformation taking $\widetilde{m}_j$ to $\widetilde{m}_i$. Then the rays

$$\alpha_\pm := \tilde{h}_\pm([i,\infty)), \quad \beta_\pm := \gamma \circ \tilde{h}_\pm([j,\infty))$$

are all contained in X. Note that $\alpha_\pm$ both limit to a point $\xi \in \partial\mathbb{H}^3$, while $\beta_\pm$ limit to $\gamma(\xi) \in \partial\mathbb{H}^3$.

The union $\alpha_- \cup \alpha_+$ compactifies to an arc in $\mathrm{cl}(X)$, since the two rays limit to the same point in $\mathbb{H}^3$. The same is true for $\beta_- \cup \beta_+$. *Hoping for a contradiction, suppose that the points in the statement of the claim are linked.* Then the pairs of endpoints of $\alpha_- \cup \alpha_+$ and $\beta_- \cup \beta_+$ are also linked on $\widetilde{m}_i = \partial\mathrm{cl}(X)$. We now have two arcs on the disk $\mathrm{cl}(X)$ with linked endpoints on $\partial\mathrm{cl}(X)$, so the two arcs must intersect. As $\alpha_\pm$, $\beta_\pm$ are all disjoint, the only intersection can be on $\partial\mathbb{H}^3$, so their endpoints at infinity must all agree, i.e., $\gamma(\xi) = \xi$.

Since $\gamma(\xi) = \xi$, all the rays $\gamma^k \circ \tilde{h}_+$ limit to ξ, where $k \in \mathbb{Z}$. Hence, all these (quasigeodesic) rays are pairwise mutually homoclinic, and for each pair k, l, the rays $\gamma^k \circ \tilde{h}_+$ and $\gamma^l \circ \tilde{h}_+$ eventually intersect the same components of $\widetilde{m}$, in the same order, although their initial behaviour may be different. In analogy with the setup of Claim 6.5, let $d_{k,l}$ be the liminf of the distances from $\gamma^k \circ \tilde{h}_+$ to $\gamma^l \circ \tilde{h}_+$ along the components of $\widetilde{m}$ that they both intersect.

We claim that there are k, l such that $d_{k,l} < \epsilon$, where ϵ is the constant from Claim 6.5. Indeed, for $N > \mathrm{length}(m)/\epsilon$, it is impossible to pack N points at least ϵ apart in any component of $\widetilde{m}$. So if we let k range over a set $F \subset \mathbb{Z}$ of size N, whenever a component of $\widetilde{m}$ intersects all $\gamma^k \circ \tilde{h}_+$, $k \in F$, two such intersections must be within ϵ of each other. There are infinitely many such components of $\widetilde{m}$ and F is finite, so we can pick $k, l \in S$ such that $\gamma^k \circ \tilde{h}_+$ and $\gamma^l \circ \tilde{h}_+$ are within ϵ on infinitely many such components.

Finally, $\gamma^k \circ \tilde{h}_+$ and $\gamma^l \circ \tilde{h}_+$ are mutually homoclinic lifts of $\tilde{h}_+$, and $d_{k,l} < \epsilon$, so the exact same argument as in Claim 6.5 shows that $\gamma^k \circ \tilde{h}_+$ and $\gamma^l \circ \tilde{h}_+$ are asymptotic on $\partial\widetilde{M}$. It follows that h_+ spirals onto a (simple) closed curve in ∂M in the homotopy class of (a primitive root of) γ^{l-k}. (Indeed, γ^{l-k} lifts to a deck transformation of the universal cover $\mathbb{H}^2 \to \partial M$, and the axis of this deck transformation is asymptotic to suitably chosen lifts of both $\gamma^k \circ \tilde{h}_+$ and $\gamma^l \circ \tilde{h}_+$.) This is a contradiction, though, since h_+ limits onto λ, which is not a simple closed curve. □

Assume now that our mutually homoclinic rays $\tilde{h}_\pm$ are not asymptotic on $\partial\widetilde{M}$, as otherwise we're in case (2) of the theorem and are done. By Claim 6.5, there is some $\epsilon > 0$ such that $d_{\widetilde{m}_i}(\tilde{h}_+(i), \tilde{h}_-(i)) \geq \epsilon$ for all i. We will show that λ is an intrinsic limit of annuli, in the sense of Lemma 5.13, which says that then λ is an intrinsic limit of meridians.

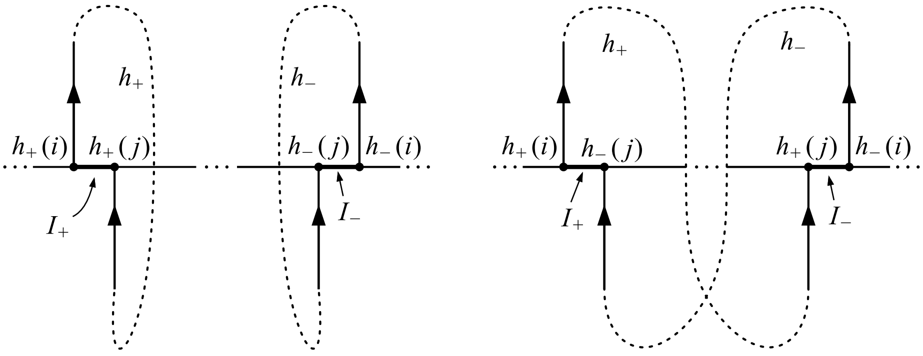


Figure 10: Above, the horizontal curve is always m.

The proof is an adaptation and correction of a surgery argument of Lecuire [42, Affirmation C.3]. As there are two gaps[9] in Lecuire's earlier argument, we give the proof in full detail below, without many citations of [42].

Claim 6.7 *Given* $0 < \delta < \epsilon$, *there are choices of* $i < j$ *such that either*:

(I) *The points* $h_+(i)$ *and* $h_+(j)$ *bound a segment* $I_+ \subset m$ *of length less than* δ, *and similarly with* $-$ *instead of* $+$. *The four velocity vectors* $h'_+(i)$, $h'_+(j)$, $h'_-(i)$, $h'_-(j)$ *all point to the same side of* m, *and the four segments* $h_+([i,j])$, $h_-([i,j])$, I_+ *and* I_- *have disjoint interiors. Therefore, the curves* $\gamma_\pm := h_\pm([i,j]) \cup I_\pm \subset \partial M$ *are simple and disjoint.*

(II) *The points* $h_+(i)$ *and* $h_-(j)$ *bound a segment* $I_+ \subset m$ *of length less than* δ, *and similarly the points* $h_-(i)$ *and* $h_+(j)$ *bound a segment* $I_- \subset m$ *of length less than* δ. *The four velocity vectors* $h'_+(i)$, $h'_+(j)$, $h'_-(i)$, $h'_-(j)$ *all point to the same side of* m, *and the four segments* $h_+([i,j])$, $h_-([i,j])$, I_- *and* I_+ *have disjoint interiors. Therefore, the curve* $\gamma \subset \partial M$ *obtained by concatenating all four segments is simple.*

See Figure 10 for a very useful picture. Note that in the picture, the velocity vectors of all paths intersecting m point to the same side of m, i.e., "up", and all 4-tuples of points are unlinked as in Claim 6.6.

Proof Start by fixing a circular order on m. Define "the right" to be the direction in m that is increasing with respect to the circular order, and define "the left" similarly. Since λ is minimal and not a simple closed curve, it has infinitely many leaves ℓ that are not boundary leaves. Fix some such ℓ, making sure that $h_\pm \not\subset \ell$ if the given rays happen to lie inside the lamination λ. The ray h_+ accumulates onto both sides of ℓ, so if we fix $p \in \ell \cap m$, the set $h_+(\mathbb{N})$ accumulates onto p from both sides, and similarly with $-$ instead of $+$. Fix an interval $J \subset m$ of length δ centred at p, and write $J = J_l \cup J_r$ as the union of the closed subintervals to the "left" and to the "right" of p. Note that $p \notin h_\pm(\mathbb{N})$, so each intersection of $h_\pm$ with J lies in exactly one of J_l or J_r.

---

[9] The first gap is that the sentence "*Quitte à extraire, la suite* $(gh^{-1})^{2n} g(\tilde{l}^1)$ *converge vers une géodésique* $\tilde{\gamma} \subset p^{-1}(\alpha_1)$ *dont la projection* $l \subset \partial M$ *est une courbe fermé.*" at the end of the proof of Affirmation C.3 isn't adequately justified; this is fixed in Claim 6.5. The second is that the assumption $d(l^2_+(y_i), l^2_+(y_j)) < \epsilon'$ in the statement of Affirmation C.3 is never actually verified, and does not come trivially from a compactness argument. This is fixed in Claim 6.7.

Let's call an index i *left-closest* if either $h_+(i)$ or $h_-(i)$ lies in J_l and is closer to p than any previous $h_\pm(k)$, $k < i$, that lies in J_l. *Right-closest* is defined similarly using J_r, and we call an index i *closest* if it is either left or right closest. Note that since $\delta < \epsilon$ we can never have both $h_+(i), h_-(i)$ in J simultaneously, so no i is both right-closest and left-closest at the same time. Since there are infinitely many i of both types, at some point there will be a transition where some i_l is left-closest, some $i_r > i_l$ is right-closest, and there are no closest indices in between.

Let i_c be the smallest closest index that is bigger than i_r. (Here, c stands for "centre", since the corresponding point on J will lie between the points we get from the indices i_l and i_r.) We now have *three* points in J, so two of the corresponding velocity vectors point to the same side of m. Let $i, j \in \{i_l, i_r, i_c\}$ be the two corresponding indices, and for concreteness, *let's assume for the moment that* $h_+(i)$ *and* $h_+(j)$ *are the corresponding points in* J, deferring a discussion of the other cases to the end of the proof. Note that since the rays $h_\pm$ are mutually homoclinic and are in tight position with respect to m, the velocity vectors $h'_-(i)$ and $h'_-(j)$ point to the same sides of m as $h'_+(i)$ and $h'_+(j)$, respectively, and so all four vectors point to the same side. That is,

(a) $h'_+(i), h'_+(j), h'_-(i), h'_-(j)$ all point to the same side of m,

(b) the segment $I_+ \subset J$ bounded by $h_+(i)$ and $h_+(j)$ contains no element $h_+(k)$ or $h_-(k)$ where k is between i and j.

Let $I_- \subset m \setminus J$ be the segment that is bounded by the points $h_-(i)$ and $h_-(j)$. *Suppose for a moment that we knew that* I_- *had length less than* δ. Then for each k, it is impossible that *both* $h_+(k)$ or $h_-(k)$ lie in I_-, as we're assuming that corresponding intersections of $h_\pm$ with m stay at least $\epsilon > \delta$ apart. In particular, if k is between i, j and we apply the unlinking condition of Claim 6.6 twice, once to i, k and once to j, k, we get from this and (b) above that *neither* element $h_+(k)$ or $h_-(k)$ is contained in I_-. So, the four segments $h_+([i, j])$, $h_-([i, j])$, I_+ and I_- have disjoint interiors, and we're in the situation of case (I) in the claim, as desired.

As constructed above, however, there is unfortunately no reason to believe that the interval I_- has length less than δ. To rectify this, recall that λ actually has infinitely many nonboundary leaves ℓ^n. For each such ℓ^n and $p^n \in \ell^n \cap m$, we can repeat the above construction using constants $\delta^n \to 0$, producing points (say) $h_+(i^n), h_+(j^n)$ that lie within the length δ^n-interval $J^n \ni p^n$ and that satisfy properties (a) and (b) above. Choose the sequence $p^n \in \ell^n \cap m$ so that it is monotonic in the circular order induced on m, and let $\delta^n \to 0$ fast enough so that the associated intervals I_+^n are all disjoint, so that in the circular order on m we have

$$h_+(i^1) < h_+(j^1) < h_+(i^2) < h_+(j^2) < \cdots < h_+(i^n),$$

and where each I_+^n is the interval $[h_+(i^n), h_+(j^n)]$, rather than the complementary arc on m that has endpoints $h_+(i^n), h_+(j^n)$. Then Claim 6.6 implies that

$$h_-(i^1) > h_-(j^1) > h_-(i^2) > h_-(j^2) > \cdots > h_-(i^n).$$

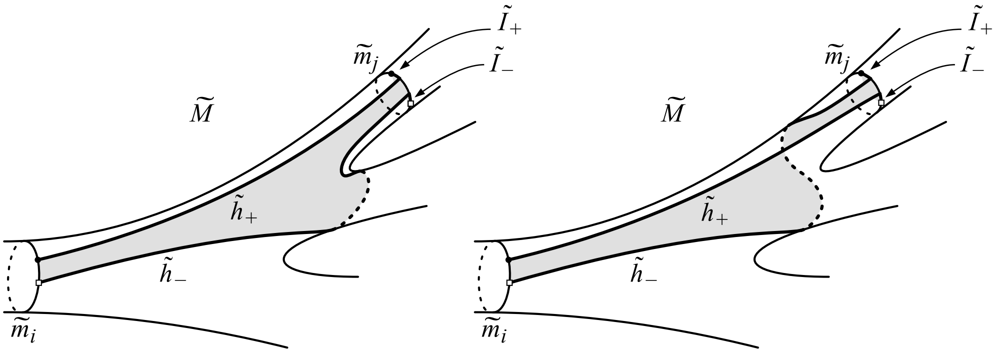


Figure 11: On the left, the two paths drawn in heavy ink project to the two simple closed curves $\gamma_\pm$ in Claim 6.7(I), shown on the left in Figure 10. The shaded region is a rectangle embedded in $\widetilde{M}$ that projects to an embedded annulus $A \hookrightarrow M$ with boundary $\gamma_- \cup \gamma_+$. On the right, the union of the two paths projects to the simple closed curve γ on M of Claim 6.7(II), and the shaded region projects to a Möbius band $B \hookrightarrow M$ with boundary γ.

Discarding finitely many n, we can assume all the point $h_+(i^n)$, $h_+(j^n)$ lie in an interval $U \subset m$ of length less than δ. Since the points $h_-(i^n)$, $h_-(j^n)$ are at least $\epsilon > \delta$ away from the corresponding $+$ points, they all lie in $m \setminus U$. Then since the interval I_-^n is defined to be disjoint from $I_+^n \subset U$, we must have $I_-^n = [h_-(j^n), h_-(i^n)]$, rather than the other interval with those endpoints. It follows that at least for large n, all the intervals I_-^n are disjoint. Since m is compact, we can then pick some n where I_-^n has length less than δ, as desired. Therefore, we are in case (I) in the statement of the claim, and are done.

In the argument above we simplified the notation by assuming that we have points $h_+(i^n), h_+(j^n) \in J^n$ satisfying conditions (a) and (b), which put us in case (I) at the end. Up to exchanging $+, -$, the only other relevant case is when our chosen points are $h_-(i^n), h_+(j^n) \in J^n$. After passing to a subsequence in n, if we are not in the case already addressed, then we may assume that our chosen points are $h_-(i^n), h_+(j^n) \in J^n$ for all n. And after exchanging $+$ with $-$ and passing to a further subsequence, we may assume

$$h_-(i^1) < h_+(j^1) < h_-(i^2) < h_+(j^2) < \cdots < h_-(i^1)$$

in the circular order on m, and that the interval $I_+^n = [h_-(i^n), h_+(j^n)]$. Everything from then on works exactly as above: if we set I_-^n to be the interval bounded by $h_+(i^n)$, $h_-(j^n)$ that is disjoint from I_+^n, then for some n we have that the length of I_-^n is less than δ, and it is easy to verify that we are in case (II) of the claim. □

We now finish the proof of Theorem 6.1. Suppose we are in case (I) of Claim 6.7. Then the two simple closed curves $\gamma_\pm$ drawn on the left in Figure 10 are the projections to M of the paths in $\widetilde{M}$ obtained by concatenating the arcs $\tilde{h}_\pm|_{[i,j]}$ with lifts $\widetilde{I}_\pm \subset \widetilde{m}_j$ of the intervals $I_\pm \subset m$; see Figure 11. We can homotope one path to the other in $\widetilde{M}$ while preserving the fact that the endpoints are points on $\widetilde{m}_i$, $\widetilde{m}_j$ that differ by the unique deck transformation taking $\widetilde{m}_i$ to $\widetilde{m}_j$. So projecting down, the simple closed curves $\gamma_\pm$ are freely homotopic in M, and hence bound an annulus $A \hookrightarrow M$. See the left part of Figure 11.

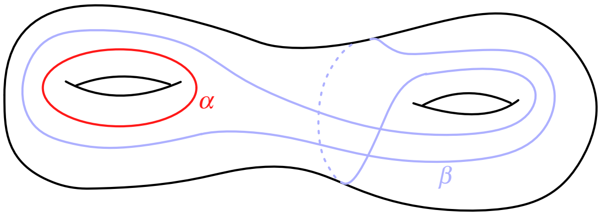


Figure 12: The two curves α and β are homotopic through the handlebody pictured, and therefore bound an essential singular annulus. The only annuli one can produce from surgery are inessential, but one can surger to obtain the "obvious" disc in the picture that separates the handlebody into two solid tori.

There is a uniform lower bound (depending on λ, m) for the angle at any intersection point of any leaf of λ with m, and the points $h_\pm(i)$ are at least ϵ away from each other in m. This implies that there is a uniform lower bound for the Hausdorff distance between $h_\pm|_{[i,j]}$ on ∂M. As long as the bound δ on the lengths of $I_\pm$ is small enough, the geodesics in the homotopy classes of $\gamma_\pm$ stay very close to $h_\pm|_{[i,j]}$, and are therefore distinct. So, the curves $\gamma_\pm$ are not homotopic in ∂M, and hence bound an *essential* annulus $A \hookrightarrow M$.

Choosing i, j to be large and δ to be very small, ∂A is contained in $S(\lambda)$ and its intersection number with λ is small. Hence, λ is an intrinsic limit of annuli, in the sense of Lemma 5.13, so we're done.

Case (II) is similar. Here, the single simple closed curve γ described in Claim 6.7(II) bounds a Möbius band $B \hookrightarrow M$; see the right side of Figure 11. Since ∂M is orientable, B is not boundary parallel, and hence by JSJ theory the boundary of a regular neighbourhood of B is an essential annulus $A \hookrightarrow M$ whose boundary consists of two disjoint curves that are both homotopic to γ on ∂M. As in case (I), we can make the intersection number of ∂A with λ arbitrarily small, so λ is an intrinsic limit of annuli, and we are done.

Remark 6.8 The proof (D) above is quite delicate. Most of this delicacy comes from Claims 6.5 and 6.7, which are needed to ensure that the annuli approximating λ that are produced immediately afterward are *embedded*. But while we are able to prove these claims using arguments involving the planarity of the closure of $\partial\widetilde{M}$ in $\mathbb{H}^3 \cup \partial_\infty \mathbb{H}^3$, one would not have to worry about these annuli being embedded if there was a strong "annulus theorem" guaranteeing that any essential singular annulus in an irreducible 3-manifold M can be surgered to give an essential embedded annulus. If this were true, Claims 6.5 and 6.7 could be replaced by a one paragraph compactness argument. Here, a *singular annulus* is a map $f : (A, \partial A) \to (M, \partial M)$ where $A = S^1 \times [0,1]$. We say f is *essential* if it is not homotopic rel ∂A into ∂M.

Such an annulus theorem follows from the JSJ decomposition when M has incompressible boundary. When M has compressible boundary, there is a similar theorem as long as the original singular annulus has a spanning arc that is not homotopic rel ∂ into ∂M; see Cannon and Feustel's article [15]. However, our proof above does not provide such annuli, and indeed such annuli do not exist in compression bodies (the M of most interest to us), since *any* proper arc in a compressionbody M is homotopic rel ∂ into ∂M.

In fact, in a general M, one *cannot* always surger essential singular annuli to produce embedded essential annuli. For instance, the two curves in Figure 12 bound an essential singular annulus that cannot be surgered to give an embedded essential annulus. However, in that example, one *can* surger to get a meridian in the handlebody, so maybe an essential singular annulus can always be surgered to give either a meridian or an essential embedded annulus? This also turns out not to be true. Suppose P is a pair of pants and let $M = P \times [0,1]$, which is homeomorphic to a genus-two handlebody. If $\gamma \to P$ is an immersed figure-8 whose image forms a spine of P, then the singular annulus $\gamma \times [0,1] \to P \times [0,1]$ is essential, but the three embedded annuli that one can obtain from it by surgery are all inessential, and no meridian can be created by surgery either. However, we expect this is the *only* counterexample. The first author of this paper has spent considerable time trying to prove this with a tower argument, but pushing down the tower is very subtle, since if the obvious constructions fail, one has to characterise the figure-8 example.

7 Hausdorff limits of meridians

Let M be an orientable hyperbolisable compact 3-manifold, and equip $\partial_{\chi<0}M$ with a hyperbolic metric. Lecuire [42, Theorem B.1] showed that every lamination λ on $\partial_{\chi<0}M$ that is a Hausdorff limit of meridians contains a homoclinic leaf that is a homoclinic geodesic. This is a more general version of Casson's criterion for handlebodies, which was stated in the introduction.

The converse is not true: certainly in order to be a Hausdorff limit of meridians μ needs to be connected. And there are even connected laminations that contain homoclinic leaves but are not commensurable to Hausdorff limits of meridians. One way to do this is to just take a lamination that contains a meridian, but is not a limit even of simple closed curves, as on the left in Figure 13. There are also more subtle examples in genus 2, as pictured on the right in Figure 13: the reason they are not limits of meridians is as follows.

Lemma 7.1 *Suppose that S is a closed, genus-two surface, and λ is a geodesic lamination on S such that there is a separating meridian μ that does not intersect λ transversely, and λ intersects transversely the two nonseparating meridians disjoint from μ. Then λ is not a Hausdorff limit of meridians.*

Proof Let $T_\pm \subset S \setminus \mu$ be the two components of $S \setminus \mu$. Hoping for a contradiction, take a sequence of meridians (m_n) that Hausdorff converges to λ. We can assume after passing to a subsequence that m_n has a μ-wave in T_+ (say) for all n. Since T_+ is a compressible punctured torus, there is a *unique* homotopy class rel μ of μ-wave in T_+, so λ contains a leaf ℓ that either intersects T_+ in an arc in this homotopy class, or is contained in T_+ and is obtained by spinning an arc in this homotopy class around μ. But then ℓ

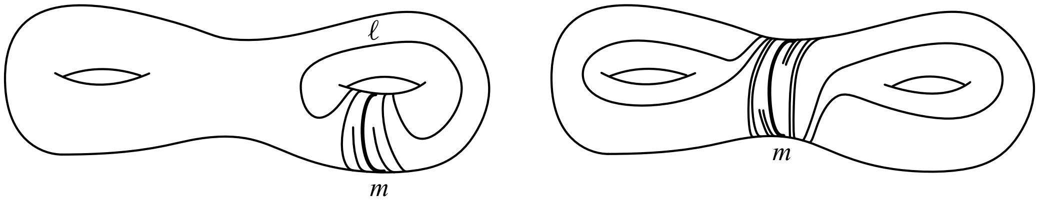


Figure 13: Two laminations on the boundary of a handlebody that have a meridian m as a leaf, but are not Hausdorff limits of meridians.

intersects nontrivially every nonperipheral minimal lamination in T_+ other than the unique nonseparating meridian μ_+ of T_+, so λ is disjoint from μ_+, contrary to assumption. □

In the examples in Figure 13, the problem is the spiralling leaves. So, maybe being a Hausdorff limit of meridians is the same as containing a homoclinic leaf if we ignore spiralling leaves? We say that two laminations μ_1, μ_2 are *commensurable* if they contain a common sublamination ν such that for both i, the difference $\mu_i \setminus \nu$ is the union of finitely many leaves. μ_1 and μ_2 are *strongly commensurable* if they contain a common ν such that for both i, the difference $\mu_i \setminus \nu$ is the union of finitely many leaves, none of which are simple closed curves.

Theorem 7.2 (Hausdorff limits of meridians) *Suppose that $S \subset \partial_{\chi<0} M$ is a connected subsurface with geodesic boundary, such that ∂S is incompressible, and that the disc set $\mathcal{D}(S, M)$ is "large", i.e., it has infinite diameter in the curve graph $C(S)$. Let λ be a geodesic lamination in* $\operatorname{int}(S)$ *that is a finite union of minimal laminations, and assume that the following does* **not** *hold:*

($\star$) *S is a closed, genus-two surface, there exists a separating meridian μ that does not intersect λ transversely, and λ intersects transversely the two nonseparating meridians disjoint from μ.*

Then λ is strongly commensurable to a Hausdorff limit of meridians in S if and only if λ is strongly commensurable to a lamination containing a homoclinic leaf, and this happens if and only if one of the following holds:

1. *λ is disjoint from a meridian on S,*
2. *some component of λ is an intrinsic limit of meridians, or*
3. *there is an essential (possibly nontrivial) interval bundle $B \subset M$ over a compact surface Y that is not an annulus or Möbius band, and there are components $\lambda_\pm \subset \lambda$ that each fill a component of $\partial_h B$ (possibly the same component, if $\partial_h B$ is connected), such that $\lambda_\pm$ are essentially homotopic through B, as in Section 2.9, and there is a compression arc α for B that is disjoint from λ.*

The case ($\star$) above really is exceptional. Here, one should imagine a picture like the example on the right in Figure 13, but with the spiralling leaves replaced with minimal laminations in the two punctured tori. At least when $\mu \subset \lambda$ we have that λ contains a homoclinic leaf, but one can see that λ is *not* strongly commensurable to a Hausdorff limit of meridians, by using Lemma 7.1.

Recall from Proposition 3.1 that when $\mathcal{D}(S, M)$ does not have infinite diameter in $C(S)$, it is either empty, consists of a single separating meridian, or consists of a single nonseparating meridian m and all separating curves that are band sums of m. In these cases, it is obvious what the Hausdorff limits of meridians are. For instance, in the last case a finite union λ of minimal laminations in S is strongly commensurable to a Hausdorff limit of meridians if and only if either $\lambda = m$ or $\lambda \subset S \setminus m$. For the "if" direction, note that if $\lambda \subset S \setminus m$ then it can be approximated by an arc with endpoints on opposite sides of m, and doing a band sum with m gives a curve that approximates λ. For the "only if" direction, just note that all meridians are either equal to m or are contained in $S \setminus m$.

7.1 The proof of Theorem 7.2

Most of the proof of Theorem 7.2 is contained in the following results. Assume that $S \subset \partial_{\chi<0} M$ is a connected subsurface with geodesic boundary, ∂S is incompressible, and $\mathcal{D}(S, M)$ is large.

Lemma 7.3 *Suppose $\lambda \subset S$ is a lamination, there is a meridian μ that does not intersect λ transversely, and that if S is a closed surface of genus 2 then μ is nonseparating. Then λ is strongly commensurable to a Hausdorff limits of meridians on S.*

The proof of Lemma 7.3 uses some ideas that the first author developed with Sebastian Hensel, whom we thank for his contribution.

Proof We may assume that λ is a finite union of minimal components. It suffices to assume μ is not a leaf of λ, as long as we prove the conclusion both for such a λ and for $\lambda \cup \mu$.

Assume first that μ is nonseparating in S. Let (c_i) be a sequence of simple closed curves on S that Hausdorff-converges to λ. One can do this by constructing for each component $\lambda_0 \subset \lambda$ a simple closed geodesic approximating λ_0, by taking an arc that runs along a leaf of λ_0 for a long time, and then closing it up the next time it passes closest to its initial endpoint in the correct direction. Let α be a simple closed curve on S that intersects μ once, and intersects all the components of λ. For each k, let γ_i^k be the geodesic homotopic to the "band sum" of μ and $T_{c_i}^k(\alpha)$, where T_{c_i} is the twist around c_i and a band sum of two curves intersecting once is the boundary of a regular neighbourhood of their union. Note that γ_i^k is a meridian for all i, k. If (k_i) is a sequence that increases quickly enough, $(\gamma_i^{k_i})$ converges to a lamination strongly commensurable to λ. And if we pick a meridian β on S that intersects both μ and λ, then $T_\mu^i \circ T_{\gamma_i^{k_i}}^i(\beta)$ Hausdorff converges to a lamination strongly commensurable to $\lambda \cup \mu$.

Now suppose μ is separating. We claim that there is another separating meridian in S that is disjoint from μ. Let m be a maximal multicurve of meridians in S that contains μ as a component. Since $\mathcal{D}(S, M)$ is large, $m \neq \mu$. If m has a separating component distinct from μ, we are done. So, suppose we have a nonseparating component $m_0 \subset m$. We can make a (separating) band sum of m_0 that is disjoint from μ unless m_0 lies in a punctured torus component of $S \setminus \mu$. So, we assume the latter is true. Since $\mathcal{D}(S, M)$ is large, it cannot be that $m = \mu \cup m_0$, since then all meridians are disjoint from m_0. So, there is another component m_1 of m, which we can assume is also nonseparating. This m_1 must lie on the opposite side of μ from m_0, and as before we're done unless the component of $S \setminus \mu$ containing m_1 is also a punctured torus. But in this case, S is a genus-two surface contrary to assumption.

Let $T \subset S \setminus \mu$ be a component that contains a nonperipheral separating meridian, which we call μ'. Let V be the other component. Write $\lambda = \lambda_T \cup \lambda_V$, where $\lambda_T \subset T$ and $\lambda_V \subset S \setminus T$. Let C be the compression body with exterior boundary equal to the component of ∂M that contains S, that one obtains by compressing the meridian μ.

We claim that there are sequences of simple closed curves (α_i), (β_i) in T such that

- (α_i) and (β_i) both Hausdorff converge to a geodesic lamination strongly commensurable to λ_T, and
- for all i, α_i and β_i bound an essential annulus in C.

To construct these sequences, start by picking a simple closed curve α in T such that α and each component of λ_T together fill T. Let β be a simple closed curve on T such that α, β, μ bound a pair of pants in T. In C, we can compress the boundary component μ of this pair of pants, so α, β bound an annulus in C. Moreover, this annulus is essential, since otherwise α, β bound an annulus in T, implying T is torus with the one boundary component μ, contradicting the fact that there is a separating nonperipheral meridian in T. Then find a sequence (c_i) of simple closed curves in T that Hausdorff converge to a geodesic lamination strongly commensurable to λ_T, take k_i to be a fast increasing sequence and set $\alpha_i = T_{c_i}^{k_i}(\alpha)$ and $\beta_i := T_{c_i}^{k_i}(\beta)$. Since α fills with every component of λ_T, the curve β intersects every component of λ_T. It follows that (α_i) and (β_i) Hausdorff converge to a geodesic lamination strongly commensurable to λ_T. And since each c_i is nonperipheral in T, each component of c_i bounds an annulus in C with a curve on the interior boundary of C, so the twist T_{c_i} extends to C, implying that α_i, β_i bound an annulus in C as desired above.

Now let C' be the compression body obtained by compressing both μ and μ', so $C \subset C' \subset M$. Note that since both curves are separating and are disjoint, Proposition 3.1 says that $\mathcal{D}(S, C')$ is large, so we can pick a meridian $m \in \mathcal{D}(S, C')$ that intersects μ and every component of λ. Fix a sequence of geodesic multicurves (d_i) in V that Hausdorff converges to λ_V. As with the twists T_{c_i} in C, the twists T_{d_i} extend to C'. And the compositions $T_{\alpha_i} \circ T_{\beta_i}^{-1}$ extend to C' because the curves bound annuli in $C \subset C'$. We then define $\gamma_i := (T_{\alpha_i} \circ T_{\beta_i}^{-1})^{k_i} \circ T_{d_i}^{k_i}(m)$ for some fast increasing $k_i \to \infty$. These γ_i are all meridians and Hausdorff converge to a lamination strongly commensurable to λ. To obtain $\lambda \cup \mu$ instead, hit γ_i with high powers of twists around μ. □

Here is a more powerful version of Lemma 7.3. The idea of the proof is more or less the same, but more complicated.

Proposition 7.4 (promoting Hausdorff limits) *Suppose that ν, η are disjoint geodesic laminations on S that are finite unions of minimal components. Suppose also that no component of ν is a meridian.*

Let X be the union of the subsurfaces with geodesic boundary that are filled by the components of ν. Suppose that there are disjoint, nonhomotopic meridians μ, μ' on S that are disjoint from η, and a sequence of homeomorphisms

$$f_i : S \to S, \quad f_i|_{S \setminus \mathrm{int}(X)} = \mathrm{id},$$

such that $\mu_i := f_i(\mu)$ and $\mu_i' := f_i(\mu')$ are both sequences of meridians that Hausdorff converge to laminations strongly commensurable to ν. Then $\nu \cup \eta$ is strongly commensurable to a Hausdorff limit of meridians in S.

Before proving the proposition, we record the following application.

Corollary 7.5 *Suppose that λ is a geodesic lamination on S that is a finite union of minimal components. If either*

- *some component $\nu \subset \lambda$ that is not a simple closed curve is an intrinsic limit of meridians,*
- *there are (possibly equal) components $\lambda_\pm \subset \lambda$, neither of which is a simple closed curve, and where each fills a component of the horizontal boundary (possibly the same component if $\partial_h B$ is connected) of*

an essential interval bundle

$$(B, \partial_H B) \hookrightarrow (M, S),$$

where $\lambda_\pm$ *are essentially homotopic through* B*, and where there is a compression arc* α *for* B *that is disjoint from* λ*,*

then λ *is strongly commensurable to a Hausdorff limit of meridians.*

Proof Suppose some component $\nu \subset \lambda$ that is not a simple closed curve is an intrinsic limit of meridians. Setting $X := S(\lambda)$ we can take (μ_i) to be any sequence of meridians in X that Hausdorff converges to a lamination strongly commensurable to ν. Moreover, since ν fills X and is a limit of meridians, the disc set $\mathcal{D}(X, M)$ is large, so for each i there is some meridian μ'_i disjoint from μ_i. Since there are only finitely many topological types of pairs of disjoint curves in X up to the pure mapping class group of X, after passing to a subsequence we can assume that all μ_i, μ'_i are of the form in the proposition. The desired conclusion follows.

In the second case, we let X be the subsurface with geodesic boundary obtained by tightening $\partial_H B$ and set $\nu = \lambda_- \cup \lambda_+$. Write the interval bundle as $\pi : B \to Y$, where Y is a compact surface with boundary. We can assume without loss of generality that α is a strict compression arc, i.e., that it is homotopic rel endpoints to a fibre $\pi^{-1}(y)$, $y \in \partial Y$. Note that since $\lambda_\pm$ are not simple closed curves, Y is not an annulus or Möbius band.

Since $\lambda_\pm$ are essentially homotopic through B, Fact 2.16 says that if our reference hyperbolic metrics are chosen appropriately, we have that $\lambda_- \cup \lambda_+ = (\pi|_{\partial_H B})^{-1}(\bar{\lambda})$ for some geodesic lamination $\bar{\lambda}$ on Y. Since $\lambda_\pm$ together fill $\partial_H B$, the lamination $\bar{\lambda}$ is minimal and fills Y. So in particular, it has no closed, one-sided leaves, and therefore if we pick a nonzero transverse measure on $\bar{\lambda}$, we have that $\bar{\lambda}$ is the projective limit of a sequence of two-sided nonperipheral simple closed curves (c_i) in Y, by Theorem 1.2 of [20]. Homotope the c_i on Y to be based simple loops at $y \in \partial Y$, let $m(c_i)$ be the associated compressible curves on S constructed in Claim 2.11, and let μ_i be the geodesic meridians on S in their homotopy classes. Then (μ_i) Hausdorff converges to a lamination strongly commensurable to $\lambda_- \cup \lambda_+$. After passing to a subsequence, we can assume that all the c_i differ by pure homeomorphisms of Y, in which case the meridians μ_i are as required in Proposition 7.4, for some μ, f_i. Note that since our compression arc is assumed to be disjoint from λ, all the μ_i are disjoint from $\eta := \lambda \setminus \lambda_\pm$, and hence so is our μ. We create disjoint meridians μ'_i similarly, by taking some c'_i on Y disjoint from c_i, and letting μ'_i be the geodesic meridian homotopic to $m(c'_i)$. It then follows from Proposition 7.4 that λ is strongly commensurable to a Hausdorff limit of meridians as desired. □

We now prove the proposition.

Proof of Proposition 7.4 Assume that μ, μ' are disjoint meridian on S that are disjoint from η, that $f_i : S \to S$ are homeomorphisms that are the identity outside of X, and that $\mu_i := f_i(\mu)$ and $\mu'_i := f_i(\mu')$ are sequences of meridians that Hausdorff converge to laminations strongly commensurable to ν.

We want to show that $\nu \cup \eta$ is strongly commensurable to a Hausdorff limit of meridians on S. We now basically repeat the argument in Lemma 7.3, so the reader should make sure that they understand that argument before continuing here.

Suppose μ (say) is nonseparating in S. Choose a simple closed curve α on S that intersects μ once and intersects essentially each component of η. Let $\alpha_i := f_i(\alpha)$, and note that α_i intersects μ_i once, and also intersects essentially each component of η. Let (c_i) be a geodesic multicurve that Hausdorff converges to η, and let γ_i^k be the geodesic homotopic to the "band sum"

$$B(\mu_i, T_{c_i}^k(\alpha_i)) = T_{c_i}^k(B(\mu_i, \alpha_i)) = T_{c_i}^k \circ f_i(B(\mu, \alpha)), \tag{4}$$

where here $B(\cdot,\cdot)$ takes in two simple closed curves that intersect once and returns the boundary of the regular neighbourhood of their union. If one of the inputs in a band sum is a meridian, then so is the output, so γ_i^k is a meridian for all i, k. The given equalities are true at least for large i. The first equality holds because μ is disjoint from η, $f_i = \mathrm{id}$ on the subsurfaces filled by the components of η, and therefore μ_i is disjoint from c_i for large i. The second equality is obvious from the definitions of μ_i, α_i.

Let (k_i) be a fast increasing sequence. After passing to a subsequence, we can assume that $(\gamma_i^{k_i})$ Hausdorff converges to a lamination λ. *We claim that λ is strongly commensurable to $\nu \cup \eta$.*

First, using the second term in (4), if k_i is huge with respect to i, then c_i is contained in a small neighbourhood of $\gamma_i^{k_i}$, and so since (c_i) Hausdorff converges to η, *we have $\lambda \supset \eta$.*

We claim that each $\gamma_i^{k_i}$ essentially intersects each component $X_0 \subset X$. If not, then from the third term in (4) it follows that $B(\mu, \alpha)$ is disjoint from X_0. But μ essentially intersects X_0, since otherwise the Hausdorff limit of the μ_i will not contain the associated component $\nu_0 \subset \nu$. So, μ, α and X_0 all lie in the punctured torus $T \subset S$ bounded by $B(\mu, \alpha)$. But since α intersects every component of η, we have that η intersects T as well, in a collection of arcs disjoint from μ. Since X_0 is disjoint from η, $X_0 = \mu$, so $\nu_0 = \mu$ is a meridian, contrary to our standing assumption.

It now follows that the Hausdorff limit λ essentially intersects each component of X. Since $\gamma_i^{k_i}$ is disjoint from μ_i and (μ_i) Hausdorff converges to a lamination containing ν, The laminations λ, ν cannot intersect transversely. Since each component of X is filled by a component of ν, *we have $\lambda \supset \nu$.*

Finally, if Y is the union of all the subsurfaces with geodesic boundary that are filled by the components of η, then as $f_i = \mathrm{id}$ outside X and $c_i \subset Y$ for large i, the intersection of $\gamma_i^{k_i}$ with $S \setminus (X \cup Y)$ is properly homotopic to the intersection of $B(\mu, \alpha)$, which is independent of i. It follows that $\lambda \setminus (\nu \cup \eta)$ *is a finite collection of nonclosed leaves*, so we are done.

We can now assume that both μ, μ' are separating, so that μ_i, μ_i' are also separating for all i. Let $T_i \subset S \setminus \mu_i$ be the component containing μ_i', and let V_i be the other component. Note that T_i is not a punctured torus, since it contains a nonperipheral separating curve. Since $\partial T_i \cap \eta = \varnothing$, we have

$$\eta = \eta_T \sqcup \eta_V,$$

where the first term is the intersection of η with T_i, and the second term is defined similarly. Note that since $f_i = \mathrm{id}$ on $S \setminus X$, all the μ_i induce the same two-element partition of the components of $S \setminus X$, so at least after passing to a subsequence the decomposition of η above is actually independent of i, which is why we have omitted the i in the notation.

Let C be the compression body whose exterior boundary is the component of ∂M containing S, and which is obtained by compressing the curve μ. Let C' be similarly obtained by compressing both μ and μ', so that $C \subset C' \subset M$. Since C' admits two nonhomotopic disjoint separating meridians, the disc set $\mathcal{D}(S, C')$ is large by Proposition 3.1, so we can pick a meridian $m \in \mathcal{D}(S, C')$ that intersects every component of $\nu \cup \eta$, as well as μ, μ'. Let $C_i \subset C_i' \subset M$ be the compression bodies obtained by compressing μ_i, μ_i'. Then f_i extends to a map $C' \to C_i'$, implying that $m_i := f_i(m)$ is a meridian in C_i'.

As in the proof of Lemma 7.3, we can pick sequences (α_i), (β_i) of simple closed curves in T_i such that (α_i) and (β_i) both Hausdorff converge to η_T, and where α_i, β_i bound an essential annulus in C_i for all i. As in the lemma, $T_{\alpha_i} \circ T_{\beta_i}^{-1}$ extends to C_i'. Let (c_i) be a sequence of multicurves in V_i that Hausdorff converges to η_V. Each component of c_i bounds an annulus in C_i' with a curve on the interior boundary of C_i', and hence the multitwist T_{c_i} extends to a homeomorphism of C_i'. For any given k, set

$$\gamma_i^k := (T_{\alpha_k} \circ T_{\beta_k}^{-1})^k \circ T_{c_k}^k(m_i).$$

We claim that for fast increasing k_i, the curves $\gamma_i^{k_i}$ Hausdorff converge to a lamination that is strongly commensurable to $\nu \cup \eta$ as desired. This is proved using the same types of arguments we employed in the nonseparating case above. In particular, recall that m was selected to intersect all components of $\nu \cup \eta$. Since f_i is supported on subsurfaces filled by components of ν, all the $m_i = f_i(m)$ intersect all components of $\nu \cup \eta$, and hence for large k_i they intersect $\alpha_{k_i}, \beta_{k_i}$. So, $\gamma_i^{k_i}$ is twisted many times around $\alpha_{k_i}, \beta_{k_i}$, and hence its Hausdorff limit contains ν. Similarly, the m_i intersect c_{k_i} for large i. Since c_k lies in V_k, it is disjoint from $\alpha_k \subset T_k$ and $\beta_k \subset T_k$, and thus the Hausdorff limit of $\gamma_i^{k_i}$ contains η. Finally, the Hausdorff limit has no other minimal components because α_k, β_k, c_k are contained in subsurfaces filled by components of $\nu \cup \eta$, and $m_i = f_i(m)$ is constant outside this subsurfaces. □

We can now start the proof of the theorem.

Proof of Theorem 7.2 Suppose that $\lambda \subset S$ is a lamination and $(\star)$ does not hold, so that it is not the case that S is a genus-two surface and λ is disjoint from a separating meridian μ, but intersects the two nonseparating meridians disjoint from μ. We want to show that λ is strongly commensurable to a Hausdorff limit of meridians if and only if it is strongly commensurable to a lamination containing a homoclinic leaf, which happens if and only if either

1. λ is disjoint from a meridian,
2. some component of λ is an intrinsic limit of meridians, or
3. there is an essential (possibly nontrivial) interval bundle $B \subset M$ over a compact surface Y that is not an annulus or Möbius band, and there are components $\lambda_\pm \subset \lambda$ that each fill a component of $\partial_H B$, such that $\lambda_\pm$ are essentially homotopic through B, as in Section 2.9, and there is a compression arc α for B that is disjoint from λ.

Hausdorff limit $\Longrightarrow$ homoclinic leaf Suppose first that λ is strongly commensurable to a Hausdorff limit of meridians λ'. Then by [42, Theorem B.1], there is a homoclinic leaf $h \subset \lambda'$, so λ is strongly commensurable to a lamination with a homoclinic leaf as desired.

Homoclinic leaf $\Longrightarrow$ (1), (2) or (3) We now assume we have a homoclinic leaf h in some lamination strongly commensurable to λ.

The two ends of h limit onto (possibly equal) components $\lambda_\pm \subset \lambda$. If one of $S(\lambda_\pm)$ has compressible boundary, there is a meridian disjoint from λ, so we are in case (1) and are done. So, $\partial S(\lambda_\pm)$ is incompressible, and we're in the situation of Theorem 6.1 and Corollary 6.2. We now break into cases.

If one of $\lambda_\pm$ is an intrinsic limit of meridians, we're in case (2) and are done. If we're in case (3) of Theorem 6.1 and Corollary 6.2, we're in case (3) of the theorem and are done, unless the given interval bundle $B \to Y$ is over an annulus or Möbius band. But in that case, letting c be a boundary component of Y, we can consider the geodesic meridian μ on S homotopic to the $m(c)$ constructed in Claim 2.11, using the compressing arc given by Corollary 6.2. This μ is disjoint from λ, so we're in case (1) of the theorem.

Finally suppose that the two ends of h are asymptotic on S, so that $\lambda_- = \lambda_+$. Let's separate further into the cases (i) and (ii) in Corollary 6.2. In case (i), using the notation of the corollary, the curve $c \cup h([-s, s])$ is a meridian disjoint from λ. So, we're in case (1) of the theorem. In case (ii), let T be a neighbourhood of $h \cup \lambda_\pm$ that is either a punctured torus or a pair of pants, depending on whether the two ends of h limit onto opposite sides of $\lambda_\pm$, or onto the same side. Because we're in case (ii), there is a meridian in T. Hence, whether T is a pair of pants or a punctured torus, one of the boundary components of T is a meridian, and is disjoint from λ so we're done.

(1), (2) or (3) $\Longrightarrow$ Hausdorff limit Suppose (1), (2) or (3) holds. We want to show λ is strongly commensurable to a Hausdorff limit of meridians. If λ is disjoint from a meridian, then we're done by Lemma 7.3. If a component of λ is an intrinsic limit of meridians, we're done by the first part of Corollary 7.5. In case (3) above, we're done by the second part of Corollary 7.5. □

8 Extending partial pseudo-Anosovs to compression bodies

Let M be a compression body with exterior boundary Σ. Let $S \subset \Sigma$ be an essential subsurface such that ∂S is incompressible. In this section, we prove:

Theorem 8.1 (extending partial pseudo-Anosovs) *Suppose that $f : \Sigma \to \Sigma$ is a partial pseudo-Anosov supported on S. Then f has a power that extends to a nontrivial subcompression body of (M, S) if and only if the attracting lamination of f is a projective limit of meridians that lie in S.*

When $S = \Sigma$, this is a theorem of Biringer, Johnson and Minsky [2]. The proof of Theorem 8.1 is basically the same as their proof, but we need to go through it anyway, to note the places that parabolics appear, and to deal with the fact that we are looking at subcompression bodies of (M, S) rather than of M. Also, before starting on the bulk of the proof in Section 8.2, we isolate part of the argument into a separate purely topological subsection, Section 8.1. This separation of the argument into distinct topological and geometric parts makes it more understandable than the original version, we think.

8.1 Dynamics on the space of marked compression bodies

Let Σ be a closed, orientable surface, and let $S \subset \Sigma$ be an essential subsurface. The *space of marked S-compression bodies* is defined to be

$$\mathrm{CBod}(S) = \{(C, h : \Sigma \to \partial_+ C)\}/{\sim},$$

where here C is a compression body, h is a homeomorphism, and

- the multicurve $h(\partial S) \subset \partial_+ C$ is incompressible,
- there is a multicurve m on S such that $h(m)$ is a cut system for C, i.e., $h(m)$ bounds a collection of disks that cut C into balls and trivial interval bundles over the interior boundary components.

We declare $(C_i, h_i : S \to \partial_+ C_i)$, $i = 1, 2$, to be equivalent (written $\sim$ as above) if there is a homeomorphism $\phi : C_1 \to C_2$ that respects the boundary markings: that is, $\phi \circ h_1$ and h_2 are homotopic maps $S \to \partial_+ C_2$.

We write $(C_1, h_1) \subset (C_2, h_2)$ if there is an embedding $\phi : (C_1, \partial_+ C_1) \hookrightarrow (C_2, \partial_+ C_2)$ that respects the boundary markings. This gives a partial ordering on $\mathrm{CBod}(S)$. We often identify Σ with $\partial_+ C$ instead of specifying the boundary marking, and simply write C for an element of $\mathrm{CBod}(S)$. So $\mathrm{CBod}(S)$ is the set of all compression bodies with exterior boundary Σ that one obtains by compressing curves in S (without compressing boundary curves) up to the obvious equivalence.

A marked S-compression body (C, h) has a *disk set* $\mathcal{D}(C) \subset \mathcal{C}(S)$, where a simple closed curve $\gamma \in \mathcal{C}(S)$ lies in the disc set if $h(\gamma)$ is a meridian in C. In fact, the disk set $\mathcal{D}(C)$ determines (C, h) up to equivalence, say by an argument similar to the last paragraph of the proof of Fact 2.4. The set $\mathrm{CBod}(S)$ can then be identified with the "set of all disk sets" in $\mathcal{C}(S)$. It then inherits a topology as a subset of the power set $\mathcal{P}(\mathcal{C}(S))$, wherein $D_n \to D$ if and only if for every $c \in \mathcal{C}(S)$, we have either $c \in D$ and $c \in D_n$ for all large n, or $c \notin D$ and $c \notin D_n$ for all large n.

Lemma 8.2 *If $C_n \to C$ in $\mathrm{CBod}(S)$, then $C \subset C_n$ for large n.*

Proof Suppose that C is obtained by compressing a finite set Γ of disjoint simple closed curves on S. For large n, we have $\Gamma \subset \mathcal{D}(C_n)$, so $\mathcal{C} \subset C_n$. □

Lemma 8.3 *$\mathrm{CBod}(S)$ is compact.*

Proof As $\mathcal{P}(\mathcal{C}(S))$ is compact, we want to show that $\mathrm{CBod}(S)$ is closed. Suppose C_n is a sequence of marked compression bodies with disk sets

$$D_n = \mathcal{D}(C_n) \subset \mathcal{C}(S),$$

and that $D_n \to D \subset \mathcal{C}(S)$. Let Γ be a maximal set of disjoint, pairwise nonhomotopic elements of D. Compressing Γ yields a marked compression body C. Since Γ is finite, $\Gamma \subset D_n$ for large n, so $\mathcal{D}(C) \subset D_n$. Thus, $\mathcal{D}(C) \subset D$.

It therefore suffices to show $D \subset \mathcal{D}(C)$. Suppose this is not the case, and pick $\beta \in D \setminus \mathcal{D}(C)$ such that the intersection number of β and Γ is minimal. By maximality of Γ, this intersection number cannot be zero. Since $\beta \in D$, if n is large we have $\beta \in \mathcal{D}(C_n)$. By an outermost disk argument, if $\gamma \in \Gamma$ is a component that intersects β, there is an arc $c \subset \gamma$ with endpoints on β and interior disjoint from β, that is homotopic rel endpoints in C_n to the two arcs $b_1, b_2 \subset \beta$ into which β is cut by ∂c. Passing to a subsequence, we can assume that c, b_1, b_2 are independent of n. Then $c \cup b_1$ and $c \cup b_2$ are both meridians in C_n for all large n, and hence lie in D. Since they intersect Γ fewer times than β does, they lie in $\mathcal{D}(C)$. But then β (which is a band sum of the two curves) also lies in $\mathcal{D}(C)$, a contradiction. □

Let $f : \Sigma \to \Sigma$ be a homeomorphism with $f = \mathrm{id}$ on $\Sigma \setminus S$. Then f acts on the space $\mathrm{CBod}(S)$ by $f \cdot (C, h) = (C, h \circ f^{-1})$. When we regard marked S-compression bodies as compression bodies with exterior boundary *equal* to Σ, we'll just write C and $f(C)$ for a marked compression body and its image. Note that $f(C) = C$ if and only if f extends to a homeomorphism of C.

Fixing $M \in \mathrm{CBod}(S)$ and f as above, let $\mathcal{A}$ be the set of accumulation points in $\mathrm{CBod}(S)$ of the f-orbit of M, and let

$$\mathcal{A}_{\min} = \{C \in \mathcal{A} \mid \nexists D \in \mathcal{A} \text{ such that } D \subsetneq C\}$$

be the subset consisting of all minimal elements of $\mathcal{A}$.

Theorem 8.4 (existence of maximal subcompression body) *The set $\mathcal{A}_{\min}$ is a finite f-orbit that contains a single element C_f such that $C_f \subset M$.*

Moreover, C_f is the unique maximal element of $\mathrm{CBod}(S)$ *such that $C_f \subset M$ and a power of f extends to C_f.*

This result was proved in [2] when $S = \Sigma$. Our proof follows the same general lines, but is topological instead of hyperbolic geometric.

We proceed with a series of lemmas.

Lemma 8.5 *The set $\mathcal{A}_{\min}$ is nonempty, finite and f-invariant.*

Proof The set $\mathcal{A}$ is nonempty, since $\mathrm{CBod}(S)$ is compact. This implies that $\mathcal{A}_{\min}$ is nonempty, for example, since the "height" of a compression body is nonnegative and decreases under strict containment; see Section 3 of [3].

By Lemma 8.2, $\mathcal{A}_{\min}$ is discrete. But $\mathcal{A}_{\min}$ is closed in $\mathcal{A}$, which is closed in $\mathrm{CBod}(S)$, which is compact. So, $\mathcal{A}_{\min}$ is compact, and must be finite. Finally, $\mathcal{A}_{\min}$ is f-invariant since $\mathcal{A}$ is and the f-action respects containment. □

Lemma 8.6 *Suppose that for $i = 1, 2$ we have $C_i \in$* $\mathrm{CBod}(S)$ *with $C_i \subset M$, and that $f^i(C_1) = C_1$ while $f^j(C_2) = C_2$. Then there is an element $C \in \mathcal{A}_{\min}$ such that $C_1, C_2 \subset C \subset M$.*

Proof Every element of $\mathcal{A}$ is the image under a power of f of an accumulation point of the sequence $f^{nij}(M)$, so since $\mathcal{A}_{\min}$ is f-invariant there is some $C' \in \mathcal{A}_{\min}$ to which a subsequence of $f^{nij}(M)$ limits. As $C_1, C_2 \subset f^{nij}(M)$ for all n, we must have $C_1, C_2 \in C'$.

By Lemma 8.2, there is some n such that $f^{nij}(M) \supset C'$. Then $C := f^{-nij}(C') \in \mathcal{A}_{\min}$ is contained in M and must contain C_1, C_2 as well. □

Lemma 8.7 *There is a unique element $C_f \in \mathcal{A}_{\min}$ that is contained in M, and $\mathcal{A}_{\min}$ is an f-orbit.*

Proof Applying the previous lemma to two copies of the trivial compression body $\Sigma \times I$ shows that $\mathcal{A}_{\min}$ has an element that is contained in M.

So, suppose that $C, D \in \mathcal{A}_{\min}$ are both contained in M. By the previous lemma, there is another element of $\mathcal{A}_{\min}$ that contains them both, which contradicts the minimality assumption unless $C = D$. Therefore, there is a unique element $C_f \in \mathcal{A}_{\min}$ that is contained in M.

To show that $\mathcal{A}_{\min}$ is an f-orbit, suppose that $C \in \mathcal{A}_{\min}$. Since C is an accumulation point, there is some n such that $f^n(M) \supset C$. Then $f^{-n}(C) \subset M$, implying that $f^{-n}(C) = C_f$ by uniqueness. □

This finishes the proof of Theorem 8.4, since Lemma 8.5 shows that a power of f extends to C_f and Lemmas 8.6 and 8.7 imply that any subcompression body of M to which a power of f extends is contained in C_f.

8.2 The proof of Theorem 8.1

Let $S \subset \Sigma = \partial M_+$ be a compact essential subsurface, with ∂S incompressible in M, and let $f : \Sigma \to \Sigma$ be a pseudo-Anosov map on S.

The "only if" direction of the theorem is trivial. Namely, suppose that some power f^k extends to a nontrivial subcompression body C of (M, S). Pick a meridian $m \subset S$ for C. Then $(f^k(m))$ is a sequence of meridians in M that lie in S, and converge to the attracting lamination of f.

For the "if" direction of the theorem, assume that no nonzero power of f extends to a nontrivial subcompression body of (M, S). We must show that the attracting lamination λ^+ is not in the limit set $\Lambda(S, M)$. The argument is similar to the proof of the main theorem in [2]. As such, we will sketch the argument in places and refer to [2] for details.

Consider the sequence $M_n = f^{-n}(M)$ of marked S-compression bodies, where we consider the exterior boundary of each M_n as identified with the surface Σ. Fix a base point $[X] \in \mathcal{T}(\Sigma)$ and give the interior of each M_n a geometrically finite hyperbolic metric such that the end adjacent to the exterior boundary $\Sigma = \partial_+ M$ is convex cocompact, and when its conformal boundary is identified with Σ, the conformal structure is $[X]$. Let

$$\rho_n : \pi_1 \Sigma \to \mathrm{PSL}_2\,\mathbb{C}, \quad N_n := \mathbb{H}^3/\rho_n(\pi_1 \Sigma),$$

be a representation (unique up to conjugacy) uniformising the interior of M_n and compatible with our markings, in the sense that ρ_n is the composition of the map $\pi_1 \Sigma \to \pi_1 M_n \cong \pi_1 N_n$ induced by inclusion and a faithful uniformising representation of $\pi_1 N_n$. Note that the kernel of ρ_n is

$$\ker(\rho_n) = f_*^{-n}(\ker(\pi_1 \Sigma \to \pi_1 M)).$$

By Theorem 8.4 and the assumption that no power of f extends to a nontrivial subcompression body of M, the only minimal accumulation point of (M_n) in $\mathrm{CBod}(S)$ is the trivial compression body. So in particular, we can choose a subsequence (M_{n_j}) that converges to the trivial compression body. By the compactness of generalised Bers slices (see [2, Theorem 4.3]), we may assume after appropriate conjugations and passing to a further subsequence that (ρ_{n_j}) converges algebraically to a representation

$$\rho_\infty : \pi_1 \Sigma \to \mathrm{PSL}_2\,\mathbb{C}, \quad N_\infty := \mathbb{H}^3/\rho_\infty(\pi_1\Sigma),$$

and that N_∞ can be identified with the interior of a compression body M_∞ with exterior boundary Σ in such a way that the end of N_∞ adjacent to Σ is convex cocompact with conformal structure $[X]$ and the representation ρ_∞ is compatible with the marking in the same way as before.

The disk set $\mathcal{D}(S, M_\infty)$ consists of all simple closed curves on S represented by elements $\gamma \in \pi_1\Sigma$ with $\rho_\infty(\gamma) = 1$. By Chuckrow's theorem (see [2, Lemma 2.11]), $\rho_\infty(\gamma) = 1$ if and only if $\rho_{n_j}(\gamma) = 1$ for all sufficiently large i. Since (M_{n_j}) converges to the trivial compression body in the topology of $\mathrm{CBod}(S)$, it follows that the surface $S \subset \Sigma = \partial_+ M_\infty$ is incompressible in M_∞.

Claim 8.8 *The repelling lamination λ^- of f is unrealisable in N_∞.*

Proof The proof is almost identical to that of [2, Lemma 6.2], so we offer a sketch and we refer the reader to their paper for details.

Fixing an M-meridian $\gamma \subset S$, the sequence $f^{-n_j}(\gamma)$ converges in the Hausdorff topology to a lamination λ_M that is the union of λ^- and finitely many leaves spiralling onto it. It suffices by [7, Theorem 2.3] to show that λ_M is unrealisable. So, hoping for a contradiction, assume λ_M *is* realisable; then λ_M is carried by a train track τ that maps nearly straightly into N_∞ (see [2]).

By algebraic convergence, τ also maps nearly straightly into N_{n_j} when j is large. Since $f^{-n_j}(\gamma) \to \lambda_M$, for large j the curve $f^{-n_j}(\gamma)$ is carried by τ. This implies that $f^{-n_j}(\gamma)$ is geodesically realisable in N_{n_j} for large j, contradicting the fact that it is homotopically trivial. □

By work of Thurston [58, Proposition 9.7.1], the π_1-injective surface $S \subset \partial_+ M_\infty$ is isotopic into a degenerate end of N_∞ with ending lamination λ^-. In particular, the peripheral curves of S represent cusps in N_∞ and every nonperipheral curve on S has hyperbolic type in N_∞. Any pair of disjoint nonperipheral simple closed curves on S can then be realised geodesically by a pleated surfaces $S \to N_\infty$ in the given homotopy class, and Thurston's compactness of pleated surfaces (see [50, Lemma 6.13]) implies the following.

Lemma 8.9 (compare with [2, Lemma 6.3]) *Let $\alpha \subset S$ be a simple closed curve. Then for every k, there is some K such that for any other simple closed curve β in S, we have*

$$d_{\mathcal{C}(S)}(\alpha, \beta) \le k \implies d_{N_\infty}(\alpha_\infty, \beta_\infty) \le K,$$

where α and β_∞ are the geodesics in N_∞ in the homotopy classes of α and β.

Hoping for a contradiction, suppose now that $\lambda^+ \in \Lambda(S, M)$. When regarded as an element of $\partial_\infty \mathcal{C}(S)$, the support of λ^+ is then an accumulation point of $\mathcal{D}(S, M) \subset \mathcal{C}(S)$. If $\alpha \in \mathcal{C}(S)$, then for $n = 1, 2, \ldots$

the sequence $(f^n(\alpha))$ is a quasigeodesic path that limits to $\lambda^+ \in \partial_\infty \mathcal{C}(S)$; see [47]. Since $\mathcal{D}(S, M)$ is a quasiconvex subset (see Theorem 2.3, due to Masur and Minsky), there is a constant C and for each n a meridian $\gamma_n \in \mathcal{D}(S, M)$ with

$$d_{\mathcal{C}(S)}(f^n(\alpha), \gamma_n) \leq C.$$

Translating the points $f^n(\alpha)$ and γ_n by f^{-n}, this becomes

$$d_{\mathcal{C}(S)}(\alpha, f^{-n}(\mathcal{D}(S, M)) \leq C.$$

By Lemma 8.9, an element $\gamma_{n_j} \in f^{-n_j}(\mathcal{D}(S, M))$ at distance at most C from α in $\mathcal{C}(S)$ can be geodesically realised in some fixed compact subset $A \subset N_\infty$. Algebraic convergence implies that for sufficiently large j this geodesic can be pulled back and tightened to a geodesic in N_{n_j}. But by construction, γ_{n_j} is a meridian in M_{n_j}, so it cannot possibly be realised geodesically in N_{n_j}, which is a contradiction.

9 Extending reducible maps to compression bodies

We present here a generalisation of [2, Theorem 1.1] that characterises which (possibly reducible) mapping classes of the boundary of a 3-manifold M have powers that extend to subcompression bodies.

In what follows, let M be a compression body with exterior boundary $\partial_+ M$. Let $S \subset \partial_+ M$ be an essential subsurface such that ∂S is incompressible. Let $f : \partial_+ M \to \partial_+ M$ be a homeomorphism that is "supported" in S, meaning that $f = \mathrm{id}$ on $\partial_+ M \setminus S$.

Definition 9.1 We say f is *pure* if there are disjoint, compact, essential f-invariant subsurfaces $S_i \subset S$, $i = 1, \dots, n$, such that $f = \mathrm{id}$ on $S_{\mathrm{id}} := S \setminus \bigcup_i S_i$, and where for each i, if we set $f_i := f|_{S_i}$, then either

(1) S_i is an annulus and f_i is a power of a Dehn twist, or

(2) f_i is a pseudo-Anosov map on S_i.

It follows from the Nielsen–Thurston classification, see [21], that every f has a power that is isotopic to a pure homeomorphism.

When f is pure, with associated restrictions $f_i : S_i \to S_i$ as above, we define a geodesic lamination $\lambda = \bigcup_i \lambda_i$ on S, where $\lambda_i \subset S_i$ as follows. If f_i is pseudo-Anosov, we let λ_i be the support of the attracting lamination of f_i. If f_i is a Dehn twist, we let λ_i be the core curve of the annulus S_i. So defined, λ is called the *attracting lamination* of the pure homeomorphism f.

Theorem 9.2 *Suppose that $S \subset \partial_+ M$ is an essential subsurface such that the multicurve ∂S is incompressible. Let $f : \partial_+ M \to \partial_+ M$ be a pure homeomorphism supported in S. Then f has a power that extends to a nontrivial subcompression body of (M, S) if and only if either*

(1) *there is a meridian in S_{id},*

(2) *for some i, the map f_i has a power that extends to a nontrivial subcompression body of (M, S_i), or*

(3) *there are (possibly equal) indices i, j such that S_i, S_j bound an essential interval bundle B in M, such that some power of $f|_{S_i \cup S_j}$ extends to B, and there is a compression arc α for B whose interior lies in S_{id}.*

Note that by Theorem 8.1, if f_i is pseudo-Anosov then (2) above is equivalent to the condition that λ_i is an intrinsic limit of meridians.

Recall from Section 2.7 that a "subcompression body of (M, S)" is a compression body obtained from $\partial_+ M$ by compressing some meridian multicurve in S. In (3), the condition that a power of $f|_{S_i \cup S_j}$ extends to B is easier to check. Indeed, if $\sigma : \partial_H B \to \partial_H B$ is the canonical involution, as defined in Section 2.5, then by Fact 2.7 we have that $f^k|_{\partial_H B}$ extends to B exactly when $\sigma \circ f_i^k$ is isotopic to f_j^k. When B is a twisted interval bundle, f_i, f_j are both pseudo-Anosovs on $\partial_H B$ and this means that as mapping classes we have $f_j = g \circ f_i$ for some finite order g commuting with both f_i, f_j; see, e.g., McCarthy's thesis [51]. When B is a trivial interval bundle, σ identifies S_i and S_j, and we have similarly that $f_j = g \circ \sigma(f_i)$ for some g commuting with both.

Proof of Theorem 9.2 Let's start with the "if" direction, since that's easier. If there is a meridian in S_{id}, then f extends to the compression body obtained by compressing that meridian. Suppose (2) holds, so that some power f_i^k extends to a nontrivial subcompression body C of (M, S_i). Then f also extends to C, since all the S_j, where $j \neq i$, bound trivial interval bundles with subsurfaces of the interior boundary of C. So we're done.

The only interesting case is if (3) holds, so that some S_i, S_j bound an essential interval bundle B in M such that some power of $f^k|_{S_i \cup S_j}$ extends to B, and there is a compression arc α for B whose interior lies in S_{id}. Here, let $S' \subset S$ be the smallest essential subsurface containing S_i, S_j and α; so, S' is obtained from a regular neighbourhood of the union of these three subsets of S by capping off any inessential boundary components with discs. Let C be the characteristic compression body of the pair (M, S'), as defined in Fact 2.4.

We claim that f^k extends to C. To see this, note that we can construct C as follows. For concreteness, first assume that the boundary components of S_i, S_j that contain the endpoints of α bound an annulus $A \supset \alpha$ on S. Then $S' = S_i \cup S_j \cup A$, the annulus A is parallel in M to component $A' \subset \mathrm{Fr}(B)$ that is an annulus with the same boundary curves as A, and C is the union of the interval bundle B, the solid torus bounded by A, A', and a trivial interval bundle over $\partial_+ M \setminus S'$. We can then extend f^k to C by letting it be the given extension of $f^k|_{S_i \cup S_j}$ on B, the identity on the solid torus, and the obvious fibre preserving extension of $f^k|_{\partial_+ M \setminus S'}$ to the adjacent interval bundle. The case that the boundary components of S_i, S_j that contain the endpoints of α do not bound an annulus on S is similar, except that instead of the solid torus above we take a thickened disk bounded by a rectangular neighbourhood of α on $S \setminus (S_i \cup S_j)$, and a rectangular neighbourhood of the homotopic arc on the frontier of B.

We now work on the "only if" direction. Passing to a power, suppose that f extends to a nontrivial subcompression body C of (M, S). We may assume that there is no proper, f-invariant essential subsurface $S' \subset S$ such that $f|_{S'}$ extends to a nontrivial subcompression body of (M, S'). If there were such a subsurface S', we could replace S by a minimal such S', therefore reducing the argument to the minimal case we are assuming we are in above.

If $f = \mathrm{id}$ on S, the fact that there is a nontrivial subcompression body of (M, S) means there is a meridian in $S = S_{\mathrm{id}}$, so we're in case (1) and are done. This case may seems silly, but observe that if f is some complicated pure homeomorphism where there's a meridian in S_{id}, the associated "minimal" case that we pass to in the previous paragraph is where S is an annular neighbourhood of some such meridian, and $f = \mathrm{id}$ on S.

Assume from now on that f is not the identity map of S.

We claim that every meridian $m \in \mathcal{D}(C, S)$ intersects every component of λ. Indeed, suppose some λ_i is disjoint from some such m and let S' be the component of $S \setminus S_i$ containing m. Since f extends to C and $S' \subset \partial C_+$ is f^k-invariant, f extends to the characteristic subcompression body C' of the pair (C, S'), defined by starting with $\partial_+ M$ and compressing all meridians of C that lie in S'; see Fact 2.4. Since m is a meridian in C', this C' is nontrivial, which contradicts the minimality assumption in the first paragraph.

Pick a meridian $m \in \mathcal{D}(C, S)$. Since m intersects all components of λ, the sequence of meridians $m_i := f^i(m)$ Hausdorff converges to a lamination λ' strongly commensurable to λ. Applying Theorem 7.2 to the pair (C, S) and using that all meridians in C intersect all components of λ, we have that either

- some component λ_a is an intrinsic limit of meridians lying in $S(\lambda_a)$, in which case Theorem 8.1 (applied to $f_a : S_a \to S_a$) says we're in case (2), or
- there are indices a, b such that S_a, S_b bound an essential interval bundle $B \subset C$, where λ_a, λ_b are essentially homotopic in B, and where there is a compression arc $\alpha \subset S$ for B that is disjoint from λ, and hence can be isotoped so that its interior lies in S_{id}.

Let's assume we're in the last case, since otherwise we're done. We want to show that some power of $f_a \cup f_b : \partial_H B \to \partial_H B$ extends to B.

First, suppose that B is a twisted interval bundle, so that $S_a = S_b$, $f_a = f_b$, $\lambda_a = \lambda_b$. Using just the index a from now on, if σ is the canonical involution of B, then Fact 2.16 implies that $\sigma(\lambda_a)$ is isotopic to λ_a. Let $A \subset \mathcal{T}(S)$ be the axis of f_a on the Teichmüller space $\mathcal{T}(S)$. By Theorem 12.1 of [22] and Theorem 2 of [45], we have that A, $\sigma(A)$ are asymptotic, so as they are both pseudo-Anosov axes they must be equal by discreteness of the action of the mapping class group. Since σ has finite order, it then fixes A pointwise. Now the group $\Gamma = \langle \sigma, f_a \rangle \subset \mathrm{Mod}(\partial_H B)$ is isomorphic to the direct product of a finite group fixing A pointwise and a cyclic group of pseudo-Anosovs, so for some positive k we have $\sigma \circ f_a^k = f_a^k$ in $\mathrm{Mod}(\partial_H B)$. By Theorem 3 of [4] we may isotope f_a^k, σ so that they commute, while preserving that $\sigma^2 = \mathrm{id}$; we can then alter the bundle map $\pi : B \to Y$ so that the new σ is still the canonical involution. It follows that f_a^k is a lift to $\partial_H B$ of a pseudo-Anosov map $g : Y \to Y$, and hence f_a^k extends to B as desired, see Fact 2.7.

Next, assume B is a trivial interval bundle, with canonical involution σ that switches S_a, S_b. As in the previous paragraph, we have that $\sigma(\lambda_b)$ is isotopic to λ_a, so $\Gamma = \langle f_a, \sigma \circ f_b \circ \sigma^{-1} \rangle \subset \mathrm{Mod}(S_a)$ is a direct product $\Gamma = F \times \langle \phi \rangle$ of a finite group F and a cyclic group generated by a pseudo-Anosov ϕ, where if we quotient by F then f_a and $\sigma \circ f_b \circ \sigma^{-1}$ both project to positive powers of ϕ. It suffices to show that they project to the *same* positive power of ϕ, for then we are done by the same argument as in the previous paragraph. For this, recall that all meridians of C intersect S_a, S_b, so these surfaces are "holes" for the disk set of C, as discussed in [49]. So with $m_i = f^i(m)$ the sequence of meridians in C constructed above, Lemma 12.20 of [49] says that for each i, the distance in the arc complex of S_a between

$$m_i \cap S_a = f_a^i(m \cap S_a) \quad \text{and} \quad \sigma(m_i \cap S_b) = (\sigma \circ f_b \circ \sigma^{-1})^i \sigma(m \cap S_b)$$

is at most 6. However, if f_a and $\sigma \circ f_b \circ \sigma^{-1}$ project to different positive powers of ϕ, their stable translation lengths on the arc complex of S_a are different, which is a contradiction. □

Acknowledgements

The authors would like to thank Jeff Brock, Juan Souto and Sebastian Hensel for useful conversations, and also the referee, who helped improve the accuracy and readability of the paper. Biringer was partially supported by NSF grant DMS-1308678.

References

[1] **I Agol**, *Tameness of hyperbolic* 3*-manifolds* (2004) arXiv math/0405568

[2] **I Biringer**, **J Johnson**, **Y Minsky**, *Extending pseudo-Anosov maps into compression bodies*, J. Topol. 6:4 (2013) 1019–1042 MR

[3] **I Biringer**, **N G Vlamis**, *Automorphisms of the compression body graph*, J. Lond. Math. Soc., II. Ser. 95:1 (2017) 94–114

[4] **J S Birman**, **H M Hilden**, *On isotopies of homeomorphisms of Riemann surfaces*, Ann. of Math. (2) 97 (1973) 424–439 MR

[5] **F Bonahon**, *Cobordism of automorphisms of surfaces*, Ann. Sci. École Norm. Sup. (4) 16:2 (1983) 237–270 MR

[6] **B H Bowditch**, *Relatively hyperbolic groups*, Internat. J. Algebra Comput. 22:3 (2012) art. id. 1250016 MR

[7] **J F Brock**, *Continuity of Thurston's length function*, Geom. Funct. Anal. 10:4 (2000) 741–797 MR

[8] **J F Brock**, *Iteration of mapping classes and limits of hyperbolic* 3*-manifolds*, Invent. Math. 143:3 (2001) 523–570 MR

[9] **J F Brock**, **R D Canary**, **Y N Minsky**, *The classification of Kleinian surface groups, II: The ending lamination conjecture*, Ann. of Math. (2) 176:1 (2012) 1–149 MR

[10] **J Brock**, **Y Minsky**, **H Namazi**, **J Souto**, *Bounded combinatorics and uniform models for hyperbolic* 3*-manifolds*, J. Topol. 9:2 (2016) 451–501 MR

[11] **D Calegari**, **D Gabai**, *Shrinkwrapping and the taming of hyperbolic* 3*-manifolds*, J. Amer. Math. Soc. 19:2 (2006) 385–446 MR

[12] **R D Canary**, *Algebraic convergence of Schottky groups*, Trans. Amer. Math. Soc. 337:1 (1993) 235–258 MR

[13] **R D Canary**, *Ends of hyperbolic* 3*-manifolds*, J. Amer. Math. Soc. 6:1 (1993) 1–35 MR

[14] **R D Canary**, *A covering theorem for hyperbolic* 3*-manifolds and its applications*, Topology 35:3 (1996) 751–778 MR

[15] **J W Cannon**, **C D Feustel**, *Essential embeddings of annuli and Möbius bands in* 3*-manifolds*, Trans. Amer. Math. Soc. 215 (1976) 219–239 MR

[16] **A J Casson**, **S A Bleiler**, *Automorphisms of surfaces after Nielsen and Thurston*, London Mathematical Society Student Texts 9, Cambridge Univ. Press (1988) MR

[17] **A J Casson**, **C M Gordon**, *Reducing Heegaard splittings*, Topology Appl. 27:3 (1987) 275–283 MR

[18] **A J Casson**, **D D Long**, *Algorithmic compression of surface automorphisms*, Invent. Math. 81:2 (1985) 295–303 MR

[19] **M T Clay**, **C J Leininger**, **J Mangahas**, *The geometry of right-angled Artin subgroups of mapping class groups*, Groups Geom. Dyn. 6:2 (2012) 249–278 MR

[20] **V Erlandsson**, **M Gendulphe**, **I Pasquinelli**, **J Souto**, *Mapping class group orbit closures for non-orientable surfaces*, Geom. Funct. Anal. 33:3 (2023) 637–693 MR

[21] **B Farb**, **D Margalit**, *A primer on mapping class groups*, Princeton Mathematical Series 49, Princeton Univ. Press (2012) MR

[22] **A Fathi**, **F Laudenbach**, **V Poenaru**, *Travaux de Thurston sur les surfaces*: *séminaire Orsay*, Astérisque 66-67, Soc. Math. France, Paris (1979) MR

[23] **V Gadre**, *The limit set of the handlebody set has measure zero* (2011) arXiv 1105.2980

[24] **U Hamenstädt**, *Train tracks and the Gromov boundary of the complex of curves*, from "Spaces of Kleinian groups" (Y N Minsky, M Sakuma, C Series, editors), London Math. Soc. Lecture Note Ser. 329, Cambridge Univ. Press (2006) 187–207 MR

[25] **U Hamenstädt**, *Hyperbolic relatively hyperbolic graphs and disk graphs*, Groups Geom. Dyn. 10:1 (2016) 365–405 MR

[26] **U Hamenstädt**, *Asymptotic dimension and the disk graph, I*, J. Topol. 12:2 (2019) 658–673 MR

[27] **U Hamenstädt**, *Asymptotic dimension and the disk graph, II*, J. Topol. 12:3 (2019) 675–684 MR

[28] **G C Hruska**, *Relative hyperbolicity and relative quasiconvexity for countable groups*, Algebr. Geom. Topol. 10:3 (2010) 1807–1856 MR

[29] **W Jaco**, *Lectures on three-manifold topology*, CBMS Regional Conference Series in Mathematics 43, Amer. Math. Soc., Providence, RI (1980) MR

[30] **W H Jaco**, **P B Shalen**, *Seifert fibered spaces in* 3*-manifolds*, Mem. Amer. Math. Soc. 220, Amer. Math. Soc., Providence, RI (1979) MR

[31] **W Jeon**, **I Kim**, **K Ohshika**, **C Lecuire**, *Primitive stable representations of free Kleinian groups*, Israel J. Math. 199:2 (2014) 841–866 MR

[32] **K Johannson**, *Homotopy equivalences of* 3*-manifolds with boundaries*, Lecture Notes in Mathematics 761, Springer (1979) MR

[33] **M Kapovich**, *Hyperbolic manifolds and discrete groups*, Progress in Mathematics 183, Birkhäuser, Boston, MA (2001) MR

[34] **S P Kerckhoff**, *The measure of the limit set of the handlebody group*, Topology 29:1 (1990) 27–40 MR

[35] **I Kim**, **C Lecuire**, **K Ohshika**, *Convergence of freely decomposable Kleinian groups*, Invent. Math. 204:1 (2016) 83–131 MR

[36] **I Kim**, **M Lee**, *Separable-stable representations of a compression body*, Topology Appl. 206 (2016) 171–184 MR

[37] **E Klarreich**, *The boundary at infinity of the curve complex and the relative Teichmüller space*, Groups Geom. Dyn. 16:2 (2022) 705–723 MR

[38] **G Kleineidam**, **J Souto**, *Algebraic convergence of function groups*, Comment. Math. Helv. 77:2 (2002) 244–269 MR

[39] **M Lackenby**, *Attaching handlebodies to* 3*-manifolds*, Geom. Topol. 6 (2002) 889–904 MR

[40] **C Lecuire**, *Structures hyperboliques convexes sur les variétés de dimension* 3, PhD thesis, École normale supérieure de Lyon (2004) Available at https://theses.fr/2004ENSL0273

[41] **C Lecuire**, *An extension of the Masur domain*, from "Spaces of Kleinian groups" (Y N Minsky, M Sakuma, C Series, editors), London Math. Soc. Lecture Note Ser. 329, Cambridge Univ. Press (2006) 49–73 MR

[42] **C Lecuire**, *Plissage des variétés hyperboliques de dimension* 3, Invent. Math. 164:1 (2006) 85–141 MR

[43] **D D Long**, *Discs in compression bodies*, Pacific J. Math. 122:1 (1986) 129–146 MR

[44] **J F Manning**, *The Bowditch boundary of* $(G, \mathcal{H})$ *when* G *is hyperbolic* (2015) arXiv 1504.03630

[45] **H Masur**, *Uniquely ergodic quadratic differentials*, Comment. Math. Helv. 55:2 (1980) 255–266 MR

[46] **H Masur**, *Measured foliations and handlebodies*, Ergodic Theory Dynam. Systems 6:1 (1986) 99–116 MR

[47] **H A Masur**, **Y N Minsky**, *Geometry of the complex of curves, I*: *Hyperbolicity*, Invent. Math. 138:1 (1999) 103–149 MR

[48] **H A Masur**, **Y N Minsky**, *Quasiconvexity in the curve complex*, from "In the tradition of Ahlfors and Bers, III" (W Abikoff, A Haas, editors), Contemp. Math. 355, Amer. Math. Soc., Providence, RI (2004) 309–320 MR

[49] **H Masur**, **S Schleimer**, *The geometry of the disk complex*, J. Amer. Math. Soc. 26:1 (2013) 1–62 MR

[50] **K Matsuzaki**, **M Taniguchi**, *Hyperbolic manifolds and Kleinian groups*, Clarendon, New York (1998) MR

[51] **J D McCarthy**, *Normalizers and centralizers of pseudo-anosov mapping classes*, preprint (1982) Available at https://users.math.msu.edu/users/mccarthy/publications/normcent.pdf

[52] **Y Minsky**, *The classification of Kleinian surface groups, I: Models and bounds*, Ann. of Math. (2) 171:1 (2010) 1–107 MR

[53] **M Mj**, *Cannon–Thurston maps for Kleinian groups*, Forum Math. Pi 5 (2017) art. id. e1 MR

[54] **J W Morgan**, *On Thurston's uniformization theorem for three-dimensional manifolds*, from "The Smith conjecture" (New York, 1979) (J W Morgan, H Bass, editors), Pure Appl. Math. 112, Academic, Orlando, FL (1984) 37–125 MR

[55] **H Namazi**, **J Souto**, *Heegaard splittings and pseudo-Anosov maps*, Geom. Funct. Anal. 19:4 (2009) 1195–1228 MR

[56] **J-P Otal**, *Courants géodesique et surfaces*, PhD thesis, Université Paris-Sud, Orsay (1989) Available at https://theses.fr/1989PA112051

[57] **P Scott**, *Strong annulus and torus theorems and the enclosing property of characteristic submanifolds of* 3*-manifolds*, Quart. J. Math. Oxford Ser. (2) 35:140 (1984) 485–506 MR

[58] **W P Thurston**, *The geometry and topology of three-manifolds*, lecture notes, Princeton University (1979) Available at https://url.msp.org/gt3m

[59] **W P Thurston**, *Hyperbolic Structures on* 3*-manifolds, III: Deformations of* 3*-manifolds with incompressible boundary* (1998) arXiv math/9801058

[60] **G S Walsh**, *The bumping set and the characteristic submanifold*, Algebr. Geom. Topol. 14:1 (2014) 283–297 MR

IAN BIRINGER ian.biringer@bc.edu
Department of Mathematics, Boston College, Chestnut Hill, MA, United States

CYRIL LECUIRE cyril.lecuire@ens-lyon.fr
ENS de Lyon site Monod, UMPA UMR 5669 CNRS, Lyon, France

Received: June 6, 2024 Revised: February 16, 2025

msp

Algebraic & Geometric Topology **26**:5 (2026) 1867–1905
DOI: 10.2140/agt.2026.26.1867
Published: May 29, 2026

Motivic real topological Hochschild spectrum

DOOSUNG PARK

We define real topological Hochschild homology of separated log schemes with involutions. We show that real topological Hochschild homology is $(\mathbb{P}^n, \mathbb{P}^{n-1})$-invariant, which leads to the definition of the motivic real topological Hochschild spectrum living in a certain $\mathbb{Z}/2$-equivariant logarithmic motivic category. We explore properties of real topological Hochschild homology that can be deduced from the logarithmic motivic homotopy theory. We also define the motivic real topological cyclic spectrum.

1 Introduction

Topological Hochschild homology THH and its cousin TC have a deep connection with algebraic K-theory via the cyclotomic trace of Bökstedt, Hsiang, and Madsen [8]. The Dundas–Goodwillie–McCarthy theorem [11] provided a computational tool for algebraic K-theory. It was also discovered that THH contains arithmetic data. Bhatt, Morrow, and Scholze [3] studied filtrations on the S^1-homotopy fixed point spectrum TC^- of THH, whose graded pieces are closely related to the prismatic cohomology of Bhatt and Scholze according to [4, §13].

Hesselholt and Madsen [15] defined real algebraic K-theory, which refines both algebraic and hermitian K-theories in a uniform manner via equivariant homotopy theory. They also defined real topological Hochschild homology THR of rings. For further developments on THR, we refer to [10; 18; 27]. A forthcoming work of Harpaz, Nikolaus, and Shah will show a real refinement of the Dundas–Goodwillie–McCarthy theorem, which would justify the real trace method as a computational tool for real or hermitian K-theory.

One feature of THH different from algebraic K-theory is that THH is not $\mathbb{A}^1$-invariant even for regular schemes, i.e., the induced map $\mathrm{THH}(X) \to \mathrm{THH}(X \times \mathbb{A}^1)$ is not an equivalence of spectra for every nonempty scheme X. The definition of THH of schemes is due to Geisser and Hesselholt [13]. Hence the motivic methods in $\mathbb{A}^1$-homotopy theory initiated by Morel and Voevodsky [23] are not directly applicable to THH.

Non-$\mathbb{A}^1$-invariance of THH is related to the fact that the sequence of spectra

$$\mathrm{THH}(Z) \xrightarrow{i_*} \mathrm{THH}(X) \xrightarrow{j^*} \mathrm{THH}(X - Z)$$

is not a fiber sequence when $i : Z \to X$ is a closed immersion of regular schemes and $j : X - Z \to X$ is its open complement. On the other hand, the localization sequence in algebraic K-theory can be read as the fiber sequence of spectra

$$\mathrm{K}(Z) \xrightarrow{i_*} \mathrm{K}(X) \xrightarrow{j^*} \mathrm{K}(X - Z).$$

MSC2020: primary 19D55; secondary 11E70, 14A21, 16E40, 55P91.

Hesselholt and Madsen [16] and Rognes, Sagave, and Schlichtkrull [30] studied localization sequences for THH in the logarithmic setting.

The author's joint work with Binda and Østvær [7] introduced the ∞-category of logarithmic motivic spectra $\mathrm{logSH}(S)$ for fs log schemes S, which aims to incorporate various non-$\mathbb{A}^1$-invariant cohomology theories into the motivic framework using logarithmic geometry. We refer to Ogus's book [25] for logarithmic geometry. For a closed immersion of schemes $Z \to X$, let (X, Z) denote the log scheme with the underlying scheme X and with the compactifying log structure associated with the open immersion $X - Z \to X$. In the construction of $\mathrm{logSH}(S)$ when S is a scheme (with the trivial log structure), the interval $\mathbb{A}^1$ is replaced with the set of log schemes $(\mathbb{P}^n, \mathbb{P}^{n-1})$ for all integers $n \geq 1$, and the Nisnevich topology is replaced with the strict Nisnevich topology in [6]. To construct a motivic spectrum in $\mathrm{logSH}(S)$ that represents an existing cohomology theory of schemes, one can take the following two steps:

Step 1. Extend the cohomology theory to fs log schemes.

Step 2. Show that the extension is $(\mathbb{P}^n, \mathbb{P}^{n-1})$-invariant for all integers $n \geq 1$ and satisfies strict Nisnevich descent.

Rognes [29] defines THH of log rings. According to [7, §8], it is possible to define THH and TC of log schemes based on Rognes' definition and to construct a $\mathbb{P}^1$-spectrum **THH** and **TC** (written as **logTHH** and **logTC** in loc. cit.) representing THH and TC. The author's joint work with Binda, Lundemo, and Østvær [5, §8] constructed a $\mathbb{P}^1$-spectrum representing Hochschild homology.

In this article, we employ this strategy for THR. We define THR of log rings with involutions, where an involution means an automorphism σ such that $\sigma \circ \sigma = \mathrm{id}$. With the aid of the author's joint work with Hornbostel [19], we define THR of separated log schemes with involutions. Then we prove that THR is $(\mathbb{P}^n, \mathbb{P}^{n-1})$-invariant for all integers $n \geq 1$. This allows us to apply motivic methods in logarithmic motivic homotopy theory to THR. We also consider real topological cyclic homology, which is a variant of THR. One way to obtain TCR is to impose the real cyclotomic structure on THR, which is due to Quigley and Shah [27]. The following two results are examples of this.

Theorem 1.1 (special case of Theorem 6.8) *Let $Z \to X$ be a closed immersion of smooth schemes over a noetherian separated scheme S. Then there exist fiber sequences of $\mathbb{Z}/2$-spectra*

$$\mathrm{THR}(\mathrm{Th}(\mathrm{N}_Z X)) \to \mathrm{THR}(X) \to \mathrm{THR}(\mathrm{Bl}_Z X, E),$$
$$\mathrm{TCR}(\mathrm{Th}(\mathrm{N}_Z X)) \to \mathrm{TCR}(X) \to \mathrm{TCR}(\mathrm{Bl}_Z X, E),$$

where E is the exceptional divisor.

We refer to (6-1) for the notation $\mathrm{THR}(\mathrm{Th}(\mathrm{N}_Z X))$. Proposition 6.5 implies that $\mathrm{THR}(\mathrm{Th}(\mathrm{N}_Z X))$ can be written in terms of THR of schemes. This sequence can be considered as the *localization sequence for* THR. The map $\mathrm{THR}(\mathrm{Th}(\mathrm{N}_Z X)) \to \mathrm{THR}(X)$ is called the *Gysin map*.

Theorem 1.2 (Theorem 6.10) *Let $Z \to X$ be a closed immersion of smooth schemes over a finite-dimensional noetherian separated scheme S. Then the induced squares of $\mathbb{Z}/2$-spectra*

$$\begin{tikzcd}
\mathrm{THR}(X) \arrow[r] \arrow[d] & \mathrm{THR}(Z) \arrow[d] & \mathrm{TCR}(X) \arrow[r] \arrow[d] & \mathrm{TCR}(Z) \arrow[d] \\
\mathrm{THR}(Z \times_X \mathrm{Bl}_Z X) \arrow[r] & \mathrm{THR}(\mathrm{Bl}_Z X) & \mathrm{TCR}(Z \times_X \mathrm{Bl}_Z X) \arrow[r] & \mathrm{TCR}(\mathrm{Bl}_Z X)
\end{tikzcd}$$

are cartesian, where $\mathrm{Bl}_Z X$ denotes the blow-up of X along Z.

The statement in this theorem involves no log schemes, but the proof uses log schemes.

Hu, Kriz, and Ormsby [20] considered $\mathbb{P}^1 \wedge \mathbb{P}^\sigma$-spectra instead of $\mathbb{P}^1$-spectra to define a motivic $\mathbb{Z}/2$-spectrum, where $\mathbb{P}^\sigma$ is the scheme $\mathbb{P}^1$ with the involution given by $[x:y] \mapsto [y:x]$. In Definition 6.12, we analogously introduce the ∞-category of prelogarithmic motivic $\mathbb{Z}/2$-spectra $\mathrm{prelogSH}^{\mathbb{Z}/2}(S)$ for any scheme S, which is helpful for constructing logarithmic motivic spectra using the fixed point functor

$$(-)^{\mathbb{Z}/2} : \mathrm{prelogSH}^{\mathbb{Z}/2}(S) \to \mathrm{logSH}(S)$$

in Definition 6.14. We also construct the "forgetful" functor

$$i^* : \mathrm{prelogSH}^{\mathbb{Z}/2}(S) \to \mathrm{logSH}(S)$$

in Definition 6.16. The reason why we add "pre" in the notation $\mathrm{prelogSH}^{\mathbb{Z}/2}(S)$ is that a further localization is desired to obtain a better behaved ∞-category of logarithmic motivic $\mathbb{Z}/2$-spectra; see Remark 6.13 for an explanation.

Together with the $\mathbb{P}^1 \wedge \mathbb{P}^\sigma$-periodicity of THR in [19, Proposition 5.1.5], we can define the periodic $\mathbb{P}^1 \wedge \mathbb{P}^\sigma$-spectra

$$\begin{aligned}
\mathbf{THR} &:= (\mathrm{THR}^{\mathbb{Z}/2}, \mathrm{THR}^{\mathbb{Z}/2}, \dots) \in \mathrm{prelogSH}^{\mathbb{Z}/2}(S), \\
\mathbf{TCR} &:= (\mathrm{TCR}^{\mathbb{Z}/2}, \mathrm{TCR}^{\mathbb{Z}/2}, \dots) \in \mathrm{prelogSH}^{\mathbb{Z}/2}(S)
\end{aligned}$$

for every finite-dimensional noetherian separated scheme S. The construction of **THR** resembles the construction of the motivic real K-theory $\mathbb{Z}/2$-spectrum [20, §4.3]. We obtain new $\mathbb{P}^1$-spectra $\mathbf{THR}^{\mathbb{Z}/2}$ and $\mathbf{TCR}^{\mathbb{Z}/2}$, while we show $i^*\mathbf{THR} \simeq \mathbf{THH}$ and $i^*\mathbf{TCR} \simeq \mathbf{TC}$ in Proposition 6.17.

Organization of the article

To define the $\mathbb{P}^1 \wedge \mathbb{P}^\sigma$-spectrum **THR**, we need the following ingredients:

1. the definition of THR of log rings with involutions,
2. isovariant étale descent property for THR,
3. invariance of THR under passing to the associated log structure,
4. $\mathbb{P}^1 \wedge \mathbb{P}^\sigma$-periodicity of THR,
5. $(\mathbb{P}^n, \mathbb{P}^{n-1})$-invariance of THR.

Due to [19], we have (2) and (4). In Section 3, we deal with (1). This requires dihedral replete bar constructions in Section 2, which is an equivariant analogue of Rognes' replete bar constructions [29, Definition 3.16]. In Section 4, we show (3) for separated log schemes. The strategy is to work in a sufficiently local situation. In Section 5, we review the notion of real cyclotomic spectra following Quigley and Shah [27]. Then we generalize many results in the previous sections to real cyclotomic spectra so that we can define TCR of finite-dimensional noetherian separated log schemes. In Section 6, we show (5) by providing an explicit computation of $\mathrm{THR}(\mathbb{P}^n, \mathbb{P}^{n-1})$ using cubes in ∞-categories. Then we discuss properties of THR that can be deduced from [7], and we construct the $\mathbb{P}^1 \wedge \mathbb{P}^\sigma$-spectra **THR** and **TC**.

2 Dihedral replete bar constructions

Rognes [29, Definition 3.16] defined the replete bar construction of a commutative monoid, which is a key input to define topological Hochschild homology of log rings. In this section, we discuss a dihedral refinement of this construction.

We refer to [19, Example 4.1.3] for a review of real simplicial sets and dihedral sets, which are defined using crossed simplicial groups of Fiedorowicz and Loday [12]. A *real simplicial set X* is a simplicial set equipped with involutions $w_q : X_q \to X_q$ (i.e., an automorphism such that $w_q \circ w_q = \mathrm{id}$) for all simplicial degrees q satisfying the relations

$$d_i w_q = w_{q-1} d_{q-i}, \quad s_i w_q = w_{q+1} s_{q-i}$$

for $0 \leq i \leq q$. Note that a real simplicial set is different from a simplicial set with involution, i.e., a simplicial object in the category of sets with involution. A *dihedral simplicial set X* is a real simplicial set equipped with automorphisms $t_q : X_q \to X_q$ for all simplicial degrees q satisfying certain relations.

Let $\oplus$ denote the coproduct in the category of commutative monoids. For commutative monoids P and Q, the coproduct $P \oplus Q$ is naturally isomorphic to the product $P \times Q$.

Definition 2.1 Let Δ^1_σ be the real simplicial set whose underlying simplicial set is Δ^1 and whose involution $w : (\Delta^1)_q \to (\Delta^1)_q$ in simplicial degree q sends the q-simplex $a_0 \cdots a_q$ to $(1-a_0)\cdots(1-a_q)$. For a commutative monoid P with involution w, let $P \otimes \Delta^1_\sigma$ be the real simplicial set whose underlying simplicial set is the tensor product $P \otimes \Delta^1$ defined by

$$(P \otimes \Delta^1)_q := \bigoplus_{i \in (\Delta^1)_q} P$$

and whose involution is given by $(x_0, \dots, x_{q+1}) \to (w(x_{q+1}), \dots, w(x_0))$ in simplicial degree q, where the indices $0, 1, \dots, q+1$ correspond to the q-simplices $0\cdots00, 0\cdots01, \dots, 1\cdots11$ in Δ^1.

Proposition 2.2 *Let P be a commutative monoid with involution. Then the map of real simplicial sets*

$$P \otimes \Delta^1_\sigma \to P$$

given by $(x_0, \dots, x_q) \mapsto x_0 + \cdots + x_q$ in simplicial degree q is a $\mathbb{Z}/2$-weak equivalence.

Proof Let sd_σ be the Segal subdivision functor [33]. We have an isomorphism of simplicial sets with involutions

$$\mathrm{sd}_\sigma(\Delta^1_\sigma) \simeq \Delta^1 \amalg_{\{1\}} \Delta^1,$$

where the involution on the right-hand side switches the factors. This induces an isomorphism of simplicial sets with involutions

$$\mathrm{sd}_\sigma(P \otimes \Delta^1_\sigma) \simeq (P \otimes \Delta^1) \oplus_P (P \otimes \Delta^1),$$

where the involution on the right-hand side is obtained by the formula $(x, y) \mapsto (w(y), w(x))$ with w in Definition 2.1. Since Δ^1 is contractible, the map $P \to P \otimes \Delta^1$ induced by $\{1\} \to \Delta^1$ is a homotopy equivalence. This yields a $\mathbb{Z}/2$-homotopy equivalence

$$P \xrightarrow{\simeq} (P \otimes \Delta^1) \oplus_P (P \otimes \Delta^1),$$

which implies the claim. □

We review the ∞-categorical formulation of equivariant homotopy theory due to Bachmann and Hoyois [2, §9]. See [19, §A.1] for a more detailed review. Let FinGpd denote the 2-category of finite groupoids. Bachmann and Hoyois constructed the functor

$$\mathrm{SH} : \mathrm{Span}(\mathrm{FinGpd}) \to \mathrm{CAlg}(\mathrm{Cat}_\infty), \quad (X \xleftarrow{f} Y \xrightarrow{p} Z) \mapsto p_\otimes f^*;$$

see [2, §C] for the notation Span. For a morphism f in FinGpd, f^* admits a right adjoint f_*. If f is a finite covering, then f^* admits a left adjoint $f_\sharp$.

For a finite groupoid X, let $\mathrm{NAlg}(\mathrm{SH}(X))$ be the ∞-category of normed X-spectra [2, Definition 9.14], and let $\mathrm{CAlg}(\mathrm{SH}(X))$ be the ∞-category of $\mathbb{E}_\infty$-rings in $\mathrm{SH}(X)$. We have the forgetful functor $\mathrm{NAlg}(\mathrm{SH}(X)) \to \mathrm{CAlg}(\mathrm{SH}(X))$. We use the notation $\wedge$ for the coproduct in $\mathrm{NAlg}(\mathrm{SH}(X))$ and $\mathrm{CAlg}(\mathrm{SH}(X))$. For a morphism of finite groupoids $f : X \to S$, we have the induced adjoint functors

$$\begin{tikzcd} \mathrm{CAlg}(\mathrm{SH}(S)) \arrow[r, shift left, "f^*"] & \mathrm{CAlg}(\mathrm{SH}(X)). \arrow[l, shift left, "f_*"] \end{tikzcd}$$

If $f : X \to S$ is a finite covering of finite groupoids, we have the induced sequence of adjoint functors

$$\begin{tikzcd} \mathrm{NAlg}(\mathrm{SH}(S)) \arrow[r, "f^*" description] & \mathrm{NAlg}(\mathrm{SH}(X)). \arrow[l, bend right, "f_\sharp"'] \arrow[l, bend left, "f_*"] \end{tikzcd}$$

On the other hand, if f has connected fibers, then the induced functor

$$f_\otimes : \mathrm{CAlg}(\mathrm{SH}(X)) \to \mathrm{CAlg}(\mathrm{SH}(S))$$

preserves colimits.

For a finite group G, the *∞-category of G-spectra* is

$$\mathrm{Sp}^G := \mathrm{SH}(BG).$$

We also set $\mathrm{NAlg}^G := \mathrm{NAlg}(\mathrm{SH}(BG))$ and $\mathrm{CAlg}^G := \mathrm{CAlg}(\mathrm{SH}(BG))$. We omit the superscripts G in this notation if G is trivial. There is an equivalence of ∞-categories $\mathrm{CAlg} \simeq \mathrm{NAlg}$.

Consider the obvious functors $\mathrm{pt} \xrightarrow{i} B(\mathbb{Z}/2) \xrightarrow{p} \mathrm{pt}$. Observe that i is a finite covering and p has connected fibers. We often use the alternative notation

$$(-)^{\mathbb{Z}/2} := p_*, \quad N^{\mathbb{Z}/2} := i_\otimes, \quad \Phi^{\mathbb{Z}/2} := p_\otimes.$$

The functor $N^{\mathbb{Z}/2}$ is the *norm functor* of Hill, Hopkins, and Ravenel [17], and the functor $\Phi^{\mathbb{Z}/2}$ is the *geometric fixed point functor*. We have the induced functors

$$\begin{aligned} &i^* : \mathrm{Sp}^{\mathbb{Z}/2} \to \mathrm{Sp}, && N^{\mathbb{Z}/2}, i_\sharp, i_* : \mathrm{Sp} \to \mathrm{Sp}^{\mathbb{Z}/2}, \\ &i^* : \mathrm{NAlg}^{\mathbb{Z}/2} \to \mathrm{CAlg}, && N^{\mathbb{Z}/2}, i_\sharp, i_* : \mathrm{CAlg} \to \mathrm{NAlg}^{\mathbb{Z}/2}, \\ &p^* : \mathrm{Sp} \to \mathrm{Sp}^{\mathbb{Z}/2}, && \Phi^{\mathbb{Z}/2}, (-)^{\mathbb{Z}/2} : \mathrm{Sp}^{\mathbb{Z}/2} \to \mathrm{Sp}, \\ &p^* : \mathrm{CAlg} \to \mathrm{CAlg}^{\mathbb{Z}/2}, && \Phi^{\mathbb{Z}/2}, (-)^{\mathbb{Z}/2} : \mathrm{CAlg}^{\mathbb{Z}/2} \to \mathrm{CAlg}. \end{aligned}$$

We refer to [19, Proposition A.2.7] for fundamental relations among these functors. We note that the pair of functors $(i^*, \Phi^{\mathbb{Z}/2})$ is conservative.

Let $f : X \to S$ be a morphism of finite groupoids. For every map $R \to A$ in $\mathrm{NAlg}(\mathrm{SH}(S))$ and map $f^*R \to B$ in $\mathrm{NAlg}(\mathrm{SH}(X))$, we have the natural map

$$A \wedge_R f_*B \to f_*(f^*A \wedge_{f^*R} B)$$

given by the composite

$$A \wedge_R f_*B \xrightarrow{\mathrm{ad}} f_*f^*(A \wedge_R f_*B) \xrightarrow{\simeq} f_*(f^*A \wedge_{f^*R} f^*f_*B) \xrightarrow{\mathrm{ad}'} f_*(f^*A \wedge_{f^*R} B),$$

where ad (resp. ad′) denotes the map obtained by the unit (resp. counit).

Lemma 2.3 *Let $r : \mathbb{Z}/2 \to \mathrm{pt}$ be the obvious morphism of finite groupoids. Then the natural map*

$$A \wedge_R r_*B \to r_*(r^*A \wedge_{r^*R} B)$$

is an equivalence for every map $R \to A$ in CAlg and map $r^*R \to B$ in $\mathrm{NAlg}(\mathrm{SH}(\mathbb{Z}/2))$.

Proof As observed in [2, Example 9.15], we have an equivalence of ∞-categories $\mathrm{NAlg}(\mathrm{SH}(\mathbb{Z}/2)) \simeq \mathrm{CAlg}(\mathrm{SH}(\mathbb{Z}/2))$. We additionally have an equivalence of ∞-categories $\mathrm{SH}(\mathbb{Z}/2) \simeq \mathrm{SH}(\mathrm{pt}) \times \mathrm{SH}(\mathrm{pt})$ by [2, Lemma 9.6]; hence $\mathrm{CAlg}(\mathrm{SH}(\mathbb{Z}/2)) \simeq \mathrm{CAlg} \times \mathrm{CAlg}$. The functors $i_0^*, i_1^* : \mathrm{NAlg}(\mathrm{SH}(\mathbb{Z}/2)) \to \mathrm{CAlg}$ induced by the inclusions $i_0, i_1 : \mathrm{pt} \to \mathbb{Z}/2$ can be identified with the two projections $\mathrm{CAlg} \times \mathrm{CAlg} \to \mathrm{CAlg}$. Using $ri_0 = ri_1$, we see that the functor $r^* : \mathrm{CAlg} \to \mathrm{NAlg}(\mathrm{SH}(\mathbb{Z}/2))$ can be identified with the diagonal functor $\mathrm{CAlg} \to \mathrm{CAlg} \times \mathrm{CAlg}$, and the functor $r_* : \mathrm{NAlg}(\mathrm{SH}(\mathbb{Z}/2)) \to \mathrm{CAlg}$ can be identified with the direct sum functor $\oplus : \mathrm{CAlg} \times \mathrm{CAlg} \to \mathrm{CAlg}$. Use these explicit descriptions to show the claim. □

Lemma 2.4 *For $R, A \in \mathrm{NAlg}^{\mathbb{Z}/2}$ and $B \in \mathrm{CAlg}$, the natural map*

$$A \wedge_R i_*B \to i_*(i^*A \wedge_{i^*R} B) \tag{2-1}$$

is an equivalence.

Proof It suffices to show that (2-1) becomes equivalences after applying i^* and $\Phi^{\mathbb{Z}/2}$. With the aid of the forgetful functor $\mathrm{NAlg}^{\mathbb{Z}/2} \to \mathrm{CAlg}^{\mathbb{Z}/2}$, we obtain $\Phi^{\mathbb{Z}/2}(A \wedge_R i_* B) \simeq \Phi^{\mathbb{Z}/2} A \wedge_{\Phi^{\mathbb{Z}/2} R} \Phi^{\mathbb{Z}/2} i_* B$. By [19, Proposition A.2.7(3), (5)], $\Phi^{\mathbb{Z}/2} i_* \simeq 0$. Hence both sides of (2-1) vanish after applying $\Phi^{\mathbb{Z}/2}$.

Apply [19, Proposition A.1.9] to the cartesian square

$$\begin{tikzcd}
\mathbb{Z}/2 \arrow[r, "r"] \arrow[d, "r"'] & \mathrm{pt} \arrow[d, "i"] \\
\mathrm{pt} \arrow[r, "i"] & B(\mathbb{Z}/2)
\end{tikzcd}$$

to obtain a natural equivalence $r_* r^* \simeq i^* i_*$. We have natural equivalences

$$r_* r^*((-) \wedge_{(-)} (-)) \simeq r_*(r^*(-) \wedge_{r^*(-)} r^*(-)) \simeq (-) \wedge_{(-)} r_* r^*(-),$$

where the second one is due to Lemma 2.3. We have the induced commutative diagram

$$\begin{tikzcd}
i^*(A \wedge_R i_* B) \arrow[r, "\mathrm{ad}"] \arrow[d, "\simeq"'] & i^* i_* i^*(A \wedge_R i_* B) \arrow[r, "\simeq"] & i^* i_*(i^* A \wedge_{i^* R} i^* i_* B) \arrow[r, "\mathrm{ad}'"] \arrow[d, "\simeq"] & i^* i_*(i^* A \wedge_{i^* R} B) \arrow[d, "\simeq"] \\
i^* A \wedge_{i^* R} i^* i_* B \arrow[rr, "\mathrm{ad}"] \arrow[rrd, "\mathrm{ad}"'] & & r_* r^*(i^* A \wedge_{i^* R} i^* i_* B) \arrow[r, "\mathrm{ad}'"] \arrow[d, "\simeq"] & r_* r^*(i^* A \wedge_{i^* R} B) \arrow[d, "\simeq"] \\
& & i^* A \wedge_{i^* R} r_* r^* i^* i_* B \arrow[r, "\mathrm{ad}'"] & i^* A \wedge_{i^* R} r_* r^* B
\end{tikzcd}$$

Hence to show that (2-1) becomes an equivalence after applying i^*, it suffices to show that the composite of the upper vertical maps in this commutative diagram is an equivalence:

$$\begin{tikzcd}
i^* i_* B \arrow[r, "\mathrm{ad}"] \arrow[d, "\simeq"'] & r_* r^* i^* i_* B \arrow[r, "\mathrm{ad}'"] \arrow[d, "\simeq"] & r_* r^* B \arrow[d, "\simeq"] \\
r_* r^* B \arrow[r, "\mathrm{ad}"] & r_* r^* r_* r^* B \arrow[r, "\mathrm{ad}'"] & r_* r^* B
\end{tikzcd}$$

This is true since the composite of the lower horizontal maps is an equivalence by the counit-unit identity. □

For a simplicial set (resp. real simplicial set) X, we use the notation $\mathbb{S}[X] := \Sigma^\infty X_+$, which is a spectrum (resp. $\mathbb{Z}/2$-spectrum). If P is a commutative monoid, then $\mathbb{S}[P]$ is an object of CAlg. If P is a commutative monoid with involution, then $\mathbb{S}[P]$ is a commutative monoid in the model category of orthogonal $\mathbb{Z}/2$-spectrum in the sense of [19, Definition A.2.2]. Due to [19, Remark A.2.6], we can regard $\mathbb{S}[P]$ as an object of $\mathrm{NAlg}^{\mathbb{Z}/2}$.

For a commutative monoid P with involution, we denote by $\mathrm{N}^{\mathrm{di}} P$ the *dihedral nerve of* P in [19, Definition 4.2.2], and by $\mathrm{B}^{\mathrm{di}} P$ its dihedral geometric realization. In simplicial degree q, we have $(\mathrm{N}^{\mathrm{di}} P)_q := P^{\times(q+1)}$. We obtain the cyclic nerve $\mathrm{N}^{\mathrm{cy}} i^* P$ [34, §2.3] of $i^* P$ if we forget the involution structure on $\mathrm{N}^{\mathrm{di}} P$. There is a map of dihedral sets

$$\mathrm{N}^{\mathrm{di}} P \to P \tag{2-2}$$

sending $(x_0, \dots, x_q)$ to $x_0 + \dots + x_q$ in simplicial degree q.

Definition 2.5 Let $\mathcal{C}$ be an ordinary category. An *object of $\mathcal{C}$ with involution* is an object of the category $\mathcal{C}_{\mathbb{Z}/2} := \mathrm{Fun}(B(\mathbb{Z}/2), \mathcal{C})$. Let

$$i^* : \mathrm{Fun}(B(\mathbb{Z}/2), \mathcal{C}) \to \mathcal{C}$$

be the forgetful functor, and let $i_\sharp$ (resp. i_*) be its left adjoint (resp. right adjoint) if it exists.

Example 2.6 If P is a commutative monoid, then i_*P is the commutative monoid $P \times P$ with the involution w given by $w(x,y) := (w(y), w(x))$. Furthermore, there is a natural isomorphism $i_\sharp P \simeq i_* P$. We have a similar description of i_*A for any commutative ring A, but $i_\sharp A \not\simeq i_* A$.

If X is a scheme, then $i_\sharp X$ is the scheme $X \amalg X$ with the involution switching the two components.

Lemma 2.7 *Let $P \to Q$ be a map of commutative monoids with involutions. Then there is a natural equivalence of $\mathbb{Z}/2$-spectra*

$$\mathbb{S}[P \amalg P] \wedge_{\mathbb{S}[P]} \mathbb{S}[Q] \simeq \mathbb{S}[(P \amalg P) \oplus_P Q],$$

where the involution on $P \amalg P$ in the formulation switches the components.

Proof There is a natural isomorphism of sets with involutions

$$(P \amalg P) \oplus_P Q \simeq Q \amalg Q,$$

where the involution on the right-hand side switches the components. By the explicit description of the functors i^* and i_* in terms of orthogonal spectra in [19, Construction A.2.4], we have natural equivalences

$$\mathbb{S}[P \amalg P] \simeq i_* i^* \mathbb{S}[P] \quad \text{and} \quad \mathbb{S}[Q \amalg Q] \simeq i_* i^* \mathbb{S}[Q].$$

Lemma 2.4 yields a natural equivalence

$$i_* i^* \mathbb{S}[P] \wedge_{\mathbb{S}[P]} \mathbb{S}[Q] \simeq i_*(i^* \mathbb{S}[P] \wedge_{i^*\mathbb{S}[P]} i^* \mathbb{S}[Q]).$$

Combine what we have discussed above to obtain the desired equivalence. □

Proposition 2.8 *Let P be a commutative monoid with involution. Then there is a natural equivalence of $\mathbb{Z}/2$-spectra*

$$\mathbb{S}[\mathrm{B}^{\mathrm{di}} P] \simeq \mathbb{S}[P] \wedge_{\mathbb{S}[i_\sharp i^* P]} \mathbb{S}[P],$$

where the homomorphism $i_\sharp i^* P \to P$ in the formulation is obtained by the counit of the adjunction pair $(i_\sharp, i^*)$.

Proof There is an isomorphism of real simplicial sets

$$(P \otimes \Delta^1_\sigma) \oplus_{i_\sharp i^* P} P \simeq \mathrm{N}^{\mathrm{di}} P, \tag{2-3}$$

where the map $i_\sharp i^* P \to P \otimes \Delta^1_\sigma$ is given by $(x,y) \mapsto (x, 0, \ldots, 0, y)$ in simplicial degree q. The composite

$$i_\sharp i^* P \to P \otimes \Delta^1_\sigma \to P$$

coincides with the counit homomorphism. Since $P \otimes \Delta^1_\sigma$ is degreewise the disjoint union of finitely many copies of $i_\sharp i^* P$ and $i_\sharp i^* P \amalg i_\sharp i^* P$ with the switching involution as an $i_\sharp i^* P$-set, Lemma 2.7 yields a natural equivalence

$$\mathbb{S}[P \otimes \Delta^1_\sigma] \wedge_{\mathbb{S}[i_\sharp i^* P]} \mathbb{S}[P] \simeq \mathbb{S}[(P \otimes \Delta^1_\sigma) \oplus_{i_\sharp i^* P} P].$$

Use Proposition 2.2 and (2-3) to finish the proof. □

For a commutative monoid P, let P^{gp} denote its group completion.

Definition 2.9 Let P be a commutative monoid with involution. The *dihedral replete nerve of P* is the dihedral set

$$\mathrm{N}^{\mathrm{drep}} P := \mathrm{N}^{\mathrm{di}} P^{\mathrm{gp}} \times_{P^{\mathrm{gp}}} P, \tag{2-4}$$

where the map $\mathrm{N}^{\mathrm{di}} P^{\mathrm{gp}} \to P^{\mathrm{gp}}$ in this formulation is given by (2-2) for P^{gp}. The *dihedral replete bar construction of P*, denoted by $\mathrm{B}^{\mathrm{drep}} P$, is the dihedral geometric realization of $\mathrm{N}^{\mathrm{drep}} P$.

We obtain the replete bar construction $\mathrm{B}^{\mathrm{di}} i^* P$ [29, Definition 3.16] of $i^* P$ if we forget the involution structure on $\mathrm{B}^{\mathrm{drep}} P$.

Proposition 2.10 *Let P and Q be commutative monoids with involutions. Then there is a natural isomorphism of dihedral sets*

$$\mathrm{N}^{\mathrm{drep}}(P \times Q) \simeq \mathrm{N}^{\mathrm{drep}} P \times \mathrm{N}^{\mathrm{drep}} Q.$$

Proof By [19, Proposition 4.2.4], we have a natural isomorphism of dihedral sets

$$\mathrm{N}^{\mathrm{di}}(P^{\mathrm{gp}} \times Q^{\mathrm{gp}}) \times_{P^{\mathrm{gp}} \times Q^{\mathrm{gp}}} (P \times Q) \simeq (\mathrm{N}^{\mathrm{di}} P^{\mathrm{gp}} \times \mathrm{N}^{\mathrm{di}} Q^{\mathrm{gp}}) \times_{P^{\mathrm{gp}} \times Q^{\mathrm{gp}}} (P \times Q).$$

The left side is isomorphic to $\mathrm{N}^{\mathrm{drep}}(P \times Q)$, and the right side is isomorphic to $\mathrm{N}^{\mathrm{drep}} P \times \mathrm{N}^{\mathrm{drep}} Q$. □

For a commutative monoid P with involution, let $N^\sigma P$ denote the *real nerve of P* [19, Definition 4.2.1].

Proposition 2.11 *Let P be a commutative monoid with involution. Then there is a natural isomorphism of dihedral sets*

$$\mathrm{N}^{\mathrm{drep}} P \simeq P \times \mathrm{N}^\sigma P^{\mathrm{gp}}.$$

Proof Immediate from [19, Proposition 4.2.6] and (2-4). □

Definition 2.12 For a commutative monoid P with involution w, let $(i_\sharp i^* P)^{\mathrm{ex}}$ denote the commutative monoid $P \oplus P^{\mathrm{gp}}$ with the involution w given by

$$w(x,y) := (w(x), w(x) - w(y)).$$

We have the commutative triangle

$$\begin{tikzcd} & (i_\sharp i^* P)^{\mathrm{ex}} \arrow[dr, "\mu^{\mathrm{ex}}"] & \\ i_\sharp i^* P \arrow[ur, "\gamma"] \arrow[rr, "\mu"] & & P \end{tikzcd}$$

such that $\mu(x,y):=x+y$, $\mu^{\mathrm{ex}}(x,y):=x$, and $\gamma(x,y):=(x+y,y)$ for $x,y\in P$. We note that μ is the counit homomorphism. The construction of $(i_\sharp i^*P)^{\mathrm{ex}}$ is an equivariant analogue of the exactification in [25, Proposition I.4.2.19].

Proposition 2.13 *Let P be a commutative monoid with involution. Then there is a natural equivalence of $\mathbb{Z}/2$-spectra*

$$\mathbb{S}[\mathrm{B}^{\mathrm{drep}}P]\simeq \mathbb{S}[P]\wedge_{\mathbb{S}[(i_\sharp i^*P)^{\mathrm{ex}}]}\mathbb{S}[P].$$

Proof Let $\mathrm{E}P^{\mathrm{gp}}$ denote the total simplicial set of P^{gp}, and let Q be the real simplicial set whose underlying simplicial set is $P\times \mathrm{E}P^{\mathrm{gp}}$, and whose involution is given by

$$(x,g_0,\dots,g_q)\in P\times (P^{\mathrm{gp}})^{\times(q+1)}\mapsto (w(x),w(x)-w(g_q),\dots,w(x)-w(g_0))$$

in simplicial degree q for every integer $q\geq 0$. Consider the maps of real simplicial sets

$$(i_\sharp i^*P)^{\mathrm{ex}}\to Q\to P\times \mathrm{N}^\sigma P^{\mathrm{gp}}$$

given by $(x,g)\mapsto(x,g,\dots,g)$ and $(x,g_0,\dots,g_q)\mapsto(x,g_1-g_0,\dots,g_q-g_{q-1})$ in simplicial degree q.

There is a real simplicial isomorphism

$$Q\oplus_{(i_\sharp i^*P)^{\mathrm{ex}}}P\simeq P\times \mathrm{N}^\sigma P^{\mathrm{gp}}.$$

Since Q is degreewise the disjoint union of finite copies of $(i_\sharp i^*P)^{\mathrm{ex}}$ and $(i_\sharp i^*P)^{\mathrm{ex}}\amalg (i_\sharp i^*P)^{\mathrm{ex}}$ with the switching involution as an $(i_\sharp i^*P)^{\mathrm{ex}}$-set, Lemma 2.7 yields a natural equivalence of $\mathbb{Z}/2$-spectra

$$\mathbb{S}[Q]\wedge_{\mathbb{S}[(i_\sharp i^*P)^{\mathrm{ex}}]}\mathbb{S}[P]\simeq \mathbb{S}[P\times \mathrm{N}^\sigma P^{\mathrm{gp}}]. \tag{2-5}$$

Let us show that the map $Q\to P$ given by

$$(x,g_0,\dots,g_q)\mapsto x$$

in simplicial degree q is a $\mathbb{Z}/2$-homotopy equivalence. Its underlying map of simplicial sets is the projection $P\times \mathrm{E}P^{\mathrm{gp}}\to P$, which is a homotopy equivalence. The $\mathbb{Z}/2$-fixed point of the Segal subdivision $\mathrm{sd}_\sigma Q$ is in simplicial degree q the set

$$\{(x,g_0,\dots,g_q,w(x)-w(g_q),\dots,w(x)-w(g_0)):x\in P^{\mathbb{Z}/2},g_0,\dots,g_q\in P^{\mathrm{gp}}\}.$$

From this description, we see that the induced map

$$(\mathrm{sd}_\sigma Q)^{\mathbb{Z}/2}\to(\mathrm{sd}_\sigma P)^{\mathbb{Z}/2}$$

can be identified with the projection $P\times \mathrm{E}P^{\mathrm{gp}}\to P$. Hence the map $Q\to P$ is a $\mathbb{Z}/2$-homotopy equivalence. Combine this with Proposition 2.11 and (2-5) to obtain the desired equivalence. □

Remark 2.14 For a commutative monoid P with involution, we can regard $\mathbb{S}[\mathrm{B}^{\mathrm{di}}P]$ and $\mathbb{S}[\mathrm{B}^{\mathrm{drep}}P]$ as objects of $\mathrm{NAlg}^{\mathbb{Z}/2}$ by Propositions 2.8 and 2.13.

Proposition 2.15 *Let P be a commutative monoid with involution. Then there are natural equivalences of $\mathbb{Z}/2$-spectra*

$$i^*\mathbb{S}[\mathrm{B}^{\mathrm{di}}P] \simeq \mathbb{S}[\mathrm{B}^{\mathrm{cy}}i^*P] \quad \textit{and} \quad i^*\mathbb{S}[\mathrm{B}^{\mathrm{drep}}P] \simeq \mathbb{S}[\mathrm{B}^{\mathrm{rep}}i^*P].$$

Proof Observe that we obtain $\mathrm{B}^{\mathrm{cy}}i^*P$ (resp. $\mathrm{B}^{\mathrm{rep}}i^*P$) by forgetting the involution structure on $\mathrm{B}^{\mathrm{di}}P$ (resp. $\mathrm{B}^{\mathrm{drep}}P$). □

3 THR of log rings with involutions

For $R \in \mathrm{NAlg}^{\mathbb{Z}/2}$, the *real topological Hochschild homology of R* is defined to be

$$\mathrm{THR}(R) := R \wedge_{N^{\mathbb{Z}/2}i^*R} R,$$

where the maps $N^{\mathbb{Z}/2}i^*R \to R$ in the formulation are the counit of the adjunction pair $(N^{\mathbb{Z}/2}, i^*)$. If A is a commutative ring with involution, then the equivariant Eilenberg–Mac Lane spectrum $\mathrm{H}A$ can be realized as a commutative orthogonal $\mathbb{Z}/2$-ring spectrum by [31, Example 11.12], so we have $\mathrm{H}A \in \mathrm{NAlg}^{\mathbb{Z}/2}$ together with [19, Remark A.2.3]. We set $\mathrm{THR}(A) := \mathrm{THR}(\mathrm{H}A)$. For a commutative monoid P with involution, there is a natural equivalence of $\mathbb{Z}/2$-spectra

$$\mathrm{THR}(\mathbb{S}[P]) \simeq \mathbb{S}[\mathrm{B}^{\mathrm{di}}P]; \tag{3-1}$$

see [18] and also [10, Proposition 5.9]. One can also show this using Proposition 2.8 and the natural equivalence of $\mathbb{Z}/2$-spectra $N^{\mathbb{Z}/2}i^*\mathbb{S}[P] \simeq \mathbb{S}[i_\sharp i^*P]$.

Recall from [25, Definition III.1.2.3] that a *log ring*[1] (A, P) is a commutative ring A equipped with a homomorphism $P \to A$ of commutative monoids, where the monoid operation on A is the multiplication. A *homomorphism of log rings* $(A, P) \to (B, Q)$ is a pair of homomorphisms $A \to B$ and $P \to Q$ such that the square

$$\begin{tikzcd} P \arrow[r] \arrow[d] & A \arrow[d] \\ Q \arrow[r] & B \end{tikzcd}$$

commutes. We regard a commutative ring A as a log ring $(A, \{1\})$.

Rognes [29, Definition 8.11, Remark 8.12] defined the *topological Hochschild homology of a log ring* (A, P) as the coproduct, in CAlg,

$$\mathrm{THH}(A, P) := \mathrm{THH}(A) \wedge_{\mathbb{S}[\mathrm{B}^{\mathrm{cy}}P]} \mathbb{S}[\mathrm{B}^{\mathrm{rep}}P], \tag{3-2}$$

where the map $\mathbb{S}[\mathrm{B}^{\mathrm{cy}}P] \simeq \mathrm{THH}(\mathbb{S}[P]) \to \mathrm{THH}(A)$ in the formulation is obtained by applying THH to the induced map $\mathbb{S}[P] \to \mathrm{H}A$.

---

[1] A log ring is often called a *prelog ring* in the literature; see, e.g., [29, Definition 2.1].

Definition 3.1 The *real topological Hochschild homology of a log ring* (A, P) *with involution* is the coproduct, in $\mathrm{NAlg}^{\mathbb{Z}/2}$,

$$\mathrm{THR}(A, P) := \mathrm{THR}(A) \wedge_{\mathbb{S}[\mathrm{B}^{\mathrm{di}} P]} \mathbb{S}[\mathrm{B}^{\mathrm{drep}} P], \tag{3-3}$$

where the map $\mathbb{S}[\mathrm{B}^{\mathrm{di}} P] \simeq \mathrm{THR}(\mathbb{S}[P]) \to \mathrm{THR}(A)$ in this formulation is obtained by applying THR to the induced map $\mathbb{S}[P] \to \mathrm{H}A$.

We obviously have $\mathrm{THR}(A, \{1\}) \simeq \mathrm{THR}(A)$ for every commutative ring A.

Definition 3.2 Let $P \to M$ be a homomorphism of commutative monoids with involutions. For notational convenience, we set

$$\mathrm{THR}(\mathbb{S}[M], P) := \mathbb{S}[\mathrm{B}^{\mathrm{di}} M] \wedge_{\mathbb{S}[\mathrm{B}^{\mathrm{di}} P]} \mathbb{S}[\mathrm{B}^{\mathrm{drep}} P].$$

Proposition 3.3 *Let $Q \to M$ be a homomorphism of commutative monoids with involutions. Then there is a natural equivalence of $\mathbb{Z}/2$-spectra*

$$\mathrm{THR}(A[M], P \oplus Q) \simeq \mathrm{THR}(A, P) \wedge \mathrm{THR}(\mathbb{S}[M], Q).$$

Proof By [19, Proposition 4.2.4] and Proposition 2.10, there are natural isomorphisms of dihedral sets

$$\mathrm{N}^{\mathrm{di}}(P \oplus Q) \simeq \mathrm{N}^{\mathrm{di}} P \times \mathrm{N}^{\mathrm{di}} Q \text{ and } \mathrm{N}^{\mathrm{drep}}(P \oplus Q) \simeq \mathrm{N}^{\mathrm{drep}} P \times \mathrm{N}^{\mathrm{drep}} Q.$$

Apply $\mathbb{S}[-]$ to these, and use (3-3) to obtain the desired equivalence. □

Let S^σ be S^1 with the involution given by $e^{i\theta} \in S^1 \mapsto e^{-i\theta}$.

Example 3.4 By [19, (4.12)], we have an equivalence of $\mathbb{Z}/2$-spectra

$$\mathrm{THR}(\mathbb{S}[\mathbb{N}], \mathbb{N}) \simeq \bigoplus_{d=0}^{\infty} \mathbb{S}[S^\sigma].$$

Furthermore, [19, Propositions 4.2.11 and 4.2.12] implies that the induced map

$$\mathrm{THR}(\mathbb{S}[\mathbb{N}]) \to \mathrm{THR}(\mathbb{S}[\mathbb{N}], \mathbb{N})$$

can be written as the componentwise induced map

$$\mathbb{S} \oplus \bigoplus_{d=1}^{\infty} \mathbb{S}[S^\sigma] \to \bigoplus_{d=0}^{\infty} \mathbb{S}[S^\sigma].$$

In the remaining part of this section, we investigate how THH and THR interact with the functors i^* and i_*.

Proposition 3.5 *Let (A, P) be a log ring with involution. Then there is a natural equivalence of spectra*

$$i^*\mathrm{THR}(A, P) \simeq \mathrm{THH}(i^*A, i^*P).$$

Proof We have a natural equivalence $i^*\mathrm{THR}(A) \simeq \mathrm{THH}(i^*A)$ by [19, Proposition 3.4.7]. Since the functor $i^* : \mathrm{NAlg}^{\mathbb{Z}/2} \to \mathrm{CAlg}$ preserves colimits, we have a natural equivalence

$$i^*\mathrm{THR}(A,P) \simeq i^*\mathrm{THR}(A) \wedge_{i^*\mathbb{S}[\mathrm{B}^{\mathrm{di}}P]} i^*\mathbb{S}[\mathrm{B}^{\mathrm{drep}}P].$$

Together with Proposition 2.15, we obtain the desired equivalence. □

For a commutative monoid P, let $\underline{P}$ denote the commutative monoid P with the trivial involution.

Construction 3.6 Let (A,P) be a log ring. We have the composite map of spectra

$$i^*\mathrm{THR}(i_*A,\underline{P}) \xrightarrow{\simeq} \mathrm{THH}(A\oplus A,P) \to \mathrm{THH}(A,P),$$

where log structure homomorphism $\underline{P} \to i_*A$ sends $p \in P$ to $(\alpha(p),\alpha(p))$, where $\alpha : P \to A$ is the log structure homomorphism. and the first arrow is obtained by Proposition 3.5, and the second arrow is induced by the summation homomorphism $A \oplus A \to A$. By adjunction, we obtain a map of $\mathbb{Z}/2$-spectra

$$\mathrm{THR}(i_*A,\underline{P}) \to i_*\mathrm{THH}(A,P). \tag{3-4}$$

Proposition 3.7 *Let (A,P) be a log ring. Then* (3-4) *is an equivalence.*

Proof Lemma 2.4 and Proposition 2.15 yield natural equivalences of $\mathbb{Z}/2$-spectra

$$\begin{aligned} i_*\mathrm{THH}(A,P) &\simeq i_*(\mathrm{THH}(A) \wedge_{i^*\mathbb{S}[\mathrm{B}^{\mathrm{di}}\underline{P}]} i^*\mathbb{S}[\mathrm{B}^{\mathrm{drep}}\underline{P}]) \\ &\simeq i_*\mathrm{THH}(A) \wedge_{\mathbb{S}[\mathrm{B}^{\mathrm{di}}\underline{P}]} \mathbb{S}[\mathrm{B}^{\mathrm{drep}}\underline{P}]. \end{aligned}$$

By [19, Propositions 2.1.4, 2.3.3], we have a natural equivalence of $\mathbb{Z}/2$-spectra

$$i_*\mathrm{THH}(A) \simeq \mathrm{THR}(i_*A).$$

Combine these with the definition of $\mathrm{THR}(i_*A,\underline{P})$ to conclude. □

Proposition 3.8 *The functor* THR *from the category of log rings with involutions to $\mathbb{Z}/2$-spectra preserves filtered colimits.*

Proof One can directly check that the endofunctors $P \mapsto i_\sharp i^*P, (i_\sharp i^*P)^{\mathrm{ex}}$ on the category of commutative monoids with involutions preserve colimits. By Propositions 2.8 and 2.13, the functors $P \mapsto \mathbb{S}[\mathrm{B}^{\mathrm{di}}P], \mathbb{S}[\mathrm{B}^{\mathrm{drep}}P]$ from the category of commutative monoids with involutions to $\mathrm{NAlg}_{\mathbb{Z}/2}$ preserve colimits.

On the other hand, the functor $A \mapsto \mathrm{THR}(A)$ from the category of commutative rings with involutions to $\mathrm{NAlg}_{\mathbb{Z}/2}$ preserves filtered colimits since the Eilenberg–Mac Lane functor preserves filtered colimits, $N^{\mathbb{Z}/2}$ and i^* preserve colimits, and $\wedge$ is the coproduct. It follows that the functor $(A,P) \mapsto \mathrm{THR}(A,P)$ from the category of commutative rings with involutions to $\mathrm{NAlg}_{\mathbb{Z}/2}$ preserves filtered colimits. The forgetful functor $\mathrm{NAlg}_{\mathbb{Z}/2} \to \mathrm{Sp}_{\mathbb{Z}/2}$ preserves filtered colimits as observed in [19, §1], which finishes the proof. □

4 THR of log schemes with involutions

So far, we have discussed THR of log rings (A, P) with involutions. The purpose of this section is to define $\mathrm{THR}(X)$ for every separated log scheme X and to show that a canonical map

$$\mathrm{THR}(A, P) \to \mathrm{THR}(\mathrm{Spec}(A, P))$$

is an equivalence of $\mathbb{Z}/2$-spectra.

Let us briefly review basic notation and terminology in log geometry. We refer to Ogus's book [25] for the details. For a commutative monoid P, let P^* denote its submonoid of units. We set $\overline{P} := P/P^*$. We say that P is *integral* if the induced homomorphism $P \to P^{\mathrm{gp}}$ is injective.

- For a log scheme X, let $\underline{X}$ be its underlying scheme, and let $\mathcal{M}_X$ be its structure sheaf of monoids.
- A morphism of log schemes $f : Y \to X$ is *strict* if the induced morphism $Y \to X \times_{\underline{X}} \underline{Y}$ is an isomorphism.
- For a log ring (A, P), let $\mathrm{Spec}(A, P)$ denote its associated log scheme in the sense of Definition III.1.2.3 in [25].
- For a commutative monoid P, we set $\mathbb{A}_P := \mathrm{Spec}(\mathbb{Z}[P], P)$, whose structure homomorphism $P \to \mathbb{Z}[P]$ sends $p \in P$ to p.
- A *chart P of a log scheme X* is a commutative monoid P together with a strict morphism $X \to \mathbb{A}_P$ of log schemes.
- A log scheme X is *integral* if $\mathcal{M}_X$ is a sheaf of integral monoids.
- A log scheme X is *fine saturated* (or simply *fs*) if X admits strict étale locally a chart P such that P is a fine saturated monoid.

Let G be a finite group. For a point x of a G-scheme X, the *scheme-theoretic stabilizer of X at x* is

$$G_x := \ker\bigl(\{g \in G : gx = x\} \to \mathrm{Aut}(k(x))\bigr).$$

A morphism of G-schemes $Y \to X$ is an *isovariant étale cover* if the induced homomorphism $G_y \to G_{f(x)}$ is an isomorphism for every point $y \in Y$ and f is étale and surjective after forgetting the G-action.

For a commutative ring A with involution, an *A-algebra with involution* is a commutative ring B with involution equipped with a $\mathbb{Z}/2$-equivariant map $A \to B$.

Proposition 4.1 *Let (A, P) be a log ring with involution. Then the presheaf* $\mathrm{THR}(-, P)$ *on the opposite category of A-algebras with involution is an isovariant étale hypersheaf.*

Proof Let $B \to C$ be an isovariant étale homomorphism of A-algebras with involutions. By Theorem 3.2.3 of [19], there are natural equivalences of $\mathbb{Z}/2$-spectra

$$\mathrm{THR}(B, P) \wedge_{\mathrm{H}B} \mathrm{H}C \simeq \mathrm{THR}(B, P) \wedge_{\mathrm{H}\iota(B^{\mathbb{Z}/2})} \mathrm{H}\iota(C^{\mathbb{Z}/2}) \simeq \mathrm{THR}(C, P),$$

where ιM denotes the commutative monoid with the trivial involution for a commutative monoid M. Argue as in the proof of [19, Theorem 3.4.3] to show the claim. □

Lemma 4.2 *Let $\theta : P \to Q$ be a homomorphism of integral monoids. If $\bar{\theta} : \bar{P} \to \bar{Q}$ is an isomorphism, then the induced homomorphism of monoids*

$$\eta : P \oplus_{P^*} Q^* \to Q \tag{4-1}$$

is an isomorphism.

Proof Since $\bar{\theta}$ is an isomorphism, η is surjective. To show that η is injective, assume $\eta(p, v) = \eta(p', v')$ with $p, p' \in P$ and $v, v' \in Q^*$. This implies $\theta(p) + v = \theta(p') + v'$. Since $\bar{\theta}$ is an isomorphism, there exists $u \in P^*$ satisfying $p = p' + u$. Together with the assumption that Q is integral, we have $v' = v + \theta(u)$. Use [25, Proposition I.1.1.5(3)] to see that $(p, v) = (p', v')$ in $P \oplus_{P^*} Q^*$. Hence η is an isomorphism. □

Lemma 4.3 *Let $\theta : P \to Q$ be a homomorphism of integral monoids with involution. If $\bar{\theta} : \bar{P} \to \bar{Q}$ is surjective, then the induced map*

$$\mathrm{THR}(\mathbb{S}[Q], P) \to \mathrm{THR}(\mathbb{S}[Q], Q)$$

is an equivalence of $\mathbb{Z}/2$-spectra.

Proof Due to Propositions 2.8 and 2.13, it suffices to show that the induced map

$$(\mathbb{S}[Q] \wedge_{\mathbb{S}[i_\sharp i^* Q]} \mathbb{S}[Q]) \wedge_{(\mathbb{S}[P] \wedge_{\mathbb{S}[i_\sharp i^* P]} \mathbb{S}[P])} (\mathbb{S}[P] \wedge_{\mathbb{S}[(i_\sharp i^* P)^{\mathrm{ex}}]} \mathbb{S}[P]) \to \mathbb{S}[Q] \wedge_{\mathbb{S}[(i_\sharp i^* Q)^{\mathrm{ex}}]} \mathbb{S}[Q]$$

is an equivalence of $\mathbb{Z}/2$-spectra. The left-hand side is equivalent to

$$\mathbb{S}[Q] \wedge_{(\mathbb{S}[i_\sharp i^* Q] \wedge_{\mathbb{S}[i_\sharp i^* P]} \mathbb{S}[(i_\sharp i^* P)^{\mathrm{ex}}])} \mathbb{S}[Q].$$

Hence it suffices to show that the induced map

$$\mathbb{S}[i_\sharp i^* Q] \wedge_{\mathbb{S}[i_\sharp i^* P]} \mathbb{S}[(i_\sharp i^* P)^{\mathrm{ex}}] \to \mathbb{S}[(i_\sharp i^* Q)^{\mathrm{ex}}]$$

is an equivalence of $\mathbb{Z}/2$-spectra. Since P is integral, $u + x = x$ implies $u = 0$ for $u \in P^*$ and $x \in P$. It follows that $i^* P$ is a free $i^* P^*$-set, so the induced map

$$\mathbb{S}[i_\sharp i^* Q^*] \wedge_{\mathbb{S}[i_\sharp i^* P^*]} \mathbb{S}[i_\sharp i^* P] \to \mathbb{S}[i_\sharp i^* Q]$$

is an equivalence of $\mathbb{Z}/2$-spectra. Hence it suffices to show that the induced map

$$\mathbb{S}[i_\sharp i^* Q^*] \wedge_{\mathbb{S}[i_\sharp i^* P^*]} \mathbb{S}[(i_\sharp i^* P)^{\mathrm{ex}}] \to \mathbb{S}[(i_\sharp i^* Q)^{\mathrm{ex}}] \tag{4-2}$$

is an equivalence of $\mathbb{Z}/2$-spectra.

Consider the localization P_F with respect to the face $F := \theta^{-1}(Q^*)$ of P. By [25, Theorem I.4.5.7, Proposition I.4.6.3(3), Remark I.4.6.6], $i^* P_F$ is a filtered colimit of free $i^* P$-sets. This implies that $\mathbb{S}[i_\sharp i^* P_F]$ is a filtered colimit of free $\mathbb{S}[i_\sharp i^* P]$-modules. It follows that the induced map

$$\mathbb{S}[i_\sharp i^* P_F] \wedge_{\mathbb{S}[i_\sharp i^* P]} \mathbb{S}[(i_\sharp i^* P)^{\mathrm{ex}}] \to \mathbb{S}[(i_\sharp i^* P_F)^{\mathrm{ex}}]$$

is an equivalence. Replace $P \to Q$ with $P_F \to Q$ to reduce to the case when $\bar{\theta}$ is an isomorphism. Then we have $Q \simeq P \oplus_{P^*} Q^*$ by Lemma 4.2.

We have the induced cocartesian square

$$\begin{tikzcd} i^*P^* \oplus i^*P^* \arrow[r, "\alpha"] \arrow[d] & i^*P \oplus i^*P^{\mathrm{gp}} \arrow[d] \\ i^*Q^* \oplus i^*Q^* \arrow[r] & i^*Q \oplus i^*Q^{\mathrm{gp}} \end{tikzcd} \tag{4-3}$$

where α sends (p, p') to $(p+p', p')$. Observe that $i^*P \oplus i^*P^{\mathrm{gp}}$ is a free $i^*P^* \oplus i^*P^*$-set. Apply $\mathbb{S}[-]$ to this square to see that the induced map

$$\mathbb{S}[i^*Q \oplus i^*Q] \wedge_{\mathbb{S}[i^*P \oplus i^*P]} \mathbb{S}[i^*P \oplus i^*P^{\mathrm{gp}}] \to \mathbb{S}[i^*Q \oplus i^*Q^{\mathrm{gp}}]$$

is an equivalence of spectra, i.e., (4-2) becomes an equivalence after applying i^*. Hence it remains to show that (4-2) becomes an equivalence after applying $\Phi^{\mathbb{Z}/2}$.

We claim that θ is exact, i.e., the induced square

$$\begin{tikzcd} P \arrow[r, "\eta"] \arrow[d, "\theta"'] & P^{\mathrm{gp}} \arrow[d, "\theta"] \\ Q \arrow[r, "\eta'"] & Q^{\mathrm{gp}} \end{tikzcd}$$

is cartesian. Assume that $\eta'(q) = \theta(p')$ for some $q \in Q$ and $p' \in P^{\mathrm{gp}}$. Using [25, Proposition I.1.1.5(3)] for $Q \simeq P \oplus_{P^*} Q^*$, we can write $q = (p, v)$ for some $p \in P$ and $v \in Q^*$. Since the group completion functor is a left adjoint, we have an isomorphism $Q^{\mathrm{gp}} \simeq P^{\mathrm{gp}} \oplus_{P^*} Q^*$. Using [25, Proposition I.1.1.5(3)] for this, we have $(\eta(p), v) = (p', 0)$ in Q'^{gp}, and hence we have $v = \theta(u)$ for some $u \in P^*$. This implies that $q = \theta(p+u)$ and $p' = \eta(p+u)$, so θ is exact.

Observe that the $\mathbb{Z}/2$-fixed point monoids of $i_\sharp i^*P^*$ and $i_\sharp i^*Q^*$ are isomorphic to i^*P^* and i^*Q^*. Let w denote the involutions on P and Q, and let M and N be the $\mathbb{Z}/2$-fixed point monoids of $(i_\sharp i^*P)^{\mathrm{ex}}$ and $(i_\sharp i^*Q)^{\mathrm{ex}}$. Observe that N is a submonoid of $Q \oplus Q^{\mathrm{gp}}$ consisting of (q, y) such that $\eta'(q) = y + w(y)$. We have a similar description for M too. Using $Q^{\mathrm{gp}} \simeq P^{\mathrm{gp}} \oplus_{P^*} Q^*$, we can find $x \in P^{\mathrm{gp}}$ and $v \in Q^*$ such that $y = \theta(x) + v$. Since θ is exact, there exists a unique $p \in P$ such that $\eta(p) = x + w(x)$ and $\theta(p) = q - v - w(v)$. This implies that the induced homomorphism $M \oplus_{i^*P^*} i^*Q^* \to N$ is surjective.

On the other hand, if $(p, x) \in M$ and $v \in Q^*$ satisfies $(\theta(p), \theta(x)) = (v + w(v), v)$ in N, then we have $p \in \theta^{-1}(Q^*)$. By [25, Proposition I.1.1.5(3)] for $Q \cong P \oplus_{P^*} Q^*$, we have $\theta^{-1}(Q^*) = P^*$, so we have $p \in P^*$. This implies that the induced homomorphism $M \oplus_{i^*P^*} i^*Q^* \to N$ is injective and hence an isomorphism. Since M is a free i^*P^*-set, we deduce that the induced map

$$\mathbb{S}[i^*Q] \wedge_{\mathbb{S}[i^*P]} \mathbb{S}[M] \to \mathbb{S}[N]$$

is an equivalence of spectra, i.e., (4-2) becomes an equivalence after applying $\Phi^{\mathbb{Z}/2}$. □

A morphism of log schemes with involutions $f : Y \to X$ is a *strict isovariant étale cover* if f is strict and its underlying morphism of schemes with involutions $\underline{f} : \underline{Y} \to \underline{X}$ is an isovariant étale cover. The *strict isovariant étale topology* is the topology generated by strict isovariant étale covers. Let sisoét be the shorthand for this topology.

Proposition 4.4 *Let X be a separated integral log scheme with involution. Then there exists a strict isovariant étale covering $\{U_i \to X\}_{i\in I}$ such that each $\underline{U_i}$ is an affine scheme with involution.*

Proof Let w be the involution on X. Choose an open neighborhood U_x of x in X such that $\underline{U_x}$ is an affine scheme. If $x = w(x)$, then the underlying scheme of $U_x \cap w(U_x)$ is an affine scheme with involution since X is separated. If $x \neq w(x)$, then the underlying scheme of $U_x \amalg w(U_x)$ is an affine scheme with involution. We finish the proof using [14, Corollary 2.19]. □

Let lSch denote the category of separated integral log schemes. Following Definition 2.5, we obtain the category $\mathrm{lSch}_{\mathbb{Z}/2}$, and we have the forgetful functor $i^* : \mathrm{lSch}_{\mathbb{Z}/2} \to \mathrm{lSch}$ with a left adjoint $i_\sharp$.

Definition 4.5 Let X be a separated integral log scheme with involution. Consider the presheaf $\mathrm{THR}|_X$ of $\mathbb{Z}/2$-spectra given by

$$\mathrm{THR}|_X(U) := \mathrm{THR}(\Gamma(U, \mathcal{O}_U), \Gamma(U, \mathcal{M}_U))$$

for every strict isovariant étale morphism $U \to X$. The *real topological Hochschild homology of X* is

$$\mathrm{THR}(X) := (L_{\mathrm{ениso\acute{e}t}}\mathrm{THR}|_X)(X) \in \mathrm{Sp}^{\mathbb{Z}/2},$$

where $L_{\mathrm{siso\acute{e}t}} : \mathrm{PSh}(\mathrm{lSch}_{\mathbb{Z}/2}, \mathrm{Sp}^{\mathbb{Z}/2}) \to \mathrm{Sh}_{\mathrm{siso\acute{e}t}}(\mathrm{lSch}_{\mathbb{Z}/2}, \mathrm{Sp}^{\mathbb{Z}/2})$ denotes the strict isovariant étale sheafification functor; see [21, Lemma 1.3.4.3] for sheafification functors.

Observe that THR is a strict isovariant étale sheaf by definition.

Theorem 4.6 *Let (A, P) be an integral log ring with involution. Then the induced map of $\mathbb{Z}/2$-spectra*

$$\mathrm{THR}(A, P) \to \mathrm{THR}(\mathrm{Spec}(A, P))$$

is an equivalence.

Proof Consider the presheaf $\mathrm{THR}|_{(A,P)}$ of $\mathbb{Z}/2$-spectra given by

$$\mathrm{THR}|_{(A,P)}(B) := \mathrm{THR}(B, P)$$

for every A-algebra B with involution such that $\mathrm{Spec}(A) \to \mathrm{Spec}(B)$ is isovariant étale. We have the induced map of presheaves

$$\mathrm{THR}|_{(A,P)} \to \mathrm{THR}|_{\mathrm{Spec}(A,P)}.$$

Observe that the left-hand side is an isovariant étale sheaf by Proposition 4.1. It suffices to show that this map of presheaves becomes an equivalence after sheafification since taking the global sections produces the desired equivalence.

For this, it suffices to show that the map of stalks is an equivalence. For every sheaf $\mathcal{F}$ and a point x of X, the stalk $\mathcal{F}_x$ is given by the filtered colimit $\mathrm{colim}_{U\ni x}\,\mathcal{F}(U)$. Hence by Propositions 3.8 and 4.4, we reduce to the case when A is a local ring with involution. In this case, we need to show that the induced map

$$\mathrm{THR}(A, P) \to \mathrm{THR}(A, P^a)$$

is an equivalence, where P^a is the logification of P, which is given by $P^a := P \oplus_{\theta^{-1}(A^*)} A^*$ if $\theta : P \to A$ is the structure map. This is a consequence of Lemma 4.3 for $P \to P^a$. □

In the next results, we explain how THR is related with THH under the functors i^* and i_*. We review the definition of THH of schemes in [7, Definition 8.3.7] as follows. We consider the presheaf of spectra $\mathrm{THH}|_X$ on $X_{\text{ét}}$ given by

$$\mathrm{THH}|_X(U) := \mathrm{THH}(\Gamma(U, \mathcal{O}_U), \Gamma(U, \mathcal{M}_U))$$

for $U \in X_{\text{ét}}$. The topological Hochschild homology of X is

$$\mathrm{THH}(X) := (L_{\text{ét}}\mathrm{THH}|_X)(X) \in \mathrm{Sp},$$

where $L_{\text{ét}}$ denotes the étale sheafification functor. The induced map

$$\mathrm{THH}(A, P) \simeq \mathrm{THH}(\mathrm{Spec}(A, P))$$

is an equivalence of spectra for every integral log ring (A, P) as Theorem 4.6.

Proposition 4.7 *Let X be a separated integral log scheme with involution. Then there is a natural equivalence of spectra*

$$\mathrm{THH}(i^*X) \simeq i^*\mathrm{THR}(X).$$

Proof By Proposition 4.4, we reduce to the case when $X = \mathrm{Spec}(A, P)$ for some log ring (A, P) with involution. Proposition 3.5 finishes the proof. □

Proposition 4.8 *Let X be a separated integral log scheme. Then there is a natural equivalence of $\mathbb{Z}/2$-spectra*

$$\mathrm{THR}(i_\sharp X) \simeq i_*\mathrm{THH}(X).$$

Proof As above, we reduce to the case when $X = \mathrm{Spec}(A, P)$ for some log ring (A, P). In this case, there is a natural isomorphism

$$i_\sharp \,\mathrm{Spec}(A, P) \simeq \mathrm{Spec}(i_* A, \underline{P}).$$

Proposition 3.7 finishes the proof. □

Corollary 4.9 *Let X be a separated integral log scheme with involution. Then there is a natural equivalence of spectra*

$$\mathrm{THR}(i_\sharp X)^{\mathbb{Z}/2} \simeq \mathrm{THH}(X).$$

Proof Combine Proposition 4.8 with [19, Proposition A.2.7(2)] to conclude. □

5 TCR of log schemes with involutions

Throughout this section, p is a prime number. This section is relying on [27; 28] due to Quigley and Shah, which we review as follows. Keep in mind that our C_p and $\mathbb{Z}/2$ correspond to their μ_p and C_2. See [26, §3] for another review.

A $\mathbb{Z}/2$-∞-*category* (resp. $\mathbb{Z}/2$-∞-*space*) is a presheaf of ∞-categories (resp. spaces) on the orbit category $\mathcal{O}_{\mathbb{Z}/2}$. Recall that $\mathcal{O}_{\mathbb{Z}/2}$ can be described as the diagram

$$\begin{tikzcd} (\mathbb{Z}/2)/e \arrow[loop left, "w"] \arrow[r, "\mathrm{res}"] & (\mathbb{Z}/2)/(\mathbb{Z}/2). \end{tikzcd}$$

Observe that we can naturally view a $\mathbb{Z}/2$-space as a $\mathbb{Z}/2$-category. A $\mathbb{Z}/2$-functor of $\mathbb{Z}/2$-categories is a morphism of presheaves. For $\mathbb{Z}/2$-categories $\mathcal{C}$ and $\mathcal{D}$, let $\mathrm{Fun}_{\mathbb{Z}/2}(\mathcal{C},\mathcal{D})$ be the ∞-category of $\mathbb{Z}/2$-functors $\mathcal{C}\to\mathcal{D}$.

Consider the $\mathbb{Z}/2$-∞-categories $\underline{\mathrm{Sp}}^{\mathbb{Z}/2}$ and $\underline{\mathrm{NAlg}}^{\mathbb{Z}/2}$ in [26, Example 3.2]. We have the equations $\underline{\mathrm{Sp}}^{\mathbb{Z}/2}((\mathbb{Z}/2)/e)=\mathrm{Sp}$, $\underline{\mathrm{Sp}}^{\mathbb{Z}/2}((\mathbb{Z}/2)/(\mathbb{Z}/2))=\mathrm{Sp}^{\mathbb{Z}/2}$, and $\underline{\mathrm{Sp}}^{\mathbb{Z}/2}(\mathrm{res})=i^*$. We also have a similar description for $\underline{\mathrm{NAlg}}^{\mathbb{Z}/2}$. Let BS^σ and BC_p^σ be the classifying $\mathbb{Z}/2$-spaces of S^σ and C_p^σ, where C_p^σ denotes the $\mathbb{Z}/2$-subspace of S^σ whose underlying space is C_p. We have the ∞-categories

$$\begin{aligned}(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} &:= \mathrm{Fun}_{\mathbb{Z}/2}(BS^\sigma,\underline{\mathrm{Sp}}^{\mathbb{Z}/2}),\\ (\mathrm{Sp}^{\mathbb{Z}/2})^{BC_p^\sigma} &:= \mathrm{Fun}_{\mathbb{Z}/2}(BC_p^\sigma,\underline{\mathrm{Sp}}^{\mathbb{Z}/2}).\end{aligned}$$

We similarly have the ∞-categories $(\mathrm{NAlg}^{\mathbb{Z}/2})^{BS^\sigma}$ and $(\mathrm{NAlg}^{\mathbb{Z}/2})^{BC_p^\sigma}$.

For equivariant homotopy theory, we refer to [32, §3] for the model-categorical approach (the group G can be a compact Lie group) and [2, §9] for the ∞-categorical approach (G is only a profinite group). For a compact Lie group G, let Sp^G be the ∞-category of G-spectra, which is the underlying ∞-category of orthogonal G-spectra. According to [28, Theorem A, Remark 1.8], we have the forgetful functors

$$\begin{aligned} j^*&:\mathrm{Sp}^{O(2)}\to(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma},\\ j^*&:\mathrm{Sp}^{D_{2p}}\to(\mathrm{Sp}^{\mathbb{Z}/2})^{BC_p^\sigma},\end{aligned}$$

which admit right adjoints j_*. We have the fixed point functors

$$\begin{aligned}(-)^{C_p}&:\mathrm{Sp}^{O(2)}\to\mathrm{Sp}^{O(2)/C_p}\simeq\mathrm{Sp}^{O(2)},\\ (-)^{C_p}&:\mathrm{Sp}^{D_{2p}}\to\mathrm{Sp}^{\mathbb{Z}/2}.\end{aligned}$$

We have the norm functor

$$N_{\mathbb{Z}/2}^{D_{2p}}:\mathrm{Sp}^{\mathbb{Z}/2}\to\mathrm{Sp}^{D_{2p}}.$$

We have the geometric fixed point functors

$$\begin{aligned}\Phi^{C_p}&:\mathrm{Sp}^{O(2)}\to\mathrm{Sp}^{O(2)/C_p}\simeq\mathrm{Sp}^{O(2)},\\ \Phi^{C_p}&:\mathrm{Sp}^{D_{2p}}\to\mathrm{Sp}^{\mathbb{Z}/2}.\end{aligned}$$

Recall from [28, Definition 1.6, Remark 1.8] that the *parametrized Tate constructions* are

$$\begin{aligned}(-)^{tC_p^\sigma} &:= j^*\Phi^{C_p} j_* : (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \to (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma},\\ (-)^{tC_p^\sigma} &:= \Phi^{C_p} j_* : (\mathrm{Sp}^{\mathbb{Z}/2})^{BC_p^\sigma} \to \mathrm{Sp}^{\mathbb{Z}/2}.\end{aligned}$$

For $X \in \mathrm{Sp}^{\mathbb{Z}/2}$, we often use the notation $X^{\otimes C_p^\sigma} := j^* N_{\mathbb{Z}/2}^{D_{2p}} X$. We have the parametrized Tate diagonal given by the composite natural map

$$\Delta : X \xrightarrow{\simeq} \Phi^{C_p} N_{\mathbb{Z}/2}^{D_{2p}} X \to \Phi^{C_p} j_* j^* N_{\mathbb{Z}/2}^{D_{2p}} X = (X^{\otimes C_p^\sigma})^{tC_p^\sigma}.$$

A *real cyclotomic spectrum* X is an object of $(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma}$ equipped with maps

$$\varphi_p : X \to X^{tC_p^\sigma}$$

in $(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma}$ for all primes p. In [27, Definition 1.20], the ∞-category of real cyclotomic spectra is defined to be the lax equalizer

$$\mathbb{R}\mathrm{CycSp} := \mathrm{LEq}\Big((\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \underset{\prod_p (-)^{tC_p^\sigma}}{\overset{\mathrm{id}}{\rightrightarrows}} \textstyle\prod_p (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma}\Big).$$

See [27, Remark 1.21] for the $\mathbb{Z}/2$-∞-category of real cyclotomic spectra $\underline{\mathbb{R}\mathrm{CycSp}}$. We refer to [27, Remark 4.3] for the functor

$$\mathrm{TCR} : \mathbb{R}\mathrm{CycSp} \to \mathrm{Sp}^{\mathbb{Z}/2}.$$

By [26, Proposition 3.7], the forgetful functor

$$q^* : (\mathrm{NAlg}^{\mathbb{Z}/2})^{BS^\sigma} \to \mathrm{NAlg}^{\mathbb{Z}/2}$$

admits a left adjoint $q_\otimes$. For $A \in \mathrm{NAlg}^{\mathbb{Z}/2}$, following [27, §5] (see also [26, Definition 3.11 and Proposition 3.12], we set

$$\mathrm{THR}(A) := q_\otimes A.$$

Recall from [27, §5] the following facts:

(1) There exists a unique map

$$\varphi_p : \mathrm{THR}(A) \to \mathrm{THR}(A)^{tC_p^\sigma}$$

such that the induced square

$$\begin{tikzcd} A \arrow[r] \arrow[d, "\Delta"'] & \mathrm{THR}(A) \arrow[d, "\varphi_p"] \\ (A^{\otimes C_p^\sigma})^{tC_p^\sigma} \arrow[r] & \mathrm{THR}(A)^{tC_p^\sigma} \end{tikzcd} \tag{5-1}$$

commutes, where the lower horizontal map is induced by the inclusion $C_p \to S^1$.

(2) φ_p is a map in $\mathrm{NAlg}^{\mathbb{Z}/2}$.

(3) As a consequence of (2), $\mathrm{THR}(A)$ is an object of $\mathrm{NAlg}(\mathbb{R}\mathrm{CycSp})$, where $\mathrm{NAlg}(\mathbb{R}\mathrm{CycSp})$ denotes the ∞-category of normed algebras in $\underline{\mathbb{R}\mathrm{CycSp}}$ defined as in [2, Definition 9.14]. We write down the definition as follows. For $X \in \mathrm{FinGpd}$, let Fin_X denote the category of finite coverings of X in FinGpd. We have a functor

$$\underline{\mathbb{R}\mathrm{CycSp}}^{\otimes} : \mathrm{Span}(\mathrm{Fin}_{B(\mathbb{Z}/2)}) \to \mathrm{Cat}_\infty$$

in [27, Point 3 after Remark 4.3]. A *normed algebra in* $\underline{\mathbb{R}\mathrm{CycSp}}$ is a section of $\underline{\mathbb{R}\mathrm{CycSp}}^{\otimes}$ over $\mathrm{Span}(\mathrm{Fin}_{B(\mathbb{Z}/2)})$ that is cocartesian over the backward morphisms in $\mathrm{Span}(\mathrm{Fin}_{B(\mathbb{Z}/2)})$. Recall that a backward morphism in a span ∞-category is a morphism of the form $X \leftarrow Y \xrightarrow{\mathrm{id}} Y$; see [2, Appendix C].

Proposition 5.1 *Let P be a commutative monoid with involution. Then for every prime p, there exists a natural equivalence of $\mathbb{Z}/2$-spaces*

$$(\mathrm{B}^{\mathrm{di}} P)^{C_p} \simeq \mathrm{B}^{\mathrm{di}} P. \tag{5-2}$$

Proof Let sd_p be the p-fold subdivision functor in [8]. For every integer $q \geq 0$, the set of q-simplices in $(\mathrm{sd}_p \mathrm{N}^{\mathrm{di}} P)^{C_p}$ is the set

$$\{(x_0, \dots, x_q, \dots, x_0, \dots, x_q) : x_0, \dots, x_q \in P\},$$

which is isomorphic to the set of q-simplices in $\mathrm{N}^{\mathrm{di}} P$. This isomorphism is also compatible with the involutions, so we obtain the desired equivalence. □

Proposition 5.2 *Let P be a commutative monoid with involution. Then for every prime p, there exists a natural equivalence of $\mathbb{Z}/2$-spaces*

$$(\mathrm{B}^{\mathrm{drep}} P)^{C_p} \simeq \mathrm{B}^{\mathrm{drep}} P. \tag{5-3}$$

Proof The isomorphism $\mathrm{N}^{\mathrm{di}} P^{\mathrm{gp}} \cong (\mathrm{sd}_p \mathrm{N}^{\mathrm{di}} P^{\mathrm{gp}})^{C_p}$ obtained by the proof of Proposition 5.1 can be restricted to an isomorphism $\mathrm{N}^{\mathrm{drep}} P \cong (\mathrm{sd}_p \mathrm{N}^{\mathrm{drep}} P)^{C_p}$. This produces the desired equivalence. □

Proposition 5.3 *Let P be a commutative monoid with involution. Then the Frobenius $\varphi_p : \mathrm{THR}(\mathbb{S}[P]) \to \mathrm{THR}(\mathbb{S}[P])^{tC_p^\sigma}$ is equivalent to the composite*

$$\mathbb{S}[\mathrm{B}^{\mathrm{di}} P] \xrightarrow{\simeq} \mathbb{S}[(\mathrm{B}^{\mathrm{di}} P)^{C_p}] \xrightarrow{\simeq} \Phi^{C_p} \mathbb{S}[\mathrm{B}^{\mathrm{di}} P] \to \mathbb{S}[\mathrm{B}^{\mathrm{di}} P]^{tC_p^\sigma}, \tag{5-4}$$

where the first map is obtained by (5-2).

Proof Consider the commutative square

$$\begin{tikzcd}
P \arrow[r] & \mathrm{B}^{\mathrm{di}} P \\
(P^{\oplus C_p})^{C_P} \arrow[u, "\simeq"] \arrow[r] & (\mathrm{B}^{\mathrm{di}} P)^{C_p} \arrow[u, "\simeq"']
\end{tikzcd}$$

where the right vertical arrow is obtained by (5-2), and the lower horizontal arrow is induced by the inclusion $C_p \to S^1$. After taking $\mathbb{S}[-]$, we obtain the commutative square

$$\begin{tikzcd}
\mathbb{S}[P] \arrow[r] & \mathbb{S}[\mathrm{B}^{\mathrm{di}}P] \\
\Phi^{C_p} N^{D_{2p}}_{\mathbb{Z}/2}\mathbb{S}[P] \arrow[r] \arrow[u, "\simeq"] & \Phi^{C_p}\mathbb{S}[\mathrm{B}^{\mathrm{di}}P] \arrow[u, "\simeq"']
\end{tikzcd}$$

Using the natural transformation $\Phi^{C_p} \to \Phi^{C_p} j_* j^*$, we obtain the commutative square

$$\begin{tikzcd}
\mathbb{S}[P] \arrow[r] \arrow[d] & \mathbb{S}[\mathrm{B}^{\mathrm{di}}P] \arrow[d] \\
(\mathbb{S}[P]^{\otimes C_p^\sigma})^{tC_p^\sigma} \arrow[r] & \mathbb{S}[\mathrm{B}^{\mathrm{di}}P]^{tC_p^\sigma}
\end{tikzcd}$$

Compare this with (5-1) to conclude. □

Definition 5.4 Let P be a commutative monoid with involution. We have the Frobenius $\varphi : \mathbb{S}[\mathrm{B}^{\mathrm{drep}}P] \to \mathbb{S}[\mathrm{B}^{\mathrm{drep}}P]^{tC_p^\sigma}$ given by the composite

$$\mathbb{S}[\mathrm{B}^{\mathrm{drep}}P] \xrightarrow{\simeq} \mathbb{S}[(\mathrm{B}^{\mathrm{drep}}P)^{C_p}] \xrightarrow{\simeq} \Phi^{C_p}\mathbb{S}[\mathrm{B}^{\mathrm{drep}}P] \to \mathbb{S}[\mathrm{B}^{\mathrm{drep}}P]^{tC_p^\sigma}, \tag{5-5}$$

where the first map is obtained by (5-3). Since the canonical map $\mathrm{B}^{\mathrm{di}}P \to \mathrm{B}^{\mathrm{drep}}P$ is a map of commutative monoids in the category of topological $O(2)$-spaces, we obtain a natural map

$$\mathbb{S}[\mathrm{B}^{\mathrm{di}}P] \to \mathbb{S}[\mathrm{B}^{\mathrm{drep}}P] \tag{5-6}$$

in $(\mathrm{NAlg}^{\mathbb{Z}/2})^{BS^\sigma}$. Compare (5-4) and (5-5) to see that (5-6) is compatible with φ_p. Hence we can promote (5-6) to a map in $\mathrm{NAlg}(\mathbb{R}\mathrm{CycSp})$.

Let (A, P) be a log ring with involution. We take the coproduct

$$\mathrm{THR}(A, P) := \mathrm{THR}(A) \wedge_{\mathbb{S}[\mathrm{B}^{\mathrm{di}}P]} \mathbb{S}[\mathrm{B}^{\mathrm{drep}}P]$$

in $\mathrm{NAlg}(\mathbb{R}\mathrm{CycSp})$. The *real topological cyclic homology of* (A, P) is

$$\mathrm{TCR}(A, P) := \mathrm{TCR}(\mathrm{THR}(A, P)).$$

Proposition 5.5 *The forgetful functor*

$$\mathbb{R}\mathrm{CycSp} \to \mathrm{Sp}^{\mathbb{Z}/2}$$

is conservative, exact, symmetric monoidal, and preserves colimits and finite limits.

Proof We refer to [27, Remark 4.3]. □

Definition 5.6 A morphism of log schemes with involutions is a *strict equivariant Nisnevich cover* if it is strict and its underlying morphism of schemes with involutions is an equivariant Nisnevich cover in the sense of Voevodsky [9, §3.1]. See also [14, §2].

The *strict equivariant Nisnevich topology* is the topology generated by strict equivariant Nisnevich covers. Let seNis be the shorthand for this topology. Observe that the strict equivariant Nisnevich topology is coarser than the strict isovariant étale topology.

Proposition 5.7 *Let (A, P) be a finite-dimensional noetherian log ring with involution. Then the presheaf* $\mathrm{THR}(-, P)$ *on the opposite category of A-algebras of finite type with involution is an equivariant Nisnevich sheaf of real cyclotomic spectra.*

Proof For an equivariant Nisnevich distinguished square Q (see [14, §2.1]) consisting of A-algebras of finite type with involution, we need to show that $\mathrm{THR}(Q, P)$ is a cocartesian square of real cyclotomic spectra. By Proposition 5.5, it suffices to show that $\mathrm{THR}(Q, P)$ is a cocartesian square of $\mathbb{Z}/2$-spectra. This is a consequence of Proposition 4.1. □

Remark 5.8 We do not know whether $\mathrm{THR}(-, P)$ in Proposition 5.7 is an isovariant étale sheaf of real cyclotomic spectra. An affirmative answer would remove the finite-dimensional noetherian assumption in Theorem 5.12 below.

Proposition 5.9 *Let X be a finite-dimensional noetherian separated log scheme with involution. Then there exists an equivariant Nisnevich covering $\{U_i \to X\}_{i \in I}$ such that each $\underline{U_i}$ is an affine scheme with involution.*

Proof This is a consequence of [14, Lemma 2.20]. □

Proposition 5.10 *The functor* THR *from the category of log rings with involutions to $\mathbb{Z}/2$-spectra preserves filtered colimits.*

Proof The is a consequence of Propositions 3.8 and 5.5. □

Definition 5.11 Let X be a finite-dimensional noetherian separated integral log scheme with involution. Consider the presheaf $\mathrm{THR}|_X$ of real cyclotomic spectra given by

$$\mathrm{THR}|_X(U) := \mathrm{THR}(\Gamma(U, \mathcal{O}_U), \Gamma(U, \mathcal{M}_U))$$

for every strict isovariant étale morphism $U \to X$. Consider

$$\mathrm{THR}(X) := (L_{\mathrm{seNis}} \mathrm{THR}|_X)(X) \in \mathbb{R}\mathrm{CycSp},$$

where L_{seNis} denotes the strict equivariant Nisnevich sheafification functor.

Observe that THR is a strict equivariant Nisnevich sheaf of real cyclotomic spectra by definition. Furthermore, if we forget the real cyclotomic structure on $\mathrm{THR}(X)$, then we recover $\mathrm{THR}(X)$ in Definition 4.5 by Theorem 5.12 below.

The *real topological cyclic homology of X* is

$$\mathrm{TCR}(X) := \mathrm{TCR}(\mathrm{THR}(X)).$$

Theorem 5.12 *Let (A, P) be a finite-dimensional noetherian integral log ring with involution. Then the induced map*

$$\mathrm{THR}(A, P) \to \mathrm{THR}(\mathrm{Spec}(A, P))$$

is an equivalence of real cyclotomic spectra. In particular, the induced map

$$\mathrm{TCR}(A, P) \to \mathrm{TCR}(\mathrm{Spec}(A, P))$$

is an equivalence of $\mathbb{Z}/2$-spectra.

Proof Argue as in the proof of Theorem 4.6, but use instead Propositions 5.7, 5.9, and 5.10 to reduce to the case when A is a local ring. In this case, we need to show that the induced map

$$\mathrm{THR}(A, P) \to \mathrm{THR}(A, P^a)$$

is an equivalence of real cyclotomic spectra, where $P^a := P \oplus_{\theta^{-1}(A^*)} A^*$ if $\theta : P \to A$ is the structure map. This is a consequence of Theorem 4.6 and Proposition 5.5. □

The purpose of the remaining part of this section is to show Proposition 5.30, which is needed for Proposition 6.17. For this, we will check that various squares commute.

The map of $\mathbb{Z}/2$-spaces $q : * \to BS^\sigma$ induces the forgetful functor

$$q^* : (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \to \mathrm{Sp}^{\mathbb{Z}/2},$$

which is conservative by [26, Proposition 3.8] and admits a left adjoint $q_\sharp$ and right adjoint q_* by [26, Proposition 3.7].

Proposition 5.13 *There are natural equivalences*

$$q^* q_\sharp \simeq \Sigma^\infty S^\sigma_+ \wedge (-), \tag{5-7}$$

$$q^* q_* \simeq \Sigma^{-\sigma} \Sigma^\infty S^\sigma_+ \wedge (-). \tag{5-8}$$

Proof For (5-7), consider the cartesian square

$$\begin{tikzcd} S^\sigma \arrow[r, "r"] \arrow[d, "r"'] & * \arrow[d, "q"] \\ * \arrow[r, "q"] & BS^\sigma \end{tikzcd}$$

where $r : S^\sigma \to *$ is the unique map. By [28, Lemma 4.3], we have a natural equivalence $q^* q_\sharp \simeq r_\sharp r^*$. To obtain (5-7), observe that the projection formula [28, Lemma 5.44] yields a natural equivalence $r_\sharp r^* \simeq \Sigma^\infty S^\sigma_+ \wedge (-)$.

We have $\mathrm{fib}(\mathrm{id} \to q^* q_\sharp) \simeq \Sigma^\sigma$. By taking right adjoints, we have $\mathrm{cofib}(q^* q_* \to \mathrm{id}) \simeq \Sigma^{-\sigma}$, which yields (5-8). □

Remark 5.14 We similarly have the conservative forgetful functor

$$q^* : \mathrm{Sp}^{BS^1} \to \mathrm{Sp}$$

with a left adjoint $q_\sharp$ and right adjoint q_*. There are also natural equivalences

$$q^*q_\sharp \simeq \Sigma^\infty S^1_+ \wedge (-),$$
$$q^*q_* \simeq \Sigma^{-1}\Sigma^\infty S^1_+ \wedge (-).$$

Proposition 5.15 *There are induced commutative squares*

$$\begin{tikzcd}
\mathrm{Sp}^{\mathbb{Z}/2} \arrow[r, "q_\sharp"] \arrow[d, "i^*"'] & (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[d, "i^*"] \\
\mathrm{Sp} \arrow[r, "q_\sharp"] & \mathrm{Sp}^{BS^1}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathrm{Sp}^{\mathbb{Z}/2} \arrow[r, "q_*"] \arrow[d, "i^*"'] & (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[d, "i^*"] \\
\mathrm{Sp} \arrow[r, "q_*"] & \mathrm{Sp}^{BS^1}
\end{tikzcd}$$

Proof We focus on the first square since the proofs are similar. Since $q^* : \mathrm{Sp}^{BS^1} \to \mathrm{Sp}$ is conservative, it suffices to show that the composite natural transformation

$$q^*q_\sharp i^* \to q^*i^*q_\sharp \xrightarrow{\simeq} i^*q^*q_\sharp$$

is an isomorphism. Using Proposition 5.13 and Remark 5.14, this natural transformation can be identified with the natural transformation

$$\Sigma^\infty S^1_+ \wedge i^* \to i^*(\Sigma^\infty S^\sigma_+ \wedge (-)).$$

This is an equivalence since $i^*\Sigma^\infty S^\sigma \simeq \Sigma^\infty S^1$. □

Let $v : BS^\sigma \to B(S^\sigma/C_p^\sigma) \simeq BS^\sigma$ be the functor induced by the quotient map $S^\sigma \to S^\sigma/C_p^\sigma$. We have the induced functor

$$v^* : (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \to (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma},$$

whose left adjoint is $(-)_{hC_p^\sigma}$ and right adjoint is $(-)^{hC_p^\sigma}$, which correspond to $(-)_{hC_2 S^1}$ and $(-)^{hC_2 S^1}$ used in [28, Example 5.57]. Similarly, we have the induced functor

$$v^* : \mathrm{Sp}^{BS^1} \to \mathrm{Sp}^{BS^1},$$

whose left adjoint is $(-)_{hC_p}$ and right adjoint is $(-)^{hC_p}$, which are used in [24, Theorem 1.3].

Proposition 5.16 *There are induced commutative squares*

$$\begin{tikzcd}
\mathrm{Sp}^{BS^1} \arrow[r, "v^*"] \arrow[d, "i_\sharp"'] & \mathrm{Sp}^{BS^1} \arrow[d, "i_\sharp"] \\
(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[r, "v^*"] & (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathrm{Sp}^{BS^1} \arrow[r, "v^*"] \arrow[d, "i_*"'] & \mathrm{Sp}^{BS^1} \arrow[d, "i_*"] \\
(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[r, "v^*"] & (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma}
\end{tikzcd}$$

Proof We focus on the first square since the proofs are similar. We have the induced commutative diagram

$$\begin{tikzcd}
i_\sharp q^* v^* \arrow[r] \arrow[d, "\simeq"'] & q^* i_\sharp v^* \arrow[r] & q^* v^* i_\sharp \arrow[d, "\simeq"] \\
i_\sharp q^* \arrow[rr] & & q^* i_\sharp
\end{tikzcd}$$

The vertical arrows are equivalences since the composite $* \xrightarrow{q} BS^\sigma \xrightarrow{v} BS^\sigma$ agrees with q and the same claim holds for BS^1 too. The lower horizontal and upper right horizontal arrows are equivalences by Proposition 5.15. Hence the upper left horizontal arrow is an equivalence. To conclude, use the fact that $q^* : \mathrm{Sp}^{BS^1} \to \mathrm{Sp}$ is conservative. □

Let $u : BS^\sigma \to *$ be the unique map. We have the induced functor $u^* : \mathrm{Sp}^{\mathbb{Z}/2} \to (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma}$ whose left adjoint is $(-)_{hS^\sigma}$ and right adjoint is $(-)^{hS^1}$. Similarly, we have the induced functor $u^* : \mathrm{Sp} \to \mathrm{Sp}^{BS^1}$ whose left adjoint is $(-)_{hS^1}$ and right adjoint is $(-)^{hS^1}$.

Proposition 5.17 *There are induced commutative squares*

$$\begin{tikzcd}
\mathrm{Sp} \arrow[r, "u^*"] \arrow[d, "i_\sharp"'] & \mathrm{Sp}^{BS^1} \arrow[d, "i_\sharp"] \\
\mathrm{Sp}^{\mathbb{Z}/2} \arrow[r, "u^*"] & (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathrm{Sp} \arrow[r, "u^*"] \arrow[d, "i_*"'] & \mathrm{Sp}^{BS^1} \arrow[d, "i_*"] \\
\mathrm{Sp}^{\mathbb{Z}/2} \arrow[r, "u^*"] & (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma}
\end{tikzcd}$$

Proof Argue as in Proposition 5.16. □

Proposition 5.18 *There are induced commutative squares*

$$\begin{tikzcd}
(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[r, "(-)^{tC_p^\sigma}"] \arrow[d, "i^*"'] & (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[d, "i^*"] \\
\mathrm{Sp}^{BS^1} \arrow[r, "(-)^{tC_p}"] & \mathrm{Sp}^{BS^1}
\end{tikzcd}
\qquad
\begin{tikzcd}
(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[r, "(-)^{tS^\sigma}"] \arrow[d, "i^*"'] & \mathrm{Sp}^{\mathbb{Z}/2} \arrow[d, "i^*"] \\
\mathrm{Sp}^{BS^1} \arrow[r, "(-)^{tS^1}"] & \mathrm{Sp}
\end{tikzcd}$$

Proof Consider the appropriate adjoint squares of the squares in Proposition 5.16, and compare the cofiber sequences $(-)_{hC_p} \to (-)^{hC_p} \to (-)^{tC_p}$ and $(-)_{hC_p^\sigma} \to (-)^{hC_p^\sigma} \to (-)^{tC_p^\sigma}$ to obtain the left square. Argue similarly for the right square, but use Proposition 5.17 instead. □

Construction 5.19 The motivic purity transformation [1] can be adapted to the equivariant homotopy theory as follows. Let $f : X \to S$ be a finite covering in FinGpd. Consider the induced diagram

$$\begin{tikzcd}
X \arrow[r, "a"] & X \times_S X \arrow[r, "p_2"] \arrow[d, "p_1"'] & X \arrow[d, "f"] \\
& X \arrow[r, "f"] & S
\end{tikzcd}$$

where a is the diagonal morphism, and p_1 (resp. p_2) is the first (resp. second) projection. We set $\Sigma_f := p_{2\sharp}a_*$. We have the natural transformation

$$f_\sharp \xrightarrow{\mathfrak{p}_f} f_*\Sigma_f : \mathrm{SH}(X) \to \mathrm{SH}(S)$$

given by the composite

$$f_\sharp \xrightarrow{\simeq} f_\sharp p_{1*}a_* \to f_* p_{2\sharp}a_*.$$

Recall that $i : * \to B(\mathbb{Z}/2)$ denotes the unique morphism.

Proposition 5.20 *The functor* Σ_i *is equivalent to the identity functor, and the natural transformation* $\mathfrak{p}_i : i_\sharp \to i_*\Sigma_i$ *is an equivalence.*

Proof Let $r : \mathbb{Z}/2 \to *$ be the unique map of finite groupoids, and let $a : * \to \mathbb{Z}/2$ be the inclusion to the first point. The functor $a_* : \mathrm{Sp} \to \mathrm{Sp} \times \mathrm{Sp}$ is $(\mathrm{id}, 0)$, and the functor $r_\sharp : \mathrm{Sp} \times \mathrm{Sp} \to \mathrm{Sp}$ is the direct sum functor. Hence $\Sigma_i := r_\sharp a_*$ is equivalent to the identity functor.

The pair of functors $(i^*, \Phi^{\mathbb{Z}/2})$ is conservative. Since $\Phi^{\mathbb{Z}/2} i_\sharp \simeq \Phi^{\mathbb{Z}/2} i_* \simeq 0$ by Proposition A.2.7(3), (5) in [19], it suffices to show that $i^*\mathfrak{p}_i$ is an equivalence. Using Proposition 5.15, it suffices to show that $\mathfrak{p}_r$ is an equivalence. This can be shown directly using the following description in [19, Proposition A.1.5]: for a finite set X with n elements, $\mathrm{SH}(X)$ is equivalent to the product of n copies of Sp. □

Construction 5.21 Consider the composite functor

$$\Sigma_i : \mathrm{Sp}^{BS^1} \xrightarrow{a_*} \mathrm{Sp}^{BS^1} \times \mathrm{Sp}^{BS^1} \xrightarrow{r_\sharp} \mathrm{Sp}^{BS^1}, \tag{5-9}$$

where $a_* := (\mathrm{id}, 0)$, and $r_\sharp$ is the direct sum functor. As in Construction 5.19, we have the natural transformation

$$i_\sharp \xrightarrow{\mathfrak{p}_f} i_*\Sigma_i : (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \to \mathrm{Sp}^{BS^1}. \tag{5-10}$$

Proposition 5.22 *The functor* (5-9) *is equivalent to the identity functor, and the natural transformation* (5-10) *is an equivalence.*

Proof The descriptions of a_* and $r_\sharp$ show that (5-9) is equivalent to the identity functor. Since $q^* : \mathrm{Sp}^{BS^1} \to \mathrm{Sp}$ is conservative, the second claim is reduced to Proposition 5.20 by Proposition 5.15. □

Proposition 5.23 *There are induced commutative squares*

$$\begin{tikzcd}
\mathrm{Sp}^{BS^1} \arrow[r, "(-)_{hC_p}"] \arrow[d, "i_\sharp"'] & \mathrm{Sp}^{BS^1} \arrow[d, "i_\sharp"] \\
(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[r, "(-)_{hC_p^\sigma}"] & (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathrm{Sp}^{BS^1} \arrow[r, "(-)^{hC_p}"] \arrow[d, "i_*"'] & \mathrm{Sp}^{BS^1} \arrow[d, "i_*"] \\
(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[r, "(-)^{hC_p^\sigma}"] & \mathrm{Sp}^{\mathbb{Z}/2}
\end{tikzcd}$$

$$\begin{tikzcd}
\mathrm{Sp}^{BS^1} \arrow[r, "(-)_{hS^1}"] \arrow[d, "i_\sharp"'] & \mathrm{Sp} \arrow[d, "i_\sharp"] \\
(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[r, "(-)_{hS^\sigma}"] & \mathrm{Sp}^{\mathbb{Z}/2}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathrm{Sp}^{BS^1} \arrow[r, "(-)^{hS^1}"] \arrow[d, "i_*"'] & \mathrm{Sp} \arrow[d, "i_*"] \\
(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[r, "(-)^{hS^\sigma}"] & \mathrm{Sp}^{\mathbb{Z}/2}
\end{tikzcd}$$

Proof The right squares commute since their left adjoint squares commute. Together with Propositions 5.20 and 5.22, we see that the left squares commute too. □

Proposition 5.24 *There are induced commutative squares*

$$\begin{tikzcd}
\mathrm{Sp}^{BS^1} \arrow[r, "(-)^{tC_p}"] \arrow[d, "i_\sharp"'] & \mathrm{Sp}^{BS^1} \arrow[d, "i_\sharp"] & \mathrm{Sp}^{BS^1} \arrow[r, "(-)^{tC_p}"] \arrow[d, "i_*"'] & \mathrm{Sp}^{BS^1} \arrow[d, "i_*"] \\
(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[r, "(-)^{tC_p^\sigma}"] & (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} & (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[r, "(-)^{tC_p^\sigma}"] & (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma}
\end{tikzcd}$$

Proof Argue as in Proposition 5.18, but use Proposition 5.23 instead. □

Proposition 5.25 *There is an adjunction*

$$i^* : \mathbb{R}\mathrm{CycSp} \rightleftarrows \mathrm{CycSp} : i_*$$

satisfying the following properties:

(1) *For* $Y \in \mathbb{R}\mathrm{CycSp}$, *the Frobenius* $i^*Y \to (i^*Y)^{tC_p}$ *is identified with* $i^*\varphi$ *if* $\varphi : Y \to Y^{tC_p^\sigma}$ *is the Frobenius.*

(2) *For* $X \in \mathrm{CycSp}$, *the Frobenius* $i_*X \to (i_*X)^{tC_p^\sigma}$ *is identified with* $i_*\varphi$ *if* $\varphi : X \to X^{tC_p}$ *is the Frobenius.*

(3) i^* *is symmetric monoidal.*

Proof The functor $i^* : \mathbb{R}\mathrm{CycSp} \to \mathrm{CycSp}$ is obtained by taking lax equalizers to the rows of the diagram

$$\begin{tikzcd}
(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[r, shift left, "\mathrm{id}"] \arrow[r, shift right, "\prod_p(-)^{tC_p^\sigma}"'] \arrow[d, "i^*"'] & \prod_p(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[d, "i^*"] \\
\mathrm{Sp}^{BS^1} \arrow[r, shift left, "\mathrm{id}"] \arrow[r, shift right, "\prod_p(-)^{tC_p}"'] & \prod_p \mathrm{Sp}^{BS^1}
\end{tikzcd} \tag{5-11}$$

whose two squares commute by Proposition 5.18. The two squares in the induced diagram

$$\begin{tikzcd}
\mathrm{Sp}^{BS^1} \arrow[r, shift left, "\mathrm{id}"] \arrow[r, shift right, "\prod_p(-)^{tC_p}"'] \arrow[d, "i_*"'] & \prod_p \mathrm{Sp}^{BS^1} \arrow[d, "i_*"] \\
(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[r, shift left, "\mathrm{id}"] \arrow[r, shift right, "\prod_p(-)^{tC_p^\sigma}"'] & \prod_p(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma}
\end{tikzcd} \tag{5-12}$$

commute by Proposition 5.24. Hence we see that a right adjoint i_* is obtained by taking lax equalizers to the rows of (5-12) since we can check the counit-unit identities pointwise. This implies the claims (1) and (2).

Recall from [24, Construction IV.2.1(ii)] the following fact: Let $\mathcal{C}$, $\mathcal{D}$, and $\mathcal{E}$ be symmetric monoidal ∞-categories. If $F : \mathcal{C} \to \mathcal{D}$ (resp. $G : \mathcal{C} \to \mathcal{D}$) is a symmetric (resp. lax symmetric) monoidal functor, then having a symmetric monoidal functor $\mathcal{E} \to \mathrm{LEq}(F, G)$ is equivalent to having a symmetric monoidal functor $H : \mathcal{E} \to \mathcal{F}$ and a lax symmetric monoidal transformation $F \circ H \to G \circ H$.

The canonical functor $\mathbb{R}\mathrm{CycSp} \to (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma}$ is symmetric monoidal by the above paragraph, and one can show that the forgetful functor $i^* : (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \to \mathrm{Sp}^{BS^1}$ is symmetric monoidal using the descriptions of the symmetric monoidal structures on Sp^{BS^1} and $(\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma}$ obtained by [24, Construction IV.2.1(1); 28, Example 5.13]. Hence the composite functor $\mathbb{R}\mathrm{CycSp} \to \mathrm{Sp}^{BS^1}$ is symmetric monoidal too. Furthermore, we have the lax symmetric natural transformation of the two composite functors

$$\begin{tikzcd} \mathbb{R}\mathrm{CycSp} \arrow[r] & (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma} \arrow[r, "i^*"] & \mathrm{Sp}^{\mathbb{Z}/2} \arrow[r, shift left, "\mathrm{id}"] \arrow[r, shift right, "\prod_p (-)^{tC_p}"'] & \prod_p \mathrm{Sp}^{BS^1} \end{tikzcd}$$

using (5-11) and the above paragraph. We deduce the claim (3) using the above paragraph again. □

Proposition 5.26 *For* $X \in \mathrm{CycSp}$ *and* $Y \in \mathbb{R}\mathrm{CycSp}$, *there are natural equivalences*

$$\begin{aligned} i^*\mathrm{TCR}(Y) &\simeq \mathrm{TC}(i^*Y), \\ i_*\mathrm{TC}(X) &\simeq \mathrm{TCR}(i_*X). \end{aligned}$$

Proof Using the descriptions of $i^*Y \to (i^*Y)^{tC_p}$ and $i_*X \to (i_*X)^{tC_p^\sigma}$ in Proposition 5.25, the claim follows from [24, Corollary 1.5; 27, Theorem C(3)], and Propositions 5.17 and 5.23. □

Proposition 5.27 *Let* (A, P) *be a log ring with involution. Then there is a natural equivalence of cyclotomic spectra*

$$i^*\mathrm{THR}(A, P) \simeq \mathrm{THH}(i^*A, i^*P)$$

and hence an equivalence of spectra

$$i^*\mathrm{TCR}(A, P) \simeq \mathrm{TCR}(i^*A, i^*P).$$

Proof We have a natural equivalence of spectra $i^*\mathrm{THR}(A) \simeq \mathrm{THH}(i^*A)$ by [19, Proposition 3.4.7]. Apply i^* to (5-1), and compare this with the square in [24, p. 342] to show that $i^*\mathrm{THR}(A) \xrightarrow{i^*\varphi_p} i^*(\mathrm{THR}(A)^{tC_p^\sigma})$ can be identified with $\mathrm{THH}(A) \xrightarrow{\varphi_p} \mathrm{THH}(A)^{tC_p}$. We also need Proposition 5.18 here. Together with Proposition 5.25(2), we have an equivalence of cyclotomic spectra $i^*\mathrm{THR}(A) \simeq \mathrm{THH}(A)$.

On the other hand, if we apply i^* to (5-4) and (5-5), then we get φ_p for $\mathbb{S}[\mathrm{B}^{\mathrm{cy}}P]$ and $\mathbb{S}[\mathrm{B}^{\mathrm{rep}}P]$ by Proposition 2.15 by Proposition 5.18. Together with Proposition 5.25(2), we have equivalences of cyclotomic spectra $i^*\mathbb{S}[\mathrm{B}^{\mathrm{di}}P] \simeq \mathbb{S}[\mathrm{B}^{\mathrm{cy}}P]$ and $i^*\mathbb{S}[\mathrm{B}^{\mathrm{drep}}P] \simeq \mathbb{S}[\mathrm{B}^{\mathrm{rep}}P]$. Proposition 5.25(3) finishes the proof. □

Proposition 5.28 *There is a commutative square*

$$\begin{tikzcd} \mathrm{CycSp} \arrow[r, "i_*"] \arrow[d] & \mathbb{R}\mathrm{CycSp} \arrow[d] \\ \mathrm{Sp} \arrow[r, "i_*"] & \mathrm{Sp}^{\mathbb{Z}/2} \end{tikzcd}$$

where the vertical arrows are the forgetful functors.

Proof By Proposition 5.16, it suffices to show that there is a commutative square

$$\begin{tikzcd}
\mathrm{CycSp} \arrow[r, "i_*"] \arrow[d] & \mathbb{R}\mathrm{CycSp} \arrow[d] \\
\mathrm{Sp}^{BS^1} \arrow[r, "i_*"] & (\mathrm{Sp}^{\mathbb{Z}/2})^{BS^\sigma}
\end{tikzcd}$$

where the vertical arrows are the forgetful functors. This can be obtained using (5-12). □

Proposition 5.29 *Let (A, P) be a log ring. Then there is a natural equivalence of real cyclotomic spectra*

$$\mathrm{THR}(i_*A, \underline{P}) \simeq i_*\mathrm{THH}(A, P).$$

Proof Argue as in Construction 3.6 but use Proposition 5.27 instead to construct a natural map of real cyclotomic spectra $\mathrm{THR}(i_*A, \underline{P}) \to i_*\mathrm{THR}(A, P)$. By Proposition 5.5, it suffices to show that this becomes an equivalence of $\mathbb{Z}/2$-spectra after applying the forgetful functor $\mathbb{R}\mathrm{CycSp} \to \mathrm{Sp}^{\mathbb{Z}/2}$. This is a consequence of Propositions 3.7 and 5.28. □

Proposition 5.30 *Let X be a finite-dimensional noetherian separated integral log scheme. Then there is a natural equivalence of real cyclotomic spectra*

$$\mathrm{THR}(i_\sharp X) \simeq i_*\mathrm{THH}(X)$$

and hence an equivalence of $\mathbb{Z}/2$-spectra

$$\mathrm{TCR}(i_\sharp X) \simeq i_*\mathrm{TC}(X)$$

and an equivalence of spectra

$$\mathrm{TCR}(i_\sharp X)^{\mathbb{Z}/2} \simeq \mathrm{TC}(X).$$

Proof Argue as in Proposition 4.8, but use Propositions 5.9 and 5.29 to obtain the first equivalence. For the remaining equivalences, use Proposition 5.26 and [19, Proposition A.2.7(2)]. □

We will use the notation

$$\mathrm{THR}^{\mathbb{Z}/2}(-) := \mathrm{THR}(-)^{\mathbb{Z}/2}, \quad \mathrm{TCR}^{\mathbb{Z}/2}(-) := \mathrm{TCR}(-)^{\mathbb{Z}/2}.$$

6 Motivic representability

Throughout this section, we fix a finite-dimensional noetherian separated scheme S (with the trivial log structure and involution). The purpose of this section is to represent $\mathrm{THH}^{\mathbb{Z}/2}$ and $\mathrm{TCR}^{\mathbb{Z}/2}$ in $\mathrm{logSH}(S)$ and THR and TCR in a $\mathbb{Z}/2$-equivariant analogue of $\mathrm{logSH}(S)$.

We begin with recalling the ingredients for defining $\mathrm{logSH}(S)$. Let SmlSm/S denote the category of fs log schemes Y of finite type over S such that $Y \to S$ is log smooth and $\underline{Y} \to S$ is smooth. A morphism in SmlSm/S is a *strict Nisnevich cover* if it is strict and its underlying morphism of schemes is a Nisnevich cover. We have the ∞-category of strict Nisnevich sheaves of spectra $\mathrm{Sh}_{\mathrm{sNis}}(\mathrm{SmlSm}/S, \mathrm{Sp})$.

For a smooth scheme X of finite type over S with a strict normal crossing divisor D, let (X, D) denote the fs log scheme with the underlying scheme X and the compactifying log structure associated with the open immersion $X - D \to X$. By [6, Lemma A.5.10], every object of SmlSm$/S$ arises as this form. We often regard $\mathbb{P}^{n-1}$ as a closed complement of $\mathbb{A}^n$ in $\mathbb{P}^n$, and we form $(\mathbb{P}^n, \mathbb{P}^{n-1})$. We set $\square := (\mathbb{P}^1, \infty)$.

Now, the ∞*-category of logarithmic motivic* S^1*-spectra* is defined to be the localization

$$\mathrm{logSH}_{S^1}(S) := (\mathbb{P}^\bullet, \mathbb{P}^{\bullet-1})^{-1}\mathrm{Sh}_{\mathrm{sNis}}(\mathrm{SmlSm}/S, \mathrm{Sp}),$$

where $(\mathbb{P}^\bullet, \mathbb{P}^{\bullet-1})^{-1}$ denotes the class of projections $(\mathbb{P}^n, \mathbb{P}^{n-1}) \times X \to X$ for all $X \in \mathrm{SmlSm}/S$ and integers $n \geq 1$. This is one of the various models of $\mathrm{logSH}_{S^1}(S)$ in [7, §3.4]. The ∞*-category of logarithmic motivic spectra* is defined to be the ∞-category of $\mathbb{P}^1$-spectra in $\mathrm{logSH}_{S^1}(S)$, that is,

$$\mathrm{logSH}(S) := \mathrm{Sp}_{\mathbb{P}^1}(\mathrm{logSH}_{S^1}(S)).$$

Remark 6.1 In this section, we often compute $\mathrm{THR}(X \times Y)$ in terms of $\mathrm{THR}(X)$ for noetherian separated fs log schemes X and Y such that there is a Zariski covering $\{\mathrm{Spec}(\mathbb{Z}[M_i], P_i) \to Y\}_{i \in I}$ for some homomorphisms of monoids $P_i \to M_i$. In view of Proposition 3.3, the base X is irrelevant for the computation. We will often argue as if everything takes place on the spectral scheme $\mathrm{Spec}(\mathbb{S})$ instead of X for notational convenience.

Recall that S^σ is S^1 with the involution given by $e^{i\theta} \in S^1 \mapsto e^{-i\theta}$. We have the functor

$$\Sigma^\sigma := \Sigma^\infty S^\sigma \wedge (-) : \mathrm{Sp}^{\mathbb{Z}/2} \to \mathrm{Sp}^{\mathbb{Z}/2}.$$

Let us recall the computation of THR for the projective spaces with the trivial involutions as follows.

Proposition 6.2 *Let X be a noetherian separated fs log scheme. Then there exists a natural equivalence of $\mathbb{Z}/2$-spectra*

$$\mathrm{THR}(X \times \mathbb{P}^n) \simeq \begin{cases} \mathrm{THR}(X) \oplus \bigoplus_{j=1}^{\lfloor n/2 \rfloor} i_*\mathrm{THH}(X) & \text{if } n \text{ is even,} \\ \mathrm{THR}(X) \oplus \bigoplus_{j=1}^{\lfloor n/2 \rfloor} i_*\mathrm{THH}(X) \oplus \Sigma^{n(\sigma-1)}\mathrm{THR}(X) & \text{if } n \text{ is odd.} \end{cases}$$

Proof Following Remark 6.1, we can argue as if everything takes place on $\mathbb{S}$ instead of X. Then the claim is due to [19, Theorem 5.2.6]. □

Our next goal is to show that $\mathrm{THR}^{\mathbb{Z}/2}$ and $\mathrm{TCR}^{\mathbb{Z}/2}$ are representable in $\mathrm{logSH}_{S^1}(S)$. For this, we need the notion of cubes [22, Definition 6.1.1.2], which we recall as follows.

For a set I, let $\mathbf{P}(I)$ denote the set of subsets of I. We impose the partially ordered set structure on $\mathbf{P}(I)$ with respect to the inclusion, and we regard $\mathbf{P}(I)$ as the associated category. For an ∞-category $\mathcal{C}$, an *I-cube in $\mathcal{C}$* is a functor

$$Q : \mathbf{P}(I) \to \mathcal{C}.$$

If the cardinal of I is an integer n, then an I-cube is called an *n-cube*. The *total cofiber of Q* is

$$\mathrm{tcofib}(Q) := \mathrm{cofib}(\mathrm{colim}(Q|_{\mathbf{P}(I)-\{I\}}) \to Q(I)),$$

whenever the colimit and cofiber exist. The *total fiber of Q* is

$$\operatorname{tfib}(Q) := \operatorname{fib}(Q(\varnothing) \to \lim(Q|_{\mathbf{P}(I)-\{\varnothing\}})),$$

whenever the limit and fiber exist.

Theorem 6.3 *Let X be a noetherian separated fs log scheme with involution. Then the induced map*

$$\mathrm{THR}(X) \to \mathrm{THR}(X \times (\mathbb{P}^n, \mathbb{P}^{n-1}))$$

is an equivalence of real cyclotomic spectra, where we impose the trivial involution on $(\mathbb{P}^n, \mathbb{P}^{n-1})$.

Proof By Proposition 5.5, it suffices to show that the map is an equivalence of $\mathbb{Z}/2$-spectra.

Following Remark 6.1, we will argue for $\mathbb{S}$ instead of X. Let $\underline{U_0}, \dots, \underline{U_n}$ be the standard cover of $\mathbb{P}^n_{\mathbb{S}}$, and we set $U_i := (\mathbb{P}^n_{\mathbb{S}}, \mathbb{P}^{n-1}_{\mathbb{S}}) \times_{\mathbb{P}^n_{\mathbb{S}}} \underline{U_i}$. For every nonempty subset $I \subset [n]$, we set $U(I) := U_I := \bigcap_{i \in I} U_i$. We also set $U_\varnothing := \mathbb{S}$. We regard U as an $(n+1)$-cube. We need to show $\operatorname{tcofib}(\mathrm{THR}(U)) \simeq 0$.

For notational convenience, we consider the commutative monoids

$$\begin{aligned}
P_j &:= \{(x_1, \dots, x_n) \in \mathbb{Z}^n : x_j \geq 0\} \quad \text{for } j = 1, \dots, n, \\
P_0 &:= \{(x_1, \dots, x_n) \in \mathbb{Z}^n : x_1 + \dots + x_n \leq 0\}, \\
F_0 &:= \{(x_1, \dots, x_n) \in \mathbb{Z}^n : x_1 + \dots + x_n = 0\}.
\end{aligned}$$

For every subset $J \subset [n]$, we set $P_J := \bigcap_{j \in J} P_j$. By convention, we have $P_\varnothing := \mathbb{Z}^n$. There is a canonical equivalence

$$\mathrm{THR}(U_I) \simeq \begin{cases} \mathrm{THR}(\mathbb{S}[P_{[n]-I}]) & \text{if } 0 \in I, \\ \mathrm{THR}(\mathbb{S}[P_{[n]-I}], P_{[n]-I} \cap F_0) & \text{if } 0 \notin I. \end{cases}$$

Each P_J is isomorphic to products of finite copies of $\mathbb{Z}$ and $\mathbb{N}$. The computations of $\mathrm{B}^{\mathrm{di}}\mathbb{N}$ and $\mathrm{B}^{\mathrm{di}}\mathbb{Z}$ in [19, (4.12), Propositions 4.2.11, 4.2.12] and the fact B^{di} preserves finite products [19, Proposition 4.2.4] yield a natural equivalence of $\mathbb{Z}/2$-spectra

$$\mathrm{THR}(U_I) \simeq \mathbb{S}[\Phi(I)]$$

with

$$\begin{aligned}
\Phi(I) &:= \coprod_{(x_1,\dots,x_n) \in \mathbb{Z}^n} \Phi(I; (x_1, \dots, x_n)), \\
\Phi(I; (x_1, \dots, x_n)) &:= \begin{cases} \Phi_1(I; x_1) \times \dots \times \Phi_n(I; x_n) & \text{if } (x_1, \dots, x_n) \in P_{[n]-I}, \\ \varnothing & \text{otherwise,} \end{cases} \\
\Phi_i(I; x_i) &:= \begin{cases} * & \text{if } x_i = 0 \text{ and } i \in I, \\ S^\sigma & \text{otherwise.} \end{cases}
\end{aligned}$$

Since there is a natural decomposition

$$\operatorname{tcofib}(\Phi) \simeq \coprod_{x \in \mathbb{Z}^n} \operatorname{tcofib}(\Phi(-; x)),$$

it suffices to show $\operatorname{tcofib}(\Phi(-; x)) \simeq 0$ for every $x \in \mathbb{Z}^n$.

If $x \neq (0,\dots,0)$, then there exists $i \in [n]$ such that $x \notin P_i$. The induced map $\Phi(I;x) \to \Phi(I \cup \{i\};x)$ is a $\mathbb{Z}/2$-homeomorphism whenever $i \notin I$. Together with the categorical result [7, Proposition A.6.5], we have $\mathrm{tcofib}(\Phi(-;x)) \simeq 0$.

If $x = (0,\dots,0)$, then the induced map $\Phi(I;x) \to \Phi(I \cup \{0\};x)$ is a $\mathbb{Z}/2$-homeomorphism whenever $0 \notin I$. We have $\mathrm{tcofib}(\Phi(-;x)) \simeq 0$ similarly. □

The THH and TC parts of the following result are proved in [7, Theorem 8.4.4].

Theorem 6.4 *Let S be a noetherian separated scheme. Then the sheaves*

$$\mathrm{THH}, \mathrm{THR}^{\mathbb{Z}/2}, \mathrm{TC}, \mathrm{TCR}^{\mathbb{Z}/2} \in \mathrm{Sh}_{\mathrm{sNis}}(\mathrm{SmlSm}/S, \mathrm{Sp})$$

are in $\mathrm{logSH}_{S^1}(S)$.

Proof Immediate from Theorem 6.3. □

Since THR is not a sheaf of spectra but a sheaf of $\mathbb{Z}/2$-spectra, we cannot directly apply the results in [6; 7] to THR. Instead, we will often use the fact that the pair $(i^*, (-)^{\mathbb{Z}/2})$ is conservative to reduce to the case of THH and $\mathrm{THR}^{\mathbb{Z}/2}$.

For any vector bundle $\mathcal{E} \to X$ with $X \in \mathrm{SmlSm}/S$, the *real topological Hochschild homology of the Thom space* $\mathrm{Th}(\mathcal{E})$ is defined to be

$$\mathrm{THR}(\mathrm{Th}(\mathcal{E})) := \mathrm{fib}(\mathrm{THR}(\mathcal{E}) \to \mathrm{THR}(\mathrm{Bl}_Z\mathcal{E}, E)), \tag{6-1}$$

where Z is the zero section of $\mathcal{E}$, and E is the exceptional divisor.

Proposition 6.5 *With the above notation, there are natural equivalences of $\mathbb{Z}/2$-spectra*

$$\mathrm{THR}(\mathrm{Th}(\mathcal{E})) \simeq \mathrm{fib}\big(\mathrm{THR}(\mathbb{P}(\mathcal{E} \oplus \mathcal{O})) \to \mathrm{THR}(\mathbb{P}(\mathcal{E}))\big),$$
$$\mathrm{TCR}(\mathrm{Th}(\mathcal{E})) \simeq \mathrm{fib}\big(\mathrm{TCR}(\mathbb{P}(\mathcal{E} \oplus \mathcal{O})) \to \mathrm{TCR}(\mathbb{P}(\mathcal{E}))\big).$$

Proof We focus on the case of THR since the proofs are similar.

Let Y and Y' be the blow-up of $\mathcal{E}$ and $\mathbb{P}(\mathcal{E} \oplus \mathcal{O})$ along the zero section, and let E and E' be the exceptional divisors. By the proof of [6, Proposition 7.4.5], there exists a commutative diagram

$$\begin{tikzcd}
(Y,E) \arrow[r] \arrow[d] & (Y',E') \arrow[d] & \mathbb{P}(\mathcal{E}) \arrow[l] \\
\mathcal{E} \arrow[r] & \mathbb{P}(\mathcal{E} \oplus \mathcal{O}) &
\end{tikzcd}$$

such that the induced maps

$$\mathrm{fib}\big(\mathcal{F}(\mathbb{P}(\mathcal{E} \oplus \mathcal{O})) \to \mathcal{F}(Y',E')\big) \to \mathrm{fib}(\mathcal{F}(\mathcal{E}) \to \mathcal{F}(Y,E)),$$
$$\mathcal{F}(Y',E') \to \mathcal{F}(\mathbb{P}(\mathcal{E}))$$

are equivalences for $\mathcal{F} \in \mathrm{logSH}_{S^1}(S)$. In particular, these are equivalences for THH and $\mathrm{THR}^{\mathbb{Z}/2}$ by Theorem 6.4. From this, we obtain the desired equivalence. □

Lemma 6.6 *Let $\mathcal{E}_1 \to X_1$ and $\mathcal{E}_2 \to X_2$ be vector bundles, where $X_1, X_2 \in \mathrm{SmlSm}/S$. We regard $T := \mathcal{E}_1 \times_S \mathcal{E}_2$ as a vector bundle over $X_1 \times_S X_2$. We set*

$$T_1 := (\mathrm{Bl}_{Z_1}\mathcal{E}_1, E_1) \times_S \mathcal{E}_2, \quad T_2 := \mathcal{E}_1 \times_S (\mathrm{Bl}_{Z_2}\mathcal{E}_2, E_2), \quad T_{12} := T_1 \times_S T_2,$$

where Z_1 and Z_2 are the zero sections of $\mathcal{E}_1$ and $\mathcal{E}_2$. Then there are natural equivalences of $\mathbb{Z}/2$-spectra

$$\mathrm{THR}(\mathrm{Th}(T)) \simeq \mathrm{tfib}(\mathrm{THR}(C_3)), \quad \mathrm{TCR}(\mathrm{Th}(T)) \simeq \mathrm{tfib}(\mathrm{TCR}(C_3)),$$

where C_3 is the cartesian square

$$\begin{tikzcd} T_{12} \arrow[r] \arrow[d] & T_1 \arrow[d] \\ T_2 \arrow[r] & T \end{tikzcd}$$

Proof Again, we focus on THR. Consider the fs log schemes $T_I \in \mathrm{SmlSm}/S$ for every subset I of $\{1,2,4\}$ in [6, Construction 7.4.14], and consider the induced squares

$$C_1 := \begin{tikzcd} T_4 \arrow[r] \arrow[d] & T_4 \arrow[d] \\ T_4 \arrow[r] & T \end{tikzcd} \qquad C_2 := \begin{tikzcd} T_{124} \arrow[r] \arrow[d] & T_{14} \arrow[d] \\ T_{24} \arrow[r] & T \end{tikzcd}$$

We have the induced maps of squares $C_1 \leftarrow C_2 \rightarrow C_3$. By Theorem 6.4 and the proof of Proposition 7.4.15 in [6], the induced maps

$$\mathrm{tfib}(\mathcal{F}(C_1)) \to \mathrm{tfib}(\mathcal{F}(C_2)) \leftarrow \mathrm{tfib}(\mathcal{F}(C_3))$$

are equivalences for $\mathcal{F} := \mathrm{THH}, \mathrm{THR}^{\mathbb{Z}/2}$. This implies that we have a natural equivalence

$$\mathrm{tfib}(\mathrm{THR}(C_1)) \simeq \mathrm{tfib}(\mathrm{THR}(C_3)),$$

which yields the desired equivalence. □

Proposition 6.7 *For $X \in \mathrm{SmlSm}/S$ and integer $n \geq 0$, there are natural equivalences of $\mathbb{Z}/2$-spectra*

$$\mathrm{THR}(\mathrm{Th}(X \times \mathbb{A}^n)) \simeq \Sigma^{n(\sigma-1)}\mathrm{THR}(X), \quad \mathrm{TCR}(\mathrm{Th}(X \times \mathbb{A}^n)) \simeq \Sigma^{n(\sigma-1)}\mathrm{TCR}(X),$$

where we regard $X \times \mathbb{A}^n$ as the rank-n trivial bundle over X.

Proof Again, we focus on THR. We proceed by induction on n. The claim is trivial if $n = 0$. Assume $n > 0$. We set $X_1 := X$, $\mathcal{E}_1 := X \times \mathbb{A}^{n-1}$, $X_2 := S$, and $\mathcal{E}_2 := S \times \mathbb{A}^1$. The square C_3 in Lemma 6.6 becomes

$$\begin{tikzcd} (\mathrm{Bl}_X(X \times \mathbb{A}^{n-1}), E_{n-1}) \times (\mathbb{A}^1, 0) \arrow[r] \arrow[d] & (\mathrm{Bl}_X(X \times \mathbb{A}^{n-1}), E_{n-1}) \times \mathbb{A}^1 \arrow[d] \\ X \times \mathbb{A}^{n-1} \times (\mathbb{A}^1, 0) \arrow[r] & X \times \mathbb{A}^{n-1} \times \mathbb{A}^1 \end{tikzcd}$$

where E_{n-1} is the exceptional divisor. By [19, Proposition 5.2.4], Proposition 3.3, and Lemma 6.6, we have a natural equivalence

$$\mathrm{THR}(\mathrm{Th}(X\times\mathbb{A}^n))\simeq\mathrm{THR}(\mathrm{Th}(X\times\mathbb{A}^{n-1}))\wedge\mathrm{fib}(\mathrm{THR}(\mathbb{S}[\mathbb{N}])\to\mathrm{THR}(\mathbb{S}[\mathbb{N}],\mathbb{N})).$$

Combine this with Example 3.4 and use the induction hypothesis to conclude. □

Next, we are concerned about the Gysin cofiber sequence in logarithmic motivic homotopy theory. For this, we recall the deformation to the normal cone construction in the logarithmic setting as follows. Assume that $X\in\mathrm{SmlSm}/S$ has the form $(\underline{X},Z_1+\cdots+Z_n)$, and let Z be a strict closed subscheme of X such that $\underline{Z}$ is strict normal crossing with $Z_1+\cdots+Z_n$ in the sense of [6, Definition 7.2.1]. The *blow-up of X along Z* is

$$\mathrm{Bl}_Z X:=(\mathrm{Bl}_{\underline{Z}}\underline{X},\widetilde{Z}_1+\cdots+\widetilde{Z}_n),$$

where $\widetilde{Z}_i$ is the strict transform of Z_i for $1\le i\le n$. The *normal bundle of Z in X* is defined to be

$$\mathrm{N}_Z X:=\mathrm{N}_{\underline{Z}}\underline{X}\times_{\underline{X}}X,$$

where $\mathrm{N}_{\underline{Z}}\underline{X}$ denotes the normal bundle of $\underline{Z}$ in $\underline{X}$. The *deformation space associated with $Z\to X$* is defined to be

$$\mathrm{D}_Z X:=\mathrm{Bl}_{Z\times\{0\}}(X\times\square)-\mathrm{Bl}_{Z\times\{0\}}(X\times\{0\}).$$

Theorem 6.8 *With the above notation, there exist natural cofiber sequences of $\mathbb{Z}/2$-spectra*

$$\mathrm{THR}(\mathrm{Th}(\mathrm{N}_Z X))\to\mathrm{THR}(X)\to\mathrm{THR}(\mathrm{Bl}_Z X,E),$$
$$\mathrm{TCR}(\mathrm{Th}(\mathrm{N}_Z X))\to\mathrm{TCR}(X)\to\mathrm{TCR}(\mathrm{Bl}_Z X,E),$$

where E is the exceptional divisor.

Proof Again, we focus on THR. We have the induced maps

$$\mathrm{fib}(\mathrm{THR}(X)\to\mathrm{THR}(\mathrm{Bl}_Z X,E))\leftarrow\mathrm{fib}\big(\mathrm{THR}(\mathrm{D}_Z X)\to\mathrm{THR}(\mathrm{Bl}_{Z\times\square}(\mathrm{D}_Z X),E^D)\big)$$
$$\to\mathrm{fib}\big(\mathrm{THR}(\mathrm{N}_Z X)\to\mathrm{THR}(\mathrm{Bl}_Z(\mathrm{N}_Z X),E^N)\big),$$

where E, E^D, and E^N are the exceptional divisors. It suffices to show that the corresponding maps for THH and $\mathrm{THR}^{\mathbb{Z}/2}$ are equivalences of spectra since the pair of functors $(i^*,(-)^{\mathbb{Z}/2})$ is conservative. This is a consequence of [6, Theorem 7.5.4] (see also [7, Theorem 3.2.14]) and Theorem 6.4. □

Recall from [7, Definition 8.5.3] that we have the log motivic spectra

$$\mathbf{THH}:=(\mathrm{THH},\mathrm{THH},\dots),$$
$$\mathbf{TC}:=(\mathrm{TC},\mathrm{TC},\dots),$$

whose bonding maps $\mathrm{THH}\to\Omega_{\mathbb{P}^1}\mathrm{THH}$ and $\mathrm{TC}\to\Omega_{\mathbb{P}^1}\mathrm{TC}$ are obtained by the projective bundle formula.

Example 6.9 Let X be a scheme, and let $\mathcal{E} \to X$ be a rank-n vector bundle. There exists a Thom equivalence

$$\mathrm{THH}(\mathrm{Th}(\mathcal{E})) \simeq \mathrm{THH}(X);$$

see [7, Proposition 8.6.9]. This is a consequence of the fact that **THH** is an orientable logarithmic motivic spectrum [7, Definition 7.1.3, Theorem 8.6.7]. Unlike this, in general, we do not have an equivalence

$$\mathrm{THR}(\mathrm{Th}(\mathcal{E})) \simeq \Sigma^{n(\sigma-1)}\mathrm{THR}(X),$$

i.e., $\mathrm{THR}(\mathrm{Th}(\mathcal{E})) \not\simeq \mathrm{THR}(\mathrm{Th}(X \times \mathbb{A}^n))$ by Proposition 6.7.

For example, if $X := \mathbb{P}^n_{\mathbb{S}}$ and $Z := \mathbb{P}^{n-1}_{\mathbb{S}}$ with even $n \geq 2$, then Proposition 6.2 and Theorems 6.3 and 6.8 yield

$$\mathrm{THR}(\mathrm{Th}(\mathrm{N}_Z X)) \simeq \bigoplus_{j=1}^{\lfloor n/2 \rfloor} i_*\mathbb{S}.$$

This is not equivalent to

$$\Sigma^{\sigma-1}\mathrm{THR}(Z) \simeq \Sigma^{\sigma-1} \bigoplus_{j=1}^{\lfloor (n-1)/2 \rfloor} i_*\mathbb{S} \oplus \Sigma^{n(\sigma-1)}$$

that is obtained by Proposition 6.2. We have a similar conclusion if n is odd too.

We also obtain the following descent result with respect to blow-ups along smooth centers (with the trivial involutions).

Theorem 6.10 *Let $Z \to X$ be a closed immersion of smooth schemes over S. Then the induced squares of $\mathbb{Z}/2$-spectra*

$$\begin{tikzcd}
\mathrm{THR}(X) \arrow[r] \arrow[d] & \mathrm{THR}(Z) \arrow[d] & \mathrm{TCR}(X) \arrow[r] \arrow[d] & \mathrm{TCR}(Z) \arrow[d] \\
\mathrm{THR}(Z \times_X \mathrm{Bl}_Z X) \arrow[r] & \mathrm{THR}(\mathrm{Bl}_Z X) & \mathrm{TCR}(Z \times_X \mathrm{Bl}_Z X) \arrow[r] & \mathrm{TCR}(\mathrm{Bl}_Z X)
\end{tikzcd}$$

are cartesian.

Proof Again, we focus on THR. The corresponding squares of spectra for THH and $\mathrm{THR}^{\mathbb{Z}/2}$ are cartesian by [6, Theorem 7.3.3] (see also [7, Theorem 3.3.5]) and Theorem 6.4. This implies the claim since the pair of functors $(i^*, (-)^{\mathbb{Z}/2})$ is conservative. □

Let $\mathbb{P}^\sigma$ be the scheme $\mathbb{P}^1$ with the involution given by $[x : y] \mapsto [y : x]$. Let 1 be the base points of $\mathbb{P}^1$ and $\mathbb{P}^\sigma$, and then we form the endofunctors $\Omega_{\mathbb{P}^1}$, $\Omega_{\mathbb{P}^\sigma}$, and $\Omega_{\mathbb{P}^1 \wedge \mathbb{P}^\sigma} \simeq \Omega_{\mathbb{P}^\sigma}\Omega_{\mathbb{P}^1}$ on $\mathrm{PSh}((\mathrm{SmlSm}/S)_{\mathbb{Z}/2}, \mathrm{Sp})$ and $\mathrm{PSh}((\mathrm{SmlSm}/S)_{\mathbb{Z}/2}, \mathrm{Sp}^{\mathbb{Z}/2})$.

Proposition 6.11 *Let X be a noetherian separated fs log scheme. Then there are natural equivalences of $\mathbb{Z}/2$-spectra*

$$\mathrm{THR}(X) \simeq \Omega_{\mathbb{P}^1 \wedge \mathbb{P}^\sigma}\mathrm{THR}(X), \quad \mathrm{TCR}(X) \simeq \Omega_{\mathbb{P}^1 \wedge \mathbb{P}^\sigma}\mathrm{TCR}(X).$$

Proof Again, we focus on THR. Following Remark 6.1, we reduce to [19, Proposition 5.1.5]. □

Definition 6.12 The *∞-category of prelogarithmic motivic $\mathbb{Z}/2$-spectra over S* is defined to be

$$\mathrm{prelogSH}^{\mathbb{Z}/2}(S) := \mathrm{Sp}_{\mathbb{P}^1 \wedge \mathbb{P}^\sigma}\big((\mathbb{P}^\bullet, \mathbb{P}^{\bullet-1})^{-1}\mathrm{Sh}_{\mathrm{seNis}}((\mathrm{SmlSm}/S)_{\mathbb{Z}/2}, \mathrm{Sp})\big).$$

Remark 6.13 Let $\mathbb{A}^w$ denote the affine line $\mathbb{A}^1$ with the involution w given by $x \mapsto -x$. We have the $\mathbb{A}^1$-homotopy

$$\mathbb{A}^1 \times \mathbb{A}^w \to \mathbb{A}^w \tag{6-2}$$

given by $(x, y) \mapsto xy$. Due to this, the projections $X \times \mathbb{A}^w \to X$ are automatically inverted if we invert the projections $X \times \mathbb{A}^1 \to X$ for all $X \in (\mathrm{Sm}/S)_{\mathbb{Z}/2}$.

However, this phenomenon is not generalized to the logarithmic setting since there exists no morphism of fs log schemes

$$\square \times \square^w \to \square^w$$

extending (6-2), where $\square^w$ denotes the fs log scheme $\square$ with the involution $[x, y] \mapsto [-x, y]$. This indicates that inverting the projections $X \times \square^w \to X$ for all $X \in (\mathrm{SmlSm}/S)_{\mathbb{Z}/2}$ is at least required for a better behaved ∞-category of logarithmic motivic $\mathbb{Z}/2$-spectra.

Definition 6.14 The *prelogarithmic motivic $\mathbb{Z}/2$-fixed point functor* is defined to be the functor

$$(-)^{\mathbb{Z}/2} : \mathrm{prelogSH}^{\mathbb{Z}/2}(S) \to \mathrm{logSH}(S)$$

sending a $\mathbb{P}^1 \wedge \mathbb{P}^\sigma$-spectrum $(\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2, \dots)$ to the $\mathbb{P}^1$-spectrum

$$(\mathcal{F}_0, \Omega_{\mathbb{P}^\sigma}\mathcal{F}_1, \Omega_{\mathbb{P}^{2\sigma}}\mathcal{F}_2, \dots)$$

restricting to the objects of the category SmlSm/S with the trivial involutions, where the bonding map $\Omega_{\mathbb{P}^{n\sigma}}\mathcal{F}_n \to \Omega_{\mathbb{P}^1 \wedge \mathbb{P}^{(n+1)\sigma}}\mathcal{F}_{n+1}$ is induced by the bounding map $\mathcal{F}_n \to \Omega_{\mathbb{P}^1 \wedge \mathbb{P}^\sigma}\mathcal{F}_{n+1}$.

Definition 6.15 The *motivic real topological Hochschild $\mathbb{Z}/2$-spectrum* and *motivic real topological cyclic $\mathbb{Z}/2$-spectrum* are the $\mathbb{P}^1 \wedge \mathbb{P}^\sigma$-spectra

$$\begin{aligned}\mathbf{THR} &:= (\mathrm{THR}^{\mathbb{Z}/2}, \mathrm{THR}^{\mathbb{Z}/2}, \dots) \in \mathrm{prelogSH}^{\mathbb{Z}/2}(S),\\ \mathbf{TCR} &:= (\mathrm{TCR}^{\mathbb{Z}/2}, \mathrm{TCR}^{\mathbb{Z}/2}, \dots) \in \mathrm{prelogSH}^{\mathbb{Z}/2}(S),\end{aligned}$$

whose bonding maps are given by Proposition 6.11. We have the induced motivic spectra

$$\mathbf{THR}^{\mathbb{Z}/2}, \mathbf{TCR}^{\mathbb{Z}/2} \in \mathrm{logSH}(S).$$

Definition 6.16 For a strict Nisnevich sheaf of spectra $\mathcal{F}$ on $(\mathrm{SmlSm}/S)_{\mathbb{Z}/2}$, let $i^*\mathcal{F}$ be the strict Nisnevich sheaf of spectra on SmlSm/S given by

$$i^*\mathcal{F}(X) := \mathcal{F}(i_\sharp X).$$

Observe that for every integer $n \geq 1$, $i^*\mathcal{F}$ is $(\mathbb{P}^n, \mathbb{P}^{n-1})$-invariant if $\mathcal{F}$ is $(\mathbb{P}^n, \mathbb{P}^{n-1})$-invariant. We have the "forgetful" functor

$$i^* : \mathrm{prelogSH}^{\mathbb{Z}/2}(S) \to \mathrm{logSH}(S)$$

sending a $\mathbb{P}^1 \wedge \mathbb{P}^\sigma$-spectrum $(\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2, \dots)$ to the $\mathbb{P}^1$-spectrum

$$(i^*\mathcal{F}_0, \Omega_{\mathbb{P}^\sigma} i^*\mathcal{F}_1, \Omega_{\mathbb{P}^{2\sigma}} i^*\mathcal{F}_2, \dots)$$

with the induced bonding maps.

Proposition 6.17 *There are equivalences, in* $\mathrm{logSH}(S)$,

$$i^*\mathbf{THR} \simeq \mathbf{THH},$$
$$i^*\mathbf{TCR} \simeq \mathbf{TC}.$$

Proof This is a consequence of Corollary 4.9 and Proposition 5.30. □

Remark 6.18 In this section, we have considered mainly THR and TCR, but we can similarly show the analogous results for $\mathrm{TCR}^- := \mathrm{THR}^{hS^\sigma}$, TCR_{hS^σ}, and $\mathrm{TPR} := \mathrm{THR}^{tS^\sigma}$.

Acknowledgements

This research was conducted in the framework of the DFG-funded research training group GRK 2240: *Algebro-Geometric Methods in Algebra, Arithmetic and Topology*. We thank Jens Hornbostel for many helpful conversations on this topic. We are grateful to the referee for detailed and precise comments leading to improvements of this text.

References

[1] **J Ayoub**, *Les six opérations de Grothendieck et le formalisme des cycles évanescents dans le monde motivique, I*, Astérisque 314, Soc. Math. France, Paris (2007) MR

[2] **T Bachmann**, **M Hoyois**, *Norms in motivic homotopy theory*, Astérisque 425, Soc. Math. France, Paris (2021) MR

[3] **B Bhatt**, **M Morrow**, **P Scholze**, *Topological Hochschild homology and integral p-adic Hodge theory*, Publ. Math. Inst. Hautes Études Sci. 129 (2019) 199–310 MR

[4] **B Bhatt**, **P Scholze**, *Prisms and prismatic cohomology*, Ann. of Math. (2) 196:3 (2022) 1135–1275 MR

[5] **F Binda**, **T Lundemo**, **D Park**, **P A Østvær**, *A Hochschild–Kostant–Rosenberg theorem and residue sequences for logarithmic Hochschild homology*, Adv. Math. 435 (2023) art. id. 109354 MR

[6] **F Binda**, **D Park**, **P A Østvær**, *Triangulated categories of logarithmic motives over a field*, Astérisque 433, Soc. Math. France, Paris (2022) MR

[7] **F Binda**, **D Park**, **P A Østvær**, *Logarithmic motivic homotopy theory* (2023) arXiv 2303.02729v2

[8] **M Bökstedt**, **W C Hsiang**, **I Madsen**, *The cyclotomic trace and algebraic K-theory of spaces*, Invent. Math. 111:3 (1993) 465–539 MR

[9] **P Deligne**, *Voevodsky's lectures on motivic cohomology* 2000/2001, from "Algebraic topology" (N A Baas, E M Friedlander, B Jahren, P A Østvær, editors), Abel Symp. 4, Springer (2009) 355–409 MR

[10] **E Dotto**, **K Moi**, **I Patchkoria**, **S P Reeh**, *Real topological Hochschild homology*, J. Eur. Math. Soc. 23:1 (2021) 63–152 MR

[11] **B I Dundas**, **T G Goodwillie**, **R McCarthy**, *The local structure of algebraic K-theory*, Algebra and Applications 18, Springer (2013) MR

[12] **Z Fiedorowicz**, **J-L Loday**, *Crossed simplicial groups and their associated homology*, Trans. Amer. Math. Soc. 326:1 (1991) 57–87 MR

[13] **T Geisser**, **L Hesselholt**, *Topological cyclic homology of schemes*, from "Algebraic K-theory" (Seattle, WA, 1997) (W Raskind, C Weibel, editors), Proc. Sympos. Pure Math. 67, Amer. Math. Soc., Providence, RI (1999) 41–87 MR

[14] **J Heller**, **A Krishna**, **P A Østvær**, *Motivic homotopy theory of group scheme actions*, J. Topol. 8:4 (2015) 1202–1236 MR

[15] **L Hesselholt**, **I Madsen**, *Real algebraic K-theory*, preprint Available at https://web.math.ku.dk/~larsh/papers/s05/book.pdf

[16] **L Hesselholt**, **I Madsen**, *On the K-theory of local fields*, Ann. of Math. (2) 158:1 (2003) 1–113 MR

[17] **M A Hill**, **M J Hopkins**, **D C Ravenel**, *On the nonexistence of elements of Kervaire invariant one*, Ann. of Math. (2) 184:1 (2016) 1–262 MR

[18] **A Høgenhaven**, *Real topological cyclic homology of spherical group rings* (2016) arXiv 1611.01204v1

[19] **J Hornbostel**, **D Park**, *Real topological Hochschild homology of schemes*, J. Inst. Math. Jussieu 23:3 (2024) 1461–1518 MR

[20] **P Hu**, **I Kriz**, **K Ormsby**, *The homotopy limit problem for Hermitian K-theory, equivariant motivic homotopy theory and motivic real cobordism*, Adv. Math. 228:1 (2011) 434–480 MR

[21] **J Lurie**, *Spectral algebraic geometry*, preprint (2016) Available at https://url.msp.org/Lurie-SAG

[22] **J Lurie**, *Higher algebra*, preprint (2017) Available at https://url.msp.org/Lurie-HA

[23] **F Morel**, **V Voevodsky**, $\mathbb{A}^1$*-homotopy theory of schemes*, Inst. Hautes Études Sci. Publ. Math. :90 (1999) 45–143 MR

[24] **T Nikolaus**, **P Scholze**, *On topological cyclic homology*, Acta Math. 221:2 (2018) 203–409 MR

[25] **A Ogus**, *Lectures on logarithmic algebraic geometry*, Cambridge Studies in Advanced Mathematics 178, Cambridge Univ. Press (2018) MR

[26] **D Park**, *Syntomic cohomology and real topological cyclic homology* (2023) arXiv 2311.06593v1

[27] **J D Quigley**, **J Shah**, *On the equivalence of two theories of real cyclotomic spectra* (2021) arXiv 2112.07462v2

[28] **J D Quigley**, **J Shah**, *On the parametrized Tate construction* (2022) arXiv 2110.07707v2

[29] **J Rognes**, *Topological logarithmic structures*, from "New topological contexts for Galois theory and algebraic geometry" (Banff, Canada, 2008) (A Baker, B Richter, editors), Geom. Topol. Monogr. 16, Geom. Topol., Coventry (2009) 401–544 MR

[30] **J Rognes**, **S Sagave**, **C Schlichtkrull**, *Localization sequences for logarithmic topological Hochschild homology*, Math. Ann. 363:3-4 (2015) 1349–1398 MR

[31] **S Schwede**, *Lecture notes on equivariant stable homotopy theory* (2016) Available at https://www.math.uni-bonn.de/people/schwede/equivariant.pdf

[32] **S Schwede**, *Global homotopy theory*, New Mathematical Monographs 34, Cambridge Univ. Press (2018) MR

[33] **G Segal**, *Configuration-spaces and iterated loop-spaces*, Invent. Math. 21 (1973) 213–221 MR

[34] **F Waldhausen**, *Algebraic K-theory of topological spaces, II*, from "Algebraic topology" (Aarhus, Denmark, 1978) (J L Dupont, I H Madsen, editors), Lecture Notes in Math. 763, Springer (1979) 356–394 MR

DOOSUNG PARK dpark@uni-wuppertal.de
Department of Mathematics and Informatics, Bergische Universität Wuppertal, Wuppertal, Germany

Received: June 6, 2024 Revised: April 30, 2025

Algebraic & Geometric Topology **26**:5 (2026) 1907–1921
DOI: 10.2140/agt.2026.26.1907
Published: May 29, 2026

A note on the involutive invariants of splices

KRISTEN HENDRICKS, MATTHEW STOFFREGEN AND IAN ZEMKE

A natural family of potentially 2-torsion elements in the integer homology cobordism group consists of splices of knots with their mirrors. We show that such 3-manifolds have locally trivial involutive Floer homology. We show some related families of splices also have locally trivial involutive Floer homology. Our arguments show that many gauge-theoretic invariants also vanish on these 3-manifolds.

1 Introduction

The integer homology cobordism group $\Theta^3_{\mathbb{Z}}$ is the group of oriented homology three-spheres up to the equivalence relation of homology cobordism. In 2013, C. Manolescu used a Pin(2)-equivariant version of Seiberg–Witten Floer homology to show that $\Theta^3_{\mathbb{Z}}$ contains no element Y of order two whose Rokhlin invariant $\mu(Y)$ is 1 [18], which due to previous work of Galewski and Stern [7] and Matumoto [19] was sufficient to disprove the remaining outstanding cases of the triangulation conjecture. It remains unknown whether $\Theta^3_{\mathbb{Z}}$ has any torsion elements; in particular, whether it contains a torsion element of order two. In order to produce an element of order two, it suffices to exhibit an oriented integer homology sphere Y with an orientation-reversing diffeomorphism $Y \cong -Y$ with the property that Y takes a nontrivial value under any invariant of homology cobordism.

Three-manifolds obtained by splicing knot complements have attracted attention as a potential source of examples of elements of order two in $\Theta^3_{\mathbb{Z}}$. If $K_0 \subseteq Y_0$ and $K_1 \subseteq Y_1$ are two knots, a *splice* of K_0 and K_1 is a 3-manifold obtained by gluing $Y_0 \setminus \nu(K_0)$ and $Y_1 \setminus \nu(K_1)$ using an orientation-reversing diffeomorphism ϕ of their boundaries. (Some authors require ϕ to swap the meridian with the Seifert longitude, but we consider more general diffeomorphisms ϕ).

In this note, we consider two natural constructions of splices which produce homology 3-spheres with orientation-reversing diffeomorphisms:

- (Type-1) splices of a knot $K \subseteq Y$ with its mirror $K \subseteq -Y$ such that there is a diffeomorphism of the splice which swaps $Y \setminus \nu(K)$ and $-Y \setminus \nu(K)$;
- (Type-2) splices of two knots K_0, K_1 in S^3 such that there is a diffeomorphism of the splice which fixes the subsets $S^3 \setminus \nu(K_0)$ and $S^3 \setminus \nu(K_1)$ setwise, but is orientation reversing on each.

Not all knots or gluing maps will yield a splice which admits such a symmetry. We enumerate all the requirements in Section 2. In this paper, we will refer to the above families as *Type*-1 and *Type*-2 *symmetric splices*, respectively. We will see in Section 2.1 that Type-1 splices must have (Y, K) *reversible*,

---

MSC2020: 57K18, 57K31, 57R58.

i.e., there must be a diffeomorphism $(Y, K) \cong (Y, -K)$, where $-K$ denotes the knot with the opposite string orientation. In Section 2.2 we will see that in a Type-2 symmetric splice, one of K_0 and K_1 must be negatively amphichiral and the other positively amphichiral. We will also enumerate all possible gluing maps.

Involutive Heegaard Floer homology is a shadow theory of Pin(2)-equivariant Seiberg–Witten Floer homology, introduced by Manolescu and Hendricks in 2015 and elaborated by Hendricks and Zemke with Manolescu [12; 13], which is conjecturally equivalent to $\mathbb{Z}/4\mathbb{Z}$-equivariant Seiberg–Witten Floer homology. Although involutive Heegaard Floer homology does not possess the technical power of a Pin(2)-equivariant theory, it enjoys the greater computability of the Heegaard Floer invariants, including a conveniently computable surgery formula [11], and has been a key tool in recent developments regarding the structure of the homology cobordism group [4; 9; 10; 11].

In this note we show that the homology cobordism involutive invariants of Type-1 and Type-2 splices are typically trivial. Recall for a homology sphere Z, these invariants consist of a pair $(\mathrm{CF}^-(Z), \iota)$ a chain complex together with a homotopy involution, together called an iota-complex; for more details, see Section 3.

Theorem 1.1 (1) *Suppose that K is a knot in an integer homology* 3*-sphere Y. If Z is a Type*-1 *symmetric splice of (Y, K) with $(-Y, -K)$ then the iota-complex $(\mathrm{CF}^-(Z), \iota)$ is locally trivial.*

(2) *If Z is a Type*-2 *symmetric splice of (S^3, K_0) and (S^3, K_1) such that $\mathrm{CFK}^-(K_0)$ and $\mathrm{CFK}^-(K_1)$ are (noninvolutively) locally trivial, then the iota-complex $(\mathrm{CF}^-(Z), \iota)$ is locally trivial.*

Remark 1.2 (1) It is not clear to the authors whether Theorem 1.1(2) extends to all amphichiral K_0 and K_1, nor whether it can be extended to knots in homology 3-spheres other than S^3.

(2) As we mentioned above, in a symmetric splice of Type-2, one of K_0 and K_1 must be negative amphichiral, and the other positive amphichiral. To the best of our knowledge, all known amphichiral knots have locally trivial (noninvolutive) knot Floer complex $\mathrm{CFK}^-(K)$. Also note if K is *strongly* negative amphichiral, i.e., if the pair (S^3, K) admits an orientation-reversing diffeomorphism ϕ which has exactly two fixed points, both of which lie along K, then Kawauchi's result [15] implies that K is rationally slice, and hence has locally trivial $\mathrm{CFK}^-(K)$.

The key topological input to Theorem 1.1 is the following result:

Proposition 1.3 (1) *If Z is a Type*-1 *symmetric splice of (Y, K) with $(-Y, -K)$, then there is a negative definite* Spin *cobordism from Z to $\mathbb{RP}^3$ which has $b_2^- = 1$ and $b_1 = 0$.*

(2) *If Z is a Type*-1 *symmetric splice of (Y, K) with $(-Y, -K)$, then there is a negative definite* (*non*-Spin) *filling of Z with $b_2^- = 2$ and $b_1 = 0$.*

(3) *If Z is a Type*-2 *symmetric splice of (Y_0, K_0) with (Y_1, K_1), then there is a negative definite* Spin *cobordism from Z to $(Y_0 \# Y_1)_{-2}(K_0 \# K_1)$ with $b_2^- = 1$ and $b_1 = 0$.*

Remark 1.4 In unpublished work, Mike Miller Eismeier independently proved Proposition 1.3(2) and used it to show that certain instanton-theoretic gauge-theoretic invariants are trivial on such splices.

We now sketch some ideas in the proof of Theorem 1.1, assuming Proposition 1.3. For Type-1 splices, the negative definite Spin cobordism W from Z to $\mathbb{RP}^3$ has a unique self-conjugate Spin^c structure $\mathfrak{s}$. Furthermore, the Heegaard Floer grading shift $d(W,\mathfrak{s})$ is equal to the correction term $d(\mathbb{RP}^3,\mathfrak{s}|_{\mathbb{RP}^3})$. Since $\mathbb{RP}^3$ is a Heegaard Floer L-space, the cobordism map $F_{W,\mathfrak{s}}$ can be viewed as a local map from $(\mathrm{CF}^-(Z),\iota)$ to the trivial ι-complex. Since $Z \cong -Z$, we can dualize the map $F_{W,\mathfrak{s}}$ to get a local map in the opposite direction.

For Type-2 splices, the cobordism W from Z to $S^3_{-2}(K_0 \# K_1)$ also has a unique self-conjugate Spin structure $\mathfrak{s}$. In this case, $\mathfrak{s}$ restricts to the Spin^c structure identified with $[1] \in \mathbb{Z}/2 \cong \mathrm{Spin}^c(S^3_{-2}(K_0 \# K_1))$ under the standard identification. We use [11, Theorem 1.6(2)], which implies that since $\mathrm{CFK}^-(K_0)$ and $\mathrm{CFK}^-(K_1)$ are (noninvolutively) locally trivial, the ι-complex $\big(\mathrm{CF}^-(S^3_{-2}(K_0 \# K_1),[1]),\iota\big)$ is locally trivial up to an overall grading shift.

1.1 Other gauge-theoretic invariants

We note that the topological argument yielding Theorem 1.1 applies equally well to the Pin(2)-equivariant Seiberg–Witten Floer spectra:

Proposition 1.5 *The* Pin(2)*-equivariant Seiberg–Witten Floer spectra of symmetric splices of Type*-1 *are locally trivial.*

The same argument also implies the vanishing of Lin's invariants $\alpha(Z)$, $\beta(Z)$, $\gamma(Z)$ [16] for Type-1 symmetric splices.

Note that Proposition 1.3(2) can be applied to other invariants of gauge theory. In particular if A is a partially ordered set and

$$\omega : \Theta^3_{\mathbb{Z}} \to A$$

is a homology cobordism invariant which is monotonic under negative definite cobordisms with $b_1 = 0$, then Proposition 1.3(2) implies that if Z is a Type-1 symmetric splice, then $\omega(Z) \le \omega(S^3)$ and $\omega(S^3) \le \omega(Z)$, so in particular $\omega(Z) = \omega(S^3)$. This holds for many gauge-theoretic invariants of homology cobordism. For example, this argument applies to the r_s-invariants of Nozaki, Sato, Taniguchi [21] and Daemi's Γ invariant [1]. Compare also [2].

Organization

This note is organized as follows. In Section 2 we discuss the geometry of splices; in particular, we classify which symmetric splices of Types 1 and 2 are integer homology spheres and have a natural orientation-reversing diffeomorphism. In Section 3 we briefly recall relevant aspects of Heegaard Floer theory, focusing on its interaction with homology cobordism and concordance, for the reader's convenience. In Section 4 we conclude with the proof of Theorem 1.1.

2 Symmetric splices

In this section we recall some background on splices. Let $K_0 \subseteq Y_0$ and $K_1 \subseteq Y_1$ be oriented knots in integer homology 3-spheres. Let $\phi \in \mathbb{Z}^2 \to \mathbb{Z}^2$ be a 2×2 matrix with determinant -1. The map ϕ determines an orientation-reversing diffeomorphism

$$\phi : \partial(Y_0 \setminus \nu(K_0)) \to \partial(Y_1 \setminus \nu(K_1)),$$

where we view the first component of $\mathbb{Z}^2$ being an oriented meridian of K_i, and the second component being the oriented Seifert longitude. We define

$$\mathrm{Sp}_\phi(K_0, K_1) := (Y_0 \setminus \nu(K_0)) \cup_\phi (Y_1 \setminus \nu(K_1)).$$

In this section, we will be interested in knots that have various symmetries. We use the following standard terminology:

Definition 2.1 Let K be a knot in an oriented 3-manifold Y.

(1) (Y, K) is *reversible* if $(Y, K) \cong (Y, -K)$.

(2) (Y, K) is *negative amphichiral* if $(Y, K) \cong (-Y, -K)$.

(3) (Y, K) is *positive amphichiral* if $(Y, K) \cong (-Y, K)$.

In the above, $\cong$ means orientation-preserving diffeomorphic.

2.1 Type-1 symmetric splices

We now focus on Type-1 splices, i.e., splices of (Y, K) and $(-Y, -K)$ which admit orientation-reversing diffeomorphisms which switch $Y \setminus \nu(K)$ with $-Y \setminus \nu(K)$ but fix $\mathbb{T}^2 := \partial(Y \setminus \nu(K))$ setwise. We will write $\mathrm{Sp}_\phi(K, mK)$ for such splices. In this section we prove the following:

Proposition 2.2 *Suppose K is a knot in an integer homology 3-sphere Y, and $\phi \in \mathrm{GL}_2^-(\mathbb{Z})$. Assume that K is not an unknot in Y. Then the 3-manifold $\mathrm{Sp}_\phi(K, mK)$ admits an orientation-reversing diffeomorphism g which fixes $\partial(Y \setminus \nu(K))$ setwise and such that the image of each of $Y \setminus \nu(K)$ and $-Y \setminus \nu(K)$ is the other, if and only if the following are satisfied:*

(1) *K is reversible.*

(2) *$\phi = \left(\begin{smallmatrix} n & \pm 1 \\ \pm(1+n^2) & n \end{smallmatrix}\right)$ for some $n \in \mathbb{Z}$. In this case we define $\phi = \phi_n^\pm$.*

Proof It will be evident from the proof that if (Y, K) and ϕ are as in the statement, then $\mathrm{Sp}_\phi(K, mK)$ will have a symmetry g as above. Hence, we assume that $\mathrm{Sp}_\phi(K, mK)$ admits an orientation-reversing diffeomorphism g, as in the statement, and we will show that K and ϕ have the stated properties.

For our proof, it is somewhat easier write the gluing map in terms of a different basis. Note that if (μ, λ) is our oriented basis for $\partial Y \setminus \nu(K)$, then ϕ is written in terms of the basis $(\mu, -\lambda)$ for $-Y \setminus \nu(K)$. For our purposes, it is more helpful to write ϕ in terms of the basis (μ, λ) for both $Y \setminus \nu(K)$ and $-Y \setminus \nu(K)$, using the same longitude and meridian for both. Let us write ψ for the map ϕ in this basis. Note that $\det(\psi) = 1$.

Additionally, to simplify the notation, we will view $\mathrm{Sp}_\phi(K, mK)$ as the union of two copies of $Y \setminus \nu(K)$, which we denote by X_0 and X_1. We write

$$\mathrm{Sp}_\phi(K, mK) = \frac{X_0 \sqcup X_1}{\sim},$$

where $x \in \partial X_0$ is identified with $\psi(x) \in \partial X_1$. By assumption g is induced by some diffeomorphisms $g_{10} : X_0 \to X_1$ and $g_{01} : X_1 \to X_0$ as

$$\begin{tikzcd} X_0 \arrow[r, bend left, "g_{10}"] & X_1 . \arrow[l, bend left, "g_{01}"] \end{tikzcd}$$

The diffeomorphism $g_{10} \sqcup g_{01}$ descends to the quotient if and only if

$$g_{10}(x) = (\psi \circ g_{01} \circ \psi)(x) \tag{2-1}$$

for all $x \in \partial X_0$.

We now claim that

$$g_{01}|_{\partial X_1}, g_{10}|_{\partial X_0} \in \{\mathrm{id}, -\mathrm{id}\}. \tag{2-2}$$

To see this, note that both must map λ to $\pm\lambda$, because they must preserve the kernel of the map $H_1(\partial X_i) \to H_1(X_i)$. Less obviously, they must also send μ to $\pm\mu$. Homology considerations imply that g_{01} and g_{10} map μ to $\pm\mu + j\lambda$ for some $j \in \mathbb{Z}$. This would imply that, up to composition with the elliptic involution, $g_{01}|_{\partial X_0}$ is an j-fold composition of a Dehn twist parallel to the Seifert longitude, and similarly for g_{10}. It follows from [20, Theorem 1] that this can only happen if λ bounds a disk in $Y \setminus \nu(K)$, i.e., K is an unknot, which we exclude by hypothesis.

Since $g_{01}|_{\partial X_1}, g_{10}|_{\partial X_0} \in \{\mathrm{id}, -\mathrm{id}\}$, these maps are central in $\mathrm{GL}_2(\mathbb{Z})$, and hence (2-1) implies that $\psi^2 = \pm\,\mathrm{id}$.

We now consider the map ψ in more detail. The Mayer–Vietoris exact sequence reads

$$H_1(\mathbb{T}^2) \to H_1(Y \setminus \nu(K)) \oplus H_1(Y \setminus \nu(K)) \to H_1(Z) \to 0.$$

In particular, we see that for Y to be a homology sphere, we need $\psi(\lambda) = \pm\mu + n\lambda$, for some $n \in \mathbb{Z}$. That is, we can write ψ as a matrix as

$$\psi = \begin{pmatrix} n_1 & \pm 1 \\ * & n_2 \end{pmatrix}.$$

The condition that $\det \psi = 1$ imposes the restriction that

$$\psi = \begin{pmatrix} n_1 & \pm 1 \\ \mp(1 - n_1 n_2) & n_2 \end{pmatrix}. \tag{2-3}$$

It is straightforward to see that there are no such matrices of the above form which square to the identity matrix. On the other hand, the matrix in (2-3) squares to $-\mathrm{id}$ if and only if $n_1 = -n_2$. Setting $n = n_1$ and then changing to the basis gives the expression for ϕ in the statement.

Next, we observe that (2-1) now implies that $g_{10}|_{\partial X_0} = -g_{01}|_{\partial X_1}$. Therefore one map is the identity, while the other is the elliptic involution. Therefore $Y \setminus \nu(K)$ admits an orientation-preserving diffeomorphism which restricts to the elliptic involution on the boundary. Equivalently, there is a diffeomorphism of pairs $(Y, K) \cong (Y, -K)$. □

Lemma 2.3 *Suppose that* (Y, K) *is reversible. Then* $\mathrm{Sp}_{\phi_n^+}(K, mK) \cong \mathrm{Sp}_{\phi_{-n}^-}(K, mK)$.

Proof Since K is reversible, the elliptic involution of the boundary extends to an orientation-preserving diffeomorphism of $Y \setminus \nu(K)$. Therefore

$$\mathrm{Sp}_{\phi_n^+}(K, mK) \cong \mathrm{Sp}_{-\phi_n^+}(K, mK) = \mathrm{Sp}_{\phi_{-n}^-}(K, mK). \qquad \square$$

2.2 Type-2 symmetric splices

We now consider Type-2 splices. We say that a pair (Y, K) is *negative amphichiral* if $(-Y, -K) \cong (Y, K)$, and we say that (Y, K) is *positive amphichiral* if $(-Y, K) \cong (Y, K)$.

Proposition 2.4 *Let* (Y_0, K_0) *and* (Y_1, K_1) *be knots and* $\phi \in \mathrm{GL}_2^-(\mathbb{Z})$, *and furthermore suppose that* $\mathrm{Sp}_\phi(K_0, K_1)$ *is an integer homology* 3*-sphere. Then* $\mathrm{Sp}_\phi(K_0, K_1)$ *admits an orientation-reversing diffeomorphism* g *which preserves the subspaces* $Y_i \setminus \nu(K_i)$ *setwise if and only if the following hold*:

(1) *One of* (Y_0, K_0), (Y_1, K_1) *is negative amphichiral and the other is positive amphichiral.*

(2) $\phi = \pm\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.

Proof We write

$$X_0 = Y_0 \setminus \nu(K_0) \quad \text{and} \quad X_1 = Y_1 \setminus \nu(K_1).$$

It will follow from the course of our proof that if ϕ and (Y_i, K_i) are as in the statement, then there is an orientation-reversing diffeomorphism g as in the statement. Hence we will assume that such a g exists, and prove that it has the stated form.

We assume that g is induced by a pair of maps, g_0 and g_1, as

$$g_0 \circlearrowright X_0, \quad X_1 \circlearrowleft g_1.$$

The maps g_0 and g_1 induce a map on the quotient space if and only if

$$\phi \circ g_0 = g_1 \circ \phi. \tag{2-4}$$

The proof of (2-2) shows that for $i = 0, 1$, we have $g_i|_{\partial X_i} \in \{e, -e\}$ where

$$e = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Equation (2-4) implies that

$$\phi \circ e = \pm e \circ \phi.$$

It is easy to see that this restricts ϕ to be one of four matrices

$$\phi = \pm e \quad \text{and} \quad \phi = \pm \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Note that if $\phi = \pm e$, then the splice $\mathrm{Sp}_\phi(K_0, K_1)$ has $b_1 = 1$, so we exclude this case and restrict to the second case. We observe that in the latter case, we have

$$\phi \circ e = -e \circ \phi.$$

In particular, we conclude from (2-4) that $g_0 = -g_1$. Note that this corresponds exactly to one of K_0 and K_1 being negative amphichiral, and the other being positive amphichiral. □

2.3 Factorizations in $\mathrm{SL}_2(\mathbb{Z})$

In this section, we describe some straightforward algebra which will be used in the subsequent section on Kirby calculus.

We consider the elements $\psi_n^+ \in \mathrm{SL}_2(\mathbb{Z})$, given by

$$\psi_n^+ = \begin{pmatrix} n & 1 \\ -(1+n^2) & -n \end{pmatrix}.$$

We define the following elements of $\mathrm{SL}_2(\mathbb{Z})$:

$$T_n := \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad H := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Lemma 2.5 *The map ψ_n^+ may be written as*

$$\psi_n^+ = HT_{-n}HT_nH.$$

The proof is a straightforward computation, which we leave to the reader.

2.4 Kirby calculus

We can now translate Lemma 2.5 into Kirby calculus. Our main result is the following:

Proposition 2.6 *Let $K \subseteq Y$ be a knot. The manifold $\mathrm{Sp}_{\phi_n^+}(K, mK)$ has a Kirby diagram as shown in Figure 1.*

We begin with the following elementary topological lemma.

Lemma 2.7 *Let (Y_0, K_0) and (Y_1, K_1) be knots with Morse framings λ_0 and λ_1, respectively. Let ϕ be the gluing map which identifies the meridian μ_0 with μ_1, and which maps λ_0 to $-\lambda_1$. (Here, $-\lambda_2$ denotes the Morse framing λ_2 with the parametrization reversed). Then*

$$(Y_0 \setminus \nu(K_0)) \cup_\phi (Y_1 \setminus \nu(K_1))$$

is equal to $(Y_0 \# Y_1)_{\lambda_0+\lambda_1}(K_0 \# K_1)$.

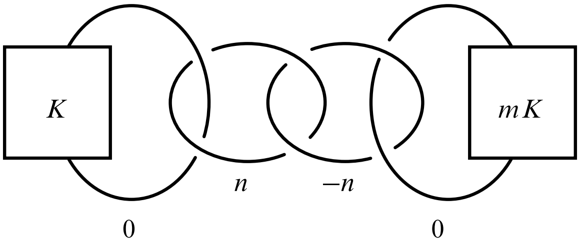


Figure 1: The manifold $\mathrm{Sp}_{\phi_n^+}(K, mK)$. We view the box labeled by K as being inside Y.

See [6, Lemma 6.1; 8, Lemma 7.1] for a proof. See also [26, Section 1.1.7].

The above lemma extends in a straightforward manner to link complements when we take the connected sum along a single component. We note that the Hopf link has complement $\mathbb{T}^2 \times [0,1]$. The meridian of the first component of the Hopf link is equal to the longitude of the second up to sign, and vice versa. Therefore from a factorization of the gluing diffeomorphism ψ_n^+, as in Lemma 2.5, we may read off a Kirby calculus description of $\mathrm{Sp}_{\phi_n^+}(K, mK)$. Namely, we start with K, which we give framing 0. Reading the factorization

$$\psi_n^+ = HT_{-n}HT_nH$$

from right to left we form a Kirby calculus presentation inductively as follows:

(1) Start with K, given framing 0.

(2) For each H, we take the connected sum with a $(0,0)$-framed Hopf link.

(3) For each T_n, we add n to the framing of the most-recently added component in this process.

(4) We finish by taking the connected sum of the final unknot, which has framing 0, with mK.

Note that since we are assuming that K is reversible, we do not need to worry about the sign of the clasps that we add when taking the connected sum with Hopf links.

We now describe some Kirby calculus moves in our description of $\mathrm{Sp}_{\phi_n^+}(K, mK)$ which will be helpful later on.

Lemma 2.8 *Let K be a reversible knot in an integer homology sphere Y, and $n \in \mathbb{Z}$. Then there are reversible knots $K' \subseteq Y'$ and $K'' \subseteq Y''$, where Y' and Y'' are also integer homology 3-spheres, so that*

$$\mathrm{Sp}_{\phi_n^+}(K, mK) = \mathrm{Sp}_{\phi_{n+1}^+}(K', mK') = \mathrm{Sp}_{\phi_{n-1}^+}(K'', mK'').$$

Proof The proof is to take the Kirby calculus description in Figure 1, and blow up the clasps between K and the unknotted component clasped with it and between mK and the unknotted component clasped with it. In this manner, n can be increased or decreased. Then $Y' = Y_{+1}(K)$ and K' is the dual knot of K. We give K' blackboard framing 1 in Figure 2, but note that this corresponds to the Seifert framing for K' if we view it as living in $Y_{+1}(K)$. Hence, we obtain a description of the same form as Figure 1, except with n replaced by $n+1$. □

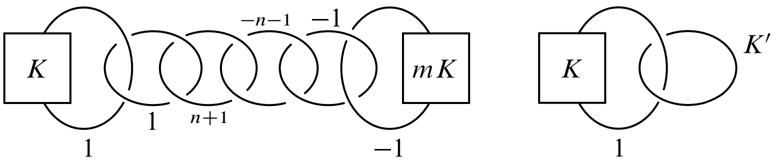


Figure 2: Left: an alternate surgery description of $\mathrm{Sp}_{\phi_n^+}(K, mK)$, obtained by blowing up two clasps in Figure 1. This identifies $\mathrm{Sp}_{\phi_n^+}(K, mK)$ with $\mathrm{Sp}_{\phi_{n+1}^+}(K', mK')$ where $K' \subseteq Y' = Y_{+1}(K)$ is the dual knot of K. Right: the knot $K' \subseteq Y_{+1}(K)$.

3 Heegaard Floer invariants of concordance and homology cobordism

In this section, we review some background on Heegaard Floer invariants of homology cobordism and knot concordance. We focus on Hendricks and Manolescu's *involutive Heegaard Floer homology* [12], which we review in Section 3.1, as well as the notion of a knot-like complex, which we review in Section 3.2. We presume the reader is familiar with ordinary Heegaard Floer homology for three-manifolds [23; 24] and knots [22; 25].

3.1 Iota-complexes and involutive Heegaard Floer homology

In this section we briefly introduce the structure of the involutive Heegaard Floer invariants, with a focus on the properties of local equivalence.

We begin with certain algebraic definitions. Throughout, $\mathbb{F}$ denotes the field with two elements, U is a variable of degree -2, and $\mathbb{F}[U]_d$ is the graded module such that $\mathrm{gr}(1) = d$.

Definition 3.1 An *iota-complex* (or *ι-complex*) (C, ι) is a chain complex C, which is free and finitely generated over $\mathbb{F}[U]$, equipped with an endomorphism ι. Here $\mathbb{F}$ is the field of 2 elements, and U is a formal variable with grading -2. Furthermore, the following hold:

(1) C is equipped with a $\mathbb{Z}$-grading, compatible with the action of U. We call this grading the *Maslov* or *homological* grading.

(2) There is a grading-preserving isomorphism $U^{-1}H_*(C) \cong \mathbb{F}[U, U^{-1}]$.

(3) ι is a grading-preserving chain map and $\iota^2 \simeq \mathrm{id}$.

Given two iota-complexes (C_1, ι_1) and (C_2, ι_2), a homogeneously graded $\mathbb{F}[U]$-chain map $f : C_1 \to C_2$ is said to be an *ι-homomorphism* if $\iota_2 \circ f + f \circ \iota_1 \simeq 0$. Two iota-complexes (C_1, ι_1) and (C_2, ι_2) are called *ι-equivalent* if there is a homotopy equivalence $\Phi : C_1 \to C_2$ which is an ι-homomorphism.

Heegaard Floer homology associates to any closed oriented 3-manifold Y equipped with a Spin^c structure $\mathfrak{s}$ an $\mathbb{F}[U]$-chain complex $\mathrm{CF}^-(Y, \mathfrak{s})$, well defined up to homotopy equivalence. If $\mathfrak{s}$ is self-conjugate, involutive Heegaard Floer homology considers the additional data of a homotopy involution ι on $\mathrm{CF}^-(Y, \mathfrak{s})$. In the case that Y is a rational homology 3-sphere, $(\mathrm{CF}^-(Y, \mathfrak{s}), \iota)$ is an iota-complex. Hendricks and Manolescu [12] prove that pair $(\mathrm{CF}^-(Y, \mathfrak{s}), \iota)$ is well defined up to the notion of iota-equivalence described above.

Continuing with algebra, the tensor product of iota-complexes (C_1, ι_1) and (C_2, ι_2) is given by

$$(C_1, \iota_1) \otimes (C_2, \iota_2) := (C_1 \otimes_{\mathbb{F}[U]} C_2, \iota_1 \otimes \iota_2). \tag{3-1}$$

Moreover, Hendricks, Manolescu, and Zemke [13] establish that

$$(\mathrm{CF}^-(Y_1 \# Y_2, \mathfrak{s}_1 \# \mathfrak{s}_2), \iota) \simeq (\mathrm{CF}^-(Y_1, \mathfrak{s}_1), \iota_1) \otimes (\mathrm{CF}^-(Y_2, \mathfrak{s}_2), \iota_2),$$

where $\simeq$ denotes homotopy-equivalence of iota-complexes.

Definition 3.2 Suppose (C, ι) and (C', ι') are two iota-complexes.

1. A *local map* from (C, ι) to (C', ι') is a grading-preserving ι-homomorphism $F : C \to C'$, which induces an isomorphism from $U^{-1}H_*(C)$ to $U^{-1}H_*(C')$.
2. We say that (C, ι) are (C', ι') are *locally equivalent* if there is a local map from (C, ι) to (C', ι'), as well as a local map from (C', ι') to (C, ι). We say that (C, ι) is *locally trivial* if it is locally equivalent to $(\mathbb{F}[U]_0, \mathrm{Id})$.

The set of local equivalence classes forms an abelian group, denoted by $\mathfrak{I}$, with product given by the operation $\otimes$ in (3-1) [13, Section 8]. Inverses are given by dualizing both the chain complex C and the map ι with respect to $\mathbb{F}[U]$. The map

$$Y \mapsto [(\mathrm{CF}^-(Y), \iota)]$$

determines a homomorphism from $\Theta^3_{\mathbb{Z}}$ to $\mathfrak{I}$ [13, Theorem 1.8].

The local equivalence classes of nonzero integer surgeries on knots are computed in [11, Theorem 1.6(2)]. For our purposes, the important case is the following.

Lemma 3.3 [11, Theorem 1.6(2)] *For $n > 0$, the local equivalence class of $(\mathrm{CF}^-(S^3_{2n}(K), [n]), \iota)$ has the form*

$$\begin{tikzcd}
A_n^-(K) \arrow[dr, "v"'] & & A_n^-(K) \arrow[dl, "v"] \\
& B_n^-(K) &
\end{tikzcd}$$

where $A_n^-(K)$ and $B_n^-(K)$ are subcomplexes of the knot Floer complex of K, v is a particular map between them, and the involution swaps the two copies of $A_n^-(K)$, and fixes $B_n^-(K)$. The gradings on the above are induced by the Maslov grading on the knot Floer complex, shifted up by the Heegaard Floer correction term $d(L(2n,1),[n])$ of the lens space $L(2n,1)$ in the corresponding Spin^c *structure.*

One straightforward corollary is the following:

Corollary 3.4 *For $n > 0$, if $K \subseteq S^3$ is a knot such that $d(A_n^-(K)) = 0$, then $\big(\mathrm{CF}^-(S^3_{2n}(K), [n]), \iota\big)$ is locally equivalent to $(\mathbb{F}[U]_d, \mathrm{id})$, where $d = d(L(2n,1),[n])$.*

Proof We note that $B_n^-(K) \simeq \mathbb{F}[U]$, so using the same logic as in the proof of [11, Proposition 3.24], we may replace it with a copy of $\mathbb{F}[U]$. By the classification theorem for finitely generated chain complexes over $\mathbb{F}[U]$, we can write $A_n^-(K)$ as a sum of one tower $\mathbb{F}[U]$, as well as some number of 2-step complexes of the form $\mathbb{F}[U] \xrightarrow{U^i} \mathbb{F}[U]$. Write $\boldsymbol{x}_l$ and $\boldsymbol{x}_r$ for tower generators of the two copies of $A_n^-(K)$, which is to say, generators for the copy of $\mathbb{F}[U]$ in the basis chosen. The map v sends $\boldsymbol{x}_l$ and $\boldsymbol{x}_r$ to a nonzero element of the tower $\mathbb{F}[U]$. Since $d(A_n^-(K)) = 0$, we conclude that $v(\boldsymbol{x}_l) = v(\boldsymbol{x}_r) = 1$. On a two step subcomplex of the left copy of $A_n^-(K)$, say with generators $\boldsymbol{a}$ and $\boldsymbol{b}$ such that $\partial(\boldsymbol{a}) = U^i\boldsymbol{b}$, we must have $v(\boldsymbol{b}) = 0$ because v is a chain map. If $v(\boldsymbol{a})$ is nonzero, then perform a change of basis, adding a multiple of $\boldsymbol{x}_l$ to $\boldsymbol{a}$. After this change of basis, $v(\boldsymbol{a}) = 0$. We do the same change of basis to the right copy of $A_n^-(K)$. After this change of basis, it becomes apparent that the complex in the statement of Lemma 3.3 is locally equivalent to

$$\begin{tikzcd}
\mathbb{F}[U]_d \arrow[dr, "1"'] & & \mathbb{F}[U]_d \arrow[dl, "1"] \\
& \mathbb{F}[U]_d &
\end{tikzcd}$$

where the involution is reflection, and d denotes $d(L(2n,1),[n])$. The above is homotopy equivalent to $(\mathbb{F}[U]_d, \mathrm{id})$. □

3.2 Knot-like complexes

We now recall the standard notion of knot-like complexes in Heegaard Floer theory. There are many different variations on the definition in the literature due to many different authors. The earliest version is Hom's notion of ε-equivalence [14]. See also [3; 5; 28] for other variations.

Definition 3.5 A *knot-like* complex C is a finitely generated, free chain complex over a 2-variable polynomial ring $\mathbb{F}[U,V]$ satisfying the following:

(1) C is equipped with a $\mathbb{Z}\times\mathbb{Z}$-valued bigrading, denoted by $(\mathrm{gr}_w, \mathrm{gr}_z)$, which has the property that $(\mathrm{gr}_w - \mathrm{gr}_z)/2$ is integrally valued. The variable U has bigrading $(-2,0)$ and the variable V has bigrading $(0,-2)$.

(2) There is a grading-preserving isomorphism $(U,V)^{-1}H_*(C) \cong \mathbb{F}[U,V,U^{-1},V^{-1}]$.

(3) ∂ has bigrading $(-1,-1)$.

A local map from C_0 to C_1 (where C_i are knot-like complexes) consists of a grading-preserving $\mathbb{F}[U,V]$ linear chain map $F : C_0 \to C_1$ such that F induces an isomorphism from $(U,V)^{-1}H_*(C_0)$ to $(U,V)^{-1}H_*(C_1)$. We say that C_0 and C_1 are *locally equivalent* if there exist local maps from C_0 to C_1 and from C_1 to C_0. A knot-like complex C is *locally trivial* if it is locally equivalent to a rank-one complex $\mathbb{F}[U,V]$ wherein $1 \in \mathbb{F}[U,V]$ is concentrated in grading $(0,0)$. Note that C is locally trivial if and only if there is an isomorphism

$$C \cong \mathbb{F}[U,V] \oplus A,$$

where A is a summand of C such that $(U,V)^{-1}H_*(A) \cong 0$.

If $K \subseteq S^3$, the full version of the knot Floer complex $\mathrm{CFK}^-(K)$ is a knot-like complex. If K is a slice knot, then $\mathrm{CFK}^-(K)$ is locally trivial.

4 Proofs of the main results

We first observe that if Z is a symmetric splice, since Z has order at most two in $\Theta^3_{\mathbb{Z}}$, we must have $d(Z)=0$ since d is a homomorphism.

Lemma 4.1 *Suppose that Z is the homology sphere obtained by splicing (Y,K) and $(-Y,-K)$ using the gluing map ϕ_0^+ from Proposition 2.2, where Y is a homology 3-sphere. Then there is a negative definite* Spin *cobordism W from Z to $\mathbb{RP}^3$. The 4-manifold W has a unique self-conjugate* Spin^c *structure $\mathfrak{s}$. Further, letting $d(W,\mathfrak{s})$ denote the grading shift of the map associated to the cobordism $(W,\mathfrak{s})$, we have*

$$d(W,\mathfrak{s})=d(\mathbb{RP}^3,\mathfrak{s}|_{\mathbb{RP}^3})-d(Z).$$

Proof We begin with the Kirby calculus presentation from Figure 1. We can blow-down one of the unknots to obtain a Kirby calculus description of Z as Dehn surgery on a knot $K\#H\#-K\subseteq Y\#-Y$, where H denotes a Hopf link. That is, we add a clasp between K and $-K$. (Note that the sign of the clasp is not important since (Y,K) is reversible). The two components are given framing 0. We now blow up the clasp to obtain the 3-component link $K\cup U\cup mK$ with clasps between the components. Each component is given framing -1. There is a cobordism X from Z to a manifold Z' by performing -1 surgery on a meridian of the unknot U, where this knot is given Seifert framing -2 inside of Z. The result is -2 surgery on $K\#-K\subseteq Y\#-Y$. The pair $(Y\#-Y,K\#-K)$ is homology concordant to (S^3,U). Therefore Z' admits a homology cobordism to $\mathbb{RP}^3$, viewed as -2 surgery on the unknot. Let W denote the composition of these two cobordisms. The cobordism W is shown in Figure 3.

We observe that $d(Z)=0$ since $Z\cong -Z$. On the other hand, we compute that the shift in grading for the Spin structure on W is

$$\tfrac{1}{4}(-2\chi(W)-3\sigma(W))=\tfrac{1}{4}.$$

We note that the d-invariants of the two Spin^c structures (= Spin structures) on $\mathbb{RP}^3$ are $\frac{1}{4}$ and $-\frac{1}{4}$. Since the Maslov grading takes values in a single coset of $\mathbb{Q}/\mathbb{Z}$ in each Spin^c structure, it follows that the Spin structure on W restricts to the Spin^c structure on $\mathbb{RP}^3$ which has d-invariant $\frac{1}{4}$. □

Using Lemma 4.1, we now prove Theorem 1.1(1), which concerns symmetric splices of Type-1:

Proof of Theorem 1.1(1) Let Z be a symmetric splice of Type-1. By Proposition 2.2, Z can be written as a splice $\mathrm{Sp}_{\phi_n^{\pm}}(K,mK)$ for some pair (Y,K), where Y is a homology sphere and (Y,K) is reversible. By Lemmas 2.3 and 2.8, we may assume that $n=0$ by changing (Y,K) appropriately. Applying Lemma 4.1 gives a local map from $(\mathit{CF}^-(Z),\iota)$ to $(\mathit{CF}^-(\mathbb{RP}^3,\mathfrak{s}),\iota)$ for the Spin structure $\mathfrak{s}$ with $d(\mathbb{RP}^3,\mathfrak{s})=\frac{1}{4}$. Since the grading shift of the cobordism map is also $\frac{1}{4}$, and $(\mathit{CF}^-(\mathbb{RP}^3,\mathfrak{s}),\iota)\cong(\mathbb{F}[U]_{1/4},\mathrm{id})$, we conclude that there is a local map from $(\mathit{CF}^-(Z),\iota)$ to the trivial complex. Dualizing and using the fact that $Z\cong -Z$ gives a local map in the opposite direction. □

Proof of Theorem 1.1(2) The proof is similar to the proof of Theorem 1.1(1). By Proposition 2.4, the 3-manifold Z can be written as $\mathrm{Sp}_{\phi_0^+}(K_0,K_1)$ where $K_0,K_1\subseteq S^3$ are positive and negative amphichiral knots, respectively. By adapting the argument from Theorem 1.1(1), we obtain a negative definite Spin

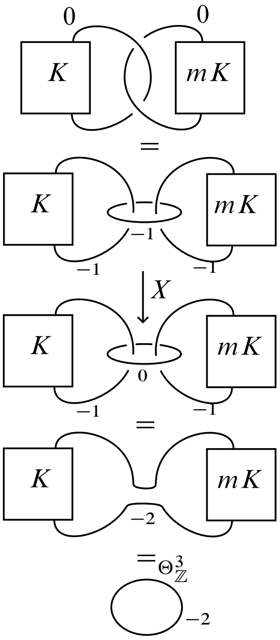


Figure 3: The cobordism W from Z to $\mathbb{RP}^3$.

cobordism from Z to $S^3_{-2}(K_0 \# K_1)$ which shifts the Maslov grading by $\frac{1}{4}$. Assuming (up to relabeling) that K_0 is positive amphichiral and K_1 is negative amphichiral, we observe that

$$S^3_{-2}(K_0 \# K_1) \cong -S^3_{+2}(K_0 \# -K_1),$$

where $-K_1$ denotes K_1 with its string orientation reversed. Therefore, Corollary 3.4 implies that

$$\big(\mathit{CF}^-(S^3_{+2}(K_0 \# -K_1), [1]), \iota\big) \sim_{\mathrm{loc}} (\mathbb{F}[U]_{-1/4}, \mathrm{id}).$$

Dualizing, we obtain that

$$\big(\mathit{CF}^-(S^3_{-2}(K_0 \# K_1), [1]), \iota\big) \sim_{\mathrm{loc}} (\mathbb{F}[U]_{1/4}, \mathrm{id}).$$

It follows that there is a local map from $(\mathit{CF}^-(Z), \iota)$ to $(\mathbb{F}[U], \mathrm{id})$. Since $Z \cong -Z$, we conclude that $(\mathit{CF}^-(Z), \iota)$ is locally trivial. □

We now prove Proposition 1.3, most of which we have already proven:

Proof of Proposition 1.3 Part (1) follows from Lemma 4.1, above. Part (3) is similar, and is described in our proof of Theorem 1.1(2). Finally Part (2) is obtained by composing the cobordism from Part (1) with the natural negative definite cobordism from $\mathbb{RP}^3$ to $\varnothing$, namely the disk bundle over S^2 with Euler number -2. □

Proof of Proposition 1.5 The argument is essentially identical to the proof of Part (1) of Theorem 1.1, but where the notation is adjusted to be for Seiberg–Witten Floer spectra in the setting of [18]. In particular, local equivalence of Pin(2)-equivariant spectra is defined just as in Definition 3.2 above, except Pin(2)-equivariant spectra take the place of iota-complexes (see [27, Definition 2.7] and surrounding discussion).

Let Z be a symmetric splice of Type-1. Lemma 4.1 gives a local map

$$\Sigma^{\frac{1}{16}\mathbb{H}}\,\mathrm{SWF}(Z) \to \mathrm{SWF}(\mathbb{RP}^3, \mathfrak{s}),$$

with $\mathfrak{s}$ as in the proof of Theorem 1.1. We refer the reader to [17] for the definition of the (formal) fractional suspension. Meanwhile, $\mathrm{SWF}(\mathbb{RP}^3, \mathfrak{s}) = S^{\frac{1}{16}\mathbb{H}}$, and so we have a local map $\mathrm{SWF}(Z) \to S^0$. Using that $Z \cong -Z$, we have a local map $\mathrm{SWF}(-Z) \to S^0$; furthermore, for general integer homology spheres X, we have $\mathrm{SWF}(X)$ and $\mathrm{SWF}(-X)$ are Spanier–Whitehead dual. As a consequence, if there is a local map $\mathrm{SWF}(-Z) \to S^0$ then there is a local map $S^0 \to \mathrm{SWF}(Z)$. Thus $S^0 \le \mathrm{SWF}(Z) \le S^0$ in local equivalence, and so $\mathrm{SWF}(Z)$ is locally trivial as a Pin(2)-spectrum. □

Acknowledgements

We are grateful to Jen Hom for helpful conversations, and to the referee for helpful comments and corrections. Hendricks was partially supported by NSF grant DMS-2019396. Stoffregen was partially supported by NSF grant DMS-2203828. Zemke was partially supported by NSF grant DMS-2204375 and a Sloan Research Fellowship.

References

[1] **A Daemi**, *Chern–Simons functional and the homology cobordism group*, Duke Math. J. 169:15 (2020) 2827–2886 MR

[2] **A Daemi**, **C Scaduto**, *Equivariant aspects of singular instanton Floer homology*, Geom. Topol. 28:9 (2024) 4057–4190 MR

[3] **I Dai**, **J Hom**, **M Stoffregen**, **L Truong**, *More concordance homomorphisms from knot Floer homology*, Geom. Topol. 25:1 (2021) 275–338 MR

[4] **I Dai**, **J Hom**, **M Stoffregen**, **L Truong**, *An infinite-rank summand of the homology cobordism group*, Duke Math. J. 172:12 (2023) 2365–2432 MR

[5] **I Dai**, **J Hom**, **M Stoffregen**, **L Truong**, *Homology concordance and knot Floer homology*, Math. Ann. 390:4 (2024) 6111–6186 MR

[6] **S Fukuhara**, **N Maruyama**, *A sum formula for Casson's λ-invariant*, Tokyo J. Math. 11:2 (1988) 281–287 MR

[7] **D E Galewski**, **R J Stern**, *Classification of simplicial triangulations of topological manifolds*, Ann. of Math. (2) 111:1 (1980) 1–34 MR

[8] **C M Gordon**, *Dehn surgery and satellite knots*, Trans. Amer. Math. Soc. 275:2 (1983) 687–708 MR

[9] **K Hendricks**, **J Hom**, **T Lidman**, *Applications of involutive Heegaard Floer homology*, J. Inst. Math. Jussieu 20:1 (2021) 187–224 MR

[10] **K Hendricks**, **J Hom**, **M Stoffregen**, **I Zemke**, *On the quotient of the homology cobordism group by Seifert spaces*, Trans. Amer. Math. Soc. Ser. B 9 (2022) 757–781 MR

[11] **K Hendricks**, **J Hom**, **M Stoffregen**, **I Zemke**, *Surgery exact triangles in involutive Heegaard Floer homology* (2025) arXiv 2011.00113

[12] **K Hendricks**, **C Manolescu**, *Involutive Heegaard Floer homology*, Duke Math. J. 166:7 (2017) 1211–1299 MR

[13] **K Hendricks**, **C Manolescu**, **I Zemke**, *A connected sum formula for involutive Heegaard Floer homology*, Selecta Math. (N.S.) 24:2 (2018) 1183–1245 MR

[14] **J Hom**, *The knot Floer complex and the smooth concordance group*, Comment. Math. Helv. 89:3 (2014) 537–570 MR

[15] **A Kawauchi**, *Rational-slice knots via strongly negative-amphicheiral knots*, Commun. Math. Res. 25:2 (2009) 177–192 MR

[16] **F Lin**, *A Morse–Bott approach to monopole Floer homology and the triangulation conjecture*, Mem. Amer. Math. Soc. 1221, Amer. Math. Soc., Providence, RI (2018) MR

[17] **C Manolescu**, *Seiberg–Witten–Floer stable homotopy type of three-manifolds with* $b_1 = 0$, Geom. Topol. 7 (2003) 889–932 MR

[18] **C Manolescu**, *Pin(2)-equivariant Seiberg–Witten Floer homology and the triangulation conjecture*, J. Amer. Math. Soc. 29:1 (2016) 147–176 MR

[19] **T Matumoto**, *Triangulation of manifolds*, from "Algebraic and geometric topology, II" (Stanford, CA, 1976) (R J Milgram, editor), Proc. Sympos. Pure Math. XXXII, Amer. Math. Soc., Providence, RI (1978) 3–6 MR

[20] **D McCullough**, *Homeomorphisms which are Dehn twists on the boundary*, Algebr. Geom. Topol. 6 (2006) 1331–1340 MR

[21] **Y Nozaki**, **K Sato**, **M Taniguchi**, *Filtered instanton Floer homology and the homology cobordism group*, J. Eur. Math. Soc. 26:12 (2024) 4699–4761 MR

[22] **P Ozsváth**, **Z Szabó**, *Holomorphic disks and knot invariants*, Adv. Math. 186:1 (2004) 58–116 MR

[23] **P S Ozsváth**, **Z Szabó**, *Holomorphic disks and three-manifold invariants*: *properties and applications*, Ann. of Math. (2) 159:3 (2004) 1159–1245

[24] **P Ozsváth**, **Z Szabó**, *Holomorphic disks and topological invariants for closed three-manifolds*, Ann. of Math. (2) 159:3 (2004) 1027–1158 MR

[25] **J A Rasmussen**, *Floer homology and knot complements*, PhD thesis, Harvard University (2003) MR arXiv math/0306378

[26] **N Saveliev**, *Invariants for homology* 3*-spheres*, Encyclopaedia of Mathematical Sciences 140, Springer (2002) MR

[27] **M Stoffregen**, *Pin(2)-equivariant Seiberg–Witten Floer homology of Seifert fibrations*, Compos. Math. 156:2 (2020) 199–250 MR

[28] **I Zemke**, *Connected sums and involutive knot Floer homology*, Proc. Lond. Math. Soc. (3) 119:1 (2019) 214–265 MR

KRISTEN HENDRICKS kristen.hendricks@rutgers.edu
Department of Mathematics, Rutgers University, Piscataway, NJ, United States

MATTHEW STOFFREGEN stoffre1@msu.edu
Department of Mathematics, Michigan State University, East Lansing, MI, United States

IAN ZEMKE izemke@uoregon.edu
Department of Mathematics, University of Oregon, Eugene, OR, United States

Received: July 15, 2024 Revised: March 21, 2025

Geometry & Topology Publications, an imprint of **mathematical sciences publishers**

msp
Algebraic & Geometric Topology **26**:5 (2026) 1923–1963
DOI: 10.2140/agt.2026.26.1923
Published: May 29, 2026

Lagrangian metric geometry with Riemannian bounds

JEAN-PHILIPPE CHASSÉ

We study collections of exact Lagrangian submanifolds respecting some uniform Riemannian bounds, which we equip with a metric naturally arising in symplectic topology (e.g., the Lagrangian Hofer metric or the spectral metric). We exhibit many metric and symplectic properties of these spaces, such that they have compact completions and that they contain only finitely many Hamiltonian isotopy classes. We then use this to exclude many unusual phenomena from happening in these bounded spaces. Taking limits in the bounds, we also conclude that there are at most countably many Hamiltonian isotopy classes of exact Lagrangian submanifolds in a Liouville manifold. Under some mild topological assumptions, we get analogous results for monotone Lagrangian submanifolds with a fixed monotonicity constant. Finally, in the process of showing these results, we get new results on the Riemannian geometry of cotangent bundles and surfaces which might be of independent interest.

1 Introduction and main results

In [11; 12], the author showed that all the familiar metrics d between Lagrangian submanifolds in symplectic topology (e.g., the Lagrangian Hofer metric or the spectral metric) behave like the classical Hausdorff metric when restricted to the space $\mathscr{L}_k^e$ of λ-exact Lagrangian submanifolds of a Liouville manifold $(M, \omega = d\lambda)$ which are "geometrically bounded by k". Broadly speaking, being geometrically bounded by k ensures that all Lagrangian submanifolds considered are contained in the same compact and have curvature and volume uniformly bounded — we give the precise definition in Section 2.1. In this paper, we continue the study of the metric spaces $(\mathscr{L}_k^e, d)$. Most notably, we will be concerned with issues of compactness and local connectedness.

Furthermore, we extend our results to the analogous space $\mathscr{L}_k^{m(\rho)}$ of ρ-monotone Lagrangian submanifolds which "bound enough disks" — the ambient symplectic manifold M is either closed or convex at infinity in that case. We will give the precise definition of what we mean here by "bounding enough disks" below — see Section 2.1 — but we note right away that this condition is automatically satisfied if either the Lagrangian submanifolds or the symplectic manifold we consider are simply connected.

In what follows, $\mathscr{L}_k^\star$ will denote either the spaces $\mathscr{L}_k^e$ or $\mathscr{L}_k^{m(\rho)}$. We will also denote by $\mathscr{L}_\infty^\star$ the space of *all* Lagrangian submanifolds respecting $\star$. Moreover, the metric d that we consider will be a so-called Chekanov-type metric [11] and be bounded from above by the Lagrangian Hofer metric. In practice, one can think of d as one of the following metrics.

- $\boldsymbol{d = d_H}$ This is the case of the Lagrangian Hofer metric, which is due to Chekanov [14].
- $\boldsymbol{d = \gamma}$ This is the case of the spectral metric, originally due to Viterbo [43] for Lagrangian submanifolds of T^*L Hamiltonian isotopic to the zero-section. The metric may be extended to all exact Lagrangian

---

MSC2020: primary 53D12; secondary 51F99, 53C17, 53C22.

submanifolds of T^*L by work from Fukaya, Seidel, and Smith [18; 19], Abouzaid [2], and Kragh [30]. In general, it has also been defined for weakly exact Lagrangian submanifolds by Leclercq [31] and monotone ones with nonvanishing quantum homology by Kislev and Shelukhin [28], following work of Leclercq and Zapolsky [32].

- $\boldsymbol{d = \gamma_{\mathrm{ext}}}$ This is a variant of the usual spectral metric, as defined by Kislev and Shelukhin [28].
- $\boldsymbol{d = \hat{d}_S^{\mathscr{F},\mathscr{F}'}}$ These are the shadow metrics appearing in work of Biran, Cornea, and Shelukhin [7; 15].
- $\boldsymbol{d = \widehat{D}^{\mathscr{F},\mathscr{F}'}}$ There are possibly many other weighted fragmentation pseudometrics — as defined by Biran, Cornea, and Zhang [8] — that belong to this class.

If $d = \gamma$ and $\star = m(\rho)$, we make a slight abuse of notation and still denote by $\mathscr{L}_k^{m(\rho)}$ the space of all ρ-monotone k-geometrically bounded Lagrangian submanifolds which both bound enough disks and have nonvanishing quantum homology. Otherwise, we also take the convention that if d is not properly defined between L and L', then $d(L, L') = +\infty$.

With this notation settled down, we enunciate the main principles that summarise our results.

(1) The space $(\mathscr{L}_k^\star, d)$ has nice metric properties, which highly restricts the symplectic phenomena which can happen within that space.

(2) Metric properties of $(\mathscr{L}_k^\star, d)$ induce topological properties on the limit space $(\mathscr{L}_\infty^\star, d)$.

We now explain our main results and how their corollaries showcase the two general principles above. Note that some of these corollaries are not as direct as others; we will properly prove all of them later in the paper.

Theorem A *On $\mathscr{L}_k^\star$, all possible choices of d in the above list induce the same topology and have homeomorphic completions $\widehat{\mathscr{L}}_k^\star$. Moreover, that completion $\widehat{\mathscr{L}}_k^\star$ is compact.*

In term of the first principle, we get the following two corollaries.

Corollary *The subspaces $\mathscr{L}_k^{L_0} := (\mathrm{Ham}(M) \cdot L_0) \cap \mathscr{L}_k^\star$, where $L_0 \in \mathscr{L}_\infty^\star$, have finite diameter in d. If $M = T^*L$ and $d = \gamma$, then the same holds on $\mathscr{L}_k^e$.*

Note that there are many cases where it is known that $\mathrm{Ham}(M) \cdot L_0 = \mathscr{L}_\infty^{L_0}$ has infinite diameter in d, e.g., $\mathscr{L}_\infty^{S^1}(T^*S^1) = \mathscr{L}_\infty^e(T^*S^1)$ has infinite Hofer diameter [27]. On the other hand, it is conjectured — and has been proven in a many cases [23; 38; 39; 44] — that $\mathscr{L}_\infty^e(T^*Q)$ has finite diameter in the spectral metric. Therefore, without any Riemannian bound, the finiteness is highly dependent on L_0, M, and d. Note that we previously proved such a boundedness result [12] on a neighbourhood of some L in $\mathscr{L}_k^e$. The improvement here is thus in going from a neighbourhood to the whole space.

As we shall see below, we also manage to extend our study to spaces of graphs of Hamiltonian diffeomorphisms of closed monotone manifolds, whether these graphs bound enough disks or not. Through this, we can rule out something like Ostrover's example [35] from happening in the corresponding $\mathscr{L}_k$ spaces.

Corollary *Suppose that M is closed and monotone. On the subspace of Hamiltonian diffeomorphisms $\varphi \in \mathrm{Ham}(M)$ whose graph has geometry bounded by k in $M \times M$, the Hofer norm $\|\cdot\|_H$ is bounded.*

Moreover, the Hofer norm on diffeomorphisms and the Lagrangian Hofer distance between their graphs induce the same topology on that space.

In term of the second principle, we get the following result from Theorem A.

Corollary *The limit space* $(\mathscr{L}^\star_\infty, d)$ *is separable, i.e., it admits a countable dense subset.*

Following separate discussions with Humilière and with Shelukhin, it seems that this result was well-accepted folklore on $\mathscr{L}^{L_0}_\infty = \mathrm{Ham}(M) \cdot L_0$. Therefore, the innovation of the above corollary seems to be on the number of Hamiltonian isotopy classes of $\mathscr{L}^\star_\infty$; this is precisely what we explore below through our second theorem.

Theorem B *The space* $\mathscr{L}^\star_k$ *contains only finitely many Hamiltonian isotopy classes. Furthermore, there is an* $A = A(k) > 0$ *such that*

$$d(L, L') \geq A,$$

whenever $L, L' \subseteq \mathscr{L}^\star_k$ *are* ***not*** *Hamiltonian isotopic.*

Obviously, the A-bound is trivial when $d = d_H$ — and potentially when $d = \gamma$ — as we have taken the convention that $d_H(L, L') = \infty$ whenever L and L' are not Hamiltonian isotopic. However, there are examples of Lagrangian submanifolds which are not Hamiltonian isotopic but are a finite distance apart in a shadow metric [7].

We note that Theorem B already fits within the motif of the first principle. However, we can go even further in this direction, as the corollaries below show.

First, we get a result on the vanishing of entropy of symplectomorphisms preserving $\mathscr{L}^\star_k$ in some form. We refer the reader to Section 4.4 and the references therein for the definition of barcode and categorical entropy.

Corollary *Let* ψ *be a symplectomorphism of* M *such that* $\psi(\mathscr{L}^\star_\infty) = \mathscr{L}^\star_\infty$*. If* L *is such that the sequence* $\{\psi^\nu(L)\}$ *is fully contained in some* $\mathscr{L}^\star_k$*, then there is some* N *such that* $\psi^N(L)$ *is Hamiltonian isotopic to* L*. Furthermore, for such* ψ *and* L*, the barcode entropy* $\hbar(\psi; L, L')$ *vanishes for any* $L' \in \mathscr{L}^\star_\infty$*.*

More generally, if the Lagrangian submanifolds $L_1, \dots, L_\ell$ *split-generate the derived Fukaya category* $\mathrm{DFuk}^\star(M)$ *of* M *and each sequence* $\{\psi^\nu(L_i)\}$ *is contained in a single* $\mathscr{L}^\star_k$*, then the categorical entropy* $h_{\mathrm{cat}}(\psi)$ *of* ψ *vanishes.*

The latter part on categorical entropy makes maybe more sense when formulated as its contraposition: if $h_{\mathrm{cat}}(\psi) > 0$, for any set of Lagrangian submanifolds $\{L_1, \dots, L_\ell\}$ split-generating the derived Fukaya category, there is some i such that the sequence $\{\psi^\nu(L_i)\}$ is not contained in any $\mathscr{L}^\star_k$. Note that we may suppose that such a Lagrangian submanifold L_i induces a nontrivial object in $\mathrm{DFuk}^\star(M)$. Therefore, this means that symplectomorphisms with positive categorical entropy must greatly deform some Lagrangian submanifold that is important in the Fukaya category, e.g., the norm of the second fundamental form or the volume of that Lagrangian must explode under iterations by such a symplectomorphism.

Secondly, we get a statement that relates the path-connected component of the completions $\widehat{\mathscr{L}}^\star_k$ to the Hamiltonian isotopy classes of $\mathscr{L}^\star_\infty$.

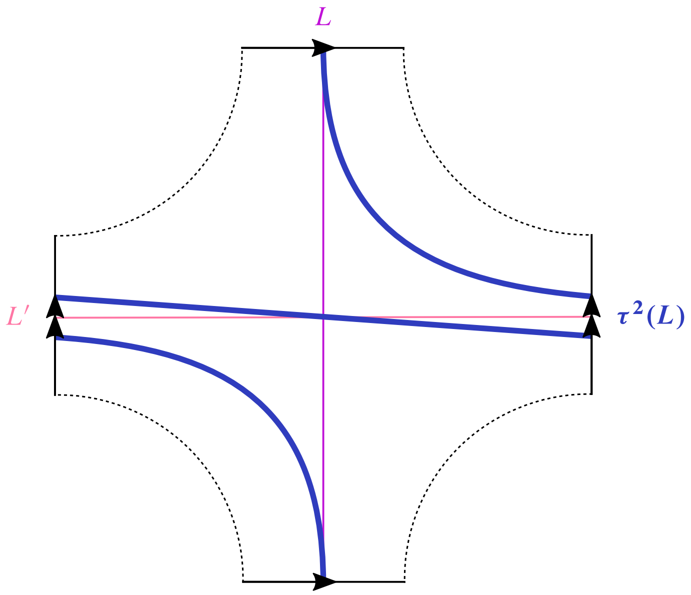


Figure 1: The circle L (thin, purple) and its image under two Dehn twists (thick, blue) along L' (thin, pink) inside some Liouville subdomain of $T^*S^1 \# T^*S^1$.

Corollary *Let $L, L' \in \mathscr{L}^\star_\infty$. Suppose that there exists a d-continuous path $t \mapsto L_t$ from L to L' in some $\widehat{\mathscr{L}}^\star_k$. Then, L is Hamiltonian isotopic to L'.*

We contrast this with a recent result from Arnaud, Humilière, and Viterbo [4], which shows that such a path always exists in the completion of $\mathscr{L}^e_\infty(T^*N)$ if $d = \gamma$. Therefore, our result can be seen as very slightly reducing the gap between that result and the nearby Lagrangian conjecture.

In terms of the second principle, we get the following new result.

Corollary *There are at most countably many Hamiltonian isotopy classes in $\mathscr{L}^\star_\infty$.*

Note that a Liouville manifold can have infinitely many Hamiltonian isotopy classes of exact Lagrangian submanifolds. For example, if M is the plumbing $T^*S^1 \# T^*S^1$ of two copies of T^*S^1, L is the zero-section in the first copy of T^*S^1, and $\tau : M \to M$ is the Dehn twist along the second copy L', then $\tau^\nu(L)$ is clearly in a different Hamiltonian isotopy class for each $\nu \geq 0$. See Figure 1 for an illustration.

Likewise, given $\rho > 0$, there are infinitely many Hamiltonian isotopy classes of ρ-monotone tori in $\mathbb{C}^3$ [5]. These automatically bound enough disks since $\pi_1(\mathbb{C}^3) = 0$. Therefore, as our approach does not perceive the topology of M, we cannot expect a better bound.

Remark 1.1 The construction by Ganatra, Pardon, and Shende of the wrapped Fukaya category [20] involves a choice of a countable set of exact isotopy classes of Lagrangian submanifolds which are cylindrical at infinity. While it is assumed in the paper that one can pick all exact isotopy classes by such a set, it is never proved. The above corollary shows that we can always make sure that this set contains every exact isotopy class of closed Lagrangian submanifolds, so that the wrapped Fukaya category always contains the full compact Fukaya category.

Moreover, we expect the techniques presented here to adapt to the cylindrical setting by working in a fixed sector, with exact Lagrangian submanifolds which extend to cylindrical ones in the completion, and with Riemannian metrics which have some standard form near the boundary. This would ensure that the wrapped Fukaya category of a Liouville sector can always be constructed in such a manner as to include all exact isotopy classes.

We also get a result on the local structure of the $\mathscr{L}_k^\star$ spaces in some cases.

Theorem C *Given an exact or monotone Lagrangian submanifold L of M, there is a Riemannian metric g on M making L totally geodesic and such that, for every $k > 0$, it admits a system of contractible neighbourhoods in $\mathscr{L}_k^\star$. If $\dim M = 2$, the latter part holds for every metric and every $k > 0$ such that $L \in \mathscr{L}_k^\star$.*

Though this result does not yield as many direct symplectic applications as the ones before, we can still use it to estimate the Hofer and spectral distances in some cases. More precisely, it is known [33] that, for graphs of exact 1-forms in T^*L, the Hofer and spectral distances agree and are given by

$$d_H(\operatorname{graph} df, \operatorname{graph} dg) = \gamma(\operatorname{graph} df, \operatorname{graph} dg) = \max|f-g| - \min|f-g|.$$

In particular, for every $f \in C^\infty(L)$, $t \mapsto \operatorname{graph}(t\,df)$ is a minimal geodesic in these metrics. However, when one embeds a neighbourhood of L in T^*L in M via a Weinstein neighbourhood, there could be a shorter path going through M. Nonetheless, through Theorem C and some estimates on Hausdorff-geodesics, we get the following estimate on how far from a minimal geodesic $t \mapsto \operatorname{graph}(t\,df)$ can be.

Corollary *Let $\Psi : D_r^* L \to M$ be a Weinstein neighbourhood of some $L \in \mathscr{L}_\infty^\star(M)$. Suppose either that the Riemannian metric g on M is as in Theorem C or that $\dim M = 2$. For every $k \geq 1$, there are constants $C > 0$ and $r' \in (0, r]$ with the following property. Whenever $f : L \to \mathbb{R}$ is such that $\Psi(\operatorname{graph} df) \in \mathscr{L}_k^\star$ and $|df| \leq r'$, we have that*

$$d(\Psi(t \operatorname{graph} df), \Psi(s \operatorname{graph} df)) \geq C(t-s)^2 \max|df|^2$$

for every $t, s \in [0,1]$.

Organisation of the paper The rest of the paper is divided into four main parts and an appendix.

In Section 2, we first define the objects that we will be working with in the paper and set down some notation. More precisely, we define what is meant by “geometrically bounded by k”, “bounding enough disks”, and “being of Chekanov type”. We then move on to recall some prior results of the author which are central in proving the main results. Finally, we explain how these results cover the case of monotone Lagrangian submanifolds bounding enough disks and of monotone graphs, as these cases were not directly covered in the prior papers.

In Section 3, we move to prove (pre)compactness and local contractibility of the $\mathscr{L}_k^\star$ spaces, that is, Theorems A and C. These correspond, respectively, to Sections 3.1 and 3.2. We end this part with

Section 3.3, which aims to understand local Hausdorff-geodesics in $\mathscr{L}_k^\star$ — this will be an essential step in proving our local Hofer/spectral estimate above. This also allows us to consider some variations of the Hausdorff metric on $\mathscr{L}_k^\star$ and to conclude that they also induce the same topology on these spaces.

In Section 4, we finally prove symplectic properties of the $\mathscr{L}_k^\star$ spaces. More precisely, we prove Theorem B and all corollaries following the first principle appearing in the introduction.

In Section 5, we study the limit space $\mathscr{L}_\infty^\star$ and prove the corollaries following the second principle in the introduction. We conclude that section with an analysis of an alternate limit space: $\varinjlim \mathscr{L}_k^\star$, the inductive limit of the $\mathscr{L}_k^\star$. This highlights the shortcomings of trying to study $\mathscr{L}_\infty^\star$ through the $\mathscr{L}_k^\star$ spaces.

Finally, we conclude the paper with an appendix which includes all the results in Riemannian geometry that we will need throughout the paper. As these results do not seem to have appeared in the literature before — and Lemma A.5 even appeared as a conjecture — we believe that they could be of independent interest. This is also why they have been compiled as an independent appendix. More precisely, Section A.1 covers the new results on the Sasaki metric on TN, Section A.2 those on Riemann surfaces, and Section A.3 those on comparison of Riemannian invariants of Lagrangian submanifolds.

2 Preliminaries

We now lay down the foundations upon which the rest of the paper will be built. More precisely, in Section 2.1, we give the main definitions that will be used throughout the paper. In Section 2.2, we enunciate previous results from the author that will be essential for the rest of the paper. We hope that this will improve the readability of this paper, as the notation here is slightly different than in [11; 12]. We close things with Section 2.3, where it is shown that both monotone Lagrangian submanifolds bounding enough disks and monotone graphs fit within the formalism of the prior results.

2.1 Definitions and notation

In this paper, we will sometimes allow our manifolds to be C^k for $k < \infty$, have boundary, or to be noncompact, but will always make it explicit when we do so. That is, when there are no mentions of it, the manifolds are assumed to be C^∞, connected, and closed.

Let (M, ω) be a symplectic manifold that is either closed or noncompact but convex at infinity. Let J be an ω-compatible almost complex structure. When M is noncompact, we assume that J is convex at infinity and that the associated Riemannian metric $g_J = \omega(\cdot\,, J\cdot)$ is complete. We also pick an exhaustion by compacts $W_1 \subsetneq W_2 \subsetneq \cdots$ of M. When M is compact, we take the convention $W_k = M$ for all k.

We now describe what was meant in the introduction by "geometrically bounded by k". For every $k \in \mathbb{N}$, we set

$$\mathscr{L}_k := \{\text{Lagrangians } L \subseteq \operatorname{Int}(W_k) \mid \|B_L\| < k,\ L \text{ is strictly } (k+1)^{-1}\text{-tame}\},$$

where $\|B_L\|$ denotes the maximum of the pointwise operator norm of the second fundamental form of L. Strict tameness is defined as follows.

Definition 2.1 Let (M, g) be a Riemannian manifold, and let L be a submanifold. Let $\varepsilon \in (0, 1]$. We say that L is *strictly ε-tame* if

$$\inf_{x \neq y} \frac{d_M(x, y)}{\min\{1, d_L(x, y)\}} > \varepsilon,$$

where d_M is the distance function on M induced by g, and d_L is the distance function on L induced by $g|_L$.

In general, this space is too big for our approach, and thus we will instead be studying the following subspaces. If M is exact with $\omega = d\lambda$, we can consider

$$\mathscr{L}_k^e := \{L \in \mathscr{L}_k \mid L \text{ is } \lambda\text{-exact}\}$$

and, if M is monotone,

$$\mathscr{L}_k^{m(\rho)} := \{L \in \mathscr{L}_k \mid L \text{ is } \rho\text{-monotone and bounds enough disks}\},$$

where $\rho \geq 0$. By ρ-monotone, we mean that $\omega = \rho\mu$ on $\pi_2(M, L)$, where μ is the Maslov index, and that the minimal Maslov number of L is at least 2. Finally, we introduce the class of Lagrangian submanifolds bounding enough disks as follows.

Definition 2.2 We say that a Lagrangian submanifold L of M *bounds enough disks* if the image of the composition

$$\pi_2(M, L) \xrightarrow{\partial} \pi_1(L) \xrightarrow{h} H_1(L; \mathbb{Z})$$

has finite cokernel. Here, h is the abelianisation homomorphism.

Remark 2.3 Using the fact that $H^1(L; \mathbb{R}) = \mathrm{Hom}(\pi_1(L), \mathbb{R}) = \mathrm{Hom}(H_1(L; \mathbb{Z})^{\mathrm{free}}, \mathbb{R})$, we see that L bounds enough disks if and only if $\partial^* : H^1(L; \mathbb{R}) \to \mathrm{Hom}(\pi_2(M, L), \mathbb{R})$ is injective. In particular, this condition is automatically satisfied if either $H^1(L; \mathbb{R}) = 0$ or $H^1(M; \mathbb{R}) = 0$.

Example Every contractible loop in an oriented closed surface bounds enough disks. On the other hand, graphs of closed 1-forms in T^*N bound enough disks if and only if $H_1(N; \mathbb{R}) = 0$.

Moreover, it will later be of interest to study graphs of Hamiltonian diffeomorphisms. Therefore, when the symplectic manifold is a product $(M \times M, -\omega \oplus \omega)$, where (M, ω) is closed and monotone, we define

$$\mathscr{L}_k^\Gamma := \{L \in \mathscr{L}_k(M \times M) \mid L = \operatorname{graph} \varphi,\ \varphi \in \mathrm{Ham}(M, \omega)\}.$$

If a result is applicable to $\mathscr{L}_k^e$, $\mathscr{L}_k^{m(\rho)}$, and $\mathscr{L}_k^\Gamma$, we will say that it is true for $\mathscr{L}_k^\star$. We will denote the space without Riemannian bounds by $\mathscr{L}_\infty^\star$, i.e., $\mathscr{L}_\infty^\star = \bigcup_k \mathscr{L}_k^\star$. We will also sometimes write $\mathscr{L}_k^\star(M)$ when we want to make the ambient manifold M more apparent.

Finally, we need to give a precise definition of the family of metrics in which d is allowed to be. Explicitly, we will ask that d be of Chekanov type and dominated by the Lagrangian Hofer metric. The former notion was defined in [12], but we recall here the definition. We however refer to [12] for the proof that the metrics enunciated in the introduction are indeed of Chekanov type.

Definition 2.4 Let $\mathscr{F} \subseteq \mathscr{L}^\star(M)$. We say that a pseudometric $d^{\mathscr{F}}$ on $\mathscr{L}^\star(M)$ is of *Chekanov type* if for all compatible almost complex structure J, all $\delta > 0$, and all $L, L' \in \mathscr{L}^\star(M)$, there exist a finite — possibly empty — subset $\{F_1, \dots, F_k\} \subseteq \mathscr{F}$ with the following property.

For any C^0- and Hofer-small Hamiltonian perturbations $\widetilde{L}, \widetilde{L}', \widetilde{F}_1, \dots, \widetilde{F}_k$ of the Lagrangian submanifolds above making them pairwise transverse, and for any $x \in \widetilde{L} \cup \widetilde{L}'$, there exists a nonconstant J-holomorphic polygon $u : S_r \to M$ that

1. has boundary along $\widetilde{L}$, $\widetilde{L}'$ and $\widetilde{F}_1, \dots, \widetilde{F}_k$;
2. passes through x;
3. respects the bound
$$\omega(u) \le d^{\mathscr{F}}(L, L') + \delta.$$

Let $\mathscr{F}' \subseteq \mathscr{L}^\star(M)$ be such that
$$\Bigl(\overline{\bigcup_{F \in \mathscr{F}} F}\Bigr) \cap \Bigl(\overline{\bigcup_{F' \in \mathscr{F}'} F'}\Bigr)$$
is discrete. We will call $\hat{d}^{\mathscr{F},\mathscr{F}'} := \max\{d^{\mathscr{F}}, d^{\mathscr{F}'}\}$ a *Chekanov-type metric* if $d^{\mathscr{F}}$ and $d^{\mathscr{F}'}$ are both Chekanov-type pseudometrics.

For the rest of the paper, we set $d = \hat{d}^{\mathscr{F},\mathscr{F}'}$, a metric of Chekanov type associated with families $(\mathscr{F}, \mathscr{F}')$ and which is dominated by the Lagrangian Hofer metric d_H, i.e., there exists $C > 0$ such that $d \le C d_H$. In particular, this latter condition ensures that C^1-convergence of a sequence of Lagrangian submanifolds implies convergence in the metric d [33].

2.2 Useful results from previous work

We now recall some results from previous work for ease of reference later on. We will not write said results in full generality, but only in the setting which is required for this paper.

Below, we make use of the Hausdorff metric δ_H, defined between two compact subsets A and B of M as
$$\begin{aligned} \delta_H(A, B) &:= \max\{s(A; B), s(B; A)\}, \\ s(A; B) &:= \inf\{\varepsilon > 0 \mid A \subseteq B_\varepsilon(B)\} = \max_{x \in A} d_M(x, B), \end{aligned}$$
where $B_\varepsilon(A) := \bigcup_{x \in A} B_\varepsilon(x)$, i.e., it is the ε-neighbourhood of A, and d_M is the distance function on M.

Theorem 2.5 [11] *There exist constants $C_1, R_1 > 0$ with the following property. For all L and L' in $\mathscr{L}_k^\star$ such that $d(L, L') < R_1$, we have that*
$$\delta_H(L, L') \le C_1 \sqrt{d(L, L')},$$
where δ_H denotes the classical Hausdorff distance.

Theorem 2.6 [12] *For all $k \in \mathbb{N}$ and all L in $\mathscr{L}_k$, there exist constants $C_2, R_2 > 0$ with the following property. Whenever $L' \in \mathscr{L}_k$ is such that $\delta_H(L, L') < R_2$, there exists a $C^{1,\alpha'}$-small closed 1-form σ of L such that $L' = \operatorname{graph}\sigma$ in a Weinstein neighbourhood of L. Furthermore, if σ is exact, then*

$$d(L, L') \leq C_2 \delta_H(L, L').$$

Moreover, if a sequence $\{L_i\} \subseteq \mathscr{L}_k$ has Hausdorff limit N, then N is an embedded $C^{1,\alpha}$-Lagrangian submanifold, and there exist diffeomorphisms $f_i : N \xrightarrow{\sim} L_i$ for i large such that $f_i \to \mathbb{1}$ in the $C^{1,\alpha'}$-topology, $0 < \alpha' < \alpha < 1$.

Remark 2.7 In [12], the limit N was possibly immersed, even in the exact case. However, this was because we were working with the weaker condition of having bounded volume, instead of being ε-tame. The latter condition ensures that Hausdorff limits are indeed embedded — this is essentially because of Shen's compacity result [40].

More precisely, that result ensures that N is the image of an immersion $f : L \looparrowright M$ that is the $C^{1,\alpha'}$ limit of $\{\iota_i \circ \varphi_i\}$, where $\iota_i : L_i \hookrightarrow M$ is the inclusion and $f_i : L \xrightarrow{\sim} L_i$ is a diffeomorphism. It also ensures that $(\iota_i \circ f_i)^* g_J$ converges to $f^* g_J$, which is of class $C^{1,\alpha}$. Therefore, if $x \neq y \in L$ are such that $f(x) = f(y)$, then $d_M\big(\iota_i(f_i(x)), \iota_i(f_i(y))\big)$ tends to 0. Note that $d_{L_i}\big(\iota_i(f_i(x)), \iota_i(f_i(y))\big)$ is equal to $d_L(x, y)$ in the metric $(\iota_i \circ f_i)^* g_J$, which stays bounded away from 0 since those metrics converge to $f^* g_J$. Indeed, suppose that these distances tend to 0, and let $c_i : [0, 1] \to L$ be a minimal $(\iota_i \circ f_i)^* g_J$-geodesic from x to y. Then

$$|\dot{c}_i|_{f^* g_J} \leq M |\dot{c}_i|_{(\iota_i \circ f_i)^* g_J} = M d_L^{(\iota_i \circ f_i)^* g_J}(x, y) \leq D$$

since $(\iota_i \circ f_i)^* g_J \to f^* g_J$ and $d_L^{(\iota_i \circ f_i)^* g_J}(x, y) \to 0$. Therefore, up to a subsequence, $\{c_i\}$ must converge in both $W^{1,1}$- and C^0-topologies to some path c from x to y. But then,

$$\varliminf_i d_L^{(\iota_i \circ f_i)^* g_J}(x, y) \geq \int_0^1 \varliminf_i |\dot{c}_i|_{(\iota_i \circ f_i)^* g_J} = \int_0^1 |\dot{c}|_{f^* g_J} \geq d_L^{f^* g_J}(x, y) > 0$$

by Fatou's lemma. Hence, there is a contradiction, and $d_{L_i}\big(\iota_i(f_i(x)), \iota_i(f_i(y))\big)$ must be uniformly bounded away from 0. But that is not possible if all the L_i are strictly ε-tame for some $\varepsilon > 0$, since $d_M\big(\iota_i(f_i(x)), \iota_i(f_i(y))\big) \to 0$. Therefore, f must be an embedding.

In the nonexact case, this is also why the additional condition that L' and L have the same first Betti number is no longer required here: if a sequence $\{L_i\}$ with Betti number $b_1(L_i) = b$ were to Hausdorff-converge to a Lagrangian submanifold L with $b_1(L) < b$, then the limit f above would be a nontrivial cover onto its image — see Proposition 1 of [12] — and thus not an embedding.

2.3 Applying Theorem 2.6

To study $\mathscr{L}_k^\star$, it will be important to know that the 1-form σ appearing in Theorem 2.6 is exact. When L and L' are exact and the Weinstein neighbourhood $\Psi : D_r^* L \to M$ is exact, i.e., $\Psi^* \lambda = \lambda_0 + dF$ for some $F : T^* L \to \mathbb{R}$, this is self-evident. We now extend this to monotone Lagrangian submanifolds bounding enough disks.

Lemma 2.8 *Let $L, L' \in \mathscr{L}_k^{m(\rho)}$. Suppose that $L' = \operatorname{graph}\sigma$ in a Weinstein neighbourhood Ψ of L. Then, σ is exact.*

Proof Suppose that σ is not exact. Then, there exists a loop $\gamma : S^1 \to L$ such that $\langle \sigma, \gamma \rangle \neq 0$. In particular, it must be that γ represents a nonzero class in $H_1(L;\mathbb{Z})^{\mathrm{free}}$. Since L bounds enough disks, there is thus a disk D in M whose boundary lies on L and is some iterate γ^m of γ.

Let $u : S^1 \times [0,1] \to T^*L$ be the cylinder given by $u(t,s) = \bigl(\gamma(mt), s\sigma(\gamma(mt))\bigr)$. Note that u has nonzero area,

$$\int_{S^1\times[0,1]} u^*\omega_0 = \int_{S^1} (\sigma \circ \gamma^m)^*\lambda_0 = \int_{S^1} (\gamma^m)^*\sigma = m\int_{S^1} \gamma^*\sigma \neq 0,$$

where we have made use of the fact that the zero-section is λ_0-exact and of the tautological property of λ_0, i.e., $\sigma^*\lambda_0 = \sigma$. Since Ψ preserves the symplectic form, $C = \Psi(u(S^1 \times [0,1]))$ has nonzero area in M.

Then, we have that

$$\omega(D) = \rho\mu(D),$$

but also

$$\omega(D) + \omega(C) = \omega(D \# C) = \rho\mu(D \# C) = \rho\mu(D),$$

which is obviously a contradiction. Therefore, σ must be exact. □

We now show that Theorem 2.6 also applies to the space of graphs $\mathscr{L}_k^{\Gamma} = \mathscr{L}_k^{\Gamma}(M \times M)$, even when that space is equipped with the metric

$$d_H'(\operatorname{graph}\varphi_1, \operatorname{graph}\varphi_2) := \|\varphi_1\varphi_2^{-1}\|_H,$$

where $\|\cdot\|_H$ is the Hofer norm on M. It is not clear at all that this is the case since Theorem 2.6 requires that $d \le d_H$, but we have here that $d = d_H' \ge d_H$. However, we have that C^1-close graphs are also d_H'-close — this is the key to applying Theorem 2.6.

Proposition 2.9 *For all L in $\mathscr{L}_k^{\Gamma}$, there exist constants $C_2, R_2 > 0$ with the following property. Whenever $L' \in \mathscr{L}_k^{\Gamma}$ is such that $\delta_H(L, L') < R_2$, there exists a $C^{2,\alpha'}$-small function f of L such that $L' = \operatorname{graph} df$ in a Weinstein neighbourhood of L.*

Furthermore, if a sequence of graphs $\{L_i\} \subseteq \mathscr{L}_k^{\Gamma}$ Hausdorff-converges to another graph L, then $L_i \to L$ in d_H'.

Proof Since $L = (\mathbb{1} \times \varphi)(\Delta)$ for some $\varphi \in \operatorname{Ham}(M)$ and $(\mathbb{1} \times \varphi)(\mathscr{L}_k^{\Gamma}) \subseteq \mathscr{L}_{k'}^{\Gamma}$ for some k', it suffices to prove the statement for $L = \Delta$ and R_2 such that $B_{R_2}(\Delta)$ is contained in a small enough Weinstein neighbourhood of Δ. The proof follows from the typical flux argument, but we give here the details.

Fix a Weinstein neighbourhood $\Psi : D_R^*M \to M \times M$ of the diagonal, and let $\{L_i = \operatorname{graph}\varphi_i\} \subseteq \mathscr{L}_k^{\Gamma}$ be such that $L_i \subseteq \Psi(D_{r_i}^* L)$, where $\{r_i\} \subseteq (0, R]$ is a decreasing sequence converging to 0. Then, $L_i \to \Delta$ in the Hausdorff metric (see Lemma 1 of [12]), and there are 1-forms σ_i on M such that $L_i = \Psi(\operatorname{graph}\sigma_i)$ for i large by Theorem 2.6.

In fact, the theorem implies that $\{\sigma_i\}$ must C^1-converge to 0. Therefore, for i large, $\Psi(\operatorname{graph} t\sigma_i)$ is a graph for all $t \in [0,1]$, and we can define $\{\psi_i^t\}_{t\in[0,1]} \subseteq \operatorname{Symp}(M)$ via $\operatorname{graph}\psi_i^t = \Psi(\operatorname{graph} t\sigma_i)$. A direct computation gives $\operatorname{Flux}(\{\psi_i^t\}) = [\sigma_i]$. But fix a Hamiltonian isotopy $\{\varphi_i^t\}$ with $\varphi_i^1 = \varphi_i$, then the concatenation $\Psi_i = \psi_i \mathbin{\#} \overline{\varphi_i}$ of $\{\psi_i^t\}$ with $\{\varphi_i^{1-t}\}$ is a loop in $\operatorname{Symp}(M)$. Therefore,

$$\operatorname{Flux}(\Psi_i) = \operatorname{Flux}(\{\psi_i^t\}) - \operatorname{Flux}(\{\varphi_i^t\}) = [\sigma_i]$$

is in the flux group Γ_ω of M. By the flux conjecture (proved by Ono [34]), Γ_ω is discrete. Since $\sigma_i \to 0$, we must thus have that $[\sigma_i] = 0$, i.e., $\sigma_i = df_i$, for i large.

To get the neighbourhood, just note that if such a neighbourhood did not exist, we could construct a sequence $\{L_i\}$ converging to the diagonal, but such that every σ_i is nonexact, which would be a contradiction.

The last statement follows directly from what we have already said: $L_i = \operatorname{graph}\varphi_i \to L = \operatorname{graph}\varphi$ in the Hausdorff metric, then it must be that $\varphi_i \to \varphi$ in the C^1-topology, and thus also in the Hofer norm. □

3 Topological and metric properties of the geometrically bounded spaces

We now prove some topological and metric properties of $\mathscr{L}_k^\star$ and their various metric completions. More precisely, in Section 3.1, we prove Theorem A and some of its direct consequences. In Section 3.2, we then prove Theorem C. We also use some comparison results to get a weaker version of that theorem which holds in any Riemannian metric. We conclude in Section 3.3 with an analysis of local Hausdorff geodesics. We also use this to conclude that many natural variations of the Hausdorff on $\mathscr{L}_k^\star$ are in fact equivalent on those spaces.

3.1 Compactness and metric completions

We now move to prove results about the compactness of the various metric completions of $\mathscr{L}_k^\star$.

It follows directly from Theorems 2.5 and 2.6 that every $L \in \mathscr{L}_k^\star$ has a neighbourhood where both metrics are equivalent. We thus get directly the following result.

Corollary 3.1 *The topology on $\mathscr{L}_k^\star$ induced by δ_H is equivalent to the one induced by d.*

The new observation is that we can actually extend this equivalence to the completions.

Proposition 3.2 *The metric completions of $\mathscr{L}_k^\star$ in δ_H and d are homeomorphic.*

Proof Recall that the metric completion $\widehat{\mathscr{L}}_k^\star$ of $\mathscr{L}_k^\star$ in d is defined as the space of Cauchy sequences in $\mathscr{L}_k^\star$ up to equivalence. Two Cauchy sequences $\{L_i\}$ and $\{L_j'\}$ are called equivalent if for all $\varepsilon > 0$ there exist $I \in \mathbb{N}$ such that $d(L_i, L_j') < \varepsilon$ for all $i, j \geq I$. For our result, it thus suffices to show that d and δ_H have the same Cauchy sequences and the same notion of equivalence between them.

By Theorem 2.5, d-Cauchy sequences are also δ_H-Cauchy sequences. Likewise, when two d-Cauchy sequences are d-equivalent, they are also δ_H-equivalent.

Suppose that $\{L_i\}$ is a δ_H-Cauchy sequence. Denote by N its Hausdorff limit, and fix $\varepsilon > 0$. By Theorem 2.6, we know that N is actually an embedded Lagrangian $C^{1,\alpha}$-submanifold. Therefore, we can take a sequence $\{L_i'\}$ such that

(1) $d_H(L_i, L_i') \leq \varepsilon$;

(2) $\{L_i'\}$ Hausdorff-converges to a smooth submanifold N';

(3) $\delta_H(N, N') \leq \varepsilon$.

For example, this can be done by taking a sequence of generic Hamiltonians $\{H_i\}$ with $\|H_i\| \leq \varepsilon$ which is also C^1-Cauchy and then taking $L_i' = \varphi^{H_i}(L_i)$. If $\star = \Gamma$, we also can suppose that N' is a graph.

But since N' is smooth, $L_i' \xrightarrow{\delta_H} N'$ implies that $L_i' \xrightarrow{d} N'$ by Theorem 2.6 — or by Proposition 2.9 when working with graphs and $d = d_H'$. Therefore, $\{L_i'\}$ is also d-Cauchy, and we get that

$$d(L_i, L_j) \leq d_H(L_i, L_i') + d(L_i', L_j') + d_H(L_j', L_j) \leq 3\varepsilon$$

for i and j large. The sequence $\{L_i\}$ is thus itself d-Cauchy. The proof that both metrics have the same notion of equivalence between Cauchy sequences is analogous. □

Noting that the space of closed subsets of the compact W_k is compact in the Hausdorff metric (see, for example, [24]) and that the completion of $\mathscr{L}_k^\star$ in the Hausdorff metric can be identified with its closure in this space, we get directly the following corollary.

Corollary 3.3 *The metric completion $\widehat{\mathscr{L}}_k^\star$ of $\mathscr{L}_k^\star$ in d is compact.*

This thus completes the proof of Theorem A. Note that the compactness result also implies that the uncompleted space $\mathscr{L}_k^\star$ is precompact in $\widehat{\mathscr{L}}_k^\star$, so that we get from the generalised Heine–Borel theorem the following.

Corollary 3.4 *The space $\mathscr{L}_k^\star$ is totally bounded in d, i.e., for every $\varepsilon > 0$, every cover of $\mathscr{L}_k^\star$ by d-balls of radius ε admits a finite subcover.*

As we shall see below, this is the statement that will be useful for the corollaries appearing in the introduction.

3.2 Local contractibility

We finally show some local path-connectedness properties for $\mathscr{L}_k^\star$ and $\widehat{\mathscr{L}}_k^\star$ with some additional hypotheses on M. In fact, in each case, we end up proving something stronger: the spaces are locally contractible.

Note that this subsection makes heavy use of Riemannian-geometric results, which we have decided to keep for a dedicated appendix at the end of this paper. This thus makes this part much more readable if one is willing to accept those technical results.

The Sasaki metric case We first show that given L, the metric may be chosen so that we have a system of contractible neighbourhoods, i.e., we prove the part of Theorem C that is not specific to dimension 2.

Proposition 3.5 *Suppose that $g = g_J$ is such that, in a Weinstein neighbourhood of $L \in \mathscr{L}_k^\star$, it corresponds to the Sasaki metric of $g|_L$. Then, L possesses a system of contractible neighbourhoods in $\mathscr{L}_k^\star$.*

Proof Given Theorem 2.6, we can identify a sufficiently small neighbourhood of L to a set of $C^{1,\alpha'}$-small exact forms. Furthermore, as the neighbourhood gets smaller, the $C^{1,\alpha'}$-norm of the forms tends to 0 — this is a consequence of the second part of Theorem 2.6.

Therefore, it suffices to prove that such a set of forms must be star-shaped about the origin. The result thus follows from Lemmata A.1, A.3, and A.7. For convenience of computation, these lemmata are stated for vector fields on L instead of forms, but this is equivalent given the musical isomorphisms of $g|_L$. □

Note that Proposition 3.5 extends to the metric completion of $\mathscr{L}_k^\star$.

Corollary 3.6 *Let L and g be as in Proposition 3.5. The Lagrangian L also possesses a system of contractible neighbourhoods in the metric completion $\widehat{\mathscr{L}}_k^\star$ in d.*

Proof By the proof of Proposition 3.5, we know that if $L' = \operatorname{graph} dH$ is smooth and in a small enough neighbourhood, then $tL' = \operatorname{graph} t\,dH$ stays in that neighbourhood for all $t \in [0, 1]$.

Suppose now that $L' \in \widehat{\mathscr{L}}_k^\star$ is not smooth but in the (completion of the) same neighbourhood. By definition, there is thus a sequence $\{L_i\} \subseteq \mathscr{L}_k^\star$ in that neighbourhood such that $L_i \to L$ in the Hausdorff metric. But then, tL_i stays in it for all $t \in [0, 1]$. By continuity of multiplication by a scalar, we have that $tL_i \to tL'$ in the Hausdorff metric. Therefore, tL' stays in the same neighbourhood as L', which gives the result. □

The two-dimensional case Even though we expect any L in $\mathscr{L}_k^\star$ to have a system of contractible neighbourhoods for any metric g, the computations involved quickly become too complex to handle. An exception to this is when $\dim M = 2$. Namely, we can prove the following, which corresponds to the second part of Theorem C.

Proposition 3.7 *Let $\dim M = 2$ and $k \in \mathbb{N}$. Every Lagrangian $L \in \mathscr{L}_k^\star$ admits a system of contractible neighbourhoods.*

To do so, we employ a similar approach to the Sasaki case. We can do this because, in dimension 2, a tubular neighbourhood of L admits some fairly nice coordinates given by

$$\varphi\colon (0,\ell) \times (-r,r) \to M, \quad (s,t) \mapsto \exp_{\gamma(s)}(tJ\dot{\gamma}(s)),$$

where $\gamma : [0,\ell) \to L$ is a parametrisation such that $|\dot{\gamma}| \equiv 1$. Note that

$$\begin{aligned}
\varphi^* g &= |W|^2 ds^2 + dt^2, \\
\varphi^* J &= |W| \frac{\partial}{\partial t} \otimes ds - \frac{1}{|W|} \frac{\partial}{\partial s} \otimes dt, \\
\varphi^* \omega &= |W| ds \wedge dt = d\left(\left(-\int_0^t |W|\, d\tau\right) ds\right),
\end{aligned}$$

where $W(s,t)$ is the value at time t of the unique Jacobi field along the geodesic $t \mapsto \varphi(s,t)$ such that $W(s,0) = \dot{\gamma}(s)$ and $\dot{W}(s,0) = -\kappa(s)\dot{\gamma}(s)$. Here, κ is the (signed) geodesic curvature of L, which is defined via the relation $\ddot{\gamma} := \nabla_{\dot{\gamma}}\dot{\gamma} = \kappa J\dot{\gamma}$.

In these coordinates, a graph $L' = \{(s, \xi(s)) \mid s \in [0,\ell)\} =: \operatorname{graph}\xi$ is in the same Hamiltonian isotopy class in $S^1 \times (-r,r)$ as $L = \operatorname{graph} 0$ if and only if $s \mapsto \int_0^{\xi(s)} |W|\,dt$ admits a primitive, which in turn is equivalent to

$$\int_0^\ell \int_0^{\xi(s)} |W|\,dt\,ds = 0. \tag{1}$$

In particular, even if $\operatorname{graph}\xi$ is exact, one should not expect $\operatorname{graph}\alpha\xi$ to be for $\alpha \in (0,1)$. We can however circumvent this problem by slightly adjusting our approach.

Lemma 3.8 *Suppose that* $\|\xi\| := \max|\xi| < \frac{r}{2}$. *Then, for any* $\alpha \in [0,1]$, *there exists a unique real number* $c(\alpha)$ *in* $\left(-\frac{r}{2}, \frac{r}{2}\right)$ *such that* $\xi_\alpha := \alpha\xi + c(\alpha)$ *defines an exact graph. Furthermore,* c *depends continuously on* α *and* $|c(\alpha)| \le \alpha\|\xi\|$. *In particular, the path* $\alpha \mapsto \operatorname{graph}\xi_\alpha$ *is Hausdorff-continuous and stays in the image of* φ.

Moreover, if $\|\xi\| < \frac{r}{3}$, *then*

$$|c(\alpha) - c(\alpha')| \le \|\xi\|\,|\alpha - \alpha'|.$$

Proof Fix $\alpha \in [0,1]$. For each $s \in [0,\ell)$, the map $\tau \mapsto \int_0^{\alpha\xi(s)+\tau} |W|\,dt$ is increasing on $\left(-\frac{r}{2}, \frac{r}{2}\right)$. Indeed, $|W| > 0$ for all $(s,t) \in [0,\ell) \times (-r,r)$ since φ is a chart, and $|\alpha\xi(s) + \tau| \le \|\xi\| + \frac{r}{2} < r$. Furthermore, the map is positive if $\tau > \alpha\|\xi\|$ and negative if $\tau < -\alpha\|\xi\|$ for the same reason. Therefore, the same holds for the function $\tau \mapsto \int_0^\ell \int_0^{\alpha\xi(s)+\tau} |W|\,dt\,ds$. In particular, there is a unique solution $\tau = c(\alpha)$ in $\left(-\frac{r}{2}, \frac{r}{2}\right)$ to (1), that is such that

$$\int_0^\ell \int_0^{\alpha\xi(s)+c(\alpha)} |W|\,dt\,ds = 0.$$

The continuity of c follows directly from the fact that $\int_0^\ell \int_0^{\alpha\xi(s)+\tau} |W|\,dt\,ds$ depends continuously on α. The estimate on the value of c follows from the fact that this integral is positive if $\tau > \alpha\|\xi\|$ and negative if $\tau < -\alpha\|\xi\|$.

The final estimate follows from applying the same logic as above to the function

$$\tau \mapsto \int_0^\ell \int_0^{(\alpha-\alpha')\xi(s)+\tau-c(\alpha')} |W|\,dt$$

and noting that $\tau = c(\alpha)$ must be its unique zero in $\left(-\frac{r}{3}, \frac{r}{3}\right)$. □

Therefore, we are precisely in the setting of Section A.2, and Proposition 3.7 follows directly from applying Lemmata A.9 and A.11 to the path $\alpha \mapsto \xi_\alpha$.

The general case We now partially extend the local path-connectedness result to other metrics than the Sasaki ones. The results are of course weaker in this context, but they still point in the same direction as the locally Sasaki case. More precisely, we get the following.

Proposition 3.9 *For every $k > 0$ and $L \in \mathscr{L}_k^\star$, there are $a \geq 1$ and $b \geq 0$ with the following property. The Lagrangian submanifold L possesses a system of neighbourhoods U in $\mathscr{L}_k^\star$ such that the inclusion $U \hookrightarrow \mathscr{L}_{ak+b}^\star$ is nullhomotopic.*

This follows directly from combining the Sasaki case (Proposition 3.5) with the comparison results Lemmata A.12 and A.13 below.

3.3 Local geodesics

We now turn our attention to the geodesics in the Hausdorff metric in a neighbourhood of a Lagrangian submanifold $L \in \mathscr{L}_k^\star$.

We first recall the definition of a geodesic in a general metric space.

Definition 3.10 Let (X, d) be a metric space. The *length* of curve $c : [a, b] \to X$ is given by

$$\ell(c) := \sup_{a=t_0<\cdots<t_\ell=b} \sum_{i=1}^{\ell} d(c(t_{i-1}), c(t_i)) \in [0, \infty].$$

We say that c has *constant speed* if there exist $\lambda \geq 0$ such that $\ell(c|_{[t,s]}) = \lambda|t-s|$ for all $a \leq t \leq s \leq b$. In that case, we call λ its *speed*.

A *geodesic* is a curve $c : [a, b] \to X$ which has constant speed and is locally minimising in d, i.e., for every $t_0 \in [a, b]$, there is some $\varepsilon > 0$ such that if $t \leq s \in [a, b] \cap (t_0 - \varepsilon, t_0 + \varepsilon)$, then

$$\ell(c|_{[t,s]}) = d(c(t), c(s)).$$

If the above equality holds for all $t, s \in [a, b]$, then we call c a *minimising geodesic*.

We begin by describing certain geodesics in the Hausdorff distance in a small enough neighbourhood of any submanifold.

Proposition 3.11 *Let N be a submanifold of a complete Riemannian manifold M with tubular neighbourhood U, i.e., there is a neighbourhood V of the zero-section of $TN^\perp$ such that the exponential gives a diffeomorphism $V \xrightarrow{\sim} U$. Suppose that $V = B_\varepsilon(N)$ for some $\varepsilon > 0$, that is, V is the ε-neighbourhood of the zero-section in $TN^\perp$. If $N' \subseteq U$ is a submanifold such that $N' = \exp\sigma(N)$ for some section σ of $TN^\perp$, then*

$$\delta_H(tN', sN') = |t-s| \max|\sigma|$$

for all $t, s \in [0, 1]$, where $tN' := \exp t\sigma(N)$.

We get directly from this a characterisation of the radial Hausdorff-geodesics.

Corollary 3.12 *If N and N' are as above, the path defined by $c(t) = tN'$ is a minimal geodesic in the space of submanifolds of M equipped with the Hausdorff metric. In particular, if the Riemannian metric corresponds to the Sasaki metric on a Weinstein neighbourhood of $L \in \mathscr{L}_k^\star$, then L possesses a system of geodesically star-shaped neighbourhoods in $(\mathscr{L}_k^\star, \delta_H)$.*

Proof Take $0 \le a < b \le 1$. Then, the length of $c|_{[a,b]}$ is given by

$$\begin{aligned} \ell(c|_{[a,b]}) &= \sup_{a=t_0<\dots<t_\ell=b} \sum_{i=1}^{\ell} \delta_H(c(t_{i-1}), c(t_i)) \\ &= \sup_{a=t_0<\dots<t_\ell=b} \sum_{i=1}^{\ell} (t_i - t_{i-1}) \max|\sigma| \\ &= (b-a)\max|\sigma| \\ &= \delta_H(c(a), c(b)), \end{aligned}$$

which proves the first part of the result.

The statement on geodesically star-shaped then follows directly, knowing that tL' stays in $\mathscr{L}_k^\star$ by Proposition 3.5. □

Proof of Proposition 3.11 We first note that, for every $t, s \in [0,1]$,

$$\begin{aligned} s(tN'; sN') &= \max_{x\in N} d_M(\exp t\sigma(x), sN') \\ &= \max_{x\in N} \inf_{y\in sN'} d_M(\exp t\sigma(x), y) \\ &\le \max_{x\in N} d_M(\exp t\sigma(x), \exp s\sigma(x)) = |t-s|\max|\sigma|, \end{aligned} \tag{2}$$

since the exponential on V is a radial isometry. By exchanging the role of s and t above, we get that $\delta_H(tN', sN') = \max\{s(tN'; sN'), s(sN'; tN')\} \le |t-s|\max|\sigma|$.

Suppose that $s < t$, and let $x_0 \in N$ be such that $|\sigma(x_0)| = \max|\sigma|$. Suppose that there exists $y \in N$ such that $d_M(\exp t\sigma(x_0), \exp s\sigma(y)) < d_M(\exp t\sigma(x_0), \exp s\sigma(x_0)) = (t-s)|\sigma(x_0)|$. Then,

$$\begin{aligned} t|\sigma(x_0)| &= (t-s)|\sigma(x_0)| + s|\sigma(x_0)| \\ &\ge d_M(\exp t\sigma(x_0), \exp s\sigma(x_0)) + s|\sigma(y)| \\ &> d_M(\exp t\sigma(x_0), \exp s\sigma(y)) + d_M(y, \exp s\sigma(y)) \\ &\ge d_M(\exp t\sigma(x_0), y). \end{aligned}$$

This means that

$$d_M(\exp t\sigma(x_0), N) \le d_M(\exp t\sigma(x_0), y) < t|\sigma(x_0)|.$$

Let $\gamma : [0,1] \to M$ be a minimal geodesic from N to $\exp t\sigma(x_0)$. Since N is closed, $\gamma'(0) \in TN^\perp$, so that $\gamma(t) = \exp(t\gamma'(0))$. But then, $|\gamma'(0)| = d_M(\exp t\sigma(x_0), N) < t|\sigma(x_0)| < \varepsilon$. Therefore, $\gamma'(0)$ and $t\sigma(x_0)$ are two vectors in V whose image under the exponential map is $\exp t\sigma(x_0)$, which is a contradiction with the hypothesis that $\exp|_V$ be a diffeomorphism onto its image. The inequality (2) is thus in fact an equality, which proves the lemma. □

In the two-dimensional case, things are not as straightforward. Indeed, it is easy to see that the above proof gives that

$$\delta_H(\operatorname{graph}\xi_\alpha, \operatorname{graph}\xi_{\alpha'}) = \max|\xi_\alpha - \xi_{\alpha'}|,$$

which means that we should not expect $\alpha \mapsto \xi_\alpha$ to be a Hausdorff-geodesic in general. However, the above equality together with the estimate on the Lipschitz constant of $c(\alpha)$ in Lemma 3.8 gives the following.

Lemma 3.13 *Suppose that* $\dim M = 2$*. Every* $L \in \mathscr{L}_k^\star$ *has a neighbourhood* U *in* $\mathscr{L}_k^\star$ *such that*

$$\delta_H(\operatorname{graph}\xi_\alpha, \operatorname{graph}\xi_{\alpha'}) \le 2|\alpha-\alpha'| \max|\xi| = 2|\alpha-\alpha'|\delta_H(L, \operatorname{graph}\xi),$$

whenever $\operatorname{graph}\xi \in U$*.*

Variations on the Hausdorff metric Following a result of Sosov [42], the Hausdorff metric between L and L' is given by the infimum over all δ_H-continuous paths of closed subsets from L to L'. In fact, this infimum is even realised by a geodesic. This is because we have chosen the Riemannian metric on M so that (M, d_M) is a complete, geodesic metric space. Therefore, his definition of a geodesic corresponds to ours.

In this context, it is thus natural to consider what happens when we take the infimum over paths in a smaller set. More precisely, we are interested in the two following variants of the usual Hausdorff metric on $\mathscr{L}_k^\star$:

$$\begin{aligned}\delta_H^{\mathrm{Man}}(L_0, L_1) &:= \inf\{\ell(c) \mid c(i) = L_i,\ c(t) \text{ is an } n\text{-dimensional manifold } \forall t \in [0,1]\},\\ \delta_H^{(\star,k)}(L_0, L_1) &:= \inf\{\ell(c) \mid c(i) = L_i,\ c(t) \in \mathscr{L}_k^\star\ \forall t \in [0,1]\}.\end{aligned}$$

Here, all c's are δ_H-continuous, and all manifolds are smooth, closed, and connected. Note that

$$\delta_H \le \delta_H^{\mathrm{Man}} \le \delta_H^{(\star,k)}.$$

The first part of this subsection shows that, at least locally, these inequalities are equalities in good cases.

Proposition 3.14 *Every* $L \in \mathscr{L}_k^\star$ *has a neighbourhood* U *in* $\mathscr{L}_k^\star$ *such that for all* $L' \in U$*, the following holds.*

(i) $\delta_H^{\mathrm{Man}}(L, L') = \delta_H(L, L')$.

(ii) *If the Riemannian metric of* M *corresponds to the Sasaki metric on a Weinstein neighbourhood of* L*, then* $\delta_H^{(\star,k)}(L, L') = \delta_H(L, L')$*.*

(iii) *If* $\dim M = 2$*, then* $\delta_H^{(\star,k)}(L, L') \le 2\delta_H(L, L')$*.*

Proof We take U to be a tubular neighbourhood of L. By making U smaller if necessary, we may suppose that all $L' \in \mathscr{L}_k^\star$ such that $L' \subseteq U$ are graphs by Theorem 2.6. Therefore, (i) and (ii) follow directly from Corollary 3.12. Likewise, (iii) follows from Lemma 3.13. □

In particular, we get the following characterisation of the topologies induced by the variations of the Hausdorff metric.

Corollary 3.15 *The metrics* δ_H^{Man} *and* δ_H *induce the same topology on* $\mathscr{L}_k^\star$*. If* $\dim M = 2$*, then the same holds for* $\delta_H^{(\star,k)}$ *and* δ_H*.*

4 Symplectic properties of the geometrically bounded spaces

We now show some symplectic properties of the Lagrangian submanifolds in $\mathscr{L}_k^\star$ which derive from the topological and metric properties proved above. More precisely, we first prove Theorem B in Section 4.1. The rest of the subsections are then dedicated to proving the many corollaries following the first principle appearing in the introduction.

4.1 Hamiltonian isotopy classes

We explain how the connected components of $\mathscr{L}_k^\star$ are related to the isotopy classes of the Lagrangian submanifolds they contain.

Proposition 4.1 *For each $k > 0$, there exists $A > 0$ with the following property. If $L, L' \in \mathscr{L}_k^\star$ are not Hamiltonian isotopic, then*

$$d(L, L') \geq A.$$

Proof Let $\{L_i\}$ and $\{L_i'\}$ be sequences in $\mathscr{L}_k^\star$ such that $d(L_i, L_i')$ tends to zero, that is, they are equivalent in d. By Proposition 3.2, they must then be also equivalent in d_H. In particular, $d(L_i, L_i')$ is finite for i large, which implies that they are Hamiltonian isotopic for i large. Thus, such an $A > 0$ must exist. □

Finally, we can similarly get a fairly powerful result on the possible Hamiltonian isotopy classes in $\mathscr{L}_k^\star$.

Proposition 4.2 *There are finitely many Hamiltonian isotopy classes in $\mathscr{L}_k^\star$.*

Proof Suppose the contrary. Then, there exists a sequence $\{L_i\} \subseteq \mathscr{L}_k^\star$ with L_i not Hamiltonian isotopic to L_j if $i \neq j$. By Corollary 3.3, we may pass to a converging subsequence. But by Proposition 4.1 above, we will eventually get $d(L_i, L_j) < A$, so that L_i must be Hamiltonian isotopic to L_j for i and j large, and we have a contradiction. □

We close this section with a simple, but important observation: it is necessary to fix a Liouville form λ when $\star = e$ or a monotonicity constant ρ for $\star = m(\rho)$. Likewise, we truly need the "bounding enough disks" condition for our results to hold. Indeed, in each case when one of these conditions is broken, we get a counterexample to Proposition 4.2.

- On the flat cylinder T^*S^1, each parallel is a totally geodesic 1-tame Lagrangian submanifold which is exact for some primitive of the usual symplectic form. However, that primitive is different for each parallel, i.e., only one of them can belong to $\mathscr{L}^e(T^*S^1)$.

- These parallels can also be seen as monotone Lagrangian submanifold for any $\rho \geq 0$. However, they bound no disk at all, and thus do not respect the condition of bounding enough disks, i.e., they never belong to $\mathscr{L}^{m(\rho)}(T^*S^1)$.

- In $\mathbb{R}^2$ with its usual structure, the Hamiltonian isotopy class of a circle is determined by the area it encloses. Clearly, for any $k > 0$, there is a continuum of possible areas enclosed by a circle in $\mathscr{L}_k(\mathbb{R}^2)$. Furthermore, each of these circles are monotone. However, they are all so for different monotonicity constants, i.e., only one class can belong to any $\mathscr{L}^{m(\rho)}(\mathbb{R}^2)$.

Remark 4.3 In dimension 2, Proposition 4.2 follows directly from Proposition 4.1 and Corollary 3.3, since we then know that $\widehat{\mathscr{L}}_k^\star$ is locally path connected by Proposition 3.7. Indeed, this ensures that the connected components of $\widehat{\mathscr{L}}_k^\star$ are open, and they are thus in finite number, by compactness of the space. But by Proposition 4.1, each connected component is contained in a unique Hamiltonian class, so that the latter must also be in finite number.

4.2 Boundedness of *d*

We now prove the corollary on the boundedness of d when it is restricted to one Hamiltonian orbit.

Corollary 4.4 *For every $k \geq 1$, there is some $B > 0$ with the following property. Let $L, L' \in \mathscr{L}_k^\star$. Suppose that either L and L' are Hamiltonian isotopic or that $M = T^*N$, $\star = e$, and $d = \gamma$. Then,*

$$d(L, L') \leq B.$$

Proof If $L, L' \in \mathscr{L}_k^{L_0}$ for some $L_0 \in \mathscr{L}_\infty^\star$, then $d_H(L, L') < \infty$ by definition. Since d is dominated by d_H, we thus also have that $d(L, L') < \infty$. Therefore, total boundedness of $\mathscr{L}_k^{L_0}$ (as proved in Corollary 3.4) implies boundedness.

When $M = T^*N$, $\star = e$, and $d = \gamma$, the same argument works because we then have $d(L, L') < \infty$ for all $L, L' \in \mathscr{L}_k^\star$. □

Remark 4.5 The improvement here, compared to the version of the Viterbo conjecture appearing in [12], is that the bound on $\gamma(L)$ stands for all exact Lagrangian submanifolds in the unit codisk bundle, not just in a codisk bundle of small enough radius. However, the constant A now explicitly depends on k. We have however not simply rescaled the previous estimate: the present bound applies to Lagrangian submanifolds which are not graphs, which was not the case previously.

4.3 Graphs and Ostrover's example

In [35], Ostrover constructs, for every closed symplectic manifold M such that $\pi_2(M) = 0$ and any $c > 0$ small enough, a sequence of Hamiltonian diffeomorphisms $\{\varphi_i^c\} \subseteq \mathrm{Ham}(M)$ such that

(1) $\|\varphi_i^c\|_H \xrightarrow{i\to\infty} \infty$;

(2) $d_H(\Delta, \operatorname{graph} \varphi_i^c) \equiv c$,

where $\Delta \subseteq M \times M$ is the diagonal and d_H is the Lagrangian Hofer metric of $M \times M$. In particular, if we set $\varphi_i := \varphi_i^{1/i}$, we get a sequence of Hamiltonian diffeomorphisms which Hofer-converges to infinity, but whose graphs Lagrangian–Hofer-converges to the diagonal. In this subsection, we want to show that such a phenomenon is impossible in the world of geometrically bounded Lagrangian graphs. That is, we show the second corollary in the introduction.

Note that contrary to all other subsections of this paper, we *do not* require that our monotone Lagrangian submanifolds bound enough disks. We recall that

$$\mathscr{L}_k^\Gamma = \{L \in \mathscr{L}_k(M \times M) \mid L = \operatorname{graph} \varphi,\ \varphi \in \mathrm{Ham}(M)\},$$

where M is equipped with some Riemannian metric, and $M \times M$, with the resulting product metric.

From Proposition 2.9, Theorem 2.6 applies to $\mathscr{L}_k^\Gamma$ equipped with the metric d'_H induced by the Hofer norm, i.e., defined by

$$d'_H(\operatorname{graph}\varphi_1, \operatorname{graph}\varphi_2) := \|\varphi_1\varphi_2^{-1}\|_H,$$

where $\|\cdot\|_H$ is the Hofer norm of M. Since $d'_H \geq d_H$, Theorem 2.5 also trivially applies. In particular, we can make use of Corollary 3.4. We thus get the following, since $\mathscr{L}_k^\Gamma$ contains a unique Hamiltonian isotopy class.

Corollary 4.6 *On $\mathscr{L}_k^\Gamma$, d_H and d'_H induce the same topology, Furthermore, d'_H is bounded. In particular, an example* à la *Ostrover does not exist in $\mathscr{L}_k^\Gamma$.*

Note that we have a defined notion of C^1-distance between graphs and of C^1-bounds on them through the diffeomorphisms that define them. We suspect that the spaces resulting from these bounds also obey a result analogous to Corollary 4.6 above.

However, working with curvature bounds of the graphs allows the limit in the completion to be represented by Lagrangian submanifolds of $M \times M$ which are not graphs. In particular, we get the following.

Corollary 4.7 *There are elements in the metric completion of $(\mathrm{Ham}(M), \|\cdot\|_H)$ which can uniquely be represented by nongraphical $(C^{1,\alpha})$ Lagrangian submanifolds of $M \times M$.*

It would be quite interesting to be able to detect which elements of the completion have this property.

Remark 4.8 One could ask the same question as above but with the Hofer norm replaced by the spectral one. However, this is a trivial question: it is known [32] that in the monotone setting, the spectral norm of a Hamiltonian diffeomorphism in M is equal to the spectral distance of its graph to the diagonal in $M \times M$.

4.4 Order of a symplectomorphism and categorical entropy

We now move on to the first corollary of Theorem B. That result (or more precisely, Proposition 4.2) directly implies the following.

Corollary 4.9 *Let $L \in \mathscr{L}_k^\star$, and let ψ be a symplectomorphism of M. If there exist $k \geq 1$ such that $\psi^\nu(L) \in \mathscr{L}_k^\star$ for all $\nu \geq 1$, then there exist N such that $\psi^N(L)$ is Hamiltonian isotopic to L.*

In fact, we can make the above statement somewhat quantitative through the various notions of entropy. First of all, we note that we have a criterion for the vanishing of barcode entropy — we refer the reader to [10; 16] for the definitions.

Corollary 4.10 *Let L and ψ be as in Corollary 4.9. If ψ is Hamiltonian and L' is another exact or monotone Lagrangian submanifold, then the relative barcode entropy $\hbar(\psi; L, L')$ vanishes. If L, L', and ψ are all exact, then the same holds for the slow relative barcode entropy $\hbar^{\mathrm{sl}}(\psi; L, L')$.*

Proof By Proposition 4 of [12], $\psi^\nu(L) \in \mathscr{L}_k^\star$ for all ν implies a universal bound on the volume of the $\psi^\nu(L)$. The result on the usual barcode entropy then follows directly from the proof of Theorem 2.4 of [10]. The result on slow barcode entropy follows instead from the proof of Theorem A of [16]. □

Perhaps more interestingly however, we gather from Corollary 4.9 a geometrical criterion for the vanishing of the so-called categorical entropy of a symplectomorphism, which we define below. We make use of the definition using multiple generators of [6], instead of the original definition using a single split generator [17], but it is shown in the former paper that these are equivalent. The reason for this is that we want to work with actual Lagrangian submanifolds, not abstract twisted complexes or modules over the Fukaya category.

We first introduce the following notation. Let $\mathscr{C}$ be a (nongraded) triangulated category. For a morphism $f : A \to B$ in $\mathscr{C}$, we denote by $\operatorname{Cone}(f)$ its cone, i.e.,the unique-up-to-isomorphism object turning $A \to B \to \operatorname{Cone}(f) \to A$ into a distinguished triangle of $\mathscr{C}$. More generally, we define by induction $\operatorname{Cone}(f_1, \dots, f_m)$ to be the cone of the map $f_m : A_m \to \operatorname{Cone}(f_1, \dots, f_{m-1})$.

Definition 4.11 Let $\mathscr{C}$ be a nongraded triangulated category, and let $A, G_1, \dots, G_\ell$ be objects of $\mathscr{C}$. The *complexity* of A with respect to $G_1, \dots G_\ell$ is given by

$$\delta(G_1, \dots, G_\ell; A) := \inf\{m \mid A \oplus A' = \operatorname{Cone}(f_1, \dots, f_m),\ A' \in \operatorname{Ob}(\mathscr{C}),\ \operatorname{dom} f_i \in \{G_1, \dots, G_\ell\}\}.$$

Furthermore, if $G_1, \dots G_\ell$ split-generate $\mathscr{C}$ and Φ is an endofunctor of $\mathscr{C}$, we define its *categorical entropy* to be

$$h_{\mathrm{cat}}(\Phi) := \lim_{\nu \to \infty} \frac{\delta(G_1, \dots, G_\ell; \Phi^\nu(G_1 \oplus \dots \oplus G_\ell))}{\nu} \in [0, +\infty].$$

In other words, complexity measures how many iterated cones are needed to get A from $G_1, \dots, G_\ell$ up to some splitting, whilst categorical entropy measures how much Φ "complexifies" the generators of $\mathscr{C}$. As the notation suggests, the definition of categorical entropy is independent of the choice of split-generators.

Remark 4.12 In the definition of complexity above, we could replace the $A \oplus A' = \operatorname{Cone}(f_1, \dots, f_m)$ condition by simply $A = \operatorname{Cone}(f_1, \dots, f_m)$ and then work with generators to define categorical entropy. This is perfectly valid but leads to a number which is — in general — larger than what we have defined here. Since we are interested in a criterion for the vanishing of entropy, it is a more general approach to work with split-generation.

In the symplectic context, we take $\mathscr{C}$ to be the derived Fukaya category $\operatorname{DFuk}^\star(M)$ generated by $\mathscr{L}^\star_\infty$ — this is well defined in both the exact [37] and monotone [41] settings. Then, any symplectomorphism ψ preserving $\mathscr{L}^\star_\infty$ will induce an endofunctor of $\operatorname{DFuk}^\star(M)$ — we will call the categorical entropy of that functor the categorical entropy of ψ. The following result follows directly from Corollary 4.9 since Hamiltonian isotopic Lagrangian submanifolds induce isomorphic objects in the derived Fukaya category.

Corollary 4.13 *Suppose that ψ is a symplectomorphism of M preserving $\mathscr{L}^\star_\infty$ such that, for a set of generator $L_1, \dots L_\ell$ of $\operatorname{DFuk}^\star(M)$, $\psi^\nu(L_i) \in \mathscr{L}^\star_k$ for some k, for every ν and every i. Then, $h_{\mathrm{cat}}(\psi) = 0$.*

In other words, if $h_{\mathrm{cat}}(\psi) > 0$, then there is some Lagrangian submanifold L which is a factor of a split-generator G of $\operatorname{DFuk}^\star(M)$ such that the sequence $\{\psi^\nu(L)\}$ is not contained in any $\mathscr{L}^\star_k$. Note that we may suppose that such a Lagrangian submanifold L induces a nontrivial object in $\mathrm{DFuk}^\star(M)$. Thus, ψ must then deform some Floer-theoretically essential Lagrangian submanifold.

Remark 4.14 There is currently work in progress from Ambrosioni, Biran, and Cornea which defines weighted versions of categorical entropy coming from the triangulated persistence structure of the derived Fukaya category [8], which associates to cones an associated weight. Because our spaces are all totally bounded, we expect that their notion of entropy is also well behaved in our setting, so that we can expect similar results as above.

Remark 4.15 Lemmata A.12 and A.13 below imply that, for any symplectomorphism ψ preserving $\mathscr{L}^\star_\infty$ and any $L \in \mathscr{L}^\star_\infty$, the quantity

$$\eta(\psi; L) := \limsup_{\nu\to\infty} \frac{\log^+\bigl(\underline{k}(\psi^\nu(L))\bigr)}{\nu} \in [0, +\infty],$$

where $\underline{k}(L) := \inf\{k \mid L \in \mathscr{L}^\star_k\}$ and $\log^+(x) := \max\{0, \log(x)\}$, is independent on the Riemannian metric g on M or on the choice of compacts W_k. However, the geometric meaning of this quantity is still unclear to us.

For example, it is well known that the quantity

$$\Gamma(\psi; L) := \limsup_{\nu\to\infty} \frac{\log^+\bigl(\mathrm{Vol}(\psi^\nu(L))\bigr)}{\nu} \in [0, +\infty]$$

is a lower bound to topological entropy [45] and an upper bound to barcode entropy with any $L' \in \mathscr{L}^\star_\infty$ [10]. Furthermore, we have shown (Proposition 4 of [12]) that being in $\mathscr{L}_k$ implies respecting a volume bound. However, that volume bound is generally not polynomial in k, so that there is no obvious link between η and Γ — and thus between η and entropy.

Even in the case when $n = 1$, in which case the volume bound reduces to

$$\mathrm{Vol}(B_{(k+1)^{-1}}(W_k)) \geq 2\lfloor \mathrm{Diam}(L)\rfloor \min\{\mathrm{Diam}(L), (k+1)^{-1}\}, \tag{3}$$

with $\mathrm{Vol}(L) = 2\,\mathrm{Diam}(L)$, this only implies that

$$\eta(\psi; L) \geq \Gamma(\psi; L) - \limsup_{\nu\to\infty} \frac{\log(\mathrm{Vol}(W_{\underline{k}(\psi^\nu(L))}))}{\nu}. \tag{4}$$

But the right-hand side vanishes. Indeed, we have the freedom to choose in (3) any W_k containing $\psi^\nu(L)$. In particular, let W_k be the "smallest" possible choice: the tubular neighbourhood of $\psi^\nu(L)$ of radius r for r small. Then, $\mathrm{Vol}(W_k)$ behaves like $\mathrm{Vol}(\psi^\nu(L))r$ — see, for example, Theorem 9.23 of [21] — and the superior limit is simply $\Gamma(\psi; L)$.

4.5 Connected components of $\widehat{\mathscr{L}}^\star_k$

We now move to the second corollary of Theorem B. In fact, we prove the following slightly stronger statement.

Corollary 4.16 *If L and L' are smooth Lagrangian submanifolds belonging to the same connected component of $\mathscr{L}_k^\star$ or of $\widehat{\mathscr{L}}_k^\star$, then they are Hamiltonian isotopic.*

In particular, L and L' are Hamiltonian isotopic if there is a d-continuous path in $\widehat{\mathscr{L}}_k^\star$ from L to L'.

Proof Note that the set $\mathscr{L}_k^{L_0} = (\mathrm{Ham}(M) \cdot L_0) \cdot \mathscr{L}_k^\star$ must be a clopen of $\mathscr{L}_k^\star$. Indeed, the fact that $d(L, L') \geq A > 0$ whenever $L' \in \mathscr{L}_k^\star$ is not Hamiltonian isotopic to L implies that $\mathscr{L}_k^{L_0}$ contains all of its limit points in $\mathscr{L}_k^\star$, i.e., $\mathscr{L}_k^{L_0}$ is closed. But that fact also implies that $\mathscr{L}_k^{L_0}$ is equal to the union of the metric balls of radius $\frac{A}{2}$ centred at points on $\mathscr{L}_k^L$, so that it must also be open. The conclusion then follows from the fact that a clopen always fully contains the connected components of its points.

For the last statement, simply note that the path-connected component of a point is always contained in its connected component. □

Remark 4.17 Every exact Lagrangian isotopy $\{L_t\}_{t\in[0,1]}$ respects the hypotheses of Corollary 4.16. However, these hypotheses are strictly more general. For example, it is proven in [26] that if $H^1(N;\mathbb{R})=0$, then the Floer barcode is C^0-continuous. That is, if $t \mapsto \varphi_t$ is a C^0-continuous path in the C^0-completion of the group of symplectomorphisms of T^*N and $L, L' \subseteq T^*N$ are exact, then the Floer barcode $\mathcal{B}(\varphi_t(L), L')$ depends continuously of t in the bottleneck distance. In particular, this means that such a C^0-continuous path $t \mapsto \varphi_t(L)$ through Lagrangian submanifolds of $\mathscr{L}_k^e$ respects the hypotheses of Corollary 4.16.

4.6 Hofer geodesics

We now finally move on to the corollary of Theorem C in the introduction.

We recall that it has been proven by Milinković [33] that the Hofer and spectral distances between two graphs in T^*L are both given by

$$d_H(\operatorname{graph} df, \operatorname{graph} dg) = \gamma(\operatorname{graph} df, \operatorname{graph} dg) = \max|f-g| - \min|f-g|. \tag{5}$$

In particular, if L' is a graph, then the path $t \mapsto tL'$ is a minimising geodesic from L to L' in the Hofer metric of T^*L.

However, when T^*L is embedded in a symplectic manifold M via a Weinstein neighbourhood Ψ of some Lagrangian submanifold L, (5) is reduced to a simple bound in the Hofer metric of M. In particular, we are no longer guaranteed that $t \mapsto \Psi(tL')$ is minimising when $L' \subseteq T^*L$ is a graph. Note that it is, however, still a geodesic by Theorem 2 of [25]. Indeed, if $L' = \operatorname{graph} df$, then $t \mapsto \Psi(tL')$ is generated by some extension of $\Psi_*(\pi^* f)$, whose extrema along $\Psi(tL')$ are attained at fixed points in $\bigcap_t \Psi(tL') = \Psi(\mathrm{Crit}(f))$. Therefore, the isotopy is always *locally* minimising.

In the case when L is either exact or monotone, we can use the results of Sections 3.2 and 3.3 to give a lower estimate on how far the path $t \mapsto \Psi(tL')$ is from being minimising. More precisely, we prove the following.

Corollary 4.18 *Let L be a Lagrangian submanifold of M that is either exact or monotone, and let g be a metric on M which corresponds to the Sasaki metric of L on a Weinstein neighbourhood $\Psi : D_r^* L \to M$.*

For every $k \geq 1$, there are constants $C > 0$ and $r' \in (0, r]$ with the following property. Whenever $f : L \to \mathbb{R}$ is such that $\Psi(\operatorname{graph} df) \in \mathscr{L}_k$ and $|df| \leq r'$, we have that

$$d_H(\Psi(t \operatorname{graph} df), \Psi(s \operatorname{graph} df)) \geq C(t-s)^2 \max|df|^2$$

for every $t, s \in [0, 1]$, where d_H is the Hofer distance in M.

Proof Note that $\Psi(t \operatorname{graph} df) \in \mathscr{L}_k$ for all $t \in [0, 1]$ by Proposition 3.5. Therefore, the bound follows directly from Theorem 2.5 by using Proposition 3.11 to compute $\delta_H(\Psi(t \operatorname{graph} df), \Psi(s \operatorname{graph} df))$. □

Remark 4.19 In [12], the precise C is computed in terms of k and the sectional curvature and injectivity radius of M. In fact, by choosing r' small enough, C can be made to only depend on the values of these invariants on $\Psi(D_r^* L)$. However, those values still depend on more than just k — except when L is flat. For example, the sectional curvature and the injectivity radius in the Sasaki metric are not uniformly bounded in T^*L when L is not flat (see [29]), so that C must depend heavily on r. On the other hand, we can replace the dependency of C on k for one depending on $\|d\varphi\|$, where φ is the Hamiltonian diffeomorphism generated by $-\pi^* f$ and $\pi : T^*L \to L$ is the natural projection (see [13]).

5 Properties in the limit

There is a natural question of whether the properties of the $\mathscr{L}_k^\star$'s survive in $\mathscr{L}_\infty^\star$. This is however not a simple matter to see which properties can be transported to the limit: the sequence $\{L_i\} \subseteq \mathscr{L}_\infty^e(T^*S^1)$ in Figure 2 exemplifies how the relation between δ_H and d is not as clear in $\mathscr{L}_\infty^\star$. Indeed, one can easily convince oneself that such a sequence Hofer-converges but does not Hausdorff-converge to L_0 (see [11] for a more detailed analysis of this example).

This sequence suggests that $\mathscr{L}_k^\star$ is a fairly pathological subspace of $\mathscr{L}_\infty^\star$, since it indicates that every open subset of $\mathscr{L}_\infty^\star$ intersect all $\mathscr{L}_k^\star$ with large k. In particular, $\{\mathscr{L}_k^\star\}_{k\geq 1}$ is far from being an exhaustion of $\mathscr{L}_\infty^\star$ by compact sets, which complicates things.

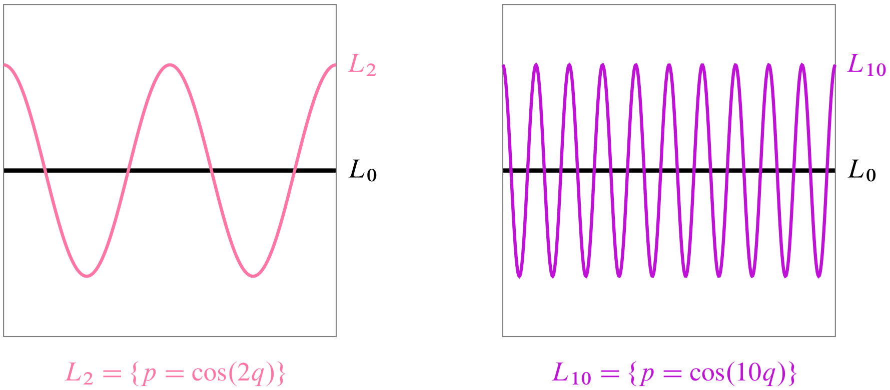


Figure 2: The sequence $\{L_i\}$ and its Hofer limit L_0.

Nonetheless, we make here an attempt to extract properties. More precisely, Sections 5.1 and 5.2 are dedicated to proving the two corollaries in the introduction which follow the second principle. That is, in Section 5.1, we show that $\mathscr{L}^\star_\infty$ is separable and, in Section 5.2, that it contains at most countably many Hamiltonian isotopy classes. We end this part with Section 5.3, which is a study of another possible limit space of the $\mathscr{L}^\star_k$'s. Even though that space is better suited to the $\mathscr{L}^\star_k$ spaces, we show that it basically has the C^2-topology, and thus captures mostly metric phenomena rather than symplectic ones.

5.1 Topological and metric properties

We investigate the implications of Section 3 to the space $\mathscr{L}^\star_\infty$ of all Lagrangian submanifolds in M respecting $\star$ when that space is equipped with the metric d. This corresponds to the second corollary of Theorem A in the introduction.

Proposition 5.1 *The metric space* $(\mathscr{L}^\star_\infty, d)$ *— and thus its completion — is separable.*

Proof We first note that every totally bounded metric space X is separable. Indeed, for every $m \geq 1$, we can cover X by a finite number of balls of radius $\frac{1}{m}$. Let $x_1, \dots, x_{N_m}$ be the centres of these balls. By construction,

$$\mathcal{B} := \bigcup_{m=1}^{\infty} \{x_1, \dots, x_{N_m}\}$$

is then a countable dense subset of X.

In particular, $\mathscr{L}^\star_k$ admits a countable dense subset $\mathcal{B}_k$ for all $k \geq 1$ by Corollary 3.4. Therefore, $\bigcup_k \mathcal{B}_k$ is the required countable dense subset of $\mathscr{L}^\star_\infty = \bigcup_k \mathscr{L}^\star_k$ □

Owing to the equivalence of many topological properties on metric spaces, we directly get the following.

Corollary 5.2 *The metric space* $(\mathscr{L}^\star_\infty, d)$ *and its completion are second countable, paracompact, and hereditarily Lindelöf, i.e., for every subspace* A, *an open cover of* A *admits a countable subcover.*

Remark 5.3 It was pointed out to us by Vincent Humilière and Egor Shelukhin that the Hamiltonian orbit of any Lagrangian L is always separable in the Lagrangian Hofer metric d_H. This follows from the fact that $C^\infty_c([0,1] \times M)$ is separable in the C^1-norm and that $d_H(\varphi^1_H(L), \varphi^1_G(L)) \leq \|\bar{H} \# G\|_H \leq C\|H - G\|_{C^1}$. Therefore, Proposition 5.1 is more of a statement on the behaviour of the Hamiltonian isotopy classes of $\mathscr{L}^\star_\infty$. We will explore them more in depth below.

5.2 Symplectic properties

We now explore the possible Hamiltonian isotopy classes of $\mathscr{L}^\star_\infty$. This corresponds to the third corollary of Theorem B in the introduction.

Proposition 5.4 *There are at most countably many Hamiltonian isotopy classes in* $\mathscr{L}^\star_\infty$.

Proof We construct a sequence which enumerates all Hamiltonian isotopy classes as follows. For $k = 1$, let $\{L_1, \dots, L_{N_1}\}$ be a collection of Lagrangian submanifolds such that L_i and L_j are not Hamiltonian isotopic if $i \neq j$ and such that any $L \in \mathscr{L}_1^\star$ is Hamiltonian isotopic to one of the L_i. By Proposition 4.2, such a $N_1 < \infty$ exists. Then, $\{L_1, \dots, L_{N_{k+1}}\}$ is built from $\{L_1, \dots, L_{N_k}\}$ by adding representatives of the Hamiltonian isotopy classes of $\mathscr{L}_{k+1}^\star - \mathscr{L}_k^\star$ is a similar fashion. Since every $L \in \mathscr{L}_\infty^\star$ must be contained in $\mathscr{L}_k^\star$ for some k, it is clear that $\{L_1, L_2, \dots\}$ is in bijection with the Hamiltonian isotopy classes in $\mathscr{L}_\infty^\star$. □

Since every Hamiltonian isotopy class is contained in a single path-connected component of $\mathscr{L}_\infty^\star$, we also get the following.

Corollary 5.5 *The space $(\mathscr{L}_\infty^\star, d)$ has at most countably many path-connected components.*

5.3 Another limit space

As noted above, we sadly lose many properties of the $\mathscr{L}_k^\star$ spaces when we go to the limit space $\mathscr{L}_\infty^\star$. This is ultimately because the $\mathscr{L}_k^\star$ spaces are quite pathological in $\mathscr{L}_\infty^\star$. There is however another natural topology on the set $\bigcup_k \mathscr{L}_k^\star$ which circumvents this issue: the limit $\varinjlim \mathscr{L}_k^\star$ of the inductive system $\mathscr{L}_1^\star \subseteq \mathscr{L}_2^\star \subseteq \cdots$. In other words, $\varinjlim \mathscr{L}_k^\star = \bigcup_k \mathscr{L}_k^\star$ as a set, and a subset $U \subseteq \bigcup_k \mathscr{L}_k^\star$ is open if and only if $U \cap \mathscr{L}_k^\star$ is open in $\mathscr{L}_k^\star$ for all k. In particular, this means that $\varinjlim \mathscr{L}_k^\star = \mathscr{L}_\infty^\star$ as sets, but the topology on the metric spaces $(\mathscr{L}_\infty^\star, d)$ and $(\mathscr{L}_\infty^\star, \delta_H)$ is coarser than that of $\varinjlim \mathscr{L}_k^\star$. Note that by Lemmata A.12 and A.13 below, the topology on $\varinjlim \mathscr{L}_k^\star$ is independent of the choice of Riemannian metric.

To exemplify how this topology is better behaved in some regards, we show that its connected components are much simpler than those of $\mathscr{L}_\infty^\star$.

Proposition 5.6 *The connected components and path-connected components of $\varinjlim \mathscr{L}_k^\star$ agree, and they are precisely the Hamiltonian isotopy classes.*

Proof First note that a given Hamiltonian isotopy class is always contained in a single path-connected component of $\varinjlim \mathscr{L}_k^\star$. To see this, suppose that $L, L' \in \varinjlim \mathscr{L}_k^\star$ are Hamiltonian isotopic, and take a Hamiltonian isotopy $\{\varphi_t\}$ such that $\varphi_1(L) = L'$. By smoothness of the isotopy, the path $c(t) = \varphi_t(L)$ is fully contained in $\mathscr{L}_k^\star$ for some k and is Hausdorff-continuous. Therefore, L and L' are in the same path-connected component of $\mathscr{L}_k^\star$, and thus of $\varinjlim \mathscr{L}_k^\star$.

On the other hand, given $L \in \varinjlim L_k$, its connected component in $\varinjlim L_k$ must contain the Hamiltonian isotopy class $\mathscr{L}_\infty^L$ of L. Indeed, in Corollary 4.16, we have shown that $\mathscr{L}_k^L = \mathscr{L}_\infty^L \cap \mathscr{L}_k^\star$ is clopen in $\mathscr{L}_k^\star$ for all k, so that $\mathscr{L}_\infty^L$ must also be clopen in $\varinjlim L_k$. But a clopen must contain the connected component of its elements, which proves the inclusion.

Since each path-connected component is contained in a single connected component, this proves the result. □

We now compare the limit topology with other ones to better understand it. The following example shows that these topologies are strictly coarser than the limit topology. The example is for $M = T^*S^1$, but it can easily be generalised to any symplectic manifold by using a Darboux chart adapted to a given Lagrangian submanifold.

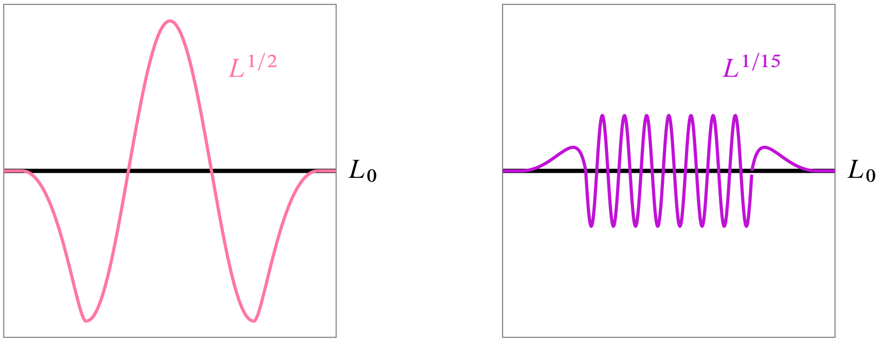


Figure 3: The behaviour of $\{L^s\}$ as s tends to 0.

Example On $M = T^*S^1$, consider the 1-parameter family of Hamiltonian $\{H^s\}_{s>0}$ defined via

$$H^s(q,p) = s^{3/2}\beta(q)\sin\frac{q}{s},$$

where we have identified S^1 with $\mathbb{R}/\mathbb{Z}$ and $\beta : [0,1] \to [0,1]$ is zero near $\{0,1\}$ and takes value 1 on $\left[\frac{1}{4},\frac{3}{4}\right]$. Set $L^0 = \{p = 0\}$ and $L^s = \varphi_1^{H^s}(L^0) = \operatorname{graph}(-dH^s)$. See Figure 3 for a visualisation.

On the one hand, it is easy to see that $s \mapsto L^s$, $s \in [0,1]$, defines a path which is continuous with respect to both δ_H and d_H (and thus also d). In particular, the set $\{L^s\}_{s\in[0,1]} \subseteq \mathscr{L}^e_\infty$ is compact in all these metrics. On the other hand, it is not contained in any $\mathscr{L}^e_k$, since

$$\lim_{s\to 0}\|B_{L^s}\| = \lim_{s\to 0}\max_{q\in S^1}\frac{|(H^s)'''(q)|}{\left(1+((H^s)''(q))^2\right)^{3/2}} = \lim_{s\to 0} s^{-3/2} = \infty,$$

where we have used that the second fundamental form B_{L^s} takes on a particularly simple form for graphs in the flat cylinder. Therefore, it cannot be compact in the limit topology (see Lemma 5.7 below).

Lemma 5.7 *Let $\{X_k \subseteq X\}$ be an increasing sequence of compact subspaces of a Hausdorff space X. The space $\varinjlim X_k$ is Hausdorff. Moreover, a subset A of $\varinjlim X_k$ is compact if and only if it is closed and $A \subseteq X_k$ for some k.*

Proof We first prove the Hausdorffness. If $x \neq y \in \varinjlim X_k = \bigcup_k X_k$, then there are open subsets U and V of X such that $x \in U$ and $y \in V$, but $U \cap V = \varnothing$. But then, the restrictions $U \cap X_k$ and $V \cap X_k$ are open for each k, so that they are also open in $\varinjlim X_k$. Thus, x and y are also separated in $\varinjlim X_k$, and $\varinjlim X_k$ is indeed Hausdorff.

We now prove the equivalence. One direction is obvious. Let thus A be compact in $\varinjlim X_k$. Since the limit space is Hausdorff, A must be closed. Suppose however that A is not contained in any X_k. Then, for every k, there is some $x_k \in A - X_k$. In particular, the set $S = \{x_k\} \subseteq A$ is such that $S \cap X_k$ is finite for all k. Therefore, that intersection is closed, since X_k is Hausdorff. By definition of the limit topology, this thus means that S itself is closed. In fact, this logic shows that every subset of S is closed, i.e., S is a closed infinite discrete subset of A. But this is impossible if A is compact, hence the contradiction. □

The above lemma allows us to completely characterise the limit topology.

Proposition 5.8 *A sequence* $\{L_i\} \subseteq \varinjlim \mathscr{L}_k^\star$ *converges to some* $L \in \varinjlim \mathscr{L}_k^\star$ *if and only if there exists some* $k \in \mathbb{N}$ *such that* $\{L_i\} \subseteq \mathscr{L}_k^\star$ *and* $L_i \to L$ *in the Hausdorff topology — or in any of the many equivalent topology on* $\mathscr{L}_k^\star$*.*

Indeed, the result is a direct consequence of Lemma 5.7 just above and of the following simple fact from point-set topology.

Lemma 5.9 *Let* $\{x_i\}_{i\in\mathbb{N}}$ *be a sequence in a Hausdorff space* X *such that none of its elements is a limit point. The sequence converges to some* $x \in X$ *if and only if the subspace* $\overline{S} = \{x_i\}_{i\in\mathbb{N}} \cup \{x\}$ *is compact.*

Proof Suppose that $x_i \to x$, and let $\{U_a\}_{a\in A}$ be an open cover of $\overline{S}$. Then, there is some $a_0 \in A$ such that $x \in U_{a_0}$. But by convergence, there is some $N \in \mathbb{N}$ such that $x_i \in U_{a_0}$ for all $i > N$. It then suffices to pick a_i such that $x_i \in U_{a_i}$ for each $i \leq N$ to get a finite subcover $\{U_{a_i}\}_{i=0}^N$.

Suppose now that $\overline{S}$ is compact, and let U_0 be an open neighbourhood of x. Note that for each x_i, there is some open U_i such that $U_i \cap \overline{S}$ is finite. Indeed, otherwise, x_i would be a limit point of the sequence, which would be a contradiction with the hypothesis on $\{x_i\}$. Since X is Hausdorff, we may suppose that $U_i \cap \overline{S} = \{x_i\}$. Therefore, there must be only a finite number of i such that $x_i \notin U_0$, otherwise $\{U_i\}_{i=0}^\infty$ would be an open cover of $\overline{S}$ with no finite subcover. □

Given Proposition 5.8, we can see where the limit topology sits with regard to the various C^k-topologies.

Corollary 5.10 *The limit topology is (strictly) finer than the* $C^{1,\alpha}$*-topology, for any* $0 < \alpha < 1$*, but coarser than the* C^2*-topology.*

Proof By Theorem 2.6, the $C^{1,\alpha}$-topology is equivalent to the Hausdorff topology on $\mathscr{L}_k^\star$ for any $0 < \alpha < 1$. Therefore, every convergent sequence in the limit topology is also convergent (with the same limit) in the $C^{1,\alpha}$-topology. To see that the inclusion of topology is strict, just use the example above, but replace $s^{3/2}$ by $s^{2+\alpha}$ in the definition of H^s.

The C^2-topology is coarser because every C^2-converging sequence has uniformly bounded second fundamental form (direct computation), is uniformly ε-tame (this is the idea of Lemma A.11), and obviously also converges in the Hausdorff topology to the same limit. □

Appendix Some results in Riemannian geometry

This appendix compiles all the results in Riemannian geometry which were required throughout the paper, but that the author could not find in the literature. We suspect that many of these results are known to experts, but this is not the case for all of them: at least Lemma A.5 has appeared as a conjecture in a paper of Albuquerque [3].

Below, Section A.1 compiles the results on the Sasaki metric, Section A.2 on Riemannian surfaces, and Section A.3 on comparison results between Riemannian invariants of Lagrangian submanifolds.

A.1 Results on the Sasaki metric

Behaviour along graphs We begin by proving some useful results on the behaviour of some Riemannian invariants of graphs in TL of vector fields under the transformation $(x, v) \mapsto (x, tv)$. More precisely, we show that the norm of the second fundamental form and the tameness constant must the nondecreasing in t if the vector field is (locally) a gradient of a C^2-small function.

We begin by studying the norm of the curvature. The proof is elementary but still subtle.

Lemma A.1 *Equip L with a Riemannian metric $g = \langle \cdot, \cdot \rangle$ and TL with its associated Sasaki metric. Let $\xi = \operatorname{grad} H \in \mathfrak{X}(L)$. If $|\xi|$ and $|\nabla\xi|$ are sufficiently small, then the function $t \mapsto \|B_{t\xi}\|$ is nondecreasing for $t \in [0, 1]$.*

Proof For $X \in T_x L$, we denote by X^h and X^v its horizontal and vertical lifts in $T_{(x,y)}TL$, respectively. Then, we have that

$$T_{\xi(x)}\xi(L) = \{\widetilde{X} = X^h + (\nabla_X \xi)^v \mid X \in T_x L\}, \tag{6}$$

$$T^{\perp}_{\xi(x)}\xi(L) = \{\widetilde{Z} = Z^v - ((\nabla\xi)^* Z)^h \mid Z \in T_x L\}, \tag{7}$$

where $\langle (\nabla\xi)^* Z, Y \rangle = \langle Z, \nabla_Y \xi \rangle$ for all $Y \in T_x L$ (see, for example, [1]). Denoting by $\widetilde{\nabla}$ the Levi-Civita connection on TL, by ∇ the Levi-Civita connection on L, and by R the Riemann curvature tensor on L, we get that

$$\begin{aligned}
\widetilde{\nabla}_{X^h} Y^h &= (\nabla_X Y)^h - \tfrac{1}{2}(R(X, Y)\xi)^v, \\
\widetilde{\nabla}_{X^h} Y^v &= (\nabla_X Y)^v + \tfrac{1}{2}(R(\xi, Y)X)^h, \\
\widetilde{\nabla}_{X^v} Y^h &= \tfrac{1}{2}(R(\xi, X)Y)^h, \\
\widetilde{\nabla}_{X^v} Y^v &= 0
\end{aligned}$$

for all $X, Y \in \mathfrak{X}(L)$. Therefore, the expression for the second fundamental form of $\xi(L)$ is

$$B_\xi(\widetilde{X}, \widetilde{X}, \widetilde{Z}) = \overbrace{\langle \nabla^2_X \xi - \nabla_{\nabla_X X}\xi, Z \rangle}^{\alpha} - \overbrace{\langle \nabla_{R(\xi, \nabla_X \xi)X}\xi, Z \rangle}^{\beta}$$

for all $\widetilde{X} \in T_{\xi(x)}\xi(L)$ and all $\widetilde{Z} \in T^{\perp}_{\xi(x)}\xi(L)$.

From (6) and the definition of the Sasaki metric, we have that $|\widetilde{X}|^2 = |X|^2 + |\nabla_X \xi|^2$. Likewise, from (7), we have that $|\widetilde{Z}|^2 = |Z|^2 + |\nabla_Z \xi|^2$; this is because

$$|(\nabla\xi)^* Z|^2 = \operatorname{Hess} H(Z, (\nabla\xi)^* Z) = \langle (\nabla\xi)^* Z, \nabla_Z \xi \rangle = \operatorname{Hess} H(Z, \nabla_Z \xi) = |\nabla_Z \xi|^2,$$

since the Hessian of a function is symmetric. Therefore, for every $t \in [0, 1]$, the map

$$TTL \to TTL, \quad ((x, y), \widetilde{Y}) \mapsto \bigl((x, ty), (1 + (t^2 - 1)|\nabla_Y \xi|^2)^{-1/2}\widetilde{Y}\bigr),$$

sends $\xi(x)$ to $t\xi(x)$ and sends diffeomorphically the unit sphere of $T_{\xi(x)}\xi(L)$, respectively of $T^{\perp}_{\xi(x)}\xi(L)$, onto the one of $T_{t\xi(x)}t\xi(L)$, respectively of $T^{\perp}_{t\xi(x)}t\xi(L)$. Here, Y denotes the sum of the projections of $\widetilde{Y}$ onto the horizontal and vertical distributions, after their identification with TL. Note that, on these spheres,

$|\widetilde{Y}|^2 = |Y|^2 + |\nabla_Y \xi|^2 = 1$ so that $|Y| \le 1$ and $|\nabla_Y \xi| < 1$. In particular, $s_Y(t) := (1 + (t^2 - 1)|\nabla_Y \xi|^2)^{-1/2}$ is well defined for $t \in [0, 1]$.

Therefore, it does suffice to prove that the map

$$t \mapsto \bigl|\bigl(B_{t\xi}(s_X(t)\widetilde{X}, s_X(t)\widetilde{X}, s_Z(t)\widetilde{Z})\bigr)\bigr| = s_X^2 s_Z t |\alpha - t^2 \beta| \tag{8}$$

is nondecreasing for $t \in [0, 1]$, for all $\widetilde{X} \in T_{\xi(x)}\xi(L)$ and all $\widetilde{Z} \in T^{\perp}_{\xi(x)}\xi(L)$ such that $|\widetilde{X}| = |\widetilde{Z}| = 1$. Indeed, by the previous discussion, this will imply that the map $t \mapsto |B_{t\xi}|_x^{\mathrm{op,sym}}$ — sending t to the operator norm of $B_{t\xi}$ on the subspace of $T_{\xi(x)} t\xi(L) \otimes T_{\xi(x)} t\xi(L) \otimes T^{\perp}_{\xi(x)} t\xi(L)$ generated by elements of the form $\widetilde{X} \otimes \widetilde{X} \otimes \widetilde{Z}$ — is nondecreasing. Since $B_{t\xi}$ is symmetric in its first two entries, this is just the operator norm on the whole space, and we will get the result.

If $\alpha = 0$, we may suppose that $\beta \neq 0$, otherwise (8) is just the zero function, and the statement is trivial. In that case, (8) looks like $6|\beta|(1 - |\nabla_X \xi|^2)^{-1}(1 - |\nabla_Z \xi|^2)^{-1/2} t^3 + \mathcal{O}(t^4)$ near $t = 0$. In particular, it is increasing near $t = 0$. But (8) only possibly has critical points at

$$t = 0 \quad \text{and} \quad t = \pm \sqrt{\frac{3(1 - |\nabla_X \xi|^2)(1 - |\nabla_Z \xi|^2)}{|\nabla_X \xi|^2 |\nabla_Z \xi|^2 - |\nabla_X \xi|^2| - 2|\nabla_Z \xi|^2}}.$$

For $|\nabla \xi|$ small enough, the latter values are not real, and thus (8) is increasing.

Suppose now that $\alpha \neq 0$. By changing the sign of $\widetilde{Z}$ if necessary, we may assume that $\alpha > 0$. Since $|\beta| \le \|R\| \, |\xi| \, |\nabla \xi|^2$ and α depends only on derivatives of ξ, we thus have that $|\alpha - t^2\beta| = \alpha - t^2\beta$ for all $t \in [0, 1]$ if $|\xi|$ is small enough. But the function $t \mapsto s_X^2 s_Z t(\alpha - t^2\beta)$ converges with all derivatives to the function $t \mapsto t\alpha$ as $|\nabla \xi| \to 0$. Since that function is increasing, the derivative of (8) is positive for all $t \in [0, 1]$ for $|\nabla \xi|$ small enough. Therefore, the function is increasing over the interval for $|\xi|$ and $|\nabla \xi|$ small enough. □

Remark A.2 Given the proof of Lemma A.1, it appears that how small we must take $|\xi|$ and $|\nabla \xi|$ depends on $\widetilde{X}$ and $\widetilde{Z}$. However, since the infimum to get the operator norm $\|B_{t\xi}\|$ is taken over the unit sphere, which is compact, it is in fact a minimum. Therefore, how small we take $|\xi|$ and $|\nabla \xi|$ can be made independent of $\widetilde{X}$ and $\widetilde{Z}$.

We now move on to studying the tameness constant. This time, the proof is fairly straightforward.

Lemma A.3 *Let* $\xi \in \mathfrak{X}(L)$, *and equip* TL *with a Sasaki metric. Define*

$$\varepsilon_\xi := \inf_{x \neq y \in \xi(L)} \frac{d_{TL}(x, y)}{\min\{1, d_\xi(x, y)\}} \in (0, 1],$$

where d_ξ *denotes the intrinsic distance in* $\xi(L)$. *Then,*

$$\lim_{|\nabla \xi| \to 0} \varepsilon_\xi = 1.$$

In particular, $\varepsilon_\xi > (k + 1)^{-1}$ *for* $|\nabla \xi|$ *small enough.*

Proof Consider a path $\gamma : [0, \ell] \to L$ such that $|\dot{\gamma}| \equiv 1$. Define $\tilde{\gamma} := \xi \circ \gamma$. Then, $|\dot{\tilde{\gamma}}|^2 = 1 + |\nabla_{\dot{\gamma}}\xi|^2$, so that

$$\ell \le \int_0^\ell |\dot{\tilde{\gamma}}|\, dt \le \ell\sqrt{1 + |\nabla\xi|^2}.$$

Taking the infimum of the above inequality over all paths γ such that $\gamma(0) = x$ and $\gamma(\ell) = y$ for given $x, y \in L$, we get that

$$d_L(x, y) \le d_\xi(\xi(x), \xi(y)) \le \sqrt{1 + |\nabla\xi|^2}\, d_L(x, y), \tag{9}$$

since all smooth paths $\tilde{\gamma}$ in $\xi(L)$ may be parametrised to be of the form $\tilde{\gamma} = \xi \circ \gamma$ with $|\dot{\gamma}| \equiv 1$.

On the other hand, take a path $\tilde{\gamma} : [0, \tilde{\ell}] \to TL$ such that $|\dot{\tilde{\gamma}}| \equiv 1$, $\gamma(0) = \xi(x)$, and $\gamma(L) = \xi(y)$ for given $x, y \in L$. Then, we may write $\dot{\tilde{\gamma}} = \dot{\gamma}^h + Y^v$ for $\gamma := \pi \circ \tilde{\gamma}$ and a vector field Y of L along γ, where $\pi : TL \to L$ is the canonical projection. We thus get

$$\tilde{\ell} = \int_0^{\tilde{\ell}} \sqrt{|\dot{\gamma}|^2 + |Y|^2}\, dt \ge \int_0^{\tilde{\ell}} |\dot{\gamma}|\, dt.$$

Taking the infimum over all possible $\tilde{\gamma}$, we get that

$$d_{TL}(\xi(x), \xi(y)) \ge d_L(x, y), \tag{10}$$

since every path in L from x to y admits a lift to TL from $\xi(x)$ to $\xi(y)$ (e.g., $\xi \circ \gamma$).

Putting (9) and (10) together, we thus get

$$\inf_{x \ne y} \frac{d_L(x, y)}{\min\bigl\{1, d_L(x, y)\sqrt{1 + |\nabla\xi|^2}\bigr\}} \le \varepsilon_\xi \le 1,$$

which implies the result. □

Remark A.4 The approaches in the proofs of Lemmata A.1 and A.3 are different, because $\|B_\xi\|$ depends on higher derivatives of ξ, while ε_ξ does not. Therefore, just taking a limit in the expression for $\|B_\xi\|$ would not lead to 0, and we could not conclude anything. We need to actually understand the behaviour of $\|B_\xi\|$ in a $C^{1,\alpha'}$-neighbourhood of 0, not just in the limit $|\xi| \to 0$.

Behaviour of the geodesics To adapt Lemma A.3 to a metric which is only locally Sasaki — which is required in Section 3.2 — we will need the following technical results. As far as we know, these results have not appeared in the literature. This is also why Lemma A.5 is established in such generality: it has appeared before as a conjecture of Albuquerque [3] and could be of general interest to the Riemannian geometry community.

Lemma A.5 *The space TN is complete in the Sasaki metric associated with (N, g) if and only if (N, g) is complete.*

Proof One direction is obvious: if TN is complete and $v \in T_xN$, the exponential of $tv \in T_xN \subseteq T_xTN$ exists in TN for all $t \in \mathbb{R}$. However, N is totally geodesic in TN [36], so that the geodesic $t \mapsto \exp_x^{TN}(tv)$ must stay in N. It is thus the exponential of tv in N, and N must be complete.

Suppose now that N is complete. We recall that completeness of a Riemannian manifold M is equivalent to it having an exhaustion by compact sets $\{K_i\}$ such that, if $\{x_i\}$ is a sequence with $x_i \notin K_i$, then $d(y, x_i) \to \infty$ for some point $y \in M$ — this is part of the classical Hopf–Rinow theorem; see, for example, Theorem 7.2.8 in [9].

Let $y \in N$ and $K_i = D_i N|_{B_i^N(y)}$, where $B_i^N(y)$ is the (closed) ball of radius i in N centred at y. Let $\{x_i\}$ be such that $x_i \notin K_i$. We now study two possible types of subsequences of $\{x_i\}$.

(1) Suppose there is a subsequence, still denoted by $\{x_i\}$, such that none of the subsequences of $\{\pi(x_i)\}$ are contained in any of the $B_i^N(y)$. Then, the sequence of natural numbers given by

$$n_i = \min\{j \mid \pi(x_i) \in B_j^N(y)\}$$

converges to infinity. By completeness of N, this must mean that $d(y, \pi(x_i)) \to \infty$. But since π is a Riemannian submersion, it is nonexpansive in d, so that

$$d(y, \pi(x_i)) = d(\pi(y), \pi(x_i)) \le d(y, x_i).$$

Therefore, $d(y, x_i) \to \infty$.

(2) Suppose there is a subsequence, still denoted by $\{x_i\}$, and a $R > 0$ such that $\{\pi(x_i)\}$ is contained in $B_R^N(y)$. Since $x_i \notin K_i$, this forces that $x_i \notin D_i N$ for large enough i. Then, let γ_i be a minimal geodesic of N from $\pi(x_i)$ to y. Let $x_i' := P_{\gamma_i}(x_i)$ be the parallel transport of x_i along γ_i. Note that the horizontal lift $\tilde{\gamma}_i$ of γ_i starting at x_i ends at x_i' by construction. Therefore, we have that

$$d(x_i, x_i') \le \ell(\tilde{\gamma}_i) = \ell(\gamma_i) = d(\pi(x_i), y) \le R.$$

On the other hand, since parallel transport is an isometry on the fibres, the fact that $x_i \notin D_i N$ ensures that $x_i' \notin D_i N$. Since x_i' is in the fibre over y, this thus implies that $d(y, x_i') > i$. Therefore, the triangle inequality gives that

$$d(y, x_i) \ge d(y, x_i') - d(x_i, x_i') > i - R,$$

and $d(y, x_i) \to \infty$.

Since the original $\{x_i\}$ sequence can be written as the union of subsequences of either type, we conclude that $d(y, x_i) \to \infty$, so that TN is complete. □

Lemma A.6 *If $\alpha : [0, \ell] \to TL$ is a geodesic in the Sasaki metric, then the function $t \mapsto |\alpha(t)|^2$ is either constant or a (strictly) convex parabola. In particular, the disk bundle $D_r L$ of radius r is geodesically convex.*

Proof See α as a vector field Y along the path $x := \pi \circ \alpha : [0, \ell] \to L$. Note that $|\alpha| = |Y|$. Then, the geodesic on TL is equivalent to two equations on L [36]:

$$\begin{cases} \nabla_{\dot{x}} \dot{x} + R(Y, \nabla_{\dot{x}} Y)\dot{x} = 0; \\ \nabla_{\dot{x}}^2 Y = 0. \end{cases}$$

But then, this means that

$$\begin{aligned}\frac{d}{dt}|Y|^2 &= 2\langle \nabla_{\dot{x}} Y, Y\rangle;\\ \frac{d^2}{dt^2}|Y|^2 &= 2\langle \nabla_{\dot{x}}^2 Y, Y\rangle + 2|\nabla_{\dot{x}} Y|^2 = 2|\nabla_{\dot{x}} Y|^2;\\ \frac{d}{dt}|\nabla_{\dot{x}} Y|^2 &= 2\langle \nabla_{\dot{x}}^2 Y, \nabla_{\dot{x}} Y\rangle = 0.\end{aligned}$$

From the last equation, we get that $|\nabla_{\dot{x}} Y|$ is independent of time. If $|\nabla_{\dot{x}} Y| \equiv 0$, then $\nabla_{\dot{x}} Y \equiv 0$, so that the first equation implies that $|Y|$ is constant. If $|\nabla_{\dot{x}} Y| > 0$, then the middle equation implies that the second time derivative of $|Y|^2$ is a positive constant, thus giving that it is a strictly convex parabola. □

As mentioned above, we need to also study metrics which are only locally Sasaki. By this, we mean that (M, ω) is a symplectic manifold with a compatible almost complex structure J and that $g = \omega(\cdot\,, J\cdot)$. We also suppose that L is a Lagrangian submanifold of M. By locally Sasaki, we mean that there is a diffeomorphism Ψ from a neighbourhood of L in TL to a neighbourhood of L in M such that $\Psi^* g$ is the Sasaki metric. This makes sense since J identifies TL with $TL^{\perp} \subseteq TM$.

In this case, Lemma A.1 obviously still applies, but Lemma A.3 needs to be adapted, as we could have $d_{TL} \neq d_M$. In other words, we have to deal with the fact that the minimal geodesic in M between two points of $L' = \xi(L)$ might not be entirely contained in the neighbourhood of L where g is equal to a Sasaki metric. In particular, that minimal geodesic could be shorter than one would expect in TL so that the ratio d_M/d_L might no longer tend to 1 as $|\nabla\xi| \to 0$. We prove that this in fact cannot happen.

Lemma A.7 *Let M, L, and g as above. On a small enough neighbourhood of L, we have that $d_M = d_{TL}$.*

Proof We first prove that $d_M(x, y) = d_{TL}(x, y)$ whenever $x, y \in L$. Let thus x and y be in L. Let $\alpha : [0, \ell] \to TL$ be a minimal geodesic in TL from x to y. Note that it follows from Lemma A.6 that α is fully contained in L, so that $d_{TL}(x, y) = d_L(x, y)$. In particular, α is also a geodesic of M. Therefore, it is locally minimising in M, and we can take

$$\ell' := \sup\{t \in [0, \ell] \mid d_M(x, \alpha(s)) = s \ \forall s < t\} \in (0, \ell].$$

Suppose that α is not minimising, i.e., $\ell' < \ell$, and let $\gamma : [0, \ell'] \to M$ a minimal geodesic in M from x to $y' := \alpha(\ell')$ which is different from α — nonminimality of α ensures that it exists.

From classical facts from Riemannian geometry (see Proposition 13.2.12 of [9] for example), exactly one of two things can happen: either there is a 1-parameter family γ_s of geodesics from x to y' with $\gamma_0 = \alpha$ and $\gamma_1 = \gamma$, or γ and α are the only two minimising geodesic from x and y' and $\gamma'(\ell') = -\alpha'(\ell')$. In the first case, note that all γ_s have the same length since geodesics are critical points of the length functional. Therefore, for s small enough, γ_s is a minimal geodesic not contained in L, but fully contained in the neighbourhood of L where g is the Sasaki metric. This is of course a contradiction with Lemma A.6, since 0 and ℓ would then both have to be strict minima of $t \mapsto |\gamma_s(t)|^2$. In the second case, γ is then tangent to L at $t = \ell'$. But since L is totally geodesic, γ must then be fully contained in L, and we again get a contradiction. Therefore, the result holds on L.

We now consider x and y close to L in M. Let α, γ, ℓ', and y' be defined analogously as above. From Lemma A.5, α exists and, if we take x and y to be in a neighbourhood of the form $\Psi(D_r L)$, then α stays in that neighbourhood by Lemma A.6. Again, we have two possibilities for how γ and $\alpha|_{[0,\ell']}$ connect. If γ_s is a 1-parameter family, then we still get a contradiction for some s: since γ_1 leaves the Weinstein neighbourhood and γ_s always has the same endpoints, there is some s such that γ_s is still in that neighbourhood, but such that $t \mapsto |\gamma_s(t)|$ has a maximum (in contradiction with Lemma A.6).

However, the second possibility — that α and γ form a geodesic loop in M — does not *a priori* lead to a contradiction. Thus take sequences $\{x_i\}$ and $\{y_i\}$ such that $\lim_i d_M(x_i, L) = \lim_i d_M(y_i, L) = 0$, but such that $d_M(x_i, y_i) < d_{TL}(x_i, y_i)$ for all i. Define α_i, γ_i, ℓ_i', and y_i' analogously as before. In particular, if v_i is the unit vector such that $\exp_{x_i}(tv_i) = \alpha(t)$ for $t \in [0, \ell_i]$, then $t \mapsto \exp_{x_i}(tv_i)$, $t \in [2, \ell_i']$, is the geodesic loop $\alpha_i \# \overline{\gamma_i}$. Since $\{x_i\}$, $\{y_i\}$, and $\{v_i\}$ are all contained in a compact, we may pass to a subsequence, so that $\lim_i x_i = x \in L$, $\lim y_i = y \in L$, and $\lim v_i = v \in T_x M$. But then, $t \mapsto \exp_x(tv)$, $t \in [0, 2\ell']$, is a geodesic loop in M which is fully contained in L over $[0, \ell]$, but that eventually leaves it. Therefore, we get a last contradiction, and we must have $d_M(x, y) = d_{TL}(x, y)$ whenever x and y are close to L. □

A.2 Result on Riemann surfaces

We suppose that (M, g, J, ω) is a Riemann surface. Then, a tubular neighbourhood of a curve L admits some fairly nice coordinates given by

$$\varphi\colon (0,\ell) \times (-r, r) \to M, \quad (s,t) \mapsto \exp_{\gamma(s)}(tJ\dot\gamma(s)),$$

where $\gamma : [0, \ell) \to L$ is a parametrisation such that $|\dot\gamma| \equiv 1$. Note that

$$\begin{aligned} \varphi^* g &= |W|^2 ds^2 + dt^2, \\ \varphi^* J &= |W| \frac{\partial}{\partial t} \otimes ds - \frac{1}{|W|}\frac{\partial}{\partial s} \otimes dt, \end{aligned}$$

where $W(s,t)$ is the value at time t of the unique Jacobi field along the geodesic $t \mapsto \varphi(s,t)$ such that $W(s,0) = \dot\gamma(s)$ and $\dot W(s,0) = -\kappa(s)\dot\gamma(s)$. Here, κ is the (signed) geodesic curvature of L, which is defined via the relation $\ddot\gamma := \nabla_{\dot\gamma}\dot\gamma = \kappa J\dot\gamma$.

In these coordinates, we will call $L' = \{(s, \xi(s)) \mid s \in [0,\ell)\} =: \operatorname{graph}\xi$ the graph of $\xi : [0,\ell] \to \mathbb{R}$. Because our applications are aimed towards Section 3.2, we will be interest in 1-parameter families of the form $\{\xi_\alpha = \alpha\xi + c(\alpha)\}_{\alpha\in[0,1]}$, where $\|\xi\| < \frac{r}{2}$ and c is such that

(1) $c(0) = c(1) = 0$;

(2) $|c(\alpha)| \le \alpha\|\xi\|$.

In particular, $\operatorname{graph}\xi_\alpha$ stays in the chart defined by φ for all $\alpha \in [0,1]$.

Before moving on with the results on ξ_α, we prove the following lemma on the behaviour of $|W|$ for small values of $|t|$, which will be quite useful later on.

Lemma A.8 *Let $K : [0,\ell) \to \mathbb{R}$ be the pullback of the Gaussian curvature of M along γ. We have that*

$$|W|^2 = 1 - 2\kappa t + (\kappa^2 - K)t^2 + \mathcal{O}(t^3).$$

Proof As noted above, we have that $W(s,0) = \dot\gamma(s)$ and $\dot W(s,0) = -\kappa(s)\dot\gamma(s)$. The lemma then follows directly from

$$\begin{aligned} |W|^2|_{t=0} &= 1, \\ \frac{\partial}{\partial t}|W|^2|_{t=0} &= 2\langle \dot W, W\rangle|_{t=0} = -2\kappa, \\ \frac{\partial^2}{\partial t^2}|W|^2|_{t=0} &= 2|\dot W|^2|_{t=0} + 2\langle \ddot W, W\rangle|_{t=0} \\ &= 2\kappa^2 - 2\langle R(J\dot\gamma,\dot\gamma)J\dot\gamma,\dot\gamma\rangle = 2(\kappa^2 - K), \end{aligned}$$

where the last line follows from the fact that W satisfies the Jacobi equation. □

With this in hand, we can estimate the geodesic curvature of graph ξ_α in terms of that of L and graph ξ. More precisely, we want to prove the following.

Lemma A.9 *For every $k \ge 0$ and every $k' > k$, there exists $\delta > 0$ with the following property. If $\|B_L\| \le k$ and $L' = \operatorname{graph}\xi$ with $\|\xi\|, \|\xi'\| < \delta$, then $\|B_\alpha\| := \|B_{\operatorname{graph}\xi_\alpha}\| \le \max\{k', \|B_{L'}\|\}$ for all $\alpha \in [0,1]$.*

The lemma itself relies on the following computation.

Lemma A.10 *The geodesic curvature of graph ξ at a point is given by*

$$|B| = \frac{|W|}{(|W|^2 + |\xi'|^2)^{3/2}}\left|\xi'' + \frac{1}{2}\frac{\partial}{\partial t}|W|^2 - \frac{\xi'}{|W|^2}\left(\frac{1}{2}\frac{\partial}{\partial s}|W|^2 + \xi'\frac{\partial}{\partial t}|W|^2\right)\right|,$$

where $|W|(s) := |W(s,\xi(s))|$.

Proof This is a direct computation, but we give here the important steps. The only nonzero Christoffel symbols of φ^*g are

$$\Gamma^s_{ss} = \frac{1}{2|W|^2}\frac{\partial}{\partial s}|W|^2, \quad \Gamma^s_{st} = \Gamma^s_{ts} = \frac{1}{2|W|^2}\frac{\partial}{\partial t}|W|^2, \quad \Gamma^t_{ss} = \frac{1}{2}\frac{\partial}{\partial t}|W|^2.$$

Therefore, if we set $\Gamma(s) := (s,\xi(s))$, we get that

$$\begin{aligned} \dot\Gamma &= \frac{\partial}{\partial s} + \xi'\frac{\partial}{\partial t}, \\ J\dot\Gamma &= |W|\frac{\partial}{\partial t} - \frac{\xi'}{|W|}\frac{\partial}{\partial s}, \\ \ddot\Gamma &= \frac{1}{|W|^2}\left(\frac{1}{2}\frac{\partial}{\partial s}|W|^2 + \xi'\frac{\partial}{\partial t}|W|^2\right)\frac{\partial}{\partial s} + \left(\xi'' + \frac{1}{2}\frac{\partial}{\partial t}|W|^2\right)\frac{\partial}{\partial t}. \end{aligned}$$

Since the geodesic curvature is given by

$$|B| = \frac{|\langle \ddot\Gamma, J\dot\Gamma\rangle|}{|\dot\Gamma|^3},$$

this gives the above formula. □

Proof of Lemma A.9 The proof is somewhat tedious but quite elementary. We first combine Lemmata A.8 and A.10 to get the identity

$$|B|^2 = (1 + R_1)(\xi'' - \kappa + R_2)^2, \tag{11}$$

where R_1 and R_2 are smooth functions of s. Furthermore, at a given s, they only depend on L, M, g, and the values $\xi(s)$ and $\xi'(s)$, and in such a manner that $R_i \to 0$ as $\xi(s), \xi'(s) \to 0$.

Note that (11) implies that if $\|B\| \le k$ and $\|\xi\|, \|\xi'\| \le \frac{r}{2}$, then $\|\xi''\|$ is bounded by some constant C depending only on L, M, g, and the constants k and r. This allows us to write

$$|B_\alpha|^2 = (\alpha\xi'' - \kappa)^2 + \alpha \cdot \mathcal{O}(|\xi|, |\xi'|) \tag{12}$$

for C^1-close graphs with bounded ξ''-dependence in the error term. That is, $|B_{\mathrm{graph}\,\xi^i_\alpha}|^2 - (\alpha(\xi^i)'' - \kappa)^2$ tends uniformly to 0 if $\{\xi^i\}$ C^1-converges to 0. Here, we have made use of the fact that $\xi''_\alpha = \alpha\xi''$, $\xi'_\alpha = \alpha\xi'$, and $|\xi_\alpha| \le 2\alpha|\xi|$ — this is part of the hypotheses on c. The rest of the proof then consists of studying the behaviour of the parabola $\alpha \mapsto (\alpha\xi'' - \kappa)^2$ under small perturbations.

Fix $\varepsilon \in \left(0, \frac{1}{2}\right)$. We break down the analysis into a few subcases.

(1) $\boldsymbol{|\xi''| \le \varepsilon}$ Then, $(\alpha\xi'' - \kappa)^2 \le (|\kappa| + \varepsilon)^2 \le (k + \varepsilon)^2$. By supposing $|\xi|$ and $|\xi'|$ small enough, we may suppose the error term in (12) to be smaller than ε, so that

$$|B_\alpha|^2 \le (k + \varepsilon)^2 + \varepsilon.$$

(2) $\boldsymbol{|\xi''| > \varepsilon}$ There are three subcases (see Figure 4 for a visualisation).

(a) $\boldsymbol{\frac{2\kappa}{\xi''} + \varepsilon \ge 1}$ In this case, $\frac{\kappa}{\xi''} > 0$, and we have that the maximum of $\alpha \mapsto (\alpha\xi'' - \kappa)^2$ on the interval $\left[0, \frac{2\kappa}{\xi''}\right]$ is reached at $\alpha = 0$. Therefore, its maximum value on $\left[0, \frac{2\kappa}{\xi''}\right]$ is $\kappa^2 \le k^2$.

On the interval $\left[\frac{2\kappa}{\xi''}, \frac{2\kappa}{\xi''} + \varepsilon\right]$, the parabola is increasing with derivative

$$2(\alpha\xi'' - \kappa)\xi'' = 2|\alpha\xi'' - \kappa|\,|\xi''| \le 2(\alpha|\xi''| + |\kappa|)|\xi''| \le 2C(C + k).$$

Therefore, the maximum value of the parabola over that subinterval is bounded from above by $k^2 + 2C(C + k)\varepsilon$. Thus, for $|\xi|$ and $|\xi'|$ as in case (1), we get

$$|B_\alpha|^2 \le k^2 + (2C(C + k) + 1)\varepsilon$$

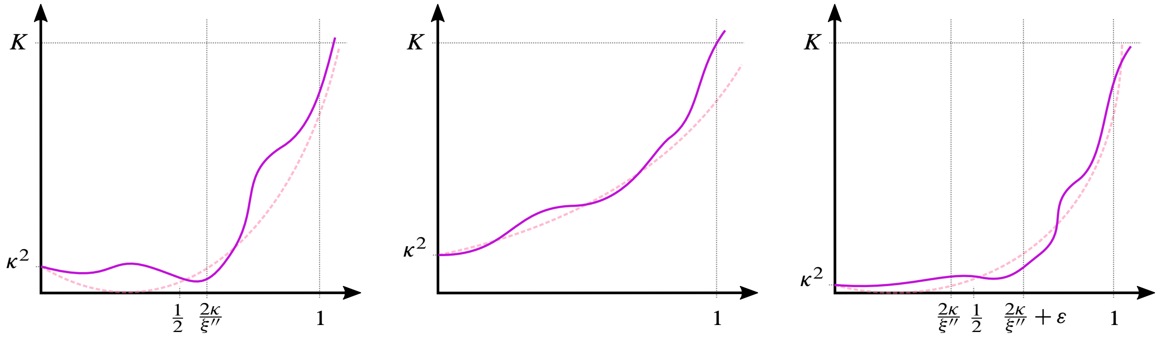
since $[0, 1] \subseteq \left[0, \frac{2\kappa}{\xi''} + \varepsilon\right]$.

(b) $\boldsymbol{\frac{2\kappa}{\xi''} + \varepsilon \le 0}$ In this case, the parabola is increasing over $[0, 1]$ with derivative

$$2(\alpha\xi'' - \kappa)\xi'' = 2|\alpha\xi'' - \kappa|\,|\xi''| \ge 2|\kappa|\,|\xi''| \ge \varepsilon|\xi''|^2 > \varepsilon^3.$$

By supposing $|\xi|$ and $|\xi'|$ small enough, we may suppose the α-derivative of the error term in (12) to be smaller than $\frac{\varepsilon^3}{2}$. Therefore, $\alpha \mapsto |B_\alpha|$ is still increasing, and we have that

$$|B_\alpha| \le |B_1| = |B_{L'}|.$$



case (a) case (b) case (c)

Figure 4: Possible graph of $|B_\alpha|$ (solid, purple) with idealised parabola (dashed, pink) and important values (pointed, grey). We denote by K whichever bound we get in the corresponding case.

(c) $\boldsymbol{0 < \frac{2\kappa}{\xi''} + \varepsilon < 1}$ This case combines the approaches of (a) and (b). Indeed, the estimates of (a) still hold for $\alpha \in \left[0, \frac{2\kappa}{\xi''} + \varepsilon\right] \subseteq [0,1]$. Likewise, over the interval $\left[\frac{2\kappa}{\xi''} + \varepsilon, 1\right]$, the parabola is increasing with derivative

$$2(\alpha\xi'' - \kappa)\xi'' = 2|\alpha\xi'' - \kappa|\,|\xi''| \geq 2|\varepsilon\xi'' + \kappa|\,|\xi''| > \varepsilon|\xi''|^2 > \varepsilon^3.$$

Therefore, for $|\xi|$ and $|\xi'|$ as in (b), we have that

$$|B_\alpha| \leq \max\{k^2 + (2C(C+k)+1)\varepsilon, |B_{L'}|\}.$$

The result then follows by taking ε such that

$$\max\{(k+\varepsilon)^2 + \varepsilon, k^2 + (2C(C+k)+1)\varepsilon\} \leq (k')^2. \qquad \square$$

We now move on to the analysis of the tameness constant of $\operatorname{graph}\xi_\alpha$. The approach that we take here is similar in spirit to the one taken in the proof of Lemma A.3.

Lemma A.11 *Given* $L' = \operatorname{graph}\xi \subseteq \varphi([0,\ell] \times (-r,r))$, *denote by* d_ξ *the distance function on* $\operatorname{graph}\xi$ *induced by the Riemannian metric* $\varphi^* g|_{\operatorname{graph}\xi}$. *Take*

$$\varepsilon_\xi := \inf_{x \neq x' \in \operatorname{graph}\xi} \frac{d_M(x,x')}{\min\{1, d_\xi(x,x')\}} \in (0,1].$$

We have that $\varepsilon_\xi \to \varepsilon_0$ *as* $\xi \xrightarrow{C^1} 0$.

Proof Fix $x \neq x' \in \operatorname{graph}\xi$, and take $\Gamma(s) = (s, \xi(s))$. Without loss of generality, we may suppose that $x = \Gamma(s_0)$ and $x' = \Gamma(s_1)$ for $0 < s_1 - s_0 \leq \frac{\ell}{2}$. Therefore, $d_\xi(x,x')$ is simply the length $\ell(\Gamma)$ of Γ along the interval $[s_0, s_1]$. Note that

$$|\dot{\Gamma}|^2 = |W|^2 + |\xi'|^2.$$

By Lemma A.8, we thus have that

$$d_\xi(x,x') \le \int_{s_0}^{s_1} (1 + C(\|\xi\| + \|\xi'\|))\,ds = (1 + C(\|\xi\| + \|\xi'\|))(s_1 - s_0)$$

for some constant $C \ge 0$ depending only on M, L, g, and r.

On the other hand, the path $t \mapsto (s, t\xi(s))$, $t \in [0,1]$, has length $|\xi(s)|$ for any s. Therefore, we get that

$$d_M(x,x') \ge d_M(\gamma(s_0),\gamma(s_1)) - d_M(\gamma(s_0),x) - d_M(x',\gamma(s_1)) \ge d_M(\gamma(s_0),\gamma(s_1)) - 2\|\xi\|.$$

Combining these two estimates, we get that

$$\frac{d_M(\Gamma(s_0),\Gamma(s_1))}{\min\{1, d_\xi(\Gamma(s_0),\Gamma(s_1))\}} \ge \frac{d_M(\gamma(s_0),\gamma(s_1)) - 2\|\xi\|}{(1 + C(\|\xi\| + \|\xi'\|))\min\{1, d_L(\gamma(s_0),\gamma(s_1))\}},$$

since $s_1 - s_0 = d_L(\gamma(s_0),\gamma(s_1))$. We similarly get

$$\frac{d_M(\Gamma(s_0),\Gamma(s_1))}{\min\{1, d_\xi(\Gamma(s_0),\Gamma(s_1))\}} \le \frac{d_M(\gamma(s_0),\gamma(s_1)) + 2\|\xi\|}{(1 - C(\|\xi\| + \|\xi'\|))\min\{1, d_L(\gamma(s_0),\gamma(s_1))\}}.$$

However, note that both functions on the right-hand side uniformly — in s_0 and s_1 — converge to the ratio $d_M / \min\{1, d_L\}$. Therefore, the same holds for the left-hand side. Since uniform limits and infima commute, the result ensues. □

A.3 Comparison between metrics

To simplify our computations, we only prove our comparison theorems between metrics of the form $g = \omega(-, J-)$, where J is an ω-compatible almost complex structure and g is complete. Thus, in what follows, g and g' are two such metrics coming from the same symplectic form ω. We however expect analogous results to hold in full generality.

Lemma A.12 *Let $W \subseteq M$ be compact. There are $a \ge 1$ and $b \ge 0$ such that every Lagrangian $L \subseteq W$ with $\|B_L\| < k$ respects $\|B'_L\|' < ak + b$, where $\|B'\|'$ is the norm in g' of the second fundamental form of L in g'.*

Proof Let A be the endomorphism of TM such that $g(AV, W) = g'(V, W)$ for all $V, W \in \mathfrak{X}(M)$. We denote by ∇ and ∇' the Levi-Civita connections of g and g', respectively. A fairly straightforward computation — see Lemma 3.1 of [22] for example — gives

$$2g'(\nabla'_V W, Z) = 2g(\nabla_V W, AZ) + g((\nabla_V A)Z, W) + g((\nabla_W A)Z, V) - g((\nabla_Z A)V, W)$$

for all $V, W, Z \in \mathfrak{X}(M)$. But it is easy to see that $A = -JJ'$. Therefore, we get that

$$\begin{aligned}|V|^2 = g'(A^{-1}V, V) \le |A^{-1}|' \cdot (|V|')^2 = |J|' \cdot (|V|')^2,\\ |AZ|^2 = |J'Z|^2 \le |J|' \cdot (|J'Z|')^2 = |J|' \cdot (|Z|')^2\end{aligned}$$

for all $V, Z \in \mathfrak{X}(M)$. Here, the prime on the norm indicates that it is the norm associated with g', not g. Finally, note also that A sends the normal bundle of a Lagrangian submanifold L in g to the normal bundle of that Lagrangian submanifold in g'.

Putting all this together, we finally get, for any Lagrangian submanifold L,

$$\|B'_L\|' \le \|J\|'^{3/2}\big(\|B_L\| + \tfrac{3}{2}\|\nabla A^{-1}\|\big), \tag{13}$$

which implies the result. The norm symbols $\|\cdot\|$ and $\|\cdot\|'$ here indicate that it is the supremum of the norms associated with g and g', respectively, over all of W or all of L. □

Lemma A.13 *Let $W \subseteq M$ be compact. There is $a \ge 1$ such that every submanifold $N \subseteq W$ which is* (*strictly*) *k^{-1}-tame in g is* (*strictly*) *$a^{-1}k^{-1}$-tame in g'.*

Proof Take $D := \frac{1}{2}\operatorname{Diam}(W, g)$, $D' = \frac{1}{2}\operatorname{Diam}(W, g')$ and $V := \overline{B_D(W)} \cup \overline{B'_{D'}(W)}$, i.e., V is the union of the (closure of the) D-neighbourhood of W in g and D'-neighbourhood of W in g'. Therefore, any path between points in W leaving V have length greater than $\operatorname{Diam}(W, g)$ in g and than $\operatorname{Diam}(W, g')$ in g'. In particular, no such path is a minimal geodesic in either metric. Since V is compact, there is $C \ge 1$ such that

$$C^{-1}g' \le g \le Cg'$$

on V.

Let $x, y \in W$, and let γ be a minimal geodesic in g' from x to y. By the above paragraph, it stays in V. Therefore,

$$d'_M(x, y) = \ell'(\gamma) \ge C^{-1}\ell(\gamma) \ge C^{-1}d_M(x, y),$$

where d_M and d'_M are the distance functions induced on M by g and g', respectively, and $\ell(\gamma)$ and $\ell'(\gamma)$ denote the lengths of γ in g and g', respectively. We analogously get that $d'_M \le Cd_M$ on W.

But note that whenever $N \subseteq W$, the same argument gives that $C^{-1}d'_N \le d_N \le d'_N$ if d_N and d'_N the distance functions induced on N by $g|_{TN}$ and $g'|_{TN}$, respectively. Therefore, we have that

$$\frac{d'_M(x, y)}{\min\{1, d'_N(x, y)\}} \ge \frac{C^{-1}d_M(x, y)}{\min\{1, Cd_N(x, y)\}} \ge C^{-2}\frac{d_M(x, y)}{\min\{1, d_N(x, y)\}}$$

for all $x \ne y \in N$, which gives the result. □

Acknowledgements

I would like to express gratitude to Octav Cornea for pushing me to write this paper. I would also like to thank both him and Paul Biran for uncountably many interesting discussions during the research process that led to this paper. Finally, I would like to thank Jean-François Barraud for some interesting exchanges on entropy and the nearby Lagrangian conjecture, and Daniil Mamaev for pointing out to me the countability assumption in the construction of the wrapped Fukaya category. Thank you also to the referee for a careful reading of the paper.

The author is partially supported by the Swiss National Science Foundation (grant 200021_204107).

References

[1] **M T K Abbassi**, **A Yampolsky**, *Transverse totally geodesic submanifolds of the tangent bundle*, Publ. Math. Debrecen 64:1-2 (2004) 129–154 MR

[2] **M Abouzaid**, *Nearby Lagrangians with vanishing Maslov class are homotopy equivalent*, Invent. Math. 189:2 (2012) 251–313 MR

[3] **R Albuquerque**, *Notes on the Sasaki metric*, Expo. Math. 37:2 (2019) 207–224 MR

[4] **M-C Arnaud**, **V Humilière**, **C Viterbo**, *Higher dimensional Birkhoff attractors* (2024) arXiv 2404.00804

[5] **D Auroux**, *Infinitely many monotone Lagrangian tori in* $\mathbb{R}^6$, Invent. Math. 201:3 (2015) 909–924 MR

[6] **H Bae**, **D Choa**, **W Jeong**, **D Karabas**, **S Lee**, *On categorical entropy from the viewpoint of symplectic topology* (2022) arXiv 2203.12205

[7] **P Biran**, **O Cornea**, **E Shelukhin**, *Lagrangian shadows and triangulated categories*, Astérisque 426, Soc. Math. France, Paris (2021) MR

[8] **P Biran**, **O Cornea**, **J Zhang**, *Triangulation*, *persistence*, *and Fukaya categories* (2023) arXiv 2304.01785

[9] **M P do Carmo**, *Riemannian geometry*, Birkhäuser, Boston, MA (1992) MR

[10] **E Çineli**, **V L Ginzburg**, **B Z Gürel**, *On the growth of the Floer barcode*, J. Mod. Dyn. 20 (2024) 275–298 MR

[11] **J-P Chassé**, *Convergence and Riemannian bounds on Lagrangian submanifolds*, International Journal of Mathematics 34:5 (2023) art. id. 2350024 MR

[12] **J-P Chassé**, *Hausdorff limits of submanifolds of symplectic and contact manifolds*, Differential Geom. Appl. 94 (2024) art. id. 102123 MR

[13] **J-P Chassé**, **R Leclercq**, *A Hölder-type inequality for the Hausdorff distance between Lagrangians*, J. Fixed Point Theory Appl. 27:2 (2025) art. id. 28 MR

[14] **Y V Chekanov**, *Invariant Finsler metrics on the space of Lagrangian embeddings*, Math. Z. 234:3 (2000) 605–619 MR

[15] **O Cornea**, **E Shelukhin**, *Lagrangian cobordism and metric invariants*, J. Differential Geom. 112:1 (2019) 1–45 MR

[16] **A Dawid**, *Floer barcode growth in the component of a Dehn–Seidel twist*, Master's thesis, ETH Zürich (2023)

[17] **G Dimitrov**, **F Haiden**, **L Katzarkov**, **M Kontsevich**, *Dynamical systems and categories*, from "The influence of Solomon Lefschetz in geometry and topology" (L Katzarkov, E Lupercio, F J Turrubiates, editors), Contemp. Math. 621, Amer. Math. Soc., Providence, RI (2014) 133–170 MR

[18] **K Fukaya**, **P Seidel**, **I Smith**, *Exact Lagrangian submanifolds in simply-connected cotangent bundles*, Invent. Math. 172:1 (2008) 1–27 MR

[19] **K Fukaya**, **P Seidel**, **I Smith**, *The symplectic geometry of cotangent bundles from a categorical viewpoint*, from "Homological mirror symmetry" (A Kapustin, M Kreuzer, K-G Schlesinger, editors), Lecture Notes in Phys. 757, Springer (2009) 1–26 MR

[20] **S Ganatra**, **J Pardon**, **V Shende**, *Covariantly functorial wrapped Floer theory on Liouville sectors*, Publ. Math. Inst. Hautes Études Sci. 131 (2020) 73–200 MR

[21] **A Gray**, *Tubes*, 2nd edition, Progress in Mathematics 221, Birkhäuser, Basel (2004) MR

[22] **Y Groman**, **J P Solomon**, *J-holomorphic curves with boundary in bounded geometry*, J. Symplectic Geom. 14:3 (2016) 767–809 MR

[23] **S Guillermou**, **N Vichery**, *Viterbo's spectral bound conjecture for homogeneous spaces* (2022) arXiv 2203.13700

[24] **J Henrikson**, *Completeness and total boundedness of the Hausdorff metric*, MIT Undergraduate Journal of Mathematics 1 (1999)

[25] **H Iriyeh**, **T Otofuji**, *Geodesics of Hofer's metric on the space of Lagrangian submanifolds*, Manuscripta Math. 122:4 (2007) 391–406 MR

[26] **A Jannaud**, *Dehn–Seidel twist, C^0 symplectic topology and barcodes* (2021) arXiv 2101.07878

[27] **M Khanevsky**, *Hofer's metric on the space of diameters*, J. Topol. Anal. 1:4 (2009) 407–416 MR

[28] **A Kislev**, **E Shelukhin**, *Bounds on spectral norms and barcodes*, Geom. Topol. 25:7 (2021) 3257–3350 MR

[29] **O Kowalski**, *Curvature of the induced Riemannian metric on the tangent bundle of a Riemannian manifold*, J. Reine Angew. Math. 250 (1971) 124–129 MR

[30] **T Kragh**, *Parametrized ring-spectra and the nearby Lagrangian conjecture*, Geom. Topol. 17:2 (2013) 639–731 MR

[31] **R Leclercq**, *Spectral invariants in Lagrangian Floer theory*, J. Mod. Dyn. 2:2 (2008) 249–286 MR

[32] **R Leclercq**, **F Zapolsky**, *Spectral invariants for monotone Lagrangians*, J. Topol. Anal. 10:3 (2018) 627–700 MR

[33] **D Milinković**, *Geodesics on the space of Lagrangian submanifolds in cotangent bundles*, Proc. Amer. Math. Soc. 129:6 (2001) 1843–1851 MR

[34] **K Ono**, *Floer–Novikov cohomology and the flux conjecture*, Geom. Funct. Anal. 16:5 (2006) 981–1020 MR

[35] **Y Ostrover**, *A comparison of Hofer's metrics on Hamiltonian diffeomorphisms and Lagrangian submanifolds*, Commun. Contemp. Math. 5:5 (2003) 803–811 MR

[36] **S Sasaki**, *On the differential geometry of tangent bundles of Riemannian manifolds*, Tohoku Math. J. (2) 10 (1958) 338–354 MR

[37] **P Seidel**, *Fukaya categories and Picard–Lefschetz theory*, EMS Zurich Lectures in Advanced Mathematics 10, European Mathematical Society (2008)

[38] **E Shelukhin**, *Symplectic cohomology and a conjecture of Viterbo*, Geom. Funct. Anal. 32:6 (2022) 1514–1543 MR

[39] **E Shelukhin**, *Viterbo conjecture for Zoll symmetric spaces*, Invent. Math. 230:1 (2022) 321–373 MR

[40] **Z M Shen**, *A convergence theorem for Riemannian submanifolds*, Trans. Amer. Math. Soc. 347:4 (1995) 1343–1350 MR

[41] **N Sheridan**, *On the Fukaya category of a Fano hypersurface in projective space*, Publ. Math. Inst. Hautes Études Sci. 124 (2016) 165–317 MR

[42] **E N Sosov**, *On Hausdorff intrinsic metric*, Lobachevskii J. Math. 8 (2001) 185–189 MR

[43] **C Viterbo**, *Symplectic topology as the geometry of generating functions*, Math. Ann. 292:4 (1992) 685–710 MR

[44] **C Viterbo**, *Inverse reduction inequalities for spectral numbers and applications* (2022) arXiv 2203.13172

[45] **Y Yomdin**, *Volume growth and entropy*, Israel J. Math. 57:3 (1987) 285–300 MR

JEAN-PHILIPPE CHASSÉ jean-philippe.chasse@umontreal.ca
CRM, Université de Montréal, C.P. 6128 Succ. Centre-Ville, Montréal, QC H3C 3J7, Canada

Received: August 1, 2024 Revised: May 16, 2025

Guidelines for Authors

Submitting a paper to Algebraic & Geometric Topology

Papers must be submitted using the upload page at the AGT 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 Algebraic & Geometric Topology

At the time of submission you need only supply a PDF file. Once accepted for publication, the paper must be supplied in LaTeX. More information on preparing articles in LaTeX for publication in AGT is available on the AGT website.

`arXiv` papers

If your paper has previously been deposited on `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.

ALGEBRAIC & GEOMETRIC TOPOLOGY

Volume 26 Issue 5 (pages 1597–1963) 2026
