Pacific Journal of Mathematics

Volume 309 No. 2 December 2020

PACIFIC JOURNAL OF MATHEMATICS

Founded in 1951 by E. F. Beckenbach (1906–1982) and F. Wolf (1904–1989)

msp.org/pjm

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

The subscription price for 2020 is US $520/year for the electronic version, and $705/year for print and electronic.
Subscriptions, requests for back issues and changes of subscriber address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

The Pacific Journal of Mathematics (ISSN 1945-5844 electronic, 0030-8730 printed) at the University of California, c/o Department of Mathematics, 798 Evans Hall #3840, Berkeley, CA 94720-3840, is published twelve times a year. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PJM peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 2, 2020
https://doi.org/10.2140/pjm.2020.309.257

THIN SUBGROUPS ISOMORPHIC TO GROMOV–PIATETSKI-SHAPIRO LATTICES

SAMUEL A. BALLAS

In this paper we produce many examples of thin subgroups of special linear groups that are isomorphic to the fundamental group nonarithmetic hyperbolic manifolds. Specifically, we show that the nonarithmetic lattices in SO(n, 1, $\mathbb{R}$) constructed by Gromov and Piatetski-Shapiro can be embedded into SL($n+1$, $\mathbb{R}$) so that their images are thin subgroups.

Introduction

Let G be a semisimple Lie group and let Γ be a finitely generated subgroup. We say that Γ is a *thin subgroup of* G if there is a lattice $\Lambda \subset G$ containing Γ such that

- Γ has infinite index in Λ,
- Γ is Zariski dense in G.

Intuitively, such groups are very sparse in the sense that they have infinite index in a lattice, but at the same time are dense in an algebraic sense. Note, that if one relaxes the first condition above, then Γ would be a lattice, so another way of thinking of thin groups is as infinite index analogues of lattices in semisimple Lie groups.

Over the last several years, thin groups have been the subject of much research, much of which has been motivated by the observation that many theorems and conjectures in number theory can be phrased in terms of counting primes in orbits of groups that are "abelian analogues of thin groups." Here are two examples. First, let $G=\mathbb{R},\ b, m\in\mathbb{N}$ such that $(b,m)=1,\ \Delta=\mathbb{Z}$ and $\Gamma=m\mathbb{Z}$. The orbit $b+\Gamma$ is an arithmetic progression and Dirichlet's theorem on primes in arithmetic progressions is equivalent to this orbit containing infinitely many primes. Next, let $G=\mathbb{R}^2$, $\Delta=\mathbb{Z}^2,\ \Gamma=\langle(1,1)\rangle$ and $b=(1,3)\in\mathbb{Z}^2$. The orbit $b+\Gamma=\{(m,m+2)\mid m\in\mathbb{Z}\}$ and the twin prime conjecture is equivalent to the statement that this orbit contains infinitely many points whose components are both prime. Note that in the first case Γ is a lattice in G, but in the second case Γ has infinite index in Δ and is an analogue of a thin group (sans Zariski density) in G.

MSC2010: 22E40, 57M50.
Keywords: thin groups, nonarithmetic lattices.

This orbital perspective was used by Brun to attack the twin primes conjecture using "combinatorial sieving" techniques. Although the full conjecture remains unproven these techniques did yield some powerful results. For instance, using these methods, Chen [1978] was able to prove that there are infinitely many pairs n and $n+2$ such that one is prime and the other is the product of at most 2 primes. More details of this perspective are explained in the excellent surveys of Bourgain [2014] and Lubotzky [2012].

Inspired by these results, Bourgain, Gamburd, and Sarnak [Bourgain et al. 2010] developed complementary "affine sieving" techniques to analyze thin group orbits. In this context, the thinness property of the group gives enough control of orbits to execute these counting arguments. Again, much of this is described in Lubotzky's survey [2012].

Given these connections it is desirable to produce examples of thin groups and understand what types of groups are thin. Presently, there are many constructions of thin groups. For instance, in recent work of Fuchs and Rivin [2017] it is shown that if one "randomly" selects two matrices in $\mathrm{SL}(n, \mathbb{Z})$ then with high probability, the group they generate is a thin subgroup of $\mathrm{SL}(n, \mathbb{R})$. However, the groups constructed in this way are almost always free groups. There are also several constructions that allow one to produce thin subgroups isomorphic to fundamental groups of closed surfaces in a variety of algebraic groups (see [Cooper and Futer 2019; Kahn et al. 2018; Kahn and Markovic 2012; Kahn and Wright 2018], for instance). Given these examples one may ask which isomorphism classes of groups are thin? More precisely, if G is a semisimple Lie group and H is an abstract finitely generated group then we say that H can be *realized as a thin subgroup of* G if there is an embedding $\iota : H \to G$ whose image is a thin subgroup of G. With this definition in hand we can rephrase the previous question as: given a semisimple algebraic group G, what isomorphism types of groups can be realized as thin subgroups of G? Recent work of the author and D. Long [Ballas and Long 2020] shows that there are many additional isomorphism types of groups that can arise as thin subgroups of special linear groups. More precisely, in [Ballas and Long 2020] it is shown that fundamental groups of arithmetic hyperbolic n-manifolds of "orthogonal type" can be realized as thin subgroups. In the present work, we extend the techniques of [Ballas and Long 2020] to produce infinitely many examples of nonarithmetic hyperbolic n-manifolds whose fundamental groups can be realized as thin subgroups of $\mathrm{SL}_{n+1}(\mathbb{R})$. Our main result is:

Theorem 1. *For each $n \geq 3$, there is an infinite collection $\mathcal{C}_n$ of nonarithmetic hyperbolic n-manifolds with the property that if $M^n \in \mathcal{C}_n$ then $\pi_1(M)$ can be realized as a thin subgroup of $\mathrm{SL}_{n+1}(\mathbb{R})$. Furthermore, the collection $\mathcal{C}_n$ contains representatives from infinitely many commensurability classes of both compact and noncompact manifolds.*

It should be noted that the collection $\mathcal{C}_n$ appearing in Theorem 1 can be described fairly explicitly, and roughly speaking consists of the hyperbolic manifolds coming from the nonarithmetic lattices in $\mathrm{SO}(n, 1, \mathbb{R})$ constructed by Gromov–Piatetski-Shapiro in [Gromov and Piatetski-Shapiro 1988].

Outline of paper. In Section 1 we recall the Gromov–Piatetski-Shapiro construction of nonarithmetic lattices in $\mathrm{SO}(n, 1, \mathbb{R})$ and define the collection $\mathcal{C}_n$ appearing in Theorem 1. In Section 2 we show that the fundamental group of any element of $\mathcal{C}_n$ can be embedded in several lattices in $\mathrm{SL}_{n+1}(\mathbb{R})$. Finally, in Section 3 we prove Theorem 1 by showing that the images of the previously mentioned embeddings are thin subgroups.

1. Gromov–Piatetski-Shapiro lattices

Gromov and Piatetski-Shapiro [1988], describe a method for constructing infinitely many nonarithmetic lattices in $\mathrm{SO}(n, 1, \mathbb{R})$. In this section we describe their construction and the construction of the lattices appearing in Theorem 1.

Let K be a totally real number field of degree $d+1$ with ring of integers $\mathcal{O}_K$. There are $d+1$ embeddings $\{\sigma_0, \ldots, \sigma_d\}$ of K into $\mathbb{R}$. Using the embedding σ_0 we will implicitly regard K as a subset of $\mathbb{R}$. In this way, it makes sense to say that elements of F are positive or negative. Let $s_K : K^\times \to \mathbb{Z}_{\geq 0}$, where $s_K(a) = |\{i \geq 1 \mid \sigma_i(a) > 0\}|$. In other words, $s_K(a)$ counts the nonidentity embeddings for which a has positive image.

Next, let $\alpha, \beta, a_2, \ldots, a_{n+1} \in \mathcal{O}_K$ be positive elements such that

- β/α is not a square in K,
- $s_K(\alpha) = s_K(\beta) = s_K(\alpha_i) = d$ for $1 \leq i \leq n$,
- $s_K(a_{n+1}) = 0$.

Next, define quadratic forms

$$J_1 = \alpha x_1^2 + \sum_{i=2}^{n} a_i x_i^2 - a_{n+1} x_{n+1}^2, \qquad J_2 = \beta x_1^2 + \sum_{i=2}^{n} a_i x_i^2 - a_{n+1} x_{n+1}^2 \tag{1-1}$$

If $A \subset \mathbb{R}$ is a subring containing 1 then we define

$$\mathrm{SO}(J_i, A) = \left\{ B \in \mathrm{SL}_{n+1}(A) \mid J_i(Bv) = J_i(v)\ \forall v \in \mathbb{R}^{n+1} \right\}.$$

Using this notation, define $\Gamma_1 = \mathrm{SO}(J_1, \mathcal{O}_K)$ and $\Gamma_2 = h\,\mathrm{SO}(J_2, \mathcal{O}_K)h^{-1}$, where $h = \mathrm{Diag}(\sqrt{\beta/\alpha}, \ldots, 1)$. Note that both Γ_1 and Γ_2 are lattices in $\mathrm{SO}(J_1, \mathbb{R})$, however, since β/α is not a square in K it follows from [Gromov and Piatetski-Shapiro 1988, see Corollary 2.7 and §2.9] that these lattices are not commensurable.

There is a model for hyperbolic n-space given by

$$\mathbb{H}^n = \left\{ v \in \mathbb{R}^{n+1} \mid J_1(v) = -1,\ v_{n+1} > 0 \right\}.$$

The identity component $\mathrm{SO}(J_1, \mathbb{R})^\circ$ of $\mathrm{SO}(J_1, \mathbb{R})$ consists of the orientation preserving isometries of $\mathbb{H}^n$ (see [Ratcliffe 2006, §3.2] for details). By passing to finite index subgroups we can assume that $\Gamma_i \subset \mathrm{SO}(J_1, \mathbb{R})^\circ$, and so $\mathbb{H}^n/\Gamma_i$ is a finite volume hyperbolic orbifold for $i = 1, 2$.

The lattice $\Gamma_2 \subset \mathrm{SO}(J_1, L)$, where $L = K(\sqrt{\beta/\alpha})$. Note that because α and β are positive and $s_K(\alpha) = s_K(\beta) = d$ it follows that L is also totally real. Furthermore, for every $\gamma \in \Gamma_2$, $\mathrm{tr}(\gamma) \in \mathcal{O}_K \subset \mathcal{O}_L$. The following lemma then shows that by passing to a subgroup of finite index we may assume that $\Gamma_2 \subset \mathrm{SO}(J_1, \mathcal{O}_L)$. This result seems well known to experts, but we include a proof for the sake of completeness.

Lemma 1.1. *Let $k \subset \mathbb{C}$ be a number field and let $\mathcal{O}_k$ be the ring of integers of k. If $\Gamma \subset \mathrm{GL}_n(k)$ acts irreducibly on $\mathbb{C}^n$ and has the property that $\mathrm{tr}(\gamma) \in \mathcal{O}_k$ for each $\gamma \in \Gamma$ then there is a finite index subgroup $\Gamma' \subset \Gamma$ such that $\Gamma' \subset \mathrm{GL}_n(\mathcal{O}_k)$.*

Proof. If $A \subset k$ is a subring then let $A\Gamma = \left\{ \sum_i a_i \gamma_i \mid a_i \in A, \gamma_i \in \Gamma \right\}$. Note that in this definition all sums have finitely many terms. By [Bass 1980, Proposition 2.2], $\mathcal{O}_k\Gamma$ is an order in the central simple algebra $k\Gamma$. The order $\mathcal{O}_k\Gamma$ is contained in some maximal order $\mathcal{D}$ in $M_n(k)$ ($n \times n$ matrices over k). Let $\mathcal{D}^1 \subset \mathrm{SL}_n(k)$ be the norm 1 elements of $\mathcal{D}$. Then $M_n(\mathcal{O}_k)$ is also an order in $M_n(k)$ whose group of norm 1 elements is $\mathrm{SL}_n(\mathcal{O}_k)$. It is a standard result using restriction of scalars that groups of norm 1 elements in maximal orders of $M_n(k)$ are commensurable. Roughly speaking this is a consequence of the fact that the intersection of two orders is again an order and the unit groups of these orders are irreducible lattices in $\mathrm{SL}_n(\mathbb{R}) \times \mathrm{SL}_n(\mathbb{R})$ (see [Morris 2015, §5.1 and Example 5.1 #7]). It follows that $\mathcal{D}^1 \cap \mathrm{SL}_n(\mathcal{O}_k)$ has finite index in $\mathcal{D}^1$ and so $\Gamma \cap \mathrm{SL}_n(\mathcal{O}_k)$ has finite index in Γ. □

Note that since Γ_2 is a lattice in $\mathrm{SO}(J_1, \mathbb{R})$ it acts irreducibly on $\mathbb{C}^{n+1}$, and so by applying Lemma 1.1 we may assume that $\Gamma_2 \subset \mathrm{SO}(J_1, \mathcal{O}_L)$.

Denote by $\mathrm{SO}(n-1, 1, \mathbb{R})$ the subgroup of $\mathrm{SO}(J_1, \mathbb{R})$ that preserves both complementary components in $\mathbb{R}^{n+1}$ of the hyperplane P given by the equation $x_1 = 0$. The intersection $P \cap \mathbb{H}^n$ is a model for hyperbolic $(n-1)$-space, $\mathbb{H}^{n-1}$ and the group $\mathrm{SO}(n-1, 1, \mathbb{R})$ can be identified with the subgroup of orientation preserving isometries of $\mathbb{H}^{n-1}$. Next, let $\hat{\Gamma} = \Gamma_1 \cap \Gamma_2 \cap \mathrm{SO}(n-1, 1, \mathbb{R})$. Since each $\Gamma_i \cap \mathrm{SO}(n-1, 1, \mathbb{R})$ is sublattice of the lattice $\mathrm{SO}(n-1, 1, \mathcal{O}_L)$ in $\mathrm{SO}(n-1, 1, \mathbb{R})$, it follows that $\hat{\Gamma}$ is also a lattice in $\mathrm{SO}(n-1, 1, \mathbb{R})$. It follows that $\mathbb{H}^{n-1}/\hat{\Gamma}$ is a hyperbolic $(n-1)$-orbifold. By passing to finite index subgroups we may arrange the following properties:

(1) Γ_i is torsion-free and contained in the identity component of $\mathrm{SO}(J_1, \mathbb{R})$. This component is isomorphic to $\mathrm{Isom}^+(\mathbb{H}^n)$, and so $M_i := \mathbb{H}^n/\Gamma_i$ is a finite volume

hyperbolic manifold (apply Selberg's lemma and the fact that $\mathrm{SO}(J,\mathbb{R})^\circ$ has finite index in $\mathrm{SO}(J,\mathbb{R})$).

(2) Since $\Sigma = \mathbb{H}^{n-1}/\hat{\Gamma}$ is a totally geodesic we may assume that Σ is a hyperbolic $(n-1)$-manifold and this manifold is embedded in both M_1 and M_2 (see [Bergeron 2000, Theorem 1]).

(3) If M_i is noncompact then all cusps of M_i are diffeomorphic to an $(n-1)$-torus times an interval (apply [McReynolds et al. 2013, Theorem 3.1])

(4) The complement $\widehat{M}_i = M_i \backslash \Sigma$ is connected for $i = 1, 2$ (see [Bergeron 2000, Theorem 2]).

The manifold $\widehat{M}_i$ is a convex submanifold of M_i and so $\widehat{M}_i = V_i/\hat{\Gamma}_i$, where V_i is a component of the preimage of $\widehat{M}_i$ in $\mathbb{H}^n$ under the universal covering projection $\mathbb{H}^n \to \mathbb{H}^n/\Gamma_i = M_i$, and $\hat{\Gamma}_i$ is a subgroup of Γ_i that stabilizes V_i. The manifold $\widehat{M}_i$ is a hyperbolic manifold with totally geodesic boundary equal to two isometric copies of Σ, and so it is possible to glue $\widehat{M}_1$ and $\widehat{M}_1$ along Σ to form the finite volume hyperbolic manifold N (see [Morris 2015, §6.5] for details). The manifold N can be realized as $\mathbb{H}^n/\Delta$ where, after appropriately conjugating $\hat{\Gamma}_i$ in Γ_i, we may assume that

$$\Delta = \langle \hat{\Gamma}_1, \hat{\Gamma}_2, s\rangle. \tag{1-2}$$

Here s comes from a "graph of spaces" description of N and can thus be written as a product $s = s_2 s_1$, where s_i is the isometry corresponding to an appropriate lift to V_i a curve in M_i whose algebraic intersection with Σ is 1 (See Figure 1). In [Gromov and Piatetski-Shapiro 1988, §2.9] it is shown that Δ is a nonarithmetic lattice in $\mathrm{SO}(J_1,\mathbb{R})$. If $N = \mathbb{H}^n/\Delta$ then we call N an *interbreeding of* M_1 *and* M_2.

Since $\Gamma_1, \Gamma_2 \subset \mathrm{SO}(J_1, \mathcal{O}_L)$ it follows that $\Delta \subset \mathrm{SO}(J_1, \mathcal{O}_L)$. As a result, we call the field L the *field of definition of* Δ. Let $\mathcal{C}_n$ be the collection of hyperbolic n-manifolds coming from the above interbreeding construction.

We close this section by proving the following result:

Proposition 1.2. *The collection $\mathcal{C}_n$ contains representatives of infinitely many commensurability classes of both closed and noncompact hyperbolic n-manifolds satisfying the properties* (1)–(4) *from above.*

To prove this we will need the following invariant, originally due to Vinberg [1971]. Let Γ be a Zariski dense subgroup of a Lie group H with Lie algebra $\mathfrak{h}$. The adjoint action of Γ on $\mathfrak{h}$ gives a representation $\mathsf{Ad} : \Gamma \to \mathfrak{gl}(\mathfrak{h})$. In [Vinberg 1971] it is shown that the field $\mathbb{Q}(\{\mathrm{tr}(\mathsf{Ad}(\gamma)) \mid \gamma \in \Gamma\})$ is an invariant of the commensurability class of Γ in H. This field is called the *adjoint trace field of* Γ.

Next, let $N = \mathbb{H}^n/\Delta \in \mathcal{C}_n$, then Δ is a lattice in $\mathrm{SO}(J_1,\mathbb{R})$, which is Zariski dense by the Borel density theorem. The following lemma allows us to compute

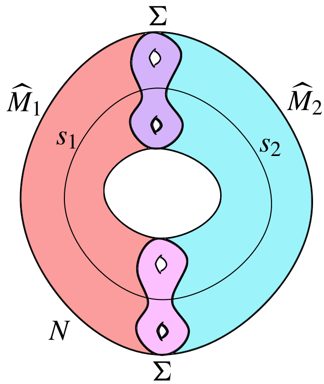


Figure 1. An graph of spaces description of the manifold N.

the adjoint trace field of Δ. It is an immediate corollary of a theorem of Mila (see [Mila 2019, Theorem 4.7]) once it is observed that L is the smallest extension of K over which the forms J_1 and J_2 are isometric.

Lemma 1.3. *Let $N = \mathbb{H}^n/\Delta \in \mathcal{C}_n$ and let L be the field of definition of Δ. Then L is the adjoint trace field of Δ.*

Proof of Proposition 1.2. From [Gromov and Piatetski-Shapiro 1988], it follows that $N = \mathbb{H}^n/\Delta$ is compact if and only if the field K used to construct Δ is not equal to $\mathbb{Q}$. For each choice of a totally real field K and a pair $\alpha, \beta \in K$ so that α/β is not a square in K we can produce an element $N \in \mathcal{C}_n$ via the interbreeding construction. By varying the choices of α and β we can produce infinitely many distinct $L = K(\sqrt{\beta/\alpha})$ for each choice of K. It follows from Lemma 1.3 that the corresponding N are representatives of infinitely many commensurability classes of both compact and noncompact hyperbolic n-manifolds. □

2. Lattices in $\mathrm{SL}_{n+1}(\mathbb{R})$

In this section we describe the lattices $\Delta \subset \mathrm{SL}_{n+1}(\mathbb{R})$ in which our thin groups will ultimately live. Let J_1 be one of the forms constructed in Section 1 and let L be the corresponding (totally real) field of definition. Let $M = L(\sqrt{r})$, where $r \in L$ is positive, square-free, and $s_L(r) = 0$. The number field M is a quadratic extension of L and we let $\tau : M \to M$ be the unique nontrivial Galois automorphism of M over L. In this context, we can extend the quadratic form J_1 on L^{n+1} to a "Hermitian" form on M^{n+1}. Let $N_{M/L} : M \to L$ given by $N_{M/L}(x) = x\tau(x)$ be the norm of the field extension M/L. Next let $x = (x_1, \dots, x_{n+1}) \in M^{n+1}$ and define $H_1 : M^{n+1} \to L$ as

$$H_1(x) = \alpha N_{M/L}(x_1) + \sum_{i=1}^{n} a_i N_{M/L}(x_i) - a_{n+1} N_{M/L}(x_{n+1}).$$

Note that this defines a Hermitian form in the sense that if $x \in M^{n+1}$ and $\lambda \in M$ then $H_1(\lambda x) = N_{M/L}(\lambda) H_1(x)$. Furthermore, since L is the fixed field of τ it follows that H_1 reduces to J_1 when restricted to L^{n+1}.

Next, we can define a unitary analogue of $\mathrm{SO}(J_1, \mathcal{O}_M)$ as

$$\mathrm{SU}(J_1, \tau, \mathcal{O}_M) = \{ A \in \mathrm{SL}_{n+1}(\mathcal{O}_M) \mid H_1(Av) = H_1(v)\ \forall v \in M^{n+1} \}.$$

It is well known (see [Morris 2015, §6.8], for example) that $\mathrm{SU}(J_1, \tau, \mathcal{O}_M)$ is an arithmetic lattice in $\mathrm{SL}_{n+1}(\mathbb{R})$.

Let $N = \mathbb{H}^n/\Delta$ be one of the manifolds from $\mathcal{C}_n$. By construction, the manifold N contains the embedded totally geodesic hypersurface $\Sigma = \mathbb{H}^{n-1}/\hat{\Gamma}$, and so it is possible to deform Δ inside of $\mathrm{SL}_{n+1}(\mathbb{R})$ using the bending construction of Johnson and Millson [1987].

Specifically, let $c_t = \mathrm{Diag}(e^{-nt}, e^t, \dots, e^t) \in \mathrm{SL}_{n+1}(\mathbb{R})$. It is easy to check that c_t centralizes $\mathrm{SO}(n-1, 1, \mathbb{R})$. Since Σ is assumed to be nonseparating, we see that write Δ as an HNN extension $\Delta \cong \hat{\Delta}*_s$, where $\hat{\Delta}$ is isomorphic to the fundamental group of $N \backslash \Sigma$ and s is a free letter. In this context, we may view $\hat{\Delta} \subset \mathrm{SO}(J_1, \mathcal{O}_L)$ and $s \in \mathrm{SO}(J_1, \mathcal{O}_L)$ and observe that as a subgroup of $\mathrm{SO}(J_1, \mathcal{O}_L)$ we can write $\Delta = \langle \hat{\Delta}, s \rangle$. We now define a new family of subgroups $\Delta_t = \langle \Delta, c_t s \rangle \subset \mathrm{SL}_{n+1}(\mathbb{R})$. Using basic theory of HNN extensions, it is easy to see that, since c_t centralizes the fundamental group of Σ, as an abstract group Δ_t is a quotient of Δ. However, by using the following result due to Benoist [2005] in the compact case and Marquis [2012] in the noncompact case, we can actually say much more.

Proposition 2.1. *For each t, the group Δ_t is isomorphic to Δ.*

Next, we show for certain values of t the group Δ_t is contained in one of the unitary lattices constructed above. Specifically, if $N = \mathbb{H}^n/\Delta$ is contained in $\mathcal{C}_n$, let J_1 and L be such that $\Delta \subset \mathrm{SO}(J_1, \mathcal{O}_L)$. Recall that the field L is totally real of degree $d+1$ over $\mathbb{Q}$ and so there are $d+1$ embeddings $\{\sigma_0 = \mathrm{Id}, \dots, \sigma_d\}$ of L into $\mathbb{R}$. We can use Lemma 3.1 of [Ballas and Long 2020] to produce a unit $u \in \mathcal{O}_L^\times$ with the property that $|u| > 2$ and $0 < |\sigma_i(u)| < 1$ for $1 \le i \le d$. Let $p(x) = x^2 - ux + 1$ and let $M = L(v)$, where v is one of the roots of $p(x)$. It is easy to check that the discriminant of $p(x)$ is $u^2 - 4$ and so $M = L(\sqrt{u^2 - 4})$. By construction $s_L(u^2 - 4) = 0$, and so $\mathrm{SU}(J_1, \tau, \mathcal{O}_M)$ is an arithmetic lattice in $\mathrm{SL}_{n+1}(\mathbb{R})$, where $\tau : M \to M$ is the nontrivial Galois automorphism of M over L. The next lemma says that by carefully choosing t, we can arrange that $\Delta_t \subset \mathrm{SU}(J_1, \tau, \mathcal{O}_M)$.

Lemma 2.2. *Let u be as above. Then if $t = \log(u)$ then $\Delta_t \subset \mathrm{SU}(J_1, \tau, \mathcal{O}_M)$.*

This is basically Lemma 3.4 of [Ballas and Long 2020], but the proof is short so we include it here for the sake of completeness.

Proof. Recall from above that there is a subgroup $\hat{\Delta} \subset \mathrm{SO}(J_1, \mathcal{O}_L)$ and $s \in \mathrm{SO}(J_1, \mathcal{O}_L)$ so that $\Delta = \langle \hat{\Delta}, s \rangle$ and $\Delta_t = \langle \hat{\Delta}, c_t s \rangle$, where

$$c_t = \mathrm{Diag}(e^{-nt}, e^t, \ldots, e^t) \in \mathrm{SL}_{n+1}(\mathbb{R}).$$

Since $\mathrm{SO}(J_1, \mathcal{O}_L) \subset \mathrm{SU}(J_1, \tau, \mathcal{O}_N)$ the proof will be complete if we can show that $c_t \in \mathrm{SU}(J_1, \tau, \mathcal{O}_M)$.

If $t = \log(u)$ then $c_t = \mathrm{Diag}(u^{-n}, u, \ldots, u)$. Furthermore, since $\tau(u)$ is the other root of $p(x)$ it follows that $u\tau(u) = 1$, or in other words $\tau(u) = u^{-1}$. It follows that $c_t^* = \mathrm{Diag}(u^n, u^{-1}, \ldots, u^{-1})$. A simple computation then shows that for each $v \in M^{n+1}$, $H_1(c_t v) = H_1(v)$, and so $c_t \in \mathrm{SU}(J_1, \tau, \mathcal{O}_M)$. □

By combining Lemma 2.2 and Proposition 2.1 we get the following corollary:

Corollary 2.3. *For each $N = \mathbb{H}^n/\Delta \in \mathcal{C}_n$ there are infinitely many lattices $\Lambda \subset \mathrm{SL}_{n+1}(\mathbb{R})$ that contain a subgroup Δ' isomorphic to Δ.*

3. Certifying thinness

The main goal of this section is to complete the proof of Theorem 1. The proof consist of proving that the subgroups constructed in the previous section are thin.

Proof of Theorem 1. Recall, that if $N = \mathbb{H}^n/\Delta \in \mathcal{C}_n$ from Corollary 2.3 it follows that we can find a lattice $\Lambda \subset \mathrm{SL}_{n+1}(\mathbb{R})$ and a subgroup $\Delta' \subset \Lambda$ that is isomorphic to Δ.

Since Δ' was obtained from Δ via a bending construction if follows from [Ballas and Long 2020, Proposition 4.1] that Δ' is Zariski dense in $\mathrm{SL}_{n+1}(\mathbb{R})$. The proof will be complete if we can show that Δ' has infinite index in Λ. Suppose for contradiction that this index is finite. Since Λ is a lattice in $\mathrm{SL}_{n+1}(\mathbb{R})$ this implies that Δ' is also a lattice in $\mathrm{SL}_{n+1}(\mathbb{R})$. However, Δ' is isomorphic to Δ and Δ is a lattice in the Lie group $\mathrm{SO}(n, 1)^\circ$. However, $\mathrm{SO}(n, 1)^\circ$ and $\mathrm{SL}_{n+1}(\mathbb{R})$ are not isomorphic and so this contradicts the Mostow rigidity theorem (see [Morris 2015, Theorem 15.1.2]). □

Acknowledgements

The author would like to thank Darren Long for several helpful conversations during the preparation of this work and Matt Stover for providing references that greatly simplified the proof of Lemma 1.3. The author was partially supported by the NSF grant DMS-1709097.

References

[Ballas and Long 2020] S. A. Ballas and D. D. Long, "Constructing thin subgroups of $\mathrm{SL}(n+1, \mathbb{R})$ via bending", *Algebr. Geom. Topol.* **20**:4 (2020), 2071–2093. MR Zbl

[Bass 1980] H. Bass, "Groups of integral representation type", *Pacific J. Math.* **86**:1 (1980), 15–51. MR Zbl

[Benoist 2005] Y. Benoist, "Convexes divisibles, III", *Ann. Sci. École Norm. Sup.* (4) **38**:5 (2005), 793–832. MR Zbl

[Bergeron 2000] N. Bergeron, "Premier nombre de Betti et spectre du laplacien de certaines variétés hyperboliques", *Enseign. Math.* (2) **46**:1-2 (2000), 109–137. MR Zbl

[Bourgain 2014] J. Bourgain, "Some Diophantine applications of the theory of group expansion", pp. 1–22 in *Thin groups and superstrong approximation*, edited by E. Breuillard and H. Oh, Math. Sci. Res. Inst. Publ. **61**, Cambridge Univ. Press, 2014. MR Zbl

[Bourgain et al. 2010] J. Bourgain, A. Gamburd, and P. Sarnak, "Affine linear sieve, expanders, and sum-product", *Invent. Math.* **179**:3 (2010), 559–644. MR Zbl

[Chen 1978] J. R. Chen, "On the representation of a large even integer as the sum of a prime and the product of at most two primes, II", *Sci. Sinica* **21**:4 (1978), 421–430. MR

[Cooper and Futer 2019] D. Cooper and D. Futer, "Ubiquitous quasi-Fuchsian surfaces in cusped hyperbolic 3-manifolds", *Geom. Topol.* **23**:1 (2019), 241–298. MR Zbl

[Fuchs and Rivin 2017] E. Fuchs and I. Rivin, "Generic thinness in finitely generated subgroups of $\mathrm{SL}_n(\mathbb{Z})$", *Int. Math. Res. Not.* **2017**:17 (2017), 5385–5414. MR Zbl

[Gromov and Piatetski-Shapiro 1988] M. Gromov and I. Piatetski-Shapiro, "Nonarithmetic groups in Lobachevsky spaces", *Inst. Hautes Études Sci. Publ. Math.* 66 (1988), 93–103. MR

[Johnson and Millson 1987] D. Johnson and J. J. Millson, "Deformation spaces associated to compact hyperbolic manifolds", pp. 48–106 in *Discrete groups in geometry and analysis* (New Haven, CT, 1984), edited by R. Howe, Progr. Math. **67**, Birkhäuser, Boston, 1987. MR Zbl

[Kahn and Markovic 2012] J. Kahn and V. Markovic, "Immersing almost geodesic surfaces in a closed hyperbolic three manifold", *Ann. of Math.* (2) **175**:3 (2012), 1127–1190. MR Zbl

[Kahn and Wright 2018] J. Kahn and A. Wright, "Nearly Fuchsian surface subgroups of finite covolume Kleinian groups", preprint, 2018. arXiv

[Kahn et al. 2018] J. Kahn, F. Labourie, and S. Mozes, "Surface subgroups in uniform lattices of some semi-simple Lie groups", preprint, 2018, available at http://flab.perso.math.cnrs.fr/preprints/SGL.pdf.

[Lubotzky 2012] A. Lubotzky, "Expander graphs in pure and applied mathematics", *Bull. Amer. Math. Soc.* (*N.S.*) **49**:1 (2012), 113–162. MR Zbl

[Marquis 2012] L. Marquis, "Exemples de variétés projectives strictement convexes de volume fini en dimension quelconque", *Enseign. Math.* (2) **58**:1-2 (2012), 3–47. MR Zbl

[McReynolds et al. 2013] D. B. McReynolds, A. W. Reid, and M. Stover, "Collisions at infinity in hyperbolic manifolds", *Math. Proc. Cambridge Philos. Soc.* **155**:3 (2013), 459–463. MR Zbl

[Mila 2019] O. Mila, *The trace field of hyperbolic gluings*, Ph.D. thesis, Universität Bern, 2019.

[Morris 2015] D. W. Morris, *Introduction to arithmetic groups*, Deductive Press, 2015. MR Zbl

[Ratcliffe 2006] J. G. Ratcliffe, *Foundations of hyperbolic manifolds*, 2nd ed., Graduate Texts in Mathematics **149**, Springer, 2006. MR Zbl

[Vinberg 1971] E. B. Vinberg, "Rings of definition of dense subgroups of semisimple linear groups", *Izv. Akad. Nauk SSSR Ser. Mat.* **35** (1971), 45–55. In Russian; translated in *Math. USSR-Izv.* **5**:1 (1971), 45–55. MR Zbl

Received November 18, 2019. Revised November 23, 2020.

SAMUEL A. BALLAS
DEPARTMENT OF MATHEMATICS
FLORIDA STATE UNIVERSITY
TALLAHASSEE, FL
UNITED STATES

ballas@math.fsu.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 2, 2020
https://doi.org/10.2140/pjm.2020.309.267

VALUE DISTRIBUTION PROPERTIES FOR THE GAUSS MAPS OF THE IMMERSED HARMONIC SURFACES

XINGDI CHEN, ZHIXUE LIU AND MIN RU

We study the value distribution theory for the immersed harmonic surfaces and K-QC harmonic surfaces. We first investigate the value distribution properties for the generalized Gauss map Φ of an immersed harmonic surface, similar to the result of Fujimoto and Ru in the minimal surfaces case. After building a relation between Φ and the classical Gauss map $\mathfrak{n}$ for the K-QC harmonic surfaces, we derive that, for a complete harmonic and K-quasiconformal surface immersed in $\mathbb{R}^3$, if its unit normal $\mathfrak{n}$ omits seven directions in S^2 and any three of which are not contained in a plane in $\mathbb{R}^3$, then the surface must be flat. In the last section, under an additional condition, we give an estimate of the Gauss curvature for the K-QC harmonic surfaces, generalizing the result of the minimal surfaces in the case that the unit normal $\mathfrak{n}$ omits a neighborhood of some fixed direction.

1. Introduction

Since R. Osserman and S. S. Chern [Chern 1965; Chern and Osserman 1967; Osserman 1964] initiated the study of the value distribution properties for the Gauss map of complete minimal surfaces immersed in $\mathbb{R}^n$, it has grown into a very rich theory due to the works of F. Xavier [1981], H. Fujimoto [1993], Osserman and Ru [1997] and M. Ru [1991] etc. On the other hand, as early as the late 60s, T. K. Milnor [1967; 1968] started to consider whether the theory carries over in an interesting way to the larger class of harmonically immersed surfaces. Many similar results have been obtained (see [Alarcón and López 2013; Connor et al. 2015; 2018; Dioos and Sakaki 2019; Jensen and Rigoli 1988; Kalaj 2013; Milnor 1979; 1980; 1983]). In particular, it was first observed by Milnor [1983] that, instead of the Gauss map $\mathfrak{n}$, the map Φ (in this paper, we call it the *generalized* Gauss map; see the next section for the definition) carries the same value distribution properties as the Gauss map $\mathfrak{n}$ in the minimal surface case. The observation is that if the induced metric ds^2 on M (from the standard metric in $\mathbb{R}^3$) is complete, then

MSC2020: primary 53C42, 53C43; secondary 30C65, 32H25.

Keywords: harmonic immersion, quasiconformal mapping, value distribution theory, Hopf differential, conformal metric, Gauss map.

the conformal metric $\|\Phi\|^2$ (which is called *the associated Klotz metric*) is also complete, so it allows us to study the value distribution properties for Φ. In this paper, we first study the value distribution properties for the map Φ. In particular, we obtain a result which is similar to Fujimoto and Ru's result in the minimal surface case (see Theorem 3.3). We then use the results we obtained to study the value distribution property for the Gauss map $\mathfrak{n}$ of harmonically immersed surfaces by comparing $\mathfrak{n}$ with Φ (see Theorem 4.7). In the last section, under an additional condition, we give an estimate of the Gauss curvature for the K-QC harmonic surfaces, generalizing the result of the minimal surfaces in the case that the unit normal $\mathfrak{n}$ omits a neighborhood of some fixed direction.

2. Immersed harmonic surfaces

We study the maps $X : M \to \mathbb{R}^n$, with $n \geq 3$, where M is a complex Riemann surface, and X is a regular and immersed map. The surface X is called *an immersed harmonic surface* if X is harmonic. Under a local coordinate $z = u + \sqrt{-1}v$ for the Riemann surface M, it is well known that X is harmonic if and only if

$$\triangle X = 4\partial^2 X/\partial z \partial \bar{z} \equiv 0$$

where $\frac{\partial}{\partial z} = \frac{1}{2}(\partial/\partial u - \sqrt{-1}\partial/\partial v)$ and $\frac{\partial}{\partial \bar{z}} = \frac{1}{2}(\partial/\partial u + \sqrt{-1}\partial/\partial v)$. Thus X is harmonic if and only if

$$\phi := \frac{\partial X}{\partial z} = (\phi_1, \ldots, \phi_n) \tag{1}$$

is holomorphic, where $\phi_i = \frac{\partial}{\partial z} X^i$ for $i = 1, \ldots, n$ when we write $X = (X^1, \ldots, X^n)$. Note that although ϕ_i are only locally defined, the holomorphic one-forms $\Phi_i := \phi_i dz$ are globally defined on M. Thus the map $\Phi := [\Phi_1 : \cdots : \Phi_n] : M \to \mathbb{P}^{n-1}(\mathbb{C})$ is well-defined and holomorphic. We call it the *generalized Gauss map* of the harmonic surface X.

Let ds^2 be the metric on M induced through X from the standard inner product on $\mathbb{R}^n$. In terms of local coordinate (u, v), the first fundamental form of ds^2 is given by

$$ds^2 = E du^2 + 2F du dv + G dv^2 \tag{2}$$

with

$$E = X_u \cdot X_u, \quad F = X_u \cdot X_v, \quad G = X_v \cdot X_v.$$

Let $z = u + \sqrt{-1}v$. Then we have

$$X_u = \phi + \overline{\phi}, \quad X_v = \sqrt{-1}(\phi - \overline{\phi}).$$

Utilizing the complex local coordinate $(z, \bar{z})$, we can rewrite (2) as

$$ds^2 = h dz^2 + 2\|\phi\|^2 |dz|^2 + \overline{h dz^2}, \tag{3}$$

where

$$h = \phi \cdot \phi = \frac{E - G - 2\sqrt{-1}F}{4}, \quad \|\phi\|^2 = \phi \cdot \overline{\phi} = \frac{E + G}{4}. \tag{4}$$

We call the quadratic differential $\eta := h dz^2$ the *Hopf differential*. It is clear that

$$|h| < \|\phi\|^2. \tag{5}$$

We call the metric $\Gamma := \frac{1}{2}\|\phi\|^2|dz|^2$ the *associated conformal metric* of ds^2, also known as the *Klotz metric*. From (3) and (5), if ds^2 is complete, then the associated Klotz metric $\|\phi\|^2|dz|^2$ is also complete (see also Lemma 1 in [Milnor 1976]). The immersion X is said to be *weakly complete* if the associated Klotz metric $\|\phi\|^2|dz|^2$ is complete.

Denote by $\mathfrak{K}(I)$ the intrinsic curvature (Gauss curvature) of the induced metric ds^2 above and $\mathfrak{K}(\Gamma)$ the Gauss curvature with respect to the Klotz metric $\Gamma := \frac{1}{2}\|\phi\|^2|dz|^2$. By Lemma 1 in [Milnor 1980], there exists a positive function $\mu \leq 1$ such that

$$\mathfrak{K}(\Gamma) \leq \mu \mathfrak{K}(I). \tag{6}$$

For any choice of a unit normal vector field $\mathfrak{n}$, one has an associated second fundamental form

$$II(\mathfrak{n}) = L du^2 + 2M du dv + N dv^2$$

with $\triangle X \cdot \mathfrak{n} = L + N \equiv 0$. Thus $\det(II(\mathfrak{n})) = -(L^2 + M^2) \leq 0$. It follows that $\mathfrak{K}(I) \leq 0$ since

$$\mathfrak{K}(I) = \frac{\sum_{j=1}^{n-2} \det(II(\mathfrak{n}_j))}{EG - F^2} \tag{7}$$

for any choices $(\mathfrak{n}_j)$ of $n-2$ mutually orthogonal unit normal vector fields.

3. Value distribution properties of the generalized Gauss map

Let

$$X = (X^1, \ldots, X^n) : M \to \mathbb{R}^n$$

be an immersed harmonic surface with the induced metric, where M is a Riemann surface. Let $\Phi = [\Phi_1 : \cdots : \Phi_n] : M \to \mathbb{P}^{n-1}(\mathbb{C})$, where $\Phi_i := \left(\frac{\partial}{\partial z} X^i\right) dz$, be the generalized Gauss map. In this section, we study the value distribution properties for the generalized Gauss map Φ. We begin with the following lemma.

Lemma 3.1. *Let $X = (X^1, \ldots, X^n) : M \to \mathbb{R}^n$ be an immersed harmonic surface with the induced metric, where M is a Riemann surface. Let*

$$\Phi = [\Phi_1 : \cdots : \Phi_n] : M \to \mathbb{P}^{n-1}(\mathbb{C}),$$

where $\Phi_i := \left(\frac{\partial}{\partial z} X^i\right) dz$, be the generalized Gauss map. If Φ is constant, then $X(M)$ lies in a 2-plane.

Proof. As above, denote by $\mathfrak{K}(\Gamma)$ the Gauss curvature with respect to the Klotz metric $\Gamma := \frac{1}{2}\|\phi\|^2|dz|^2$. Then

$$\mathfrak{K}(\Gamma) = -\frac{\|\phi'\|^2\|\phi\|^2 - (\phi' \cdot \overline{\phi})(\overline{\phi'} \cdot \phi)}{(\|\phi\|^2)^3}.$$

Since Φ is constant, we conclude that ϕ is constant, so $\mathfrak{K}(\Gamma) \equiv 0$. From (6), we get

$$0 = \mathfrak{K}(\Gamma) \leq \mu \mathfrak{K}(I),$$

where $\mu > 0$. Using the fact that $\mathfrak{K}(I) \leq 0$, we get $\mathfrak{K}(I) \equiv 0$. Now, from (7) and $\det(II(\mathfrak{n}_j)) \leq 0$ for all j, we conclude that $II(\mathfrak{n}_j) \equiv 0$ for all j, so $X(M)$ lies in a 2-plane. □

We recall the following result due to Milnor [1983] (see also Theorem 2.1 in [Jensen and Rigoli 1988]).

Theorem 3.2 [Milnor 1983, Theorem 3]. *Let $X = (X^1, \dots, X^n) : M \to \mathbb{R}^n$ be a complete immersed harmonic surface with the induced metric, where M is a Riemann surface. Let $\Phi = [\Phi_1 : \cdots : \Phi_n] : M \to \mathbb{P}^{n-1}(\mathbb{C})$, where $\Phi_i := \left(\frac{\partial}{\partial z} X^i\right) dz$ for $i = 1, \dots, n$, be the generalized Gauss map. Then either $X(M)$ is a 2-plane or else $\Phi(M)$ comes arbitrarily close to every hyperplane $\sum_{k=1}^n a_k w_k = 0$ in $\mathbb{P}^{n-1}(\mathbb{C})$.*

We note that Theorem 3.2 corresponds to Chern's theorem [1965] in the theory of minimal surfaces. It is known that Chern's result has been extended to a much sharper result by Fujimoto [1990] and Ru [1991] for the Gauss maps of minimal surfaces. In the following, we have the result corresponding to the result of Fujimoto and Ru.

Theorem 3.3. *Let M be an open Riemann surface and*

$$X = (X^1, \dots, X^n) : M \to R^n$$

be a harmonic immersion. Let $\Phi : M \to \mathbb{P}^{n-1}(\mathbb{C})$ be the generalized Gauss map. Assume that X is weakly complete with respect to the reduced metric. If Φ omits more than $\frac{1}{2}n(n+1)$ hyperplanes in $\mathbb{P}^{n-1}(\mathbb{C})$ in general position, then $X(M)$ lies in a 2-plane.

Remark 3.4. As we noted above, if the induced metric ds^2 is complete, then the associated Klotz metric $\|\phi\|^2|dz|^2$ is also complete (we say that X is weakly complete in this case).

In order to prove Theorem 3.3, we need to introduce some basic concepts and the following auxiliary results.

Let $H = \{[z_0 : z_1 : \cdots : z_k] \mid a_0 z_0 + \cdots + a_k z_k = 0\}$ be a hyperplane in $\mathbb{P}^k(\mathbb{C})$; here $\boldsymbol{a} = (a_0, \dots, a_k) \in \mathbb{C}^{k+1} \setminus \{\mathbf{0}\}$ is called *the normal vector* associated to H.

Hyperplanes $H_1, \dots, H_q$ are said to be *in m-subgeneral position* (with $m \geq k$) if and only if for every injective map $\mu : \{0, 1, \dots, m\} \to \{1, \dots, q\}$, the linear span of those corresponding normal vectors $\boldsymbol{a}_{\mu(0)}, \dots, \boldsymbol{a}_{\mu(m)}$ is $\mathbb{C}^{k+1}$. When $m = k$, then we just say the $H_1, \dots, H_q$ are in general position in $\mathbb{P}^k(\mathbb{C})$. It is clear that if the hyperplanes $H_1, \dots, H_q$ in $\mathbb{P}^m(\mathbb{C})$ are in general position, then, for $k \leq m$ and regarding $\mathbb{P}^k(\mathbb{C}) \subset \mathbb{P}^m(\mathbb{C})$, the restricted hyperplanes $H_1 \cap \mathbb{P}^k(\mathbb{C}), \dots, H_q \cap \mathbb{P}^k(\mathbb{C})$ are in m-subgeneral position.

We need the following lemma.

Lemma 3.5 [Chen 1987; Nochka 1983]. *Let $\{H_j\}_{j=1}^q$ be a set of hyperplanes in $\mathbb{P}^k(\mathbb{C})$ in m-subgeneral position. Then there exist a function $\varpi : J = \{1, \dots, q\} \to \mathbb{R}$ and a number $\theta > 0$ with the following properties*:

- $0 < \varpi(j) \leq 1$ *for all $j \in J$.*
- $q - 2m + k - 1 = \theta\big(\sum_{j=1}^q \varpi(j) - k - 1\big)$.
- $1 \leq \frac{m+1}{k+1} \leq \theta \leq \frac{2m-k+1}{k+1}$.

We call $\varpi(j)$ the Nochka weight associated to the hyperplane $H_j\,(1 \leq j \leq q)$.

Next, we recall the definition of the *derived curves*. Let F be a nondegenerate holomorphic map of Δ_R into $\mathbb{P}^k(\mathbb{C})$ (i.e., $F(\Delta_R)$ is not contained in any proper subspaces of $\mathbb{P}^k(\mathbb{C})$), where $\Delta_R := \{z \mid |z| < R\} \subset \mathbb{C}$ and $0 < R \leq \infty$. Take a reduced representation $\tilde{F} = (f_0, f_1, \dots, f_k)$ of F, i.e., $\tilde{F} : \Delta_R \to \mathbb{C}^{k+1} \setminus \{\mathbf{0}\}$ and $\mathbb{P}(\tilde{F}) = F$, where $\mathbb{P}$ is the natural projection. Let $\|\tilde{F}\| = \big(\sum_{j=0}^k |f_j|^2\big)^{\frac{1}{2}}$. Take the s-th derivative $\tilde{F}^{(s)} = (f_0^{(s)}, f_1^{(s)}, \dots, f_k^{(s)})$ and define

$$\tilde{F}_s = \tilde{F}^{(0)} \wedge \tilde{F}^{(1)} \wedge \cdots \wedge \tilde{F}^{(s)} : \Delta_R \to \bigwedge^{s+1} \mathbb{C}^{k+1} \tag{8}$$

for each $0 \leq s \leq k$. Obviously, $\tilde{F}_{k+1} \equiv 0$. Let $F_s = \mathbb{P}(\tilde{F}_s)$. We call the map F_s the *s-th derived curve of F*.

For holomorphic functions $f_0, f_1, \dots, f_k$, we denote

$$W(f_0, f_1, \dots, f_k) := \det(f_j^{(s)}, 0 \leq j, s \leq k).$$

Let $\{e_0, e_1, \dots, e_k\}$ be the standard basis of $\mathbb{C}^{k+1}$. Then we can write, for $0 \leq s \leq k$,

$$\tilde{F}_s = \sum_{0 \leq i_0 < \cdots < i_s \leq k} W(f_{i_0}, \dots, f_{i_s}) e_{i_0} \wedge \cdots \wedge e_{i_s}.$$

Hence,

$$|\tilde{F}_s|^2 := \sum_{0 \leq i_0 < \cdots < i_s \leq k} |W(f_{i_0}, \dots, f_{i_s})|^2.$$

From above, it is easy to see that if F is nondegenerate, then $\tilde{F}_s \not\equiv 0$ for any $0 \leq s \leq k - 1$.

For a hyperplane H_j in $\mathbb{P}^k(\mathbb{C})$ with the normal vector $\boldsymbol{a}_j = (a_{j0}, \dots, a_{jk})$, we define, for $0 \le s \le k$ and $1 \le j \le q$,

$$|\tilde{F}_s(H_j)|^2 = \sum_{0 \le i_1 < \cdots < i_s \le k} \Big| \sum_{t \ne i_1, \dots, i_s} a_{jt} W(f_t, f_{i_1}, \dots, f_{i_s}) \Big|^2. \tag{9}$$

Notice that

$$|\tilde{F}(H_j)| = |\tilde{F}_0(H_j)| = |a_{j0} f_0 + a_{j1} f_1 + \cdots + a_{jk} f_k|.$$

From (9), we see that $\tilde{F}_s(H_j) \equiv 0$ if and only if it holds for all $i_1, \dots, i_s$ that

$$\sum_{t \ne i_1, \dots, i_s} a_{jt} W(f_t, f_{i_1}, \dots, f_{i_s}) \equiv 0.$$

Thus if F is nondegenerate, then $\tilde{F}_s(H_j) \not\equiv 0$ for all $0 \le s \le k-1$ and $1 \le j \le q$. Indeed, if $\tilde{F}_s(H_j) \equiv 0$ for some s and j, then

$$W(\tilde{F}(H_j), f_{i_1}, \dots, f_{i_s}) \equiv 0$$

for all $i_1, \dots, i_s$. This implies that $\tilde{F}(H_j),\ f_{i_1}, \dots, f_{i_s}$ are linearly independent, which contradicts the nondegeneracy of f.

The following result is due to Ru [1991, Main Lemma]. It plays an important role in the proof of our theorem.

Lemma 3.6 [Ru 1991, Main Lemma]. *Let $F : \Delta_R \to \mathbb{P}^k(\mathbb{C})$ be a nondegenerate holomorphic map with its reduced representation $\tilde{F}$. Let $\{H_j\}_{j=1}^q$ be a set of hyperplanes in $\mathbb{P}^k(\mathbb{C})$ in m-subgeneral position, and $\varpi(j)$ be their associated Nochka weights. If $q > 2m - k + 1$ and*

$$N > \frac{2qk(k+2)}{\sum_{j=1}^q \varpi(j) - (k+1)},$$

then there exists a positive constant C such that

$$|\tilde{F}|^{\chi} \frac{|\tilde{F}_k|^{1+\frac{2q}{N}} \prod_{j=1}^q \big(\prod_{s=0}^{k-1} |\tilde{F}_s(H_j)|\big)^{\frac{4}{N}}}{\prod_{j=1}^q |\tilde{F}(H_j)|^{\varpi(j)}} \le C \left(\frac{2R}{R^2 - |z|^2} \right)^{\frac{1}{2}k(k+1) + \frac{2q}{N} \sum_{s=0}^k s^2},$$

where $\chi = \sum_{j=1}^q \varpi(j) - (k+1) - \frac{2q}{N}(k^2 + 2k - 1)$.

We also need the following lemma.

Lemma 3.7 [Fujimoto 1993, Lemma 1.6.7]. *Let $d\sigma^2$ be a conformal flat metric on an open Riemann surface M. Then for each point $p \in M$, there exists a local diffeomorphism Ψ of a disk $\Delta_R = \{w \in \mathbb{C} \mid |w| < R\} (0 < R \le \infty)$ onto an open neighborhood of p with $\Psi(0) = p$ such that Ψ is a local isometry (i.e., the pullback $\Psi^*(d\sigma^2)$ is equal to the standard Euclidean metric ds_E^2 on Δ_R),* *and there exists a point a_0 with $|a_0| = 1$, such that the Ψ-image Γ_{a_0} of the line $L_{a_0} = \{w = a_0 t : 0 < t < R\}$ is divergent in M.*

Now we are ready to prove Theorem 3.3.

Proof of Theorem 3.3. Assume that the holomorphic map $\Phi : M \to \mathbb{P}^{n-1}(\mathbb{C})$ omits the hyperplanes $H_1, \dots, H_q$, which are in general position in $\mathbb{P}^{n-1}(\mathbb{C})$ with $q > \frac{1}{2}n(n+1)$. From Lemma 3.1, it suffices to prove that Φ is constant. By taking the universal cover of M if necessary, one can assume that M is simply connected. It follows from the uniformization theorem that M is conformally equivalent to $\mathbb{C}$ or the unit disk Δ. By Nochka's result [1983] (see also [Chen 1987]) about the Cartan conjecture, we know that Φ is constant when M is conformally equivalent to $\mathbb{C}$. So the result holds in this case.

Therefore it suffices to consider the case where M is the unit disc Δ. Assume that Φ is not constant. We want to derive a contradiction. From the assumption that Φ is not constant, there exists k $(1 \le k \le n-1)$ such that the image of Φ is contained in $\mathbb{P}^k(\mathbb{C}) \subset \mathbb{P}^{n-1}(\mathbb{C})$, but not in any subspace whose dimension is lower than k. In other words, $\Phi : \Delta \to \mathbb{P}^k(\mathbb{C})$ is a nondegenerate map. Let

$$\tilde{\Phi} = (\phi_0, \phi_1, \dots, \phi_k),$$

then, by the assumption, the metric $\|\tilde{\Phi}(z)\|^2 |dz|^2$ is complete. Let $\tilde{H}_j := H_j \cap \mathbb{P}^k(\mathbb{C})$, $1 \le j \le q$. Then these hyperplanes are in $(n-1)$-subgeneral position in $\mathbb{P}^k(\mathbb{C})$. One may assume that $\tilde{H}_j$ is given by

$$\tilde{H}_j : a_{j0} z_0 + a_{j1} z_1 + \cdots + a_{jk} z_k = 0 \quad (1 \le j \le q).$$

Since $\Phi : \Delta \to \mathbb{P}^k(\mathbb{C})$ is nondegenerate, from the discussion above, $\tilde{\Phi}_k(z) \not\equiv 0$ and none of the $\tilde{\Phi}_s(\tilde{H}_j)$, $0 \le s \le k$, $1 \le j \le q$, vanishes identically. Thus, by (9) for each $\tilde{\Phi}_s(\tilde{H}_j)$, there exist $i_1, i_2, \dots, i_s$ such that

$$\psi_{js} := \sum_{t \ne i_1, \dots, i_s} a_{jt} W(\phi_t, \phi_{i_1}, \dots, \phi_{i_s}) \tag{10}$$

does not vanish identically. Note that every ψ_{js} is a holomorphic function and has only isolated zeros.

Let $\varpi(j)$ be the Nochka weight associated to the hyperplane $\{\tilde{H}_j\}$ for $1 \le j \le q$. By Lemma 3.5, one has

$$q - 2(n-1) + k - 1 = \theta\left(\sum_{j=1}^{q} \varpi(j) - k - 1\right),$$

and

$$\theta \le \frac{2(n-1) - k + 1}{k+1}.$$

Since $q > \frac{1}{2}n(n+1)$, and noticing that $n(n+1)/2 \ge (k+1)(n - k/2 - 1) + n$

always holds for any $0 \le k \le n-1$, it is easy to verify that

$$\frac{2(q-2(n-1)+k-1)}{k(2(n-1)-k+1)} = \frac{2(q-2n+k+1)}{2kn-k^2-k} > 1,$$

for all $1 \le k \le n-1$. Hence

$$\begin{aligned}\frac{2\big(\sum_{j=1}^{q}\varpi(j)-k-1\big)}{k(k+1)} &= \frac{2(q-2(n-1)+k-1)}{\theta k(k+1)}\\ &\ge \frac{2(q-2(n-1)+k-1)}{k(2(n-1)-k+1)} > 1,\end{aligned}$$

which yields that

$$\sum_{j=1}^{q}\varpi(j)-k-1-\frac{1}{2}k(k+1) > 0.$$

Let

$$\chi := \sum_{j=1}^{q}\varpi(j)-(k+1)-\frac{2q}{N}(k^2+2k-1),$$

$$\lambda_0 := \frac{1}{\chi}\bigg(\frac{1}{2}k(k+1)+\frac{2q}{N}\sum_{s=0}^{k}s^2\bigg).$$

Choose some N such that

$$\frac{\sum_{j=1}^{q}\varpi(j)-k-1-\frac{k}{2}(k+1)}{\frac{2}{q}+\sum_{s=0}^{k}(k-s)^2+k^2+2k-1} < \frac{2q}{N} < \frac{\sum_{j=1}^{q}\varpi(j)-k-1-\frac{k}{2}(k+1)}{\sum_{s=0}^{k}(k-s)^2+k^2+2k-1},$$

which implies that

$$0 < \lambda_0 < 1, \qquad \frac{4}{N\chi(1-\lambda_0)} > 1.$$

We define a new metric

$$d\sigma^2 = \left(\frac{\prod_{j=1}^{q}|\tilde{\Phi}(\tilde{H}_j)|^{\varpi(j)}}{|\tilde{\Phi}_k|^{1+\frac{2q}{N}}\prod_{j=1}^{q}\big(\prod_{s=0}^{k-1}|\psi_{js}|\big)^{\frac{4}{N}}}\right)^{\frac{2}{(1-\lambda_0)\chi}}|dz|^2 \tag{11}$$

on the set $M' := \Delta \setminus \big\{p \in \Delta \mid \text{either } \tilde{\Phi}_k = 0 \text{ or } \prod_{j=1}^{q}\prod_{s=0}^{k-1}|\psi_{js}| = 0\big\}$.

Notice that $d\sigma^2$ is a flat metric on M'. Fix a point $p_0 \in M'$; by Lemma 3.7 there exists a local diffeomorphism Ψ of a disk $\Delta_R = \{w \in \mathbb{C} : |w| < R\}$ $(0 < R \le \infty)$ onto an open neighborhood of p_0 with $\Psi(0) = p_0$ such that Ψ is a local isometry. Furthermore, there exists a point a_0 with $|a_0| = 1$ and the Ψ-image Γ_{a_0} of the line $L_{a_0} = \{w = a_0 t : 0 < t < R\}$ is divergent in M'. Again, by Nochka's result [1983] (see also [Chen 1987]), we know that $R < \infty$ since Φ is nonconstant. We claim that the Ψ-image Γ_{a_0} actually is divergent to the boundary of Δ. To this end, we assume the contrary, that the curve Γ_{a_0} is divergent to a point z_0 which either satisfies $\tilde{\Phi}_k(z_0) = 0$ or $\psi_{js}(z_0) = 0$ for some s and j. Thus, we have

$$\lim \inf_{z \to z_0} |\tilde{\Phi}_k|^{(N+2q)\delta_0/2} \prod_{1 \le j \le q, 1 \le s \le k-1} |\psi_{js}|^{2\delta_0} \cdot v > 0,$$

where

$$v = \left(\frac{\prod_{j=1}^q |\tilde{\Phi}(\tilde{H}_j)|^{\varpi(j)}}{|\tilde{\Phi}_k|^{1+\frac{2q}{N}} \prod_{j=1}^q \left(\prod_{s=0}^{k-1} |\psi_{js}|\right)^{\frac{4}{N}}} \right)^{\frac{2}{(1-\lambda_0)\chi}}$$

and

$$\delta_0 = \frac{4}{N\chi(1-\lambda_0)} > 1.$$

Thus,

$$\begin{aligned} R &= \int_{L_{a_0}} \Psi^* \, d\sigma = \int_{\Gamma_{a_0}} d\sigma \\ &= \int_{\Gamma_{a_0}} \left(\frac{\prod_{j=1}^q |\tilde{\Phi}(\tilde{H}_j)|^{\varpi(j)}}{|\tilde{\Phi}_k|^{1+\frac{2q}{N}} \prod_{j=1}^q \left(\prod_{s=0}^{k-1} |\psi_{js}|\right)^{\frac{4}{N}}} \right)^{\frac{1}{(1-\lambda_0)\chi}} |dz| \\ &\ge c \int_{\Gamma_{a_0}} \frac{1}{|z - z_0|^{\delta_0}} |dz| = \infty, \end{aligned}$$

which yields a contradiction.

Therefore $\Gamma_{a_0} = \Psi(L_{a_0})$ is divergent to the boundary of Δ. To compute the length of Γ_{a_0} with respect to the Klotz metric $\|\tilde{\Phi}\|^2 |dz|^2$ where $\tilde{\Phi} = (\phi_0, \dots, \phi_k)$, we introduce the following functions defined on $\{w \mid |w| < R\}$:

$$f_s(w) := \phi_s(\Psi(w)) \quad (0 \le s \le k),$$

and $F(w) := (f_0(w), f_1(w), \dots, f_k(w))$. For $1 \le j \le q, \ 0 \le s \le k$, we define

$$F(H_j) := a_{j0} f_0 + \cdots + a_{jk} f_k, \quad F_k := W(f_0, f_1, \dots f_k)$$

and

$$\varphi_{js} := \sum_{t \ne i_1, \dots, i_s} a_{jt} W(f_t, f_{i_1}, \dots, f_{i_s}),$$

where $(i_1, \dots, i_s)$ is the index in the definition of ψ_{js} in (10). Noticing the fact that, for $0 \le s \le k$,

$$F_s(w) = (F \wedge F' \wedge \cdots \wedge F^{(s)})(w) = (\tilde{\Phi} \wedge \cdots \wedge \tilde{\Phi}^{(s)})(z) \left(\frac{dz}{dw} \right)^{s(s+1)/2},$$

we have, from (11), that

$$\Psi^* d\sigma = \Psi^* \left(\frac{\prod_{j=1}^q |\tilde{\Phi}(\tilde{H}_j)|^{\varpi(j)}}{|\tilde{\Phi}_k|^{1+\frac{2q}{N}} \prod_{j=1}^q \left(\prod_{s=0}^{k-1} |\psi_{js}|\right)^{\frac{4}{N}}} \right)^{\frac{1}{(1-\lambda_0)\chi}} \left|\frac{dz}{dw}\right| |dw|$$

$$= \left(\frac{\prod_{j=1}^q |F(\tilde{H}_j)|^{\varpi(j)}}{|F_k|^{1+\frac{2q}{N}} \prod_{j=1}^q \left(\prod_{s=0}^{k-1} |\varphi_{js}|\right)^{\frac{4}{N}}} \right)^{\frac{1}{(1-\lambda_0)\chi}} \times \left|\frac{dz}{dw}\right|^{\frac{\left(1+\frac{2q}{N}\right)\frac{k(k+1)}{2}+\frac{2q}{N}\sum_{s=0}^{k-1} s(s+1)}{(1-\lambda_0)\chi}+1} |dw|$$

$$= \left(\frac{\prod_{j=1}^q |F(\tilde{H}_j)|^{\varpi(j)}}{|F_k|^{1+\frac{2q}{N}} \prod_{j=1}^q \left(\prod_{s=0}^{k-1} |\varphi_{js}|\right)^{\frac{4}{N}}} \right)^{\frac{1}{(1-\lambda_0)\chi}} \left|\frac{dz}{dw}\right|^{\frac{\lambda_0}{1-\lambda_0}+1} |dw|$$

$$= \left(\frac{\prod_{j=1}^q |F(\tilde{H}_j)|^{\varpi(j)}}{|F_k|^{1+\frac{2q}{N}} \prod_{j=1}^q \left(\prod_{s=0}^{k-1} |\varphi_{js}|\right)^{\frac{4}{N}}} \right)^{\frac{1}{(1-\lambda_0)\chi}} \left|\frac{dz}{dw}\right|^{\frac{1}{1-\lambda_0}} |dw|.$$

Using the isometry property of Ψ, i.e., $|dw| = \Psi^* d\sigma$, we get

$$\left|\frac{dw}{dz}\right| = \left(\frac{\prod_{j=1}^q |F(\tilde{H}_j)|^{\varpi(j)}}{|F_k|^{1+\frac{2q}{N}} \prod_{j=1}^q \left(\prod_{s=0}^{k-1} |\varphi_{js}|\right)^{\frac{4}{N}}} \right)^{\frac{1}{\chi}}. \tag{12}$$

Now, denote by $l(\Gamma_{a_0})$ the length of the curve Γ_{a_0} with respect to the Klotz metric $\|\tilde{\Phi}\|^2|dz|^2$; then from (12),

$$l(\Gamma_{a_0}) = \int_{\Gamma_{a_0}} \|\tilde{\Phi}\| |dz| = \int_{L_{a_0}} \Psi^*(\|\tilde{\Phi}\| |dz|)$$

$$= \int_{L_{a_0}} \|\tilde{\Phi}(\Psi(w))\| \left|\frac{dz}{dw}\right| |dw|$$

$$= \int_{L_{a_0}} \|F\| \left(\frac{|F_k|^{1+\frac{2q}{N}} \prod_{j=1}^q \left(\prod_{s=0}^{k-1} |\varphi_{js}|\right)^{\frac{4}{N}}}{\prod_{j=1}^q |F(\tilde{H}_j)|^{\varpi(j)}} \right)^{\frac{1}{\chi}} |dw|$$

$$\leq \int_{L_{a_0}} \left(\frac{\|F\|^{\chi} |F_k|^{1+\frac{2q}{N}} \prod_{j=1}^q \left(\prod_{s=0}^{k-1} |F_s(\tilde{H}_j)|\right)^{\frac{4}{N}}}{\prod_{j=1}^q |F(\tilde{H}_j)|^{\varpi(j)}} \right)^{\frac{1}{\chi}} |dw|.$$

In the above inequality, we use the fact that $|\varphi_{js}| \le |F_s(\tilde{H}_j)|$ for all $0 \le s \le k$, $1 \le j \le q$. By Lemma 3.6, it can be concluded that, noticing that $0 < \lambda_0 < 1$,

$$l(\Gamma_{a_0}) \le C \int_0^R \left(\frac{2R}{R^2 - |w|^2} \right)^{\lambda_0} |dw| < \infty,$$

which contradicts the completeness of the Klotz metric $\|\tilde{\Phi}\|^2|dz|^2$. Thus Φ is a constant map. □

4. Value distribution properties for the Gauss map of the harmonic and QC-harmonic surface immersed in $\mathbb{R}^3$

In this section, we study the value distribution properties for the Gauss map (i.e., its unit normal) of harmonic and K-quasiconformal surfaces immersed in $\mathbb{R}^3$. Our method is to compare the Gauss map of the surface (i.e., its unit-normal) with the generalized Gauss map Φ and apply the results obtained in the previous section. The classical Bernstein theorem says that a minimal graph over a plane is planar. W. H. Meeks III and H. Rosenberg [2005] showed that a complete embedded minimal surface over a plane is either a plane or a helicoid. It is known that the classical Bernstein theorem still holds for the K-quasiconformal harmonic graph, but fails if we only assume the graph is harmonic. So in order to get the desired value distribution properties for the Gauss map of harmonic surfaces immersed in $\mathbb{R}^3$, the additional condition that X be K-quasiconformal seems necessary.

Let M be an open Riemann surface and

$$X = (X^1, X^2, X^3) : M \to \mathbb{R}^3$$

be a harmonic immersion.Write

$$\|\nabla X\|^2 = E + G = 4\|\phi\|^2, \tag{13}$$

where $\|\nabla X\|^2$ is the Hilbert–Schmidt norm defined by

$$\|\nabla X\|^2 := \|\partial X/\partial u\|^2 + \|\partial X/\partial v\|^2.$$

Also the Jacobian of X is given by

$$J_X = \|X_u \times X_v\| = \sqrt{EG - F^2} = 2\sqrt{\|\phi\|^4 - |h|^2}, \tag{14}$$

where ϕ is given by (1), and h and $\|\phi\|$ are defined by (4).

Let $\mathfrak{n}$ be a normal vector on a harmonic surface $X : M \to \mathbb{R}^3$, and $\mathfrak{b}$ be a unit vector in $\mathbb{R}^3$. The following proposition aims to give a relation between the intersection of $\mathfrak{n}$ and $\mathfrak{b}$ and the projective distance of the generalized Gauss map Φ to a hyperplane with the normal $\mathfrak{b}$. Instead of the proof which is originally from Lemma 1.1 in [Osserman 1964] we use the mixed product to deal with it.

Proposition 4.1. *Let $X : M \to \mathbb{R}^3$ be a harmonic surface. Then every normal vector makes an angle α with the unit vector $\mathfrak{b}$ at a given point if and only if*

$$\frac{|\mathfrak{b}\cdot\phi|^2}{\|\phi\|^2}\frac{\|\phi\|^4}{\|\phi\|^4-|h|^2}-\frac{(\mathfrak{b}\cdot\phi)^2\bar{h}+(\mathfrak{b}\cdot\overline{\phi})^2h}{2\|\phi\|^2}\frac{\|\phi\|^2}{\|\phi\|^4-|h|^2}=\frac{1}{2}\sin^2\alpha, \tag{15}$$

where ϕ is given by (1), and h and $\|\phi\|$ are defined by (4).

Proof. The unit normal vector is given by $\mathfrak{n}=(X_u\times X_v)/J_X$. For a unit vector $\mathfrak{b}$, we denote by $|A|$ the mixed product of the three vectors $\mathfrak{b}$, X_u and X_v, where A is a matrix determined by three row vectors $\mathfrak{b}$, X_u and X_v. Then it follows from determinant expansion of $|AA^T|$ in its first row that

$$\begin{aligned}|\mathfrak{n}\cdot\mathfrak{b}|^2&=\frac{|X_u\times X_v\cdot\mathfrak{b}|^2}{EG-F^2}=\frac{|AA^T|}{EG-F^2}\\&=1-\frac{(\mathfrak{b}\cdot X_u)^2G+(\mathfrak{b}\cdot X_v)^2E-2(\mathfrak{b}\cdot X_u)(\mathfrak{b}\cdot X_v)F}{EG-F^2}.\end{aligned}$$

Utilizing the relations

$$X_u=\phi+\overline{\phi}\quad\text{and}\quad X_v=\sqrt{-1}(\phi-\overline{\phi}),$$

we have, from the above relation, that

$$|\mathfrak{n}\cdot\mathfrak{b}|^2=1-\frac{(\mathfrak{b}\cdot\phi)^2(G-E-2iF)+(\mathfrak{b}\cdot\overline{\phi})^2(G-E+2iF)}{EG-F^2}-\frac{2|b\cdot\phi|^2(E+G)}{EG-F^2}.$$

By the relations (4), (13) and (14), the above equality is equivalent to

$$|\mathfrak{n}\cdot\mathfrak{b}|^2=1+\frac{(\mathfrak{b}\cdot\phi)^2\bar{h}+(\mathfrak{b}\cdot\overline{\phi})^2h}{\|\phi\|^2}\frac{\|\phi\|^2}{\|\phi\|^4-|h|^2}-\frac{2|b\cdot\phi|^2}{\|\phi\|^2}\frac{\|\phi\|^4}{\|\phi\|^4-|h|^2}. \tag{16}$$

Since $\mathfrak{n}$ and $\mathfrak{b}$ are two unit vectors, $|\mathfrak{n}\cdot\mathfrak{b}|=|\cos\alpha|$. Hence, the normal vector $\mathfrak{n}$ makes an angle α with the unit vector $\mathfrak{b}$ if and only if the equality (15) holds. □

Remark 4.2. If $X : M \to \mathbb{R}^3$ is a minimal surface with a conformal minimal immersion X, then we have $h\equiv 0$. In this case, Proposition 4.1 implies that

$$\frac{|\phi\cdot\mathfrak{b}|^2}{\|\phi\|^2}=\frac{1}{2}\sin^2\alpha.$$

This relation shows that in the case of minimal surfaces, a normal vector $\mathfrak{n}$ makes an angle of at least α with a given vector $\mathfrak{b}$ if and only if its generalized Gauss map Φ has a positive projective distance to a hyperplane H with the unit normal $\mathfrak{b}$. If we take $\mathfrak{b}$ to be the x^3-axis, then our proposition is Lemma 1.1 in [Osserman 1964] in the $\mathbb{R}^3$ case.

Remark 4.3. By the fact that $|h| < \|\phi\|^2$ for a harmonic surface, the relation (15) gives us the following inequality

$$\frac{1-|\mathfrak{n}\cdot\mathfrak{b}|^2}{2} = \frac{1}{2}\sin^2\alpha \geq \frac{|\mathfrak{b}\cdot\phi|^2}{\|\phi\|^2}\frac{\|\phi\|^2}{\|\phi\|^2+|h|} \geq \frac{1}{2}\frac{|\mathfrak{b}\cdot\phi|^2}{\|\phi\|^2},$$

which shows that the inequality $|\mathfrak{b}\cdot\phi|^2/\|\phi\|^2 \geq \epsilon > 0$ implies that $|\mathfrak{n}\cdot\mathfrak{b}| \leq \eta < 1$, that is, it will force normals to avoid some neighborhood of the unoriented direction determined by the vector $\mathfrak{b}$. Conversely, it is also true that

$$\frac{|\mathfrak{b}\cdot\phi|^2}{\|\phi\|^2} \geq \frac{1-|\mathfrak{n}\cdot\mathfrak{b}|^2}{2}\frac{\|\phi\|^2-|h|}{\|\phi\|^2}.$$

However, the result that $|\mathfrak{b}\cdot\phi|^2/\|\phi\|^2$ has a positive lower bound cannot be derived from this inverse inequality under the condition that $|\mathfrak{n}\cdot\mathfrak{b}| \leq \eta < 1$, that is, the condition that normals avoid some neighborhood of the unoriented direction determined by a vector $\mathfrak{b}$ does not imply that $|\mathfrak{b}\cdot\phi|^2/\|\phi\|^2 \geq \epsilon > 0$. The harmonic rotational horn with the generalized Gauss map $\Phi = [1dz : \sqrt{-1}dz : 1/zdz]$, which is given by A. Alarcón and F. J. López [2013], is such a counterexample. To verify it, one can choose the unit vector $\mathfrak{b} = 1/\sqrt{u^2+v^2}(v,u,0)$, then $\mathfrak{n}\cdot\mathfrak{b} = 0$ and $|\mathfrak{b}\cdot\phi|^2/\|\phi\|^2 \to 0$ as $|z| \to 0$.

From Remark 4.3, we see that Theorem 3.2 would not tell any useful information about the unit-normal $\mathfrak{n}$. Indeed the Bernstein type theorem fails for some harmonic immersed surfaces. In order to derive some useful consequence from the previous results about Φ, we need to derive the lower bound for $|\mathfrak{b}\cdot\phi|^2/\|\phi\|^2$ when $|\mathfrak{n}\cdot\mathfrak{b}| \leq \eta < 1$. This requires the assumption that X is K-quasiconformal.

An immersion $X = (X^1, X^2, X^3) : M \to \mathbb{R}^3$ is called *K-quasiconformal* if it satisfies the inequality

$$\|\nabla X\|^2 \leq \left(K + \frac{1}{K}\right) J_X, \tag{17}$$

which is equivalent to

$$\|\phi\|^2 \leq \frac{K^2+1}{2K}\sqrt{\|\phi\|^4 - |h|^2}. \tag{18}$$

Note that we adopt the definition of quasiconformality given by D. Kalaj [2013] (see also [Alarcón and López 2013]). If $K = 1$, then the above inequality can be changed into the two relations

$$|X_u| = |X_v| \quad \text{and} \quad X_u \cdot X_v = 0, \tag{19}$$

where we say that X is an isothermal parametrization (isothermal coordinate) of the surface M. If a Riemann surface M admits a K-quasiconformal harmonic immersion X into $\mathbb{R}^3$, we call such a surface a *K-quasiconformal harmonic surface*, and say its immersion is a *K-quasiconformal harmonic immersion*.

Lemma 4.4. *Let M be a surface with a harmonic immersion X. Then X is K-quasiconformal in the sense of the definition given by (17) if and only if its Hopf differential η and conformal metric $\|\Phi\|$ satisfy*

$$|\eta| \leq Q_K \|\Phi\|^2, \tag{20}$$

where $Q_K = (K^2-1)/(K^2+1)$. Particularly, if $K = 1$ then X is a conformal immersion. Furthermore, the metric ds^2 and its associated conformal metric $\|\Phi\|^2$ satisfy the inequality

$$2(1-Q_K)\|\Phi\|^2 \leq ds^2 \leq 2(1+Q_K)\|\Phi\|^2. \tag{21}$$

The inequalities of the above relations hold simultaneously if and only if X is a conformal immersion.

Proof. The inequality (18) implies that

$$|h| \leq Q_K \|\phi\|^2, \tag{22}$$

i.e., $|\eta| \leq Q_K \|\Phi\|^2$ since $\eta := h\,dz^2$. Conversely, the inequality (20) (i.e., inequality (22)) implies that

$$\|\nabla X\|^2 \leq \left(K + \frac{1}{K}\right) J_X.$$

Therefore, X is K-QC in the sense of definition given by D. Kalaj. Finally, it is easy to see that the inequality (20) (i.e., inequality (22)) implies both the upper and lower bound of (21) from the first fundamental form (3). □

Under the assumption that X is K-quasiconformal, we give an estimate of the quantity

$$\frac{\frac{1}{2}(1-|\mathfrak{n}\cdot\mathfrak{b}|^2)}{|\phi\cdot\mathfrak{b}|^2/\|\phi\|^2}$$

in the quasiconformal distortion constant K for a K-quasiconformal harmonic surface in $\mathbb{R}^3$.

Lemma 4.5. *Let M be a K-quasiconformal harmonic surface in $\mathbb{R}^3$. Then for every unit normal vector $\mathfrak{n}$ and every unit vector $\mathfrak{b}$ at the same point p on M,*

$$\frac{K^2+1}{2K^2}\frac{|\phi\cdot\mathfrak{b}|^2}{\|\phi\|^2} \leq \frac{1-|\mathfrak{n}\cdot\mathfrak{b}|^2}{2} \leq \frac{K^2+1}{2}\frac{|\phi\cdot\mathfrak{b}|^2}{\|\phi\|^2}. \tag{23}$$

In particular, when $K = 1$ we have

$$\frac{1-|\mathfrak{n}\cdot\mathfrak{b}|^2}{2} = \frac{|\phi\cdot\mathfrak{b}|^2}{\|\phi\|^2}. \tag{24}$$

Proof. From the relations (16), (18) and (22), we have

$$\begin{aligned}\frac{1-|\mathfrak{n}\cdot\mathfrak{b}|^2}{2} &= \frac{|\mathfrak{b}\cdot\phi|^2}{\|\phi\|^2}\frac{\|\phi\|^4}{\|\phi\|^4-|h|^2} - \frac{(\mathfrak{b}\cdot\phi)^2\bar{h}+(\mathfrak{b}\cdot\bar{\phi})^2 h}{2\|\phi\|^2}\frac{\|\phi\|^2}{\|\phi\|^4-|h|^2}\\ &\le \left(\frac{\|\phi\|^4}{\|\phi\|^4-|h|^2}+\frac{\|\phi\|^2|h|}{\|\phi\|^4-|h|^2}\right)\frac{|\mathfrak{b}\cdot\phi|^2}{\|\phi\|^2}\\ &\le (1+Q_K)\frac{\|\phi\|^4}{\|\phi\|^4-|h|^2}\frac{|\mathfrak{b}\cdot\phi|^2}{\|\phi\|^2}\\ &\le \frac{(1+Q_K)(K+1/K)^2}{4}\frac{|\mathfrak{b}\cdot\phi|^2}{\|\phi\|^2} = \frac{K^2+1}{2}\frac{|\mathfrak{b}\cdot\phi|^2}{\|\phi\|^2}.\end{aligned}$$

Hence we complete the proof of the right-hand side of the relation (23).

Similarly, we have

$$\begin{aligned}\frac{1-|\mathfrak{n}\cdot\mathfrak{b}|^2}{2} &\ge \left(\frac{\|\phi\|^4}{\|\phi\|^4-|h|^2}-\frac{\|\phi\|^2|h|}{\|\phi\|^4-|h|^2}\right)\frac{|\mathfrak{b}\cdot\phi|^2}{\|\phi\|^2}\\ &\ge \left(\frac{\|\phi\|^2}{\|\phi\|^2+|h|}\right)\frac{|\mathfrak{b}\cdot\phi|^2}{\|\phi\|^2} \ge \frac{K^2+1}{2K^2}\frac{|\mathfrak{b}\cdot\phi|^2}{\|\phi\|^2}.\end{aligned}$$

Thus the proof of Lemma 4.5 is finished. □

Combining Lemma 4.5 with Theorem 3.2, we get the following theorem.

Theorem 4.6. *Let $X : M \to \mathbb{R}^3$ be a complete harmonic and K-quasiconformal immersion with the induced metric, where M is an open Riemann surface, and let $\mathfrak{n}$ be the unit normal of M. If its Gauss map (i.e., the normal $\mathfrak{n}$) omits a neighborhood of a direction in S^2, then X must be a plane.*

The above theorem recovers the classical Bernstein theorem for graphs of harmonic and K-quasiconformal surfaces defined on $\mathbb{R}^2$. It is known that the Bernstein theorem fails for graphs of harmonic surfaces on $\mathbb{R}^2$ without the K-quasiconformal assumption. Thus the additional assumption of K-quasiconformality seems necessary and reasonable in order to study the value distribution for its Gauss map.

From Lemma 4.5, we see that $|\mathfrak{n}\cdot\mathfrak{b}| = 1$ if and only if $\phi\cdot\mathfrak{b} = 0$. For any unit-vector $\mathfrak{b} = (b_1, b_2, b_3) \in \mathbb{R}^3$, it corresponds to a hyperplane

$$H_{\mathfrak{b}} = \{b_1w_1 + b_2w_2 + b_3w_3\} \subset \mathbb{P}^2(\mathbb{C}).$$

Notice that $H_{\mathfrak{b}_j}$, $1 \le j \le q$, are in general position if and only if any three of among $\mathfrak{b}_1, \ldots, \mathfrak{b}_q$ are not contained in a plane in $\mathbb{R}^3$. Thus Lemma 4.5 and Theorem 3.3 imply the following interesting result about the value distribution of the Gauss map $\mathfrak{n}$ which extends Theorem 4.6.

Theorem 4.7. *Let $X : M \to \mathbb{R}^3$ be a complete harmonic and K-quasiconformal immersion with the induced metric, where M is an open Riemann surface, and let $\mathfrak{n}$*

be the unit normal of M. If its Gauss map (*i.e., the normal* $\mathfrak{n}$) *omits seven directions, say unit vectors* $\mathfrak{b}_1, \dots, \mathfrak{b}_7$ *in* S^2, *and any three of which are not contained in a plane in* $\mathbb{R}^3$, *then X must be a plane.*

5. Estimate of Gauss curvature for K-QC harmonic surfaces

In this section, for a K-quasiconformal harmonic immersion $X : M \to \mathbb{R}^3$, we will study its Gauss curvature estimate if the Gauss map (not the generalized Gauss map) omits a neighborhood of some direction. For the result in the minimal surface case, see [Osserman and Ru 1997].

We first derive the expression for the Gauss curvature $\mathfrak{K}$ of a harmonic immersed surface in terms of its generalized Gauss map Φ. Let $X = (X^1, X^2, X^3)$ be its harmonic immersion. Take a local coordinate $z = u + \sqrt{-1}v$. It follows from $X_u = \phi + \overline{\phi}$ and $X_v = \sqrt{-1}(\phi - \overline{\phi})$ that

$$X_{uu} = \phi' + \overline{\phi'}, \quad X_{uv} = \sqrt{-1}(\phi' - \overline{\phi'}), \quad X_{vv} = -(\phi' + \overline{\phi'})$$

and

$$\mathfrak{n} = \frac{X_u \times X_v}{\sqrt{EG - F^2}} = \frac{\sqrt{-1}(\overline{\phi} \times \phi)}{\sqrt{\|\phi\|^4 - |h|^2}} \tag{25}$$

which imply that the Gauss curvature can be expressed by

$$\mathfrak{K} = \frac{LN - M^2}{EG - F^2} = -4\frac{|X_u \times X_v \cdot \overline{\phi'}|^2}{(EG - F^2)^2} = -\frac{|\overline{\phi} \times \phi \cdot \overline{\phi'}|^2}{(\|\phi\|^4 - |h|^2)^2}.$$

Furthermore, it can be rewritten by an expansion of the determinant as

$$\begin{aligned} \mathfrak{K} = -\frac{4}{(\sqrt{EG - F^2})^3} & \Big\{ \sqrt{EG - F^2}\|\phi'\|^2 + \frac{4\bar{h}(\phi' \cdot \phi)(\overline{\phi'} \cdot \phi)}{\sqrt{EG - F^2}} \\ & + \frac{4h(\phi' \cdot \overline{\phi})(\overline{\phi'} \cdot \overline{\phi})}{\sqrt{EG - F^2}} - \frac{4\|\phi\|^2[(\phi' \cdot \overline{\phi})(\overline{\phi'} \cdot \phi) + (\phi' \cdot \phi)(\overline{\phi'} \cdot \overline{\phi})]}{\sqrt{EG - F^2}} \Big\}. \end{aligned} \tag{26}$$

In particular, if X is a conformal harmonic immersion, that is, $h = 0$, the above relation can be simplified as

$$\mathfrak{K} = -\frac{\Delta \log \lambda}{\lambda^2} = -\frac{\|\phi'\|^2\|\phi\|^2 - (\phi' \cdot \overline{\phi})(\overline{\phi'} \cdot \phi)}{(\|\phi\|^2)^3} = -\frac{\sum_{i<j\le 3}|\phi_i\phi_j' - \phi_i'\phi_j|^2}{\|\phi\|^6}, \tag{27}$$

where $\lambda = \sqrt{2}\|\phi\|$.

It will be useful to introduce the meromorphic functions $\psi_k(z)$ defined by

$$\psi_k(z) = \frac{\phi_k(z)}{\phi_3(z)}, \quad k = 1, 2, 3. \tag{28}$$

Then

$$\psi_k' = \frac{\phi_3 \phi_k' - \phi_3' \phi_k}{\phi_3^2}$$

and

$$\psi_j \psi_k' - \psi_j' \psi_k = \frac{\phi_j \phi_k' - \phi_j' \phi_k}{\phi_3^2}, \quad j, k = 1, 2, 3.$$

Hence, we can rewrite (27) as

$$\mathfrak{K} = -\frac{|\psi_1 \psi_2' - \psi_1' \psi_2|^2 + \sum_{j=1}^{2} |\psi_j'|^2}{|\phi_3|^2 (1 + \sum_{j=1}^{2} |\psi_j|^2)^3}. \tag{29}$$

Lemma 5.1. *Let M be a K-quasiconformal harmonic surface from the unit disk $\{z \mid |z| < 1\}$ into R^3 whose unit normal makes an angle of at least $\alpha > 0$ with the x^3-axis at every point of the surface. If p is the point of M corresponding to $z = 0$, then the distance d from p to the boundary of M satisfies*

$$d \le \frac{2K}{\sqrt{K^2+1}} |\csc \alpha| \, |\phi_3(0)|. \tag{30}$$

Proof. After choosing $\mathfrak{b}$ to be a x^3-axis, we have $|\phi \cdot \mathfrak{b}| = |\phi_3|$. If a normal $\mathfrak{n}$ makes at least an angle α with the x^3-axis, then $|\mathfrak{n} \cdot \mathfrak{b}| \le |\cos \alpha|$. Thus the right-hand side of relation (23) becomes

$$\frac{|\phi_3|^2}{\|\phi\|^2} \ge \frac{1}{K^2+1} \sin^2 \alpha. \tag{31}$$

Let c be an arbitrary curve going from $z = 0$ to the boundary $|z| = 1$. Thus by (21) and (31), the length of the image γ on M of the curve c is given by

$$\begin{aligned} d_\gamma = \int_\gamma ds &\le \sqrt{2(1+Q_K)} \int_c \|\phi\| |dz| \\ &\le \frac{2K}{\sqrt{K^2+1}} |\csc \alpha| \int_c |\phi_3| |dz|. \end{aligned}$$

Set $F(z) = \int_0^z \phi_3(\zeta)\, d\zeta$. The inequality (31) shows that $F'(z) \ne 0$. Then $\zeta = F(z)$ will have an inverse $z = G(\zeta)$ in some disk with center $\zeta = 0$. Let R be the largest radius such that G is holomorphic, then by Liouville's theorem we have that R is finite. Hence, there exists a point ζ_0 on the circle $|\zeta| = R$ such that G cannot be extended to a neighborhood of ζ_0. Let L be the line segment from $\zeta = 0$ to ζ_0 and $\Gamma = G(L)$. Thus $G(R\zeta)$ satisfies Schwarz's lemma, which shows $|G'(0)| \le 1/R$. Thus,

$$\begin{aligned} d &= \inf_c \int_c ds \le \frac{2K}{\sqrt{K^2+1}} |\csc\alpha| \int_c |\phi_3||dz| \\ &= \frac{2K}{\sqrt{K^2+1}} |\csc\alpha| \int_L |d\zeta| = \frac{2K}{\sqrt{K^2+1}} |\csc\alpha| R \le \frac{2K|\csc\alpha|}{\sqrt{K^2+1}|G'(0)|} \\ &= \frac{2K}{\sqrt{K^2+1}} |\csc\alpha||F'(0)| = \frac{2K}{\sqrt{K^2+1}} |\csc\alpha||\phi_3(0)|. \end{aligned}$$ □

Theorem 5.2. *Let M be a K-quasiconformal harmonic surface in $\mathbb{R}^3$. Suppose that its unit normal makes an angle of at least $\alpha > 0$ with some fixed direction at every point of M. In addition, we assume that $|(\phi' \cdot \phi)(\overline{\phi'} \cdot \phi)|/\|\phi\|^4 \le N_K$, where N_K is a constant. If $d(p)$ is the distance of p to the boundary of M, then the Gauss curvature $\mathfrak{K}(p)$ of M at p satisfies the inequality*

$$|\mathfrak{K}(p)|d(p)^2 \le \frac{4K^2\csc^2\alpha}{(K^2+1)(1-Q_K^2)^{3/2}} \left[2(K^2+1)\csc^2\alpha - 2 + \frac{(K^2-1)N_K}{K} \right]. \tag{32}$$

Proof. After a rotation we assume that the normals make an angle of at least α with the x^3-axis. Let $\tilde{M}$ be the universal covering surface of M under a universal covering transformation $z(\zeta)$. Suppose that the point $\zeta = 0$ in $\tilde{M}$ corresponds to p in M. For a conformal metric ds_z on M, we have $ds_z = \rho(z)|dz| = \rho(z(\zeta))|dz/d\zeta||d\zeta|$, which implies that

$$\mathfrak{K}_{\tilde{M}}(\zeta) = -\frac{\Delta_\zeta \log(\rho(z(\zeta))|dz/d\zeta|)}{(\rho(z(\zeta))|dz/d\zeta|)^2} = -\frac{\Delta_z \log\rho}{\rho^2} \circ z(\zeta) = \mathfrak{K}_M(z(\zeta)).$$

If we form the functions $\psi_k(\zeta)$ by (28), then from the relation (31), we have

$$\sum_{k=1}^{2} |\psi_k(\zeta)|^2 \le (K^2+1)\csc^2\alpha - 1. \tag{33}$$

Thus we have bounded holomorphic functions ψ_k, $k = 1, 2$, on $\tilde{M}$. If $\tilde{M}$ is the entire ζ-plane, Liouville's theorem says that ψ_k, $k = 1, 2$ are constants. Then the relation (29) shows that the Gauss curvature $\mathfrak{K} \equiv 0$. Thus, the relation (32) holds automatically. Now we need to consider the case that $\tilde{M}$ is the unit disk $|\zeta| < 1$. For convenience, we adopt the following notation:

$$C_k = |\psi_k(0)|, \quad D_k = |\psi_k'(0)|, \quad M_k = \sup_{|w|<1} |\psi_k(w)|. \tag{34}$$

From the application of the Schwarz-pick lemma to the function ψ_k/M_k, we obtain

$$D_k \le M_k \left(1 - \frac{C_k^2}{M_k^2}\right) = M_k \eta_k, \tag{35}$$

where

$$\eta_k = 1 - \frac{C_k^2}{M_k^2}. \tag{36}$$

Utilizing the relation (29) at the point $w = 0$ in $\tilde{M}$, we next give an estimate of Gauss curvature $\mathfrak{K}$ at a given point $p \in M$. By (33), we get

$$C_k^2 \le M_k^2 \le (K^2+1)\csc^2\alpha - 1, \tag{37}$$

and

$$D_k^2 \le ((K^2+1)\csc^2\alpha - 1)\eta_k^2. \tag{38}$$

By the Cauchy–Schwarz inequality, we have

$$|\psi_1(0)\psi_2'(0) - \psi_1'(0)\psi_2(0)|^2 = (C_1D_2 + C_2D_1)^2 \le \sum_{j=1}^{2} C_j^2 \sum_{k=1}^{2} D_k^2.$$

Thus, the combination of (37)–(38) and the above inequality yields

$$\begin{aligned}|\psi_1(0)\psi_2'(0) - \psi_2'(0)\psi_1(0)|^2 + \sum_{j=1}^{2} |\psi_j'(0)|^2 \\ &\le \sum_{k=1}^{2} D_k^2 \Big(1 + \sum_{j=1}^{2} C_j^2\Big) \\ &\le ((K^2+1)\csc^2\alpha - 1)\Big(1 + \sum_{j=1}^{2} C_j^2\Big) \sum_{j=1}^{2} \eta_j^2.\end{aligned}$$

Hence,

$$\frac{\|\phi\|^2\|\phi'\|^2 - (\phi'\cdot\overline{\phi})(\overline{\phi'}\cdot\phi)}{\|\phi\|^6} \le \frac{((K^2+1)\csc^2\alpha - 1)\sum_{j=1}^{2}\eta_j^2}{|\phi_3(0)|^2(1+\sum_{j=1}^{2} C_j^2)^2}.$$

Furthermore, it follows from (26) and Lemma 5.1 that

$$\begin{aligned}|\mathfrak{K}(0)| &\le \frac{1}{\big(\sqrt{1-Q_K^2}\big)^3}\left[\frac{\|\phi\|^2\|\phi'\|^2 - (\phi'\cdot\overline{\phi})(\overline{\phi'}\cdot\phi)}{\|\phi\|^6} + \frac{K^2-1}{K}\frac{|(\phi'\cdot\phi)(\overline{\phi'}\cdot\phi)|}{\|\phi\|^6}\right] \\ &\le \frac{1}{\big(\sqrt{1-Q_K^2}\big)^3}\left[\frac{2(K^2+1)\csc^2\alpha - 2}{|\phi_3(0)|^2} + \frac{(K^2-1)}{K}\frac{|(\phi'\cdot\phi)(\overline{\phi'}\cdot\phi)|}{\|\phi\|^6}\right] \\ &\le \frac{1}{\big(\sqrt{1-Q_K^2}\big)^3|\phi_3(0)|^2}\left[2(K^2+1)\csc^2\alpha - 2 + \frac{(K^2-1)N_K}{K}\right] \\ &\le \frac{4K^2\csc^2\alpha}{(K^2+1)\big(\sqrt{1-Q_K^2}\big)^3 d^2(0)}\left[2(K^2+1)\csc^2\alpha - 2 + \frac{(K^2-1)N_K}{K}\right]. \qquad \square\end{aligned}$$

Acknowledgements

This work was done when Chen and Liu visited the University of Houston. They wish to thank the University of Houston for its hospitality. Chen was supported by the China Scholarships Council (#201808350049), NNSF of China (#11971182), NSF of Fujian Province of China (#2019J01066). Liu was supported by the China Scholarships Council (#201806360222). Ru was supported in part by a Simons Foundation grant award (#531604).

References

[Alarcón and López 2013] A. Alarcón and F. J. López, "On harmonic quasiconformal immersions of surfaces in $\mathbb{R}^3$", *Trans. Amer. Math. Soc.* **365**:4 (2013), 1711–1742. MR Zbl

[Chen 1987] W. Chen, *Cartan's conjecture: defect relations for meromorphic maps from parabolic manifold to projective space*, Ph.D. thesis, University of Notre Dame, 1987.

[Chern 1965] S.-s. Chern, "Minimal surfaces in an Euclidean space of N dimensions", pp. 187–198 in *Differential and combinatorial topology* (Princeton, 1961), edited by S. S. Cairns, Princeton Univ. Press, 1965. MR Zbl

[Chern and Osserman 1967] S.-s. Chern and R. Osserman, "Complete minimal surfaces in Euclidean n-space", *J. Anal. Math.* **19** (1967), 15–34. MR Zbl

[Connor et al. 2015] P. Connor, K. Li, and M. Weber, "Complete embedded harmonic surfaces in $\mathbb{R}^3$", *Exp. Math.* **24**:2 (2015), 196–224. MR Zbl

[Connor et al. 2018] P. Connor, K. Li, and M. Weber, "The Gauss–Bonnet formula for harmonic surfaces", *Comm. Anal. Geom.* **26**:3 (2018), 531–570. MR Zbl

[Dioos and Sakaki 2019] B. Dioos and M. Sakaki, "A representation formula for non-conformal harmonic surfaces in $\mathbb{R}^3$", *Results Math.* **74**:1 (2019), art. id. 35. MR Zbl

[Fujimoto 1990] H. Fujimoto, "Modified defect relations for the Gauss map of minimal surfaces, II", *J. Differential Geom.* **31**:2 (1990), 365–385. MR Zbl

[Fujimoto 1993] H. Fujimoto, *Value distribution theory of the Gauss map of minimal surfaces in $\mathbb{R}^m$*, Aspects Math. **E21**, Vieweg, Braunschweig, Germany, 1993. MR Zbl

[Jensen and Rigoli 1988] G. R. Jensen and M. Rigoli, "Harmonically immersed surfaces of $\mathbb{R}^n$", *Trans. Amer. Math. Soc.* **307**:1 (1988), 363–372. Correction in **311**:1 (1989), 425–428. MR Zbl

[Kalaj 2013] D. Kalaj, "The Gauss map of a harmonic surface", *Indag. Math.* (*N.S.*) **24**:2 (2013), 415–427. MR Zbl

[Meeks and Rosenberg 2005] W. H. Meeks, III and H. Rosenberg, "The uniqueness of the helicoid", *Ann. of Math.* (2) **161**:2 (2005), 727–758. MR Zbl

[Milnor 1967] T. Klotz, "Surfaces harmonically immersed in E^3", *Pacific J. Math.* **21**:1 (1967), 79–87. MR Zbl

[Milnor 1968] T. Klotz, "A complete R_Λ-harmonically immersed surface in E^3 on which $H \neq 0$", *Proc. Amer. Math. Soc.* **19** (1968), 1296–1298. MR Zbl

[Milnor 1976] T. K. Milnor, "Restrictions on the curvatures of Φ-bounded surfaces", *J. Differential Geom.* **11**:1 (1976), 31–46. MR Zbl

[Milnor 1979] T. K. Milnor, "Harmonically immersed surfaces", *J. Differential Geom.* **14**:2 (1979), 205–214. MR Zbl

[Milnor 1980] T. K. Milnor, "Mapping surfaces harmonically into E^n", *Proc. Amer. Math. Soc.* **78**:2 (1980), 269–275. MR Zbl

[Milnor 1983] T. K. Milnor, "Are harmonically immersed surfaces at all like minimally immersed surfaces?", pp. 99–110 in *Seminar on minimal submanifolds*, edited by E. Bombieri, Ann. of Math. Stud. **103**, Princeton Univ. Press, 1983. MR Zbl

[Nochka 1983] E. I. Nochka, "On the theory of meromorphic curves", *Dokl. Akad. Nauk SSSR* **269**:3 (1983), 547–552. In Russian; translated in *Soviet Math. Dokl.* **27**:2 (1983), 377–381. MR Zbl

[Osserman 1964] R. Osserman, "Global properties of minimal surfaces in E^3 and E^n", *Ann. of Math.* (2) **80** (1964), 340–364. MR Zbl

[Osserman and Ru 1997] R. Osserman and M. Ru, "An estimate for the Gauss curvature of minimal surfaces in $\mathbb{R}^m$ whose Gauss map omits a set of hyperplanes", *J. Differential Geom.* **46**:3 (1997), 578–593. MR Zbl

[Ru 1991] M. Ru, "On the Gauss map of minimal surfaces immersed in $\mathbb{R}^n$", *J. Differential Geom.* **34**:2 (1991), 411–423. MR Zbl

[Xavier 1981] F. Xavier, "The Gauss map of a complete nonflat minimal surface cannot omit 7 points of the sphere", *Ann. of Math.* (2) **113**:1 (1981), 211–214. MR Zbl

Received December 16, 2019. Revised March 5, 2020.

XINGDI CHEN
SCHOOL OF MATHEMATICAL SCIENCES
HUAQIAO UNIVERSITY
QUANZHOU
CHINA
chxtt@hqu.edu.cn

ZHIXUE LIU
DEPARTMENT OF MATHEMATICS, SCHOOL OF SCIENCES
BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS
BEIJING
CHINA
zxliumath@163.com

MIN RU
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF HOUSTON
HOUSTON, TX
UNITED STATES
minru@math.uh.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 2, 2020
https://doi.org/10.2140/pjm.2020.309.289

SCATTERED REPRESENTATIONS OF $SL(n, \mathbb{C})$

CHAO-PING DONG AND KAYUE DANIEL WONG

Let G be $SL(n, \mathbb{C})$. The unitary dual $\widehat{G}$ was classified by Vogan in the 1980s. This paper aims to describe the Zhelobenko parameters and the spin-lowest K-types of the scattered representations of G, which lie at the heart of $\widehat{G}^d$—the set of all the equivalence classes of irreducible unitary representations of G with nonvanishing Dirac cohomology. As a consequence, we will verify a couple of conjectures of Dong for G.

1. Introduction

1.1. ***Preliminaries on complex simple Lie groups.*** Let G be a complex connected simple Lie group, and H be a Cartan subgroup of G. Let $\mathfrak{g}_0$ and $\mathfrak{h}_0$ be the Lie algebra of G and H respectively, and we drop the subscripts to stand for the complexified Lie algebras. We adopt a positive root system $\Delta^+(\mathfrak{g}_0, \mathfrak{h}_0)$, and let $\varpi_1, \dots, \varpi_{\mathrm{rank}(\mathfrak{g}_0)}$ be the corresponding fundamental weights with $\rho = \varpi_1 + \cdots + \varpi_{\mathrm{rank}(\mathfrak{g}_0)}$ being the half sum of positive roots.

Fix a Cartan involution θ on G such that its fixed points form a maximal compact subgroup K of G. Then on the Lie algebra level, we have the Cartan decomposition

$$\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0.$$

We denote by $\langle \cdot, \cdot \rangle$ the Killing form on $\mathfrak{g}_0$. This form is negative definite on $\mathfrak{k}_0$ and positive definite on $\mathfrak{p}_0$. Moreover, $\mathfrak{k}_0$ and $\mathfrak{p}_0$ are orthogonal to each other under $\langle \cdot, \cdot \rangle$. We shall denote by $\| \cdot \|$ the norm corresponding to the Killing form.

Let $H = TA$ be the Cartan decomposition of H, with $\mathfrak{h}_0 = \mathfrak{t}_0 + \mathfrak{a}_0$. We make the following identifications:

$$\mathfrak{h} \cong \mathfrak{h}_0 \times \mathfrak{h}_0, \quad \mathfrak{t} = \{(x, -x) : x \in \mathfrak{h}_0\}, \quad \mathfrak{a} \cong \{(x, x) : x \in \mathfrak{h}_0\}. \tag{1}$$

Take an arbitrary pair $(\lambda_L, \lambda_R) \in \mathfrak{h}_0^* \times \mathfrak{h}_0^*$ such that $\mu := \lambda_L - \lambda_R$ is integral. Denote by $\{\mu\}$ the unique dominant weight to which μ is conjugate under the action of

MSC2010: primary 22E46; secondary 17B56.
Keywords: Dirac cohomology, unitary representations, scattered representations.

the Weyl group W. Write $\nu := \lambda_L + \lambda_R$. We can view μ as a weight of T and ν a character of A. Put

$$I(\lambda_L, \lambda_R) := \mathrm{Ind}_B^G(\mathbb{C}_\mu \otimes \mathbb{C}_\nu \otimes \mathbf{1})_{K\text{-finite}},$$

where B is the Borel subgroup of G determined by $\Delta^+(\mathfrak{g}_0, \mathfrak{h}_0)$. It is not hard to show that $V_{\{\mu\}}$, the K-type with highest weight $\{\mu\}$, occurs exactly once in $I(\lambda_L, \lambda_R)$. Let $J(\lambda_L, \lambda_R)$ be the unique irreducible subquotient of $I(\lambda_L, \lambda_R)$ containing $V_{\{\mu\}}$. By [Zhelobenko 1974], every irreducible admissible $(\mathfrak{g}, K)$-module has the form $J(\lambda_L, \lambda_R)$. Indeed, $J(\lambda_L, \lambda_R)$ has infinitesimal character the $W \times W$ orbit of (λ_L, λ_R), and lowest K-type $V_{\{\lambda_L - \lambda_R\}}$. We will refer to the pair (λ_L, λ_R) as the *Zhelobenko parameter* for the module $J(\lambda_L, \lambda_R)$.

1.2. *Dirac cohomology.* Fix an orthonormal basis $Z_1, \dots, Z_l$ of $\mathfrak{p}_0$ with respect to the inner product on $\mathfrak{p}_0$ induced by $\langle \cdot, \cdot \rangle$. Let $U(\mathfrak{g})$ be the universal enveloping algebra of $\mathfrak{g}$, and put $C(\mathfrak{p})$ as the Clifford algebra of $\mathfrak{p}$. One checks that

$$D := \sum_{i=1}^{l} Z_i \otimes Z_i \in U(\mathfrak{g}) \otimes C(\mathfrak{p}) \tag{2}$$

is independent of the choice of the orthonormal basis $Z_1, \dots, Z_l$. The operator D, called the *Dirac operator*, was introduced by Parthasarathy [1972]. By construction, D^2 is a natural Laplacian on G, which gives rise to the Parthasarathy's Dirac inequality (see (6) below). The inequality is very effective for detecting nonunitarity of $(\mathfrak{g}, K)$-modules, but is by no means sufficient to classify all (non)unitary modules.

To sharpen the Dirac inequality, and to offer a better understanding of the unitary dual, Vogan [1997] formulated the notion of Dirac cohomology. Let $\mathrm{Ad} : K \to SO(\mathfrak{p}_0)$ be the adjoint map, $\mathrm{Spin}\,\mathfrak{p}_0$ be the spin group of $\mathfrak{p}_0$, and denote by $p : \mathrm{Spin}\,\mathfrak{p}_0 \to SO(\mathfrak{p}_0)$ the spin double covering map. Put

$$\widetilde{K} := \{(k, s) \in K \times \mathrm{Spin}\,\mathfrak{p}_0 \mid \mathrm{Ad}(k) = p(s)\}.$$

As in the case of K-types, we will refer to an irreducible $\widetilde{K}$-type with highest weight δ as V_δ.

Let π be any admissible $(\mathfrak{g}, K)$-module, and S be the spin module of $C(\mathfrak{p})$. Then $U(\mathfrak{g}) \otimes C(\mathfrak{p})$, in particular the Dirac operator D, acts on $\pi \otimes S$. Now the *Dirac cohomology* is defined as the $\widetilde{K}$-module

$$H_D(\pi) := \mathrm{Ker}\, D / (\mathrm{Ker}\, D \cap \mathrm{Im}\, D). \tag{3}$$

It is evident from the definition that Dirac cohomology is an invariant for admissible $(\mathfrak{g}, K)$-modules. To compute this invariant, the Vogan conjecture, proved by Huang and Pandžić [2002], says that whenever $H_D(\pi) \neq 0$, one would have

$$\gamma + \rho = w\Lambda, \tag{4}$$

where Λ is the infinitesimal character of π, γ is the highest weight of any $\widetilde{K}$-type in $H_D(\pi)$, and w is some element of W.

It turns out that many interesting $(\mathfrak{g}, K)$-modules π, such as some $A_q(\lambda)$-modules and all the highest weight modules, have nonzero Dirac cohomology (see [Huang et al. 2009; 2011]). One would therefore like to classify all representations with nonzero Dirac cohomology.

1.3. ***Spin-lowest K-type.*** From now on, we set π as an irreducible unitary $(\mathfrak{g}, K)$-module with infinitesimal character Λ. In order to get a clearer picture on $H_D(\pi)$, the first-named author introduced the notion of spin-lowest K-types. Given an arbitrary K-type V_δ, its spin norm is defined as

$$\|\delta\|_{\mathrm{spin}} := \|\{\delta - \rho\} + \rho\|. \tag{5}$$

Then a K-type V_τ occurring in π is called a *spin-lowest K-type* of π if it achieves the minimum spin norm among all the K-types showing up in π.

As an application of spin-lowest K-type, note that D is self-adjoint on the unitarizable module $\pi \otimes S$. By writing out D^2 carefully, and by using the *PRV-component* [Parthasarathy et al. 1967], we can rephrase *Parthasarathy's Dirac operator inequality* [Parthasarathy 1980] as follows:

$$\|\delta\|_{\mathrm{spin}} \geq \|\Lambda\|, \tag{6}$$

where V_δ is any K-type occurring in π. Moreover, one can deduce from [Huang and Pandžić 2006, Theorem 3.5.3] that $H_D(\pi) \neq 0$ if and only if the spin-lowest K-types V_τ attain the lower bound of (6). In such cases, $V_{\{\tau-\rho\}}$ will show up in $H_D(\pi)$. Put in a different way, the spin-lowest K-types of π are exactly the K-types contributing to $H_D(\pi)$ whenever the cohomology is nonvanishing (see [Dong 2013, Proposition 2.3] for more details).

1.4. ***Scattered representations.*** Based on the studies [Barbasch and Pandžić 2011; Ding and Dong 2020], we are interested in the irreducible unitarizable $(\mathfrak{g}, K)$-modules $J(\lambda, -s\lambda)$ such that

(i) the weight 2λ is dominant integral, i.e., $2\lambda = \sum_{i=1}^{\mathrm{rank}(\mathfrak{g}_0)} c_i \varpi_i$, where each c_i is a positive integer;

(ii) the element $s \in W$ is an involution such that each simple reflection s_i, $1 \leq i \leq \mathrm{rank}(\mathfrak{g}_0)$, occurs in one (thus in each) reduced expression of s;

(iii) the module has nonzero Dirac cohomology, i.e., $H_D(J(\lambda, -s\lambda)) \neq 0$, or equivalently, there exists a K-type V_τ in $J(\lambda, -s\lambda)$ such that

$$\|\tau\|_{\text{spin}} = \|(\lambda, -s\lambda)\| = \|2\lambda\| \tag{7}$$

According to [Ding and Dong 2020], there are only finitely many such representations, which are called the *scattered representations*.

These representations lie at the heart of $\widehat{G}^d$ — the set of all the irreducible unitary $(\mathfrak{g}, K)$-modules of G with nonzero Dirac cohomology up to equivalence. Namely, by [Ding and Dong 2020, Theorem A], any member of $\widehat{G}^d$ is either a scattered representation, or it is cohomologically induced from a scattered representation tensored with a suitable unitary character of the Levi factor of a certain proper θ-stable parabolic subgroup. In the latter case, one can easily trace the spin-lowest K-types along with the Dirac cohomology of the modules before and after induction. It is therefore of interest to have a good understanding of scattered representations.

1.5. *Overview.* In this manuscript, we focus on Lie groups G of Type A. For convenience, we will start from the group $GL(n, \mathbb{C})$, written as $GL(n)$ for short. In this case, Vogan classified the unitary dual. The part that we need can be described as follows.

Theorem 1.1 [Vogan 1986]. *All irreducible unitary representations of $GL(n)$ with regular half-integral infinitesimal characters are parabolically induced from a unitary character, i.e., they are of the form*

$$\mathrm{Ind}_{(\prod_{i=0}^{m} GL(a_i))U}^{GL(n)} \left(\bigotimes_{i=0}^{m} \det{}^{p_i} \otimes \mathbf{1} \right)$$

for some $a_i \in \mathbb{Z}_{>0}$ and $p_i \in \mathbb{Z}$. For simplicity, we will write the parabolically induced module $\mathrm{Ind}_{LU}^{G}(\pi \otimes \mathbf{1})$ as $\mathrm{Ind}_{L}^{G}(\pi)$ for the rest of the manuscript.

Using [Barbasch and Pandžić 2011, Theorem 2.4], all such π have nonzero Dirac cohomology. Moreover, [Barbasch et al. 2020] proved [Barbasch and Pandžić 2011, Conjecture 4.1], which says

$$H_D(\pi) = 2^{[\mathrm{rank}(\mathfrak{g}_0)/2]} V_{\{\tau - \rho\}},$$

where V_τ is the *unique* spin-lowest K-type appearing in π *with multiplicity one*. However, it is not clear what V_τ is like from the calculations in [Barbasch et al. 2020].

In Section 2, we will give an algorithm to compute V_τ for all such π (see Proposition 2.5). In Section 3, we will see how the calculations for $GL(n)$ in Section 2 can be translated to $SL(n)$, which gives a combinatorial description of scattered representations of $SL(n)$ (Proposition 3.1). As a result, we prove the following:

- The spin-lowest K-type of each scattered representation of $SL(n)$ is *unitarily small* in the sense of Salamanca-Riba and Vogan [1998] (Corollary 3.5); and
- the number of scattered representations of $SL(n)$ is equal to 2^{n-2} (Corollary 3.9).

This verifies [Ding and Dong 2020, Conjecture C] in the case of $SL(n)$, and proves [Dong 2019, Conjecture 5.2] respectively.

It is worth noting that for any nontrivial scattered representation, its spin-lowest K-type lives deeper than, and differs from the lowest K-type. We hope the effort here will shed some light on the real case in future.

2. An algorithm computing the spin-lowest K-types

In this section, we give an algorithm to find the spin-lowest K-types of the irreducible unitary modules of $GL(n)$ given by Theorem 1.1. We use a *chain*

$$\mathcal{C} := \{c, c-2, \ldots, c-(2k-2), c-2k\},$$

where $c, k \in \mathbb{Z}$ with $k > 0$, to denote the Zhelobenko parameter

$$\begin{pmatrix} \lambda \\ -w_0\lambda \end{pmatrix} = \begin{pmatrix} \frac{1}{2}c & \frac{1}{2}c-1 & \ldots & \frac{1}{2}c-(k-1) & \frac{1}{2}c-k \\ -\frac{1}{2}c+k & -\frac{1}{2}c+(k-1) & \ldots & -\frac{1}{2}c+1 & -\frac{1}{2}c \end{pmatrix}.$$

Note that the entries of $\mathcal{C}$ are precisely equal to 2λ. Also, this parameter corresponds to the one-dimensional module $\det^{c-k}$ of $GL(k+1)$. Consequently, Theorem 1.1 implies that the Zhelobenko parameters of all irreducible unitary modules with regular half-integral infinitesimal character can be expressed by the chains

$$(\lambda, -s\lambda) = \bigcup_{i=0}^{m} \mathcal{C}_i,$$

where all the entries of $\mathcal{C}_i$ are disjoint.

In order to understand the spin-lowest K-types of these modules of $GL(n)$, we make the following:

Definition 2.1. (a) Two chains $\mathcal{C}_1 = \{A, \ldots, a\}$, $\mathcal{C}_2 = \{B, \ldots, b\}$ are *linked* if the entries of $\mathcal{C}_1$ and $\mathcal{C}_2$ are disjoint satisfying

$$A > B > a \quad \text{or} \quad B > A > b.$$

(b) We say a union of chains $\bigcup_{i \in I} \mathcal{C}_i$ is *interlaced* if for all $i \neq j$ in I, there exist indices $i = i_0, i_1, \ldots, i_m = j$ in I such that $\mathcal{C}_{i_{l-1}}$ and $\mathcal{C}_{i_l}$ are linked for all $1 \leq l \leq m$. (By convention, we also let the single chain $\mathcal{C}_1$ be interlaced).

For example, the parameter $\{9, 7, 5\} \cup \{6, 4, 2\} \cup \{3, 1\}$ is interlaced, while the parameter $\{10, 8\} \cup \{9, 7\} \cup \{6, 4\} \cup \{5, 3, 1\}$ is not interlaced.

We are now in the position to describe the spin-lowest K-types of the unitary modules in Theorem 1.1 using chains.

Algorithm 2.2. Let $J(\lambda, -s\lambda)$ be an irreducible unitary module of $GL(n)$ in Theorem 1.1 with $(\lambda, -s\lambda) = \bigcup_{i=0}^{m} \mathcal{C}_i$, where

$$\mathcal{C}_i := \{k_i + (d_i - 1), \dots, k_i - (d_i - 1)\} = \{C_{i,1}, \dots, C_{i,d_i}\}$$

is a chain with average value k_i and length d_i. Then the lowest K-type is equal to (a W-conjugate of) $(\mathcal{T}_0, \dots, \mathcal{T}_m)$, where

$$\mathcal{T}_i := (\underbrace{k_i, \dots, k_i}_{d_i}).$$

By reindexing the chains when necessary, we may and we will assume that

(8) $$\text{for any } 0 \le i < j \le m, \quad k_i > k_j \text{ or } d_i < d_j \text{ if } k_i = k_j.$$

Let us change the coordinates of $\mathcal{T}_i$ and $\mathcal{T}_j$ for all pairs of linked chains $\mathcal{C}_i$ and $\mathcal{C}_j$ such that $i < j$ by the following rule:

(a) If $C_{i,1} > C_{j,1} \ge C_{j,d_j} > C_{i,d_i}$, i.e.,

$$\begin{array}{l} \{C_{i,1}, \dots, C_{i,d_i-p}, \overbrace{C_{i,d_i-p+1}, \dots, C_{i,d_i}}^{p}\}, \\ \qquad\qquad\qquad\qquad \{C_{j,1}, \dots, C_{j,d_j}\}, \end{array}$$

with $C_{j,1} = C_{i,d_i} + 2p - 1$ and $d_j \le p$, then we change the coordinates of $\mathcal{T}_i$ and $\mathcal{T}_j$ into

$$\boxed{\begin{array}{ll} \mathcal{T}_i' : & (*, \dots, *, \overbrace{k_i + p, k_i + (p-1), \dots, k_i + (p - d_j + 1), *, \dots, *}^{p}) \\ \mathcal{T}_j' : & \qquad\qquad (k_j - p, k_j - (p-1), \dots, k_j - (p - d_j + 1)), \end{array}}$$

where the entries marked by $*$ remain unchanged.

(b) If $C_{i,1} > C_{j,1} > C_{i,d_i} > C_{j,d_j}$, i.e.,

$$\begin{array}{l} \{C_{i,1}, \quad \dots, \quad C_{i,d_i-p}, \quad \overbrace{C_{i,d_i-p+1}, \quad \dots, \quad C_{i,d_i}}^{p}\} \\ \qquad\qquad\qquad\qquad \{C_{j,1}, \quad \dots\dots, \quad C_{j,p}, \quad C_{j,p+1}, \quad \dots, \quad C_{j,d_j}\} \end{array}$$

with $C_{j,1} = C_{i,d_i} + 2p - 1$ and $d_j > p$, then we change the coordinates of $\mathcal{T}_i$ and $\mathcal{T}_j$ into

$$\boxed{\begin{array}{ll} \mathcal{T}_i' : (*, \ldots, *, \overbrace{k_i+1, \ \ldots, \ k_i+p}^{p}) & \\ \mathcal{T}_j' : \qquad (k_j-1, \ \ldots, \ k_j-p, \ *, \ \ldots, \ *). & \end{array}}$$

where the entries marked by $*$ remain unchanged.

(c) If $C_{j,1} > C_{i,1} > C_{j,d_j}$, then since $k_i \geq k_j$ one also have $C_{j,1} > C_{i,1} \geq C_{i,d_i} > C_{j,d_j}$, i.e.,

$$\begin{array}{l} \qquad\quad \{C_{i,1}, \ \ldots, \ C_{i,d_i}\} \\ \{\underbrace{C_{j,1}, \quad \ldots, \quad C_{j,q}}_{q}, \quad C_{j,q+1}, \ \ldots, \ C_{j,d_j}\} \end{array}$$

with $C_{j,1} = C_{i,d_i} + 2q - 1$, then we change the coordinates of $\mathcal{T}_i$ and $\mathcal{T}_j$ into

$$\boxed{\begin{array}{l} \mathcal{T}_i' : \qquad\qquad (k_i+(q-d_0+1), \ \ldots, \ k_i+(q-1), \ k_i+q) \\ \mathcal{T}_j' : (\underbrace{*, \ \ldots, \ *, \ k_j-(q-d_0+1), \ \ldots, \ k_j-(q-1), \ k_j-q}_{q}, \ *, \ \ldots, \ *), \end{array}}$$

where the entries marked by $*$ remain unchanged.

In the above three cases, we only demonstrate the situation that $\mathcal{C}_i$ is in the first row and $\mathcal{C}_j$ is in the second row. The rule is the same when $\mathcal{C}_j$ is in the first row while $\mathcal{C}_i$ is in the second row.

After running through all pairs of linked chains, V_τ is defined as the K-type with highest weight τ given by (a W-conjugate of) $\bigcup_{i=0}^m \mathcal{T}_i'$.

Example 2.3. Consider

$$(\lambda, -s\lambda) = \begin{array}{l} \{10 \quad 8\} \ \{6\} \ \{4\} \\ \quad \{9 \quad 7 \quad 5 \quad 3 \quad 1\}. \end{array}$$

Then the lowest K-type of $J(\lambda, -s\lambda)$ is

$$\begin{array}{l} (9 \quad 9) \ (6) \ (4) \\ \quad (5 \quad 5 \quad 5 \quad 5 \quad 5). \end{array}$$

To compute V_τ, let us label the chains so that (8) holds:

$$\mathcal{T}_0 = (9 \ 9), \quad \mathcal{T}_1 = (6), \quad \mathcal{T}_2 = (5 \ 5 \ 5 \ 5 \ 5), \quad \mathcal{T}_3 = (4).$$

Then we apply (a) to the pair $\mathcal{T}_2$, $\mathcal{T}_3$, apply (b) to the pair $\mathcal{T}_0$, $\mathcal{T}_2$, and apply (c) to the pair $\mathcal{T}_1$, $\mathcal{T}_2$. This gives us

$$\begin{matrix} (9 & 10) & (8) & (2) & \\ & (4 & 3 & 5 & 7 & 5). \end{matrix}$$

Thus $\tau = (10, 9, 8, 7, 5, 5, 4, 3, 2)$.

Theorem 2.4. *Let $J(\lambda, -s\lambda)$ be a unitary module of $GL(n)$ in Theorem 1.1, and V_τ be obtained by Algorithm 2.2. Then $[J(\lambda, -s\lambda) : V_\tau] > 0$.*

Proof. Let

$$J(\lambda, -s\lambda) = \mathrm{Ind}_{\prod_{i=0}^m GL(a_i)}^{GL(n)} \left(\bigotimes_{i=0}^m V_{(k_i,\dots,k_i)} \right).$$

By rearranging the Levi factors, one can assume the chains $\mathcal{C}_0, \dots, \mathcal{C}_m$ satisfy Equation (8). We are interested in studying

$$\begin{aligned} &\left[\mathrm{Ind}_{\prod_{i=0}^m GL(a_i)}^{GL(n)} \left(\bigotimes_{i=0}^m V_{(k_i,\dots,k_i)} \right) : V_\tau \right] \\ &= \left[\bigotimes_{i=0}^m V_{(k_i,\dots,k_i)} : V_\tau|_{\prod_{i=0}^m GL(a_i)} \right] \\ &= \left[\bigotimes_{i=0}^m V_{(k_i+t,\dots,k_i+t)} : V_\tau|_{\prod_{i=0}^m GL(a_i)} \otimes \bigotimes_{i=0}^m V_{(t,\dots,t)} \right] \\ &= \left[\bigotimes_{i=0}^m V_{(k_i+t,\dots,k_i+t)} : V_\tau|_{\prod_{i=0}^m GL(a_i)} \otimes V_{(t,\dots,t)}|_{\prod_{i=1}^m GL(a_i)} \right] \\ &= \left[\bigotimes_{i=0}^m V_{(k_i+t,\dots,k_i+t)} : V_{\tau+(t,\dots,t)}|_{\prod_{i=0}^m GL(a_i)} \right] \\ &= \left[\mathrm{Ind}_{\prod_{i=0}^m GL(a_i)}^{GL(n)} \left(\bigotimes_{i=0}^m V_{(k_i+t,\dots,k_i+t)} \right) : V_{\tau+(t,\dots,t)} \right] \end{aligned}$$

So we can assume $k_i > 0$ for all i without loss of generality.

We prove the theorem by induction on the number of Levi components. The theorem obviously holds when there is only one Levi component — the irreducible module is a unitary character of $GL(n)$. Now suppose we have the hypothesis holds when there are m Levi factors, i.e.,

$$\left[\mathrm{Ind}_{\prod_{i=0}^{m-1} GL(a_i)}^{GL(n')} \left(\bigotimes_{i=0}^{m-1} V_{(k_i,\dots,k_i)} \right) : V_{\tau_{m-1}} \right] > 0,$$

where $n' = n - a_m$, and τ_{m-1} is obtained by applying Algorithm 2.2 on $\bigcup_{i=0}^{m-1} \mathcal{C}_i$. Suppose now τ_m is obtained by applying Algorithm 2.2 on $\bigcup_{i=0}^{m} \mathcal{C}_i$. Then

$$\begin{aligned}\left[\mathrm{Ind}_{\prod_{i=0}^{m} GL(a_i)}^{GL(n)}\left(\bigotimes_{i=0}^{m} V_{(k_i,\ldots,k_i)}\right) : V_{\tau_m}\right] &= \left[\mathrm{Ind}_{GL(n')\times GL(a_m)}^{GL(n)}\left(\mathrm{Ind}_{\prod_{i=0}^{m-1} GL(a_i)}^{GL(n')}\left(\bigotimes_{i=0}^{m-1} V_{(k_i,\ldots,k_i)}\right)\otimes V_{(k_m,\ldots,k_m)}\right) : V_{\tau_m}\right]\\ &\geq [\mathrm{Ind}_{GL(n')\times GL(a_m)}^{GL(n)}(V_{\tau_{m-1}}\otimes V_{(k_m,\ldots,k_m)}) : V_{\tau_m}]\\ &= c^{\tau_m}_{\tau_{m-1},(k_m,\ldots,k_m)}\end{aligned}$$

Here $c^{\lambda}_{\mu,\nu}$ is the Littlewood–Richardson coefficient, and the last step uses [Goodman and Wallach 2009, Theorem 9.2.3].

Suppose $\tau_{m-1} = \bigcup_{i=0}^{m-1} \mathcal{T}_i''$. Here these $\mathcal{T}_i''$ are obtained by applying Algorithm 2.2 on $\mathcal{C}_0, \ldots, \mathcal{C}_{m-1}$. Then τ_m is obtained from applying Algorithm 2.2 on $\mathcal{T}_i''$ and $\mathcal{T}_m = (k_m, \ldots, k_m)$ for all linked $\mathcal{C}_i$ and $\mathcal{C}_m$. More precisely, by applying Rules (a)–(c) in Algorithm 2.2, τ_m is obtained from τ_{m-1} by the following:

(i) Construct a new partition $\tau_{m-1} \cup (k_m, \ldots, k_m)$.

(ii) For each linked $\mathcal{C}_i$ and $\mathcal{C}_m$, add $(0,\ldots,0,A,A-1,\ldots,a+1,a,0,\ldots,0)$ on the rows of τ_{m-1} corresponding to $\mathcal{T}_i''$, and subtract $(0,\ldots,0,A,A-1,\ldots,a+1,a,0,\ldots,0)$ on the corresponding rows of $(k_m,\ldots,k_m)$.

(iii) τ_m is obtained by going through (ii) for all $\mathcal{C}_i$ linked with $\mathcal{C}_m$.

By the above construction of τ_m, it follows from the Littlewood–Richardson Rule as stated in [Goodman and Wallach 2009, p. 420] that

$$c^{\tau_m}_{\tau_{m-1},(k_m,\ldots,k_m)} \geq 1. \tag{9}$$

Indeed, it suffices to find *one L-R skew tableaux* of shape τ_m/τ_{m-1} and weight

$$(\underbrace{k_m, \ldots, k_m}_{d_m})$$

in the sense of [Goodman and Wallach 2009, Definition 9.3.17]. Recall that d_m is the number of entries of the chain $\mathcal{C}_m$.

To do so, we first describe the Ferrers diagram τ_m/τ_{m-1}. Suppose $\mathcal{C}_{i_1}, \ldots, \mathcal{C}_{i_l}$ are linked to $\mathcal{C}_m$ with $i_1 > \cdots > i_l$. By Step (ii) of the above algorithm, we add $(A_j, A_j - 1, \ldots, a_j + 1, a_j)$ to the rows in τ_{m-1} corresponding to the chains $\mathcal{C}_{i_j}$. Note that by our ordering of the chains, we must have

$$A_l > \cdots > a_l > A_{l-1} > \cdots > a_{l-1} > \cdots > A_1 > \cdots > a_1.$$

The rows of the Ferrers diagram τ_m/τ_{m-1} have lengths

$$(10)\quad \underbrace{A_1,\ldots,a_1}_{:=\mathcal{R}_1};\cdots;\underbrace{A_l,\ldots,a_l}_{:=\mathcal{R}_l};\ \underbrace{k_m,\ldots,k_m;(k_m-a_1),\ldots,(k_m-A_1);\ldots;(k_m-a_l),\ldots,(k_m-A_l)}_{:=\mathcal{R}_{l+1}}$$

with $\sum_{j=1}^{l+1}|\mathcal{R}_j|=d_m$, where $|\mathcal{R}_j|$ is the number of entries in $\mathcal{R}_j$.

Now we fill in the entries on each row of τ_m/τ_{m-1} as follows. Consider the standard Young tableau T whose row sizes are

$$(\underbrace{k_m,\ldots,k_m}_{d_m})$$

and the entries of the i-th row of T are all equal to i. Now let a sequence of subtableaux of T given by

$$T_1\subset T_2\subset\cdots\subset T_l\subset T_{l+1}:=T$$

such that for each $1\le j\le l$, T_j has the shape of the form

$$A_j>\cdots>a_j>\cdots>A_1>\cdots>a_1.$$

Consider the skew tableau T_j/T_{j-1} for $1\le j\le l+1$ (where we take T_0 to be the empty tableau), then the column sizes of T_j/T_{j-1} is the same as the parametrization for the tableau $\mathcal{R}_j$ marked in (10).

For each $1\le j\le l+1$, fill in the rows of the Ferrers diagram τ_m/τ_{m-1} corresponding to $\mathcal{R}_j$ in (10) by filling the t-th row of $\mathcal{R}_j$ with the t-th entries on each column of T_j/T_{j-1} counting from the top in ascending order. This will give us a *semistandard skew tableau* of shape τ_m/τ_{m-1} and weight

$$(\underbrace{k_m,\ldots,k_m}_{d_m})$$

(see [Goodman and Wallach 2009, Definition 9.3.16]), whose row word is a *reverse lattice word* by [Goodman and Wallach 2009, Definition 9.3.17]. To sum up, it is a desired L-R tableau and (9) follows. □

Proposition 2.5. *Let $J(\lambda,-s\lambda)$ be a unitary module of $GL(n)$ in Theorem 1.1, and V_τ be the K-type obtained by Algorithm 2.2. Then τ satisfies*

$$\{\tau-\rho\}=2\lambda-\rho.$$

Consequently, V_τ is a spin-lowest K-type of $J(\lambda,-s\lambda)$ by (7).

Proof. We prove the proposition by induction on the number of chains in $(\lambda, -s\lambda) = \bigcup_{i=0}^{m} \mathcal{C}_i$, where the chains are arranged so that (8) holds. Suppose that the proposition holds for $\bigcup_{i=0}^{m-1} \mathcal{C}_i$. There are two possibilities when adding $\mathcal{C}_m$:

- There exists $\mathcal{C}_i$ such that $\mathcal{C}_i$ and $\mathcal{C}_m$ is related by Rule (a) in Algorithm 2.2:

$$\begin{matrix} \{ & \mathcal{C}_i & \} \\ & & \{\mathcal{C}_m\}. \end{matrix}$$

- There exist $\mathcal{C}_j$ and $\mathcal{C}_r, \dots, \mathcal{C}_{m-1}$, such that $\mathcal{C}_j$ and $\mathcal{C}_m$ are related by Rule (b), and $\mathcal{C}_l$, $r \le l \le m-1$ and $\mathcal{C}_m$ are related by Rule (c) in Algorithm 2.2:

$$\begin{matrix} \{ \;\; \mathcal{C}_j \;\; \} & \{ \; \mathcal{C}_r \; \} & \dots & \{ \; \mathcal{C}_{m-1} \; \} \\ \{ & & \mathcal{C}_m & \qquad \} \end{matrix}$$

We will only study the second case, and the proof of the first case is simpler. Suppose the chains in the second case are interlaced in the following fashion:

$$(11)\qquad \begin{matrix} \{ \;\; \mathcal{C}_j \;\; \} & & \{ \overbrace{\mathcal{C}_r}^{d_r} \} & & \cdots\cdots & \{ \overbrace{\mathcal{C}_{m-1}}^{d_{m-1}} \} & \\ \{ \underbrace{C_{m,1}, \cdots}_{p} & \underbrace{\cdots\cdots}_{a_r} & \underbrace{\cdots \;\; \cdots}_{d_r} & \underbrace{\cdots\cdots}_{a_{r+1}} & \cdots\cdots & \underbrace{\cdots \;\; \cdots}_{d_{m-1}}, & \underbrace{\cdots, \; C_{m,d_m}}_{a_m} \} \end{matrix}$$

for some $j+1 \le r \le m-1$, and the chains $\mathcal{C}_{j+1}, \dots, \mathcal{C}_{r-1}$ — which have not been shown in (11) — are linked with $\mathcal{C}_j$ under Rule (a) of Algorithm 2.2.

To simplify the calculations below, we introduce the notation

$$(a)_d^{\epsilon} := \underbrace{a, a+\epsilon, \dots, a+(d-1)\epsilon}_{d}.$$

Then 2λ is equal to the entries in (11). Since the values of the adjacent entries within the same chain differ by 2, and the values of the interlaced entries differ by 1, one can calculate $2\lambda - \rho$ up to a translation by a constant on all coordinates as follows:

$$(12)\qquad \begin{matrix} \{\cdots (A_{r-1})_p^0\} & \{(A_r)_{d_r}^0\} & \cdots\cdots & \{(A_{m-1})_{d_{m-1}}^0\} \\ \dots \;\; \{(A_{r-1})_p^0 \;\; (A_r)_{a_r}^{-1} (A_r)_{d_r}^0 & (A_r)_{a_{r+1}}^{-1} & \cdots\cdots & (A_{m-1})_{d_{m-1}}^0 \;\; (A_{m-1})_{a_m}^{-1}\} \end{matrix}$$

where $A_x := \sum_{l=x}^{m-1} a_{l+1}$ for $r-1 \le x \le m-1$ (note that the smallest entry of (12) is 1, appearing at the rightmost entry of the bottom chain).

On the other hand, the calculation in Algorithm 2.2 gives τ as follows:

$$\begin{matrix} (\cdots \;\; (k_j)_p^0) & (k_r)_{d_r}^0 & \cdots & (k_{m-1})_{d_{m-1}}^0 \\ \dots \;\; ((k_m)_p^0 \;\; (k_m)_{a_r}^0 & (k_m)_{d_r}^0 \;\; (k_m)_{a_{r+1}}^0 & \cdots & (k_m)_{d_{m-1}}^0 (k_m)_{a_m}^0) \end{matrix} = \bigcup_{i=0}^{m} \mathcal{T}_i \to \bigcup_{i=0}^{m} \mathcal{T}_i' = \tau,$$

where $\bigcup_{i=0}^{m} \mathcal{T}_i'$ is given by

(13)
$$\begin{array}{llll} (\cdots (k_j+1)_p^1) & (k_r+(q_r-d_r+1))_{d_r}^1 & \cdots & (k_{m-1}+(q_{m-1}-d_{m-1}+1))_{d_{m-1}}^1 \\ \ldots\ ((k_m-1)_p^{-1}(k_m)_{a_r}^0 & (k_m-(q_r-d_r+1))_{d_r}^{-1}(k_m)_{a_{r+1}}^0 & \cdots & (k_m-(q_{m-1}-d_{m-1}+1))_{d_{m-1}}^{-1}(k_m)_{a_m}^0) \end{array}$$

and q_i are obtained by Rule (c) of Algorithm 2.2. For instance, $q_r = p + a_r + d_r$. Note that

$$k_j - (d_j - 1) = k_r + (d_r - 1) + 2a_r + 2.$$

Therefore,

$$k_j - d_j = k_r + d_r + 2a_r.$$

From this, one deduces easily that $k_j \geq k_r + q_r + 1$. Thus it makes sense to talk about the interval $[k_r + q_r + 1, k_j]$.

Before we proceed, we pay closer attention to the coordinates of $\mathcal{T}_j'$, which is the left-most chain on the top row of (13). More precisely, it consists of three parts:

(i) As mentioned in the paragraph after (11), by applying Rule (a) of Algorithm 2.2 between $\mathcal{C}_j$ and each of $\mathcal{C}_{j+1}, \ldots, \mathcal{C}_{r-1}$, one can check that

$$\bigcup_{i=j+1}^{r-1} \mathcal{T}_i' \subset [k_r + q_r + 1, k_j].$$

Suppose there are $\delta \geq 0$ coordinates in $\bigcup_{i=j+1}^{r-1} \mathcal{T}_i'$, then there will be exactly δ coordinates in $\mathcal{T}_j'$ having coordinates strictly greater than $k_j + p$.

(ii) By applying Algorithm 2.2 to $\mathcal{C}_j$ and $\mathcal{C}_m$, we have p coordinates $(k_j + 1)_p^1$ in $\mathcal{T}_j'$ as in (13).

(iii) The other coordinates of $\mathcal{T}_j'$ are either equal to k_j, or smaller than k_j if they are linked with $\mathcal{C}_t$ with $t < j$.

In conclusion, the coordinates of $\mathcal{T}_j'$ are given by

$$(\overbrace{\sharp \ldots \sharp}^{\delta};\ (k_j + 1)_p^1;\ \overbrace{\flat \ldots \flat}^{d_j - \delta - p}),$$

where $\sharp \ldots \sharp$ has coordinates greater than $k_j + p$, and $\flat \ldots \flat$ has coordinates smaller than $k_j + 1$.

We now arrange the coordinates of $\bigcup_{i=j}^{m} \mathcal{T}_i'$ in (13) as follows:

$$\overbrace{\sharp\ \ldots\ \sharp}^{\delta} > \overbrace{(k_j + 1)_p^1}^{p} > \overbrace{\flat \cdots \flat}^{d_j - p - \delta} > \overbrace{\bigcup_{i=j+1}^{r-1} \mathcal{T}_i'}^{\delta} > \mathcal{T}_r' > \cdots > \mathcal{T}_{m-1}' > (k_m)_{a_r}^0 = \cdots = (k_m)_{a_m}^0$$
$$> (k_m - 1)_p^{-1} > (k_m - (q_r - d_r + 1))_{d_r}^{-1} > \cdots > (k_m - (q_{m-1} - d_{m-1} + 1))_{d_{m-1}}^{-1}$$

Here elements in the blocks $\mathcal{T}'_r, \ldots, \mathcal{T}'_{m-1}$ are still kept in the increasing manner. Note that if $x < y$, then $\mathcal{T}'_x > \mathcal{T}'_y$ in terms of their coordinates.

We index the coordinates of τ shown in (13) using the above ordering, with the smallest coordinate indexed by 1:

(14)
$$\begin{matrix} (\ldots (d_m+D_r+d_j-p+1)^1_p) & ((d_m+D_{r+1}+1)^1_{d_r}) & \cdots & ((d_m+1)^1_{d_{m-1}}) \end{matrix}$$
$$\left((D_r+p)^{-1}_p (D_r+p+1)^1_{a_r} (D_r)^{-1}_{d_r} (D_r+p+a_r+1)^1_{a_{r+1}} \cdots (D_{m-1})^{-1}_{d_{m-1}} \left(D_r+p+\sum_{l=r}^{m-1} a_l+1\right)^1_{a_m}\right),$$

where $D_x := \sum_{l=x}^{m-1} d_l$ for $r \le x \le m-1$. Note that the coordinates of the last row read as

$$(D_r+p, \ldots, 2, 1) = ((D_r+p)^{-1}_p;\ (D_r)^{-1}_{d_r};\ \ldots;\ (D_{m-1})^{-1}_{d_{m-1}}),$$
$$(D_r+p+1, \ldots, d_m-1, d_m) =$$
$$\left((D_r+p+1)^1_{a_r}; \ldots; \left(D_r+p+\sum_{l=r}^{x-1} a_l+1\right)^1_{a_x}; \ldots; \left(D_r+p+\sum_{l=r}^{m-1} a_l+1\right)^1_{a_m}\right).$$

Up to a translation of a constant of all coordinates, the difference between (13) and (14) gives (a W-conjugate of) $\{\tau-\rho\}$, which is of the form:

$$\begin{matrix} (\cdots(\beta_j)^0_p) & (\beta_r)^0_{d_r} & \cdots & (\beta_{m-1})^0_{d_{m-1}} \\ \multicolumn{4}{c}{((\alpha_j)^0_p *** (\alpha_r)^0_{d_r} *** \cdots (\alpha_{m-1})^0_{d_{m-1}} ***)} \end{matrix} \tag{15}$$

Our goal is to show (12) and (15) are equal up to a translation of a constant of all coordinates. So we need to show the following:

(i) $\alpha_j = \beta_j$: We need to show

$$k_m - 1 - (D_r+p) = k_j + 1 - (d_m + D_r + d_j - p + 1).$$

In fact, we have

$$\begin{aligned} C_{m,1} &= C_{j,d_j} + 2p - 1, \\ k_m + (d_m-1) &= k_j - (d_j-1) + 2p - 1, \\ k_m - p - 1 &= k_j - d_j + p - d_m, \\ k_m - 1 - (D_r+p) &= k_j + 1 - (d_m + D_r + d_j - p + 1), \end{aligned}$$

as required.

(ii) $\alpha_x = \beta_x$ *for all* $r \le x \le m-1$: This is the same as showing

$$k_m - (q_x - d_x + 1) - D_x = k_x + (q_x - d_x + 1) - (d_m + D_{x+1} + 1).$$

As in (i), we consider

$$\begin{aligned} C_{m,1} &= C_{x,d_x} + 2q_x - 1, \\ k_m + (d_m - 1) &= k_x - (d_x - 1) + 2q_x - 1, \\ k_m - q_x + d_x - 1 &= k_x + q_x - d_m, \\ k_m - q_x + d_x - 1 - D_x + D_{x+1} + d_x &= k_x + (q_x + 1) - (d_m + 1), \\ k_m - q_x + d_x - 1 - D_x &= k_x + (q_x - d_1 + 1) - (d_m + D_{x+1} + 1), \end{aligned}$$

as we wish to show.

(*iii*) $\alpha_j - \alpha_x = A_{r-1} - A_x$ *for all* $r \leq x \leq m-1$: In other words, we need to show

$$[(k_m - 1) - (D_r + p)] - [(k_m - (q_x - d_x + 1)) - D_x] = A_{r-1} - A_x = a_r + \cdots + a_x.$$

Indeed, by looking at (11) and applying Rule (c) of Algorithm 2.2, one gets

$$\begin{aligned} p + (a_r + \cdots + a_x) + (d_r + \cdots + d_x) &= q_x, \\ q_x - p &= (A_{r-1} - A_x) + (D_r - D_{x+1}), \\ (k_m - 1) - (k_m - 1) + q_x - p - D_r + D_{x+1} &= A_{r-1} - A_x, \\ [(k_m - 1) - (D_r + p)] - (k_m - 1) + q_x + (D_x - d_x) &= A_{r-1} - A_x, \\ [(k_m - 1) - (D_r + p)] - [(k_m - (q_x - d_x + 1)) - D_x] &= A_{r-1} - A_x, \end{aligned}$$

so the result follows.

(*iv*): Collecting the $***$ entries of (15) consecutively from left to right gives

$$\underbrace{\alpha_j, \ldots, \alpha_r + 1}_{a_r}; \ldots; \underbrace{\alpha_x, \ldots, \alpha_{x+1} + 1}_{a_{x+1}}; \ldots; \underbrace{\alpha_{m-1}, \ldots, \alpha_{m-1} - (a_m - 1)}_{a_m}.$$

In order for the above expression to make sense, one needs $\alpha_x - \alpha_{x+1} = a_x$ for all $r \leq x \leq m-1$ for instance. This is indeed the case, since $\alpha_x - \alpha_{x+1} = A_x - A_{x+1}$ by (iii), and the latter is equal to a_{x+1} by the definition of A_x for $r - 1 \leq x \leq m-1$. So it suffices to check $k_m - (D_r + p + \sum_{l=r}^{x} a_l + 1) = \alpha_x$.

To see it is the case, one can check that the leftmost entry of the second row of (15) is equal to

$$\begin{aligned} \alpha_j &= k_m - 1 - (D_r + p), \\ \alpha_x + A_{r-1} - A_x &= k_m - (D_r + p + 1), \qquad \text{(by (iii))} \\ \alpha_x + \sum_{l=r}^{x} a_l &= k_m - (D_r + p + 1), \\ \alpha_x &= k_m - \left(D_r + p + \sum_{l=r}^{x} a_l + 1\right), \end{aligned}$$

as follows.

Combining (i)–(iv), (15) can be rewritten as

$$\begin{matrix} (\cdots(\alpha_j)^0_p) & & ((\alpha_r)^0_{d_r}) & & \cdots\cdots & & ((\alpha_{m-1})^0_{d_{m-1}}) & \\ & ((\alpha_j)^0_p & (\alpha_j)^{-1}_{a_r} & (\alpha_r)^0_{d_r} & (\alpha_r)^{-1}_{a_{r+1}} & \cdots\cdots & (\alpha_{m-1})^0_{d_{m-1}} & (\alpha_{m-1})^{-1}_{a_m}), \end{matrix}$$

whose coordinates are in descending order from left to right. So it is equal to $\{\tau-\rho\}$ up to a translation of a constant. Moreover, by comparing it with (12), we have shown that all coordinates of $2\lambda-\rho$ and $\{\tau-\rho\}$ differ by a constant (note that the other coordinates on the left of $\mathcal{C}_j$ are taken care of by induction hypothesis). To see they are exactly equal to each other, we calculate the *true* values of A_{m-1} and α_{m-1} in $2\lambda-\rho$ and τ respectively on the entry marked by $\circledast$ below:

$$\begin{matrix} \{\dots, & *,\dots,*\} & & \{*,\dots,*\} & & \cdots & & \{*,\dots,*\} & \\ & \{*,\dots,*; & *,\dots,*; & *,\dots,*; & *,\dots,*; & \dots & ; *,\dots,\circledast; & \underbrace{*,\dots,*}_{a_m}\} \end{matrix}$$

For $2\lambda-\rho$, $\circledast$ takes the value

$$C_{m,d_m-a_m}-\rho_{a_m+2},$$

where $\rho=(\rho_n,\dots,\rho_2,\rho_1)$ with $\rho_i=\rho_1+(i-1)$. So it can be simplified as

$$\begin{aligned} C_{m,d_m-a_m}-\rho_{a_m+2} &= k_m-(d_m-1)+2a_m-\rho_{a_m+2} \\ &= k_m-d_m+1+2a_m-\rho_1-(a_m+1) \\ &= k_m-d_m+a_m-\rho_1. \end{aligned}$$

On the other hand, for $\{\tau-\rho\}$, $\circledast$ takes the value

$$k_m-q_{m-1}-\rho_1$$

(recall that we had $\alpha_{m-1}=k_m-q_{m-1}-1$ for $\circledast$ in our previous calculation).

By looking at (11) and applying Rule (c) of Algorithm 2.2 again, one has $q_{m-1}=d_m-a_m$, hence $2\lambda-\rho$ and $\{\tau-\rho\}$ takes the same value on the $\circledast$ coordinate. Since we have seen that their coordinates differ by the same constant, one can conclude that $2\lambda-\rho=\{\tau-\rho\}$. □

Example 2.6. For the interlaced chain in Example 2.3, the translate of $2\lambda-\rho$ in (12) is equal to

$$\begin{matrix} \{10-8 & & 8-6\} & & \{6-4\} & & \{4-2\} & \\ & \{9-7 & & 7-5 & & 5-3 & & 3-1 & 1-0\} \end{matrix} = \begin{matrix} \{2 & & 2\} & \{2\} & \{2\} & \\ & \{2 & 2 & & 2 & & 2 & 1\}. \end{matrix}$$

Also, the translate of $\tau-\rho$ in (15) is given by

$$\begin{matrix} (9-8 & & 10-9) & & (8-7) & & (2-1) & \\ & (4-3 & & 3-2 & & 5-4 & & 7-6 & 5-5) \end{matrix} = \begin{matrix} (1 & & 1) & (1) & (1) & \\ & (1 & 1 & & 1 & & 1 & 0) \end{matrix}$$

Hence their coordinates differ by the same constant 1. To see $2\lambda - \rho$ and $\{\tau - \rho\}$ are equal, where $\rho = (4, 3, 2, 1, 0, -1, -2, -3, -4)$, one can look at the *true* values of them for the rightmost entry of the bottom chain:

$$2\lambda - \rho : 1 - \rho_1 = 1 - (-4) = 5; \quad \tau - \rho : 5 - \rho_5 = 5 - 0 = 5.$$

Hence $2\lambda - \rho = \{\tau - \rho\} = (6, 6, 6, 6, 6, 6, 6, 6, 5)$, and the unique $\widetilde{K}$-type in the Dirac cohomology of the corresponding unitary module is $V_{(6,6,6,6,6,6,6,6,5)}$.

3. Scattered representations of $SL(n)$

It is easy to parametrize irreducible unitary representations of $SL(n)$ using the parametrization for $GL(n)$. In such cases, we impose the condition on λ such that the sum of the coordinates is equal to 0. In other words, for each possible regular, half-integral infinitesimal character λ for $SL(n)$, one can shift the coordinates by a suitable scalar, so that it corresponds to an infinitesimal character λ' of $GL(n)$ whose smallest coordinate is equal to $\frac{1}{2}$.

Therefore, the irreducible unitary representations of $SL(n)$ are parametrized by chains with n coordinates whose smallest coordinate is equal to 1.

The following proposition characterizes which of these representations are scattered in the sense of Section 1.4:

Proposition 3.1. *Let $\pi := J(\lambda, -s\lambda)$ be an irreducible unitary representation of $SL(n)$ such that λ is dominant and half-integral. Then π is a scattered representation if and only if the translated Zhelobenko parameter $(\lambda', -s\lambda')$ can be expressed as a union of interlaced chains with smallest coordinate equal to* 1.

Proof. By the arguments in Section 1.4, one only needs to check that $s \in W$ involves all simple reflections in its reduced expression if and only if $(\lambda', -s\lambda') = \bigcup_{i=0}^m \mathcal{C}_i$ are interlaced. Indeed, $s \in W$ can be read from $\bigcup_{i=0}^m \mathcal{C}_i$ as follows: label the entries of $\bigcup_{i=0}^m \mathcal{C}_i$ in descending order, e.g.,

$$\bigcup_{i=0}^{m} \mathcal{C}_i = \begin{matrix} & \{p_{k+1}, \ldots\} \cdots \\ \{p_1, p_2, \ldots, p_k, & p_{k+2}, \ldots\} \cdots \end{matrix}$$

with $p_1 > p_2 > \cdots > p_n$, then we "flip" the entries of each chain $\mathcal{C}_i$ by

$$\{C_{i,1}, \ldots, C_{i,d_i}\} \to \{C_{i,d_i}, \ldots, C_{i,1}\}.$$

Suppose we have

$$\begin{matrix} & \{p_{s_{k+1}}, \ldots\} \cdots \\ \{p_{s_1}, p_{s_2}, \ldots, p_{s_k}, & p_{s_{k+2}}, \ldots\} \cdots \end{matrix}$$

after flipping each chain, then $s \in S_n$ is obtained by

$$s = \begin{pmatrix} 1 & 2 & \dots & n \\ s_1 & s_2 & \dots & s_n \end{pmatrix}$$

(see Example 3.2).

Define the equivalence class of interlaced chains by letting $\mathcal{C}_i \sim \mathcal{C}_j$ if and only if $i = j$, or $\mathcal{C}_i, \mathcal{C}_j$ are interlaced. So we have a partition of $\{p_1, \dots, p_n\}$ by the entries of chains in the same equivalence class. It is not hard to check that the entries on each partition have consecutive indices, i.e.,

$$\mathcal{E}_i = \{p_{a_i}, p_{a_i+1}, \dots, p_{b_i-1}, p_{b_i}\}$$

and $\bigcup_{i=0}^{m} \mathcal{C}_i$ are interlaced if and only if there is only one equivalence class.

We now prove the proposition. Suppose there exists more than one equivalence class, i.e., we have

$$\mathcal{E}_1 = \{p_1, \dots, p_a\}; \quad \mathcal{E}_2 = \{p_{a+1}, \dots, p_b\}$$

for some $1 \leq a < n$. Since the smallest element in any equivalence class must be the smallest element of a chain, and the largest element in a class must be the largest element of a chain, we have

$$\mathcal{C}_i = \{\dots, p_a\}\{p_{a+1}, \dots\} = \mathcal{C}_j.$$

By the above description of $s \in S_n$, it is obvious that $s \in S_a \times S_{n-a} \subset S_n$, which does not involve the simple reflection s_a.

Conversely, if there is only one equivalence class, we suppose on the contrary that there exists some $1 \leq a < n$ such that $s \in S_a \times S_{n-a}$. Since p_a, p_{a+1} are in the same equivalence class, then at least one of the following:

$$\{p_a, p_{a+1}\}, \quad \{p_a, p_{a+2}\}, \quad \{p_{a-1}, p_{a+1}\}$$

is in the same chain $\mathcal{C}_i$ for some $0 \leq i \leq m$. By "flipping" $\mathcal{C}_i$ in either case, there must be some $u \leq a < a+1 \leq v$ such that

$$s = \begin{pmatrix} \dots & u & \dots & v & \dots \\ \dots & v & \dots & u & \dots \end{pmatrix}.$$

The reduced expression of such s must involve the simple reflection s_a, hence we obtain a contradiction. Therefore, s must involve all simple reflections in its reduced expression. □

Example 3.2. Consider the interlaced chain with smallest coordinate 1 given in Example 2.3:

$$\begin{matrix} \{10 & & 8\} & \{6\} & \{4\} & & \\ & \{9 & 7 & 5 & 3 & 1\} \end{matrix}$$

Its corresponding irreducible representation in $SL(9)$ has Langlands parameter $(\lambda', -s\lambda')$, where

$$s = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 9 & 1 & 8 & 5 & 6 & 7 & 4 & 2 \end{pmatrix},$$

and $\lambda' = \left[\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 1\right]$, where $[a_1, \ldots, a_m]$ is defined by

$$[a_1, \ldots, a_m] := a_1\varpi_1 + \cdots + a_m\varpi_m.$$

In fact, the coordinates of λ' is simply obtained by taking the difference of the neighboring coordinates of $\lambda = \frac{1}{2}(10, 9, 8, 7, 6, 5, 4, 3, 1)$.

Example 2.3 implies that the spin-lowest K-type for $J(\lambda', -s\lambda')$ in $SL(8)$ is $V_{[1,1,1,2,0,1,1,1]}$.

Example 3.3. We explore the possibilities of chains $\bigcup_{i=0}^{m} \mathcal{C}_i$ whose corresponding Zhelobenko parameter $(\lambda', -s\lambda')$ gives a spherical representation.

In order for the lowest K-type to be trivial, we need the $\mathcal{T}_i$ in Algorithm 2.2 to have the same average value k_i for all i, that is, the mid-point of all $\mathcal{C}_i$ (if there is more than one) must be the same. This leaves the possibility of $\bigcup_{i=0}^{m} \mathcal{C}_i$ consisting of a single chain, which corresponds to the trivial representation, or there are two chains of lengths $a > b > 0$ whose entries are of different parity. Hence it must be of the form

$$\{2a-1, 2a-3, \ldots, 3, 1\} \cup \{a+(b-1), a+(b-3), \ldots, a-(b-3), a-(b-1)\},$$

where a, b are of different parity.

In other words, such representations can only occur for $SL(n)$ with $n = a + b$ is odd, and is equal to $\mathrm{Ind}_{S(GL(a)\times GL(b))}^{SL(n)}(\mathrm{triv} \otimes \mathrm{triv})$, which are the unipotent representations corresponding to the nilpotent orbit with Jordan block $(2^b 1^{a-b})$ (see [Barbasch and Pandžić 2011, §5.3]). Its Langlands parameter $(\lambda', -s\lambda')$ has

$$2\lambda' = [\underbrace{2, \ldots, 2}_{(a-b-1)/2}, \underbrace{1, \ldots, 1}_{2b}, \underbrace{2, \ldots, 2}_{(a-b-1)/2}]$$

and $s = w_0$ (see [Ding and Dong 2020, Conjecture 5.6]). Moreover, its spin-lowest K-type is given by [Barbasch and Pandžić 2011, Equation (5.5)], which matches with our calculations in Algorithm 2.2.

For the rest of this section, we give two applications of Proposition 3.1:

3.1. ***The spin-lowest K-type is unitarily small.*** To offer a unified conjectural description of the unitary dual, Salamanca-Riba and Vogan [1998] formulated the notion of unitarily small (*u-small* for short) K-type. Here we only quote them for a complex connected simple Lie group G — using the setting in the introduction, a K-type V_δ is u-small if and only if $\langle \delta - 2\rho, \varpi_i \rangle \leq 0$ for $1 \leq i \leq \mathrm{rank}(\mathfrak{g}_0)$ (see [Salamanca-Riba and Vogan 1998, Theorem 6.7]).

Lemma 3.4. *Let* $\lambda = \sum_{i=1}^{\mathrm{rank}(\mathfrak{g}_0)} \lambda_i \varpi_i \in \mathfrak{h}_0^*$ *be a dominant weight such that* $\lambda_i = \frac{1}{2}$ *or* 1 *for each* $1 \leq i \leq n$, *and* V_δ *be the* K*-type with highest weight* δ *such that*

$$\{\delta - \rho\} = 2\lambda - \rho.$$

Then $\langle \delta - 2\rho, \varpi_i \rangle \leq 0$, $1 \leq i \leq \mathrm{rank}(\mathfrak{g}_0)$. *Therefore, the* K*-type* V_δ *is u-small.*

Proof. By assumption, there exists $w \in W$ such that $\delta = w^{-1}(2\lambda - \rho) + \rho$. Thus

$$\begin{aligned} \langle \delta - 2\rho, \varpi_i \rangle &= \langle w^{-1}(2\lambda - \rho) - \rho, \varpi_i \rangle \\ &= \langle w^{-1}(2\lambda - \rho), \varpi_i \rangle - \langle \rho, \varpi_i \rangle \\ &= \langle 2\lambda - \rho, w(\varpi_i) \rangle - \langle \rho, \varpi_i \rangle. \end{aligned}$$

On the other hand, let $w = s_{\beta_1} s_{\beta_2} \cdots s_{\beta_p}$ be a reduced decomposition of w into simple root reflections. Then by [Dong and Huang 2011, Lemma 5.5],

$$\varpi_i - w(\varpi_i) = \sum_{k=1}^{p} \langle \varpi_i, \beta_k^\vee \rangle s_{\beta_1} s_{\beta_2} \cdots s_{\beta_{k-1}}(\beta_k). \tag{16}$$

Note that $s_{\beta_1} s_{\beta_2} \cdots s_{\beta_{k-1}}(\beta_k)$ is a positive root for each k. Now we have that

$$\begin{aligned} \langle \delta - 2\rho, \varpi_i \rangle &= \Big\langle 2\lambda - \rho, \varpi_i - \sum_{k=1}^{p} \langle \varpi_i, \beta_k^\vee \rangle s_{\beta_1} s_{\beta_2} \cdots s_{\beta_{k-1}}(\beta_k) \Big\rangle - \langle \rho, \varpi_i \rangle \\ &= 2\langle \lambda - \rho, \varpi_i \rangle - \sum_{k=1}^{p} \langle \varpi_i, \beta_k^\vee \rangle \langle 2\lambda - \rho, s_{\beta_1} s_{\beta_2} \cdots s_{\beta_{k-1}}(\beta_k) \rangle \\ &\leq 2\langle \lambda - \rho, \varpi_i \rangle \\ &\leq 0. \end{aligned}$$

□

Corollary 3.5. *The unique spin-lowest* K*-type* V_τ *of any scattered representation of* $SL(n)$ *is u-small. Consequently*, [Ding and Dong 2020, Conjecture C] *holds for* $SL(n)$.

Proof. Let $(\lambda, -s\lambda)$ be the Zhelobenko parameter for a scattered representation of $SL(n)$. Write $\lambda = \sum_{i=1}^{n-1} \lambda_i \varpi_i$ in terms of the fundamental weights. Then it is direct from our definition of the interlaced chains that each λ_i is either $\frac{1}{2}$ or 1 (recall Proposition 3.1 and Example 3.2). Let V_τ be the unique spin-lowest K-type of the scattered representation. Then $\{\tau - \rho\} = 2\lambda - \rho$ (see Proposition 2.5). Thus the result follows from Lemma 3.4. □

3.2. ***Number of scattered representations.*** As another application of Proposition 3.1, we compute the number of scattered representations of $SL(n)$. By the proposition, it is equal to the number of interlaced chains with n entries with the smallest entry equal to 1. We now give an algorithm of constructing new interlaced chains with smallest coordinate equal to 1 from those with one less coordinate:

Algorithm 3.6. Let $\bigcup_{i=1}^{p}\{2A_i-1,\dots,2a_i-1\}\cup\bigcup_{j=1}^{q}\{2B_j,\dots,2b_j\}$ be a union of interlaced chains with such that

- $A_{i'}>A_i$ if $i'>i$, and $B_{j'}>B_j$ if $j'>j$; and
- $2a_p-1=1$.

We construct two new interlaced chains with one extra coordinate as follows. (When $q=0$, we adopt Case I only.)

Case I: If $2A_p-1>2B_q+1$, then the two new interlaced chains are

$$\begin{array}{ccccccc} & & & \{2B_q & \dots & 2b_q\} & \dots \\ \{\mathbf{2A_p+1} & 2A_p-1 & \dots & & 2a_p-1\} & & \dots \end{array}$$

and

$$\begin{array}{ccccccc} & \{\mathbf{2A_p-2}\} & & \{2B_q & \dots & 2b_q\} & \dots \\ \{2A_p-1 & & \dots & \dots & 2a_p-1\} & & \dots \end{array}$$

Case II: If $2A_p-1=2B_q+1$, then the two new interlaced chains are

$$\begin{array}{cccccc} & & \{2B_q & \dots & 2b_q\} & \dots \\ \{\mathbf{2A_p+1} & 2A_p-1 & & \dots \quad 2a_p-1\} & & \dots \end{array}$$

and

$$\begin{array}{cccccc} \{\mathbf{2B_q+2} & & 2B_q & \dots & 2b_q\} & \dots \\ & \{2A_p-1 & & \dots \quad 2a_p-1\} & & \dots \end{array}$$

Case III: If $2A_p-1=2B_q-1$, then the two new interlaced chains are

$$\begin{array}{cccccc} & \{2B_q & & \dots & 2b_q\} & \dots \\ \{\mathbf{2A_p+1} & & 2A_p-1 & \dots & 2a_p-1\} & \dots \end{array}$$

and

$$\begin{array}{cccccc} \{\mathbf{2B_q+2} & 2B_q & & \dots & 2b_q\} & \dots \\ & & \{2A_p-1 & \dots & 2a_p-1\} & \dots \end{array}$$

Case IV: If $2A_p-1<2B_q-1$, then the two new interlaced chains are

$$\begin{array}{ccccccc} \{2B_q & & \dots & \dots & & 2B_q\} & \dots \\ & \{\mathbf{2B_q-1}\} & \{2A_p-1 & & \dots & 2a_p-1\} & \dots \end{array}$$

and

$$\begin{array}{llllllll} \{\mathbf{2B_q+2} & 2B_q & \ldots & & 2b_p\} & \ldots & & \\ & & & \{2A_p-1 & & \ldots & 2a_p-1\} & \ldots \end{array}$$

Example 3.7. Suppose we begin with an interlaced chain $\{9, 7, 5, 3, 1\} \cup \{4, 2\}$. Then the new interlaced chains with one extra coordinate are

$$\{11, 9, 7, 5, 3, 1\} \cup \{4, 2\} \quad \text{and} \quad \{9, 7, 5, 3, 1\} \cup \{8\} \cup \{4, 2\}.$$

Proposition 3.8. *All interlaced chains with $n \geq 2$ entries with smallest coordinate equal to 1 can be obtained uniquely from the chain $\{3\ 1\}$ by inductively applying the above algorithm.*

Proof. Suppose $\bigcup_{i=0}^{m} \mathcal{C}_i$ be interlaced chains with largest coordinate equal to $M \in \mathcal{C}_0$. We remove a coordinate from it by the following rule: If $\mathcal{C}_i \neq \{M-1\}$ for all i, remove the entry M from $\mathcal{C}_0$. Otherwise, remove the whole chain $\{M-1\}$ from the original interlaced chains.

One can easily check from the definition of interlaced chain that the reduced chains are still interlaced, and one can recover the original chain by applying Algorithm 3.6 on the reduced chain.

Therefore, for all interlaced chains with smallest entry 1, we can use the reduction mentioned in the first paragraph repeatedly to get an interlaced chain with only 2 entries, which must be of the form $\{3\ 1\}$, and repeated applications of Algorithm 3.6 on $\{3\ 1\}$ will retrieve the original interlaced chains (along with other chains). In other words, all interlaced chains with smallest entry 1 can be obtained by Algorithm 3.6 inductively on $\{3\ 1\}$.

We are left to show that all interlaced chains are uniquely constructed using the algorithm. Suppose on the contrary that there are two different interlaced chains that give rise to the same $\bigcup_{i=0}^{m} \mathcal{C}_i$ after applying Algorithm 3.6. By the algorithm, these two chains must be obtained from $\bigcup_{i=0}^{m} \mathcal{C}_i$ by removing its largest odd entry $M_o \in \mathcal{C}_p$ or largest even entry $M_e \in \mathcal{C}_q$. So they must be equal to

$$\bigcup_{i \neq p} \mathcal{C}_i \cup (\mathcal{C}_p \backslash \{M_o\}) \quad \text{and} \quad \bigcup_{i \neq q} \mathcal{C}_i \cup (\mathcal{C}_q \backslash \{M_e\}),$$

respectively.

Assume $M_o > M_e$ for now (and the proof for $M_e > M_o$ is similar). By applying Algorithm 3.6 to $\bigcup_{i \neq q} \mathcal{C}_i \cup (\mathcal{C}_q \backslash \{M_e\})$, we obtain two interlaced chains

$$\bigcup_{i \neq p,q} \mathcal{C}_i \cup \mathcal{C}'_p \cup (\mathcal{C}_q \backslash \{M_e\}) \quad \text{and} \quad \bigcup_{i \neq q} \mathcal{C}_i \cup (\mathcal{C}_q \backslash \{M_e\}) \cup \{M_o - 1\},$$

where

$$\mathcal{C}'_p := \{M_o + 2, \overbrace{M_o, \ldots, m_o}^{\mathcal{C}_p}\}.$$

Note that none of the above gives rise to the interlaced chains $\bigcup_{i=0}^{m} \mathcal{C}_i$: Even in the case when $M_0 - 1 = M_e$, $(\mathcal{C}_q \backslash \{M_e\}) \cup \{M_o - 1\}$ and $\mathcal{C}_q$ are different — although they have the same coordinates, the first consists of two chains while the second consists of one chain only. So we have a contradiction, and the result follows. □

Corollary 3.9. *The number of interlaced chains with n coordinates and the smallest coordinate equal to* 1 *is equal to* 2^{n-2}.

Since the scattered representations of $SL(n+1)$ are in one to one correspondence with interlaced chains with $n+1$ coordinates having smallest coordinate 1, this corollary implies that the number of scattered representations of Type A_n is equal to 2^{n-1}. This verifies [Dong 2019, Conjecture 5.2]. Moreover, by using `atlas`, the spin-lowest K-types for all scattered representations of $SL(n)$ with $n \leq 6$ are given in [Dong 2019, Tables 1–3]. One can easily check the results there match with our V_τ in Algorithm 2.2.

Example 3.10. Let us start from $SL(2, \mathbb{C})$ and the chain $\{3\ 1\}$. This chain corresponds to the trivial representation.

Now we consider $SL(3, \mathbb{C})$. By Algorithm 3.6, the chain $\{3\ 1\}$ for $SL(2)$ produces two chains

$$\{5\ 3\ 1\} \qquad \begin{matrix} & \{2\} & \\ \{3 & & 1\} \end{matrix}.$$

The first corresponds to the trivial representation, while the second gives the representation with $\lambda = \left[\frac{1}{2}, \frac{1}{2}\right]$ and

$$s = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

One computes by Algorithm 2.2 that the spin-lowest K-type $\tau = [1, 1]$.

Now let us consider $SL(4)$. By Algorithm 3.6, the chain $\{5\ 3\ 1\}$ for $SL(3)$ produces two chains

$$\{7\ 5\ 3\ 1\} \qquad \begin{matrix} & \{4\} & & \\ \{5 & & 3 & 1\} \end{matrix}.$$

The first chain corresponds to the trivial representation, while the second one gives the representation with $\lambda = \left[\frac{1}{2}, \frac{1}{2}, 1\right]$ and

$$s = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}.$$

One computes by Algorithm 2.2 that the spin-lowest K-type $\tau = [2, 0, 1]$. The other chain of $SL(3)$ shall produce

$$\begin{matrix} & & \{2\} & \\ \{5 & 3 & & 1\} \end{matrix} \qquad \begin{matrix} \{4 & & 2\} & \\ & \{3 & & 1\} \end{matrix}$$

One computes that

$$\lambda = \left[1, \tfrac{1}{2}, \tfrac{1}{2}\right], \quad s = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}, \quad \tau = [1, 0, 2];$$

and that

$$\lambda = \left[\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right], \quad s = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, \quad \tau = [1, 1, 1],$$

respectively. These four representations (and their spin-lowest K-types) match precisely with [Dong 2019, Table 1].

Acknowledgements

We thank the referee sincerely for very careful reading and nice suggestions.

Dong was supported by the National Natural Science Foundation of China (grant 11571097, 2016–2019). Wong is supported by the National Natural Science Foundation of China (grant 11901491) and the Presidential Fund of CUHK(SZ).

References

[Barbasch and Pandžić 2011] D. Barbasch and P. Pandžić, "Dirac cohomology and unipotent representations of complex groups", pp. 1–22 in *Noncommutative geometry and global analysis*, edited by A. Connes et al., Contemp. Math. **546**, Amer. Math. Soc., Providence, RI, 2011. MR Zbl

[Barbasch et al. 2020] D. Barbasch, C.-P. Dong, and K. D. Wong, "Dirac series for complex classical Lie groups", preprint, 2020. arXiv

[Ding and Dong 2020] J. Ding and C.-P. Dong, "Unitary representations with Dirac cohomology: a finiteness result for complex Lie groups", *Forum Math.* **32**:4 (2020), 941–964. MR Zbl

[Dong 2013] C.-P. Dong, "On the Dirac cohomology of complex Lie group representations", *Transform. Groups* **18**:1 (2013), 61–79. Erratum in **18**:2 (2013), 595–597. MR Zbl

[Dong 2019] C.-P. Dong, "Unitary representations with non-zero Dirac cohomology for complex E_6", *Forum Math.* **31**:1 (2019), 69–82. MR Zbl

[Dong and Huang 2011] C.-P. Dong and J.-S. Huang, "Jacquet modules and Dirac cohomology", *Adv. Math.* **226**:4 (2011), 2911–2934. MR Zbl

[Goodman and Wallach 2009] R. Goodman and N. R. Wallach, *Symmetry, representations, and invariants*, Graduate Texts in Mathematics **255**, Springer, 2009. MR Zbl

[Huang and Pandžić 2002] J.-S. Huang and P. Pandžić, "Dirac cohomology, unitary representations and a proof of a conjecture of Vogan", *J. Amer. Math. Soc.* **15**:1 (2002), 185–202. MR Zbl

[Huang and Pandžić 2006] J.-S. Huang and P. Pandžić, *Dirac operators in representation theory*, Mathematics: Theory and Applications, Birkhäuser, Boston, 2006. MR Zbl

[Huang et al. 2009] J.-S. Huang, Y.-F. Kang, and P. Pandžić, "Dirac cohomology of some Harish–Chandra modules", *Transform. Groups* **14**:1 (2009), 163–173. MR Zbl

[Huang et al. 2011] J.-S. Huang, P. Pandžić, and V. Protsak, "Dirac cohomology of Wallach representations", *Pacific J. Math.* **250**:1 (2011), 163–190. MR Zbl

[Parthasarathy 1972] R. Parthasarathy, "Dirac operator and the discrete series", *Ann. of Math.* (2) **96** (1972), 1–30. MR Zbl

[Parthasarathy 1980] R. Parthasarathy, "Criteria for the unitarizability of some highest weight modules", *Proc. Indian Acad. Sci. Sect. A Math. Sci.* **89**:1 (1980), 1–24. MR Zbl

[Parthasarathy et al. 1967] K. R. Parthasarathy, R. Ranga Rao, and V. S. Varadarajan, "Representations of complex semi-simple Lie groups and Lie algebras", *Ann. of Math.* (2) **85** (1967), 383–429. MR Zbl

[Salamanca-Riba and Vogan 1998] S. A. Salamanca-Riba and D. A. Vogan, Jr., "On the classification of unitary representations of reductive Lie groups", *Ann. of Math.* (2) **148**:3 (1998), 1067–1133. MR Zbl

[Vogan 1986] D. A. Vogan, Jr., "The unitary dual of GL(n) over an Archimedean field", *Invent. Math.* **83**:3 (1986), 449–505. MR Zbl

[Vogan 1997] J. Vogan, D. A., "Dirac operators and unitary representations", 3 talks at MIT Lie groups seminar, MIT, 1997.

[Zhelobenko 1974] D. P. Zhelobenko, *Harmonic analysis on complex semisimple Lie groups*, Izdat. "Nauka", Moscow, 1974. MR

Received October 23, 2019. Revised September 19, 2020.

CHAO-PING DONG
SCHOOL OF MATHEMATICAL SCIENCES
SOOCHOW UNIVERSITY
SUZHOU
CHINA
chaopindong@163.com

KAYUE DANIEL WONG
SCHOOL OF SCIENCE AND ENGINEERING
THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN
SHENZHEN
CHINA
kayue.wong@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 2, 2020
https://doi.org/10.2140/pjm.2020.309.313

NUMBER OF SINGULAR FIBRES OF SURFACE FIBRATIONS OVER $\mathbb{P}^1$

CHENG GONG AND YI GU

Let $f : X \to \mathbb{P}^1_{\Bbbk}$ be a nonisotrivial surface fibration of fibre genus $g > 0$ over an algebraically closed field $\Bbbk$ of positive characteristic, we study the optimum lower bound for the number of singular fibres of f with respect to the characteristic of $\Bbbk$ in this paper.

1. Introduction

Let us fix an algebraically closed field $\Bbbk$ and let $f : X \to \mathbb{P}^1_{\Bbbk}$ be a relatively minimal surface fibration of genus $g > 0$ ranging in a certain class of fibrations (e.g., the class of *nonisotrivial or semistable* fibrations). The number of singular fibres of f, denoted by $s(f)$, is one of the first invariants people are interested in. A first systematic study of the optimum lower bound of $s(f)$ is done in the well known paper by Beauville [1981]. In that paper Beauville proved the following theorem.

Theorem 1.1 [Beauville 1981]. *Let $f : X \to \mathbb{P}^1_{\Bbbk}$ be a relatively minimal surface fibration over an algebraically closed field $\Bbbk$ of characteristic p of fibre genus $g \geq 1$. Then*

(1) *if either $p = 0$ or $p > 2g+1$, we have*
- *$s(f) \geq 2$ if f is not trivial, and*
- *$s(f) \geq 3$ if f is not isotrivial;*

(2) *if $p = 0$, we have $s(f) \geq 4$ if f is semistable.*

The two inequalities on $s(f)$ in (1) *are optimum for any characteristic p mentioned.*

Later, the part (2) of Theorem 1.1 has been improved. First, Tan and Liu independently prove the following result:

Theorem 1.2 [Tan 1995; Liu 1996]. *Let $f : X \to \mathbb{P}^1$ be a relatively minimal surface fibration over $\mathbb{C}$ of fibre genus $g \geq 2$. Then we have $s(f) \geq 5$ if f is semistable.*

Gu is the corresponding author. The authors are supported by NSFC (No. 11801391), NSF of Jiangsu Province (No. BK20180832) and NSFC-ISF (No. 11761141005).
MSC2010: 14D05, 14D06.
Keywords: algebraic surface, singular fibre, positive characteristic.

Then, Nguyen [1998] proves that the part (2) of Theorem 1.1 is valid for any characteristic p.

After that, various scholars have also obtained a lot of results on the lower bound of $s(f)$ with different assumptions on f. One can refer to [Kovács 1997; Nguyen 1997; Viehweg and Zuo 2001; Tan et al. 2005; Tu 2007; Zamora 2012; Gong et al. 2013; Lu et al. 2016; 2018] for details.

The central topic of the present paper is to study the analogue of Theorem 1.1(1) in the remaining cases where $0 < p \leq 2g+1$. We show that Beauville's lower bound are not valid in these cases. The main results are the following.

Theorem 1.3. *Fixing an algebraically closed field* $\Bbbk$ *of characteristic* $p > 0$, *let* $f : X \to \mathbb{P}^1_{\Bbbk}$ *be a nonisotrivial surface fibration over* $\Bbbk$ *of fibre genus* $g \geq 1$. *Then*

(1) *if* $p \leq 2g-1$ *then the optimum lower bound of* $s(f)$ *is* 1;

(2) *if* $p = 2g+1$ *and* f *is hyperelliptic, then the optimum lower bound of* $s(f)$ *is* 2;

(3) *if* $p = 2$ *or* 3 *and* $g = 1$ *then the optimum lower bound of* $s(f)$ *is* 2.

When $p \leq 2g-1$, we give in Examples 4.1, 4.2 and 4.4 a nonisotrivial fibration f with $s(f) = 1$ fulfilling the conditions in Theorem 1.3(1).

When $p = 2g+1$, examples of nonisotrivial hyperelliptic fibration with $s(f) = 2$ are given in Examples 3.7 and 3.8. A classification of hyperelliptic fibrations with $s(f) = 1$ is also given in Theorem 3.10, such a fibration must be isotrivial.

We prove in Corollary 3.2 and Theorem 6.1 the lower bounds of $s(f)$ mentioned in Theorem 1.3 (2) and (3), respectively.

Finally, we would like to mention the following question:

Question 1.4. Does there exist a nonisotrivial surface fibration $f : X \to \mathbb{P}^1_{\Bbbk}$ of fibre genus g with only one singular fibre over an algebraically closed field $\Bbbk$ of characteristic p when $p = 2g+1$?

Remark 1.5. Such f exists if and only there is a $\mathbb{Z}/p\mathbb{Z}$-Galois cover $\pi : C \to \mathbb{P}^1_{\Bbbk}$ of smooth curves that is étale over $\mathbb{A}^1_{\Bbbk} \subseteq \mathbb{P}^1_{\Bbbk}$ and a Kodaira fibration $h : Y \to C$ along with an equivariant $\mathbb{Z}/p\mathbb{Z}$-automorphism on Y.

This paper is organized as follows.

- Section 2 is for preliminaries, we recall the monodromy theory in positive characteristics.
- In Section 3, we analyse the monodromy in case $p = 2g+1$, we give a structure theorem (Theorem 3.1) of surface fibrations with one singular fibre based on the monodromy analysis in this part. Some discussion on the hyperelliptic case is also provided.

- In Section 4, we give examples of nonisotrivial fibrations with only one singular fibre when $p < 2g+1$.
- In Section 5, we take a special look to the case $g = 2$. Some further examples of nonisotrivial fibration of genus 2 with only one singular fibre are given.
- In Section 6, we take a look at the case $g = 1$, we show that a nonisotrivial elliptic fibration have at least two singular fibres. Some examples of elliptic fibrations with one singular fibre are given.

2. Preliminaries

2A. ***Conventions.*** We keep the following conventions in the rest of the paper:

(1) $\Bbbk$ is an algebraically closed field of characteristic $p > 0$; every variety mentioned is defined over $\Bbbk$.

(2) A surface fibration is a flat morphism from a smooth projective surface to a smooth curve having connected fibres. For our specific purpose, we assume the general fibre of any surface fibration considered in this paper is smooth.

(3) We regard $\mathbb{A}^1_{\Bbbk} \subseteq \mathbb{P}^1_{\Bbbk}$ as an open subset in the standard way throughout this paper and denote by t the affine coordinate function of $\mathbb{A}^1_{\Bbbk}$.

(4) An Artin–Schreier curve is a smooth projective curve C along with a $\mathbb{Z}/p\mathbb{Z}$ cyclic cover $\pi : C \to \mathbb{P}^1_{\Bbbk}$ that is étale over $\mathbb{A}^1_{\Bbbk}$.

Definition 2.1. A relatively minimal surface fibration $f : X \to \mathbb{P}^1_{\Bbbk}$ of genus $g \geq 1$ with smooth generic fibre is called isotrivial if all smooth closed fibres of f are mutually isomorphic. In other words, f is isotrivial if the induced rational map $\mathbb{P}^1_{\Bbbk} \dashrightarrow \mathcal{M}_g$ is a constant one.

2B. ***Picard–Lefschetz monodromy in positive characteristics.***

The inertia group. Let $C/\Bbbk$ be a smooth curve and $c \in C$ be a closed point. Denote by $\mathcal{O}^{\mathrm{sh}}_{C,c} \supsetneq \mathcal{O}_{C,c}$ the strict Henselisation of the local ring $\mathcal{O}_{C,c}$ (see, e.g., [Fu 2015, §2]). Recall that the inertia group $I_c := \mathrm{Gal}(K_c^{\mathrm{sep}}/K_c)$ at this time is the absolute Galois group of $K_c := \mathrm{Frac}(\mathcal{O}^{\mathrm{sh}}_{C,c})$. Let us then fix a uniformizer π of $\mathcal{O}_{C,c}$ and denote

$$K_{c,t} := \bigcup_{(n,p)=1} K_c[\sqrt[n]{\pi}] \subsetneq K_c^{\mathrm{sep}}.$$

This field $K_{c,t}$ is a Galois extension of K_c independent on the choice of π. The tame inertia group is then defined as

$$I_{c,t} := \mathrm{Gal}(K_{c,t}/K_c).$$

By definition, we have the following exact sequence:

$$1 \to P_c := \mathrm{Gal}(K_c^{\mathrm{sep}}/K_{c,t}) \to I_c \to I_{c,t} \to 1. \tag{2-1}$$

We call the group P_c as the wild inertia subgroup.

Remark 2.2. It is well known that $I_{c,t}$ is canonically isomorphic to

$$\widehat{\mathbb{Z}}_{(p)}(1) := \varprojlim_{(n,p)=1} \mu_n(K_c) = \varprojlim_{(n,p)=1} \mu_n(\Bbbk),$$

and P_c is a pro-p group.

Picard–Lefschetz monodromy and semistable reduction of curves. Let $C/\Bbbk$ still be a smooth projective curve and $f : X \to C$ be a relatively minimal surface fibration of fibre genus g. Let ℓ be a prime different from p, then the inertia group I_c acts naturally and continuously on the space $H^1_{\acute{e}t}(X_{K_c^{\mathrm{sep}}}, \mathbb{Z}_\ell(1)) \simeq \mathbb{Z}_\ell^{2g}$. We call the associated homomorphism $\Psi : I_c \to \mathrm{GL}_{2g}(\mathbb{Z}_\ell)$ as the local monodromy homomorphism.

Theorem 2.3 (stable reduction theorem of curves). *We have the following*:

(1) [Serre and Tate 1968, Appendix; SGA 7_{I} 1972]. *There is an open subgroup $G \subseteq I_c$ such that $\Psi(G)$ consists of unipotent matrices.*

(2) [Bosch et al. 1990; Deligne and Mumford 1969]. *Suppose the fibre genus of f is at least 2, then the fibre X_c is semistable if and only if $\Psi(I_c)$ is consisting of unipotent matrices.*

(3) (Néron–Ogg–Shafarevich criterion [Bosch et al. 1990]). *Suppose the fibre genus is at least 2, then the local monodromy homomorphism is trivial if and only if X_c is semistable and its Jacobian is an Abelian variety.*

3. Number of singular fibres, I: $p = 2g + 1$

Let $f : X \to \mathbb{P}^1_\Bbbk$ be a nonisotrivial surface fibration of fibre genus $g \geq 2$. By an argument involving monodromy, Beauville [1981] proved (see Theorem 1.1(1)) that $s(f) \geq 3$ provided that $p > 2g + 1$. The assumption $p > 2g + 1$ is used in order to avoid the wild inertia subgroup action in local monodromy. In case of $p = 2g + 1$, the wild inertia subgroup will occur inevitably in local monodromy. We shall carefully deal with the wild inertia group action in this section.

3A. *A structure theorem.*

We prove the following structure theorem in this section.

Theorem 3.1. *Let $f : X \to \mathbb{P}^1_\Bbbk$ be a surface fibration of fibre genus g that is smooth over $\mathbb{A}^1_\Bbbk \subseteq \mathbb{P}^1_\Bbbk$. If $p = 2g + 1$, then there is an Artin–Schreier curve*

$\pi : C \to \mathbb{P}^1_{\Bbbk}$ of degree p with some $(n, p) = 1$ such that the relative minimal model Y of $X_n \times_{\mathbb{P}^1_{\Bbbk},\pi} C \to C$ in the following diagram is smooth over C:

$$\begin{array}{ccccccc}
Y & \overset{\text{relative minimal model}}{\dashrightarrow} & X_n \times_{\mathbb{P}^1_{\Bbbk}} C & \longrightarrow & X_n = X \times_{\mathbb{P}^1_{\Bbbk},[n]_{\mathbb{P}^1_{\Bbbk}}} \mathbb{P}^1_{\Bbbk} & \longrightarrow & X \\
\Big\downarrow{\scriptstyle h} & & \Big\downarrow & & \Big\downarrow & & \Big\downarrow{\scriptstyle f} \\
C & =\!=\!=\!= & C & \xrightarrow{\ \pi\ } & \mathbb{P}^1_{\Bbbk} & \xrightarrow{[n]_{\mathbb{P}^1_{\Bbbk}}} & \mathbb{P}^1_{\Bbbk}.
\end{array}$$

Here $[n]_{\mathbb{P}^1_{\Bbbk}} : \mathbb{P}^1_{\Bbbk} \to \mathbb{P}^1_{\Bbbk}$ is the morphism given by $t \mapsto t^n$.

Corollary 3.2. *If the morphism f in Theorem 3.1 is furthermore hyperelliptic, i.e., any smooth closed fibre of f is hyperelliptic, then f is isotrivial.*

Proof. Since $h : Y \to C$ is a complete family of smooth hyperelliptic curves, it has to be isotrivial and we are done. □

Remark 3.3. (1) One can replace the hyperelliptic assumption in the this corollary by superelliptic or some other property whose moduli space contains no complete curves.

(2) By Theorem 3.1, if f is nonisotrivial, then we obtain a Kodaira fibration $h : Y \to C$ along with an equivariant $\mathbb{Z}/p\mathbb{Z}$ automorphism induced by the base change π. We can not prove or disprove the existence of such Kodaira fibration over Artin–Schreier curves at the moment.

We need some preparations for the proof of the theorem. First we fix a prime $\ell \neq p$ such that its congruence class $[\ell] \in \mathbb{Z}/p^2\mathbb{Z}$ generates the group of units $(\mathbb{Z}/p^2\mathbb{Z})^*$ (it is a cyclic group of order $p(p-1)$). We can see that the cardinality $\sharp \mathrm{GL}_{p-1}(\mathbb{Z}/\ell\mathbb{Z}) = \prod_{j=0}^{p-2}(\ell^{p-1} - \ell^j)$ is then not divided by p^2.

Lemma 3.4. *The polynomial*

$$p(x) := x^{p-1} + x^{p-2} + \cdots + 1 = \frac{x^p - 1}{x - 1}$$

is irreducible in both $\mathbb{F}_\ell[x]$ and $\mathbb{Q}_\ell[x]$.

Proof. If $p(x)$ is irreducible over $\mathbb{F}_\ell$ then it is also irreducible over $\mathbb{Z}_\ell$ and $\mathbb{Q}_\ell$. The roots of $p(x)$ consist of primitive unit roots of order p and note the group of units in a finite field is always a cyclic group. The polynomial $p(x)$ is then either irreducible or splitting over any finite field. The splitting of $p(x)$ over $\mathbb{F}_\ell$ is now the same as $p \mid (l - 1) = \sharp(\mathbb{F}_\ell^*)$, which is not the case by our choice of ℓ. □

Corollary 3.5. *Let A be a matrix in $\mathrm{GL}_{p-1}(\mathbb{Z}_\ell)$ of order p, then the characteristic polynomial of A is $p(x)$. In particular, any matrix in $\mathrm{GL}_{p-1}(\overline{\mathbb{Q}}_\ell)$ commuting with A is semisimple.*

Proof. Since $x^p - 1 = (x-1)p(x)$ eliminates A, $p(x)$ is irreducible and A is not killed by $x-1$, both the minimal and characteristic polynomial of A are $p(x)$ for sake of dimension. □

Now let us turn to the proof of Theorem 3.1.

Proof of Theorem 3.1. Denote by I_∞ (resp. $I_{\infty,t}$, P_∞) the inertia group (resp. tame/wild inertia subgroup) at $\infty \in \mathbb{P}^1_{\Bbbk}$ and $\Psi : I_\infty \to \mathrm{GL}_{2g}(\mathbb{Z}_\ell)$ the local monodromy homomorphism associated to f. The Sylow-p subgroup (of a pro-finite group, see [Fu 2015, §4.2]) of $\mathrm{GL}_{p-1}(\mathbb{Z}_\ell)$ is a cyclic group of order p since $p^2 \nmid \sharp\mathrm{GL}_{2g}(\mathbb{Z}/\ell\mathbb{Z})$ by our choice of ℓ. In particular, in the above local monodromy homomorphism $\Psi : I_\infty \to \mathrm{GL}_{p-1}(\mathbb{Z}_\ell)$, the image $\Psi(P_\infty)$ is either trivial or a cyclic group of order p.

If the image of $\Psi(P_\infty)$ is trivial, then Ψ factors through the quotient group $I_{\infty,t}$. By Theorem 2.3(1), there is an open subgroup G of $I_{\infty,t}$ such that $\Psi(G)$ consists of unipotent matrices. Due to the structure of $I_{\infty,t}$ (see Remark 2.2), G is a normal subgroup and the index $n := [I_{\infty,t} : G]$ is prime to p. Now replacing X by X_n (or more precisely, its associated relative minimal model), we can assume the image of Ψ is trivial and hence f is at worst semistable by Theorem 2.3(2). It then follows from Szpiro's rigidity theorem [1979, Theorem 3.3] that f is trivial. We do not even need the base change $\pi : C \to \mathbb{P}^1_{\Bbbk}$ in this case.

If the image of $\Psi(P_\infty)$ is a cyclic group generated by a matrix A of order p. Note that P_∞ is normal in I_∞ so its image $\langle A\rangle$ is normal in the image $\Psi(I_\infty)$. Now for any $B \in \Psi(I_\infty)$, we have $BAB^{-1} = A^i$ for some i. As a result, B^{p-1} must commute with A and therefore B^{p-1} is semisimple by Corollary 3.5. In particular, B is itself semisimple. It then follows immediately from Theorem 2.3(1) that $\Psi(I_\infty)$ is finite. Again by the structure of $I_{\infty,t}$ (see Remark 2.2), $\Psi(I_\infty)$ is the extension of $\langle A\rangle$ by a cyclic group of order n prime to p. Then after the base change $[n]_{\mathbb{P}^1_{\Bbbk}} : \mathbb{P}^1_{\Bbbk} \to \mathbb{P}^1_{\Bbbk}$, we may furthermore assume that $\Psi(I_\infty) = \langle A\rangle$. By the following Lemma 3.6, we see that after another base change by an Artin–Schreier curve $\pi : C \to \mathbb{P}^1_{\Bbbk}$ of degree p. The local monodromy associated on C is then everywhere trivial. Then by Theorem 2.3 (2) and (3), the relative minimal model $h : Y \to C$ is at least semistable and the associated Jacobian fibration $j : \mathrm{Pic}^0_{Y/C} \to \mathbb{C}$ has everywhere good reduction. Denoting by $c = \pi^{-1}(\infty)$ the unique point lying above ∞, the fibre Y_c admits a faithful $\mathbb{Z}/p\mathbb{Z} \simeq \mathrm{Aut}(\pi)$-action on the ℓ-adic Tate module $T_\ell(\mathrm{Jac}(Y_c)) \simeq \mathbb{Z}_\ell^{2g}$. By construction, this action is nothing but the action of $\langle A\rangle \simeq \mathbb{Z}/p\mathbb{Z}$ on $\mathbb{Z}_\ell^{p-1}$ given by its matrix representation. Now we claim Y_c is smooth. Otherwise, Y_c has more than one irreducible component of positive genus. The number of such components is less than $2g-2 = p-3 < p$ and hence the $\mathbb{Z}/p\mathbb{Z}$-action fixes each irreducible component of positive genus. In particular, for any component D, $T_\ell(\mathrm{Jac}(D)) \subsetneq T_\ell(\mathrm{Jac}(Y_c))$ is a nontrivial submodule invariant under $\mathbb{Z}/p\mathbb{Z}$. This leads to a contradiction to that the matrix A is irreducible over $\mathbb{Z}_\ell$. We are done. □

Lemma 3.6. *The canonical homomorphism* $\iota : H^1_{\text{ét}}(\mathbb{A}^1_{\Bbbk}, \mathbb{F}_p) \to H^1_{\text{ét}}(K_\infty, \mathbb{F}_p)$ *is an isomorphism.*

Proof. We have the following commutative diagram of exact sequences derived from the celebrated Artin–Schreier exact sequence:

$$\begin{array}{ccccc} \Bbbk[t] & \xrightarrow{\text{id}-F} & \Bbbk[t] & \twoheadrightarrow & H^1_{\text{ét}}(\mathbb{A}^1_{\Bbbk}, \mathbb{F}_p) \\ \downarrow & & \downarrow & & \downarrow \iota \\ K_\infty & \xrightarrow{\text{id}-F} & K_\infty & \twoheadrightarrow & H^1_{\text{ét}}(K_\infty, \mathbb{F}_p) \end{array}$$

One easy observation is that we have the following exact sequence:

$$0 \to \Bbbk \to \Bbbk[t] \oplus \mathcal{O}_\infty \to K_\infty \to 0.$$

Here $\mathcal{O}_\infty = \mathcal{O}^{\text{sh}}_{\mathbb{P}^1_{\Bbbk},\infty}$ and $K_\infty = \text{Frac}(\mathcal{O}_\infty)$. As a consequence, by the 5-lemma:

$$\begin{array}{ccccc} \Bbbk & \hookrightarrow & \Bbbk[t] \oplus \mathcal{O}_\infty & \twoheadrightarrow & K_\infty \\ \downarrow \text{id}-F & & \downarrow \text{id}-F & & \downarrow \text{id}-F \\ \Bbbk & \hookrightarrow & \Bbbk[t] \oplus \mathcal{O}_\infty & \twoheadrightarrow & K_\infty \\ \downarrow & & \downarrow & & \downarrow \\ 0 & \longrightarrow & H^1_{\text{ét}}(\mathbb{A}^1_{\Bbbk}, \mathbb{F}_p) & \xrightarrow{\iota} & H^1_{\text{ét}}(K_\infty, \mathbb{F}_p) \end{array}$$

that ι is an isomorphism. Here we use the fact that $\mathcal{O}_\infty$ is strictly Henselian, hence

$$\mathcal{O}_\infty \xrightarrow{\text{id}-F} \mathcal{O}_\infty$$

is surjective. □

3B. *Examples with nontrivial fibration of* $s(f) = 2$.

Example 3.7. Suppose $p = 3$, $g = 1$, let X_0 be the relatively minimal surface over $\mathbb{A}^1_{\Bbbk}$ (with coordinate function t) associated to the following hyperelliptic fibration

$$y^2 = x^3 + x^2 + xt + t^2;$$

then X_0 has exactly one singular fibre at $t = 0$, over $\mathbb{A}^1_{\Bbbk}$. In particular, the proper relatively minimal model $X \to \mathbb{P}^1_{\Bbbk}$ associated has exactly 2 singular fibre, and it is nonisotrivial.

Proof. We can compute that $j(t) = 2/t^3$, so it is nonisotrival. □

Example 3.8. Suppose $p=5,\ g=2$, and let X_0 be the relatively minimal surface over $\mathbb{A}^1_{\Bbbk}$ associated to the following hyperelliptic fibration:

$$y^2=x^5+x^4-2t^2x^2+t^4.$$

Then X_0 has exactly one singular fibre at $t=0$, over $\mathbb{A}^1_{\Bbbk}$. In particular, the proper relatively minimal model $X\to\mathbb{P}^1_{\Bbbk}$ associated has exactly 2 singular fibre, and it is nonisotrivial.

Proof. Similar as the above example, we can use Igusa's J-invariants [Igusa 1960; Liu 1994] to deduce the nonisotriviality: $J_2=4t^2,\ J_4=0,\ J_6=t^8,\ J_{10}=4t^{14}$. □

3C. *Hyperelliptic fibration with $s(f)=1$.* We classify those hyperelliptic fibrations $f:X\to\mathbb{P}^1_{\Bbbk}$ that has a single singular fibre at ∞ when $g\geq 2$. Denote by F_p the smooth projective hyperelliptic curve associated to the following hyperelliptic equation $y^2=x^p-x$. According to [Roquette 1970, p. 158; Homma 1980/81, Theorem 2], F_p is the unique curve of genus $g=(p-1)/2$ admitting an automorphism of order p. Its actual automorphism group is

$$\mathrm{Aut}(F_p)=\mathrm{PGL}_2(\mathbb{F}_p)\times\mathbb{Z}/2\mathbb{Z}.$$

Here the direct summand $\mathbb{Z}/2\mathbb{Z}$ is associated to the hyperelliptic involution and $\mathrm{PGL}_2(\mathbb{F}_p)\subseteq\mathrm{PGL}_2(\Bbbk)=\mathrm{Aut}(\mathbb{P}^1_{\Bbbk})$ is the automorphism of the underlying $\mathbb{P}^1_{\Bbbk}$ fixing the branch locus defined by the equation x^p-x and the infinity.

Lemma 3.9. *Up to conjugate, there are two nontrivial subgroups $\Gamma_i,\ i=1,2$ of $\mathrm{PGL}_2(\mathbb{F}_p)$ as below can be realized as the Galois group of an étale cover of $\mathbb{A}^1_{\Bbbk}$:*

(1) Γ_1 *is the p-cyclic subgroup generated by $\left(\begin{smallmatrix}1,1\\0,1\end{smallmatrix}\right)$.*

(2) $\Gamma_2=\mathrm{PSL}_2(\mathbb{F}_p)\subseteq\mathrm{PGL}_2(\mathbb{F}_p)$.

Proof. By [Raynaud 1994, Corollary 2.2.2], a finite group G can be realized as a Galois group of $\mathbb{A}^1_{\Bbbk}$ if and only if $G=p(G)$, here $p(G)$ is the subgroup of G generated by all Sylow-p subgroups. First we note that the Sylow-p subgroup of $\mathrm{PGL}_2(\mathbb{F}_p)$ is a cyclic group of order p. Next, we consider the canonical faithful action of $\mathrm{PGL}_2(\mathbb{F}_p)$ on the $(p+1)$-set $\mathbb{P}^1(\mathbb{F}_p)$. One observes that there is a one-one correspondence between points in $\mathbb{P}^1(\mathbb{F}_p)$ and the set of Sylow-p subgroups of $\mathrm{PGL}_2(\mathbb{F}_p)$. In fact, the p-Sylow subgroup of $\mathrm{PGL}_2(\mathbb{F}_p)$ is a cyclic group of order p, it must have a unique fixed point in $\mathbb{P}^1(\mathbb{F}_p)$. Conversely, for any point in $\mathbb{P}^1(\mathbb{F}_p)$ its isotropic subgroup contains a unique p-subgroup. Now let Γ be a subgroup of $\mathrm{PGL}_2(\mathbb{F}_p)$ with $p(\Gamma)=\Gamma$.

Case Γ is contained in an isotropic subgroup. We may assume this point is $t=\infty$, and we have Γ_1 in this case.

Case Γ is not contained in any isotropic subgroups. In this case, Γ must act transitively on $\mathbb{P}^1(\mathbb{F}_p)$. In fact, any Sylow-$p$ group has exact one fixed point, as Γ is not contained in any isotropic subgroup, it is transitive. Now using the correspondence between Sylow-p subgroups and $\mathbb{P}^1(\mathbb{F}_p)$, Γ contains all the Sylow-p subgroups of $\mathrm{PGL}_2(\mathbb{F}_p)$. Finally it follows from the simplicity of $\mathrm{PSL}_2(\mathbb{F}_p)$, we have $p(\mathrm{PGL}_2(\mathbb{F}_p)) = \mathrm{PSL}_2(\mathbb{F}_p)$. □

Theorem 3.10. *Suppose $f : X \to \mathbb{P}^1_{\Bbbk}$ is a nontrivial hyperelliptic fibration of fibre genus g such that $p = 2g+1$. If f is smooth over $\mathbb{A}^1_{\Bbbk}$, then there is an étale Galois cover $\pi : C_0 \to \mathbb{A}^1_{\Bbbk}$ with Galois group Γ_i, $i = 1$ or 2 (as in the lemma above) such that $f_0 : X_0 = X \times_{\mathbb{P}^1_{\Bbbk}} \mathbb{A}^1_{\Bbbk} \to \mathbb{A}^1_{\Bbbk}$ is obtained by*

$$\begin{array}{ccc} C_0 \times F_p & \xrightarrow{\pi} & X_0 = C_0 \times F_p / \Gamma_i \\ \downarrow p_1 & & \downarrow f_0 \\ C_0 & \longrightarrow & \mathbb{A}^1_{\Bbbk} = C_0 / \Gamma_i \end{array}$$

Here the Γ_i-action on $C_0 \times F_p$ is the canonical diagonal one.

Proof. By Theorem 3.1, such a fibration has to be isotrivial. So there is an étale Galois cover $\pi : C_0 \to \mathbb{A}^1_{\Bbbk}$ of Galois group Γ such that $f_0' : X \times_{\mathbb{P}^1_{\Bbbk}} C_0 \to C_0$ is trivial. Namely we have an C_0-isomorphism $X \times_{\mathbb{P}^1_{\Bbbk}} C_0 \simeq F \times_{\Bbbk} C_0$. The natural Γ-action on the left hand side then induces an action on $C_0 \times F$. Such an action must be diagonal since the automorphism group of F is discrete as $g(F) \geq 2$. Shrinking Γ if necessary at first, we may actually assume the Γ-action on F side is faithful. Then one must have $\Gamma \subset \mathrm{Aut}(F)$ is nontrivial (otherwise f is trivial) and $p(\Gamma) = \Gamma$. In particular, F admits an automorphism group of order p, so F is isomorphic to F_p by [Homma 1980/81] as mentioned before. We then complete our proof by the previous lemma. □

Remark 3.11. From the theorem, we obtain a one-one correspondence between

- nontrivial hyperelliptic fibration of genus g over $\mathbb{P}^1_{\Bbbk}$ with a unique singular fibre at infinity, and
- étale Galois cover of $\mathbb{A}^1_{\Bbbk}$ with Galois group isomorphic to Γ_i, $i = 1, 2$.

Moreover, given a nontrivial smooth hyperelliptic fibration $f : X_0 \to \mathbb{A}^1_{\Bbbk}$ of genus $g = (p-1)/2$, the hyperelliptic involution $\sigma : X_0 \to X_0$ induces a flat double cover $\rho : X_0 \to X_0/\sigma \simeq \mathbb{P}^1 \times_{\Bbbk} \mathbb{A}^1_{\Bbbk}$. The branch divisor $W \subseteq X_0/\sigma \simeq \mathbb{P}^1 \times_{\Bbbk} \mathbb{A}^1_{\Bbbk}$ of ρ is by construction an étale cover of degree $p+1$ over the base $\mathbb{A}^1_{\Bbbk}$. This Galois group is the Galois group of the smallest étale Galois cover $\pi : C_0 \to \mathbb{A}^1_{\Bbbk}$ we need to trivialize $f : X_0 \to \mathbb{A}^1_{\Bbbk}$. The above theorem then tells us there are only three possibilities for the Galois group of $W/\mathbb{A}^1_{\Bbbk}$:

(1) The Galois group is trivial, which means the fibration f is also trivial.

(2) The Galois group is isomorphic to Γ_1, which gives the case Γ_1 in Theorem 3.10.

(3) The Galois group is isomorphic to Γ_2, which gives the case Γ_2 in Theorem 3.10.

Example 3.12. Suppose we have case Γ_1 in Theorem 3.10, then the equation of $f_0 : X_0 \to \mathbb{A}^1_{\Bbbk}$ can be written as

$$y^2 = x^p - x + v(t),\ v(t) \in \Bbbk[t]$$

with $v(t)$ not equal to $u(t)^p - u(t)$ for all $u(t) \in \Bbbk[t]$. The converse is also true.

Proof. Given the equation $y^2 = x^p - x + v(t)$, it is easily to see that f is smooth and the branch divisor W is disjoint union of two irreducible components: one is the infinite section and the other one defined by $x^p - x + v(t) = 0$. As a result, the Galois group of $W/\mathbb{A}^1_{\Bbbk}$ is isomorphic to Γ_1.

Conversely, if some fibration $f : X_0 \to \mathbb{A}^1_{\Bbbk}$ meets the case Γ_1 in Theorem 3.10, then $W \subseteq \mathbb{P}^1 \times_{\Bbbk} \mathbb{A}^1_{\Bbbk}$ is a disjoint union of a section and an (open) Artin–Schreier curve. By suitably change the coordinate of $\mathbb{P}^1 \times_{\Bbbk} \mathbb{A}^1_{\Bbbk}$, we can assume the section is the infinite the section. As a result, the second component is an Artin–Schreier curve contained in $\mathbb{A}^1_{\Bbbk} \times_{\Bbbk} \mathbb{A}^1_{\Bbbk}$, which must be defined by equation of formation $x^p - x + v(t)$. We are done. □

Example 3.13. The smooth hyperelliptic fibration defined by the following equation:

$$y^2 = x^{p+1} + tx + 1$$

meets the case of Γ_2 in Theorem 3.10.

Proof. We can directly work out $\mathrm{Disc}(x^{p+1} + tx + 1, x^p + t) = 1$, so the fibration is smooth over $\mathbb{A}^1_{\Bbbk}$. This time, the branch divisor W is defined by $x^{p+1} + tx + 1 = 0$ so the Galois group of $W/\mathbb{A}^1_{\Bbbk}$ is clearly not trivial or of order p. □

4. Number of singular fibres, II: $p < 2g + 1$

In this section, we provide examples of hyperelliptic fibrations $f : X \to \mathbb{P}^1_{\Bbbk}$ of genus g with $s(f) = 1$ when $p < 2g + 1$.

Example 4.1. Suppose $p \geq 3$, $m \geq 1$ and $(m, p) = 1$, and let X_0 be the relatively minimal surface over $\mathbb{A}^1_{\Bbbk}$ (with coordinate function t) associated to the following hyperelliptic function:

$$y^2 = x^{p+2m} + tx^{2m} - 1.$$

Then X_0 is smooth and nonisotrivial over $\mathbb{A}^1_{\Bbbk}$ with fibre genus $g = m + \frac{p-1}{2}$.

Proof. The variety X_0 is by construction a flat double cover of $\mathbb{P}^1_{\Bbbk} \times \mathbb{A}^1_{\Bbbk}$ where the x, t are the coordinate function of the vertical $\mathbb{P}^1_{\Bbbk}$ and horizontal $\mathbb{A}^1_{\Bbbk}$ respectively. The branch divisor of this flat double cover consists of the infinite section $B_0 = \infty \times \mathbb{A}^1_{\Bbbk}$

and another divisor B_1 defined by $x^{p+2m}+tx^{2m}-1$. Take $v(x)=x^{p+2m}+tx^{2m}-1$, then $v(x)-\frac{x}{2m}v'(x)=-1$ and hence B_1 is smooth over $\mathbb{A}^1_{\Bbbk}$. As $B_0\cap B_1=\varnothing$, X_0 is smooth over $\mathbb{A}^1_{\Bbbk}$. To show the nontriviality, we first note that X_0 is clearly not trivial over $\mathbb{A}^1_{\Bbbk}$. Then if X_0 is isotrivial, there must be an automorphism σ_t of each fibre $\mathbb{P}^1_{\Bbbk}\times t$ of $\mathbb{P}^1_{\Bbbk}\times\mathbb{A}^1_{\Bbbk}$ preserving the branch divisors for each $t\in\mathbb{A}^1_{\Bbbk}$. Since $\infty\in\mathbb{P}^1_{\Bbbk}\times t$ is the common branch point, we can assume σ_t fixes ∞. In particular, σ_t is an automorphism of $\mathbb{A}^1_{\Bbbk}\simeq\mathbb{A}^1_{\Bbbk}\times t$ of order p. However, any order p automorphism of $\mathbb{A}^1_{\Bbbk}$ cannot have any fixed point, its orbit all have cardinality p and hence the cardinality of the branch points lying in $\mathbb{A}^1_{\Bbbk}\times t\subseteq\mathbb{P}^1_{\Bbbk}\times t$ is a multiple of p. A contradiction of $p\nmid p+2m$. □

Example 4.2. Suppose $p\geq 3,\ m=2n+1,\ n=1,2,\dots$ and such that $(m,p)=1$, and let X_0' be the relatively minimal surface over $\mathbb{A}^1_{\Bbbk}$ associated to the following hyperelliptic function

$$y^2=x(x^{p+m}+tx^m-1).$$

Then X_0 is smooth, nonisotrivial over $\mathbb{A}^1_{\Bbbk}$ with fibre genus $g=\frac{m+1}{2}+\frac{p-1}{2}$.

Proof. One can prove following the same way as in the previous example. In this case, X_0 is realized as a flat double cover of $\mathbb{P}^1_{\Bbbk}\times\mathbb{A}^1_{\Bbbk}$ branching at $B_0=0\times\mathbb{A}^1_{\Bbbk}$, $B_1=\infty\times\mathbb{A}^1_{\Bbbk}$ and B_2 defined by $v(x)=x^{p+m}+tx^m-1$. As $(v(x),v'(x))=1$, B_2 is smooth over the base $\mathbb{A}^1_{\Bbbk}$ and it is clear that these three divisors are disjoint from each other, we see that X_0 is smooth over $\mathbb{A}^1_{\Bbbk}$. The nonisotriviality follows from the argument of the previous proof. In fact, if it is not, we shall similarly obtain an automorphism σ_t of $\mathbb{P}^1_{\Bbbk}\simeq\mathbb{P}^1_{\Bbbk}\times t$ fixing $0,\infty$ of order p. This is clearly not possible. □

Remark 4.3. The genus occurring in Examples 4.1 and 4.2 exhausts the integers strictly larger than $\frac{p-1}{2}$.

Example 4.4. Suppose $p=2$, and $l\geq 3$ is odd integer. Let X_0 be the relatively minimal surface over $\mathbb{A}^1_{\Bbbk}$ associated to the following hyperelliptic function

$$y^2+y=x^l(x^2+t);$$

then $f_0: X_0\to\mathbb{A}^1_{\Bbbk}$ has no singular fibre and is nonisotrivial, its fibre genus is $g=(l+1)/2$. In particular, the proper relatively minimal model $X\to\mathbb{P}^1_{\Bbbk}$ associated has exactly 1 singular fibre, and it is a nonisotrival fibration.

Proof. We consider the relative plane curve $X_1\subsetneq\mathbb{P}^2\times\mathbb{A}^1_{\Bbbk}$ over $\mathbb{A}^1_{\Bbbk}$ defined by

$$Y^2Z^l+YZ^{l+1}=X^{l+2}+tX^lZ^2.$$

Here X,Y,Z are the homogeneous coordinates of $\mathbb{P}^2$ and t is the coordinate of $\mathbb{A}^1_{\Bbbk}$. By construction, X_0 is the desingularization of X_1 and we carry out this desingularization by blow-ups. It is clear that the singular locus of X_1 is the closed subset defined by $Z = X = 0$. In particular, we only have to work over the affine plane $Y = 1$ over $\mathbb{A}^1_{\Bbbk}$ and we can write out the function of X_1 as

$$z_0^l + z_0^{l+1} = x_0^{l+2} + t x_0^l z_0^2.$$

Here $x_0 = X/Y$, and $z_0 = Z/Y$. Blowing up the ideal (x_0, z_0), we obtain the new affine defining equation,

$$z_1^l + z_1^{l+1} x_0 = x_0^2 + t x_0^2 z_1^2,$$

here $z_1 = z_1/x_0$. Again the singular locus is $z_1 = x_0 = 0$. Blowing up (z_1, x_0) we have:

$$z_1^{l-2} + z_1^l x_1 = x_1^2 + t x_1^2 z_1^2$$

with $x_1 = x_0/z_1$. Keeping on blowing up the singular locus again and again (adding $x_{i+1} = x_i/z_1$ each step), we will finally get with $m = \frac{l-1}{2}$,

$$z_1^{l-2k} + z_1^{l+1-k} x_m = x_m^2 + t x_m^2 z_1^2$$

namely,

$$z_1 + z_1^{(l+3)/2} x_m = x_m^2 + t x_m^2 z_1^2.$$

It is not only regular, but smooth over $\mathbb{A}^1_{\Bbbk}$.

Next, we show that this fibration is not isotrivial. It suffices to show that the two fibres with $t = 0$ and $t = 1$ are not isomorphic. For each t, we write F_t to indicate the associated fibre. We see from the construction:

> The double cover given by affine equation: $y^2 + y = x^l(x^2 + t)$ from the affine open subset $U_t = \operatorname{Spec}\big(\Bbbk[x, y]/(y^2 + y + x^l(x^2 + t))\big)$ of F_t to the affine line $\mathbb{A}^1_{\Bbbk} = \operatorname{Spec}(\Bbbk[x])$ is intrinsic. In fact, this morphism is the canonical morphism F_t and U_t is étale locus of the canonical morphism. In other word, to get an isomorphism of F_0 and F_1, we must have a linear transformation:
>
> $$\begin{array}{ccc} U_1 & \xrightarrow{\nu_1} & \mathbb{A}^1_{\Bbbk} = \operatorname{Spec}(\Bbbk[x_1]) \\ \downarrow \sim & & \downarrow \varphi \\ U_0 & \xrightarrow{\nu_0} & \mathbb{A}^1_{\Bbbk} = \operatorname{Spec}(\Bbbk[x_0]) \end{array}$$
>
> with φ being a linear isomorphism.

So we have the following transformation:

$$\begin{cases} x_0 = a(x_1 + b), \\ y_0 = c y_1 + v(x_1), \end{cases}$$

with $a, c \in \Bbbk^*$ and $v(x) \in \Bbbk[x]$. Note that the canonical involution transforms preserve x and maps $y_i \mapsto y_i + 1$, we have $c = 1$. Now compare the functions at $t = 1$ and $t = 0$, we have

$$a^{l+2}(x_1 + b)^{l+2} = x_1^{l+2} + x_1^l + v^2 + v.$$

Compare the leading coefficient and notice that $l + 2$ is an odd number, we have $a^{l+2} = 1$. So

$$bx^{l+1} + \left(\binom{l+2}{2}b + 1\right)x^l + (\text{low degree terms}) = v^2 + v.$$

As l is odd and $l > \frac{l+1}{2}$, one must have $\binom{l+2}{2}b + 1 = 0$. As a consequence, we have $b = 1$. By now we have

$$(x+1)^{l+2} + x^{l+2} + x^l = v^2 + v.$$

By symmetry we shall have

$$(x+1)^{l+2} + x^{l+2} + (x+1)^l = u^2 + u$$

for $u(x) = v(x+1)$. And as a consequence,

$$(x+1)^l + x^l = (u+v)^2 + (u+v).$$

One clearly checks that $u, v \in \mathbb{F}_2[x]$, in particular we have $u(1) + v(1) \in \mathbb{F}_2$. So we have

$$1 = 0^l + 1^l = (u(1) + v(1))^2 + (u(1) + v(1)) = 0,$$

a contradiction. □

5. Small genus case I: $g = 2$

Let $f : X \to \mathbb{P}^1_{\Bbbk}$ be a nonisotrivial relatively minimal fibration of genus 2 in characteristic p. We already have the following lower bounds:

(1) if $p > 5$, then f has at least 3 singular fibres,

(2) if $p = 5$, then f has at least 2 singular fibres,

from Theorem 1.1 and Corollary 3.2. In this section we give examples in characteristic $p = 2, 3$ of nonisotrivial genus 2 fibration $f : X \to \mathbb{P}^1_{\Bbbk}$ with only one singular fibre at $\infty \in \mathbb{P}^1_{\Bbbk}$.

5A. *Characteristic $p = 3$.* Suppose $p = 3$ and f is smooth over $\mathbb{A}^1_{\Bbbk}$, the Weierstrass divisor $W \subsetneq X_0 := X \times_{\mathbb{P}^1_{\Bbbk}} \mathbb{A}^1_{\Bbbk}$ is then an étale cover of $\mathbb{A}^1_{\Bbbk}$ of degree 6. It then gives a continuous homomorphism $\pi : \pi_1^{\text{ét}}(\mathbb{A}^1_{\Bbbk}, \eta) \to S_6$, the latter is the 6-th permutation group. By [Raynaud 1994, Corollary 2.2.2] and a simple group-theoretic computation, the image of π in S_6 up to conjugation can only be A_4, A_5, A_6.

Note that there are two different embeddings of A_5 in S_6, the canonical one and the exceptional one. The exceptional embedding $A_5 \subsetneq S_6$ is induced by the celebrated exceptional embedding $S_5 \to S_6$. One may find this embedding as: $S_5 \simeq \mathrm{PGL}_2(\mathbb{F}_5)$ and the latter acts naturally on the 6-element set $\mathbb{P}^1(\mathbb{F}_5)$ transitively.

We then give for each case an example of such f. Here we remind that t is the affine coordinate of the base $\mathbb{A}^1_{\Bbbk}$.

(1) $[A_4]$: The surface fibration defined by hyperelliptic equation

$$y^2 = x(x^4 - tx + 1).$$

Its Igusa's J-invariants are: $J_2 = 2,\ J_4 = 0,\ J_6 = t^4 + 1,\ J_{10} = 1.$

(2) $[A_5]$ (canonical): The surface fibration defined by hyperelliptic equation

$$y^2 = x^5 + tx^3 + 1.$$

Its Igusa's J-invariants are: $J_2 = 0,\ J_4 = 0,\ J_6 = t^6 + 2t,\ J_{10} = 2.$

(3) $[A_5]$ (exceptional): The surface fibration defined by hyperelliptic equation

$$y^2 = x^6 + tx^3 + x + t.$$

Its Igusa's J-invariants are: $J_2 = 0,\ J_4 = 0,\ J_6 = t^6 + 2t^3 + 2t,\ J_{10} = 2.$

(4) $[A_6]$: The surface fibration defined by hyperelliptic equation

$$y^2 = x^6 + tx^3 + x + 1.$$

Its Igusa's J-invariants are: $J_2 = 0,\ J_4 = 0,\ J_6 = t^6 + 2t + 2,\ J_{10} = 2.$

The smoothness over $\mathbb{A}^1_{\Bbbk}$ of each example follows from the definition of J_{10}, which is the discriminant of the right hand side polynomial of x. Namely, $J_{10} \in \Bbbk^*$ if and only if it is smooth over $\mathbb{A}^1_{\Bbbk}$. The nonisotriviality directly follows from Igusa's invariants we give here. We also mention that the Galois groups here are calculated by the computer software Magma Computational Algebra System.

5B. ***Characteristic $p = 2$.*** It has been given in Example 4.4 that the surface fibration defined by

$$y^2 + y = x^5 + tx^3$$

in characteristic 2 is a nonisotrivial genus 2 fibration smooth over $\mathbb{A}^1_{\Bbbk}$. One can work out its Igusa J-invariant: $J_2 = t^2,\ J_4 = t^4,\ J_6 = t^6,\ J_{10} = 1.$

6. Small genus case II: $g = 1$

We consider a relatively minimal genus 1 fibration $f : X \to \mathbb{P}^1_{\Bbbk}$ in this section. We call such a fibration f as an elliptic fibration if f admits an section. For any genus 1 fibration $f : X \to \mathbb{P}^1_{\Bbbk}$ we can associate an elliptic fibration $f' : Y \to \mathbb{P}^1_{\Bbbk}$, called as the associated Jacobian fibration to it. The Jacobian fibration g has the following properties:

- if f is smooth over an open subset U, so is f';
- the j-invariant of f coincides with that of f'.

Due to these two properties, when we study the lower bound of $s(f)$ for nonisotrivial f, it suffices to study the elliptic fibrations.

6A. ***Number of singular fibres for nonisotrivial fibrations.*** In this section, we discuss the number of singular fibres for a nonisotrivial elliptic fibration.

Theorem 6.1 [Beauville 1981, p. 99]. *Suppose $f : E \to \mathbb{P}^1_{\Bbbk}$ is a relatively minimal elliptic fibration and f is nonisotrivial. Then*

(1) *f has at least 3 singular fibres if $p \geq 5$;*

(2) *f has at least 2 singular fibres if $p = 2, 3$.*

Part (1) is proved in [Beauville 1981] and we follow his idea to prove (2).

Proof. We assume the contrary that f is nonisotrivial. The j-invariant then induces a surjective homomorphism:

$$j : \mathbb{P}^1_{\Bbbk} \to \mathbb{P}^1_{\Bbbk}$$

which decomposes into

$$\begin{array}{ccc} \mathbb{P}^1_{\Bbbk} & \xrightarrow{\quad h \quad} & \mathbb{P}^1_{\Bbbk} \\ & {\scriptstyle j}\searrow \quad \swarrow{\scriptstyle j'} & \\ & \mathbb{P}^1_{\Bbbk} & \end{array}$$

where h is purely inseparable and j' is separable. Let $d := \deg(j')$.

Lemma 6.2. *If $p = 2, 3$, then j' is wildly ramified at any point in $j'^{-1}(0)$.*

Proof. Replace $\mathbb{A}^1_{\Bbbk}$ by a suitable étale cover $\nu : C_0 \to \mathbb{A}^1_{\Bbbk}$, we may assume the elliptic curve $X \times_{\mathbb{P}^1_{\Bbbk}} C_0 \to C_0$ admits a level structure of level 5. Denote by $M_{1,5}$ the fine moduli space of elliptic curve with level 5 structure. We therefore have the following commutative diagram:

$$\begin{array}{ccc} C_0 & \xrightarrow{\ \nu\ } & \mathbb{A}^1_{\Bbbk} \\ \downarrow & & \downarrow{\scriptstyle j} \\ M_{1,5} & \xrightarrow{\ \tau\ } & M_1 = \mathbb{A}^1_{\Bbbk} \end{array}$$

It then follows directly that τ is wildly ramified at any point in $\tau^{-1}(0)$. □

Once the lemma is true, note that the ramification index at $j'^{-1}(\infty)$ is at least $d-1$, it then follows from Hurwitz formula,

$$-2 \geq -2d + (d-1) + d = -1.$$

Here the ramification index at infinity and 0 is at least $d-1$ and d (for sake of wild ramification), respectively. A contradiction. □

6B. ***Examples of elliptic fibrations over $\mathbb{P}^1_{\Bbbk}$ with one singular fibre.***

Proposition 6.3. *Suppose $\Bbbk$ is an algebraically closed field of characteristic 3 and $f : X \to \mathbb{P}^1_{\Bbbk}$ is an elliptic fibration smooth over $\mathbb{A}^1_{\Bbbk}$. Then either X is trivial or we can write out the equation of X as*

$$y^2 = x^3 - x + v(t),$$

for some $v(t) \in \Bbbk[t]$. Here t is the affine coordinate of the base $\mathbb{A}^1_{\Bbbk} \subsetneq \mathbb{P}^1_{\Bbbk}$. In particular, all smooth fibres are isomorphic to the unique supersingular elliptic curve $E_0 : y^2 = x^3 - x$.

Proof. Considering the quotient morphism by $\pm\,\mathrm{id}$ on $X_0 := X \times_{\mathbb{P}^1_{\Bbbk}} \mathbb{A}^1_{\Bbbk}$, it induces a flat double cover $\mu : X_0 \to P_0 := X_0/\{\pm\,\mathrm{id}\}$ ramified at the divisor $R := X_0[2]$. This divisor R is étale over $\mathbb{A}^1_{\Bbbk}$ of degree $2^2 = 4$ and contains a section, the identity section, so we write $R = R_0 + R_1$ with R_0 being the identity section. Now the branch divisor $B = \mu(R)$ is a relative divisor in $P \simeq \mathbb{P}^1 \times \mathbb{A}^1_{\Bbbk}$ isomorphic to R. So we can write $B = B_0 + B_1$ with $B_0 = \mu(R_0)$ such that B_0 is a section and B_1 is a degree 3 étale cover of $\mathbb{A}^1_{\Bbbk}$. After changing of the relative linear coordinate x of $\mathbb{P}^1 \times \mathbb{A}^1_{\Bbbk}$, we may assume B_0 is the infinite section. Then $X_0 \backslash \{\text{identity}\}$ is given by a hyperelliptic equation

$$y^2 = F(x,t) \in \Bbbk[x,t]$$

with $F(x,t)$ being the defining equation of B_1 in $\mathbb{A}^1_{\Bbbk} \times_{\Bbbk} \mathbb{A}^1_{\Bbbk}$. Here x, t are the linear coordinate of the two $\mathbb{A}^1_{\Bbbk}$-factors respectively.

Note as S_3 can not be realized as a Galois cover of $\mathbb{A}^1_{\Bbbk}$, there are two possibilities:

(A) B_1 is a disjoint union of three sections.

(B) $B_1 \to \mathbb{A}^1_{\Bbbk}$ is a Galois cover.

Case (A) directly leads us to the triviality of X. Now for (B) where B_1 is an Artin–Schreier curve of degree 3. It is well known that the embedded Artin–Schreier curves in $\mathbb{A}^1_{\Bbbk} \times \mathbb{A}^1_{\Bbbk}$ has defining equation of form

$$x^3 - x + v(t) = 0.$$

We are done then. □

Using Tate's algorithm [Silverman 1994, §IV.9], we can see the special fibre at ∞ for the equation $y^2 = x^3 - x + v(t)$ is of Kodaira type

(II) if $v(t) = t^5$;

(IV) if $v(t) = t^4$;

(II*) if $v(t) = t$;

(IV*) if $v(t) = t^2$.

Note that these four types of reduction are all the possibility of potentially good reduction type where the good reduction is realized by a cyclic cover of degree 3 (see, e.g., [Lorenzini 2010]).

Example 6.4 (Type I_n^* with potentially good reduction). Taking an algebraically closed field $\Bbbk$ of characteristic 2 and let S be the elliptic fibration over $\mathbb{A}^1_{\Bbbk}$ defined by the equation

$$y^2 + xy = x^3 + tx^2 + 1.$$

One easily works out the discriminant of this equation is 1 and hence smooth over $\mathbb{A}^1_{\Bbbk}$. By Theorem 6.1, it must have potentially good reduction at $\infty \in \mathbb{P}^1_{\Bbbk}$. Applying Tate's algorithm [Silverman 1994, §IV.9], one can show its reduction type at ∞ is I_4^*.

As is well known when $p \geq 5$, a type I_n^* fibre has potentially bad reduction (see, e.g., [Lorenzini 2010]).

References

[Beauville 1981] A. Beauville, "Le nombre minimum de fibres singulières d'une courbe stable sur $\mathbb{P}^1$", pp. 97–108 in *Séminaire sur les pinceaux de courbes de genre au moins deux*, Astérisque **86**, 1981. MR Zbl

[Bosch et al. 1990] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Ergebnisse der Mathematik (3) **21**, Springer, 1990. MR Zbl

[Deligne and Mumford 1969] P. Deligne and D. Mumford, "The irreducibility of the space of curves of given genus", *Inst. Hautes Études Sci. Publ. Math.* **36** (1969), 75–109. MR Zbl

[Fu 2015] L. Fu, *Etale cohomology theory*, Revised ed., Nankai Tracts in Mathematics **14**, World Scientific, Hackensack, NJ, 2015. MR Zbl

[Gong et al. 2013] C. Gong, X. Lu, and S.-L. Tan, "Families of curves over $\mathbb{P}^1$ with 3 singular fibers", *C. R. Math. Acad. Sci. Paris* **351**:9-10 (2013), 375–380. MR

[Homma 1980/81] M. Homma, "Automorphisms of prime order of curves", *Manuscripta Math.* **33**:1 (1980/81), 99–109. MR Zbl

[Igusa 1960] J.-i. Igusa, "Arithmetic variety of moduli for genus two", *Ann. of Math.* (2) **72** (1960), 612–649. MR Zbl

[Kovács 1997] S. J. Kovács, "On the minimal number of singular fibres in a family of surfaces of general type", *J. Reine Angew. Math.* **487** (1997), 171–177. MR Zbl

[Liu 1994] Q. Liu, "Modèles minimaux des courbes de genre deux", *J. Reine Angew. Math.* **453** (1994), 137–164. MR Zbl

[Liu 1996] K. Liu, "Geometric height inequalities", *Math. Res. Lett.* **3**:5 (1996), 693–702. MR Zbl

[Lorenzini 2010] D. Lorenzini, "Models of curves and wild ramification", *Pure Appl. Math. Q.* **6**:1, Special Issue: In honor of John Tate. Part 2 (2010), 41–82. MR Zbl

[Lu et al. 2016] X. Lu, S.-L. Tan, W.-Y. Xu, and K. Zuo, "On the minimal number of singular fibers with non-compact Jacobians for families of curves over $\mathbb{P}^1$", *J. Math. Pures Appl.* (9) **105**:5 (2016), 724–733. MR Zbl

[Lu et al. 2018] X. Lu, S.-L. Tan, and K. Zuo, "Singular fibers and Kodaira dimensions", *Math. Ann.* **370**:3-4 (2018), 1717–1728. MR Zbl

[Nguyen 1997] K. V. Nguyen, "On classification of elliptic fibrations with small number of singular fibres over a base of genus 0 and 1", *Proc. Japan Acad. Ser. A Math. Sci.* **73**:6 (1997), 103–104. MR Zbl

[Nguyen 1998] K. V. Nguyen, "A remark on semi-stable fibrations over $\mathbf{P}^1$ in positive characteristic", *Compositio Math.* **112**:1 (1998), 41–44. MR Zbl

[Raynaud 1994] M. Raynaud, "Revêtements de la droite affine en caractéristique $p > 0$ et conjecture d'Abhyankar", *Invent. Math.* **116**:1-3 (1994), 425–462. MR Zbl

[Roquette 1970] P. Roquette, "Abschätzung der Automorphismenanzahl von Funktionenkörpern bei Primzahlcharakteristik", *Math. Z.* **117** (1970), 157–163. MR Zbl

[Serre and Tate 1968] J.-P. Serre and J. Tate, "Good reduction of abelian varieties", *Ann. of Math.* (2) **88** (1968), 492–517. MR Zbl

[SGA 7_{I} 1972] A. Grothendieck, *Groupes de monodromie en géométrie algébrique, I: Exposés I–II, VI–IX* (Séminaire de Géométrie Algébrique du Bois Marie 1967–1969), Lecture Notes in Math. **288**, Springer, 1972. MR Zbl

[Silverman 1994] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics **151**, Springer, 1994. MR Zbl

[Szpiro 1979] L. Szpiro, "Sur le théorème de rigidité de Parsin et Arakelov, II", pp. 169–202 in *Journées de Géométrie Algébrique de Rennes* (Rennes, 1978), Astérisque **64**, Soc. Math. France, Paris, 1979. MR Zbl

[Tan 1995] S. L. Tan, "The minimal number of singular fibers of a semistable curve over $\mathbf{P}^1$", *J. Algebraic Geom.* **4**:3 (1995), 591–596. MR Zbl

[Tan et al. 2005] S.-L. Tan, Y. Tu, and A. G. Zamora, "On complex surfaces with 5 or 6 semistable singular fibers over $\mathbb{P}^1$", *Math. Z.* **249**:2 (2005), 427–438. MR Zbl

[Tu 2007] Y. Tu, "Surfaces of Kodaira dimension zero with six semistable singular fibers over $\mathbb{P}^1$", *Math. Z.* **257**:1 (2007), 1–5. MR Zbl

[Viehweg and Zuo 2001] E. Viehweg and K. Zuo, "On the isotriviality of families of projective manifolds over curves", *J. Algebraic Geom.* **10**:4 (2001), 781–799. MR Zbl

[Zamora 2012] A. G. Zamora, "Semistable genus 5 general type $\mathbb{P}^1$-curves have at least 7 singular fibres", *Note Mat.* **32**:2 (2012), 1–4. MR Zbl

Received August 21, 2019. Revised August 16, 2020.

CHENG GONG

DEPARTMENT OF MATHEMATICS
SOOCHOW UNIVERSITY
SUZHOU
JIANGSU PROVINCE
CHINA
cgong@suda.edu.cn

YI GU
DEPARTMENT OF MATHEMATICS
SOOCHOW UNIVERSITY
SUZHOU
JIANGSU PROVINCE
CHINA
sudaguyi2017@suda.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 2, 2020
https://doi.org/10.2140/pjm.2020.309.333

ASYMPTOTIC BEHAVIOR OF SOLUTIONS FOR SOME ELLIPTIC EQUATIONS IN EXTERIOR DOMAINS

ZONGMING GUO AND ZHONGYUAN LIU

This paper is concerned with the asymptotic behavior of solutions of the problems

$$-\Delta u = e^u \text{ in } \mathbb{R}^2\backslash B, \quad \int_{\mathbb{R}^2\backslash B} e^{u(x)}\,dx < \infty, \tag{0-1}$$

where $B = \{x \in \mathbb{R}^2 : |x| < 1\}$ is the unit ball of $\mathbb{R}^2$, and

$$\Delta^2 u = e^u \text{ in } \mathbb{R}^4\backslash B, \quad \int_{\mathbb{R}^4\backslash B} e^{u(x)}\,dx < \infty, \tag{0-2}$$

where $B = \{x \in \mathbb{R}^4 : |x| < 1\}$ is the unit ball of $\mathbb{R}^4$. It is seen that the asymptotic behavior of solutions for (0-1) and (0-2) is equivalent to the asymptotic behavior of singular solutions of the related problems (via the transformation $v(y) = u(x)$, $y = x/|x|^2$):

$$-\Delta_y v = |y|^{-4} e^v \text{ in } B\backslash\{0\}, \quad \int_{B\backslash\{0\}} |y|^{-4} e^{v(y)}\,dy < \infty \tag{0-3}$$

and

$$\Delta_y^2 v = |y|^{-8} e^v \text{ in } B\backslash\{0\}, \quad \int_{B\backslash\{0\}} |y|^{-8} e^{v(y)}\,dy < \infty, \tag{0-4}$$

respectively. We obtain the exact asymptotic behavior of solutions of (0-1) and (0-2) as $|x| \to \infty$. Meanwhile, we find that the singular solutions of the related problems (0-3) and (0-4) in $B\backslash\{0\}$ are asymptotic radial solutions and obtain the corresponding asymptotic behavior as $|y| \to 0$.

The research of Guo is supported by NSFC (No. 11171092, 11571093), the research of Liu is supported by NSFC (No. 11971147) and CPSF (No. 2019M662475).

MSC2020: 35B45, 35J40.

Keywords: asymptotic behavior of solutions, semilinear biharmonic problems, exterior domains, singular solutions.

1. Introduction

In this paper, we study the asymptotic behavior of solutions for the following problems:

$$-\Delta u = e^u \text{ in } \mathbb{R}^2\backslash B, \quad \int_{\mathbb{R}^2\backslash B} e^{u(x)}\,dx < \infty, \tag{1-1}$$

where $B = \{x \in \mathbb{R}^2 : |x| < 1\}$ is the unit ball of $\mathbb{R}^2$, and

$$\Delta^2 u = e^u \text{ in } \mathbb{R}^4\backslash B, \quad \int_{\mathbb{R}^4\backslash B} e^{u(x)}\,dx < \infty, \tag{1-2}$$

where $B = \{x \in \mathbb{R}^4 : |x| < 1\}$ is the unit ball in $\mathbb{R}^4$.

The equations in (1-1) and (1-2) have roots in conformal geometry. Let (M, g) be a complete Riemannian manifold. Associated to g, there are tensors such as the full curvature tensor R_g, the Ricci curvature tensor Ric_g and the scalar curvature S_g. The Laplace operator Δ_g is a well-known elliptic operator on M associated with the metric g. In dimension 4, the equation in (1-2) is closely related to the Q-curvature problem. The Q-curvature is similar to the scalar curvature in dimension 2. See [Chang and Yang 1995; 1997; Graham et al. 1992; Lin 1998; Martinazzi 2009; Xu 2006].

The structure of solutions of (1-1) and (1-2) in $\mathbb{R}^2$ and $\mathbb{R}^4$ respectively has been studied in [Chen and Li 1991; Lin 1998; Martinazzi 2009; Wei and Xu 1999; Wei and Ye 2008; Xu 2006]. For a solution $u \in C^4(\mathbb{R}^4)$ of the equation in (1-2), an important fact $-\Delta u \geq 0$ in $\mathbb{R}^4$ can be obtained. Using the moving-plane or moving-sphere arguments, Lin [1998] and Xu [2006] classified the solutions and obtained the asymptotic behavior of solutions as $|x| \to \infty$. Moreover, the singular solutions of the equation

$$-\Delta u = e^u \text{ in } B\backslash\{0\}, \quad \int_{B\backslash\{0\}} e^{u(x)}\,dx < \infty, \tag{1-3}$$

where B is the unit ball in $\mathbb{R}^2$, have also been studied in [Chou and Wan 1994] via the theory of complex variables. More precisely, Chou and Wan [1994] showed that the singular solutions of (1-3) are asymptotic radial solutions and obtain the asymptotic behavior of solutions as $|x| \to 0$. By the transformation

$$v(y) = u(x), \qquad y = \frac{x}{|x|^2},$$

we see that the problems (1-1) and (1-2) are equivalent to the problems

$$-\Delta_y v = |y|^{-4} e^v \text{ in } B\backslash\{0\}, \quad \int_{B\backslash\{0\}} |y|^{-4} e^{v(y)}\,dy < \infty \tag{1-4}$$

and

$$\Delta_y^2 v = |y|^{-8} e^v \text{ in } B\backslash\{0\}, \quad \int_{B\backslash\{0\}} |y|^{-8} e^{v(y)}\, dy < \infty, \tag{1-5}$$

respectively. The asymptotic behavior of solutions for (1-1) and (1-2) is equivalent to the asymptotic behavior of singular solutions for (1-4) and (1-5). We will obtain the exact asymptotic behavior of solutions of (1-1) and (1-2) as $|x| \to \infty$. Moreover, we will show that the singular solutions of (1-4) and (1-5) are asymptotic radial solutions and obtain the asymptotic behavior of the singular solutions as $|y| \to 0$ by using the theory of PDEs. We find that the study of (1-2) is more complicated than that of (1-1). To obtain the result similar to that of (1-1), we need to put an extra assumption on the solution to avoid the appearance of an extra fundamental solution of the operator Δ^2. Our main results of this paper are the following theorems:

Theorem 1.1. *Assume that $u \in C^2(\mathbb{R}^2\backslash B)$ is a solution of (1-1). Then*

$$\frac{u(x)}{\ln|x|} \to \alpha \quad as \quad |x| \to \infty, \tag{1-6}$$

where $\alpha < -2$.

Theorem 1.2. *Assume that $u \in C^4(\mathbb{R}^4\backslash B)$ is a solution of (1-2) and*

$$u(x) = o(|x|^2) \quad as \quad |x| \to \infty. \tag{1-7}$$

Then

$$\frac{u(x)}{\ln|x|} \to \alpha \quad as \quad |x| \to \infty, \tag{1-8}$$

$$-|x|^2 \Delta u(x) \to \frac{1}{2|\mathbb{S}^3|} \int_{\mathbb{R}^4\backslash B} e^{u(y)}\, dy + \kappa \quad as \quad |x| \to \infty, \tag{1-9}$$

where $\alpha < -4$, $|\mathbb{S}^3|$ is the surface area of the unit sphere, κ is a constant.

Remark 1.3. We will see from the proof that the conclusions of Theorems 1.1 and 1.2 are still true if we assume $u \in C^2(\mathbb{R}^2\backslash\bar{B})$ and $u \in C^4(\mathbb{R}^4\backslash\bar{B})$ respectively or $u \in C^2(\mathbb{R}^2\backslash\overline{B_R(0)})$ and $u \in C^4(\mathbb{R}^4\backslash\overline{B_R(0)})$ respectively for some $R > 1$, where and in the following, $B_R(0) = \{x \in \mathbb{R}^2 : |x| < R\}$ or $B_R(0) = \{x \in \mathbb{R}^4 : |x| < R\}$. Our assumptions in Theorems 1.1 and 1.2 are only for convenience of using some expressions in our calculations.

As an application of Theorem 1.2, we can consider the following problem:

$$\Delta^2 v = e^v \text{ in } B\backslash\{0\}, \quad \int_{B\backslash\{0\}} e^{v(y)}\, dy < \infty, \tag{1-10}$$

where $B = \{y \in \mathbb{R}^4 : |y| < 1\}$ and obtain the asymptotic radial symmetry result for (1-10) in the punctured ball.

Theorem 1.4. *Assume that $v \in C^4(B\backslash\{0\})$ is a singular solution of the problem (1-10) with*

$$v(y) = o(|y|^{-2}) \quad as \quad |y| \to 0.$$

Then

$$\frac{v(y)}{\ln|y|} \to \gamma \quad as \quad |y| \to 0,$$

where $\gamma > -4$.

Similar results in $\mathbb{R}^4$ are well-known in [Lin 1998; Xu 2006]. Liouville theorem for harmonic functions plays the key role in obtaining these results in $\mathbb{R}^4$. However, the corresponding Liouville theorem does not hold in $\mathbb{R}^4 \setminus B$ and the methods in [Lin 1998; Xu 2006] cannot be used here. Moreover, we cannot show $-\Delta u \geq 0$ in $\mathbb{R}^4\backslash B$ for a solution $u \in C^4(\mathbb{R}^4\backslash B)$ of (1-2). To this end, we need to overcome some technical difficulties here and use some new idea to obtain the corresponding results in $\mathbb{R}^4\backslash B$.

The organization of the paper is the following: In Section 2, we give some qualitative properties of solutions for (1-2). The main results will be obtained in Section 3. In the Appendix, we present some estimates used in Section 3.

2. Preliminaries

In this section, we study the qualitative properties of solutions for (1-2). This is crucial to the proof of Theorem 1.2.

Let $u \in C^4(\mathbb{R}^4\backslash B)$ be a solution of problem (1-2) and

$$v(y) = u(x), \qquad y = \frac{x}{|x|^2}.$$

Then $v \in C^4(B\backslash\{0\})$ satisfies the problem

$$\Delta_y^2 v = |y|^{-8} e^v \text{ in } B\backslash\{0\}, \qquad \int_{B\backslash\{0\}} |y|^{-8} e^{v(y)}\, dy < \infty, \tag{2-1}$$

where $B \subset \mathbb{R}^4$ is the unit ball. Moreover,

$$v(y) = o(|y|^{-2}) \quad \text{as} \quad |y| \to 0. \tag{2-2}$$

It is easy to see that 0 is a nonremovable singular point of v. Using the fact that

$$\int_{B\backslash\{0\}} |y|^{-8} e^{v(y)}\, dy = \int_{\mathbb{R}^4\backslash B} e^u\, dx,$$

we have

$$\infty > \int_{B\backslash\{0\}} |y|^{-8} e^{v(y)}\, dy = |\mathbb{S}^3| \int_0^1 \rho^{-5} \overline{e^v}\, d\rho \geq |\mathbb{S}^3| \int_0^1 \rho^{-5} e^{\bar{v}}\, d\rho, \tag{2-3}$$

where $\rho = |y|$, $|\mathbb{S}^3|$ is the surface area of the unit sphere and

$$\bar{v}(\rho) := \frac{1}{|\mathbb{S}^3|} \int_{\mathbb{S}^3} v(\rho, \theta)\, d\theta \quad \text{for all } \rho \in (0, 1).$$

In the following, we first consider the asymptotic behavior of $\bar{v}(\rho)$ as ρ tends to 0.

Lemma 2.1. *Let $v \in C^4(B \backslash \{0\})$ be a solution of (2-1) satisfying (2-2). Then*

$$\frac{\bar{v}(\rho)}{\ln \rho} \to \beta \quad \textit{as} \quad \rho \to 0, \tag{2-4}$$

where $\beta > 4$.

Proof. Note that $\bar{v}(\rho)$ satisfies the problem

$$\Delta^2 \bar{v} = \rho^{-8} \overline{e^v} \text{ in } (0, 1), \qquad \int_0^1 \rho^{-5} \overline{e^v}(\rho)\, d\rho < \infty.$$

By (2-2), we find

$$\bar{v}(\rho) = o(\rho^{-2}). \tag{2-5}$$

Step 1: We claim that if $\lim_{\rho \to 0} \rho^3 \bar{v}'(\rho)$ exists, then it must be 0, i.e.,

$$\lim_{\rho \to 0} \rho^3 \bar{v}'(\rho) = 0. \tag{2-6}$$

On the contrary, there is $M \neq 0$ (M maybe $\pm\infty$) such that $\lim_{\rho \to 0} \rho^3 \bar{v}'(\rho) = M$. We consider two cases:

(i) $M > 0$,

(ii) $M < 0$.

For the case (i), we have that there exist $M_0 > 0$ and $\rho_0 > 0$ such that

$$\bar{v}'(\rho) \geq M_0 \rho^{-3} \text{ for all } \rho \in (0, \rho_0). \tag{2-7}$$

Integrating (2-7) on (ρ, ρ_0), we obtain

$$\bar{v}(\rho) \leq -\tfrac{1}{2} M_0 \rho^{-2} + \bar{v}(\rho_0) + \tfrac{1}{2} M_0 \rho_0^{-2} \text{ for all } \rho \in (0, \rho_0), \tag{2-8}$$

which is a contradiction with (2-5).

For the case (ii), we have that there exist $\rho_0 > 0$ and $M_0 < 0$ such that

$$\bar{v}'(\rho) \leq M_0 \rho^{-3} \text{ for all } \rho \in (0, \rho_0). \tag{2-9}$$

By integrating (2-9) on (ρ, ρ_0), we see

$$\bar{v}(\rho) \geq -\tfrac{1}{2} M_0 \rho^{-2} + \bar{v}(\rho_0) + \tfrac{1}{2} M_0 \rho_0^{-2} \text{ for all } \rho \in (0, \rho_0).$$

This also contradicts (2-5). Thus, our claim (2-6) holds.

Step 2: We claim that there is a negative constant M satisfying

$$\lim_{\rho\to 0}\rho^3(\Delta\bar{v})'(\rho)=M. \tag{2-10}$$

Since $\bar{v}(\rho)$ satisfies the equation

$$(\rho^3(\Delta\bar{v})'(\rho))'=\rho^{-5}\overline{e^{v}} \text{ for all } \rho\in(0,1). \tag{2-11}$$

Then $f(\rho):=\rho^3(\Delta\bar{v})'(\rho)$ is an increasing function and hence $\lim_{\rho\to 0}f(\rho)=M<\infty$ exists and M maybe $-\infty$. For $\epsilon>0$ sufficiently small, by integrating (2-11) on $(\epsilon,1)$, we get

$$(\Delta\bar{v})'(1)-\epsilon^3(\Delta\bar{v})'(\epsilon)=\int_\epsilon^1\rho^{-5}\overline{e^{v}}(\rho)\,d\rho. \tag{2-12}$$

Since $\int_0^1\rho^{-5}\overline{e^{v}}(\rho)\,d\rho<\infty$, we easily see that $M>-\infty$.

We next show that $M<0$. On the contrary, we have

$$(\Delta\bar{v})'(\rho)=\left(M+\int_0^\rho t^{-5}\overline{e^{v}}(t)\,dt\right)\rho^{-3}>0 \text{ for all } \rho\in(0,1). \tag{2-13}$$

Thus $\lim_{\rho\to 0}\Delta\bar{v}(\rho)=\tilde{M}_1<\infty$ exists and $\tilde{M}_1$ maybe $-\infty$. We now consider three cases here:

(a) $\tilde{M}_1>0$,

(b) $\tilde{M}_1=0$,

(c) $\tilde{M}_1<0$.

For the case (a), we have that there exist $\rho_1>0$ and $0<M_1\le\frac{1}{2}\tilde{M}_1$ such that

$$\Delta\bar{v}(\rho)\ge M_1 \text{ for all } \rho\in(0,\rho_1).$$

Hence,

$$(\rho^3\bar{v}'(\rho))'\ge M_1\rho^3>0 \text{ for all } \rho\in(0,\rho_1)$$

and

$$\lim_{\rho\to 0}\rho^3\bar{v}'(\rho) \text{ exists.}$$

By Step 1, we find

$$\lim_{\rho\to 0}\rho^3\bar{v}'(\rho)=0.$$

On the other hand, since $\tilde{M}_1<\infty$, we see that there exist $\rho_2>0$ and $M_2\ge 2\tilde{M}_1$ such that

$$(\rho^3\bar{v}'(\rho))'\le M_2\rho^3 \text{ for all } \rho\in(0,\rho_2]. \tag{2-14}$$

Integrating (2-14) on $(0, \rho)$, we infer

$$\bar{v}'(\rho) \leq \tfrac{1}{4} M_2 \rho \text{ for all } \rho \in (0, \rho_2].$$

Thus

$$\bar{v}(\rho) \geq C > -\infty \text{ for all } \rho \in (0, \rho_2]. \tag{2-15}$$

This contradicts the fact that

$$e^C \int_0^1 \rho^{-5} \, d\rho \leq \int_0^1 \rho^{-5} e^{\bar{v}(\rho)} \, d\rho \leq \int_0^1 \rho^{-5} \overline{e^{v}}(\rho) \, d\rho < \infty.$$

For the case (b), we see that there exists $\rho_3 > 0$ such that

$$\Delta \bar{v}(\rho) \geq 0 \text{ for all } \rho \in (0, \rho_3).$$

By Step 1, we see

$$\lim_{\rho \to 0} \rho^3 \bar{v}'(\rho) = 0.$$

Similarly, there are $\rho_4 > 0$ and $M_3 > 0$ satisfying

$$(\rho^3 \bar{v}'(\rho))' \leq M_3 \rho^3 \text{ for all } \rho \in (0, \rho_4]. \tag{2-16}$$

We can also derive a contradiction from (2-16) as in the proof of the case (a).

For the case (c), we see that there exist $\rho_5 > 0$ and $-\infty < \frac{1}{2}\tilde{M}_1 < M_4 < 0$ such that

$$(\rho^3 \bar{v}'(\rho))' \leq M_4 \rho^3 < 0 \text{ for all } \rho \in (0, \rho_5]. \tag{2-17}$$

By Step 1, we get

$$\lim_{\rho \to 0} \rho^3 \bar{v}'(\rho) = 0.$$

Integrating (2-17) on $(0, \rho)$, we obtain

$$\bar{v}'(\rho) \leq \tfrac{1}{4} M_4 \rho \text{ for all } \rho \in (0, \rho_5]$$

and

$$\bar{v}(\rho_5) - \bar{v}(\rho) \leq \tfrac{1}{8} M_4 (\rho_5^2 - \rho^2) \text{ for all } \rho \in (0, \rho_5],$$

which implies

$$\bar{v}(\rho) \geq C > -\infty \text{ for all } \rho \in (0, \rho_5].$$

This is a contradiction with (2-3).

Step 3: We prove (2-4).

In view of (2-10) and (2-11), we deduce that

$$(\Delta \bar{v})'(\rho) = \left(M + \int_0^\rho s^{-5} \overline{e^{v}}(s) \, ds \right) \rho^{-3} = (M + \eta(\rho)) \rho^{-3} \text{ for } \rho \text{ near } 0, \tag{2-18}$$

where $\eta(\rho)=\int_0^\rho s^{-5}\overline{e^{v}}(s)\,ds$. Since $M<0$, we see that $\lim_{\rho\to 0}\Delta\bar v(\rho)=\gamma$ exists. As in Step 2, we infer

$$\lim_{\rho\to 0}\rho^3\bar v'(\rho)=0. \tag{2-19}$$

Integrating (2-18) on $[\rho,\rho_*]$, we obtain

$$\Delta\bar v(\rho_*)-\Delta\bar v(\rho)=-\tfrac12 M(\rho_*^{-2}-\rho^{-2})+\int_\rho^{\rho_*}\eta(s)s^{-3}\,ds \text{ for } \rho\in(0,\rho_*),$$

where $\rho_*>0$ is sufficiently small. Then

$$\Delta\bar v(\rho)=\Delta\bar v(\rho_*)+\tfrac12 M\rho_*^{-2}-\tfrac12 M\rho^{-2}-\int_\rho^{\rho_*}\eta(s)s^{-3}\,ds \text{ for } \rho\in(0,\rho_*)$$

and

$$\begin{aligned}(\rho^3\bar v'(\rho))'\\ &=\big[\Delta\bar v(\rho_*)+\tfrac12 M\rho_*^{-2}\big]\rho^3-\tfrac12 M\rho-\rho^3\int_\rho^{\rho_*}\eta(s)s^{-3}\,ds \text{ for } \rho\in(0,\rho_*).\end{aligned} \tag{2-20}$$

Integrating (2-20) on $(0,\rho]$ and using (2-19), we have

$$\begin{aligned}\bar v'(\rho)&=\tfrac14\big[\Delta\bar v(\rho_*)+\tfrac12 M\rho_*^{-2}\big]\rho-\tfrac14 M\rho^{-1}\\ &\qquad-\rho^{-3}\int_0^\rho t^3\int_t^{\rho_*}\eta(s)s^{-3}\,ds\,dt \text{ for } \rho\in(0,\rho_*).\end{aligned} \tag{2-21}$$

Integrating (2-21) on $[\rho,\rho_*]$, we deduce

$$\begin{aligned}\bar v(\rho)&=\bar v(\rho_*)-\tfrac18\big[\Delta\bar v(\rho_*)+\tfrac12 M\rho_*^{-2}\big](\rho_*^2-\rho^2)+\tfrac14 M(\ln\rho_*-\ln\rho)\\ &\qquad+\int_\rho^{\rho_*}\xi^{-3}\int_0^\xi t^3\int_t^{\rho_*}\eta(s)s^{-3}\,ds\,dt\,d\xi \text{ for } \rho\in(0,\rho_*).\end{aligned} \tag{2-22}$$

Note that

$$\int_\rho^{\rho_*}\xi^{-3}\int_0^\xi t^3\int_t^{\rho_*}\eta(s)s^{-3}\,ds\,dt\,d\xi=o_\rho(1)\ln\rho+O(1) \text{ for } \rho \text{ near } 0.$$

Then

$$\frac{\bar v(\rho)}{\ln\rho}\to\beta \quad\text{as}\quad \rho\to 0, \tag{2-23}$$

where $\beta=-\frac14 M$. Since $\int_0^1\rho^{-5}e^{\bar v(\rho)}\,d\rho<\infty$, we easily see that $\beta>4$ and this completes the proof of this lemma. □

Next, we need the following key lemma. Similar results are well-known from [Lin 1998; Wei and Xu 1999; Xu 2006] for solutions of the equation of (1-2) in $\mathbb{R}^4$ by using the fact that $-\Delta u\ge 0$ in $\mathbb{R}^4$. However, we cannot obtain such "nice" property for solutions of (1-2) in $\mathbb{R}^4\backslash B$. To do so, we will use some new arguments here, which are interesting themselves.

Lemma 2.2. *Let $u \in C^4(\mathbb{R}^4\backslash B)$ be a solution of (1-2) and (1-7) hold. Then, there is a constant C such that*

$$u(x) \le C \quad \text{for} \quad x \in \mathbb{R}^4\backslash B. \tag{2-24}$$

Moreover,

$$\Delta u(x) \to 0 \quad \text{as} \quad |x| \to \infty. \tag{2-25}$$

Proof. We divide the proof into several steps.

Step 1: We first show that

$$\overline{\lim}_{|x|\to\infty} \Delta u(x) \le 0. \tag{2-26}$$

Suppose $\overline{\lim}_{|x|\to\infty} \Delta u(x) > 0$. Then, there is a sequence $\{x_k\} \subset \mathbb{R}^4\backslash B$ with $|x_k| \to \infty$ as $k \to \infty$ and $\epsilon > 0$ independent of k, such that

$$\Delta u(x_k) \ge \epsilon > 0 \quad \text{for} \quad k \ge 1.$$

Let $w = -\Delta u$. Then

$$\Delta u + w = 0 \text{ in } \mathbb{R}^4\backslash B, \quad \Delta w + e^u = 0 \text{ in } \mathbb{R}^4\backslash B.$$

Define

$$\bar{u}_k(r) = \frac{1}{|\partial B_r(x_k)|} \int_{\partial B_r(x_k)} u(x) d\sigma, \quad 0 \le r \le \tfrac{1}{2}|x_k|.$$

Using Jensen's inequality, we have

$$\Delta \bar{u}_k + \bar{w}_k = 0 \text{ for } r \in \big[0, \tfrac{1}{2}|x_k|\big], \quad \Delta \bar{w}_k + e^{\bar{u}_k} \le 0 \text{ for } r \in \big[0, \tfrac{1}{2}|x_k|\big]. \tag{2-27}$$

Since $r^3 \bar{w}_k'(r) < 0$, we find that

$$\bar{w}_k(r) \le \bar{w}_k(0) \le -\epsilon.$$

By (2-27), we have

$$(r^3 \bar{u}_k')' \ge \epsilon r^3,$$

which implies

$$\bar{u}_k'(r) \ge \tfrac{1}{4}\epsilon r.$$

Integrating both the sides, we deduce

$$\bar{u}_k(r) \ge \bar{u}_k(0) + \tfrac{1}{8}\epsilon r^2 \text{ for all } r \in \big(0, \tfrac{1}{2}|x_k|\big].$$

Note that $\bar{u}_k(0) = u(x_k) = o(|x_k|^2)$ for k sufficiently large, we find

$$\bar{u}_k\big(\tfrac{1}{2}|x_k|\big) \ge \big(\tfrac{1}{8}\epsilon + o(1)\big)\big(\tfrac{1}{2}|x_k|\big)^2,$$

which contradicts the fact $u(x) = o(|x|^2)$ as $|x| \to \infty$.

Step 2: We show

$$\underline{\lim}_{|x|\to\infty} \Delta u(x) \geq 0.$$

On the contrary, there exist $\epsilon > 0$ and a sequence $\{x_k\} \subset \mathbb{R}^4 \setminus B$ with $|x_k| \to \infty$ as $k \to \infty$ such that

$$\Delta u(x_k) \leq -\epsilon \quad \text{for} \quad k \geq 1.$$

Setting $v_k(y) = u(x)$, $y = x - x_k$, we see that

$$\Delta_y^2 v_k = e^{v_k}, \quad \Delta_y v_k(0) = \Delta_x u(x_k) \leq -\epsilon.$$

Let

$$z_k(y) = \frac{\Delta v_k(y)}{\Delta v_k(0)}, \quad \bar{z}_k(r) = \frac{1}{|\partial B_r(0)|} \int_{\partial B_r(0)} z_k(y)\, d\sigma \text{ for } r \in \left[0, \tfrac{1}{2}|x_k|\right].$$

Then, $z_k(0) = 1$ and

$$\Delta \bar{z}_k = \frac{\overline{e^{v_k}}}{\Delta v_k(0)}.$$

Integrating on $(0, r)$ yields

$$r^3 \bar{z}_k'(r) = \frac{1}{\Delta v_k(0)|\mathbb{S}^3|} \int_{B_r(0)} e^{v_k(y)}\, dy < 0. \tag{2-28}$$

For any fixed $R > 0$, we have

$$\int_{B_R(0)} e^{v_k(y)} dy \leq \int_{B_{|x_k|/2}(x_k)} e^{u(y)}\, dy \to 0 \quad \text{as} \quad k \to \infty.$$

Therefore, it follows from (2-28) that

$$\bar{z}_k'(r) \to 0 \text{ uniformly for } r \in (0, R] \text{ as } k \to \infty,$$

which implies

$$\bar{z}_k(r) \to 1 \text{ uniformly for } r \in [0, R] \text{ as } k \to \infty. \tag{2-29}$$

On the other hand, we see that, for $r \in \left[R, \tfrac{1}{2}|x_k|\right]$,

$$-\bar{z}_k'(r) \leq \frac{r^{-3}}{|\Delta v_k(0)||\mathbb{S}^3|} \int_{B_{|x_k|/2}(0)} e^{v_k(y)} dy. \tag{2-30}$$

Integrating both sides on $[R, r]$, we find

$$0 \leq \bar{z}_k(R) - \bar{z}_k(r) \leq \frac{1}{2R^2|\Delta v_k(0)||\mathbb{S}^3|} \int_{B_{|x_k|/2}(0)} e^{v_k(y)}\, dy \to 0 \tag{2-31}$$

uniformly on $r \in \left[R, \frac{1}{2}|x_k|\right]$ as $k \to \infty$. By (2-29) and (2-31), we deduce

$$\bar{z}_k(r) \to 1 \text{ uniformly on } r \in \left[0, \tfrac{1}{2}|x_k|\right] \text{ as } k \to \infty. \tag{2-32}$$

Hence, for k sufficiently large, we have

$$\Delta \bar{v}_k(r) \le \tfrac{1}{2}\Delta v_k(0) < -\tfrac{1}{2}\epsilon \quad \text{for} \quad r \in \left[0, \tfrac{1}{2}|x_k|\right].$$

Using the similar arguments as in Step 1, we infer

$$\bar{v}_k\left(\tfrac{1}{2}|x_k|\right) - \bar{v}_k(0) < -\tfrac{1}{16}M\left(\tfrac{1}{2}|x_k|\right)^2$$

and

$$\bar{u}_k\left(\tfrac{1}{2}|x_k|\right) \le -\left(\tfrac{1}{16}M + o(1)\right)\left(\tfrac{1}{2}|x_k|\right)^2.$$

This contradicts the fact $u(x) = o(|x|^2)$ as $|x| \to \infty$.

Combining Steps 1 and 2, we can obtain

$$\lim_{|x|\to\infty} \Delta u(x) = 0. \tag{2-33}$$

Step 3: we show (2-24). On the contrary, there is a sequence $\{x_k\} \subset \mathbb{R}^4 \backslash B$ with $|x_k| \to \infty$ as $k \to \infty$ such that $u(x_k) \to \infty$ as $k \to \infty$. Setting $v_k(y) = u(x)$, $y = x - x_k$, we see from (2-33) that, for k sufficiently large,

$$\Delta_y v_k(y) \ge -\vartheta \text{ for all } y \in B_{|x_k|/2}(0),$$

where ϑ is a positive constant. Thus

$$\Delta \bar{v}_k(r) \ge -\vartheta \text{ for all } r \in \left(0, \tfrac{1}{2}|x_k|\right].$$

Then, for k sufficiently large,

$$\bar{v}_k(r) \ge \bar{v}_k(0) - \tfrac{1}{8}\vartheta r^2 \text{ for all } r \in \left(0, \tfrac{1}{2}|x_k|\right]$$

and

$$e^{\bar{v}_k(r)} \ge e^{u(x_k)}e^{-\vartheta r^2/8} \ge Me^{-\vartheta r^2/8} \text{ for all } r \in \left(0, \tfrac{1}{2}|x_k|\right], \tag{2-34}$$

for some $M > 0$ suitably large. Note that $u(x_k) \to \infty$ as $k \to \infty$. From (2-34), we have

$$\int_{B_{|x_k|/2}(x_k)} e^{u(y)}\,dy \ge M|\mathbb{S}^3| \int_0^2 r^3 e^{-\vartheta r^2/8}\,dr > 0,$$

which is a contradiction with

$$\int_{B_{|x_k|/2}(x_k)} e^{u(y)}\,dy \to 0 \quad \text{as} \quad k \to \infty. \qquad \square$$

3. Proof of the main results

In this section, we present the proof of Theorems 1.1 and 1.2. The proof of Theorem 1.1 is simple by using the result in [Chou and Wan 1994]. We mainly concentrate our attention to the proof of Theorem 1.2.

Proof of Theorem 1.1. Let $u \in C^2(\mathbb{R}^2\backslash B)$ be a solution to (1-1). Using the transformation

$$v(y) = u(x), \quad y = \frac{x}{|x|^2},$$

we see that $v \in C^2(B\backslash\{0\})$ satisfies the problem

$$-\Delta_y v = |y|^{-4} e^v \text{ in } B\backslash\{0\}, \quad \int_{B\backslash\{0\}} |y|^{-4} e^{v(y)}\, dy < \infty. \tag{3-1}$$

It is clear that 0 is a nonremovable singular point of v. To obtain the asymptotic behavior of $u(x)$ as $|x| \to \infty$, we only need to obtain the asymptotic behavior of $v(y)$ as $|y| \to 0$.

Let $w(y) = v(y) - 4\ln|y|$. We find that $w(y)$ satisfies the problem

$$-\Delta w = e^w \text{ in } B\backslash\{0\}, \quad \int_{B\backslash\{0\}} e^{w(y)}\, dy < \infty. \tag{3-2}$$

It follows from [Chou and Wan 1994, Theorem 5] that

$$\frac{w(y)}{\ln|y|} \to \beta_0 \quad \text{as} \quad |y| \to 0, \tag{3-3}$$

where $\beta_0 > -2$, which implies

$$\frac{v(y)}{\ln|y|} \to \beta \quad \text{as} \quad |y| \to 0, \tag{3-4}$$

where $\beta = \beta_0 + 4 > 2$. Therefore, (1-6) can be obtained from (3-4) and the proof of Theorem 1.1 is complete. □

Proof of Theorem 1.2. Define

$$w(x) = \frac{1}{4|\mathbb{S}^3|} \int_{\mathbb{R}^4\backslash B} \ln\Big(\frac{|x-y|}{|y|}\Big) e^{u(y)}\, dy \tag{3-5}$$

and

$$\bar{w}(r) = \frac{1}{|\partial B_r(0)|} \int_{\partial B_r(0)} w(y)\, d\sigma, \quad r > 1. \tag{3-6}$$

Then, we have

$$\Delta^2 w(x) = -e^{u(x)} \text{ and } \Delta^2(u+w)(x) = 0 \text{ for } x \in \mathbb{R}^4\backslash B.$$

Note that u is upper bounded, as in [Lin 1998], we can deduce

$$\Delta w = \frac{1}{2|\mathbb{S}^3|}\int_{\mathbb{R}^4\backslash B}\frac{1}{|x-y|^2}e^{u(y)}\,dy \tag{3-7}$$

and

$$\Delta w(x)\to 0 \quad \text{as} \quad |x|\to\infty.$$

Set $\psi = u+w$. Then $\Delta^2\psi = 0$ in $\mathbb{R}^4\backslash B$. Let $k(t,\theta)=\psi(r,\theta)$, $\bar{k}(t)=\bar{\psi}(r)$, $t=\ln r$, $r=|x|$, $r>1$. Then $\Delta^2\bar{\psi}(r)=0$, $r>1$. By Lemmas 2.1, A.1, we know

$$\frac{\bar{k}(t)}{t}\to\alpha_0-\beta \quad \text{as} \quad t\to\infty,$$

where

$$\alpha_0=\frac{1}{4|\mathbb{S}^3|}\int_{\mathbb{R}^4\backslash B}e^{u(y)}\,dy.$$

Define

$$z(t,\theta)=k(t,\theta)-\bar{k}(t).$$

Then

$$z_t^{(4)}-4z_{tt}+2\Delta_\theta z_{tt}+\Delta_\theta^2 z=0, \quad (t,\theta)\in(0,\infty)\times\mathbb{S}^3. \tag{3-8}$$

Let

$$z(t,\theta)=\sum_{i=1}^{\infty}\sum_{j=1}^{m_i}z_i^j(t)Q_i^j(\theta),$$

where $Q_i^j(\theta)$ is an eigenfunction corresponding to the eigenvalue σ_i of the problem

$$\Delta_\theta^2 Q=\sigma Q, \quad \theta\in\mathbb{S}^3.$$

It is known from [Guo et al. 2015] that $\sigma_i=\lambda_i^2$, $\lambda_i=i(2+i)$, $m_i=(1+i)^2$ is the multiplicity of σ_i and Q_i^j also satisfies

$$-\Delta_\theta Q_i^j=\lambda_i Q_i^j, \quad \theta\in\mathbb{S}^3.$$

Hence for each (i,j) with $i=1,2,\ldots$, $j=1,2,\ldots,m_i$,

$$(z_i^j)^{(4)}(t)-2(\lambda_i+2)(z_i^j)_{tt}(t)+\lambda_i^2(z_i^j)(t)=0. \tag{3-9}$$

The characteristic equation of (3-9) is

$$\tau^{(4)}-2(2+\lambda_i)\tau^2+\lambda_i^2=0$$

and the corresponding characteristic roots are given by

$$\tau_1^{(i)}=\sqrt{2+\lambda_i+2\sqrt{1+\lambda_i}}, \quad \tau_2^{(i)}=-\sqrt{2+\lambda_i+2\sqrt{1+\lambda_i}},$$
$$\tau_3^{(i)}=\sqrt{2+\lambda_i-2\sqrt{1+\lambda_i}}, \quad \tau_4^{(i)}=-\sqrt{2+\lambda_i-2\sqrt{1+\lambda_i}}.$$

Moreover,

$$\tau_2^{(i)} < \tau_4^{(i)} < 0 < \tau_3^{(i)} < \tau_1^{(i)}.$$

By the standard ODE theory, we see that there is $T \gg 1$ such that for $t > T$

$$z_i^j(t) = B_1 e^{\tau_1^{(i)} t} + B_2 e^{\tau_2^{(i)} t} + B_3 e^{\tau_3^{(i)} t} + B_4 e^{\tau_4^{(i)} t}.$$

Since $|z(t, \theta)| \le Ct$, we have $B_1 = B_3 = 0$. Thus

$$z_i^j(t) = B_2 e^{\tau_2^{(i)} t} + B_4 e^{\tau_4^{(i)} t}$$

and

$$B_2 = O(T) e^{-\tau_2^{(i)} T}, \quad B_4 = O(T) e^{-\tau_4^{(i)} T}.$$

Thus

$$z_i^j(t) = O(T) e^{\tau_2^{(i)}(t-T)} + O(T) e^{\tau_4^{(i)}(t-T)}.$$

Let $Z^2(t) = \sum_{i=1}^{\infty} \sum_{j=1}^{m_i} [z_i^j(t)]^2$. Note that $\tau_2^{(i)} < \tau_4^{(i)} < 0$, then

$$Z^2(t) \le CT \sum_{i=1}^{\infty} m_i \big(e^{2\tau_2^{(i)}(t-T)} + e^{2\tau_4^{(i)}(t-T)}\big) \le CT \sum_{i=1}^{\infty} m_i e^{2\tau_4^{(i)}(t-T)} \le CT e^{2\tau_4^{(1)}(t-T)},$$

where C is a positive constant independent of t. Here we have used the fact that for $t > T_* := 10T$,

$$\lim_{i\to\infty} \frac{m_{i+1}}{m_i} e^{2(\tau_4^{(i+1)} - \tau_4^{(i)})(t-T)} \le e^{-2(t-T)} \le \tfrac{1}{2}.$$

Note that $\|Q_i^j\|_{L^2(\mathbb{S}^3)} = 1$ for each (i, j). Hence

$$\|z\|_{L^2(\mathbb{S}^3)} \le C e^{\tau_4^{(1)}(t-T)} \le C e^{\tau_4^{(1)} t}.$$

By the interior L^∞-estimate of (3-8) in $(t-1, t+1) \times \mathbb{S}^3$, we obtain

$$|z(t, \theta)| \le C \|z\|_{L^2((t-1,t+1)\times\mathbb{S}^3)} \le C e^{\tau_4^{(1)} t}, \tag{3-10}$$

which implies

$$\max_{\theta\in\mathbb{S}^3} |z(t, \theta)| \le C e^{\tau_4^{(1)} t} \quad \text{for} \quad t \in (T_*, \infty).$$

Since $\tau_4^{(1)} = -1$, we find

$$u(x) = -w(x) + \bar{\psi}(|x|) + O(|x|^{-1}).$$

By Lemma A.1, we see

$$\frac{w(x)}{\ln|x|} \to \alpha_0 \quad \text{as} \quad |x| \to \infty.$$

Therefore

$$\lim_{|x|\to\infty}\frac{u(x)}{\ln|x|}=-\beta,\quad \beta>4. \tag{3-11}$$

Next we show

$$-|x|^2\Delta u(x)\to\frac{1}{2|\mathbb{S}^3|}\int_{\mathbb{R}^4\setminus B}e^{u(y)}\,dy+\kappa\quad\text{as}\quad|x|\to\infty, \tag{3-12}$$

where κ is a constant.

Thanks to (3-7), (3-11), we deduce

$$|x|^2\Delta w(x)\to\frac{1}{2|\mathbb{S}^3|}\int_{\mathbb{R}^4\setminus B}e^{u(y)}\,dy\quad\text{as}\quad|x|\to\infty. \tag{3-13}$$

Let $h(x)=\Delta(u+w)(x)$,

$$\bar{h}(r)=\frac{1}{|\partial B_r(0)|}\int_{\partial B_r(0)}h(y)\,d\sigma.$$

Since $\lim_{|x|\to\infty}\Delta(u+w)(x)=0$. Then, we see that $\lim_{r\to\infty}\bar{h}(r)=0$,

$$\Delta\bar{h}(r)=0 \text{ for all } r\in(1,\infty).$$

Then, there is a constant c such that

$$\bar{h}'(r)\equiv cr^{-3}\text{ for all } r\in(1,\infty). \tag{3-14}$$

By integrating (3-14) in (r,∞), we obtain

$$\bar{h}(r)=-\kappa r^{-2},\quad\text{where}\quad\kappa=\tfrac{1}{2}c. \tag{3-15}$$

To obtain (3-12), we only need to show

$$|x|^2(h(x)-\bar{h}(|x|))\to 0\quad\text{as}\quad|x|\to\infty. \tag{3-16}$$

Let

$$z(t,\theta)=h(r,\theta)-\bar{h}(r),\quad t=\ln r.$$

Then $z(t,\theta)$ satisfies

$$z_{tt}(t,\theta)+2z_t(t,\theta)+\Delta_\theta z(t,\theta)=0,\quad(t,\theta)\in(0,\infty)\times\mathbb{S}^3. \tag{3-17}$$

Set

$$z(t,\theta)=\sum_{i=1}^{\infty}\sum_{j=1}^{m_i}z_i^j(t)Q_i^j(\theta),$$

where $Q_i^j(\theta)$ is an eigenfunction corresponding to the eigenvalue λ_i of the problem

$$-\Delta_\theta Q=\lambda Q,\quad\theta\in\mathbb{S}^3.$$

It is well-known that $\lambda_i = i(2+i)$ and $m_i = (1+i)^2$ is the multiplicity of λ_i. Then $z_i^j(t)$ satisfies

$$(z_i^j)''(t) + 2(z_i^j)'(t) - \lambda_i z_i^j(t) = 0, \quad t \in (0, \infty). \tag{3-18}$$

The characteristic equation of (3-18) is

$$\tau^2 + 2\tau - \lambda_i = 0,$$

whose characteristic roots are given by

$$\tau_1^{(i)} = -1 - \sqrt{1+\lambda_i} < 0 \quad \text{and} \quad \tau_2^{(i)} = -1 + \sqrt{1+\lambda_i} > 0, \quad i = 1, 2, \ldots$$

Therefore, for $T \gg 1$ and $t > T$, we see that

$$z_i^j(t) = Ae^{\tau_1^{(i)} t} + Be^{\tau_2^{(i)} t}, \text{ where } A, B \text{ are generic constants.}$$

Using the fact that $h(x)$ is bounded and hence $z(t, \theta)$ is bounded, we deduce that $B = 0$ and

$$z_i^j(t) = Ae^{\tau_1^{(i)} t} \quad \text{with} \quad A = O(1)e^{-\tau_1^{(i)} T}.$$

Hence, for $j = 1, 2, \ldots, m_i$, we have

$$z_i^j(t) = O(1)e^{\tau_1^{(i)}(t-T)} \text{ for all } t > T.$$

Since

$$\lim_{i\to\infty} \frac{m_{i+1}}{m_i} e^{2(\tau_1^{(i+1)} - \tau_1^{(i)})(t-T)} \le e^{-2(t-T)} < \tfrac{1}{2}.$$

Thus, for $t > 10T$, we obtain

$$\|z\|^2_{L^2(\mathbb{S}^3)} \le C \sum_{i=1}^{\infty} m_i e^{2\tau_1^{(i)}(t-T)} \le Ce^{2\tau_1^{(1)} t}, \tag{3-19}$$

where $C > 0$ is independent of t.

For any fixed $(t, \theta) \in (T_* + 1, \infty) \times \mathbb{S}^3$, by the interior L^∞-estimate of (3-17) in $(t-1, t+1) \times \mathbb{S}^3$, we obtain from (3-19) that

$$|z(t, \theta)| \le C\|z\|_{L^2((t-1,t+1)\times\mathbb{S}^3)} \le Ce^{\tau_1^{(1)} t}, \tag{3-20}$$

where $C > 0$ is independent of t. Thus

$$\max_{\theta\in\mathbb{S}^3} |z(t, \theta)| \le Ce^{-3t} \text{ for } t \in (T_*, \infty). \tag{3-21}$$

Therefore, for $|x| \ge 2e^{T_*}$, we have

$$|x|^2 |h(x) - \bar{h}(|x|)| \le C|x|^{-1},$$

where $C > 0$ is independent of $|x|$. Then (3-12) holds and the proof of the theorem is complete. □

Now we are in the position to give the proof of Theorem 1.4.

Proof of Theorem 1.4. Let $w(y) = v(y) + 8\ln|y|$. Then $w(y)$ satisfies the problem

$$\Delta^2 w = |y|^{-8} e^w \text{ in } B\backslash\{0\}, \qquad \int_{B\backslash\{0\}} |y|^{-8} e^{w(y)}\, dy < \infty.$$

It is known from the proof of Theorem 1.2 and Remark 1.3 that

$$\frac{w(y)}{\ln|y|} \to \beta \quad \text{as} \quad |y| \to 0$$

with $\beta > 4$. Then we have

$$\frac{v(y)}{\ln|y|} \to \gamma \quad \text{as} \quad |y| \to 0$$

with $\gamma = \beta - 8 > -4$. This completes the proof of the theorem. □

Appendix: Some estimates

In this section, we shall present some estimates used in the proof of Theorem 1.2. The proof is similar to that in [Lin 1998]. For the reader's convenience, we give the proof.

Let u be a solution of (1-2). Define

$$\alpha_0 = \frac{1}{4|\mathbb{S}^3|} \int_{\mathbb{R}^4\backslash B} e^{u(y)}\, dy, \tag{A-1}$$

$$w(x) = \frac{1}{4|\mathbb{S}^3|} \int_{\mathbb{R}^4\backslash B} \ln\Big(\frac{|x-y|}{|y|}\Big) e^{u(y)}\, dy, \tag{A-2}$$

$$\bar{w}(r) = \frac{1}{|\partial B_r(0)|} \int_{\partial B_r(0)} w(y)\, d\sigma \text{ for } r > 1. \tag{A-3}$$

Lemma A.1. *Let u be a solution of* (1-2) *and $\bar{v}(r)$ is defined in* (A-3). *Then*

$$\frac{w(x)}{\ln|x|} \to \alpha_0 \quad as \quad |x| \to \infty, \tag{A-4}$$

$$\frac{\bar{w}(r)}{\ln r} \to \alpha_0 \quad as \quad r \to \infty. \tag{A-5}$$

Proof. We only need to show (A-4). We first show that

$$w(x) \le \alpha_0 \ln|x| + C, \text{ where } C \text{ is a positive constant.} \tag{A-6}$$

For $|x| \ge 4$, we split $\mathbb{R}^4 \setminus B = \Omega_1 \cup \Omega_2$, where $\Omega_1 = \big\{y \in \mathbb{R}^4\backslash B : |y-x| \le \frac{1}{2}|x|\big\}$, $\Omega_2 = \big\{y \in \mathbb{R}^4\backslash B : |y-x| > \frac{1}{2}|x|\big\}$. For $y \in \Omega_1$, we have

$$|y| \ge |x| - |x-y| \ge \tfrac{1}{2}|x| \ge |x-y|.$$

Then $\ln(|x-y|/|y|) \le 0$. Note that $|x-y| \le |x|+|y| \le |x||y|$ for $|x|, |y| \ge 2$. Since $\frac{3}{2} \ge |x-y|/|x| \ge \frac{1}{2}$ for $|x| \ge 4$, $|y| \le 2$. Thus, we find

$$\ln|x-y| \le \ln|x| + C.$$

Then

$$\begin{aligned} w(x) &\le \frac{1}{4|\mathbb{S}^3|}\int_{\Omega_2} \ln\Big(\frac{|x-y|}{|y|}\Big) e^{u(y)}\,dy \\ &\le \frac{1}{4|\mathbb{S}^3|}\ln|x| \int_{\Omega_2} e^{u(y)}\,dy + \frac{1}{4|\mathbb{S}^3|}\int_{\Omega_2\cap\{|y|\le 2\}} (C-\ln|y|)e^{u(y)}\,dy \\ &\le \alpha_0 \ln|x| + C. \end{aligned}$$

Next, we claim that for any $\varepsilon > 0$, there exists $R = R(\varepsilon) > 0$ such that for $|x| > R$,

$$w(x) \ge \big(\alpha_0 - \tfrac{1}{2}\varepsilon\big)\ln|x| + \frac{1}{4|\mathbb{S}^3|}\int_{B_1(x)} \ln(|x-y|)e^{u(y)}\,dy. \tag{A-7}$$

We decompose $\mathbb{R}^4\backslash B$ into A_1, A_2 and A_3, where $A_1 = \{y : 1 < |y| \le R_0\}$, $A_2 = \big\{y : |y-x| \le \frac{1}{2}|x|, |y| \ge R_0\big\}$, $A_3 = \{y : |y-x| > \frac{1}{2}|x|, |y| \ge R_0\}$. For any $\varepsilon > 0$, choosing R_0 large, and taking $|x|$ sufficiently large, we have

$$\begin{aligned} &\frac{1}{4|\mathbb{S}^3|}\int_{A_1} \ln\Big(\frac{|x-y|}{|y|}\Big)e^{u(y)}\,dy - \frac{1}{4|\mathbb{S}^3|}\ln|x|\int_{\mathbb{R}^4\backslash B} e^{u(y)}\,dy \\ &\quad\ge \frac{1}{4|\mathbb{S}^3|}\ln|x|\int_{A_1}\ln\Big(\frac{|x-y|}{|x||y|}\Big)e^{u(y)}\,dy - \frac{1}{4|\mathbb{S}^3|}\ln|x|\int_{A_1^c} e^{u(y)}\,dy \ge -\tfrac{1}{4}\varepsilon\ln|x|. \end{aligned}$$

Then

$$\frac{1}{4|\mathbb{S}^3|}\int_{A_1}\ln\Big(\frac{|x-y|}{|y|}\Big)e^{u(y)}\,dy \ge \big(\alpha_0 - \tfrac{1}{4}\varepsilon\big)\ln|x|.$$

Since $|y| < 2|x|$ in A_2, we find

$$\begin{aligned} &\frac{1}{4|\mathbb{S}^3|}\int_{A_2}\ln\Big(\frac{|x-y|}{|y|}\Big)e^{u(y)}\,dy \\ &\qquad\ge \frac{1}{4|\mathbb{S}^3|}\int_{B_1(x)}\ln(|x-y|)e^{u(y)}\,dy - \frac{1}{4|\mathbb{S}^3|}\ln(2|x|)\int_{A_2} e^{u(y)}\,dy. \end{aligned}$$

For A_3, if $|y| \le 2|x|$, $|x-y| \ge \frac{1}{2}|x| \ge \frac{1}{4}|y|$. If $|y| \ge 2|x|$, $|x-y| \ge |y|-|x| \ge \frac{1}{2}|y|$. Then $|x-y|/|y| \ge \frac{1}{4}$. Hence

$$\frac{1}{4|\mathbb{S}^3|}\int_{A_3}\ln\Big(\frac{|x-y|}{|y|}\Big)e^{u(y)}\,dy \ge -\frac{1}{4|\mathbb{S}^3|}\ln 4\int_{A_3} e^{u(y)}\,dy.$$

Therefore, our claim follows.

Let δ_0 small, R_0 sufficiently large such that

$$\int_{B_4(x)} e^{u(y)}\,dy \le \delta_0, \quad |x| \ge R_0. \tag{A-8}$$

Set h be the solution of

$$\Delta^2 h = e^u \text{ in } B_4(x), \quad h = \Delta h = 0 \text{ on } \partial B_4(x).$$

By [Lin 1998, Lemma 2.3], we find, for small $\delta_0 > 0$,

$$\int_{B_4(x)} e^{2|h(y)|}\, dy \le \sigma, \tag{A-9}$$

where $\sigma > 0$ is independent of x.

Let $\varphi = u - h$. Then

$$\Delta^2 \varphi = 0 \text{ in } B_4(x), \quad \Delta\varphi = \Delta u, \varphi = u \text{ on } \partial B_4(x).$$

Setting $\phi(y) = -\Delta\varphi(y)$, then

$$\Delta\phi = 0 \text{ in } B_4(x), \quad \phi = -\Delta u \text{ on } \partial B_4(x).$$

By Lemma 2.2, we see that Δu is bounded. Thus

$$|\phi(y)| \le c_0, \quad y \in \overline{B_2(x)}.$$

Note that

$$\Delta\varphi = -\phi \text{ in } B_4(x), \quad \varphi = u \text{ on } \partial B_4(x).$$

By the elliptic estimates, we have

$$\sup_{B_1(x)} \varphi \le C(\|\varphi^+\|_{L^1(B_2(x))} + \|\phi\|_{L^q(B_2(x))}), \quad q > 2.$$

Since $\varphi = u - h$, then $\varphi^+ \le u^+ + |h|$. Thus

$$\int_{B_2(x)} \varphi^+\, dy \le C \int_{B_2(x)} e^{\varphi^+/2}\, dy \le C\left(\int_{B_2(x)} e^{u^+(y)}\, dy\right)^{1/2} \left(\int_{B_2(x)} e^{|h(y)|}\, dy\right)^{1/2}.$$

Note $e^{u^+} \le 1 + e^u$, we find that $\sup_{B_1(x)} \varphi(x) \le C$, $u \le C + |h(y)|$, $y \in B_1(x)$. Then

$$\int_{B_1(x)} e^{2u(y)}\, dy \le C \int_{B_1(x)} e^{2|h(y)|}\, dy \le C.$$

Thus

$$\left|\int_{B_1(x)} \ln(|x-y|) e^{u(y)}\, dy\right| \le \left(\int_{B_1(x)} \ln^2|x-y|\, dy\right)^{1/2} \left(\int_{B_1(x)} e^{2u(y)}\, dy\right)^{1/2} \le C.$$

By (A-7), we deduce, for $|x|$ large enough,

$$w(x) \ge (\alpha_0 - \varepsilon) \ln|x|. \tag{A-10}$$

In view of (A-6),(A-10), we can obtain (A-4). □

Acknowledgements

The authors are very grateful to the referee for many valuable suggestions and comments, which greatly improved this paper.

References

[Chang and Yang 1995] S.-Y. A. Chang and P. C. Yang, "Extremal metrics of zeta function determinants on 4-manifolds", *Ann. of Math.* (2) **142**:1 (1995), 171–212. MR Zbl

[Chang and Yang 1997] S.-Y. A. Chang and P. C. Yang, "On uniqueness of solutions of n-th order differential equations in conformal geometry", *Math. Res. Lett.* **4**:1 (1997), 91–102. MR Zbl

[Chen and Li 1991] W. X. Chen and C. Li, "Classification of solutions of some nonlinear elliptic equations", *Duke Math. J.* **63**:3 (1991), 615–622. MR Zbl

[Chou and Wan 1994] K. S. Chou and T. Y.-H. Wan, "Asymptotic radial symmetry for solutions of $\Delta u + e^u = 0$ in a punctured disc", *Pacific J. Math.* **163**:2 (1994), 269–276. MR Zbl

[Graham et al. 1992] C. R. Graham, R. Jenne, L. J. Mason, and G. A. J. Sparling, "Conformally invariant powers of the Laplacian, I: Existence", *J. Lond. Math. Soc.* (2) **46**:3 (1992), 557–565. MR Zbl

[Guo et al. 2015] Z. Guo, X. Huang, and F. Zhou, "Radial symmetry of entire solutions of a biharmonic equation with exponential nonlinearity", *J. Funct. Anal.* **268**:7 (2015), 1972–2004. MR Zbl

[Lin 1998] C.-S. Lin, "A classification of solutions of a conformally invariant fourth order equation in $\mathbb{R}^n$", *Comment. Math. Helv.* **73**:2 (1998), 206–231. MR Zbl

[Martinazzi 2009] L. Martinazzi, "Classification of solutions to the higher order Liouville's equation on $\mathbb{R}^{2m}$", *Math. Z.* **263**:2 (2009), 307–329. MR Zbl

[Wei and Xu 1999] J. Wei and X. Xu, "Classification of solutions of higher order conformally invariant equations", *Math. Ann.* **313**:2 (1999), 207–228. MR Zbl

[Wei and Ye 2008] J. Wei and D. Ye, "Nonradial solutions for a conformally invariant fourth order equation in $\mathbb{R}^4$", *Calc. Var. Partial Differential Equations* **32**:3 (2008), 373–386. MR Zbl

[Xu 2006] X. Xu, "Classification of solutions of certain fourth-order nonlinear elliptic equations in $\mathbb{R}^4$", *Pacific J. Math.* **225**:2 (2006), 361–378. MR

Received April 30, 2020. Revised August 26, 2020.

ZONGMING GUO
gzm@htu.cn
DEPARTMENT OF MATHEMATICS
HENAN NORMAL UNIVERSITY
XINXIANG
CHINA

ZHONGYUAN LIU
liuzy@henu.edu.cn
SCHOOL OF MATHEMATICS AND STATISTICS
HENAN UNIVERSITY
KAIFENG
CHINA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 2, 2020
https://doi.org/10.2140/pjm.2020.309.353

ULRICH ELEMENTS IN NORMAL SIMPLICIAL AFFINE SEMIGROUPS

JÜRGEN HERZOG, RAHELEH JAFARI AND DUMITRU I. STAMATE

Let $H \subseteq \mathbb{N}^d$ be a normal affine semigroup, $R = K[H]$ its semigroup ring over the field K and ω_R its canonical module. The Ulrich elements for H are those h in H such that for the multiplication map by x^h from R into ω_R, the cokernel is an Ulrich module. We say that the ring R is almost Gorenstein if Ulrich elements exist in H. For the class of slim semigroups that we introduce, we provide an algebraic criterion for testing the Ulrich property. When $d = 2$, all normal affine semigroups are slim. Here we have a simpler combinatorial description of the Ulrich property. We improve this result for testing the elements in H which are closest to zero. In particular, we give a simple arithmetic criterion for when is (1, 1) an Ulrich element in H.

Introduction

Let H be an affine semigroup in $\mathbb{N}^d$ and $K[H]$ its semigroup ring over the field K. In this paper we investigate the almost Gorenstein property for $K[H]$ taking into account the natural multigraded structure of this ring, under the assumption that H is normal and simplicial.

The almost Gorenstein property appeared in [Barucci and Fröberg 1997] in the context of 1-dimensional analytical unramified rings. It was extended to 1-dimensional local rings by Goto, Matsuoka and Thi Phuong [Goto et al. 2013], and later on to rings of higher dimension by Goto, Takahashi and Taniguchi [Goto et al. 2015]. Let R be a positively graded Cohen–Macaulay K-algebra with canonical module ω_R. We let $a = -\min\{k \in \mathbb{Z} : (\omega_R)_k \neq 0\}$, which is also known as the a-invariant of R. In [Goto et al. 2015], R is called (graded) almost Gorenstein (AG for short) if there exists an exact sequence of graded R-modules

$$0 \to R \to \omega_R(-a) \to E \to 0, \tag{1}$$

where E is an Ulrich module, i.e., E is a Cohen–Macaulay graded module which is minimally generated by $e(E)$ elements. Here $e(E)$ denotes the multiplicity of E with respect to the graded maximal ideal in R.

MSC2020: primary 05E40, 13H10; secondary 13H15, 20M25.
Keywords: almost Gorenstein ring, Ulrich element, affine semigroup ring, lattice points.

Let $H \subseteq \mathbb{N}^d$ be an affine semigroup whose associated group is $\mathrm{gr}(H) = \mathbb{Z}^d$. We denote C the cone over H. Assume H is normal, i.e., $H = C \cap \mathbb{Z}^d$, equivalently, the ring R is normal. Then R is a Cohen–Macaulay ring [Hochster 1972] and a K-basis for the canonical module ω_R is given by the monomials with exponents in the relative interior of the cone C [Danilov 1978; Stanley 1978], i.e., in the set $\omega_H = \mathbb{Z}^d \cap \mathrm{relint}\, C$. In the multigraded setting that we want to consider here, there does not seem to be any distinguished element in ω_H to replace the a-invariant in the short exact sequence (1). In this sense, we propose the following.

Definition 3.1. For $\boldsymbol{b} \in \omega_H$ consider the exact sequence

$$0 \to R \to \omega_R(\boldsymbol{b}) \to E \to 0, \tag{2}$$

where $1 \in R$ is mapped to $u = \boldsymbol{x}^{\boldsymbol{b}}$ and $E = \omega_R / uR$. Then $\boldsymbol{b}$ is called an *Ulrich element* in H, if E is an Ulrich R-module.

If H admits an Ulrich element $\boldsymbol{b}$, then we call the ring $R = K[H]$ *almost Gorenstein with respect to* $\boldsymbol{b}$, or simply AG if H has an Ulrich element.

The Gorenstein property has attracted a lot of interest due to its multifaceted algebraic and homological descriptions. For rings with a combinatorial structure behind, there are often nice characterizations of the Gorenstein property. Scratching only at the surface, we mention that Gorenstein toric rings were characterized by Hibi [1992], and for special subclasses the results are more precise; see [De Negri and Hibi 1997; Hibi 1987; Hibi et al. 2019; Dinu 2020].

The almost Gorenstein property was characterized for determinantal rings in [Taniguchi 2018], numerical semigroup rings in [Nari 2013] and Hibi rings in [Miyazaki 2018].

In this paper we investigate the Ulrich elements in H under the assumption that the normal affine semigroup $H \subset \mathbb{N}^d$ is also simplicial, i.e., the cone C over H has $d = \dim_{\mathbb{R}} \mathrm{aff}(H)$ extremal rays. That will be assumed for the rest of this introduction, too.

Next we outline the main results. We denote $\boldsymbol{a}_1, \dots, \boldsymbol{a}_d$ the primitive integer vectors in H situated on each extremal ray of the cone C, respectively, and we call them the extremal rays of H. They are part of the Hilbert basis of H, denoted B_H, which is the unique minimal generating set of H.

When H is normal and simplicial it is known that the monomials $\boldsymbol{x}^{\boldsymbol{a}_1}, \dots, \boldsymbol{x}^{\boldsymbol{a}_d}$ form a maximal regular sequence on R. A first result that we prove in Proposition 2.2 is that for any $\boldsymbol{b} \neq 0$ in H the sequence $\boldsymbol{x}^{\boldsymbol{b}}, \boldsymbol{x}^{\boldsymbol{a}_1} - \boldsymbol{x}^{\boldsymbol{a}_2}, \dots, \boldsymbol{x}^{\boldsymbol{a}_1} - \boldsymbol{x}^{\boldsymbol{a}_d}$ is regular, as well. Let $J = (\boldsymbol{x}^{\boldsymbol{a}_i} - \boldsymbol{x}^{\boldsymbol{a}_j} : 1 \leq i, j \leq d)R$. The next technical result is vital in our study of Ulrich elements. Namely, in Theorem 2.4, we show that J is a reduction ideal of $\mathfrak{m}$ modulo the ideal $\boldsymbol{x}^{\boldsymbol{b}} R$, if and only if for any $\boldsymbol{c} \in B_H \setminus \omega_H$ the sum of the coordinates of $\boldsymbol{c}$ with respect to the basis $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_d$ is at least one. The normal and simplicial semigroups with that property are to be called *slim*.

In this notation, we provide the following characterization.

Theorem 3.2. *Let H be a slim semigroup and $\boldsymbol{b} \in \omega_H$. Then $\boldsymbol{b}$ is an Ulrich element in H if and only if $\mathfrak{m}\omega_R \subseteq (\boldsymbol{x}^{\boldsymbol{b}} R, J\omega_R)$.*

This result allows to produce first examples of semigroups with Ulrich elements; see Examples 3.5, 3.4.

In the next sections we focus on making more explicit the AG property in dimension two. Let H be any normal affine semigroup $H \subseteq \mathbb{N}^2$. It is automatically slim since it is simplicial and $B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2\} \subseteq \omega_H$ (see Lemma 1.1). We denote by $\boldsymbol{a}_1, \boldsymbol{a}_2$ its extremal rays. In Theorem 4.3 we prove that any element $\boldsymbol{b} \in B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2\}$ is an Ulrich element in H if and only if for all $\boldsymbol{c}_1, \boldsymbol{c}_2$ in B_H one has

$$\boldsymbol{c}_1 + \boldsymbol{c}_2 \in (\boldsymbol{a}_1 + H) \cup (\boldsymbol{a}_2 + H) \cup (\boldsymbol{b} + H).$$

Equivalently, if for all $\boldsymbol{c}_1, \boldsymbol{c}_2 \in B_H$ so that $\boldsymbol{c}_1, \boldsymbol{c}_2 \in P_H$ it follows that $\boldsymbol{c}_1 + \boldsymbol{c}_2 \in \boldsymbol{b} + H$. Here $P_H = \{\lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 : 0 \leq \lambda_1, \lambda_2 < 1\}$ is the fundamental parallelogram of H.

Based on this result, in Section 4 we find examples with zero, one, or several Ulrich elements in B_H.

We prove in Lemma 5.1 that for any $H \subseteq \mathbb{N}^2$ as above, the semigroup ideal ω_H has a unique minimal element with respect to the componentwise partial order on $\mathbb{N}^2$. We call it the bottom element of H. This definition naturally extends to higher embedding dimension, but when $d > 2$ not all normal semigroups in $\mathbb{N}^d$ have a bottom element.

However, bottom elements, when available, are good candidates to check against the Ulrich property. We prove that when $H \subseteq \mathbb{N}^d$ is a slim semigroup such that

- (Proposition 3.6) the nonzero elements in H have all the entries positive and $\boldsymbol{b} = (1, 1, \ldots, 1) \in \omega_H$, or
- (Proposition 5.5) $d = 2$ and $\boldsymbol{b}$ the bottom element in H satisfies $2\boldsymbol{b} \in P_H$,

then $\boldsymbol{b}$ is the only possible Ulrich element in H.

These results motivate us to find more direct criteria for testing the Ulrich property of the bottom element. Our attempt is successful when $d = 2$.

In the following, H is a normal affine semigroup in $\mathbb{N}^2$ with the extremal rays $\boldsymbol{a}_1 = (x_1, y_1)$ and $\boldsymbol{a}_2 = (x_2, y_2)$ with $\boldsymbol{a}_1$ closer to the x-axis than $\boldsymbol{a}_2$. Considering $\boldsymbol{b} = (u, v)$ the bottom element of H, for $i = 1, 2$ we define H_i to be the normal semigroup with the extremal rays $\boldsymbol{b}$ and $\boldsymbol{a}_i$. We denote $H_i^* = \operatorname{relint} P_{H_i} \cap \mathbb{Z}^2$ for $i = 1, 2$. We show:

Lemma 5.12. *For $\boldsymbol{b}$ the bottom element in H the following are equivalent*:

(a) *$\boldsymbol{b}$ is an Ulrich element in H.*

(b) *For $i = 1, 2$, if $\boldsymbol{p}, \boldsymbol{q} \in H_i^*$ then $\boldsymbol{p} + \boldsymbol{q} \notin H_i^*$.*

We shall say that H is AG1 if point (b) above is satisfied for $i = 1$ and we call it AG2 if it holds for $i = 2$.

Thus the bottom element is an Ulrich element in H if and only if H is AG1 and AG2. This calls for a better understanding of the points in H_1^* and H_2^*. Lemma 5.16 shows that H_i^* has $|vx_i - uy_i| - 1$ elements, for $i = 1, 2$. An immediate consequence of independent interest is the following Gorenstein criterion.

Corollary 5.17. *With notation as above, the ring $K[H]$ is Gorenstein if and only if $vx_1 - uy_1 = uy_2 - vx_2 = 1$.*

The x-coordinates of points in H_1^* are distinct integers in the interval (u, x_1). Moreover, if for any integer i we consider the integers q_i, r_i so that $iy_1 = q_i x_1 + r_i$ with $0 \le r_i < x_1$ then any integer $k \in (u, x_1)$ is the x-coordinate of some $\boldsymbol{p} \in H_1^*$ (i.e., $k \in \pi_1(H_1^*)$) if and only if $q_k = v - 1 + q_{k-u}$, or equivalently, if $r_k \ge x_1 - (vx_1 - uy_1)$. In that case, $\boldsymbol{p} = (k, q_k + 1)$. These observations (detailed in Lemma 5.18) allow us to test the AG1 property as follows.

Proposition 5.20. *The semigroup H is AG1 if and only if $r_k + r_\ell < 2x_1 - (vx_1 - uy_1)$ for all integers $k, \ell \in \pi_1(H_1^*)$ with $k + \ell < x_1$.*

When the bottom element is $(1, 1)$ (i.e., $y_1 < x_1$ and $x_2 < y_2$) we can describe recursively the points in H_1^*.

Lemma 6.1. *Assume $(1, 1) \in \omega_H$ and $H_1^* \neq \varnothing$. Let $n = |H_1^*| = x_1 - y_1 - 1$. Recursively, we define nonnegative integers $\ell_1, \dots, \ell_n$ and $s_1, \dots, s_n$ by*

$$x_1 = \ell_1(x_1 - y_1) + s_1, \quad \textit{with} \quad s_1 < x_1 - y_1,$$

and

$$y_1 + s_{i-1} = \ell_i(x_1 - y_1) + s_i \quad \textit{with} \quad s_i < x_1 - y_1,$$

for $i = 2, \dots, n$. Then

$$H_1^* = \left\{ \boldsymbol{p}_t = (c_t, d_t) : c_t = t + \sum_{i=1}^{t} \ell_i,\ d_t = \sum_{i=1}^{t} \ell_i,\ t = 1, \dots, n \right\}.$$

A similar description is available for points in H_2^*. A little bit more effort is necessary to obtain the following arithmetic criterion for the Ulrich property of $(1, 1)$. The effort is compensated with the simplicity of the statement.

Theorem 6.3. *Assume $(1, 1) \in \omega_H$. Then $(1, 1)$ is an Ulrich element in H if and only if $x_i \equiv 1 \bmod (x_i - y_i)$ for $i = 1, 2$.*

Consequently, by Corollary 6.4, if $x_1y_1x_2y_2 \neq 0$ the ring $K[H]$ is AG if and only if $x_i \equiv 1 \bmod (x_i - y_i)$ for $i = 1, 2$.

In Section 7 we discuss another extension of the Gorenstein property for affine semigroup rings. According to the definition proposed in [Herzog et al. 2019], any Cohen–Macaulay ring $K[H]$ is called nearly Gorenstein if the trace ideal $\mathrm{tr}(\omega_{K[H]})$ contains the graded maximal ideal of $K[H]$. For one-dimensional rings, the almost Gorenstein property implies the nearly Gorenstein property, but for rings of larger dimension there is no implication between these two properties. We prove in Theorem 7.1 that when H is a normal semigroup in $\mathbb{N}^2$ the ring $K[H]$ is nearly Gorenstein. Example 7.2 shows that the statement is not valid in higher embedding dimensions.

1. Background on affine semigroups and their toric rings

In this paper all semigroups considered are fully embedded, i.e., when writing $H \subseteq \mathbb{N}^d$ we shall implicitly assume that the group generated by H is $\mathrm{gr}(H) = \mathbb{Z}^d$. A subset $H \subseteq \mathbb{N}^d$ is called an affine semigroup if there exist $\boldsymbol{c}_1, \dots, \boldsymbol{c}_r \in H$ such that $H = \sum_{i=1}^r \mathbb{N}\boldsymbol{c}_i$. Moreover, H is called a normal semigroup if for all $\boldsymbol{h}$ in $\mathbb{N}^d$ and n positive integer, $n\boldsymbol{h} \in H$ implies that $\boldsymbol{h} \in H$.

Let K be any field and $H = \sum_{i=1}^r \mathbb{N}\boldsymbol{c}_i \subseteq \mathbb{N}^d$. The semigroup ring $K[H]$ is the subalgebra of the polynomial ring $K[x_1, \dots, x_d]$ generated by the monomials with exponents in H. Then H is normal if and only if the semigroup ring $K[H]$ is integrally closed in its field of fractions [Bruns and Gubeladze 2009]. The normality for H is also equivalent to the fact that H contains all the lattice points of the rational polyhedral cone C that it generates, i.e., $H = C \cap \mathbb{Z}^d$, where

$$C = \left\{ \sum_{i=1}^{r} \lambda_i \boldsymbol{c}_i : \lambda_i \in \mathbb{R}_{\geq 0},\ \text{for } i = 1, \dots, r \right\}.$$

The dimension (or rank) of H is defined as the dimension of $\mathrm{aff}(H)$, the affine subspace it generates. The latter is the same as $\mathrm{aff}(C)$.

Let $\langle \cdot\,, \cdot \rangle$ denote the usual scalar product in $\mathbb{R}^d$. Given $\boldsymbol{n} \in \mathbb{R}^d \setminus \{0\}$, the hyperplane $H_{\boldsymbol{n}} = \{z \in \mathbb{R}^d : \langle z, \boldsymbol{n} \rangle = 0\}$ is called a *support hyperplane* for C if $\langle z, \boldsymbol{n} \rangle \geq 0$ for all $z \in C$ and $H_{\boldsymbol{n}} \cap C \neq \varnothing$. In this case, the cone $H_{\boldsymbol{n}} \cap C$ is called a *face* of C and its dimension is $\dim \mathrm{aff}(H_{\boldsymbol{n}} \cap C)$. Let F be any face of the cone C. When $\dim F = 1$, the face F is called an *extremal ray*, and when $\dim F = d - 1$, it is called a *facet* of C. The normal vector to any hyperplane is determined up to multiplication by a nonzero factor; hence we may choose $\boldsymbol{n}_1, \dots, \boldsymbol{n}_s \in \mathbb{Z}^d$ to be the normals to the support hyperplanes that determine the facets of C and such that

$$C = \{z \in \mathbb{R}^d : \langle z, \boldsymbol{n}_i \rangle \geq 0,\ \text{for } i = 1, \dots, s\}.$$

The unique minimal set of generators for the semigroup H is called the *Hilbert basis* of H and we denote it as B_H.

It is known that the cone C has at least d facets and at least d extremal rays. When C has d facets (equivalently, that it has d extremal rays) the cone C and the semigroup H are called *simplicial*.

For any $d \geq 2$ we denote by $\mathcal{N}_d$ the class of normal simplicial affine semigroups which are fully embedded in $\mathbb{N}^d$.

Let $H \in \mathcal{N}_d$ and C the cone over H. On each extremal ray of C there exists a unique primitive element from H, which we call an *extremal ray* for H. Denote $\boldsymbol{a}_1, \dots, \boldsymbol{a}_d$ the extremal rays for H. These form an $\mathbb{R}$-basis in $\mathbb{R}^d$. For $\boldsymbol{z} \in \mathbb{R}^d$ such that $\boldsymbol{z} = \sum_{i=1}^d \lambda_i \boldsymbol{a}_i$ with $\lambda_i \in \mathbb{R}$, $i = 1, \dots, d$, we set $[\boldsymbol{z}]_i = \lambda_i$ for $i = 1, \dots, d$. In this notation, $\boldsymbol{z}$ is in the cone C if and only if $[\boldsymbol{z}]_i \geq 0$ for $i = 1, \dots, d$. Also, when $\boldsymbol{z} \in \mathbb{Z}^d$ one has that $\boldsymbol{z} \in H$ if and only if $[\boldsymbol{z}]_i \geq 0$ for $i = 1, \dots, d$.

The fundamental (semiopen) *parallelotope* of H is the set

$$P_H = \left\{ \boldsymbol{z} \in \mathbb{R}^d : \boldsymbol{z} = \sum_{i=1}^d \lambda_i \boldsymbol{a}_i \text{ with } 0 \leq \lambda_i < 1 \text{ for } i = 1, \dots, d \right\}.$$

Its closure in $\mathbb{R}^d$ is the set $\overline{P_H} = \{ \boldsymbol{z} \in \mathbb{R}^d : 0 \leq [\boldsymbol{z}]_i \leq 1 \text{ for } i = 1, \dots, d \}$. It is well known, and easy to see, that any $\boldsymbol{h}$ in H decomposes uniquely as $\boldsymbol{h} = \sum_{i=1}^d n_i \boldsymbol{a}_i + \boldsymbol{h}'$ with $\boldsymbol{h}' \in P_H \cap \mathbb{Z}^d$ and $n_1, \dots, n_d$ nonnegative integers. The extremal rays of H are in $B_H \setminus P_H$, but the rest of the elements in B_H belong to P_H.

Since H is simplicial, $\boldsymbol{x}^{\boldsymbol{a}_1}, \dots, \boldsymbol{x}^{\boldsymbol{a}_d}$ is a system of parameters in R; see [Goto et al. 1976, (1.11)]. As H is a normal semigroup, by [Hochster 1972], the ring $R = K[H]$ is Cohen–Macaulay of dimension d; hence any system of parameters in R is a regular sequence of maximal length. By [Danilov 1978; Stanley 1978], the canonical module ω_R of R is the ideal in R generated by the monomials $\boldsymbol{x}^{\boldsymbol{v}}$ whose exponent vector $\boldsymbol{v} = \log(\boldsymbol{x}^{\boldsymbol{v}})$ belongs to the relative interior of C, denoted by $\operatorname{relint} C$. Note that

$$\operatorname{relint} C = \left\{ \boldsymbol{c} \in \mathbb{R}^d : \boldsymbol{c} = \sum_{i=1}^d \lambda_i \boldsymbol{a}_i \text{ with } \lambda_i \in \mathbb{R}_{>0} \text{ for all } i = 1, \dots, d \right\}.$$

We set

$$\omega_H = \{ \boldsymbol{h} \in H : \boldsymbol{x}^{\boldsymbol{h}} \in \omega_R \} = \mathbb{Z}^d \cap \operatorname{relint} C,$$

which is a semigroup ideal of H, i.e., $\omega_H + H \subseteq \omega_H$. We note that $\boldsymbol{h} \in \mathbb{Z}^d$ is in ω_H if and only if $[\boldsymbol{h}]_i > 0$ for $i = 1, \dots, d$.

The ideal ω_R has a unique minimal system of monomial generators which we denote by $G(\omega_R)$. We set $G(\omega_H) = \{\log(u) : u \in G(\omega_R)\}$. Clearly, $G(\omega_H)$ is the unique minimal system of generators for ω_H. The situation when $G(\omega_H)$

is a singleton corresponds to the situation when R is a Gorenstein ring. When $B_H = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_d\}$ then R is a regular ring, there is no lattice point in the relative interior of $\overline{P_H}$, and $G(\omega_H) = \{\sum_{i=1}^d \boldsymbol{a}_i\}$.

The following easy lemma describes the minimal generators of ω_H when $H \in \mathcal{N}_2$. For completeness, we include a proof here.

Lemma 1.1. *Let H be in $\mathcal{N}_2$ with the extremal rays $\boldsymbol{a}_1, \boldsymbol{a}_2$. Then $B_H \cap \omega_H = B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2\}$. Moreover, if $\{\boldsymbol{a}_1, \boldsymbol{a}_2\} \subsetneq B_H$, then $G(\omega_H) = B_H \cap \omega_H$.*

Proof. Clearly, $B_H \cap \omega_H \subseteq B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2\}$. If $\boldsymbol{b} \in B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2\}$ then, since $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are the extremal rays, it follows that $\boldsymbol{b} \in \operatorname{relint} P_H$; hence $\boldsymbol{b} \in B_H \cap \omega_H$. Therefore, $B_H \cap \omega_H = B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2\}$.

Assume $\{\boldsymbol{a}_1, \boldsymbol{a}_2\} \subsetneq B_H$. Let $\boldsymbol{b} \in G(\omega_H)$. The only lattice points on the boundary of the parallelogram $\overline{P_H}$ are $0, \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_1 + \boldsymbol{a}_2$. None of them is in $G(\omega_H)$, under our hypothesis. Thus $\boldsymbol{b} \in \operatorname{relint} P_H$. If, on the contrary, $\boldsymbol{b} \notin B_H$, then $\boldsymbol{b}$ is the sum of at least two elements in B_H, out of which at least one is not in ω_H, i.e., the latter is $\boldsymbol{a}_1$ or $\boldsymbol{a}_2$. This implies that $\boldsymbol{b} \notin P_H$, a contradiction. Consequently, $G(\omega_H) \subseteq B_H \cap \omega_H$. The reverse inclusion is clear. □

We refer to the monographs [Bruns and Herzog 1998; Bruns and Gubeladze 2009; Villarreal 2015; Ziegler 1995; Fulton 1993] for more background on affine semigroups, their semigroup rings, rational cones and the connections with algebraic geometry.

2. A regular sequence in $K[H]$ and slim semigroups

Throughout this section H is a semigroup in $\mathcal{N}_d$ having the extremal rays $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_d$, and $R = K[H]$. The main result is Theorem 2.4 where we present equivalent conditions for the existence of a convenient reduction for the graded maximal ideal of $K[H]/(\boldsymbol{x}^{\boldsymbol{b}})$, where $\boldsymbol{b}$ is any nonzero element in H. This result motivates us to introduce the class of slim semigroups.

The following lemma plays a crucial role in the proof of Proposition 2.2 and in several other arguments in this paper.

Lemma 2.1. *Let $\boldsymbol{b} = \sum_{i=1}^d \lambda_i \boldsymbol{a}_i$, with $\lambda_1, \ldots, \lambda_d \in \mathbb{R}$.*

(a) *Let $n_i = \lfloor \lambda_i \rfloor + 1$ for $i = 1, \ldots, d$. Then $\left(\sum_{i=1}^d n_i \boldsymbol{a}_i\right) - \boldsymbol{b} \in \omega_H$.*

(b) *If $\boldsymbol{b} \in \overline{P_H}$ then $\left(\sum_{i=1}^d \boldsymbol{a}_i\right) - \boldsymbol{b} \in H$, and in particular, if $\boldsymbol{b} \in P_H$, then $\left(\sum_{i=1}^d \boldsymbol{a}_i\right) - \boldsymbol{b} \in \omega_H$.*

Proof. For (a) we note that $\left(\sum_{i=1}^d n_i \boldsymbol{a}_i\right) - \boldsymbol{b} = \sum_{i=1}^d (1 - \{\lambda_i\}) \boldsymbol{a}_i$ and $0 < 1 - \{\lambda_i\} \le 1$ for all i; hence the sum of interest is in ω_H. Here we denoted $\{\lambda_i\} = \lambda_i - \lfloor \lambda_i \rfloor$ for all i. Part (b) follows immediately. □

Proposition 2.2. *For any $\boldsymbol{b} \neq 0$ in H, the sequence $\boldsymbol{x}^{\boldsymbol{b}}, \boldsymbol{x}^{\boldsymbol{a}_1} - \boldsymbol{x}^{\boldsymbol{a}_2}, \dots, \boldsymbol{x}^{\boldsymbol{a}_1} - \boldsymbol{x}^{\boldsymbol{a}_d}$ is a regular sequence on R.*

Proof. In order to simplify notation we set $u = \boldsymbol{x}^{\boldsymbol{b}}$ and $v_i = \boldsymbol{x}^{\boldsymbol{a}_i}$ for $i = 1, \dots, d$. Let $I = (u, v_1 - v_2, \dots, v_1 - v_d)$. We may write $\boldsymbol{b} = \sum_{i=1}^d \lambda_i \boldsymbol{a}_i$ with $\lambda_i \geq 0$ for $i = 1, \dots, d$. We denote $n_i = \lfloor \lambda_i \rfloor + 1$ for all i and we set $N = \sum_{i=1}^d n_i$.

We will show that $v_i^N \in I$ for $i = 1, \dots, d$. Since $v_i - v_j \in I$ for all i and j, it follows by symmetry that it is enough to show that $v_1^N \in I$.

We write

$$\begin{aligned} v_1^N &= (v_1^{n_2} - v_2^{n_2}) \cdot v_1^{N-n_2} + v_1^{N-n_2} v_2^{n_2} \\ &= (v_1^{n_2} - v_2^{n_2}) \cdot v_1^{N-n_2} + v_1^{N-n_2-n_3} v_2^{n_2} (v_1^{n_3} - v_3^{n_3}) + v_1^{N-n_2-n_3} v_2^{n_2} v_3^{n_3} \\ &= \sum_{i=2}^{d} v_1^{N-\sum_{j=2}^i n_j} v_2^{n_2} \cdots v_{i-1}^{n_{i-1}} (v_1^{n_i} - v_i^{n_i}) + v_1^{n_1} \cdots v_d^{n_d}. \end{aligned}$$

Hence by Lemma 2.1 it follows that $v_1^N \in I$.

Since H is a normal simplicial semigroup, $v_1, \dots, v_d$ is a regular sequence in R; hence $v_1^N, \dots, v_d^N$ is a regular sequence in R, as well. Since R is a Cohen–Macaulay ring of dimension d we get that $v_1^N, \dots, v_d^N$ is also a system of parameters for R. Thus $0 < \lambda(R/I) \leq \lambda(R/(v_1^N, \dots, v_d^N)) < \infty$, which implies that $u, v_1 - v_2, \dots, v_1 - v_d$ is a system of parameters, and consequently a regular sequence for R. Here $\lambda(M)$ denotes the length of an R-module M. □

In the sequel, our aim is to find a reduction ideal for the graded maximal ideal $\mathfrak{m}$ of R, modulo the ideal $\boldsymbol{x}^{\boldsymbol{b}} R$, for any $\boldsymbol{b} \in H \setminus \{0\}$. In this order, we need the following lemma which is interesting on its own.

Lemma 2.3. *For any $\boldsymbol{b}$ in H, there exists a positive integer k such that for all $\boldsymbol{c}_1, \dots, \boldsymbol{c}_k$ in ω_H, one has $\boldsymbol{c}_1 + \cdots + \boldsymbol{c}_k \in \boldsymbol{b} + H$.*

Proof. Assume $\boldsymbol{n}_1, \dots, \boldsymbol{n}_r \in \mathbb{Z}^d$ are normal vectors to the support hyperplanes of the facets of the cone C such that

$$C = \{\boldsymbol{x} \in \mathbb{R}^d : \langle \boldsymbol{x}, \boldsymbol{n}_i \rangle \geq 0, \text{ for all } i = 1, \dots, r\}.$$

The map $\sigma : \mathbb{R}^d \to \mathbb{R}^r$ given by

$$\sigma(\boldsymbol{h}) = (\langle \boldsymbol{h}, \boldsymbol{n}_1 \rangle, \dots, \langle \boldsymbol{h}, \boldsymbol{n}_r \rangle), \text{ for all } \boldsymbol{h} \in \mathbb{R}^d$$

is clearly $\mathbb{R}$-linear and $\sigma(H) \subseteq \mathbb{N}^r$.

Let $k_0 = \max\{\langle \boldsymbol{b}, \boldsymbol{n}_j \rangle : j = 1, \dots, r\}$. For any integer $k > k_0$, any $\boldsymbol{c}_1, \dots, \boldsymbol{c}_k \in H \cap \operatorname{relint} C$, and any $1 \leq j \leq r$, the j-th component of $\sigma(\boldsymbol{c}_1 + \cdots + \boldsymbol{c}_k - \boldsymbol{b})$ equals $\left(\sum_{i=1}^k \langle \boldsymbol{c}_i, \boldsymbol{n}_j \rangle\right) - \langle \boldsymbol{b}, \boldsymbol{n}_j \rangle \geq k - k_0 > 0$; hence $\boldsymbol{c}_1 + \cdots + \boldsymbol{c}_k \in \boldsymbol{b} + H$. □

Theorem 2.4. *Let* $R = K[H]$, $J = (\boldsymbol{x}^{\boldsymbol{a}_i} - \boldsymbol{x}^{\boldsymbol{a}_j} : i, j = 1, \ldots, d)R$ *and* $0 \neq \boldsymbol{b} \in \omega_H$. *Then the following statements are equivalent*:

(i) *There exists an integer* k *such that* $\mathfrak{m}^{k+1} = J\mathfrak{m}^k$ *modulo the ideal* $\boldsymbol{x}^{\boldsymbol{b}}R$.

(ii) $\sum_{j=1}^{d} [\boldsymbol{c}]_j \geq 1$, *for any* $\boldsymbol{c} \in B_H \setminus \omega_H$.

Proof. We denote $\{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_r\} = B_H \setminus (\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_d\} \cup \omega_H)$.

(i) ⇒ (ii): For any $\boldsymbol{c}$ in H we set $l(\boldsymbol{c}) = \sum_{j=1}^{d} [\boldsymbol{c}]_j$. Clearly, $l(\boldsymbol{a}_i) = 1$ for $i = 1, \ldots, d$, so if $r = 0$, we are done. Assuming $r > 0$, we pick t such that $l(\boldsymbol{c}_t) = \min\{l(\boldsymbol{c}_j) : j = 1, \ldots, r\}$. Let $k > 0$ so that $\mathfrak{m}^{k+1} + \boldsymbol{x}^{\boldsymbol{b}}R = J\mathfrak{m}^k + \boldsymbol{x}^{\boldsymbol{b}}R$, i.e., $\mathfrak{m}^{k+1} \subseteq J\mathfrak{m}^k + \boldsymbol{x}^{\boldsymbol{b}}R$.

As $\boldsymbol{c}_t \notin \omega_H$, there exists $1 \leq i_0 \leq d$ with $[\boldsymbol{c}_t]_{i_0} = 0$; hence $(k+1)\boldsymbol{c}_t \notin \boldsymbol{b} + H$. From $\boldsymbol{x}^{(k+1)\boldsymbol{c}_t} \in J\mathfrak{m}^k + \boldsymbol{x}^{\boldsymbol{b}}R$ we get that

$$\boldsymbol{x}^{(k+1)\boldsymbol{c}_t} = \boldsymbol{x}^{\boldsymbol{a}_j}\boldsymbol{x}^{\boldsymbol{u}_1 + \cdots + \boldsymbol{u}_k},$$

for some $1 \leq j \leq d$ and $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k \in H \setminus (\omega_H \cup \{0\})$. Consequently,

$$(k+1)l(\boldsymbol{c}_t) = 1 + \sum_{i=1}^{k} l(\boldsymbol{u}_i) \geq 1 + k\min\{1, l(\boldsymbol{c}_t)\},$$

which implies $l(\boldsymbol{c}_t) \geq 1$.

(ii) ⇒ (i): Let $u = \boldsymbol{x}^{\boldsymbol{b}}$ and $v_i = \boldsymbol{x}^{\boldsymbol{a}_i}$ for $i = 1, \ldots, d$. We decompose $\boldsymbol{b} = \sum_{i=1}^{d} \lambda_i \boldsymbol{a}_i$ with $\lambda_i \geq 0$ and we set $n_i = \lfloor \lambda_i \rfloor + 1$ for all $i = 1, \ldots, d$ and $N = \sum_{i=1}^{d} n_i$.

We claim that for any positive integer t, any $i_1, \ldots, i_t \in \{1, \ldots, d\}$ and any $v \in \{v_1, \ldots, v_d\}$ one has

$$v_{i_1} \cdots v_{i_t} \in J\mathfrak{m}^{t-1} + v^t R. \tag{3}$$

Indeed, this is a consequence of the following equations:

$$\begin{aligned} v_{i_1} \cdots v_{i_t} &= (v_{i_1} - v) \cdot v_{i_2} \cdots v_{i_t} + v \cdot v_{i_2} \cdots v_{i_t} \\ &= (v_{i_1} - v) \cdot v_{i_2} \cdots v_{i_t} + v(v_{i_2} - v) \cdot v_{i_3} \cdots v_{i_t} + v^2 v_{i_3} \cdots v_{i_t} \\ &= \Sigma_{j=1}^{d} v^{j-1} \cdot (v_{i_j} - v) \cdot v_{i_{j+1}} \cdots v_{i_t} + v^t. \end{aligned}$$

Now let $i_1, \ldots, i_N \in \{1, \ldots, d\}$. In the product $v_{i_1} \ldots v_{i_N}$ we apply (3) to the first n_1 terms, then to the next n_2 terms, etc., and we obtain that

$$v_{i_1} \ldots v_{i_N} \in \prod_{i=1}^{d} (J\mathfrak{m}^{n_i - 1}, v_i^{n_i}) \subseteq \left(J\mathfrak{m}^{N-1}, \prod_{i=1}^{d} v_i^{n_i}\right) \subseteq (J\mathfrak{m}^{N-1}, uR), \tag{4}$$

where for the last inclusion we used Lemma 2.1.

For $1 \le i \le r$ we may write $\boldsymbol{c}_i = \sum_{j=1}^{d} (p_{ij}/q_i)\boldsymbol{a}_i$, where q_i, p_{ij} are nonnegative integers and $q_i > 0$. Hence, $Nq_i \cdot \boldsymbol{c}_i = \sum_{j=1}^{d} Np_{ij}\boldsymbol{a}_j$, where by the hypothesis of (ii) we have $N \le Nq_i \le \sum_{j=1}^{d} Np_{ij}$. Thus, using (4) we derive

$$(\boldsymbol{x}^{\boldsymbol{c}_i})^{Nq_i} = \prod_{j=1}^{d} (\boldsymbol{x}^{\boldsymbol{a}_j})^{Np_{ij}} \in (J\mathfrak{m}^{\sum_{j=1}^{d} Np_{ij}-1}, uR) \subseteq (J\mathfrak{m}^{Nq_i-1}, uR).$$

We set $N_1 = \max\{Nq_i : 1 \le i \le r\}$ if $r > 0$, otherwise we let $N_1 = N$. Then

$$(\boldsymbol{x}^{\boldsymbol{c}})^{N_1} \subseteq (J\mathfrak{m}^{N_1-1}, uR) \quad \text{for all } \boldsymbol{c} \in B_H \setminus \omega_H. \tag{5}$$

Let k_0 be a positive integer satisfying the conclusion of Lemma 2.3 for the element $\boldsymbol{b}$. We set $k = k_0 + N_1 \cdot |B_H \setminus \omega_H| - 2$.

Let w be any product of $k+1$ monomial generators of $\mathfrak{m}$. If the exponents of at least k_0 of them are in relint C, then by the choice of k_0 we get that $w \in uR$. Otherwise, we may write $w = (\boldsymbol{x}^{\boldsymbol{c}})^{N_1} \cdot w'$ for some $\boldsymbol{c} \in B_H \setminus \omega_H$ and $w' \in \mathfrak{m}^{k+1-N_1}$. In the latter case, using (5) we infer that $w \in J\mathfrak{m}^k + uR$. This shows that $\mathfrak{m}^{k+1} + uR = J\mathfrak{m}^k + uR$, which completes the proof. □

The semigroups H satisfying the equivalent conditions of Theorem 2.4 deserve a special name.

Definition 2.5. A semigroup $H \in \mathcal{N}_d$ is called *slim* if $\sum_{i=1}^{d} [c]_i \ge 1$, for any $\boldsymbol{c} \in B_H \setminus \omega_H$.

We will denote by $\mathcal{H}_d$ the class of slim semigroups in $\mathbb{N}^d$.

Remark 2.6. Let $H \in \mathcal{N}_d$. If $B_H \setminus \omega_H = \{\boldsymbol{a}_1, \dots, \boldsymbol{a}_d\}$, then H is slim. In particular, by Lemma 1.1, any normal semigroup in $\mathbb{N}^2$ is slim, so $\mathcal{N}_2 = \mathcal{H}_2$.

Proposition 2.7. *Let $H \in \mathcal{N}_3$. Then H is slim if and only if $[\boldsymbol{c}]_1 + [\boldsymbol{c}]_2 + [\boldsymbol{c}]_3 = 1$ for any $\boldsymbol{c} \in B_H \setminus \omega_H$.*

Proof. Assume H is slim. If $B_H \setminus \omega_H = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$, there is nothing left to prove. Assume there exists $\boldsymbol{c} \in B_H \setminus (\omega_H \cup \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\})$ such that $\sum_{i=1}^{3} [\boldsymbol{c}]_i > 1$. Without loss of generality, we may assume that $0 < [\boldsymbol{c}]_i < 1$ for $i = 1, 2$ and $[\boldsymbol{c}]_3 = 0$. Arguing as in the proof of Lemma 2.1 we obtain that $0 \ne \boldsymbol{a}_1 + \boldsymbol{a}_2 - \boldsymbol{c} \in \mathbb{N}\boldsymbol{a}_1 + \mathbb{N}\boldsymbol{a}_2 \subset H$. We may write $\boldsymbol{a}_1 + \boldsymbol{a}_2 - \boldsymbol{c} = \boldsymbol{b} + \boldsymbol{h}$, where $\boldsymbol{b} \in B_H \setminus \omega_H$ and $\boldsymbol{h} \in H$. Then $\sum_{i=1}^{3} [\boldsymbol{b}]_i = [\boldsymbol{b}]_1 + [\boldsymbol{b}]_2 \le \sum_{i=1}^{2} [\boldsymbol{a}_1 + \boldsymbol{a}_2 - \boldsymbol{c}]_i = 2 - \sum_{i=1}^{2} [\boldsymbol{c}_i] < 1$, which is false since H is slim. □

Example 2.8. The semigroup $L \in \mathcal{N}_3$ with the extremal rays $\boldsymbol{a}_1 = (11, 13, 0)$, $\boldsymbol{a}_2 = (3, 4, 0)$ and $\boldsymbol{a}_3 = (0, 0, 1)$ is not slim. Indeed, $B_L = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, (4, 5, 0), (5, 6, 0)\}$ (compare with Example 4.6) and it is easy to check that $\sum_{i=1}^{3} [(4, 5, 0)]_i = \frac{4}{5} < 1$.

3. Ulrich elements and the almost Gorenstein property

The theory of almost Gorenstein rings has its origin in the theory of the almost symmetric numerical semigroups in [Barucci and Fröberg 1997]. If R is the semigroup ring of a numerical semigroup, then the semigroup is almost symmetric, if and only if there exists an exact sequence

$$0 \to R \to (\omega_R)(-a) \to E \to 0, \tag{6}$$

where E is annihilated by the graded maximal ideal of R; see [Herzog and Watanabe 2019]. Here ω_R denotes the canonical module of R and $-a$ the smallest degree of a generator of ω_R, i.e., $-a = \min\{k : (\omega_R)_k \neq 0\}$

In [Goto et al. 2013] the 1-dimensional positively graded rings which admit such an exact sequence are called almost Gorenstein.

Goto et al. [2015, Definition 8.1] extended the concept of the almost Gorenstein property to rings of higher dimension: let R be a positively graded Cohen–Macaulay K-algebra with a-invariant a. Then R is called *graded almost Gorenstein*, if there exists an exact sequence like in (6), where E is an Ulrich module.

Ulrich modules are defined as follows: let $(R, \mathfrak{m}, K)$ be a local (or positively graded) ring with (graded) maximal ideal $\mathfrak{m}$, and let M be a (graded) Cohen–Macaulay module over R. Then the minimal number of generators $\mu(M)$ of M is bounded by the multiplicity $e(M)$ of M. The module M is called an *Ulrich module*, if $\mu(M) = e(M)$. Ulrich [1984] asked whether any Cohen–Macaulay ring admits an Ulrich module M with $\dim M = \dim R$. At present this question is still open, and has an affirmative answer for example when R is a hypersurface ring [Backelin and Herzog 1989].

In the case of almost symmetric numerical semigroup rings, the module E in the exact sequence (6) is of Krull dimension zero. A graded module M with $\dim M = 0$ is Ulrich if and only if $\mathfrak{m}M = 0$. Thus the above definition [Goto et al. 2015, Definition 8.1] is a natural extension of 1-dimensional almost Gorenstein rings to higher dimensions.

We propose the following multigraded version of the almost Gorenstein property for normal semigroup rings.

Definition 3.1. Let H be a normal affine semigroup and $R = K[H]$. For $\boldsymbol{b} \in \omega_H$ consider the exact sequence

$$0 \to R \to \omega_R(\boldsymbol{b}) \to E \to 0, \tag{7}$$

where $1 \in R$ is mapped to $u = \boldsymbol{x}^{\boldsymbol{b}}$ and $E = \omega_R/uR$. Then $\boldsymbol{b}$ is called an *Ulrich element* in H, if E is an Ulrich R-module.

If H admits an Ulrich element $\boldsymbol{b}$, then we call the ring R *almost Gorenstein with respect to* $\boldsymbol{b}$, or simply AG if H has an Ulrich element.

Theorem 3.2. *Let $H \in \mathcal{H}_d$ with extremal rays $\boldsymbol{a}_1, \dots, \boldsymbol{a}_d$. Let $\mathfrak{m}$ be the graded maximal ideal of $R = K[H]$. Let $\boldsymbol{b} \in \omega_H$, $u = \boldsymbol{x}^{\boldsymbol{b}}$ and $J = (\boldsymbol{x}^{\boldsymbol{a}_i} - \boldsymbol{x}^{\boldsymbol{a}_j} : i, j = 1, \dots, d)R$. Then $\boldsymbol{b}$ is an Ulrich element in H if and only if*

$$\mathfrak{m}\omega_R \subseteq (uR, J\omega_R). \tag{8}$$

Proof. Since uR and ω_R are Cohen–Macaulay R-modules of dimension d, we see (keeping the notation from (7)) that depth $E \geq d-1$, and since uR and ω_R are rank 1 modules, we deduce that $\mathrm{Ann}(E) \neq 0$. Therefore, $\dim E \leq d-1$, and this implies that E is a Cohen–Macaulay R-module of dimension $d-1$.

Suppose that (8) holds. By [Goto et al. 2015, Proposition 2.2.(2)], it follows that E is an Ulrich module since (8) implies that $\mathfrak{m}E = JE$ and since J is generated by $d-1 (= \dim E)$ elements, namely by the elements $f_j = \boldsymbol{x}^{\boldsymbol{a}_1} - \boldsymbol{x}^{\boldsymbol{a}_j}$ with $j = 2, \dots, d$. Thus $\boldsymbol{b}$ is an Ulrich element in H.

Conversely, assume that $\boldsymbol{b}$ is an Ulrich element. Then E is an Ulrich module, and therefore $\lambda(E/\mathfrak{m}E) = e(E)$. It follows from Theorem 2.4 that J is a reduction ideal of $\mathfrak{m}$ with respect to E. Thus by [Bruns and Herzog 1998, Lemma 4.6.5], $e(E) = e(J, E)$, where $e(J, E)$ denotes the Hilbert–Samuel multiplicity of E with respect to J. Since E is Cohen–Macaulay of dimension $d-1$, and since J is generated by the $d-1$ elements $f_2, \dots, f_d$ and $\lambda(E/JE) < \infty$, we see that $f_2, \dots, f_d$ is a regular sequence on E. Thus [Bruns and Herzog 1998, Theorem 4.7.6] implies that $e(J, E) = \lambda(E/JE)$. Hence, $\lambda(E/\mathfrak{m}E) = \lambda(E/JE)$. Since $JE \subseteq \mathfrak{m}E$, it follows that $\mathfrak{m}E = JE$, and this implies (8). □

Remark 3.3. It follows from the proof of Theorem 3.2 that if $H \in \mathcal{N}_d$ and (8) holds for some ideal $J \subset \mathfrak{m}$, generated by $d-1$ elements, then $\boldsymbol{b}$ is an Ulrich element in H.

Example 3.4 (Ulrich elements in Gorenstein and regular rings).

(a) If $K[H]$ is a Gorenstein ring and $G(\omega_H) = \{\boldsymbol{b}\}$, then $\omega_R = \boldsymbol{x}^{\boldsymbol{b}}R$; hence (8) holds and $\boldsymbol{b}$ is an Ulrich element in H.

(b) Assume $K[H]$ is a regular ring with $\boldsymbol{a}_1, \dots, \boldsymbol{a}_d$ the extremal rays of H. Set $\boldsymbol{c} = \sum_{i=1}^{d} \boldsymbol{a}_i$. Then $\boldsymbol{a}_i + \boldsymbol{c}$ is an Ulrich element in H for any $i = 1, \dots, d$. Indeed, since $\mathfrak{m} = (\boldsymbol{x}^{\boldsymbol{a}_j} : 1 \leq j \leq d)$ and $\boldsymbol{x}^{\boldsymbol{a}_j + \boldsymbol{c}} = \boldsymbol{x}^{\boldsymbol{c}}(\boldsymbol{x}^{\boldsymbol{a}_j} - \boldsymbol{x}^{\boldsymbol{a}_i}) + \boldsymbol{x}^{\boldsymbol{c} + \boldsymbol{a}_i}$ for $j = 1, \dots, d$, we have that (8) is verified for $\boldsymbol{b} = \boldsymbol{c} + \boldsymbol{a}_i$.

Example 3.5. Let $H \in \mathcal{H}_2$ having the extremal rays $\boldsymbol{a}_1 = (11, 2)$ and $\boldsymbol{a}_2 = (31, 6)$. A computation with [Normaliz] shows that the Hilbert basis of H is

$$B_H = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{b} = (16, 3), \boldsymbol{c}_1 = (21, 4), \boldsymbol{c}_2 = (26, 5)\}.$$

Moreover, $\boldsymbol{b}, \boldsymbol{c}_1, \boldsymbol{c}_2$ are the only nonzero lattice points in P_H, and they all lie on the line $y = (x-1)/5$ passing through $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. Comparing componentwise, we have

$$\boldsymbol{a}_1 \preceq \boldsymbol{b} \preceq \boldsymbol{c}_1 \preceq \boldsymbol{c}_2 \preceq \boldsymbol{a}_2.$$

We note that $\boldsymbol{a}_1 + \boldsymbol{a}_2 = \boldsymbol{b} + \boldsymbol{c}_2 = 2\boldsymbol{c}_1$. It is also straightforward to check that in $K[H]$ we have

$$\begin{aligned} &\boldsymbol{x}^{\boldsymbol{a}_1}\boldsymbol{x}^{\boldsymbol{c}_1} = (\boldsymbol{x}^{\boldsymbol{b}})^2, \quad \boldsymbol{x}^{\boldsymbol{a}_1}\boldsymbol{x}^{\boldsymbol{c}_2} = \boldsymbol{x}^{\boldsymbol{b}}\boldsymbol{x}^{\boldsymbol{c}_1}, \quad \boldsymbol{x}^{\boldsymbol{c}_1}\boldsymbol{x}^{\boldsymbol{c}_1} = \boldsymbol{x}^{\boldsymbol{b}}\boldsymbol{x}^{\boldsymbol{c}_2}, \quad \boldsymbol{x}^{\boldsymbol{c}_1}\boldsymbol{x}^{\boldsymbol{c}_2} = \boldsymbol{x}^{\boldsymbol{b}}\boldsymbol{x}^{\boldsymbol{a}_2},\\ &\boldsymbol{x}^{\boldsymbol{c}_2}\boldsymbol{x}^{\boldsymbol{c}_2} = \boldsymbol{x}^{\boldsymbol{a}_2}\boldsymbol{x}^{\boldsymbol{c}_1} = (\boldsymbol{x}^{\boldsymbol{a}_2} - \boldsymbol{x}^{\boldsymbol{a}_1})\boldsymbol{x}^{\boldsymbol{c}_1} + \boldsymbol{x}^{\boldsymbol{a}_1}\boldsymbol{x}^{\boldsymbol{c}_1} = (\boldsymbol{x}^{\boldsymbol{a}_2} - \boldsymbol{x}^{\boldsymbol{a}_1})\boldsymbol{x}^{\boldsymbol{c}_1} + (\boldsymbol{x}^{\boldsymbol{b}})^2,\\ &\boldsymbol{x}^{\boldsymbol{a}_2}\boldsymbol{x}^{\boldsymbol{c}_2} = (\boldsymbol{x}^{\boldsymbol{a}_2} - \boldsymbol{x}^{\boldsymbol{a}_1})\boldsymbol{x}^{\boldsymbol{c}_2} + \boldsymbol{x}^{\boldsymbol{a}_1}\boldsymbol{x}^{\boldsymbol{c}_2} = (\boldsymbol{x}^{\boldsymbol{a}_2} - \boldsymbol{x}^{\boldsymbol{a}_1})\boldsymbol{x}^{\boldsymbol{c}_2} + \boldsymbol{x}^{\boldsymbol{b}}\boldsymbol{x}^{\boldsymbol{c}_1}. \end{aligned}$$

Using Theorem 3.2 we conclude that $\boldsymbol{b}$ is an Ulrich element in H, and hence $K[H]$ is AG.

In the following special case, the possible Ulrich elements can be identified.

Proposition 3.6. *Let H be a semigroup in $\mathcal{H}_d$ whose nonzero elements have all the entries positive, and assume that $(1, \dots, 1) \in \omega_H$. If H has an Ulrich element $\boldsymbol{b}$, then $\boldsymbol{b} = (1, \dots, 1)$.*

Proof. We set $\boldsymbol{b}' = (1, \dots, 1)$. Assume, on the contrary, that $\boldsymbol{b} \neq \boldsymbol{b}'$. Then by the criterion in Theorem 3.2 and using the same notation, we get that $\boldsymbol{x}^{\boldsymbol{b}'} \cdot \boldsymbol{x}^{\boldsymbol{b}'} \in (\boldsymbol{x}^{\boldsymbol{b}}R, J\omega_R)$. This implies that $(2, \dots, 2) = 2\boldsymbol{b}' = \boldsymbol{b} + \boldsymbol{c}$ for some $\boldsymbol{c} \in H$, or that $2\boldsymbol{b}' = \boldsymbol{a}_i + \boldsymbol{h}$ for some $1 \le i \le d$ and $\boldsymbol{h} \in \omega_H$, $\boldsymbol{h} \neq \boldsymbol{b}'$. Since $(1, \dots, 1)$ is the smallest element of ω_H when comparing vectors componentwise, at least one component of $\boldsymbol{b}$ (respectively, of $\boldsymbol{h}$) is at least two; hence at least one component of $\boldsymbol{c}$ (respectively, of $\boldsymbol{a}_i$) is less than or equal to zero, which is false by the assumption that all the entries of nonzero elements of H are positive. □

4. The AG property for normal semigroups in dimension 2

As mentioned before, all 2-dimensional normal affine semigroups are slim. For them, in Theorem 4.3 we make more concrete the criterion for Ulrich elements given in Theorem 3.2. We first prove a couple of lemmas.

Throughout this section, unless otherwise stated, H is a semigroup in $\mathcal{N}_2 = \mathcal{H}_2$ with extremal rays $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. We denote by $\mathfrak{m}$ the graded maximal ideal of $R = K[H]$.

Lemma 4.1. *Let $\boldsymbol{b}$ be an element in $B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2\}$. For any $\boldsymbol{c} \in \omega_H$ such that $\boldsymbol{c} \notin \boldsymbol{b} + H$, there exists $t \in \{1, 2\}$ such that $\boldsymbol{c} + \boldsymbol{a}_t \in \boldsymbol{b} + H$.*

Proof. Let C be the cone with the extremal rays $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. Let $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ be vectors normal to the facets of the cone C such that $\langle \boldsymbol{a}_i, \boldsymbol{n}_i \rangle = 0$ for $i = 1, 2$ and $\boldsymbol{x} \in C$ if and only if $\langle \boldsymbol{x}, \boldsymbol{n}_1 \rangle \ge 0$ and $\langle \boldsymbol{x}, \boldsymbol{n}_2 \rangle \ge 0$.

Since $\boldsymbol{c} \notin \boldsymbol{b} + H$ it follows that $\boldsymbol{c} - \boldsymbol{b} \notin C$. We may assume that $\langle \boldsymbol{c} - \boldsymbol{b}, \boldsymbol{n}_1 \rangle < 0$, and claim then that $\boldsymbol{c} + \boldsymbol{a}_2 \in \boldsymbol{b} + H$. Indeed,

$$\langle \boldsymbol{c} + \boldsymbol{a}_2 - \boldsymbol{b}, \boldsymbol{n}_1 \rangle = \langle \boldsymbol{c}, \boldsymbol{n}_1 \rangle + \langle \boldsymbol{a}_1 + \boldsymbol{a}_2 - \boldsymbol{b}, \boldsymbol{n}_1 \rangle > 0,$$

since $\langle \boldsymbol{a}_1 + \boldsymbol{a}_2 - \boldsymbol{b}, \boldsymbol{n}_1 \rangle > 0$, by Lemma 2.1, and

$$\langle \boldsymbol{c} + \boldsymbol{a}_2 - \boldsymbol{b}, \boldsymbol{n}_2 \rangle = \langle \boldsymbol{c} - \boldsymbol{b}, \boldsymbol{n}_2 \rangle > 0,$$

since otherwise $\boldsymbol{c} - \boldsymbol{b} \in -C = \{-\boldsymbol{a} : \boldsymbol{a} \in C\}$, a contradiction to $\boldsymbol{b} \in B_H$. □

Lemma 4.2. *Let $\boldsymbol{b}$ belong to $B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2\}$. We set $I = (\boldsymbol{x}^{\boldsymbol{b}} R, (\boldsymbol{x}^{\boldsymbol{a}_1} - \boldsymbol{x}^{\boldsymbol{a}_2})\omega_R) \subset R$. Let $\boldsymbol{c} \in \omega_H$. The following conditions are equivalent*:

(a) $\boldsymbol{x}^{\boldsymbol{c}} \in I$.

(b) $\boldsymbol{c} \in (\boldsymbol{b} + H) \cup (\boldsymbol{a}_1 + \omega_H) \cup (\boldsymbol{a}_2 + \omega_H)$.

(c) $\boldsymbol{c} \in (\boldsymbol{b} + H) \cup (\boldsymbol{a}_1 + H) \cup (\boldsymbol{a}_2 + H)$.

Proof. (a) ⇒ (b): Note that $\boldsymbol{x}^{\boldsymbol{c}} \in \boldsymbol{x}^{\boldsymbol{b}} R$ if and only if $\boldsymbol{c} \in \boldsymbol{b} + H$. If $\boldsymbol{x}^{\boldsymbol{c}} \notin \boldsymbol{x}^{\boldsymbol{b}} R$, then there exist $0 \neq f$ in ω_R and g in R such that

$$\boldsymbol{x}^{\boldsymbol{c}} = (\boldsymbol{x}^{\boldsymbol{a}_1} - \boldsymbol{x}^{\boldsymbol{a}_2}) \cdot f + \boldsymbol{x}^{\boldsymbol{b}} \cdot g.$$

Therefore, there exists a monomial $\boldsymbol{x}^{\boldsymbol{a}}$ in ω_R such that $\boldsymbol{x}^{\boldsymbol{c}} = \boldsymbol{x}^{\boldsymbol{a}_1} \cdot \boldsymbol{x}^{\boldsymbol{a}}$ or $\boldsymbol{x}^{\boldsymbol{c}} = \boldsymbol{x}^{\boldsymbol{a}_2} \cdot \boldsymbol{x}^{\boldsymbol{a}}$, equivalently $\boldsymbol{c} \in (\boldsymbol{a}_1 + \omega_H) \cup (\boldsymbol{a}_2 + \omega_H)$.

(b) ⇒ (a): If $\boldsymbol{c} \in \boldsymbol{b} + H$ then clearly $\boldsymbol{x}^{\boldsymbol{c}} \in I$. Assume $\boldsymbol{c} \notin \boldsymbol{b} + H$. By symmetry, it is enough to consider the case $\boldsymbol{c} \in \boldsymbol{a}_1 + \omega_H$.

By Lemma 4.1, since $0 \neq \boldsymbol{c} - \boldsymbol{a}_1 \in \omega_H$, $\boldsymbol{c} - \boldsymbol{a}_1 \notin \boldsymbol{b} + H$ and $(\boldsymbol{c} - \boldsymbol{a}_1) + \boldsymbol{a}_1 = \boldsymbol{c} \notin \boldsymbol{b} + H$ it follows that $\boldsymbol{c} - \boldsymbol{a}_1 + \boldsymbol{a}_2 \in \boldsymbol{b} + H$.

As we may write

$$\boldsymbol{x}^{\boldsymbol{c}} = \boldsymbol{x}^{\boldsymbol{c} - \boldsymbol{a}_1} \cdot (\boldsymbol{x}^{\boldsymbol{a}_1} - \boldsymbol{x}^{\boldsymbol{a}_2}) + \boldsymbol{x}^{\boldsymbol{c} - \boldsymbol{a}_1 + \boldsymbol{a}_2},$$

we conclude that $\boldsymbol{x}^{\boldsymbol{c}} \in I$.

(b) ⇒ (c) is trivial.

For (c) ⇒ (b) it is enough to consider the case when $\boldsymbol{c} \notin \boldsymbol{b} + H$. We may assume $\boldsymbol{c} \in \boldsymbol{a}_1 + H$. If $\boldsymbol{c} \notin \boldsymbol{a}_1 + \omega_H$, then there exists a positive integer n such that either $\boldsymbol{c} - \boldsymbol{a}_1 = n\boldsymbol{a}_1$, or $\boldsymbol{c} - \boldsymbol{a}_1 = n\boldsymbol{a}_2$. In the first case we get that $\boldsymbol{c} = (n+1)\boldsymbol{a}_1 \notin \omega_H$, a contradiction. In the second case we get that $\boldsymbol{c} = \boldsymbol{a}_1 + n\boldsymbol{a}_2 \in \boldsymbol{b} + H$, since $\boldsymbol{a}_1 + \boldsymbol{a}_2 - \boldsymbol{b} \in H$ by Lemma 2.1. This is again a contradiction. Thus $\boldsymbol{c} \in \boldsymbol{a}_1 + \omega_H$. □

Theorem 4.3. *An element $\boldsymbol{b} \in B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2\}$ is an Ulrich element in H, if and only if*

$$\boldsymbol{c}_1 + \boldsymbol{c}_2 \in (\boldsymbol{b} + H) \cup (\boldsymbol{a}_1 + H) \cup (\boldsymbol{a}_2 + H) \quad \textit{for all } \boldsymbol{c}_1, \boldsymbol{c}_2 \in B_H.$$

Proof. Let $B_H = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{c}_1, \ldots, \boldsymbol{c}_m\}$. Then $\mathfrak{m} = (\boldsymbol{x}^{\boldsymbol{a}_1}, \boldsymbol{x}^{\boldsymbol{a}_2}, \boldsymbol{x}^{\boldsymbol{c}_1}, \ldots, \boldsymbol{x}^{\boldsymbol{c}_m})$ and $\omega_R = (\boldsymbol{x}^{\boldsymbol{c}_1}, \ldots, \boldsymbol{x}^{\boldsymbol{c}_m})$.

If $\boldsymbol{b}$ is an Ulrich element, then $\mathfrak{m}\omega_R \subseteq (\boldsymbol{x}^{\boldsymbol{b}} R, (\boldsymbol{x}^{\boldsymbol{a}_1} - \boldsymbol{x}^{\boldsymbol{a}_2})\omega_R)$, and therefore $\boldsymbol{x}^{\boldsymbol{c}_i} \boldsymbol{x}^{\boldsymbol{c}_j} \in (\boldsymbol{x}^{\boldsymbol{b}} R, (\boldsymbol{x}^{\boldsymbol{a}_1} - \boldsymbol{x}^{\boldsymbol{a}_2})\omega_R)$ for all i, j. Thus the desired conclusion follows from Lemma 4.2.

Conversely, let $\boldsymbol{x}^{\boldsymbol{c}} \in \mathfrak{m}\omega_R$. Then $\boldsymbol{c} = \boldsymbol{c}_i + \boldsymbol{c}_j + h$, or $\boldsymbol{c} = \boldsymbol{a}_i + \boldsymbol{c}_j + h$ for some $h \in H$. In both cases our assumptions imply that $\boldsymbol{c} \in (\boldsymbol{b}+H) \cup (\boldsymbol{a}_1 + H) \cup (\boldsymbol{a}_2 + H)$. Thus $\boldsymbol{x}^{\boldsymbol{c}} \in (\boldsymbol{x}^{\boldsymbol{b}} R, (\boldsymbol{x}^{\boldsymbol{a}_1} - \boldsymbol{x}^{\boldsymbol{a}_2})\omega_R)$, by Lemma 4.2. This shows that $\boldsymbol{b}$ is an Ulrich element in H. □

Example 4.4. Let H be the semigroup in $\mathcal{H}_2$ with the extremal rays $\boldsymbol{a}_1 = (5, 2)$ and $\boldsymbol{a}_2 = (2, 5)$. Then the Hilbert basis of H is

$$B_H = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{c}_1 = (1, 1), \boldsymbol{c}_2 = (2, 1), \boldsymbol{c}_3 = (1, 2)\}.$$

Using Theorem 4.3, we may check that none of $\boldsymbol{c}_1$, $\boldsymbol{c}_2$ or $\boldsymbol{c}_3$ is an Ulrich element in H. The same conclusion could be reached by using Proposition 3.6 together with Theorem 6.3.

Here is one immediate application of Theorem 4.3.

Proposition 4.5. *Let $H \in \mathcal{H}_2$ such that $\boldsymbol{c} + \boldsymbol{c}' \notin P_H$ for all $\boldsymbol{c}, \boldsymbol{c}' \in B_H \cap P_H$. Then any $\boldsymbol{b} \in B_H \cap P_H$ is an Ulrich element in H.*

Proof. By the hypothesis, if $\boldsymbol{c}, \boldsymbol{c}' \in B_H \cap P_H$ then $\boldsymbol{c} + \boldsymbol{c}' \in (\boldsymbol{a}_1 + H) \cup (\boldsymbol{a}_2 + H)$. Theorem 4.3 yields the conclusion. □

One may check that the semigroup H in Example 3.5 satisfies the hypothesis of Proposition 4.5; hence H admits three Ulrich elements.

In the following example there is exactly one Ulrich element in the Hilbert basis of H.

Example 4.6. For the semigroup $H \in \mathcal{H}_2$ with $\boldsymbol{a}_1 = (11, 13)$ and $\boldsymbol{a}_2 = (3, 4)$, a computation with [Normaliz] shows that $B_H = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{c}_1 = (4, 5), \boldsymbol{c}_2 = (5, 6)\}$. We note that the points $2\boldsymbol{c}_2 - \boldsymbol{c}_1 = (6, 7)$ and $2\boldsymbol{c}_2 - \boldsymbol{a}_2 = (7, 8)$ are not in ω_H since the slope of the line through the origin and each of these respective points is not in the interval $\left(\frac{13}{11}, \frac{4}{3}\right)$. Also, clearly, $2\boldsymbol{c}_2 - \boldsymbol{a}_1 = (-1, -1) \notin H$. Therefore, by Theorem 4.3 we get that $\boldsymbol{c}_1$ is not an Ulrich element in H.

On the other hand, since $2\boldsymbol{c}_1 = (8, 10) = \boldsymbol{c}_2 + \boldsymbol{a}_2$ and $B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2\} = \{\boldsymbol{c}_1, \boldsymbol{c}_2\}$, by Theorem 4.3 we conclude that $\boldsymbol{c}_2$ is an Ulrich element in H.

5. Bottom elements as Ulrich elements in dimension 2

In the multigraded situation which we consider in Definition 3.1, there is in general no distinguished multidegree with $(\omega_{K[H]})_{\boldsymbol{b}} \neq 0$. Inspired by Proposition 3.6, we are prompted to test the Ulrich property for elements in ω_H with smallest entries. First we present the following lemma.

Lemma 5.1. *For any $H \in \mathcal{H}_2$, the set ω_H has a unique minimal element with respect to the componentwise partial ordering.*

Proof. Let C be the cone of H, and let $\boldsymbol{a} = (a_1, a_2)$ and $\boldsymbol{b} = (b_1, b_2)$ be points in the relative interior of C. We claim that $\boldsymbol{a} \wedge \boldsymbol{b} = (\min\{a_1, b_1\}, \min\{a_2, b_2\}) \in \operatorname{relint} C$. This will imply the existence of the unique minimal element of ω_H.

For the proof of the claim, it is enough to consider the case when $a_1 < b_1$ and $a_2 > b_2$. Since $a_2/a_1 > b_2/a_1 > b_2/b_1$, it follows that the point in the plane with coordinates $\boldsymbol{a} \wedge \boldsymbol{b} = (a_1, b_2)$ lies inside the cone with vertex the origin and passing through the points with coordinates $\boldsymbol{a}$ and $\boldsymbol{b}$. Since the latter cone is in $\operatorname{relint} C$ it follows that $\boldsymbol{a} \wedge \boldsymbol{b} \in \operatorname{relint} C$. □

We call the unique minimal element of ω_H with respect to the componentwise partial ordering, *the bottom element* of H.

Remark 5.2. For $H \in \mathcal{H}_2$, since the elements in ω_H have only nonnegative entries, it follows that the bottom element of H is also the smallest element in $G(\omega_H)$ with respect to the componentwise order. Moreover, if $K[H]$ is not a regular ring then the bottom element of H is componentwise the smallest element in $B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2\}$; see Lemma 1.1.

In arbitrary embedding dimension we give the following definition.

Definition 5.3. For $H \in \mathcal{H}_d$, an element $\boldsymbol{b} \in \omega_H$ is called the *bottom element* of H if $\boldsymbol{c} - \boldsymbol{b} \in \mathbb{N}^d$ for all $\boldsymbol{c} \in \omega_H$.

Remark 5.4. In general, a semigroup $H \in \mathcal{H}_d$ with $d > 2$ may not have a unique minimal element in ω_H with respect to the componentwise partial ordering $\preceq$. For instance, let $d = 3$, $\boldsymbol{a}_1 = (5, 3, 1)$, $\boldsymbol{a}_2 = (1, 5, 2)$, $\boldsymbol{a}_3 = (8, 3, 5)$. Then, a calculation with [Normaliz] shows that

$$B_H = \big\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, (1, 2, 1), (2, 1, 1), (2, 2, 1), (2, 5, 2), (3, 2, 1), (3, 2, 2), \\ (3, 5, 2), (3, 5, 3), (4, 5, 2), (5, 2, 3), (5, 5, 2), (5, 5, 4), (7, 5, 5)\big\}.$$

One can check that the vectors $\boldsymbol{n}_1 = (19, 11, -37)$, $\boldsymbol{n}_2 = (-12, 17, 9)$, $\boldsymbol{n}_3 = (1, -9, 22)$ are normal to the planes generated by $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$, by $\boldsymbol{a}_1$ and $\boldsymbol{a}_3$, by $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ respectively. Also, that no element in $B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$ lies on any of the three planes just mentioned. Consequently, there are no inner lattice points on the faces of $\overline{P_H}$ and $G(\omega_H) = B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$. It follows that $\boldsymbol{b}_1 = (1, 2, 1)$ and $\boldsymbol{b}_2 = (2, 1, 1)$ are both minimal elements in ω_H with respect to $\preceq$.

Using Theorem 4.3 we show that sometimes the bottom element may be the only Ulrich element in B_H.

Proposition 5.5. *Let $\boldsymbol{b}$ be the bottom element of $H \in \mathcal{H}_2$. If $2\boldsymbol{b} \in P_H$, then $\boldsymbol{b}$ is the only possible Ulrich element in B_H.*

Proof. Assume $\boldsymbol{b}' \in B_H$ is an Ulrich element in H. Then $2\boldsymbol{b} \in (\boldsymbol{a}_1 + H) \cup (\boldsymbol{a}_2 + H) \cup (\boldsymbol{b}' + H)$, by Theorem 4.3. Since $2\boldsymbol{b} \in P_H$, we get that $2\boldsymbol{b} \in \boldsymbol{b}' + H$; hence $2\boldsymbol{b} = \boldsymbol{b}' + \boldsymbol{h}$

for some $\boldsymbol{h} \in P_H$. Moreover, comparing componentwise, $\boldsymbol{b} \preceq \boldsymbol{b}'$ and $\boldsymbol{b} \preceq \boldsymbol{h}$ since $\boldsymbol{b}$ is the bottom element for H; thus $\boldsymbol{b}' = \boldsymbol{b}$. □

Remark 5.6. In general, as Example 4.6 shows, even when the Hilbert basis of H contains a unique Ulrich element, the latter need not be the bottom element.

In the following, we discuss when the bottom element $\boldsymbol{b}$ of $H \in \mathcal{H}_2$ is Ulrich.

Notation 5.7. To avoid repetitions, in the rest of the section $H \in \mathcal{H}_2$ has the extremal rays $\boldsymbol{a}_1 = (x_1, y_1)$ and $\boldsymbol{a}_2 = (x_2, y_2)$ such that ($y_1/x_1 < y_2/x_2$ when $x_1, x_2 > 0$) or $x_2 = 0$.

We define H_1 and H_2 to be the semigroups in $\mathcal{H}_2$ with the extremal rays $\boldsymbol{a}_1$ and $\boldsymbol{b}$, respectively $\boldsymbol{a}_2$ and $\boldsymbol{b}$. We denote $\mathbb{Z}^2 \cap \operatorname{relint} P_{H_i}$ by H_i^*, for $i = 1, 2$.

By an easy argument, the following proposition presents a class of semigroups in $\mathcal{H}_2$ with Ulrich bottom element.

Proposition 5.8. *Let $\boldsymbol{b}$ be the bottom element of H. If $(x_2, y_1) \preceq \boldsymbol{b}$, then $\boldsymbol{b}$ is an Ulrich element in H.*

Proof. If $K[H]$ is a regular ring, $\boldsymbol{b}$ is an Ulrich element in H, since $G(\omega_H) = \{\boldsymbol{b}\}$.

Assume $K[H]$ is not a regular ring. Then $\boldsymbol{b} \in G(\omega_H) = B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2\}$, by Lemma 1.1(d). Let $\boldsymbol{c}_1, \boldsymbol{c}_2 \in B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2\}$, and $\boldsymbol{c}_1 + \boldsymbol{c}_2 = (c, d)$, $\boldsymbol{b} = (u, v)$.

If $(c, d) \in H_1$, then $(c, d) = r_1(x_1, y_1) + r_2(u, v)$ for some $r_1, r_2 \in \mathbb{R}_{\geq 0}$. Since $d \geq 2v \geq y_1 + v$, we have $r_1 \geq 1$ or $r_2 \geq 1$. Consequently, $(c, d) \in (\boldsymbol{b} + H_1) \cup (\boldsymbol{a}_1 + H_1) \subset (\boldsymbol{b} + H) \cup (\boldsymbol{a}_1 + H)$.

A similar argument shows that if $(c, d) \in H_2$, then $(c, d) \in (\boldsymbol{b} + H) \cup (\boldsymbol{a}_2 + H)$. The conclusion follows by Theorem 4.3. □

Example 5.9. Let H be the semigroup with extremal rays $\boldsymbol{a}_1 = (a, 1)$ and $\boldsymbol{a}_2 = (1, b)$, where $a, b \geq 2$. Since $1/a < 1 < b$ we get that $\boldsymbol{b} = (1, 1)$ is in ω_H and it is the bottom element in H. Then Proposition 5.8 implies that $\boldsymbol{b}$ is an Ulrich element in H.

Clearly, $H = H_1 \cup H_2$ and $H_1 \cap H_2 = \mathbb{N}\boldsymbol{b}$. The following lemma states some nice properties regarding H_1 and H_2.

Lemma 5.10. *Let $\boldsymbol{b}$ be the bottom element of H. Then*

(a) $\boldsymbol{p} + \boldsymbol{q} \in \boldsymbol{b} + H$ *for all* $\boldsymbol{p} \in H_1 \setminus \{0\}$ *and* $\boldsymbol{q} \in H_2 \setminus \{0\}$,

(b) $(\boldsymbol{b} + H) \cup (\boldsymbol{a}_1 + H) \cup (\boldsymbol{a}_2 + H) = H \setminus (H_1^* \cup H_2^*)$.

Proof. (a) If $\boldsymbol{p} - \boldsymbol{b} \in H$ or $\boldsymbol{q} - \boldsymbol{b} \in H$, then clearly $\boldsymbol{p} + \boldsymbol{q} \in \boldsymbol{b} + H$. Let us assume that $\boldsymbol{p} - \boldsymbol{b} \notin H$ and $\boldsymbol{q} - \boldsymbol{b} \notin H$. Let C' be the cone generated by the extremal rays $\boldsymbol{p}, \boldsymbol{q}$. Since $\boldsymbol{b} \in C'$, we have $\boldsymbol{b} = r_1\boldsymbol{p} + r_2\boldsymbol{q}$ for some $r_1, r_2 \in \mathbb{R}_{>0}$. If $r_1 > 1$, then $\boldsymbol{b} - \boldsymbol{p} = (r_1 - 1)\boldsymbol{p} + r_2\boldsymbol{q}$; hence $\boldsymbol{b} - \boldsymbol{p} \in C' \cap \omega_H$, a contradiction with $\boldsymbol{b}$ being the bottom element in H. Therefore, $r_1 \leq 1$, and also $r_2 \leq 1$ by a similar argument. Now, $\boldsymbol{p} + \boldsymbol{q} - \boldsymbol{b} = (1 - r_1)\boldsymbol{p} + (1 - r_2)\boldsymbol{q} \in C' \cap \mathbb{Z}^2 \subset H$.

(b) Note that for any $\boldsymbol{p} \in H$ we have

$$\boldsymbol{p} \in H_1 \setminus H_1^* \Leftrightarrow \boldsymbol{p} \in (\boldsymbol{b} + H_1) \cup (\boldsymbol{a}_1 + H_1),$$
$$\boldsymbol{p} \in H_2 \setminus H_2^* \Leftrightarrow \boldsymbol{p} \in (\boldsymbol{b} + H_2) \cup (\boldsymbol{a}_2 + H_2).$$

Therefore, $H \setminus (H_1^* \cup H_2^*) = (H_1 \cup H_2) \setminus (H_1^* \cup H_2^*) \subseteq (\boldsymbol{b}+H) \cup (\boldsymbol{a}_1+H) \cup (\boldsymbol{a}_2+H)$.

In order to check the reverse inclusion, let $\boldsymbol{p} \in (\boldsymbol{b} + H) \cup (\boldsymbol{a}_1 + H) \cup (\boldsymbol{a}_2 + H)$.

We first consider the case $\boldsymbol{p} \in H_1$. Then clearly, $\boldsymbol{p} \notin H_2^*$. If we assume, on the contrary, that $\boldsymbol{p} \in H_1^*$, then $\boldsymbol{p} = r_1\boldsymbol{a}_1 + r_2\boldsymbol{b}$ with $r_1, r_2 \in (0, 1)$. We decompose $\boldsymbol{b} = \alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2$ with $\alpha_1, \alpha_2 \in (0, 1]$. This gives $\boldsymbol{p} = (r_1 + r_2\alpha_1)\boldsymbol{a}_1 + r_2\alpha_2\boldsymbol{a}_2$. Since $r_2\alpha_2 < 1$ and $r_2\alpha_2 < \alpha_2$ we infer that $\boldsymbol{p} \notin (\boldsymbol{a}_2 + H) \cup (\boldsymbol{b} + H)$. Thus $\boldsymbol{p} \in \boldsymbol{a}_1 + H$ and $r_1 \geq 1$, a contradiction. Consequently, $\boldsymbol{p} \notin H_1^* \cup H_2^*$.

A similar argument works for the case $\boldsymbol{p} \in H_2$. □

Lemma 5.11. *Let $\boldsymbol{p} = (k, r) \in H_1^*$ and $\boldsymbol{q} = (\ell, s) \in H_2^*$. If $\boldsymbol{b} = (u, v)$ is the bottom element of H, then*

(a) *$u < k < x_1$ and $v \leq r \leq y_1$,*

(b) *$u \leq \ell \leq x_2$ and $v < s < y_2$.*

Proof. We only show (a), part (b) is proved similarly. Clearly, $\boldsymbol{b} \neq \boldsymbol{p} \in \omega_H$; thus $0 < u \leq k$ and $0 < v \leq r$. If $u = k$, then since $\boldsymbol{p} \neq \boldsymbol{b}$, we have $v < r$. Then $r/k > v/u > y_1/x_1$, which gives that $\boldsymbol{p} \notin H_1$, which is false. Thus $u < k$.

On the other hand, by Lemma 2.1 applied in $H_1 \in \mathcal{H}_2$ for $\boldsymbol{p}$, the point

$$(u, v) + (x_1, y_1) - (k, r) = (u + x_1 - k, v + y_1 - r) \in H_1^*,$$

hence $u < u + x_1 - k$ and $v \leq v + y_1 - r$. That gives $k < x_1$ and $r \leq y_1$. □

The following result restricts the verification of the bottom element being Ulrich to verifying that the sum of any two points in H_i^* is not in H_i^*, for $i = 1, 2$.

Lemma 5.12. *Assume $\boldsymbol{b}$ is the bottom element of H. The following conditions are equivalent*:

(a) *$\boldsymbol{b}$ is an Ulrich element in H.*

(b) *For $i = 1, 2$, if $\boldsymbol{p}, \boldsymbol{q} \in H_i^*$ then $\boldsymbol{p} + \boldsymbol{q} \notin H_i^*$.*

Proof. We know that $\boldsymbol{b} \in G(\omega_H)$ since it is the bottom element in H. If $K[H]$ is a regular ring, then $\boldsymbol{b} = \boldsymbol{a}_1 + \boldsymbol{a}_2$. Hence statement (a) holds by Example 3.4, and (b) is true since $H_1^* = H_2^* = \varnothing$.

Assume that $K[H]$ is not a regular ring, hence $\boldsymbol{b} \in B_H \setminus \{\boldsymbol{a}_1, \boldsymbol{a}_2\}$. According to Theorem 4.3, $\boldsymbol{b}$ is an Ulrich element in H if and only if for all $\boldsymbol{p}, \boldsymbol{q} \in B_H$ one has

$$\boldsymbol{p} + \boldsymbol{q} \in (\boldsymbol{b} + H) \cup (\boldsymbol{a}_1 + H) \cup (\boldsymbol{a}_2 + H). \tag{9}$$

It is of course equivalent to check (9) for all $\boldsymbol{p}$ and $\boldsymbol{q}$ nonzero in H.

If ($\boldsymbol{p} \in H_1$ and $\boldsymbol{q} \in H_2$) or ($\boldsymbol{p} \in H_2$ and $\boldsymbol{q} \in H_1$) then $\boldsymbol{p}+\boldsymbol{q} \in \boldsymbol{b}+H$, by Lemma 5.10. Thus, for (a) it suffices to check (9) for nonzero $\boldsymbol{p}, \boldsymbol{q}$ both in H_1 or both in H_2. For $i = 1, 2$, the semigroup H_i is normal and simplicial; hence any $\boldsymbol{p} \in H_i$ is of the form $\boldsymbol{p} = n_1\boldsymbol{b} + n_2\boldsymbol{a}_i + \boldsymbol{p}'$ with $n_1, n_2 \in \mathbb{N}$ and $\boldsymbol{p}' \in H_i^* \cup \{0\}$. Consequently, $\boldsymbol{b}$ is an Ulrich element in H if and only if property (b) holds. □

Definition 5.13. We say that H *is AG1* if condition (b) in Lemma 5.12 is satisfied for $i = 1$, and we call it *AG2* if the said condition is satisfied for $i = 2$.

Thus the bottom element is an Ulrich element in H if and only if H is AG1 and AG2.

Remark 5.14. Using Lemma 5.11, property AG1 means that for any $\boldsymbol{p} = (k, r)$ and $\boldsymbol{q} = (\ell, s) \in H_1^*$ such that $k+\ell < x_1$ and $r+s \le y_1$ one has that $\boldsymbol{p}+\boldsymbol{q} \notin H_1^*$.

Similarly, the AG2 condition means that when $\boldsymbol{p} = (k, r)$ and $\boldsymbol{q} = (\ell, s) \in H_2^*$ such that $k+\ell \le x_2$ and $r+s < y_2$, then $\boldsymbol{p}+\boldsymbol{q} \notin H_2^*$.

Remark 5.15. Assume $\boldsymbol{b} = (u, v)$ is the bottom element in H. If $y_1 = 0$ then $v = 1$, since otherwise the inequalities $v/u > (v-1)/u > y_1/x_1 = 0$ would give that $(u, v-1) \in \omega_{H_1}$, a contradiction to the fact that (u, v) is the bottom element in H. Then, by Lemma 5.11 we get that $H_1^* = \varnothing$; hence H satisfies condition AG1.

Similarly, if $x_2 = 0$ then $u = 1$ and $H_2^* = \varnothing$; hence H is AG2.

In order to check the AG1 and AG2 conditions we need to have a better understanding of the points in H_1^* and H_2^*. We can count their elements.

Lemma 5.16. *Let $\boldsymbol{b} = (u, v)$ be the bottom element for H. Then*

(a) $|H_1^*| = vx_1 - uy_1 - 1$ *and* $|H_2^*| = uy_2 - vx_2 - 1$,

(b) $1 \le vx_1 - uy_1 \le x_1$ *and* $1 \le uy_2 - vx_2 - 1 \le y_2$,

(c) *if $H_1^* \ne \varnothing$ then $vx_1 - uy_1 \le x_1 - u$,*

(d) *if $H_2^* \ne \varnothing$ then $uy_2 - vx_2 \le y_2 - v$.*

Proof. We only show the first parts of (a) and (b), since the second parts are proved similarly.

(a) The area of the parallelogram spanned by $\boldsymbol{b}$ and $\boldsymbol{a}_1$ equals $\det\begin{pmatrix} x_1 & u \\ y_1 & v \end{pmatrix} = vx_1 - uy_1$. Since the boundary of that parallelogram contains precisely four lattice points, the vertices, (here we use the fact that $\gcd(u, v) = \gcd(x_1, y_1) = 1$), Pick's theorem [Beck and Robins 2015, Theorem 2.8] implies that P_{H_1} has $vx_1 - uy_1 - 1$ inner lattice points, which proves the claim.

(b) The inequality $1 \le vx_1 - uy_1$ follows from (a). Since (u, v) is the bottom element of H, it follows that $(u, v-1)$ is not in ω_H and in relint P_{H_1}. As $(v-1)/u < v/u$, and $y_1/x_1 < v/u$ by our assumption, we get that $(v-1)/u \le y_1/x_1$, i.e., $vx_1 - uy_1 \le x_1$.

Parts (c) and (d) will be proved after Remark 5.19. □

One nice consequence of Lemma 5.16 is a Gorenstein criterion for $K[H]$ in terms of the coordinates of the bottom element in H.

Corollary 5.17. *If $\boldsymbol{b} = (u, v)$ is the bottom element in H, then the K-algebra $K[H]$ is Gorenstein if and only if $vx_1 - uy_1 = uy_2 - vx_2 = 1$.*

Proof. The ring $K[H]$ is Gorenstein if and only if ω_H is a principal ideal, and hence generated by $\boldsymbol{b}$, which is equivalent to saying that P_{H_1} and P_{H_2} have no inner points. By Lemma 5.16 this is the case if and only if $vx_1 - uy_1 = uy_2 - vx_2 = 1$. □

Lemma 5.18. *Let $\boldsymbol{b} = (u, v)$ be the bottom element of H. We assume that H_1^* is not the empty set. For any integer i we consider the integers q_i, r_i such that $iy_1 = q_i x_1 + r_i$ with $0 \le r_i < x_1$.*

Assume the integer k satisfies $u < k < x_1$. The following statements are equivalent:

(i) *k is the x-coordinate of some $\boldsymbol{p} \in H_1^*$.*

(ii) $\lceil ky_1/x_1 \rceil \le v + \lfloor (k-u)y_1/x_1 \rfloor$.

(iii) $\lceil ky_1/x_1 \rceil = v + \lfloor (k-u)y_1/x_1 \rfloor$.

(iv) $q_k \le v - 1 + q_{k-u}$.

(v) $q_k = v - 1 + q_{k-u}$.

(vi) $r_k \ge r_{k-u} + x_1 - (vx_1 - uy_1)$.

(vii) $r_k = r_{k-u} + x_1 - (vx_1 - uy_1)$.

(viii) $r_k \ge x_1 - (vx_1 - uy_1)$.

If any of these conditions holds, then $\boldsymbol{p} = (k, \lceil ky_1/x_1 \rceil) = (k, q_k + 1)$.

Proof. Since $H_1^* \neq \varnothing$ we have that $y_1 > 0$ and $u < x_1$ by Remark 5.15 and Lemma 5.11, respectively. We note that for any integer $u < k < x_1$, the fractions ky_1/x_1 and $v + (k-u)y_1/x_1$ are not integers. Thus $\lceil ky_1/x_1 \rceil = q_k + 1$ and $\lfloor v + (k-u)y_1/x_1 \rfloor = v + q_{k-u}$. This shows that (ii) $\Longleftrightarrow$ (iv) and (iii) $\Longleftrightarrow$ (v). Since $q_k = (ky_1 - r_k)/x_1$, simple manipulations give that (iv) $\Longleftrightarrow$ (vi) and (v) $\Longleftrightarrow$ (vii).

We also infer that the number of points in H_1^* whose x-coordinate is k equals the number of lattice points on the line $x = k$ located strictly between the lines $y = (y_1/x_1)x$ and $y = (y_1/x_1)(x-u) + v$, which is

$$\left\lfloor \frac{y_1}{x_1}(k-u) + v \right\rfloor - \left\lceil \frac{y_1}{x_1}k \right\rceil + 1 = v + q_k + \left\lfloor \frac{r_k - uy_1}{x_1} \right\rfloor - (q_k + 1) + 1 \tag{10}$$

$$= \left\lfloor \frac{vx_1 - uy_1 + r_k}{x_1} \right\rfloor \in \{0, 1\}. \tag{11}$$

The latter statement is due to the fact that $r_k < x_1$ and $vx_1 - uy_1 \le x_1$, by Lemma 5.16.

Consequently, $k \in (u, x_1)$ is the x-coordinate of some point in H_1^* if and only if the value in equation (10) is at least (and actually equal to) 1, which is equivalent to property (ii), respectively to (iii). That is, moreover, equivalent (using (11)) to

$$1 \leq \frac{vx_1 - uy_1 + r_k}{x_1},$$

which can be rewritten as $r_k \geq x_1 - (vx_1 - uy_1)$, namely statement (viii).

From (10) and (11) we obtain that if k is the x-coordinate of some point $\boldsymbol{p} \in H_1^*$, then $\boldsymbol{p} = (k, \lceil (y_1/x_1)k \rceil) = (k, q_k + 1)$. □

Remark 5.19. A similar result holds for the points in H_2^* in terms of the integers q_i', r_i' such that $ix_2 = q_i' y_2 + r_i'$ with $0 \leq r_i' < y_2$.

Now we can finish the proof of Lemma 5.16.

Proof of Lemma 5.16, continued. (c) By Lemma 5.18, for each $u < k < x_1$ there is at most one point in H_1^* whose x-coordinate is k; therefore $|H_1^*| \leq x_1 - u - 1$. Using point (a) we obtain the inequality at (c). Part (d) is proved similarly. □

It will be convenient to denote $\pi_1(H_1^*) = \{k : \text{ there exists } (k, \ell) \in H_1^*\}$. The next result is a criterion to verify the AG1 property in terms of the remainders r_i introduced in Lemma 5.18, with $i \in \pi_1(H_1^*)$. A similar statement characterizes the AG2 property in terms of the r_j''s from Remark 5.19, with $j \in \pi_2(H_2^*)$.

Proposition 5.20. *For any integer i let $r_i \equiv iy_1 \bmod x_1$ with $0 \leq r_i < x_1$. Then H is AG1 if and only if $r_k + r_\ell < 2x_1 - (vx_1 - uy_1)$ for all integers $k, \ell \in \pi_1(H_1^*)$ with $k + \ell < x_1$.*

Proof. If $H_1^* = \varnothing$ then there is nothing to prove. Assume H_1^* is not empty. If $k, \ell \in \pi_1(H_1^*)$, then by Lemma 5.18, $\boldsymbol{p}_1 = (k, \lceil ky_1/x_1 \rceil)$ and $\boldsymbol{p}_2 = (\ell, \lceil \ell y_1/x_1 \rceil)$ are the corresponding points in H_1^*. By definition, H is AG1 if and only if $\boldsymbol{p}_1 + \boldsymbol{p}_2 \notin H_1^*$ for all $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ as above. When $k + \ell \geq x_1$, Lemma 5.11 implies already that $\boldsymbol{p}_1 + \boldsymbol{p}_2 \notin H_1^*$. If $k + \ell < x_1$, then $\boldsymbol{p}_1 + \boldsymbol{p}_2 \notin H_1^*$ if and only if

$$\left\lceil \frac{ky_1}{x_1} \right\rceil + \left\lceil \frac{\ell y_1}{x_1} \right\rceil \geq (k + \ell - u)\frac{y_1}{x_1} + v, \quad \text{equivalently} \tag{12}$$

$$\frac{ky_1 - r_k}{x_1} + 1 + \frac{\ell y_1 - r_\ell}{x_1} + 1 \geq (k + \ell - u)\frac{y_1}{x_1} + v,$$

$$ky_1 - r_k + \ell y_1 - r_\ell + 2x_1 \geq (k + \ell)y_1 - uy_1 + vx_1,$$

$$2x_1 - (vx_1 - uy_1) \geq r_k + r_\ell. \tag{13}$$

Since $u < k + \ell < x_1$, the term of the right-hand side of (12) is not an integer; hence the inequality at (12) (and equivalently, at (13)) can not become an equality. □

6. A criterion for (1, 1) to be an Ulrich element

Our aim in this section is to obtain a complete classification of when $\boldsymbol{b} = (1, 1)$ is an Ulrich element. The setup in Notation 5.7 is in use. The element $(1, 1)$ is in ω_H if and only if $y_1/x_1 < 1 < y_2/x_2$. If that is the case, it is clear that $(1, 1)$ is the bottom element in H. It suffices to verify the AG1 and AG2 conditions, by Lemma 5.12.

Set $n = x_1 - y_1 - 1$, which equals $|H_1^*|$, by Lemma 5.16. If $n = 0$, then H is clearly AG1.

We consider the case $n > 0$. The next result presents an explicit way to determine H_1^*. Recursively, we define nonnegative integers $\ell_1, \ldots, \ell_n$ and $s_1, \ldots, s_n$ by

$$x_1 = \ell_1(x_1 - y_1) + s_1 \quad \text{with} \quad s_1 < x_1 - y_1,$$
$$y_1 + s_{i-1} = \ell_i(x_1 - y_1) + s_i \quad \text{with} \quad s_i < x_1 - y_1,$$

for $i = 2, \ldots, n$.

Lemma 6.1. *Assume that* $(1, 1)$ *belongs to* ω_H *and* $H_1^* \neq \varnothing$. *Then*

$$H_1^* = \left\{ \boldsymbol{p}_t = (c_t, d_t) : c_t = t + \sum_{i=1}^{t} \ell_i,\ d_t = \sum_{i=1}^{t} \ell_i,\ t = 1, \ldots, n \right\},$$

Proof. For $k = 1, \ldots, x_1 - 1$, let $ky_1 = q_k x_1 + r_k$ with integers $q_k \geq 0$ and $x_1 > r_k \geq 0$.

By Lemma 5.18, the integer $k > 1$ is the x-coordinate of an element of H_1^* if and only if $q_k = q_{k-1}$. In this case, $(k, 1 + q_k) \in H_1^*$.

Now, let $t \geq 1$. Summing up the equations $x_1 = \ell_1(x_1 - y_1) + s_1$ and $y_1 + s_{i-1} = \ell_i(x_1 - y_1) + s_i$, for $i = 2, \ldots, t$, we get

$$x_1 + (t-1)y_1 + s_1 + s_2 + \cdots + s_{t-1} = \sum_{i=1}^{t} \ell_i(x_1 - y_1) + s_1 + s_2 + \cdots + s_t,$$

and consequently,

$$\left(t - 1 + \sum_{i=1}^{t} \ell_i\right) y_1 = \left(\sum_{i=1}^{t} \ell_i - 1\right) x_1 + s_t.$$

Then

$$\left(t + \sum_{i=1}^{t} \ell_i\right) y_1 = \left(\sum_{i=1}^{t} \ell_i - 1\right) x_1 + s_t + y_1,$$

with $s_t + y_1 < x_1$. Therefore, $q_k = q_{k-1} = \left(\sum_{i=1}^{t} \ell_i - 1\right)$ for $k = t + \sum_{i=1}^{t} \ell_i$.

Note that

$$\begin{aligned} n+\sum_{i=1}^{n}\ell_i &= n+\frac{x_1-s_1}{x_1-y_1}+\sum_{i=2}^{t}\frac{y_1-s_{i-1}+s_i}{x_1-y_1} \\ &= n+\frac{x_1+(n-1)y_1-s_n}{x_1-y_1} = n+1+\frac{ny_1-s_n}{x_1-y_1} \\ &< n+1+y_1 = x_1. \end{aligned}$$

Hence $\boldsymbol{p}_t = \left(t+\sum_{i=1}^{t}\ell_i, \sum_{i=1}^{t}\ell_i\right) \in H_1^*$ for $t = 1, \ldots, n$.

We know from Lemma 5.16, that H_1^* has exactly $n = x_1 - y_1 - 1$ elements, so $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$ are the only elements of H_1^*. □

Examples 6.2. Let $x_1 = \ell_1(x_1 - y_1) + s_1$ and $y_1 + s_{i-1} = \ell_i(x_1 - y_1) + s_i$, for $i = 2, \ldots, n = x_1 - y_1 - 1$ as before.

(a) If $y_1 = 1$, then $H_1^* = \{(m, 1) : m = 2, \ldots, x_1 - 1\}$. In this case, H is AG1 by Lemma 5.12.

(b) If $x_1 - y_1 \in \{1, 2\}$ then by Lemma 5.16, H_1^* is either empty, or it consists of one element, which is different from $(0, 0)$. Hence H is AG1.

(c) If $2 < 2y_1 < x_1 < 3y_1$, then $\ell_1 = \ell_2 = 1$. Therefore, $\boldsymbol{p}_1 = (2, 1)$ and $\boldsymbol{p}_2 = (4, 2) = 2\boldsymbol{p}_1$ belong to H_1^*. Then, by definition, H is not AG1.

Next, we give a simple arithmetic criterion to check the AG1 or AG2 property:

Theorem 6.3. *Assume that* $(1, 1)$ *belongs to* ω_H. *Assuming Notation 5.7, then*

(a) *H is AG1 if and only if* $x_1 \equiv 1 \bmod (x_1 - y_1)$;

(b) *H is AG2 if and only if* $y_2 \equiv 1 \bmod (y_2 - x_2)$;

(c) $(1, 1)$ *is an Ulrich element in H if and only if* $x_i \equiv 1 \bmod (x_i - y_i)$ *for* $i = 1, 2$.

Proof. (a) Let $n = x_1 - y_1 - 1 = |H_1^*|$. If $n \in \{0, 1\}$, then H is AG1 by Examples 6.2(b). On the other hand, if $n = 0$ then $x_1 - y_1 = 1$ and clearly, $x_1 \equiv 1 \bmod (x_1 - y_1)$. When $n = 1$ we have $x_1 - y_1 = 2$. Since $\gcd(x_1, y_1) = 1$ we get that x_1 is odd; hence $x_1 \equiv 1 \bmod (x_1 - y_1)$, too.

We further prove the stated equivalence when $n \geq 2$. Let $\ell_1, \ldots, \ell_n \geq 0$ and $x_1 - y_1 > s_1, \ldots, s_n \geq 0$ such that

$$x_1 = \ell_1(x_1 - y_1) + s_1, \quad y_1 + s_{i-1} = \ell_i(x_1 - y_1) + s_i$$

for $i = 2, \ldots, n$. Then

$$H_1^* = \left\{\boldsymbol{p}_t = (c_t, d_t) : c_t = t + \sum_{i=1}^{t}\ell_i,\ d_t = \sum_{i=1}^{t}\ell_i,\ t = 1, \ldots, n\right\},$$

by Lemma 6.1. We note that since $y_1 > 0$ (see Remark 5.15) we have $x_1 > x_1 - y_1$; hence $\ell_1 \geq 1$.

Assume that $x_1 \equiv 1 \bmod (x_1 - y_1)$. Then it is easy to check that $s_i = i$ and $\ell_i = \ell_1 - 1$ for $i = 2, \ldots, n$. Consequently,

$$H_1^* = \{(t\ell_1 + 1, t(\ell_1 - 1) + 1) : t = 1, \ldots, n\},$$

and therefore, the sum of any two elements of H_1^* is not in H_1^*, i.e., H is AG1.

Conversely, assume that H is AG1. As $n > 0$ we get that $x_1 - y_1 > 1$ and $y_1 > 0$. In case $y_1 = 1$, then clearly, $x_1 \equiv 1 \bmod (x_1 - y_1)$.

We consider the case $y_1 \geq 2$. As

$$1 = \gcd(x_1, y_1) = \gcd(x_1, x_1 - y_1) = \gcd(s_1, x_1 - y_1)$$

and $x_1 - y_1 > 1$, we have that $s_1 > 0$. We need to prove that $s_1 = 1$.

Assume, on the contrary, that $s_1 \neq 1$. Then $s_1 \geq 2$. Since

$$\begin{aligned}(\ell_1 - 1)(x_1 - y_1) + s_1 = y_1 \leq y_1 + s_{i-1} = \ell_i(x_1 - y_1) + s_i,\\ y_1 + s_{i-1} < y_1 + (x_1 - y_1) = x_1 = \ell_1(x_1 - y_1) + s_1,\end{aligned}$$

we have $\ell_1 - 1 \leq \ell_i \leq \ell_1$, for $i = 2, \ldots, n$.

If $\ell_2 = \ell_1$, then $\boldsymbol{p}_2 = (2 + 2\ell_1, 2\ell_1) = 2\boldsymbol{p}_1$, which contradicts the AG1 property. Now, we consider the case $\ell_2 = \ell_1 - 1$. By subtracting the equations

$$x_1 = \ell_1(x_1 - y_1) + s_1 \quad \text{and} \quad y_1 + s_1 = \ell_2(x_1 - y_1) + s_2,$$

we get that $s_2 = 2s_1$; hence $s_2 > s_1$.

If $\ell_2 = \cdots = \ell_n$ then $s_1 < s_2 < \cdots < s_n$ is an increasing sequence of n positive integers less than $n + 1$, and hence $s_1 = 1$, which is false. Thus $\ell_i = \ell_1$ for some $i \geq 3$. Let i be the smallest index with this property, i.e., $\ell_2 = \cdots = \ell_{i-1} = \ell_1 - 1$ and $\ell_i = \ell_1$. Then

$$\begin{aligned}\boldsymbol{p}_i &= \big(i + (i-2)(\ell_1 - 1) + 2\ell_1, (i-2)(\ell_1 - 1) + 2\ell_1\big)\\ &= (1 + \ell_1, \ell_1) + \big(i - 1 + (i-2)(\ell_1 - 1) + \ell_1, (i-2)(\ell_1 - 1) + \ell_1\big)\\ &= \boldsymbol{p}_1 + \boldsymbol{p}_{i-1},\end{aligned}$$

which is a contradiction. This shows that when H is AG1, then $x_1 \equiv 1 \bmod (x_1 - y_1)$.

For part (b) we let H' be the semigroup in $\mathcal{H}_2$ with the extremal rays $\boldsymbol{a}_1' = (y_2, x_2)$ and $\boldsymbol{a}_2' = (y_1, x_1)$. We remark that H is AG2 if and only if H' is AG1, and we use (a). Part (c) is a consequence of (a) and (b). □

Corollary 6.4. *Let H be a semigroup in $\mathcal{H}_2$ with extremal rays $\boldsymbol{a}_i = (x_i, y_i)$ for $i = 1, 2$. Assume $(1, 1) \in \omega_H$ and $x_1x_2y_1y_2 \neq 0$. Then $K[H]$ is AG if and only if $x_i \equiv 1 \bmod (x_i - y_i)$ for $i = 1, 2$.*

Proof. By Proposition 3.6, the only possible Ulrich element in H is $(1, 1)$. The conclusion follows by Theorem 6.3. □

Remark 6.5. In the statement of Corollary 6.4, the assumption $x_1x_2y_1y_2 \neq 0$ can not be dropped. For instance, let $H \in \mathcal{H}_2$ with the extremal rays $\boldsymbol{a}_1 = (1, 0)$ and $\boldsymbol{a}_2 = (2, 5)$. Its Hilbert basis is

$$B_H = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{c}_1 = (1, 1), \boldsymbol{c}_2 = (2, 3), \boldsymbol{c}_3 = (1, 2)\}.$$

The bottom element in H is $\boldsymbol{c}_1$, and by Theorem 6.3 it follows that H is not AG2.

Still, H is AG. Since $2\boldsymbol{c}_1 = (2, 2) = \boldsymbol{c}_3 + \boldsymbol{a}_1$, $2\boldsymbol{c}_2 = (4, 6) = \boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{c}_1$ and $\boldsymbol{c}_1 + \boldsymbol{c}_2 = (3, 4) = \boldsymbol{a}_1 + 2\boldsymbol{c}_3$, by Theorem 4.3 we get that $\boldsymbol{c}_3$ is an Ulrich element in H.

7. Nearly Gorenstein semigroup rings

In this section we prove the nearly Gorenstein property for semigroup rings $K[H]$ when $H \in \mathcal{H}_2$.

Nearly Gorenstein rings approximate Gorenstein rings in a different way as almost Gorenstein rings. In [Herzog et al. 2019], a local (or graded) Cohen–Macaulay ring which admits a canonical module ω_R is called *nearly Gorenstein* if the trace of ω_R contains the (graded) maximal ideal of R. In the case that R is a domain, the canonical module can be realized as an ideal of R and its trace in R, which we denote by $\operatorname{tr}(\omega_R)$, is the ideal $\sum_f f\omega_R$, where the sum is taken over all f in the quotient field of R for which $f\omega_R \subseteq R$; see [Herzog et al. 2019, Lemma 1.1].

A one-dimensional almost Gorenstein ring is nearly Gorenstein, but the converse does not hold in general. In higher dimension there is in general no implication valid between these two concepts; see [Herzog et al. 2019].

Theorem 7.1. *Let H be a simplicial affine semigroup in $\mathcal{H}_2$. Then $R = K[H]$ is a nearly Gorenstein ring.*

Proof. Let $\boldsymbol{a}_1 = (c, d)$ and $\boldsymbol{a}_2 = (e, f)$ be the extremal rays of H. We may assume that $d/c < f/e$ and that R is not already a Gorenstein ring.

The vector $\boldsymbol{n}_1 = (-d, c)$ is orthogonal to $\boldsymbol{a}_1$ and $\boldsymbol{n}_2 = (f, -e)$ is orthogonal to $\boldsymbol{a}_2$. Moreover, $\boldsymbol{c}$ is in C, the cone over H, if and only if $\langle \boldsymbol{n}_1, \boldsymbol{c}\rangle \geq 0$ and $\langle \boldsymbol{n}_2, \boldsymbol{c}\rangle \geq 0$.

Let $\boldsymbol{c}_1, \dots, \boldsymbol{c}_t, \boldsymbol{c}_{t+1}, \boldsymbol{c}_{t+2}$ be the Hilbert basis of H, where $\boldsymbol{c}_{t+i} = \boldsymbol{a}_i$ for $i = 1, 2$. Then ω_R is generated by $v_i = \boldsymbol{x}^{\boldsymbol{c}_i}$ for $i = 1, \dots, t$; see Lemma 1.1.

In order to prove that R is nearly Gorenstein, it suffices to show that for each element $\boldsymbol{c}_i$ of the Hilbert basis there exist $\boldsymbol{c} \in \mathbb{Z}^2$ and an integer $k \in \{1, \dots, t\}$ such that

(i) $\boldsymbol{c} + \boldsymbol{c}_j \in C$ for $j = 1, \dots, t$, and

(ii) $\boldsymbol{c} + \boldsymbol{c}_k = \boldsymbol{c}_i$.

If $i \in \{1, \dots, t\}$, we may choose $\boldsymbol{c} = 0$ and $k = i$. It suffices to consider the cases $i = t + 1$ and $i = t + 2$. By symmetry we may assume that $i = t + 1$, and have to find $\boldsymbol{c} \in \mathbb{Z}^2$ and $k \in \{1, \dots, t\}$ such that (i) is satisfied and such that $\boldsymbol{c} + \boldsymbol{c}_k = \boldsymbol{a}_1$.

Let $k \in \{1, \dots, t\}$ be chosen such that $\langle \boldsymbol{n}_1, \boldsymbol{c}_k \rangle = \min\{\langle \boldsymbol{n}_1, \boldsymbol{c}_j \rangle : j = 1, \dots, t\}$. Set $\boldsymbol{c} = \boldsymbol{a}_1 - \boldsymbol{c}_k$. Then $\boldsymbol{c} + \boldsymbol{c}_k = \boldsymbol{a}_1$. Moreover, by the choice of k for $j = 1, \dots, t$ we have

$$\langle \boldsymbol{n}_1, \boldsymbol{c} + \boldsymbol{c}_j \rangle = \langle \boldsymbol{n}_1, \boldsymbol{a}_1 \rangle - \langle \boldsymbol{n}_1, \boldsymbol{c}_k \rangle + \langle \boldsymbol{n}_1, \boldsymbol{c}_j \rangle = 0 - \langle \boldsymbol{n}_1, \boldsymbol{c}_k \rangle + \langle \boldsymbol{n}_1, \boldsymbol{c}_j \rangle \geq 0,$$

$$\langle \boldsymbol{n}_2, \boldsymbol{c} + \boldsymbol{c}_j \rangle = \langle \boldsymbol{n}_2, \boldsymbol{a}_1 \rangle - \langle \boldsymbol{n}_2, \boldsymbol{c}_k \rangle + \langle \boldsymbol{n}_2, \boldsymbol{c}_j \rangle.$$

Since $\boldsymbol{c}_j \in H$, we have $\langle \boldsymbol{n}_2, \boldsymbol{c}_j \rangle \geq 0$. Let L be the line passing through $\boldsymbol{c}_k$ which is parallel to $L_2 = \mathbb{R}\boldsymbol{a}_2$, and L' be the line passing through $\boldsymbol{a}_1$ parallel to L_2. Since $\boldsymbol{c}_k \in P_H$, the line L has a smaller distance to L_2 than the line L'. This implies that $\langle \boldsymbol{n}_2, \boldsymbol{a}_1 \rangle > \langle \boldsymbol{n}_2, \boldsymbol{c}_k \rangle$; hence $\langle \boldsymbol{n}_2, \boldsymbol{c} + \boldsymbol{c}_j \rangle > 0$. Thus we conclude that $\boldsymbol{c} + \boldsymbol{c}_j \in C$, as desired. ☐

Theorem 7.1 is no longer valid when $\dim K[H] > 2$, as the following example shows.

Example 7.2. We consider again the semigroup $H \in \mathcal{H}_3$ from Remark 5.4. It turns out that $K[H]$ is not nearly Gorenstein for this semigroup H. One can see that $\boldsymbol{a}_1$ does not satisfy the two conditions (i) and (ii) in the proof of Theorem 7.1.

In fact, if we consider the set A of all $\boldsymbol{a}_1 - \boldsymbol{c}_i$ for $i = 1, \dots, 13$, then the third component of elements in A belongs to $\{0, -1, -2, -3, -4\}$. Adding the elements with negative third component to $(1, 2, 1)$, we get a vector with third component less than 1, which does not belong to C, the cone over H. Adding those elements in A with zero third component to either $(2, 1, 1)$ or $(1, 2, 1)$, we again get a vector which does not belong to C.

Acknowledgement

We gratefully acknowledge the use of the [Singular] and [Normaliz] software for our computations. Jafari was in part supported by a grant from IPM (No. 98130112). Stamate was partly supported by the University of Bucharest, Faculty of Mathematics and Computer Science through the 2018 and 2019 Mobility Fund.

References

[Backelin and Herzog 1989] J. Backelin and J. Herzog, "On Ulrich-modules over hypersurface rings", pp. 63–68 in *Commutative algebra* (Berkeley, CA, 1987)), edited by M. Hochster et al., Math. Sci. Res. Inst. Publ. **15**, Springer, 1989. MR Zbl

[Barucci and Fröberg 1997] V. Barucci and R. Fröberg, "One-dimensional almost Gorenstein rings", *J. Algebra* **188**:2 (1997), 418–442. MR Zbl

[Beck and Robins 2015] M. Beck and S. Robins, *Computing the continuous discretely*, 2nd ed., Springer, 2015. MR Zbl

[Bruns and Gubeladze 2009] W. Bruns and J. Gubeladze, *Polytopes, rings, and K-theory*, Springer, 2009. MR Zbl

[Bruns and Herzog 1998] W. Bruns and J. Herzog, *Cohen–Macaulay rings*, rev. ed., Cambridge Studies in Advanced Mathematics **39**, Cambridge Univ. Press, 1998. MR Zbl

[Danilov 1978] V. I. Danilov, "The geometry of toric varieties", *Uspekhi Mat. Nauk* **33**:2(200) (1978), 85–134, 247. In Russian; translated in *Russian Math. Surveys*, **33**:97 (1978) 97–154. MR Zbl

[De Negri and Hibi 1997] E. De Negri and T. Hibi, "Gorenstein algebras of Veronese type", *J. Algebra* **193**:2 (1997), 629–639. MR Zbl

[Dinu 2020] R. Dinu, "Gorenstein t-spread Veronese algebras", *Osaka J. Math.* **57**:4 (2020), 935–947. MR Zbl

[Fulton 1993] W. Fulton, *Introduction to toric varieties*, Annals of Mathematics Studies **131**, Princeton Univ. Press, 1993. MR Zbl

[Goto et al. 1976] S. Goto, N. Suzuki, and K. Watanabe, "On affine semigroup rings", *Japan. J. Math.* (*N.S.*) **2**:1 (1976), 1–12. MR Zbl

[Goto et al. 2013] S. Goto, N. Matsuoka, and T. T. Phuong, "Almost Gorenstein rings", *J. Algebra* **379** (2013), 355–381. MR Zbl

[Goto et al. 2015] S. Goto, R. Takahashi, and N. Taniguchi, "Almost Gorenstein rings—towards a theory of higher dimension", *J. Pure Appl. Algebra* **219**:7 (2015), 2666–2712. MR Zbl

[Herzog and Watanabe 2019] J. Herzog and K.-i. Watanabe, "Almost symmetric numerical semigroups", *Semigroup Forum* **98**:3 (2019), 589–630. MR Zbl

[Herzog et al. 2019] J. Herzog, T. Hibi, and D. I. Stamate, "The trace of the canonical module", *Israel J. Math.* **233**:1 (2019), 133–165. MR Zbl

[Hibi 1987] T. Hibi, "Distributive lattices, affine semigroup rings and algebras with straightening laws", pp. 93–109 in *Commutative algebra and combinatorics* (Kyoto, 1985), edited by M. Nagata and H. Matsumura, Adv. Stud. Pure Math. **11**, North-Holland, 1987. MR Zbl

[Hibi 1992] T. Hibi, "Dual polytopes of rational convex polytopes", *Combinatorica* **12**:2 (1992), 237–240. MR Zbl

[Hibi et al. 2019] T. Hibi, M. Lasoń, K. Matsuda, M. Michałek, and M. Vodička, "Gorenstein graphic matroids", 2019. arXiv

[Hochster 1972] M. Hochster, "Rings of invariants of tori, Cohen–Macaulay rings generated by monomials, and polytopes", *Ann. of Math.* (2) **96** (1972), 318–337. MR Zbl

[Miyazaki 2018] M. Miyazaki, "Almost Gorenstein Hibi rings", *J. Algebra* **493** (2018), 135–149. MR Zbl

[Nari 2013] H. Nari, "Symmetries on almost symmetric numerical semigroups", *Semigroup Forum* **86**:1 (2013), 140–154. MR Zbl

[Normaliz] W. Bruns, B. Ichim, T. Römer, R. Sieg, and C. Söger, "Normaliz: Algorithms for rational cones and affine monoids", software, https://www.normaliz.uni-osnabrueck.de.

[Singular] W. Decker, G.-M. Greuel, G. Pfister, and H. Schönemann, "Singular 4-1-2—A computer algebra system for polynomial computations", software, http://www.singular.uni-kl.de.

[Stanley 1978] R. P. Stanley, "Hilbert functions of graded algebras", *Advances in Math.* **28**:1 (1978), 57–83. MR Zbl

[Taniguchi 2018] N. Taniguchi, "On the almost Gorenstein property of determinantal rings", *Comm. Algebra* **46**:3 (2018), 1165–1178. MR Zbl

[Ulrich 1984] B. Ulrich, "Gorenstein rings and modules with high numbers of generators", *Math. Z.* **188**:1 (1984), 23–32. MR Zbl

[Villarreal 2015] R. H. Villarreal, *Monomial algebras*, 2nd ed., Chapman & Hall/CRC, 2015. MR Zbl

[Ziegler 1995] G. M. Ziegler, *Lectures on polytopes*, Graduate Texts in Mathematics **152**, Springer, 1995. MR Zbl

Received March 13, 2020.

JÜRGEN HERZOG
FAKULTÄT FÜR MATHEMATIK
UNIVERSITÄT DUISBURG-ESSEN
ESSEN
GERMANY
juergen.herzog@uni-essen.de

RAHELEH JAFARI
MOSAHEB INSTITUTE OF MATHEMATICS
KHARAZMI UNIVERSITY
TEHRAN
IRAN
and
SCHOOL OF MATHEMATICS
INSTITUTE FOR RESEARCH IN FUNDAMENTAL SCIENCES (IPM)
TEHRAN
IRAN
rjafari@ipm.ir

DUMITRU I. STAMATE
FACULTY OF MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF BUCHAREST
BUCHAREST
ROMANIA
dumitru.stamate@fmi.unibuc.ro

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 2, 2020
https://doi.org/10.2140/pjm.2020.309.381

NOTES ON EQUIVARIANT HOMOLOGY WITH CONSTANT COEFFICIENTS

SOPHIE KRIZ

For a finite group, we discuss a method for calculating equivariant homology with constant coefficients. We apply this method to completely calculate the geometric fixed points of the equivariant spectrum representing equivariant (co)homology with constant coefficients. We also treat a more complicated example of inverting the standard representation in the equivariant homology of split extraspecial groups at the prime 2.

1. Introduction

Equivariant spectra are the foundation of equivariant generalized homology and cohomology theory of G-spaces (and ultimately, G-spectra) for a finite (or more generally compact Lie) group G, which has all the formal properties of generalized nonequivariant homology and cohomology theory, including duality, along with stability under suspensions by finite-dimensional real representations. They were introduced and developed in [Lewis et al. 1986] (see also [Adams 1984; Greenlees 1985]). Equivariant homology and cohomology $H\underline{A}_G$ with constant coefficients in an abelian group A, on the other hand, can be defined on the chain level, as a part of the theory of Bredon [1967]. Both contexts are reconciled in [Lewis et al. 1981], where, more generally, equivariant Eilenberg–MacLane spectra of Mackey functors are defined. In this paper, we discuss a spectral sequence (Propositions 4, 5) which can be used to compute generalized equivariant homology of a G-CW-complex from its subquotients of constant isotropy. This spectral sequence is especially efficient in the case of $H\underline{A}_G$. For example, we shall prove that for a (finite) p-group G, the spectral sequence computing $H\underline{\mathbb{Z}/p}^G_*(X)$ always collapses to E^1 (see Theorem 7).

Note that this is false in cohomology. By [Bredon 1967], for a G-CW-complex X and an abelian group A,

$$H^*_G(X; \underline{A}) = H^*(X/G; A). \tag{1}$$

Let $G = \mathbb{Z}/2$ and $X = S^\alpha$ where α denotes the sign representation of $\mathbb{Z}/2$. Then $X/(\mathbb{Z}/2) \simeq *$. Thus, by (1), $H^n_G(X; \underline{A})$ is only nontrivial for $n = 0$, while X has a 1-cell of isotropy $\{e\}$.

MSC2020: 55N91.
Keywords: equivariant homology, geometric fixed points, extraspecial groups.

One may think, therefore, that equivariant homology with constant coefficients carries less information than cohomology. It turns out, however, that equivariant E-homologies of certain spaces give important information about a spectrum E. For example, the *geometric fixed point spectra* [Greenlees 1985; Lewis et al. 1986, II §8, 9] $\Phi^H E$, where H runs through subgroups of G, completely characterize the spectrum E. The coefficients $\Phi^G E_*$, for a finite group G, are the reduced E-homology of the smash product $S^{\infty V}$ of infinitely many copies of the one-point compactification S^V of the reduced regular representation V. Geometric fixed point spectra proved very useful in applications, for example, in [tom Dieck 1970; Hesselholt and Madsen 1997; Hill et al. 2016].

We will see that our method allows a complete computation of the coefficients of the geometric fixed point spectrum of homology with constant coefficients $\Phi^G(H\underline{A})_*$, which we denote by $\Phi^G(\underline{A})_*$, by reducing it to the case where G is an elementary abelian group. We show (Proposition 11) that $\Phi^G H\underline{A}_G = 0$ if G is not a p-group, and that

$$\Phi^G H\underline{A}_G = \Phi^{G/G'_p} H\underline{A}_{G/G'_p},$$

where G is a p-group and G'_p is its Frattini subgroup (Proposition 9). In the elementary abelian case, the computation was carried out for $A = \mathbb{Z}/p$ in my previous paper [Kriz 2015] (see also [Holler and Kriz 2017; 2020]).

Let first $p = 2$. We have

$$H^*((\mathbb{Z}/2)^n; \mathbb{Z}/2) = \mathbb{Z}/2[x_1, \dots, x_n],$$

where x_i have cohomological dimension 1. Let, for $\alpha = (\alpha_1, \dots, \alpha_n) \in (\mathbb{Z}/2)^n \smallsetminus \{0\}$,

$$x_\alpha = \alpha_1 x_1 + \dots \alpha_n x_n.$$

For $p > 2$, following the notation of [Kriz 2015], we have

$$H^*((\mathbb{Z}/p)^n; \mathbb{Z}/p) = \mathbb{Z}/p[z_i] \otimes \Lambda_{\mathbb{Z}/p}[dz_i],$$

where z_i have cohomological dimension 2 and dz_i have cohomological dimension 1. Let, for $\alpha = (\alpha_1, \dots, \alpha_n) \in (\mathbb{Z}/p)^n \smallsetminus \{0\}$,

$$z_\alpha = \alpha_1 z_1 + \dots + \alpha_n z_n,$$
$$dz_\alpha = \alpha_1 dz_1 + \dots + \alpha_n dz_n.$$

Theorem 1 [Holler and Kriz 2017; 2020; Kriz 2015]. (a) $\Phi^{(\mathbb{Z}/2)^n}(\underline{\mathbb{Z}/2})$ *is the subring of*

$$H^*((\mathbb{Z}/2)^n; \mathbb{Z}/2)[x_\alpha^{-1} \mid \alpha \neq 0]$$

generated by $y_\alpha = x_\alpha^{-1}$.

(b) *For* $p > 2$, $\Phi^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}/p})$ *is the subring of*

$$H^*((\mathbb{Z}/p)^n; \mathbb{Z}/p)[z_\alpha^{-1} | \alpha \in (\mathbb{Z}/p)^n \smallsetminus \{0\}]$$

generated by $t_\alpha = z_\alpha^{-1}$ *and* $u_\alpha = t_\alpha dz_\alpha$.

One can also explicitly describe these rings in generators and defining relations:

Theorem 2 [Holler and Kriz 2017; 2020; Kriz 2015]. (a) *For* $p = 2$, *we have*

$$\Phi_*^{(\mathbb{Z}/2)^n}(\underline{\mathbb{Z}/2}) \cong \mathbb{Z}/2[y_\alpha | \alpha \in (\mathbb{Z}/2)^n \smallsetminus \{0\}]/\sim$$

where $\sim$ *denotes the relations*

$$y_\alpha y_\beta + y_\alpha y_\gamma + y_\beta y_\gamma \sim 0$$

for $\alpha + \beta + \gamma = 0$. *The elements* y_α *are in degree* 1.

(b) *For* $p > 2$,

$$\Phi_*^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}/p}) = \mathbb{Z}/p[t_\alpha] \otimes \Lambda_{\mathbb{Z}/p}[u_\alpha]/\sim, \tag{2}$$

where $\sim$ *denotes the relations*

$$t_{i\alpha} \sim i^{-1} t_\alpha, \quad u_{i\alpha} \sim u_\alpha, \quad t_\beta t_{\alpha+\beta} + t_\alpha t_{\alpha+\beta} \sim t_\alpha t_\beta,$$
$$t_\beta u_{\alpha+\beta} - t_{\alpha+\beta} u_\beta + t_{\alpha+\beta} u_\alpha \sim u_\alpha t_\beta, \quad -u_\beta u_{\alpha+\beta} + u_\alpha u_{\alpha+\beta} \sim u_\alpha u_\beta,$$

where $i \in \mathbb{Z}/p \smallsetminus \{0\}$, *for* $\alpha, \beta, \alpha + \beta \in (\mathbb{Z}/p)^n \smallsetminus \{0\}$. *The elements* u_α *are in degree* 1 *and the elements* t_α *are of degree* 2.

This allows a complete answer for $\underline{\mathbb{Z}}$, the universal constant coefficients. Fix an element $\alpha_0 \in (\mathbb{Z}/p)^n \smallsetminus \{0\}$ and put $\tilde{u}_\alpha = u_\alpha - u_{\alpha_0}$, ($\tilde{y}_\alpha = y_\alpha - y_{\alpha_0}$ for $p = 2$). For $p = 2$, we also set $t_{\alpha_0} = y_{\alpha_0}^2$.

Theorem 3. (a) *For every prime* p, *we have that* $\Phi_*^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}})$ *is the subring of the ring* $\Phi_*^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}/p})$ *on which the Bockstein*

$$\beta : \Phi_*^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}/p}) \to \Phi_{*-1}^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}/p})$$

vanishes.

(b) *Explicitly, for* $p = 2$, *we have*

$$\Phi_*^{(\mathbb{Z}/2)^n}(\underline{\mathbb{Z}}) \cong \mathbb{Z}/2[\tilde{y}_\alpha, t_{\alpha_0} | \alpha \in (\mathbb{Z}/2)^n \smallsetminus \{0\}]/\sim,$$

where $\sim$ *denotes the relations*

$$\tilde{y}_{\alpha_0} \sim 0, \quad \tilde{y}_\alpha \tilde{y}_\beta + \tilde{y}_\alpha \tilde{y}_\gamma + \tilde{y}_\beta \tilde{y}_\gamma + t_{\alpha_0} \sim 0,$$

where $\alpha + \beta + \gamma = 0$.

(c) *For* $p > 2$,

$$\Phi_*^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}}) \cong (\mathbb{Z}/p[t_\alpha|\alpha \in (\mathbb{Z}/p)^n \smallsetminus \{0\}] \otimes \Lambda_{\mathbb{Z}/p}[\tilde{u}_\alpha|\alpha \in (\mathbb{Z}/p)^n \smallsetminus \{0\}])/\sim,$$

where $\sim$ *denotes the relations*

$$(3)\qquad \begin{gathered} t_{i\alpha} \sim i^{-1} t_\alpha, \quad \tilde{u}_{i\alpha} \sim \tilde{u}_\alpha, \quad t_\beta t_{\alpha+\beta} + t_\alpha t_{\alpha+\beta} \sim t_\alpha t_\beta, \quad \tilde{u}_{\alpha_0} \sim 0, \\ t_\beta \tilde{u}_{\alpha+\beta} - t_{\alpha+\beta}\tilde{u}_\beta + t_{\alpha+\beta}\tilde{u}_\alpha \sim t_\beta \tilde{u}_\alpha, \quad -\tilde{u}_\beta \tilde{u}_{\alpha+\beta} + \tilde{u}_\alpha \tilde{u}_{\alpha+\beta} \sim \tilde{u}_\alpha \tilde{u}_\beta, \end{gathered}$$

for $i \in \mathbb{Z}/p \smallsetminus \{0\}$ *and* $\alpha, \beta, \alpha+\beta \in (\mathbb{Z}/p)^n \smallsetminus \{0\}$.

The reductions contained in Propositions 9 and 11 are quite easy. However, if one considers the more general problem of calculating $\widetilde{H\underline{A}}_*^G(S^{\infty\gamma})$ for a general finite-dimensional representation γ of G, one gets nontrivial examples. One such example is treated in Section 4, where G is a split extraspecial 2-group and γ is the irreducible representation nontrivial on the center.

The present paper is organized as follows: In Section 2, we discuss our spectral sequences. In Section 3, we discuss the application to geometric fixed points. Section 4 contains the extraspecial group example.

2. The spectral sequences

For a G-equivariant spectrum E, we will need to use the homotopy co-fixed point (Borel homology) spectrum

$$(4)\qquad E_{hG} = (E \wedge EG_+)^G,$$

where EG is a nonequivariantly contractible free G-CW complex and for a space X, we write $X_+ = X \coprod \{*\}$. The formula (4) includes a key fact called the Adams isomorphism [Lewis et al. 1986, II §7]: G-equivariant cell spectra with G-free cells can be identified with naive (i.e., nonequivariant) cell spectra with a free cellular G-action. The Adams isomorphism says that for any cell G-spectrum E, E_{hG} is equivalent to $(E \wedge EG_+)/G$, where $E \wedge EG_+$ is considered as a naive G-spectrum.

Recall that a *family* $\mathscr{F}$ is defined as a set of subgroups of G that is closed under subconjugacies. For a family $\mathscr{F}$, we have a G-CW complex $E\mathscr{F}$ such that

$$\begin{aligned} E\mathscr{F}^H &\simeq * \quad \text{for } H \in \mathscr{F}, \\ E\mathscr{F}^G &\simeq \varnothing \quad \text{for } H \notin \mathscr{F}. \end{aligned}$$

If we denote by $\widetilde{X}$ the unreduced suspension of a G-space X, we have

$$\begin{aligned} \widetilde{E\mathscr{F}}^H &\simeq * \quad \text{for } H \in \mathscr{F}, \\ \widetilde{E\mathscr{F}}^G &\simeq S^0 \quad \text{for } H \notin \mathscr{F}. \end{aligned}$$

If V is a real G-representation, denote by $S(V)$ the unit sphere of V and by S^V the union of the 1-point compactifications of S^W for finite-dimensional subrepresentations W of V. If we set $\infty V = \bigoplus_\infty V$, then $S(\infty V)$ is a model for $E\mathscr{F}_V$, where

$\mathscr{F}_V = \{H \subseteq G \mid V^H \neq 0\}$. Thus $S^{\infty V}$ is a model for $\widetilde{E\mathscr{F}}_V$. Since $S^{\infty V} \wedge S^{\infty V} = S^{\infty V}$, for a commutative ring spectrum E,

$$\widetilde{E}_* \widetilde{E\mathscr{F}}_V = \widetilde{E}_* S^{\infty V}$$

is a commutative ring. Here $\widetilde{E}_* X$, for a G-spectrum E and a based G-CW complex X (recall that the base point is required to be G-fixed), is the equivariant reduced homology of X, i.e., $\pi_*(E \wedge \Sigma^\infty X)$ (without adding a disjoint base point). Note that this is also the $\mathbb{Z}$-graded part of the $RO(G)$-graded coefficient ring of $\alpha_V^{-1} E$, where $\alpha_V \in \pi_{-V} E$ is the class obtained from the inclusion $S^0 \to S^V$.

For a finite group G, a family $\mathscr{F}$, and an $H \in \mathscr{F}$, define the height of H inductively by

$$h_{\mathscr{F}}(H) = \max\{0, h_{\mathscr{F}}(K) \mid K \in \mathscr{F}, H \subsetneq K\} + 1.$$

Now let X be a G-CW-complex. Consider the family

$$\mathscr{F} = \mathscr{F}_X = \{H \subseteq G \mid X^H \neq \varnothing\}.$$

Let E be a G-spectrum. Then we have a spectral sequence converging to the E-homology of X, using Borel homology of parts of X of the same isotropy. There are two versions of the spectral sequence, one for the unreduced homology of X, the other for the reduced homology of its unreduced suspension. Keeping track of terms can be delicate, so we list both versions:

Proposition 4. *We have a spectral sequence*

$$E^1_{p,q} \Rightarrow E_{p+q} X,$$

where

$$E^1_{p,q} = \bigoplus_{(H),\, H \in \mathscr{F}_X,\, h_{\mathscr{F}_X}(H) = p} \left(\left(E^H \wedge \left(X^H / \bigcup_{H \subsetneq K} X^K \right) \right)_{hW(H)} \right)_{p+q}. \tag{5}$$

(*Here* $W(H) = N(H)/H$, *where* $N(H)$ *is the normalizer of* H *in* G, *and* (H) *runs through the conjugacy classes of* $H \in \mathscr{F}_X$.)

Proposition 5. *We have a spectral sequence*

$$E^1_{p,q} \Rightarrow \widetilde{E}_{p+q} \widetilde{X},$$

where

$$E^1_{0,q} = E_q(*),$$

$$E^1_{p,q} = \bigoplus_{(H), H \in \mathscr{F}_X,\, h_{\mathscr{F}_X}(H) = p} \left(\left(E^H \wedge \left(X^H / \bigcup_{H \subsetneq K} X^K \right) \right)_{hW(H)} \right)_{p+q-1}. \tag{6}$$

The proofs of these statements are sufficiently similar to only give one of them. We prove Proposition 5 which is more closely related to our applications.

Proof of Proposition 5. Define an increasing G-equivariant filtration of X by

$$F'_pX = \bigcup_{h_{\mathscr{F}}(H)\le p} X^H = \bigcup_{h_{\mathscr{F}}(H)=p} X^H.$$

Then define an increasing G-equivariant filtration of $\widetilde{X}$ by

$$\begin{aligned} F_0\widetilde{X} &= S^0, \\ F_p\widetilde{X} &= \bigcup_{h_{\mathscr{F}}(H)\le p} (\widetilde{X})^H = \bigcup_{h_{\mathscr{F}}(H)=p} (\widetilde{X})^H. \end{aligned}$$

We have a spectral sequence

$$E^1_{p,q} = \widetilde{E}_{p+q}(F_p\widetilde{X}/F_{p-1}\widetilde{X}) \Rightarrow \widetilde{E}_{p+q}(\widetilde{X}) = (E\wedge\widetilde{X})_{p+q}.$$

By definition, for $p\ge 1$,

$$F_p\widetilde{X}/F_{p-1}\widetilde{X} = \bigcup_{h_{\mathscr{F}}(H)=p} (\widetilde{X})^H \Big/ \bigcup_{h_{\mathscr{F}}(H)\le p-1} (\widetilde{X})^H.$$

On the other hand,

$$\begin{aligned} F_p\widetilde{X}/F_{p-1}\widetilde{X} &= (F_p\widetilde{X}/F_0\widetilde{X})/(F_{p-1}\widetilde{X}/F_0\widetilde{X}) \\ &= (\Sigma F'_p(X)_+)/(\Sigma F'_{p-1}(X)_+) \\ &= \Sigma(F'_p(X)/F'_{p-1}(X)). \end{aligned}$$

Thus,

$$\widetilde{E}_{p+q}(F_p\widetilde{X}/F_{p-1}\widetilde{X}) = \widetilde{E}_{p+q-1}(F'_p(X)/F'_{p-1}(X)).$$

Now, we have

$$\begin{aligned} F'_p(X)/F'_{p-1}(X) &= \bigcup_{h_{\mathscr{F}}(H)=p} X^H \Big/ \bigcup_{h_{\mathscr{F}}(H)<p} X^H \\ &= \bigvee_{h_{\mathscr{F}}(H)=p} \Big(X^H \Big/ \bigcup_{H\subsetneq K\in\mathscr{F}} X^K \Big). \end{aligned}$$

Note that

$$X^H \Big/ \bigcup_{H\subsetneq K\in\mathscr{F}} X^K$$

is a free based $W(H)$-CW complex.

On the other hand,

$$\begin{aligned}\widetilde{E}^G_{p+q-1}\Big(\bigvee_{h_{\mathscr{F}}(H)=p}\Big(X^H/\bigcup_{H\subsetneq K\in\mathscr{F}}X^K\Big)\Big) \\ &= \bigoplus_{(H),\, h_{\mathscr{F}}(H)=p}\widetilde{E}^G_{p+q-1}\Big(\bigvee_{H'=gHg^{-1}}\Big(X^H/\bigcup_{H'\subsetneq K\in\mathscr{F}}X^K\Big)\Big) \\ &= \bigoplus_{(H),\, h_{\mathscr{F}}(H)=p}\widetilde{E}^{N(H)}_{p+q-1}\Big(X^H/\bigcup_{H\subsetneq K\in\mathscr{F}}X^K\Big).\end{aligned}$$

To justify the third isomorphism above, note that G acts transitively on the conjugacy classes of H and thus

$$\bigvee_{H'=gHg^{-1}}\Big(X^H/\bigcup_{H'\subsetneq K\in\mathscr{F}}X^K\Big)$$

is the pushforward from $N(H)$ to G of

$$X^H/\bigcup_{H\subsetneq K\in\mathscr{F}}X^K.$$

Therefore,

$$\begin{aligned}E^1_{p,q} &= \bigoplus_{(H),\, h_{\mathscr{F}}(H)=p}\widetilde{E}^{N(H)}_{p+q-1}\Big(\Big(X^H/\bigcup_{H\subsetneq K}X^K\Big)\wedge EW(H)_+\Big) \\ &= \bigoplus_{(H),\, h_{\mathscr{F}}(H)=p}\Big(\Big(E^H\wedge\Big(X^H/\bigcup_{H\subsetneq K}X^K\Big)\Big)_{hW(H)}\Big)_{p+q-1},\end{aligned}$$

since $E^G=(E^K)^{G/K}$. □

We shall be especially interested in the case of classifying spaces of families. Consider, for a family $\mathscr{F}$ and a group $H\in\mathscr{F}$, the poset

(7) $$P^{\mathscr{F}}_H=\{K\in\mathscr{F}\mid K\supsetneq H\}$$

with respect to inclusion. Note that this poset has a $N(H)$-action by conjugation.

For a poset P, denote by $|P|$ the nerve (also called the classifying space or bar construction) of P, which one defines as the geometric realization of the simplicial set whose n-simplices are chains

$$x_0\le\cdots\le x_n$$

where faces are given by deletions and degeneracies by repetitions. This is a special case of the nerve of a category where n-simplices are composable n-tuples of morphisms.

Corollary 6. *We have a spectral sequence*

$$E^1_{p,q} \Rightarrow (E \wedge \widetilde{E\mathscr{F}})_{p+q} = \widetilde{E}_{p+q}\widetilde{E\mathscr{F}},$$

where

$$E^1_{0,q} = E_q(*)$$

and for $p > 0$,

$$E^1_{p,q} = \bigoplus_{(H),\ H\in\mathscr{F},h_{\mathscr{F}}(H)=p} ((E^H \wedge \widetilde{|P^{\mathscr{F}}_H|})_{hW(H)})_{p+q-1}, \tag{8}$$

where (H) *runs through the conjugacy classes of groups* $H \in \mathscr{F}$.

Proof. Apply Proposition 5 to $X = E\mathscr{F}$. Note that $\mathscr{F}_{E\mathscr{F}} = \mathscr{F}$. Let $H \in \mathscr{F} = \mathscr{F}_X$. Then

$$E\mathscr{F}^H \simeq *.$$

We may realize the system of spaces $(E\mathscr{F}^K)_{H\subsetneq K\in\mathscr{F}}$ as a $N(H)$-equivariant functor

$$F : P^{\mathscr{F}}_H \to \Delta^{Op}\text{-Set},$$

where the right-hand side denotes the category of simplicial sets, such that the canonical maps

$$\operatorname*{colim}_{y<x} F(y) \to F(x)$$

are injective for all $x \in P^{\mathscr{F}}_H$. Now if we denote by $\pi : P^{\mathscr{F}}_H \to *$ the terminal map, $\bigcup_{H\subsetneq K\in\mathscr{F}} E\mathscr{F}^K$ is the left Kan extension $\pi_\# F$. However, by our injectivity assumption, the canonical $N(H)$-equivariant map

$$L\pi_\# F \to \pi_\# F \tag{9}$$

(where L denotes the left derived functor) is a nonequivariant equivalence. Additionally, the left-hand side can be expressed as the 2-sided bar construction $B(*, P^{\mathscr{F}}_H, F)$. Since the values of F on objects are contractible, we have an $N(H)$-equivariant map

$$B(*, P^{\mathscr{F}}_H, F) \to B(*, P^{\mathscr{F}}_H, *) = |P^{\mathscr{F}}_H| \tag{10}$$

which is a nonequivariant equivalence. The equivariant maps (9), (10) induce equivalences on homotopy fixed points, since they are nonequivariant equivalences. Thus, we have an equivalence

$$\left(E^H \wedge \left(E\mathscr{F}^H / \bigcup_{H\subsetneq K} E\mathscr{F}^K\right)\right)_{hW(H)} \sim (E^H \wedge \widetilde{|P^{\mathscr{F}}_H|})_{hW(H)}$$

as required. □

Again, there is also an unbased version for $E\mathscr{F}$ instead of $\widetilde{E\mathscr{F}}$.

Theorem 7. *If G is a p-group and $E = H\underline{\mathbb{Z}/p}$, then the spectral sequence (5) collapses to E^1. Additionally, the spectral sequence (6) collapses to E^1 when $X^G = \varnothing$.*

Proof. Again, the proofs of the reduced and unreduced cases are similar. We treat the unreduced case this time.

Suppose G is a p-group and X is a G-CW-complex. Then $H\underline{\mathbb{Z}/p}^G_* X$ can be calculated on the chain level. Let $C^G(X)$ be the cellular chain complex of X in the category of G-coefficient systems in the sense of Bredon [1967], i.e., functors $\mathscr{O}_G^{Op} \to Ab$ where $\mathscr{O}_G$ is the orbit category. This is defined by

$$(C_n^G(X))(G/H) = C_n^{\text{cell}}(X^H).$$

Then we have

$$H\underline{A}_n(X) = H_n(C^G(X) \otimes_{\mathscr{O}_G} \underline{A}),$$

where $\underline{A}$ is the constant co-coefficient system for an abelian group A, i.e., the functor $\mathscr{O}_G \to Ab$ where for $f : G/H \to G/K \in Mor(\mathscr{O}_G)$, f_* is multiplication by $|K|/|H|$.

We will show that

$$C^G(X) \otimes_{\mathscr{O}_G} \underline{\mathbb{Z}/p} \cong \bigoplus_{(H)} \widetilde{C}^{\text{cell}}\Big(X^H / \bigcup_{H \subsetneq K} X^K\Big) \otimes_{\mathbb{Z}[W(H)]} \mathbb{Z}/p. \tag{11}$$

In each degree separately, (11) holds as abelian groups, since all $\mathscr{O}_G$-identifications corresponding to nonisomorphisms are trivial. For any $f : H \subsetneq K$, consider the summand of the differential d^{tot} of $C^G(X) \otimes_{\mathscr{O}_G} \underline{\mathbb{Z}/p}$

$$d_{H,K} : C_n(X^H) \otimes \mathbb{Z}/p \to C_{n-1}(X^K) \otimes \mathbb{Z}/p.$$

By conjugation, it suffices to show that these maps are 0.

The differential d^{tot} is given by

$$\bigoplus d/\sim : \bigoplus C_n(X^H) \otimes \mathbb{Z}/p/\sim \to \bigoplus C_{n-1}(X^H) \otimes \mathbb{Z}/p/\sim,$$

where $\sim$ denotes the equivalence relation generated by

$$f^* a \otimes b \sim a \otimes f_* b.$$

In particular, for $q \in C_n(X^H)$, let $c \in C_{n-1}(X^H)$ be the sum of the terms of $d(q)$ on cells in X^K, where d is the differential of $C(X^H)$. Then we have

$$d_{H,K}(q) = f^* c \otimes 1 = c \otimes f_* 1 = c \otimes \frac{|K|}{|H|} = 0.$$

Therefore, $d_{H,K} = 0$. Thus, we have proved (11), and hence the spectral sequence collapses to E^1. □

3. Geometric fixed points

In this section, we shall apply the methods of the previous section to completely calculate the coefficients of the geometric fixed point spectrum

$$\Phi^G_* H\underline{A} = \widetilde{H\underline{A}}_* \widetilde{E\mathscr{F}[G]},$$

where $\mathscr{F}[G] = \{H \mid H \subsetneq G\}$ for any finite group G.

A basic fact about posets is useful for computing examples: For functors $F : \mathscr{C} \to \mathscr{D}$ between any categories, we get continuous maps

$$|F| : |\mathscr{C}| \to |\mathscr{D}|,$$

and for natural transformations $F \to G$, we get

$$|F| \simeq |G|.$$

In particular, if $f, g : P \to Q$ are morphisms of posets and $f \leq g$, then $|f| \simeq |g|$. Thus, in particular, if P has lowest or highest element x, $|\mathrm{Id}| \simeq |\mathrm{Const}_x|$. Therefore, then, $|P|$ is contractible.

Lemma 8. *For posets $Q \subseteq P$, if there exists a morphism $f : P \to Q$ satisfying both*

(1) *for every $x \leq y \in P$, we have $f(x) \leq f(y)$,*

(2) *for every $x \in P$, we have $x \leq f(x)$ (or alternatively $x \geq f(x)$),*

then the inclusion induces a homotopy equivalence $|Q| \simeq |P|$.

Proof. Suppose we have posets P, Q and a morphism $f : P \to Q$ satisfying the assumptions. We have an inclusion

$$\iota : Q \hookrightarrow P.$$

For an $x \in Q$, $\iota(x) = x$. So, for every $x \in Q$,

$$f \circ \iota(x) = f(x) \geq x = \mathrm{Id}_Q(x).$$

Therefore $f \circ \iota \geq \mathrm{Id}_Q$. So $|f||\iota| = |f \circ \iota| \simeq \mathrm{Id}_{|Q|}$. On the other hand, for $x \in P$, $f(x) \in Q$, so

$$\iota \circ f(x) = f(x) \geq x = \mathrm{Id}_P(x).$$

So, $f \circ \iota \geq \mathrm{Id}_P$. So $|\iota||f| = |\iota \circ f| \simeq \mathrm{Id}_{|P|}$. □

Denote by G'_p the Frattini subgroup of G, i.e., the subgroup generated by the commutator subgroup and p-th powers.

Proposition 9. *Suppose G is a p-group. Then for any G-spectrum E, we have*

$$\Phi^G(E) \simeq \Phi^{G^{ab}/p}(E^{G'_p}).$$

In particular, for a constant Mackey functor $\underline{A}$, we have

$$\Phi^G(\underline{A}) \simeq \Phi^{G^{ab}/p}(\underline{A}).$$

Comment. Note that one always has an equivalence

$$\Phi^G(E) \simeq \Phi^{G/H}\Phi^H E.$$

The special feature here is that $E^H \to \Phi^H E$ induces an equivalence on G/H-geometric fixed points if G is a p-group and H is the Frattini subgroup.

We shall first prove:

Lemma 10. *Let G be a p-group and let $H \subseteq G$ be a subgroup not containing G'_p. Then*

$$|P_H^{\mathscr{F}[G]}| \simeq *.$$

Proof. Denote by Q the poset of proper subgroups of G containing $G'_p \cdot H$. We have an inclusion $Q \subseteq P_H^{\mathscr{F}[G]}$. By the Burnside basis theorem, for any subgroup $K \subsetneq G$, we have $K \cdot G'_p \subsetneq G$. Thus, we have a map of posets $\varphi : P_H^{\mathscr{F}[G]} \to Q$ given by

$$K \mapsto K \cdot G'_p.$$

Also, for $K \in Q$, $\varphi(K) \supseteq K$. Thus, $|Q| \simeq |P_H^{\mathscr{F}[G]}|$ by Lemma 8.
However, $|Q| \simeq *$ since Q has a minimal element. □

Note that this implies Proposition 9, since the quotient map

$$\widetilde{E\mathscr{F}}[G] \to \widetilde{E\mathscr{F}}[G/G'_p] \tag{12}$$

induces an isomorphism on E^1-terms of the spectral sequence (5). (The spectral sequence is not functorial in general with respect to change of groups. In the present case, however, we have a surjection of groups which preserves the height of the proper subgroup containing G'_p, while the remaining terms in the source are 0. Thus, a morphism of spectral sequences which is an isomorphism an E^1 arises.)

Also note that for the present purpose, the spectral sequence can be skipped entirely and one can simply argue that (12) is an equivalence by examining its H-fixed points for each H: The fixed point set of the left-hand side is contractible for $H \subsetneq G$ and equivalent to S^0 if $H = G$, while the right-hand side is contractible if $H \cdot G'_p \subsetneq G$ and equivalent to S^0 if $H \cdot G'_p = G$. By the Burnside basis theorem, both conditions are equivalent. This was pointed out to me during the process of revising this paper.

Proposition 11. *If G is not a p-group, then*

$$\Phi^G(\underline{A}) = 0.$$

Proof. First, suppose G is a finite group that is not a p-group.

The spectrum $\Phi^G(\underline{\mathbb{Z}})$ is a commutative ring spectrum, since we have

$$S^{\infty V} \wedge S^{\infty V} \cong S^{\infty V}.$$

Choose a prime p. Then by the first Sylow theorem, there exists a p-Sylow subgroup P of G. Then there exists an H with $P \subseteq H \subsetneq G$ that is maximal (i.e., there does not exist K such that $P \subseteq H \subsetneq K \subsetneq G$). Therefore the contribution of H to the spectral sequence will include

$$\widetilde{H}_0^{W(H)}(\widetilde{\varnothing}) \simeq H_0^{W(H)}(*) = \mathbb{Z}. \tag{13}$$

In (13), $1 \in \mathbb{Z}$ represents an element $\eta \in E^1_{1,0}$, where

$$d^1(\eta) = \pm\frac{|G|}{|H|} \in \mathbb{Z} = E^1_{0,0}.$$

Since we have $p \nmid |G|/|H|$, the gcd of all these numbers is 1, and thus, $1 \in \mathbb{Z} = E^1_{0,0}$ of the spectral sequence (8) is killed. Since $\Phi^G_*(\underline{\mathbb{Z}})$ is a commutative ring, it must be 0. So

$$\Phi^G(\underline{\mathbb{Z}}) = 0. \qquad \square$$

Now, one can apply the results of my previous paper [Kriz 2015], as well as those of [Holler and Kriz 2017; 2020], to calculate $\Phi^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}})$.

Recall that for any space or spectrum X, we can obtain maps

$$H^n(X;\mathbb{Z}/p) \xrightarrow{\beta} H^{n+1}(X;\mathbb{Z}/p), \tag{14}$$

$$H^n(X;\mathbb{Z}/p) \xrightarrow{\beta} H^{n+1}(X;\mathbb{Z}) \tag{15}$$

as the connecting maps of the long exact sequences from taking cohomology with coefficients in the following respective short exact sequences:

$$0 \to \mathbb{Z}/p \to \mathbb{Z}/(p^2) \to \mathbb{Z}/p \to 0,$$
$$0 \to \mathbb{Z} \xrightarrow{\cdot p} \mathbb{Z} \to \mathbb{Z}/p \to 0.$$

These are called the Bockstein maps, and the long exact sequence involving (15) forms an exact couple which gives rise to the Bockstein spectral sequence, in which (14) is d^1.

Consider first the case of $p > 2$. Recalling the notation of Theorems 1 and 2, the Bockstein acts by

$$\beta(dz_i) = z_i, \quad \beta z_i = 0.$$

Also recall that we have

$$\beta(ab) = \beta a \cdot b + (-1)^{|a|} a \cdot \beta b. \tag{16}$$

Thus we get

$$\beta(t_\alpha) = 0, \quad \beta(u_\alpha) = 1$$

Note that β preserves the relations of (2). For example,

$$\begin{aligned}&\beta(-u_\beta u_{\alpha+\beta}+u_\alpha u_{\alpha+\beta}-u_\alpha u_\beta)\\ &\;=-\beta(u_\beta)u_{\alpha+\beta}+u_\beta\beta(u_{\alpha+\beta})+\beta(u_\alpha)u_{\alpha+\beta}-u_\alpha\beta(u_{\alpha+\beta})-\beta(u_\alpha)u_\beta+u_\alpha\beta(u_\beta)=0.\end{aligned}$$

Now by computing directly on the chain level the equivariant $(\mathbb{Z}/p)^n$-homology of $S^{\infty V}$, where V is the reduced regular representation, the standard $(\mathbb{Z}/p)^n$-CW decomposition of $S^{\infty V}$ (thought of as a colimit of smash products of representation spheres of nontrivial 1-dimensional complex representations α) has only one 0-cell beside the base point, to which (using one α) p 1-cells are attached on which $(\mathbb{Z}/p)^n$ acts transitively. Thus, $\Phi_0^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}}) = \mathbb{Z}/p$. Since $\Phi_*^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}})$ is a ring, it has characteristic p (i.e., every element is annihilated by p). Thus the Bockstein spectral sequence collapses to E^2, or in other words

$$H_*(\Phi_*^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}/p}), \beta) = 0. \tag{17}$$

Hence, we have an exact sequence

$$0 \to \Phi_*^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}}) \to \Phi_*^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}/p}) \xrightarrow{\beta} \Phi_*^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}/p}). \tag{18}$$

Therefore, $\Phi_*^{(\mathbb{Z}/p)^n}(\underline{\mathbb{Z}})$ contains the elements t_α and $\sum_i a_i u_{\alpha_i}$, where $\sum_i a_i = 0 \in \mathbb{Z}/p$. Choosing an $\alpha_0 \in (\mathbb{Z}/p)^n \smallsetminus \{0\}$, since we are in characteristic p, the elements $\sum_i a_i u_{\alpha_i}$ where $\sum_i a_i = 0 \in \mathbb{Z}/p$ are linear combinations of $\tilde{u}_\alpha = u_\alpha - u_{\alpha_0}$. One easily verifies the relations (3). For example,

$$\begin{aligned}&-\tilde{u}_\beta\tilde{u}_{\alpha+\beta} + \tilde{u}_\alpha\tilde{u}_{\alpha+\beta} - \tilde{u}_\alpha\tilde{u}_\beta\\ &\quad= -(u_\beta - u_{\alpha_0})(u_{\alpha+\beta} - u_{\alpha_0}) + (u_\alpha - u_{\alpha_0})(u_{\alpha+\beta} - u_{\alpha_0}) - (u_\alpha - u_{\alpha_0})(u_\beta - u_{\alpha_0})\\ &\quad= -u_\beta u_{\alpha+\beta} + u_{\alpha_0}u_{\alpha+\beta} + u_\beta u_{\alpha_0} - u_{\alpha_0}^2 + u_\alpha u_{\alpha+\beta} - u_{\alpha_0}u_{\alpha+\beta} - u_\alpha u_{\alpha_0}\\ &\qquad\qquad + u_{\alpha_0}^2 - u_\alpha u_\beta + u_{\alpha_0}u_\beta + u_\alpha u_{\alpha_0} - u_{\alpha_0}^2\\ &\quad= 0.\end{aligned}$$

Let $\widetilde{R}_n$ denote the ring $\mathbb{Z}/p[t_\alpha, \tilde{u}_\alpha]$ modulo the relations (3). Write (18) as

$$0 \to R_{\mathbb{Z}} \to R_{\mathbb{Z}/p} \xrightarrow{\beta} R_{\mathbb{Z}/p}.$$

We therefore have a homomorphism of rings

$$\varphi : \widetilde{R}_n \to R_{\mathbb{Z}}.$$

We want to prove that this is an isomorphism.

Now, let us consider $p = 2$. Choose again a representative $\alpha_0 \in (\mathbb{Z}/2)^n \smallsetminus \{0\}$. Again, the $R_{\mathbb{Z}}$ contains elements of the form $\sum_i a_i y_i$ with $\sum_i a_i = 0$ which are generated by $\tilde{y}_\alpha = y_\alpha - y_{\alpha_0}$. Also, the elements $t_\alpha = y_\alpha^2 \in R_{\mathbb{Z}}$ (by (16)), but note that $\tilde{y}_\alpha^2 = y_\alpha^2 + y_{\alpha_0}^2$ (since we are in characteristic 2), so we only need to include t_{α_0} in the generators. Now similarly as for $p > 2$, one proves the relations.

$$\tilde{y}_{\alpha_0} = 0, \quad \tilde{y}_\alpha \tilde{y}_\beta + \tilde{y}_\alpha \tilde{y}_{\alpha+\beta} + \tilde{y}_\beta \tilde{y}_{\alpha+\beta} + t_{\alpha_0} = 0. \tag{19}$$

Let $\widetilde{R}_n$ denote the quotient of the ring $\mathbb{Z}/2[\tilde{y}_\alpha, t_{\alpha_0}]$ by relations (19).

Again, we have a homomorphism of rings

$$\begin{aligned} \varphi : \widetilde{R}_n &\to R_{\mathbb{Z}}, \\ \tilde{y}_\alpha &\mapsto y_\alpha - y_{\alpha_0}, \\ t_{\alpha_0} &\mapsto y_{\alpha_0}^2. \end{aligned}$$

We can prove that φ is an isomorphism by calculating the Poincaré series of $\widetilde{R}_n$, checking that it is the same as the Poincaré series of $R_{\mathbb{Z}}$ and exhibiting an additive basis of $\widetilde{R}_n$ that is linearly independent in $R_{\mathbb{Z}/2}$.

Recall from [Holler and Kriz 2017; 2020] that the Poincaré series of $R_{\mathbb{Z}/p}$ is

$$P(R_{\mathbb{Z}/p}) = \frac{1}{(1-x)^n} \prod_{i=1}^{n} (1 + (p^{i-1} - 1)x).$$

Thus, by (17), we have an exact sequence of graded $\mathbb{Z}/p$-vector spaces

$$0 \to R_{\mathbb{Z}} \to R_{\mathbb{Z}/p} \xrightarrow{\beta} R_{\mathbb{Z}/p}[1] \xrightarrow{\beta} R_{\mathbb{Z}}[2] \to 0$$

and hence

$$P(R_{\mathbb{Z}}) = \frac{1}{1+x} P(R_{\mathbb{Z}/p}) = \frac{1}{(1-x^2)(1-x)^{n-1}} \prod_{i=1}^{n} (1 + (p^{i-1} - 1)x). \tag{20}$$

For $p = 2$, we know that the Poincaré series of $\widetilde{R}_1$ is $1/(1-x^2)$. Now we can treat the α as elements of $(\mathbb{Z}/2)^n \smallsetminus \{0\}$ and assume that $\alpha_0 = (1, 0, \ldots, 0)$. For $n = 2$, we only have $\tilde{y}_{(0,1)}, \tilde{y}_{(1,1)}, \tilde{y}_{(1,0)} = 0$. By the relations, we have $\tilde{y}_{(0,1)} \tilde{y}_{(1,1)} = t_{\alpha_0}$. Therefore this ring has additive basis $\tilde{y}_{(0,1)}^{m \geq 1} t_{\alpha_0}^{m'}$, $\tilde{y}_{(1,1)}^{k \geq 1} t_{\alpha_0}^{k'}$, and $t_{\alpha_0}^{\ell}$ $(m', k', \ell \geq 0)$, which give the terms $x/((1-x)(1-x^2))$ twice and $1/(1-x^2)$ in the Poincaré series. Therefore, the Poincaré series of the ring is

$$P(\widetilde{R}_2) = \frac{1}{1-x^2} + 2 \cdot \frac{x}{(1-x)(1-x^2)} = \frac{1+x}{(1-x)(1-x^2)}.$$

After this, we can continue by induction since the relations imply

$$P(\widetilde{R}_n) = P(\widetilde{R}_{n-1}) \cdot (1 + (2^{n-1} - 1)x) \cdot \frac{1}{1-x}$$

similarly as in [Holler and Kriz 2020]: The additive basis is formed by the additive basis of $\widetilde{R}_{n-1}$ times $\tilde{y}_{(0,\ldots,0,1)}^{\geq 0}$ or $\tilde{y}_{(\alpha',1)}^{\geq 1}$, where $\alpha' \in (\mathbb{Z}/2)^{n-1} \smallsetminus \{0\}$. These elements are linearly independent in $R_{\mathbb{Z}/2}$ by performing a similar induction there (which was done in [Holler and Kriz 2020]).

The case of $p > 2$ is completely analogous. We define the homomorphism of rings

$$\varphi : \widetilde{R}_n \to R_{\mathbb{Z}},$$
$$\tilde{u}_\alpha \mapsto u_\alpha - u_{\alpha_0},$$
$$t_\alpha \mapsto t_\alpha,$$

which again we can check is a ring homomorphism by computing the relations in the target. Again, we check φ is an isomorphism by checking that for every n, the Poincaré series of $\widetilde{R}_n$ agrees with (20). We have $\widetilde{R}_1$ is generated by the basis of $(t_{(1)})^{m \geq 0}$. Thus, as before, $P(\widetilde{R}_1) = 1/(1-x^2)$. This time, for every $n \geq 2$, an additive basis of $\widetilde{R}_n$ is given by the additive basis of $\widetilde{R}_{n-1}$ times $t^{\geq 0}_{(0,\dots,0,1)} \cdot \tilde{u}^{\epsilon}_{(0,\dots,0,1)}$, where $\epsilon \in \{0, 1\}$, or times $t^{\geq 1}_{(\alpha',1)}$, or times $t^{\geq 0}_{(\alpha',1)} \cdot \tilde{u}_{(\alpha',1)}$, where $\alpha' \in (\mathbb{Z}/p)^{n-1} \smallsetminus \{0\}$. This gives

$$P(\widetilde{R}_n) = P(\widetilde{R}_{n-1}) \cdot (1 + (p^{n-1} - 1)x) \cdot \frac{1}{1-x},$$

and we can proceed by induction.

4. Another example

For the rest of the paper, we will consider equivariant homology with constant coefficients $\underline{\mathbb{Z}/p}$ for a prime p. If $G = (\mathbb{Z}/p)^n$ is an elementary abelian group, S is a set of 1-dimensional representations (real or complex depending on whether $p = 2$ or $p > 2$), and $\gamma = \bigoplus_{\alpha \in S} \alpha$, then we completely calculated in [Kriz 2015] the ($\mathbb{Z}$-graded) coefficients of

$$\widetilde{H\underline{\mathbb{Z}/p}}^G_*(S^{\infty\gamma}) = \widetilde{H\underline{\mathbb{Z}/p}}^G_* \widetilde{E\mathscr{F}_\gamma}. \tag{21}$$

By the above method, for any p-group G and any set S of nontrivial irreducible 1-dimensional representations of G/G'_p, $\gamma = \bigoplus_{\alpha \in S} \alpha$, we have

$$\widetilde{H\underline{\mathbb{Z}/p}}^G_*(S^{\infty\gamma}) = \widetilde{H\underline{\mathbb{Z}/p}}^{G/G'_p}_*(S^{\infty\gamma}).$$

However, for a G-representation γ which does not factor through G/G'_p the calculation of (21) can be nontrivial. In this section, as an example, we consider the case where G is the split extraspecial group (as described below) at $p = 2$ and V is the irreducible real representation nontrivial on the center. This group is the central product of n copies of D_8. We write $V_n = (\mathbb{Z}/2 \oplus \mathbb{Z}/2)^n$, where the generators of the i-th copy of $\mathbb{Z}/2 \oplus \mathbb{Z}/2$ are denoted by $v_{1,i}$, $v_{2,i}$. Define $q(v_{1,i}) = q(v_{2,i}) = 0$, $q(v_{1,i} + v_{2,i}) = 1$, and let q be additive between different i summands. This is a split quadratic form on the $\mathbb{F}_2$-vector space V_n with associated symplectic form

$$b(x, y) = q(x + y) + q(x) + q(y).$$

A vector subspace $W \subseteq V$ is called *isotropic* when b is 0 on W and is called *q-isotropic* if q is 0 on W. The split extraspecial group $\widetilde{V}_n$ is an extension

$$1 \to \mathbb{Z}/2 \to \widetilde{V}_n \to V_n \to 1,$$

where for $v \in V_n,\ 2v \neq 0$ if and only if $q(v) \neq 0$ and for $v, w \in V_n,\ vwv^{-1}w^{-1} = b(v, w)$.

Clearly $\widetilde{V}_n$ is isomorphic to a central product of n copies of D_8 and the (real) irreducible representation γ nontrivial on the center is obtained as the tensor product of the dihedral representation of the n copies of D_8. We shall apply Proposition 4 to

$$\widetilde{H\underline{\mathbb{Z}/2}}_*^{\widetilde{V}_n}(S^{\infty\gamma}).$$

The family $\mathscr{F}_\gamma$ consists of elementary abelian subgroups of $\widetilde{V}_n$ disjoint with the center which project to q-isotropic subspaces of V_n. Two subgroups are conjugate if and only if they project to the same q-isotropic subspaces U of V_n. We shall refer to these subgroups as *decorations* of U, and call them *decorated* q-isotropic subspaces.

We shall need to consider the following modular representations of the split extraspecial 2-group $\widetilde{V}_n$: All these representations will factor through the Frattini quotient V_n. By the representation $\underline{2}_i$, we mean a tensor product of the regular representation of $\mathbb{Z}/2\{v_{1,i}\}$ with the trivial representation on $\mathbb{Z}/2\{v_{2,i}\}$, where $v_{1,j}, v_{2,j}$ act trivially for $j \neq i$. (For counting purposes, equivalently, 1 and 2 can be reversed). The representation $\underline{3}_i$ is the kernel of the augmentation from the regular representation on $\mathbb{Z}/2\{v_{1,i}, v_{2,i}\}$ to the trivial representation (with the other coordinates also acting trivially).

Let $\underline{P}_n$ be the poset of elementary abelian subgroups of $\widetilde{V}_n$ which project to a nontrivial q-isotropic subspace of V_n. We will refer to these subgroups as *decorated q-isotropic subspaces* of V_n.

Theorem 12. *For $n > 1$, $\widetilde{H}_k(|\overline{P}_n|) = 0$ except for $k = n-1$. As a $\widetilde{V}_n$-representation, $\mathscr{H}_n := \widetilde{H}_{n-1}(|\overline{P}_n|)$ is given recursively as follows*:

$$\mathscr{H}_1 = \underline{3}_1,$$

$$\begin{aligned}\mathscr{H}_{n+1} = \underline{3}_{n+1} \otimes \mathscr{H}_n \\ \oplus \underline{2}_{n+1} \otimes (2^{2n-1} - 2^{n-1})\mathscr{H}_n \\ \oplus \underline{2}_{n+1} \otimes (2^{2n-1} + 2^{n-1} - 1)(2\mathscr{H}_n - \underline{2}_n 2^{2n-2}\mathscr{H}_{n-1}).\end{aligned} \tag{22}$$

The subtraction in the last term is to be interpreted recursively as follows: *We set $\mathscr{H}_0 = 1$. Then one copy of $\underline{2}_n\mathscr{H}_{n-1}$ is "subtracted" from the first summand of (22) with n replaced by $n-1$, to leave a copy of $\mathscr{H}_{n-1}$. The remaining $2^{2n-2} - 1$ copies of $\underline{2}_n\mathscr{H}_{n-1}$ are subtracted from the second summand of the formula (22) with n replaced by $n-1$ (thus, only one copy of the $2\mathscr{H}_n$ is involved in the subtraction).*

Comment. The expression for $\mathscr{H}_n$ given by the theorem is a direct sum of representations of the form

$$\bigotimes_{i\in S_1} \underline{2}_i \otimes \bigotimes_{j\in S_2} \underline{3}_j$$

for $S_1 \cap S_2 = \varnothing$. I do not know if homology groups of $\widetilde{V}_n$ with coefficients in these representations are all completely known. For $|S_2| \leq 2$, they can be deduced from the computation of Quillen [1971].

Proof. We proceed by induction on n. In the case of $n = 1$, the isotropic subspaces are $\langle v_1\rangle$ and $\langle v_2\rangle$, and there are two lifts to $\widetilde{V}_1$, and each pair of lifts is given by one $\mathbb{Z}/2$-summand of $V_1 = \mathbb{Z}/2 \oplus \mathbb{Z}/2$. Note that we have $\underline{3}$, because we must consider reduced homology.

Now to pass from $|\overline{P}_n|$ to $|\overline{P}_{n+1}|$, consider first the poset $\overline{Q}_n$ of decorated q-isotropic subspaces of V_{n+1} which intersect nontrivially with V_n. Then the inclusion $|\overline{P}_n| \subset |\overline{Q}_n|$ is an equivalence by Lemma 8, considering the map the other way given by

$$W \mapsto W \cap V_n.$$

Now a q-isotropic subspace W of V_{n+1} with $W \cap V_n = 0$ has $\dim(W) \leq 2$. Let $\overline{R}_{n+1}$ denote the union of Q_n and the set of decorated isotropic subspaces $W \subset V_{n+1}$ with $\dim(W) = 2$, $W \cap V_n = 0$. Then the poset $(\overline{R}_{n+1})_{\geq W} \cap Q_n$ for any such space W consists of copies of the poset $\underline{4}\,\overline{P}_{n-1}$ (here we use $\underline{4}$ to denote 4 additional independent decorations). The factors (22) correspond to the choices of W, consisting of one nonzero q-isotropic vector $w \in V_n$, and one vector w' with $q(w') = 1$, $b(w, w') = 1$. There are 2^{2n-2} choices of w' for each w. Then $W = \langle v_1 + w, v_1 + v_2 + w'\rangle$. This leads to a based cofibration

$$|\overline{P}_n| \to |\overline{R}_{n+1}| \to \bigvee_{\underline{4}(2^{2n-1}+2^{n-1}-1)2^{2n-2}} \Sigma|\overline{P}_{n-1}|. \tag{23}$$

Now $\overline{P}_{n+1}$ is the union of $\overline{R}_{n+1}$ with the set of all decorated q-isotropic subspaces $W \subset V_{n+1}$ with $W \cap V_n = 0$, $\dim(W) = 1$. For such a space W, $(\overline{P}_{n+1})_{\geq W} \cap R_{n+1}$ consists of

$$\underline{2}(2(2^{2n-1} + 2^{n-1}) + (2^{2n-1} - 2^{n-1})) \tag{24}$$

copies of $\overline{P}_n$. The $\underline{2}$ corresponds to additional decorations. The $2(2^{2n-1} + 2^{n-1})$ summands correspond to q-isotropic vectors of the form $w + v_1$, $w + v_2$, for $w \in V_n$, the $2^{2n-1} - 2^{n-1}$ summand correspond to q-isotropic vectors of the form $w + v_1 + v_2$. Thus, we obtain a based cofibration sequence

$$|\overline{R}_{n+1}| \to |\overline{P}_{n+1}| \to \bigvee_{\underline{2}(2(2^{2n-1}+2^{n-1})+(2^{2n-1}-2^{n-1}))} \Sigma|\overline{P}_n|. \tag{25}$$

Now from (23) and (25), we can easily eliminate $|\overline{R}_{n+1}|$, as we see that the copies of $\Sigma|\overline{P}_n|$ in (25) corresponding to $w = 0$ project identically to $\Sigma|\overline{P}_n| \subset \Sigma|\overline{R}_{n+1}|$ under the connecting map. Thus, we obtain a cofibration sequence of the form

$$(26)\quad |\overline{P}_{n+1}| \to \bigvee_{\underline{2}(2(2^{2n-1}+2^{n-1}-1)+2^{2n-1}-2^{n-1})} \Sigma|\overline{P}_n| \to \bigvee_{\underline{4}(2^{2n-1}+2^{n-1}-1)2^{2n-2}} \Sigma^2|\overline{P}_{n-1}|.$$

The second map (26) is shown to be onto in reduced homology using the sums of terms indicated in the statement of the theorem. (In particular, we consider, for a q-isotropic vector $w+v$, with $0 \neq w \in V_{n+1}$, all choices of vectors w' such that $\langle w+v_1, w'+v_1+v_2\rangle$ is q-isotropic. Note that this includes but is not equal to, for $n > 1$, all $\langle w+v_1, u\rangle$ q-isotropic.)

The dichotomy between canceling the first or second summand in (22) comes from distinguishing whether the projection of w to V_{n-1} is 0 or not. □

Comment. The same method shows that the reduced homology of the poset of undecorated q-isotropic subspaces of V_n is concentrated in degree $n-1$ and has rank $2^{n(n-1)}$. This poset (for $n \geq 2$) is the Tits building of $\Omega^+_{2n}(2)$ (the adjoint Chevally group of type D_n at the prime 2), and this fact therefore also follows from the Solomon–Tits theorem [Solomon 1969].

The number of q-isotropic subspaces U_k of dimension k of V_n is

$$v_{n,k} = 2^{k(k-1)/2} \cdot \frac{(2^{2n-1}+2^{n-1}-1)(2^{2n-3}+2^{n-2}-1)\dots(2^{2n-2k+1}+2^{n-k}-1)}{(2^k-1)(2^{k-1}-1)\dots(2-1)}$$

The Weyl group of U_k is $\widetilde{V}_{n-k}$. Thus, we have proved the following:

Theorem 13. *We have*

$$\widetilde{H\underline{\mathbb{Z}/2}}^G_0(S^{\infty\gamma}) = \mathbb{Z}/2.$$

For $i > 0$,

$$\widetilde{H\underline{\mathbb{Z}/2}}^G_i(S^{\infty\gamma}) = v_{n,k} \bigoplus_{k=0}^{n} H_{i-n+k-1}(\widetilde{V}_{n-k}, \mathscr{H}_{n-k}).$$ □

References

[Adams 1984] J. F. Adams, "Prerequisites (on equivariant stable homotopy) for Carlsson's lecture", pp. 483–532 in *Algebraic topology* (Aarhus, 1982), edited by I. Madsen and B. Oliver, Lecture Notes in Math. **1051**, Springer, 1984. MR Zbl

[Bredon 1967] G. E. Bredon, *Equivariant cohomology theories*, Lecture Notes in Mathematics **34**, Springer, 1967. MR Zbl

[tom Dieck 1970] T. tom Dieck, "Bordism of G-manifolds and integrality theorems", *Topology* **9** (1970), 345–358. MR Zbl

[Greenlees 1985] J. P. C. Greenlees, *Adams spectral sequences in equivariant topology*, Ph.D. thesis, University of Cambridge, 1985.

[Hesselholt and Madsen 1997] L. Hesselholt and I. Madsen, "On the K-theory of finite algebras over Witt vectors of perfect fields", *Topology* **36**:1 (1997), 29–101. MR Zbl

[Hill et al. 2016] M. A. Hill, M. J. Hopkins, and D. C. Ravenel, "On the nonexistence of elements of Kervaire invariant one", *Ann. of Math.* (2) **184**:1 (2016), 1–262. MR Zbl

[Holler and Kriz 2017] J. Holler and I. Kriz, "On RO(G)-graded equivariant "ordinary" cohomology where G is a power of $\mathbb{Z}/2$", *Algebr. Geom. Topol.* **17**:2 (2017), 741–763. MR

[Holler and Kriz 2020] J. Holler and I. Kriz, "On the coefficients of $(\mathbb{Z}/p)^n$-equivariant ordinary cohomology with coefficients in $\mathbb{Z}/p$", 2020. arXiv

[Kriz 2015] S. Kriz, "Equivariant cohomology and the super reciprocal plane of a hyperplane arrangement", preprint, 2015, available at http://www-personal.umich.edu/∼skriz/hyperplanes20027.pdf.

[Lewis et al. 1981] G. Lewis, J. P. May, and J. McClure, "Ordinary $RO(G)$-graded cohomology", *Bull. Amer. Math. Soc.* (*N.S.*) **4**:2 (1981), 208–212. MR Zbl

[Lewis et al. 1986] L. G. Lewis, Jr., J. P. May, M. Steinberger, and J. E. McClure, *Equivariant stable homotopy theory*, Lecture Notes in Mathematics **1213**, Springer, 1986. MR

[Quillen 1971] D. Quillen, "The mod 2 cohomology rings of extra-special 2-groups and the spinor groups", *Math. Ann.* **194** (1971), 197–212. MR Zbl

[Solomon 1969] L. Solomon, "The Steinberg character of a finite group with BN-pair", pp. 213–221 in *Theory of Finite Groups* (Cambridge, MA, 1968), Benjamin, New York, 1969. MR Zbl

Received July 5, 2020. Revised November 15, 2020.

SOPHIE KRIZ
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MICHIGAN
ANN ARBOR, MI
UNITED STATES
skriz@umich.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 2, 2020
https://doi.org/10.2140/pjm.2020.309.401

LOCAL NORMAL FORMS FOR MULTIPLICITY FREE $U(n)$ ACTIONS ON COADJOINT ORBITS

JEREMY LANE

Actions of $U(n)$ on $U(n+1)$ coadjoint orbits via embeddings of $U(n)$ into $U(n+1)$ are an important family of examples of multiplicity free spaces. They are related to Gelfand–Zeitlin completely integrable systems and multiplicity free branching rules in representation theory. This paper computes the Hamiltonian local normal forms of all such actions, at arbitrary points, in arbitrary $U(n+1)$ coadjoint orbits. The results are described using combinatorics of interlacing patterns; gadgets that describe the associated Kirwan polytopes.

1. Introduction

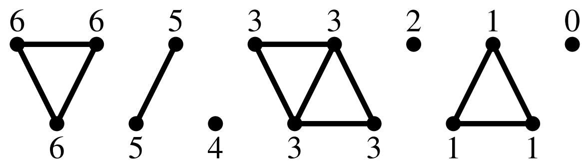
A Hamiltonian action of a compact connected Lie group K on compact symplectic manifold (M, ω) with an equivariant moment map is a *multiplicity free space* if the ring of K-invariant functions $C^\infty(M)^K$ is a commutative Poisson subalgebra [Guillemin and Sternberg 1984a]. The moment map of a multiplicity free space identifies the orbit space, M/K, with a convex polytope called the *Kirwan polytope* [1984]. Compact multiplicity free spaces are classified by their Kirwan polytope and the principal isotropy subgroup of the action [Knop 2011]. The local classification of multiplicity free spaces (in a neighbourhood of an orbit) is a crucial step in the proof of the classification theorem for compact multiplicity free spaces. It is equivalent to the classification of smooth affine spherical varieties for $G = K^{\mathbb{C}}$. Smooth affine spherical varieties are classified by their weight monoids [Losev 2009].

One particularly concrete family of examples of multiplicity free spaces is provided by the action of a unitary group, $U(n)$, on a coadjoint orbit of the unitary group $U(n+1)$ via an embedding of $U(n)$ into $U(n+1)$ (Section 3A). The Kirwan polytopes of these spaces can be described as the set of points $(\mu_1, \dots, \mu_n) \in \mathbb{R}^n$ that satisfy the so-called *interlacing inequalities*,

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \mu_n \geq \lambda_{n+1}, \tag{1}$$

MSC2010: primary 53D20; secondary 14M27, 37J35.
Keywords: coadjoint orbits, multiplicity free spaces, local normal form, Gelfand–Zeitlin, integrable systems.

where $\lambda_1, \ldots, \lambda_{n+1} \in \mathbb{R}$ are fixed parameters determined by the coadjoint orbit. The main result of this paper (Theorem 3.3) is the computation of the local classifying data of these spaces at arbitrary points in arbitrary $U(n+1)$ orbits. This result has two interesting features. First, the classifying data are described in terms of combinatorial gadgets called *interlacing patterns* that encode the combinatorics of the Kirwan polytope (see Section 3B). An example of an interlacing pattern is illustrated below. It corresponds to certain points in $U(8)$ coadjoint orbits diffeomorphic to $U(8)/U(2) \times U(1) \times U(2) \times U(1) \times U(1) \times U(1)$.



The second interesting feature is the proof (given in Section 4). Rather than using the classification of smooth affine spherical varieties, the classifying data are computed directly by elementary means. Following several standard reductions, the main step in this proof is the explicit computation of the isotropy representations (Section 4A). It is shown that they are certain products of standard representations and trivial representations of factors of the isotropy subgroup, which has a block diagonal form. The block diagonal factors of the isotropy subgroup that act by standard representations correspond to "parallelogram shapes" that appear in the interlacing pattern. For example, the isotropy subgroup corresponding to the interlacing pattern above is $U(1) \times U(1) \times 1 \times U(2) \times U(1)$ and the isotropy representation is $\{0\} \oplus \mathbb{C} \oplus \{0\} \oplus \mathbb{C}^2 \oplus \{0\}$ (see Example 5). The computation of this representation relies on the relationship between the combinatorics of interlacing patterns and divisibility properties of characteristic polynomials of certain Hermitian matrices.

Motivation for this work is provided by the *Gelfand–Zeitlin*[1] commutative completely integrable systems [Guillemin and Sternberg 1983]. Although Gelfand–Zeitlin systems have been studied extensively in recent years (see, e.g., [Alekseev et al. 2018; Bouloc et al. 2018; Cho et al. 2020; Lane 2020]), very little is known about their local normal forms as integrable systems near singular fibers (see Example 6). An ongoing program aims to use the results of this paper to prove topological and symplectic local normal forms for Gelfand–Zeitlin systems. The multiplicity free spaces studied in this paper, as well as the associated Gelfand–Zeitlin systems, have analogues for orthogonal groups and orthogonal coadjoint orbits. The local models of those multiplicity free spaces can be computed in a similar fashion.

[1] Also spelled Gelfand–Cetlin and Gelfand–Tsetlin.

2. Hamiltonian group actions and local normal forms

This section fixes conventions, notation, and recalls the statement of the Marle–Guillemin–Sternberg local normal form. Standard references are [Audin 2004; Guillemin and Sternberg 1984c] modulo conventions.

2A. ***Hamiltonian group actions.*** Let K be a connected Lie group. Denote its Lie algebra by $\mathfrak{k}$, the dual vector space by $\mathfrak{k}^*$, and the dual pairing by $\langle\cdot,\cdot\rangle$. Let Ad and Ad^* denote the adjoint and coadjoint actions respectively, i.e., $\langle \mathrm{Ad}^*_k\,\xi, X\rangle = \langle \xi, \mathrm{Ad}_{k^{-1}} X\rangle$ for $k \in K$, $\xi \in \mathfrak{k}^*$, and $X \in \mathfrak{k}$. Given a left action of K on a manifold M, the fundamental vector field of $X \in \mathfrak{k}$ is

$$\underline{X}_p = \frac{d}{dt}\Big|_{t=0} \exp(tX)\cdot p, \qquad p \in M.$$

Let (M, ω) be a symplectic manifold. A left action of K on M is *Hamiltonian* if there exists an equivariant map $\Phi : M \to \mathfrak{k}^*$ such that

$$\iota_{\underline{X}}\omega = d\langle \Phi, X\rangle.$$

A map Φ with this property is called a *moment map*. The tuple (M, ω, Φ) is a *Hamiltonian K-manifold*. Hamiltonian K-manifolds (M, ω, Φ) and (M', ω', Φ') are *isomorphic* if there exists a K-equivariant, symplectic diffeomorphism $\varphi : (M, \omega) \to (M', \omega')$ such that $\Phi' \circ \varphi = \Phi$.

Example 1 (coadjoint orbits). Let $\mathcal{O} \subset \mathfrak{k}^*$ be an orbit of the coadjoint action of K. Given $\xi \in \mathcal{O}$, the tangent space $T_\xi\mathcal{O} \subset \mathfrak{k}^*$ is the set of elements of the form $\mathrm{ad}^*_X\,\xi$, $X \in \mathfrak{k}$. The *Kostant–Kirillov–Souriau* symplectic form ω_{KKS} on $\mathcal{O}$ is defined pointwise by the formula

$$(\omega_{\mathrm{KKS}})_\xi(\mathrm{ad}^*_X\,\xi, \mathrm{ad}^*_Y\,\xi) = \langle \xi, [X, Y]\rangle.$$

The inclusion map $\iota : \mathcal{O} \to \mathfrak{k}^*$ is a moment map for the coadjoint action of K on $(\mathcal{O}, \omega_{\mathrm{KKS}})$. △

Example 2 (homomorphisms). Let (M, ω, Φ) be a Hamiltonian K-manifold, H be a Lie group, and $\varphi : H \to K$ be a Lie group homomorphism. Let $(d\varphi)^* : \mathfrak{k}^* \to \mathfrak{h}^*$ denote the linear map dual to $d\varphi : \mathfrak{h} \to \mathfrak{k}$. Then the action of H on M defined via the action of K and the homomorphism φ is Hamiltonian and $(d\varphi)^* \circ \Phi$ is a moment map. △

Let $U(n)$ denote the group of $n \times n$ unitary matrices, with Lie algebra $\mathfrak{u}(n)$, and let $\mathcal{H}_n$ denote the set of $n \times n$ Hermitian matrices, $X = X^\dagger$, where $X \mapsto X^\dagger$ denotes conjugate transpose. Fix the isomorphism

$$\mathcal{H}_n \to \mathfrak{u}(n)^*, \quad X \mapsto \left(A \mapsto \frac{1}{\sqrt{-1}}\,\mathrm{Tr}(XA)\right). \tag{2}$$

It is equivariant with respect to the action of $U(n)$ on $\mathcal{H}_n$ by conjugation, $k \cdot X = kXk^\dagger$.

Example 3 (representations). Identify $\mathbb{C}^n \cong M_{n\times 1}(\mathbb{C})$. The standard symplectic form on $\mathbb{C}^n$ is

$$\omega_{\mathrm{std}}(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2\sqrt{-1}}(\boldsymbol{x}^\dagger \boldsymbol{y} - \boldsymbol{y}^\dagger \boldsymbol{x}), \quad \boldsymbol{x}, \boldsymbol{y} \in M_{n\times 1}(\mathbb{C}). \tag{3}$$

The action of $U(n)$ on $\mathbb{C}^n$ by the standard representation is Hamiltonian with moment map

$$\Phi(\boldsymbol{x}) = -\frac{1}{2}\boldsymbol{x}\boldsymbol{x}^\dagger. \tag{4}$$

More generally, suppose that V is a real vector space equipped with a linear symplectic form ω_V. Let $\rho : K \to Sp(V, \omega_V)$ be a representation of K on V by symplectic transformations. Then the action of K on (V, ω_V) defined by ρ is Hamiltonian with moment map Φ_V defined by the condition

$$\frac{1}{2}\omega_V(d\rho(X)\boldsymbol{v}, \boldsymbol{v}) = \langle \Phi_V(\boldsymbol{v}), X\rangle \quad \text{for all } \boldsymbol{v} \in V. \tag{5}$$ △

Example 4 (isotropy representations). Let (M, ω, Φ) be a Hamiltonian K-manifold. Given $p \in M$, let $K \cdot p$ denote the orbit of the action of K through p and let $K_p \le K$ denote the *isotropy subgroup*; the subgroup of elements that fix p. Let $K_{\Phi(p)}$ denote the isotropy subgroup of $\Phi(p)$. Then $K_p \le K_{\Phi(p)}$. The *symplectic slice at $p \in M$* is the vector space

$$W_p = T_p(K \cdot p)^\omega / (T_p(K \cdot p) \cap T_p(K \cdot p)^\omega),$$

where $T_p(K \cdot p)^\omega$ denotes the subspace of elements $X \in T_pM$ such that $\omega_p(X, Y) = 0$ for all $Y \in T_p(K \cdot p)$. The restriction of ω_p to $T_p(K \cdot p)^\omega$ descends to a symplectic form on W_p denoted $\overline{\omega}_p$. The linearization of the action of K_p, a.k.a. the *isotropy representation*, preserves the subspaces $T_p(K \cdot p)^\omega$ and $T_p(K \cdot p) \cap T_p(K \cdot p)^\omega$, so it descends to an action of K_p on $(W_p, \overline{\omega}_p)$ by symplectic transformations. Thus $(W_p, \overline{\omega}_p, \Phi_W)$ is a Hamiltonian K_p-manifold, where Φ_W is defined as in Example 3. △

2B. ***Marle–Guillemin–Sternberg local normal forms.*** Given a connected Lie group K, *Marle–Guillemin–Sternberg data* (MGS data) is a tuple (ξ, L, W, ω_W) where $\xi \in \mathfrak{k}^*$, L is a Lie subgroup of K_ξ, and (W, ω_W) is a symplectic vector space equipped with a representation of L by symplectic transformations.

Given MGS data (ξ, L, W, ω_W), [Guillemin and Sternberg 1984c; Marle 1985] construct a Hamiltonian K-manifold, denoted $M(\xi, L, W, \omega_W)$, with the following properties. Let $\mathfrak{m} = \mathfrak{k}_\xi / \mathfrak{l}$ and identify $\mathfrak{m}^*$ with a L-invariant complement of $\mathfrak{l}^*$ in $\mathfrak{k}_\xi^*$.

As a manifold, $M(\xi, L, W, \omega_W)$ is the total space of the vector bundle

(6) $$K \times_L (\mathfrak{m}^* \times W) \to K/L$$

associated to the principal bundle $L \to K \to K/L$ and the representation $\mathfrak{m}^* \times W$. The symplectic structure on $M(\xi, L, W, \omega_W)$ is determined by the data (ξ, L, W, ω_W) (see [Guillemin and Sternberg 1984b; 1984c; Marle 1985] for more details). With respect to this diffeomorphic description of $M(\xi, L, W, \omega_W)$, the Hamiltonian action of K and the corresponding moment map are

(7) $$k' \cdot [k, \eta, w] = [k'k, \eta, w], \qquad \Phi([k, \eta, w]) = \mathrm{Ad}_k^*(\eta + \Phi_W(w) + \xi).$$

Let (M, ω, Φ) be a Hamiltonian K-manifold. The *Marle–Guillemin–Sternberg data of a point* $p \in M$ is $(\Phi(p), K_p, W_p, \overline{\omega}_p)$, where K_p is the isotropy subgroup of p and $(W_p, \overline{\omega}_p)$ is the symplectic slice at p equipped with the isotropy representation of K_p as described in Example 4.

Theorem 2.1 (Marle–Guillemin–Sternberg local normal forms [Guillemin and Sternberg 1984b; Marle 1985]). *Let (M, ω, Φ) be a Hamiltonian K-manifold. For all $p \in M$ there exists K-invariant neighbourhoods $U \subset M$ of the orbit $K \cdot p$ and $U' \subset M(\Phi(p), K_p, W_p, \overline{\omega}_p)$ of the orbit $K \cdot [e, 0, 0]$ and an isomorphism of Hamiltonian K-manifolds $\varphi : U \to U'$ such that $\varphi(p) = [e, 0, 0]$.*

Hamiltonian K-manifolds (M, ω, Φ) and (M', ω', Φ') are *equivalent* if there exists an automorphism ψ of K, a symplectomorphism $F : (M, \omega) \to (M', \omega')$, and an Ad_K^*-fixed element $\xi \in \mathfrak{k}^*$ such that

(1) $\psi(k) \cdot F(m) = F(k \cdot m)$, and

(2) $\Phi + \xi = (d\psi)^* \circ \Phi' \circ F$.

Marle–Guillemin–Sternberg data (ξ, L, W, ω_W) and $(\xi', L', W', \omega_{W'})$ for K are *equivalent* if the corresponding model spaces are equivalent as Hamiltonian K-manifolds. For instance, if p and p' are in the same K-orbit, then the MGS data of p and p' are equivalent.

3. Statement of the main theorem

The following notation will be useful in the remainder of the paper. Given a sequence of real numbers $\underline{\tau} = (\tau_1, \dots, \tau_n)$, let $[\underline{\tau}]$ denote the set of elements in $\underline{\tau}$. Let $\underline{\tau}_i$ denote the i-th element of $[\underline{\tau}]$ in decreasing order. Let $m(\underline{\tau})$ denote the size of $[\underline{\tau}]$. Let $n_\tau(\underline{\tau})$ denote the number of times τ occurs in $\underline{\tau}$. Let $n_i(\underline{\tau})$ denote the number of times $\underline{\tau}_i$ occurs in $\underline{\tau}$.

3A. ***Multiplicity free $U(n)$ actions on $U(n+1)$ coadjoint orbits.*** Given a non-increasing sequence of real numbers $\underline{\lambda} = (\lambda_1, \dots, \lambda_{n+1})$, let $\mathcal{O}_\Lambda$ denote the set of matrices in $\mathcal{H}_{n+1}$ with eigenvalues $\lambda_1, \dots, \lambda_{n+1}$. Then $\mathcal{O}_\Lambda$ is the orbit of

$$\Lambda := \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_{n+1} \end{pmatrix} \tag{8}$$

under the action of $U(n+1)$ by conjugation and the map $k \mapsto k\lambda k^\dagger$ descends to a $U(n+1)$-equivariant diffeomorphism

$$U(n+1)/U(n_1(\underline{\lambda})) \times \cdots \times U(n_{m(\underline{\lambda})}(\underline{\lambda})) \to \mathcal{O}_\Lambda. \tag{9}$$

The map (2) defines a $U(n)$-equivariant diffeomorphism of $\mathcal{O}_\Lambda$ with a coadjoint orbit of $U(n+1)$. Let ω_Λ denote the symplectic form on $\mathcal{O}_\Lambda$ defined by this identification and the Kostant–Kirillov–Souriau symplectic form defined in Example 1. For all $p \in \mathcal{O}_\Lambda$,

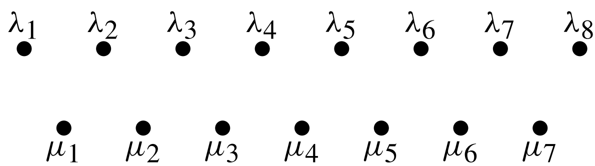
$$(\omega_\Lambda)_p([X, p], [Y, p]) = \frac{1}{\sqrt{-1}} \operatorname{Tr}(p[X, Y]) \quad \text{for all } X, Y \in \mathfrak{u}(n+1). \tag{10}$$

With respect to (2), $(\mathcal{O}_\Lambda, \omega_\Lambda, \iota : \mathcal{O}_\Lambda \to \mathcal{H}_{n+1})$ is a Hamiltonian $U(n+1)$-manifold, where ι denotes inclusion. Let $K = U(n)$ and let $\varphi : K \to U(n+1)$ be an embedding of K as a Lie subgroup of $U(n+1)$. With respect to the identification (2), $(d\varphi)^*$ is a linear projection $\mathcal{H}_{n+1} \to \mathcal{H}_n$. By Example 2, $(\mathcal{O}_\Lambda, \omega_\Lambda, \Phi)$ is a Hamiltonian K-manifold with moment map

$$\Phi = (d\varphi)^* \circ \iota : \mathcal{O}_\Lambda \to \mathcal{H}_n. \tag{11}$$

It is well-known that $(\mathcal{O}_\Lambda, \omega_\Lambda, \Phi)$ are multiplicity free spaces for all possible choices of $\underline{\lambda}$ and φ (this follows from Lemma 4.1 below).

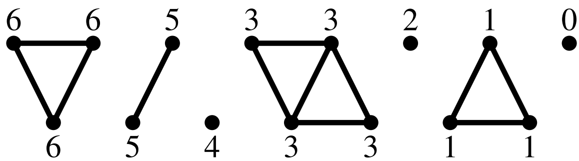
3B. ***Interlacing patterns.*** Let $\underline{\lambda} = (\lambda_1, \dots, \lambda_{n+1})$ and $\underline{\mu} = (\mu_1, \dots, \mu_n)$ be non-increasing sequences of numbers that satisfy the interlacing inequalities (1). The inequalities (1) are represented by attaching labels to a fixed set of $2n+1$ vertices arranged on a triangular grid as illustrated by the following example.

(12)


If a vertex labelled x appears to the left of a vertex labelled y, then $x \geq y$. The labels on the top row correspond to $\underline{\lambda}$ and the labels on the bottom row correspond to $\underline{\mu}$.

The (labelled) *interlacing pattern* of a pair of sequences $(\underline{\lambda}, \underline{\mu})$ that satisfy (1) is the labelled undirected plane graph obtained by adding straight edges to the diagram above according to the following rule: two vertices are connected by an edge if and only if they are nearest neighbours and their labels are equal. For example, the following is the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ where $\underline{\lambda} = (6, 6, 5, 3, 3, 2, 1, 0)$ and $\underline{\mu} = (6, 5, 4, 3, 3, 1, 1)$.

(13)



Three types of connected components can occur in interlacing patterns: ∇*-shapes*, Δ*-shapes*, and ▱*-shapes*. In the example (13): the components labelled 6, 2, and 0 are ∇-shapes, the components labelled 4 and 1 are Δ-shapes, and the components labelled 5 and 3 are ▱-shapes. By convention, an isolated vertex on the top row is a ∇-shape and an isolated vertex on the bottom row is a Δ-shape.

If $\underline{\lambda} = (\lambda_1, \ldots, \lambda_{n+1})$ is fixed, then the set of pairs $(\underline{\lambda}, \underline{\mu})$ that satisfy (1) (equivalently, the set of labelled interlacing patterns whose labels on the top row are given by $\underline{\lambda}$) is in bijection with elements of the polytope

$$\Delta_{\underline{\lambda}} := \{\underline{\mu} = (\mu_1, \ldots, \mu_n) \in \mathbb{R}^n \mid (\underline{\lambda}, \underline{\mu}) \text{ satisfies (1)}\}.$$

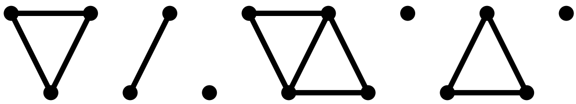
Given $(\mathcal{O}_\Lambda, \omega_\Lambda, \Phi)$ as in the previous section, a point $p \in \mathcal{O}_\Lambda$ determines a pair $(\underline{\lambda}, \underline{\mu})$ that satisfies (1), where $\underline{\mu} = (\mu_1, \ldots, \mu_n)$ denotes the eigenvalues of $\Phi(p)$ arranged in nonincreasing order. Thus, every $p \in \mathcal{O}_\Lambda$ has an associated labelled interlacing pattern. As observed in [Guillemin and Sternberg 1983], the polytope $\Delta_{\underline{\lambda}}$ defined above is the Kirwan polytope of $(\mathcal{O}_\Lambda, \omega_\Lambda, \Phi)$, i.e.,

$$\Delta_{\underline{\lambda}} = \{(\mu_1, \ldots, \mu_n) \in \mathbb{R}^n \mid \mu_1 \geq \cdots \geq \mu_n,\ \exists p \in \mathcal{O}_\Lambda \text{ with eigenvalues } \mu_1, \ldots, \mu_n\}.$$

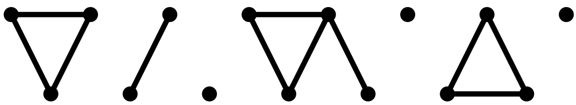
The notation $\underset{\nabla\text{-shape}}{\lambda \in [\underline{\lambda}]}$ denotes the set of all $\lambda \in [\underline{\lambda}]$ such that the connected component of the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ labelled by λ is a ∇-shape. Similar notation is used for other sets. For example, any pair $(\underline{\lambda}, \underline{\mu})$ satisfying (1) satisfies the identity

$$\sum_{i=1}^{n+1} \lambda_i - \sum_{i=1}^{n} \mu_i = \sum_{\substack{\lambda \in [\underline{\lambda}] \\ \nabla\text{-shape}}} \lambda - \sum_{\substack{\mu \in [\underline{\mu}] \\ \Delta\text{-shape}}} \mu. \tag{14}$$

Remark 3.1. An *unlabelled interlacing pattern* is an undirected plane graph that can be obtained from a labelled interlacing pattern by erasing the labels. In other words, the edges in an unlabelled interlacing pattern must correspond to a configuration of equalities and strict inequalities that is allowed by (1). For instance, the following is an unlabelled interlacing pattern.

(15) 

On the other hand, the following is not an unlabelled interlacing pattern.

(16) 

If $\underline{\mu}$ and $\underline{\mu}'$ are contained in the relative interior of the same face of $\Delta_{\underline{\lambda}}$, then the unlabelled interlacing patterns of $(\underline{\lambda}, \underline{\mu})$ and $(\underline{\lambda}, \underline{\mu}')$ are the same. Thus the set of unlabelled interlacing patterns obtained by erasing labels from labelled interlacing patterns of pairs $(\underline{\lambda}, \underline{\mu})$, $\underline{\lambda}$ fixed, is in natural bijection with the set of faces of $\Delta_{\underline{\lambda}}$. The partial order on faces of $\Delta_{\underline{\lambda}}$ corresponds to an obvious partial order on the set of all such unlabelled interlacing patterns. Thus, they encode $\Delta_{\underline{\lambda}}$ as an abstract polytope. It is also straightforward to read the local moment cone of a point $\underline{\mu} \in \Delta_{\underline{\lambda}}$ from the unlabelled interlacing pattern of $(\underline{\lambda}, \underline{\mu})$. The intersection of this local moment cone with the standard lattice in $\mathbb{R}^n$ is the weight monoid of the corresponding smooth affine spherical variety that appears in the classification of [Knop 2011].

Remark 3.2. The interlacing patterns described here occur as rows in larger diagrams, also called interlacing patterns, that describe points and faces of Gelfand–Zeitlin polytopes as well as fibers of Gelfand–Zeitlin systems (see, e.g., [An et al. 2018; Cho et al. 2020; Pabiniak 2014; Bouloc et al. 2018]). Some authors use an equivalent combinatorial gadget called *ladder diagrams* and introduce terminology such as W-blocks, M-blocks, and N-blocks that is equivalent to the notions of $\triangledown$*-shapes*, $\vartriangle$*-shapes*, and $\square$*-shapes* used here.

3C. ***Statement of the main theorem.*** Let $K = U(n)$ and let $(\underline{\lambda}, \underline{\mu})$ be a pair of nonincreasing sequences $\underline{\lambda} = (\lambda_1, \dots, \lambda_{n+1})$ and $\underline{\mu} = (\mu_1, \dots, \mu_n)$ that satisfy the interlacing inequalities (1). Let $\mathrm{M} := \mathrm{diag}(\mu_1, \dots, \mu_n)$. The stabilizer subgroup K_{M} for the conjugation action of K is a block diagonal subgroup isomorphic to $U(n_1(\underline{\mu})) \times \cdots \times U(n_{m(\underline{\mu})}(\underline{\mu}))$. Define

(17)
$$W_{(\underline{\lambda},\underline{\mu})} := \bigoplus_{\substack{\mu \in [\underline{\mu}] \\ \square\text{-shape}}} \mathbb{C}^{n_\mu(\underline{\mu})},$$

and the block-diagonal subgroup

$$L_{(\underline{\lambda},\underline{\mu})} := L_1 \times \cdots \times L_{m(\underline{\mu})} \leq U(n_1(\underline{\mu})) \times \cdots \times U(n_{m(\underline{\mu})}(\underline{\mu})) = K_{\mathrm{M}}, \tag{18}$$

where

$$L_i = \left\{ \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & k \end{array} \right) \mid k \in U(n_i(\underline{\mu}) - 1) \right\} \leq U(n_i(\underline{\mu})) \tag{19}$$

if the component of the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ labelled $\underline{\mu}_i$ is a ⏢-shape, and $L_i = U(n_i(\underline{\mu}))$ otherwise. Equip $W_{(\underline{\lambda},\underline{\mu})}$ with the representation of $L_{(\underline{\lambda},\underline{\mu})}$ where the factor L_i acts by the standard representation on the corresponding factor $\mathbb{C}^{n_i(\underline{\mu})}$ if the component of the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ labelled $\underline{\mu}_i$ is a ▱-shape, and it acts trivially otherwise.

Example 5. Consider the interlacing pattern in equation (13). It follows that $\mathrm{M} = \mathrm{diag}(6, 5, 4, 3, 3, 1, 1)$,

$$L_{(\underline{\lambda},\underline{\mu})} = \left\{ \left(\begin{array}{c|c|c|c|c} k_6 & & & & \\ \hline & k_5 & & & \\ \hline & & 1 & & \\ \hline & & & k_3 & \\ \hline & & & & \begin{matrix} 1 & \\ & k_1 \end{matrix} \end{array} \right) \;\middle|\; k_6, k_5, k_1 \in U(1),\ k_3 \in U(2) \right\} \tag{20}$$

$$W_{(\underline{\lambda},\underline{\mu})} = \{0\} \oplus \mathbb{C} \oplus \{0\} \oplus \mathbb{C}^2 \oplus \{0\}.$$

The representation of $L_{(\underline{\lambda},\underline{\mu})}$ on $W_{(\underline{\lambda},\underline{\mu})}$ is $(k_6, k_5, k_3, k_1)\cdot(z_5, z_3) = (k_5 z_5, k_3 z_3)$. △

For $\mu \in \underline{\mu}$, define $r_\mu \geq 0$ such that

$$r_\mu^2 = -\Bigg(\prod_{\substack{\lambda \in [\underline{\lambda}] \\ \text{▽-shape}}} (\mu - \lambda) \Bigg) \Bigg(\prod_{\substack{\tau \in [\underline{\mu}] \\ \text{⏢-shape} \\ \tau \neq \mu}} \frac{1}{(\mu - \tau)} \Bigg) \tag{21}$$

if the connected component of the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ labelled μ is a ⏢-shape, and $r_\mu = 0$ otherwise. If the connected component of the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ labelled μ is a ⏢-shape, then $r_\mu^2 > 0$.

Provided that the component of the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ labelled $\mu = \underline{\mu}_i$ is not a ⏢-shape, define

$$C_i := C_\mu := \sum_{i=1}^{n+1} \lambda_i - \sum_{i=1}^{n} \mu_i - \mu + \sum_{\substack{\tau \in [\underline{\mu}] \\ \text{⏢-shape}}} \frac{r_\tau^2}{\mu - \tau}. \tag{22}$$

Finally, define a linear symplectic form on $W_{(\underline{\lambda},\underline{\mu})}$ by the formula

(23)
$$\omega_{(\underline{\lambda},\underline{\mu})}(\boldsymbol{u},\boldsymbol{w}) := \frac{1}{\sqrt{-1}} \sum_{\substack{\mu\in[\underline{\mu}] \\ \text{▱-shape}}} \frac{-\boldsymbol{u}_\mu^\dagger \boldsymbol{w}_\mu + \boldsymbol{w}_\mu^\dagger \boldsymbol{u}_\mu}{C_\mu},$$

for all $\boldsymbol{u}, \boldsymbol{w} \in W_{(\underline{\lambda},\underline{\mu})}$, where $\boldsymbol{u}_\mu$ denotes the projection of $\boldsymbol{u}$ to the factor $\mathbb{C}^{n_\mu(\underline{\mu})}$.

Theorem 3.3. *Let $K = U(n)$ and let $(\mathcal{O}_\Lambda, \omega_\Lambda, \Phi)$ be the Hamiltonian K-manifold associated to a nonincreasing sequence $\underline{\lambda} = (\lambda_1, \dots, \lambda_{n+1})$ and an embedding $\varphi : K \to U(n+1)$ as in Section 3A. Then, the Marle–Guillemin–Sternberg local normal form data of $p \in \mathcal{O}_\Lambda$ is equivalent to*

(24)
$$(\mathrm{M}, L_{(\underline{\lambda},\underline{\mu})}, W_{(\underline{\lambda},\underline{\mu})}, \omega_{(\underline{\lambda},\underline{\mu})}),$$

where $(\underline{\lambda}, \underline{\mu})$ is determined by p as in Section 3B and M, $L_{(\underline{\lambda},\underline{\mu})}$, $W_{(\underline{\lambda},\underline{\mu})}$, *and* $\omega_{(\underline{\lambda},\underline{\mu})}$ *are as defined above.*

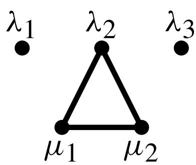
The proof of Theorem 3.3, given in Section 4, describes an explicit linear isomorphism between the isotropy representation at p and the symplectic representation $(W_{(\underline{\lambda},\underline{\mu})}, \omega_{(\underline{\lambda},\underline{\mu})})$.

Remark 3.4. It is straightforward to check that as $L_{(\underline{\lambda},\underline{\mu})}$-representations,

(25)
$$\mathfrak{m}^* \cong \bigoplus_{\substack{\mu\in[\underline{\mu}] \\ \text{⏢-shape}}} (\mathbb{R} \times \mathbb{C}^{n_\mu(\underline{\mu})-1}),$$

where if the component of the interlacing pattern labelled $\underline{\mu}_i$ is a ⏢-shape, then the factor $L_i \cong U(n_i(\underline{\mu}) - 1)$ acts on the corresponding factor $\mathbb{R} \times \mathbb{C}^{n_i(\underline{\mu})-1}$ as the product of the trivial representation and the standard representation. Otherwise the factor L_i acts trivially. The moment map of the local normal form $M(\mathrm{M}, L_{(\underline{\lambda},\underline{\mu})}, W_{(\underline{\lambda},\underline{\mu})}, \omega_{(\underline{\lambda},\underline{\mu})})$ is easily computed by combining Example 3 and (7).

Example 6. Let $\lambda_1 > \lambda_2 > \lambda_3$ and let $p \in \mathcal{O}_\Lambda$ such that the eigenvalues of $\Phi(p)$ are $\mu_1 = \mu_2 = \lambda_2$. The interlacing pattern of p is



It follows from Theorem 3.3 that the orbit through p is a Lagrangian $U(2)/U(1) \cong S^3$ and a neighbourhood of this orbit is isomorphic to a neighbourhood of the zero section in T^*S^3, equipped with the Hamiltonian action of $U(2)$ by cotangent lift of the action of $U(2)$ on S^3. This particular example was derived by Alamiddine

[2009], who used it to show that the Gelfand–Zeitlin systems on regular $U(3)$ coadjoint orbits are isomorphic, in a neighbourhood of this Lagrangian S^3 fiber, to an integrable system for the normalized geodesic flow on T^*S^3 for the round metric on S^3. △

4. Proof of Theorem 3.3

Let $K = U(n)$ and fix an arbitrary nonincreasing sequence $\underline{\lambda} = (\lambda_1, \ldots, \lambda_{n+1})$. Several standard reductions are in order.

First, any two embeddings $K \to U(n+1)$ endow $\mathcal{O}_\Lambda$ with equivalent Hamiltonian K-manifold structures: the restricted coadjoint actions differ by the coadjoint action of an element $g \in U(n+1)$. Thus, it is sufficient to compute the MGS data with respect to the embedding

$$\varphi : K \to U(n+1), \quad k \mapsto \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & k \end{array}\right). \tag{26}$$

With respect to (2),

$$(d\varphi)^* : \mathcal{H}_{n+1} \to \mathcal{H}_n, \quad (d\varphi)^*(X) = X^{(n)}, \tag{27}$$

where $X^{(n)}$ is the bottom right principal $n \times n$ submatrix of X. Thus $\Phi(X) = X^{(n)}$.

Second, it is sufficient to compute the MGS data for points of the form

$$p = \left(\begin{array}{c|c} c & z^\dagger \\ \hline z & \mathrm{M} \end{array}\right) = \left(\begin{array}{c|ccccc} c & \bar{z}_1 & \bar{z}_2 & \cdots & \bar{z}_{n-1} & \bar{z}_n \\ \hline z_1 & \mu_1 & & & & \\ z_2 & & \mu_2 & & & \\ \vdots & & & \ddots & & \\ z_{n-1} & & & & \mu_{n-1} & \\ z_n & & & & & \mu_n \end{array}\right), \tag{28}$$

$$z_i \in \mathbb{C} \quad \text{and} \quad c = \sum_{i=1}^{n+1} \lambda_i - \sum_{i=1}^{n} \mu_i,$$

where $\mu_1 \geq \cdots \geq \mu_n$. Indeed, every point in $\mathcal{O}_\Lambda$ can be brought to this form by the action $U(n)$, so its MGS data is equivalent to the MGS data of a point of this form. Note that $p \in \Phi^{-1}(\mathrm{M})$ if and only if p is of the form (28).

Before giving the final reduction, recall from [Guillemin and Sternberg 1983] that the condition $p \in \mathcal{O}_\Lambda$, for p of the form (28), is equivalent to the following equality of characteristic polynomials,

$$\prod_{i=1}^{n+1}(x - \lambda_i) = (x - c)\prod_{i=1}^{n}(x - \mu_i) - \sum_{i=1}^{n} |z_i|^2 \prod_{\substack{j=1 \\ i \neq j}}^{n}(x - \mu_j). \tag{29}$$

Rewrite p in block form

$$p = \left(\begin{array}{c|c|c|c|c} c & z_1^\dagger & z_2^\dagger & \cdots & z_m^\dagger \\ \hline z_1 & \underline{\mu}_1 I_{n_1(\underline{\mu})} & & & \\ \hline z_2 & & \underline{\mu}_2 I_{n_2(\underline{\mu})} & & \\ \hline \vdots & & & \ddots & \\ \hline z_m & & & & \underline{\mu}_m I_{n_m(\underline{\mu})} \end{array}\right), \quad z_i \in M_{n_i(\underline{\mu})\times 1}(\mathbb{C}). \tag{30}$$

where $m = m(\underline{\mu})$. If $\mu = \underline{\mu}_i$, let $z_\mu = z_i$ denote the corresponding block. Then (29) becomes

$$\prod_{\lambda\in[\underline{\lambda}]} (x-\lambda)^{n_\lambda(\underline{\lambda})} = (x-c)\prod_{\mu\in[\underline{\mu}]}(x-\mu)^{n_\mu(\underline{\mu})} - \sum_{\mu\in[\underline{\mu}]} \|z_\mu\|^2 (x-\mu)^{n_\mu(\underline{\mu})-1} \prod_{\substack{\tau\in[\underline{\mu}]\\ \tau\neq\mu}} (x-\tau)^{n_\tau(\underline{\tau})}. \tag{31}$$

The following lemma is well-known. Its proof is left as an exercise using the fact that $p \in \mathcal{O}_\Lambda$ if and only if p satisfies (31).

Lemma 4.1. *Let p be of the form (30). Then $p \in \mathcal{O}_\Lambda$ if and only if for all $\mu \in \underline{\mu}$, $\|z_\mu\|^2 = r_\mu^2$. Moreover, the action of K_{M} on $\Phi^{-1}(\mathrm{M})$ is transitive.*

The final reduction concerns the isotropy subgroup. Given $(\underline{\lambda}, \underline{\mu})$, define $\tilde{p} \in \mathcal{O}_\Lambda$ of the form (30) such that for all $\mu \in [\underline{\mu}]$,

$$z_\mu = \begin{pmatrix} r_\mu \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{32}$$

By construction, $K_{\tilde{p}} = L_{(\underline{\lambda},\underline{\mu})}$. The MGS data of every other point $p \in \Phi^{-1}(\mathrm{M})$ is equivalent to that of $\tilde{p}$ by Lemma 4.1.

Remark 4.2. Many of the facts mentioned in this section are also useful for studying Gelfand–Zeitlin systems [Guillemin and Sternberg 1983; Cho et al. 2020].

4A. ***The isotropy representation.*** Continuing from the previous section, this section computes the isotropy representations at the points $\tilde{p} \in \Phi^{-1}(\mathrm{M})$ as described in (30), (32) and Lemma 4.1.

Lemma 4.3. *Let $p \in \Phi^{-1}(\mathrm{M})$ and let c, z be defined as in (30). The subspace $T_p(K\cdot p)^\omega$ consists of all matrices of the form*

$$\left(\begin{array}{c|c} 0 & (c-\mathrm{M})\boldsymbol{x}^\dagger + z^\dagger X^\dagger \\ \hline (c-\mathrm{M})\boldsymbol{x} + Xz & 0 \end{array}\right), \quad X \in \mathfrak{k},\ \boldsymbol{x} \in M_{n\times 1}(\mathbb{C}) \tag{33}$$

such that

$$
\begin{aligned}
0 &= \boldsymbol{x}^\dagger \boldsymbol{z} + \boldsymbol{z}^\dagger \boldsymbol{x}, \\
0 &= \boldsymbol{x}\boldsymbol{z}^\dagger + \boldsymbol{z}\boldsymbol{x}^\dagger + [X, \mathrm{M}].
\end{aligned} \tag{34}
$$

The subspace $T_p(K \cdot p) \cap T_p(K \cdot p)^\omega$ consists of all matrices of the form

$$
\left(\begin{array}{c|c} 0 & \boldsymbol{z}^\dagger Y^\dagger \\ \hline Y\boldsymbol{z} & 0 \end{array}\right), \quad Y \in \mathfrak{k}_{\mathrm{M}}. \tag{35}
$$

Proof. Denote

$$
\eta := \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & Y \end{array}\right), \quad \xi := \left(\begin{array}{c|c} x_0 & -\boldsymbol{x}^\dagger \\ \hline \boldsymbol{x} & X \end{array}\right), \quad X, Y \in \mathfrak{k},\ x_0 \in \sqrt{-1}\mathbb{R},\ \boldsymbol{x} \in M_{n\times 1}(\mathbb{C}).
$$

The tangent space $T_p\mathcal{O}_\Lambda$ consists of elements of the form $[\xi, p]$. Since diagonal elements of $\mathfrak{u}(n+1)$ act trivially, set $x_0 = 0$. Then elements of $T_p\mathcal{O}_\Lambda$ have block form

$$
[\xi, p] = \left(\begin{array}{c|c} -\boldsymbol{x}^\dagger \boldsymbol{z} - \boldsymbol{z}^\dagger \boldsymbol{x} & (c-\mathrm{M})\boldsymbol{x}^\dagger + \boldsymbol{z}^\dagger X^\dagger \\ \hline (c-\mathrm{M})\boldsymbol{x} + X\boldsymbol{z} & \boldsymbol{x}\boldsymbol{z}^\dagger + \boldsymbol{z}\boldsymbol{x}^\dagger + [X, \mathrm{M}] \end{array}\right), \quad X \in \mathfrak{k},\ \boldsymbol{x} \in M_{n\times 1}(\mathbb{C}).
$$

Elements of $T_p(K \cdot p)$ have block form

$$
[\eta, p] = \left(\begin{array}{c|c} 0 & \boldsymbol{z}^\dagger Y^\dagger \\ \hline Y\boldsymbol{z} & [Y, \mathrm{M}] \end{array}\right), \quad Y \in \mathfrak{k}.
$$

Recall,

$$
T_p(K \cdot p)^\omega = \left\{[\xi, p] \in T_p\mathcal{O}_\Lambda \mid (\omega_\Lambda)_p([\xi, p], [\eta, p]) = 0 \text{ for all } Y \in \mathfrak{k}\right\}.
$$

By (10),

$$
\begin{aligned}
\sqrt{-1}(\omega_\Lambda)_p([\xi, p], [\eta, p]) &= \mathrm{Tr}(p[\xi, \eta]) \\
&= -\mathrm{Tr}(\boldsymbol{z}^\dagger Y \boldsymbol{x}) - \mathrm{Tr}(\boldsymbol{z}\boldsymbol{x}^\dagger Y) + \mathrm{Tr}(\mathrm{M}[X, Y]) \\
&= \mathrm{Tr}(([\mathrm{M}, X] - \boldsymbol{x}\boldsymbol{z}^\dagger - \boldsymbol{z}\boldsymbol{x}^\dagger)Y).
\end{aligned}
$$

Let $\sqrt{-1}E_{i,i}$, $E_{i,j} - E_{j,i}$, and $\sqrt{-1}(E_{i,j} + E_{j,i})$ be standard basis elements for $\mathfrak{k}$ (where $E_{i,j}$ denotes the matrix whose i, j-entry is 1 and all other entries are 0). Plugging these elements in for Y yields a system of equations,

$$
\begin{aligned}
0 &= x_i\bar{z}_i + z_i\bar{x}_i && \text{for all } i, \\
0 &= (\mu_j - \mu_i)(X_{j,i} + X_{i,j}) - (x_j\bar{z}_i + z_j\bar{x}_i - x_i\bar{z}_j - z_i\bar{x}_j) && \text{for all } i \neq j, \\
0 &= (\mu_j - \mu_i)(X_{j,i} - X_{i,j}) - (x_j\bar{z}_i + z_j\bar{x}_i + x_i\bar{z}_j + z_i\bar{x}_j) && \text{for all } i \neq j,
\end{aligned} \tag{36}
$$

(where $X_{i,j}$ denotes the i, j entry of X) which in turn is equivalent to the system of equations

$$
\begin{aligned}
0 &= x_i\bar{z}_i + z_i\bar{x}_i && \text{for all } i, \\
0 &= (\mu_j - \mu_i)X_{j,i} - (x_j\bar{z}_i + z_j\bar{x}_i) && \text{for all } i \neq j.
\end{aligned} \tag{37}
$$

This system of equations is equivalent to the system of matrix equations (34). It follows from (34) that the block diagonal parts of $[\xi, p] \in T_p(K \cdot p)^{\omega}$ are zero, so $[\xi, p]$ has the form (33) subject to the equations (34). By properties of equivariant moment maps, $T_p(K \cdot p) \cap T_p(K \cdot p)^{\omega} = T_p(K_{\mathrm{M}} \cdot p)$ [Guillemin and Sternberg 1984c]. Elements of $T_p(K_{\mathrm{M}} \cdot p)$ have block form of (35), which completes the proof. □

Equations (34) dictate the form of the vectors $(c - \mathrm{M})\boldsymbol{x} + X\boldsymbol{z}$, as the next two lemmas demonstrate.

Lemma 4.4. *Let $p \in \Phi^{-1}(\mathrm{M})$ and let $\boldsymbol{z}$ be defined as in (30). Let $X \in \mathfrak{k}$ and $\boldsymbol{x} \in M_{n \times 1}(\mathbb{C})$ such that*

$$0 = \boldsymbol{x}\boldsymbol{z}^{\dagger} + \boldsymbol{z}\boldsymbol{x}^{\dagger} + [X, \mathrm{M}]. \tag{38}$$

If the component of the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ labelled μ is not a $\triangle$-shape, then

$$(X\boldsymbol{z})_{\mu} = \Bigg(\sum_{\substack{\tau \in [\underline{\mu}] \\ \triangle\text{-shape}}} \frac{r_{\tau}^2}{\mu - \tau} \Bigg) \boldsymbol{x}_{\mu}.$$

Proof. Let $\mu \neq \nu$ distinct elements of $\underline{\mu}$. Let $X_{\mu,\nu}$, $\boldsymbol{x}_{\mu}$, $\boldsymbol{z}_{\mu}$, etc. denote the corresponding blocks of X, $\boldsymbol{x}$, and $\boldsymbol{z}$. By (38), the μ, ν block of X is given by the formula

$$X_{\mu,\nu} = \frac{1}{\mu - \nu} (\boldsymbol{x}_{\mu}\boldsymbol{z}_{\nu}^{\dagger} + \boldsymbol{z}_{\mu}\boldsymbol{x}_{\nu}^{\dagger}) \quad \text{for all } \mu \neq \nu.$$

By Lemma 4.1, if the component of the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ labelled μ is not a $\triangle$-shape, then $\boldsymbol{z}_{\mu} = 0$. Thus

$$\begin{aligned} (X\boldsymbol{z})_{\mu} &= \sum_{\substack{\tau \in [\underline{\mu}] \\ \tau \neq \mu}} X_{\mu,\tau}\boldsymbol{z}_{\tau} = \sum_{\substack{\tau \in [\underline{\mu}] \\ \tau \neq \mu}} \frac{1}{\mu - \tau} \boldsymbol{x}_{\mu}\boldsymbol{z}_{\tau}^{\dagger}\boldsymbol{z}_{\tau} \\ &= \Bigg(\sum_{\substack{\tau \in [\underline{\mu}] \\ \triangle\text{-shape}}} \frac{\|\boldsymbol{z}_{\tau}\|^2}{\mu - \tau} \Bigg) \boldsymbol{x}_{\mu} = \Bigg(\sum_{\substack{\tau \in [\underline{\mu}] \\ \triangle\text{-shape}}} \frac{r_{\tau}^2}{\mu - \tau} \Bigg) \boldsymbol{x}_{\mu}. \end{aligned}$$

□

Recall the definition of C_{μ} from (22).

Lemma 4.5. *Let p, X, and $\boldsymbol{x}$ as in Lemma 4.4 such that (38) holds. Assume that the component of the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ labelled μ is not a $\triangle$-shape. Then, $C_{\mu} = 0$ if and only if the component of the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ labelled μ is a $\bigtriangledown$-shape.*

Proof. First, note that it is sufficient to prove

$$\prod_{\substack{\lambda\in[\underline{\lambda}]\\ \text{▽-shape}}} (x-\lambda) = (x-c)\prod_{\substack{\mu\in[\underline{\mu}]\\ \text{△-shape}}} (x-\mu) - \sum_{\substack{\mu\in[\underline{\mu}]\\ \text{△-shape}}} r_\mu^2 \prod_{\substack{\tau\in[\underline{\mu}]\\ \text{△-shape}\\ \tau\neq\mu}} (x-\tau). \tag{39}$$

Indeed, since the component of the interlacing pattern labelled μ is not a △-shape, plugging in $x=\mu$ yields

$$\prod_{\substack{\lambda\in[\underline{\lambda}]\\ \text{▽-shape}}} (\mu-\lambda) = \left(\mu - c - \sum_{\substack{\tau\in[\underline{\mu}]\\ \text{△-shape}}} \frac{r_\tau^2}{\mu-\tau}\right) \prod_{\substack{\tau\in[\underline{\mu}]\\ \text{△-shape}}} (\mu-\tau) = -C_\mu \prod_{\substack{\tau\in[\underline{\mu}]\\ \text{△-shape}}} (\mu-\tau) \tag{40}$$

and the factor

$$\prod_{\substack{\tau\in[\underline{\mu}]\\ \text{△-shape}}} (\mu-\tau) \tag{41}$$

is nonzero.

Second, applying Lemma 4.1 ($r_\mu = 0$ when the component labelled μ is not a △-shape) and rearranging, observe that

$$\begin{aligned}
&(x-c)\prod_{\mu\in[\underline{\mu}]} (x-\mu)^{n_\mu([\underline{\mu}])} - \sum_{\mu\in[\underline{\mu}]} r_\mu^2 (x-\mu)^{n_\mu([\underline{\mu}])-1} \prod_{\substack{\tau\in[\underline{\mu}]\\ \tau\neq\mu}} (x-\tau)^{n_\tau([\underline{\mu}])} \\
&= (x-c)\prod_{\substack{\mu\in[\underline{\mu}]\\ \text{△-shape}}} (x-\mu)^{n_\mu([\underline{\mu}])} \prod_{\substack{\tau\in[\underline{\mu}]\\ \text{▽, ▱-shape}}} (x-\tau)^{n_\tau([\underline{\mu}])} \\
&\quad - \sum_{\substack{\mu\in[\underline{\mu}]\\ \text{△-shape}}} r_\mu^2 (x-\mu)^{n_\mu([\underline{\mu}])-1} \prod_{\substack{\tau\in[\underline{\mu}]\\ \text{△-shape}\\ \tau\neq\mu}} (x-\tau)^{n_\tau([\underline{\mu}])} \prod_{\substack{\tau\in[\underline{\mu}]\\ \text{▽, ▱-shape}}} (x-\tau)^{n_\tau([\underline{\mu}])} \\
&= \left((x-c)\prod_{\substack{\mu\in[\underline{\mu}]\\ \text{△-shape}}} (x-\mu) - \sum_{\substack{\mu\in[\underline{\mu}]\\ \text{△-shape}}} r_\mu^2 \prod_{\substack{\tau\in[\underline{\mu}]\\ \text{△-shape}\\ \tau\neq\mu}} (x-\tau)\right) \\
&\qquad \cdot \prod_{\substack{\tau\in[\underline{\mu}]\\ \text{△-shape}}} (x-\tau)^{n_\tau([\underline{\mu}])-1} \prod_{\substack{\tau\in[\underline{\mu}]\\ \text{▽, ▱-shape}}} (x-\tau)^{n_\tau([\underline{\mu}])}.
\end{aligned} \tag{42}$$

Then (39) follows by combining (42) and (31), which completes the proof. □

For $p \in \Phi^{-1}(\mathrm{M})$, let $V_p \subset \mathbb{C}^n$ denote the image of injective linear map

$$T : T_p(K \cdot p)^{\omega} \to \mathbb{C}^n,$$

$$\left(\begin{array}{c|c} 0 & (c-\mathrm{M})\boldsymbol{x}^{\dagger} + \boldsymbol{z}^{\dagger} X^{\dagger} \\ \hline (c-\mathrm{M})\boldsymbol{x} + X\boldsymbol{z} & 0 \end{array}\right) \mapsto (c-\mathrm{M})\boldsymbol{x} + X\boldsymbol{z} \tag{43}$$

and let $U_p \subset V_p$ denote the image of $T_p(K \cdot p) \cap T_p(K \cdot p)^{\omega}$. Specialize to the case of $\tilde{p}$ and recall that $K_{\tilde{p}} = L_{(\underline{\lambda},\underline{\mu})}$. The map T is $K_{\tilde{p}}$-equivariant with respect to the action of $K_{\tilde{p}}$ on $\mathbb{C}^n$ as a block-diagonal subgroup of $K = U(n)$ acting by the standard representation. Decompose $\mathbb{C}^n = \bigoplus_{i=1}^m \mathbb{C}^{n_i(\underline{\mu})}$, $m = m(\underline{\mu})$. The subspaces $V_{\tilde{p}}$ and $U_{\tilde{p}}$ have the forms $\bigoplus_{i=1}^m V_i$ and $\bigoplus_{i=1}^m U_i$, respectively, for some subspaces $U_i \subset V_i \subset \mathbb{C}^{n_i(\underline{\mu})}$. The map T descends to an isomorphism of $K_{\tilde{p}}$-representations,

$$W_{\tilde{p}} = T_{\tilde{p}}(K \cdot \tilde{p})^{\omega} / (T_{\tilde{p}}(K \cdot \tilde{p}) \cap T_{\tilde{p}}(K \cdot \tilde{p})^{\omega}) \cong \bigoplus_{i=1}^{m} V_i/U_i. \tag{44}$$

The representation of $K_{\tilde{p}} = L_1 \times \cdots \times L_m$ on the right is given in each component by the inclusion $L_i \subset U(n_i(\underline{\mu}))$ and the standard representation of $U(n_i(\underline{\mu}))$ on $\mathbb{C}^{n_i(\underline{\mu})}$. This representation of L_i preserves the subspaces $U_i \subset V_i$ so it induces a representation on V_i/U_i.

Recall that if the component of the interlacing pattern labelled $\underline{\mu}_i$ is a $\square$-shape, then $L_i = U(n_i(\underline{\mu}))$.

Proposition 4.6. *For all $i = 1, \ldots, m$, $m = m(\underline{\mu})$, there is an isomorphism of L_i representations*

$$V_i/U_i \cong \begin{cases} \mathbb{C}^{n_i(\underline{\mu})} & \text{if the component of the interlacing pattern} \\ & \quad \text{of } (\underline{\lambda}, \underline{\mu}) \text{ labelled } \underline{\mu}_i \text{ is a } \square\text{-shape,} \\ \{0\} & \text{otherwise,} \end{cases}$$

where $\mathbb{C}^{n_i(\underline{\mu})}$ denotes the standard representation of $U(n_i(\underline{\mu}))$.

Proof. In general,

$$U_i = \{(Y\boldsymbol{z})_i \mid Y \in \mathfrak{k}_{\mathrm{M}}\} = \{Y_{i,i}\boldsymbol{z}_i \mid Y_{i,i} \in \mathfrak{u}(n_i(\underline{\mu}))\}.$$

If the component of the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ labelled $\underline{\mu}_i$ is a $\triangle$-shape, then, by Lemma 4.1, $\boldsymbol{z}_i \neq 0$, so $U_i = \mathbb{C}^{n_i(\underline{\mu})}$ and $V_i/U_i \cong \{0\}$. If the component of the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ labelled $\underline{\mu}_i$ is not a $\triangle$-shape, then, $\boldsymbol{z}_i = 0$, so $U_i = \{0\}$.

It remains to determine the subspace V_i when the component of the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ labelled $\underline{\mu}_i$ is not a $\triangle$-shape. In this case, it follows by Lemma 4.4 that the block

$$((c-\mathrm{M})\boldsymbol{x} + X\boldsymbol{z})_i = (c-\mathrm{M})\boldsymbol{x}_i + (X\boldsymbol{z})_i = C_i\boldsymbol{x}_i,$$

where $C_i = C_{\underline{\mu}_i}$ as defined in (22). By Lemma 4.3,

$$V_i = \{((c-\mathrm{M})\boldsymbol{x} + X\boldsymbol{z})_i \mid X \in \mathfrak{k}, \boldsymbol{x} \in \mathbb{C}^n, \boldsymbol{x}\boldsymbol{z}^\dagger + \boldsymbol{z}\boldsymbol{x}^\dagger + [X, \mathrm{M}]\} = \{C_i \boldsymbol{x}_i \mid \boldsymbol{x}_i \in \mathbb{C}^{n_i(\underline{\mu})}\}. \tag{45}$$

By Lemma 4.5, $C_i = 0$ if and only if the component of the interlacing pattern of $(\underline{\lambda}, \underline{\mu})$ labelled $\underline{\mu}_i$ is a $\triangledown$-shape. This completes the proof. □

Thus $\bigoplus_{i=1}^m V_i/U_i$ is isomorphic to the $L_{(\underline{\lambda},\underline{\mu})}$-representation $W_{(\underline{\lambda},\underline{\mu})}$.

Proposition 4.7. *The linear symplectic structure on $W_{(\underline{\lambda},\underline{\mu})}$ defined via the symplectic form $\overline{\omega}_{\tilde{p}}$ and the isomorphism (44) equals the linear symplectic form $\omega_{(\underline{\lambda},\underline{\mu})}$ defined in (23).*

Proof. Denote

$$\eta := \left(\begin{array}{c|c} 0 & -\boldsymbol{y}^\dagger \\ \hline \boldsymbol{y} & Y \end{array}\right), \quad \xi := \left(\begin{array}{c|c} 0 & -\boldsymbol{x}^\dagger \\ \hline \boldsymbol{x} & X \end{array}\right), \quad X, Y \in \mathfrak{k},\ \boldsymbol{x}, \boldsymbol{y} \in M_{n\times 1}(\mathbb{C}).$$

Then, using Lemma 4.4,

$$\begin{aligned} &\sqrt{-1}(\omega_\Lambda)_{\tilde{p}}([\xi, \tilde{p}], [\eta, \tilde{p}]) \\ &= \mathrm{Tr}(\tilde{p}[\xi, \eta]) = \mathrm{Tr}([\tilde{p}, \xi]\eta) \\ &= -\mathrm{Tr}\left(\left(\begin{array}{c|c} 0 & (c-\mathrm{M})\boldsymbol{x}^\dagger + \boldsymbol{z}^\dagger X^\dagger \\ \hline (c-\mathrm{M})\boldsymbol{x} + X\boldsymbol{z} & 0 \end{array}\right)\left(\begin{array}{c|c} 0 & -\boldsymbol{y}^\dagger \\ \hline \boldsymbol{y} & Y \end{array}\right)\right) \\ &= -((c-\mathrm{M})\boldsymbol{x}^\dagger + \boldsymbol{z}^\dagger X^\dagger)\boldsymbol{y} + \mathrm{Tr}(((c-\mathrm{M})\boldsymbol{x} + X\boldsymbol{z})\boldsymbol{y}^\dagger) \\ &= -((c-\mathrm{M})\boldsymbol{x}^\dagger + \boldsymbol{z}^\dagger X^\dagger)\boldsymbol{y} + \mathrm{Tr}(\boldsymbol{y}^\dagger((c-\mathrm{M})\boldsymbol{x} + X\boldsymbol{z})) \\ &= -(c-\mathrm{M})(\boldsymbol{x}^\dagger\boldsymbol{y} - \boldsymbol{y}^\dagger\boldsymbol{x}) - \boldsymbol{z}^\dagger X^\dagger \boldsymbol{y} + \boldsymbol{y}^\dagger X\boldsymbol{z} \\ &= -(c-\mathrm{M})(\boldsymbol{x}^\dagger\boldsymbol{y} - \boldsymbol{y}^\dagger\boldsymbol{x}) - (X\boldsymbol{z})^\dagger \boldsymbol{y} + \boldsymbol{y}^\dagger X\boldsymbol{z} \\ &= -(c-\mathrm{M})(\boldsymbol{x}^\dagger\boldsymbol{y} - \boldsymbol{y}^\dagger\boldsymbol{x}) + \sum_{i=1}^m \left(\sum_{\substack{\triangle\text{-shape} \\ j\neq i}} \frac{r_j^2}{\mu_i - \mu_j}\right)(-\boldsymbol{x}_i^\dagger\boldsymbol{y}_i + \boldsymbol{y}_i^\dagger\boldsymbol{x}_i). \end{aligned} \tag{46}$$

Viewing $[\xi, \tilde{p}]$ and $[\eta, \tilde{p}]$ as representatives of vectors in the isotropy representation,

$$\begin{aligned} (\overline{\omega}_\lambda)_{\tilde{p}}([\xi, \tilde{p}], [\eta, \tilde{p}]) &= \frac{1}{\sqrt{-1}} \sum_{\substack{\square\text{-shape} \\ i=1}}^m \left(c - \mu_i + \sum_{\substack{\triangle\text{-shape} \\ j\neq i}} \frac{r_j^2}{\mu_i - \mu_j}\right)(-\boldsymbol{x}_i^\dagger\boldsymbol{y}_i + \boldsymbol{y}_i^\dagger\boldsymbol{x}_i) \\ &= \frac{1}{\sqrt{-1}} \sum_{\substack{\square\text{-shape} \\ i=1}}^m C_i(-\boldsymbol{x}_i^\dagger\boldsymbol{y}_i + \boldsymbol{y}_i^\dagger\boldsymbol{x}_i). \end{aligned} \tag{47}$$

Applying the isomorphism $T : W_{\tilde{p}} \to W_{(\underline{\lambda},\underline{\mu})}$, $[\xi, p] \mapsto \boldsymbol{u} = (C_i \boldsymbol{x}_i)_i$, $[\eta, p] \mapsto \boldsymbol{v} = (C_i \boldsymbol{y}_i)_i$ yields

$$\omega_{(\underline{\lambda},\underline{\mu})}(\boldsymbol{u}, \boldsymbol{w}) = \frac{1}{\sqrt{-1}} \sum_{\substack{\square\text{-shape} \\ i=1}}^{m} \frac{-\boldsymbol{u}_i^\dagger \boldsymbol{w}_i + \boldsymbol{w}_i^\dagger \boldsymbol{u}_i}{C_i}. \qquad \square$$

Acknowledgements

The author would like to thank Yael Karshon who some years ago provided him with notes from a lecture on Gelfand–Zeitlin systems by N.T. Zung that inspired this paper. The author would also like to thank the Fields Institute and the organizers of the thematic program on Toric Topology and Polyhedral Products for the support of a Fields Postdoctoral Fellowship during writing of this paper.

References

[Alamiddine 2009] I. Alamiddine, *Géométrie de systèmes Hamiltoniens intégrables : le cas du système de Gelfand–Ceitlin*, Ph.D. Thesis, Université Toulouse, 2009, available at http://www.theses.fr/2009TOU30073.

[Alekseev et al. 2018] A. Alekseev, J. Lane, and Y. Li, "The $U(n)$ Gelfand–Zeitlin system as a tropical limit of Ginzburg–Weinstein diffeomorphisms", *Philos. Trans. Roy. Soc. A* **376**:2131 (2018), 20170428, 20. MR Zbl

[An et al. 2018] B. H. An, Y. Cho, and J. S. Kim, "On the f-vectors of Gelfand–Cetlin polytopes", *European J. Combin.* **67** (2018), 61–77. MR Zbl

[Audin 2004] M. Audin, *Torus actions on symplectic manifolds*, revised ed., Progress in Mathematics **93**, Birkhäuser, Basel, 2004. MR Zbl

[Bouloc et al. 2018] D. Bouloc, E. Miranda, and N. T. Zung, "Singular fibres of the Gelfand–Cetlin system on $\mathfrak{u}(n)^*$", *Philos. Trans. Roy. Soc. A* **376**:2131 (2018), 20170423, 28. MR

[Cho et al. 2020] Y. Cho, Y. Kim, and Y.-G. Oh, "Lagrangian fibers of Gelfand–Cetlin systems", *Adv. Math.* **372** (2020), 107304, 57. MR Zbl

[Guillemin and Sternberg 1983] V. Guillemin and S. Sternberg, "The Gelfand–Cetlin system and quantization of the complex flag manifolds", *J. Funct. Anal.* **52**:1 (1983), 106–128. MR

[Guillemin and Sternberg 1984a] V. Guillemin and S. Sternberg, "Multiplicity-free spaces", *J. Differential Geom.* **19**:1 (1984), 31–56. MR Zbl

[Guillemin and Sternberg 1984b] V. Guillemin and S. Sternberg, "A normal form for the moment map", pp. 161–175 in *Differential geometric methods in mathematical physics* (Jerusalem, 1982), edited by S. Sternberg, Math. Phys. Stud. **6**, Reidel, Dordrecht, 1984. MR Zbl

[Guillemin and Sternberg 1984c] V. Guillemin and S. Sternberg, *Symplectic techniques in physics*, Cambridge University Press, 1984. MR Zbl

[Kirwan 1984] F. Kirwan, "Convexity properties of the moment mapping, III", *Invent. Math.* **77**:3 (1984), 547–552. MR Zbl

[Knop 2011] F. Knop, "Automorphisms of multiplicity free Hamiltonian manifolds", *J. Amer. Math. Soc.* **24**:2 (2011), 567–601. MR Zbl

[Lane 2020] J. Lane, "The geometric structure of symplectic contraction", *Int. Math. Res. Not.* **2020**:12 (2020), 3521–3539. MR Zbl

[Losev 2009] I. V. Losev, "Proof of the Knop conjecture", *Ann. Inst. Fourier* (*Grenoble*) **59**:3 (2009), 1105–1134. MR Zbl

[Marle 1985] C.-M. Marle, "Modèle d'action hamiltonienne d'un groupe de Lie sur une variété symplectique", *Rend. Sem. Mat. Univ. Politec. Torino* **43**:2 (1985), 227–251. MR Zbl

[Pabiniak 2014] M. Pabiniak, "Gromov width of non-regular coadjoint orbits of $U(n)$, $SO(2n)$ and $SO(2n+1)$", *Math. Res. Lett.* **21**:1 (2014), 187–205. MR Zbl

Received February 21, 2020. Revised October 1, 2020.

JEREMY LANE
DEPARTMENT OF MATHEMATICS
MCMASTER UNIVERSITY
HAMILTON, ON
CANADA
lanej5@math.mcmaster.ca

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 2, 2020
https://doi.org/10.2140/pjm.2020.309.421

FUNCTIONAL DETERMINANT ON PSEUDO-EINSTEIN 3-MANIFOLDS

ALI MAALAOUI

Given a three-dimensional pseudo-Einstein CR manifold $(M, T^{1,0}M, \theta)$, we establish an expression for the difference of determinants of the Paneitz type operators A_θ, related to the problem of prescribing the Q'-curvature, under the conformal change $\theta \mapsto e^w\theta$ with $w \in \mathcal{P}$ the space of pluriharmonic functions. This generalizes the expression of the functional determinant in four-dimensional Riemannian manifolds established in (*Proc. Amer. Math. Soc.* 113:3 (1991), 669–682). We also provide an existence result of maximizers for the scaling invariant functional determinant as in (*Ann. of Math.* (2) 142:1 (1995), 171–212).

1. Introduction and statement of the results

There has been extensive work on the study of spectral invariants of differential operators defined on a Riemannian manifold (M, g) and the relations to their conformal invariants; see for instance [Branson and Ørsted 1986; 1991a; 1991b]. As an example, if we consider the two-dimensional case with the pair of the Laplace operator $-\Delta_g$, and the associated invariant which is the scalar curvature R_g, we know that under conformal change of the metric $g \mapsto \tilde{g} = e^{2w}g$, one has the relation

$$R_{\tilde{g}}e^{2w} = -\Delta_g w + R_g.$$

It is also known that the spectrum of $-\Delta_g$ is discrete and can be written as $0 < \lambda_1 \leq \lambda_2 \leq \cdots$ and the corresponding zeta function is then defined by

$$\zeta_{-\Delta_g}(s) = \sum_{j=1}^{\infty} \frac{1}{\lambda_j^s}.$$

This series converges uniformly for $s > 1$ and can be extended to a meromorphic function in $\mathbb{C}$ with 0 as a regular value. The determinant of the operator $-\Delta_g$ can then be written as

$$\det(-\Delta_g) = e^{-\zeta'_{-\Delta_g}(0)}.$$

MSC2020: primary 58J35, 58J50; secondary 32V05, 32V20.
Keywords: pseudo-Einstein CR manifolds, functional determinant, the P'-operator.

The celebrated Polyakov formula [1981a; 1981b] states that if $\tilde{g} = e^{2w} g$, then

$$\ln\left(\frac{\det(-\Delta_{\tilde{g}})}{\det(-\Delta_g)}\right) = \frac{-1}{12\pi} \int_M |\nabla w|^2 + 2Rw \, dv_g,$$

for metrics with the same volume. The scaling invariant functional determinant F_2 can then be written as

$$F_2(w) = \frac{-1}{12\pi}\left(\int_M |\nabla w|^2 + 2Rw \, dv_g - \left(\int_M R \, dv_g\right) \ln\left(\fint_M e^{2w} \, dv_g\right)\right).$$

Notice that the right-hand side is a familiar quantity. It is the Beckner–Onofri energy [Beckner 1993], and we know that

$$\int_{S^2} |\nabla w|^2 + 2Rw \, dv_g - \left(\int_M R \, dv_g\right) \ln\left(\fint_{S^2} e^{2w} \, dv_g\right) \geq 0.$$

This notion of determinant was extended to dimension four for conformally invariant operators, keeping in mind that the substitute of the pair $(-\Delta_g, R)$ in dimension two is the pair (P_g, Q_g) in dimension four, where P_g is the Paneitz operator and Q_g is the Riemannian Q-curvature [Branson and Ørsted 1991b; Esposito and Malchiodi 2019]. In addition, two new terms appear in the scaling invariant functional determinant expression. Indeed, if (M, g) is a 4-dimensional manifold and A_g is a nonnegative self-adjoint conformally covariant operator, then there exist β_1, β_2 and $\beta_3 \in \mathbb{R}$ such that the scaling invariant functional determinant F_4 reads as

$$F_4(w) = \beta_1 I(w) + \beta_2 II(w) + \beta_3 III(w), \tag{1}$$

where

$$\begin{cases} I(w) = 4\int_M w|W_g|^2 \, dv_g - \left(\int_M |W_g|^2 \, dv_g\right) \ln\left(\fint_M e^{4w} \, dv_g\right), \\ II(w) = \int_M w P_g w + 4Q_g w \, dv_g - \left(\int_M Q_g \, dv_g\right) \ln\left(\fint_M e^{4w} \, dv_g\right), \\ III(w) = 12\int_M (\Delta_g w + |\nabla w|^2)^2 \, dv_g - 4\int_M w \Delta_g R_g + R_g |\nabla w|^2 \, dv_g. \end{cases}$$

In the case of the sphere S^4, we see that the second term II corresponds again to the four-dimensional Beckner–Onofri energy. The existence and uniqueness of maximizers of this expression was heavily investigated and we refer the reader to [Chang and Yang 1995; Gursky and Malchiodi 2012; Esposito and Malchiodi 2019; Okikiolu 2001] for an in-depth study of this functional in the Riemannian setting.

Now let us move to the CR setting. We consider a 3-dimensional CR manifold $(M, T^{1,0}M, J, \theta)$ and we recall that the substitute for the pair (P_g, Q_g) is (P_θ, Q_θ), where P_θ is the CR Paneitz operator and Q_θ is the CR Q-curvature [Fefferman and Hirachi 2003; Gover and Graham 2005]. The problem with this pair is that the total Q-curvature is always zero. In fact in pseudo-Einstein manifolds the Q-curvature vanishes identically. Hence, we do not have a decent substitute for the CR Beckner–Onofri inequality. Fortunately, if we restrict our study to pseudo-Einstein manifolds and variations in the space of pluriharmonic functions $\mathcal{P}$, then we have a better substitute for the pair (P_g, Q_g), namely (P'_θ, Q'_θ). These quantities were first introduced on odd dimensional spheres in [Branson et al. 2013] and then on pseudo-Einstein manifolds in [Case et al. 2016; Case and Yang 2013; Hirachi 2014]. In particular one has a Beckner–Onofri type inequality involving the operator P'_θ acting on pluriharmonic functions as proved in [Branson et al. 2013]. We also recall that the total Q'-curvature corresponds to a geometric invariant, namely the Burns–Epstein invariant $\mu(M)$ [Burns and Epstein 1988; Chêng and Lee 1990].

One is tempted to see what the spectral invariants of the operator P' are or the restriction of P' to the space $\mathcal{P}$ of pluri-harmonic functions and link them to geometric quantities such as the total Q'-curvature.

We recall that the quantity Q'_θ changes as follows: if $\tilde{\theta} = e^w\theta$ with $w \in \mathcal{P}$, then

$$P'_\theta w + Q'_\theta = Q'_{\tilde{\theta}} e^{2w} + \tfrac{1}{2} P_\theta(w^2), \tag{2}$$

which we can write as

$$P'_\theta w + Q'_\theta = Q'_{\tilde{\theta}} e^{2w} \mod \mathcal{P}^\perp.$$

We let $\tau_\theta : L^2 \to \mathcal{P}$ be the orthogonal projection on $\mathcal{P}$ with respect to the inner product induced by θ and set $A_\theta = \tau_\theta P'_\theta \tau_\theta$. Then equation (2) can be rewritten as

$$A_\theta w + \tau_\theta(Q'_\theta) = \tau_\theta(Q'_{\tilde{\theta}} e^{2w}).$$

Prescribing the quantity $\bar{Q}'_\theta = \tau_\theta(Q'_\theta)$ was thoroughly investigated in [Maalaoui 2019b; Case et al. 2016; Ho 2019] mainly because of the property that

$$\int_M \bar{Q}'_\theta \, dv_\theta = \int_M Q'_\theta \, dv_\theta = -\frac{\mu(M)}{16\pi^2}.$$

We recall that in [Maalaoui 2019a], we proved that the dual of the Beckner–Onofri inequality, namely the logarithmic Hardy–Littlewood–Sobolev inequality, can be linked to the regularized zeta function of the operator A_θ evaluated at one. This was proved in the Riemannian setting in [Morpurgo 1996; 2002; Okikiolu 2008].

In this paper, we will generalize the expression (1) by studying the determinant of the operator A_θ. In all that follows we assume that $(M, T^{1,0}M, J, \theta)$ is an embeddable pseudo-Einstein manifold such that P'_θ is nonnegative and has trivial kernel. First we show that:

Theorem 1.1 (conformal index). *Let ζ_{A_θ} be the spectral zeta function of the operator A_θ. Then $\zeta_{A_\theta}(0)$ is a conformal invariant in $\mathcal{P}$. Moreover,*

$$\zeta_{A_\theta}(0) = \frac{-1}{24\pi^2} \int_M Q'_\theta \, dv_\theta - 1.$$

In order to compute the determinant of the operator A_θ we introduce the quantities $\widetilde{A}_1(w)$, $\widetilde{A}_2(w)$ and $\widetilde{A}_3(w)$ defined by

$$
\begin{cases}
\widetilde{A}_1(w) := \int_M wA_\theta w + 2Q'_\theta w\, d\nu - \frac{1}{c_1}\ln\big(⨍_M e^{2w}\, d\nu\big),\\
\widetilde{A}_2(w) := 2\int_M R\big(\Delta_b w + \frac{1}{2}|\nabla_b w|^2\big) - \big(\Delta_b w + \frac{1}{2}|\nabla_b w|^2\big)^2\, d\nu,\\
\widetilde{A}_3(w) := 2\int_M w_{,0}R - \frac{1}{3}w_{,0}|\nabla_b w|^2 - w_{,0}\Delta_b w\, d\nu.
\end{cases}
\tag{3}
$$

One can also write $\widetilde{A}_2(w)$ as

$$
\widetilde{A}_2(w) = 2\int_M R\left(\frac{\Delta_b e^{\frac{1}{2}w}}{2e^{\frac{1}{2}w}}\right) - \left(\frac{\Delta_b e^{\frac{1}{2}w}}{2e^{\frac{1}{2}w}}\right)^2 d\nu.
$$

Then we have the following:

Theorem 1.2. *There exists c_2 and $c_3 \in \mathbb{R}$ such that for all $w \in \mathcal{P}$, we have*

$$
\ln\left(\frac{\det(A_\theta)}{\det(A_{e^w\theta})}\right) = -\frac{1}{24\pi^2}\widetilde{A}_1(w) + c_2\widetilde{A}_2(w) - c_3\widetilde{A}_3(w).
\tag{4}
$$

Notice that the expression (4) is not scaling invariant, because for $\tilde{\theta} = c^2\theta$, with c a positive constant, we have

$$
\det(A_{\tilde{\theta}}) = c^{-4\zeta_\theta(0)}\det(A_\theta).
$$

So we fix the volume V of (M,θ) and define the scaling invariant functional

$$
S_{A_\theta} = \left(\frac{\mathrm{Vol}(\theta)}{V}\right)^{\zeta_{A_\theta}(0)} \det(A).
$$

Now we can define the scaling invariant functional $F : W_H^{2,2}(M) \cap \mathcal{P} \to \mathbb{R}$ by

$$
F(w) = \ln(S_{A_\theta}) - \ln(S_{A_{e^w\theta}}),
$$

where $W_H^{2,2}(M)$ is Folland–Stein space. Then one can write the following expression of F,

$$
F(w) = c_1 II(w) + c_2 III(w) + c_3 IV(w),
$$

where

$$
\begin{cases}
II(w) = \int_M wA_\theta w + 2Q'_\theta w\, d\nu - \int_M Q'_\theta\, d\nu \ln\big(⨍ e^{2w} d\nu\big).\\
III(w) = \widetilde{A}_2(w).\\
IV(w) = -\widetilde{A}_3(w).
\end{cases}
\tag{5}
$$

Notice that the functional II is the CR Beckner–Onofri functional studied first in [Branson et al. 2013] on the standard sphere S^3. In particular one has on the standard sphere

$$
II(w) \geq 0.
$$

This functional was also investigated in [Case et al. 2016] and its critical points correspond to the pseudo-Einstein structures with constant $\overline{Q'}$-curvature. The functional III is also similar to the Riemannian one defined in [Chang and Yang 1995] and its critical points are pseudo-Einstein contact forms $\tilde{\theta}$ satisfying

$$\tilde{\tau}(\widetilde{\Delta_b}\widetilde{R}) = 0.$$

The functional IV is a bit different, in fact if we let $\mathcal{H}$ defined by

$$\mathcal{H}(w) = R_{,0} - \tfrac{1}{3}|\nabla_b w|^2_{,0} - \tfrac{2}{3}\operatorname{div}_b(w_{,0}\nabla_b w) + \Delta_b w_{,0} - (\Delta_b w)_{,0},$$

then the critical points of IV satisfy

$$\tilde{\tau}(e^{-2w}\mathcal{H}(w)) = 0.$$

We set

$$a = \frac{\int_M Q'_\theta \, d\nu}{16\pi^2}.$$

Since the coefficients c_2 and c_3 are still unknown for the operator A_θ and the corresponding functional S_{A_θ}, we will be considering them as parameters in our setting. Then we show the following for the functional F:

Theorem 1.3. *Assume that $c_2 > 0$ and $c_3 \geq 0$. Then there exists a constant μ depending on θ, such that if*

$$c_3 < \mu\left(\sqrt{25c_2^2 + \frac{1}{3\pi^2}c_2(1-a)} - 5c_2\right), \tag{6}$$

then F has a maximizer $w_\infty \in W^{2,2}_H(M) \cap \mathcal{P}$ under the constraint $\int_M e^{2w}\, d\nu = 1$. Moreover, this maximizer satisfies the Euler–Lagrange equation

$$\tau_{\theta_\infty}\left[-\frac{1}{24\pi^2}\widetilde{Q'_\theta} + c_2\widetilde{\Delta_b}\widetilde{R} + c_3 e^{-2w}\mathcal{H}(w)\right] = cte,$$

where the tilde refers to quantities computed using the contact form $\theta_\infty = e^{w_\infty}\theta$.

Notice the condition (6) implies in particular that $\int_M Q'_\theta d\nu < 16\pi^2$. Hence, as a consequence, we have that $(M, T^{1,0}M, \theta)$ is not equivalent to the standard sphere. We point out that in order to verify the sharpness of (6) one needs to check specific examples which is not as easy as in the Riemannian case, since we are dealing with CR pluriharmonic functions and we lack explicit examples of manifolds where one can have an explicit expression of the spectrum of the P'-operators.

2. Heat coefficients and conformal invariance

Let $(M, T^{1,0}M, \theta)$ be a pseudo-Einstein 3-manifold and P'_θ its P'-operator defined by

$$P'_\theta f = 4\Delta_b^2 f - 8\operatorname{Im}\big(\nabla^1(A_{11}\nabla^1 f)\big) - 4\operatorname{Re}\big(\nabla^1(R\nabla_1 f)\big). \tag{7}$$

Denote by $\tau : L^2(M) \to \mathcal{P}$ the orthogonal projection on the space of pluriharmonic functions with respect to the L^2-inner product induced by θ. We consider the operator $A_\theta = \tau P'_\theta \tau$ and for the conformal change $\tilde{\theta} = e^w \theta$, with $w \in \mathcal{P}$, we let

$$A_{\tilde{\theta}} = \tau_{\tilde{\theta}}(e^{-2w} A_\theta),$$

where $\tau_{\tilde{\theta}}$ is the orthogonal projection with respect to the L^2-inner product induced by $\tilde{\theta}$.

In order to evaluate and manipulate the spectral invariants, we need to study the expression of the heat kernel of the operator A_θ. Unfortunately, this operator is not elliptic or subelliptic (as an operator on $C^\infty(M)$), and does not have an invertible principal symbol in the sense of $\Psi_H(M)$-calculus (see [Ponge 2007]). In fact A_θ can be seen as a Toeplitz operator, and one might adopt the approach introduced in [Boutet de Monvel and Guillemin 1981] in order to study it. But instead, we will modify the operator in order to be able to use the classical computations done for the heat kernel.

Consider the operator $\mathcal{L} = A_\theta + \tau^\perp L \tau^\perp$, where L is chosen so that $\mathcal{L}$ has an invertible principal symbol in $\Psi^4_H(M)$. Notice that $\tau\mathcal{L} = \mathcal{L}\tau = A_\theta$. Based on [Ponge 2007], if $\mathcal{K}$ is the heat kernel of $\mathcal{L}$ one has the following expansion near zero:

$$\mathcal{K}(t, x, x) \sim \sum_{j=0}^{\infty} \tilde{a}_j(x) t^{\frac{1}{4}(j-4)} + \ln(t) \sum_{j=1}^{\infty} t^j \tilde{b}_j(x).$$

Since $e^{-t\mathcal{L}} = e^{-tA_\theta}\tau + e^{-tL}\tau^\perp$, we have that the kernel K of e^{-tA_θ} which is the restriction to $\mathcal{P}$ of $\mathcal{K}$, reads as

$$K(t, x, x) \sim \sum_{j=0}^{\infty} t^{\frac{1}{4}(j-4)} a_j(x) + \ln(t) \sum_{j=1}^{\infty} t^j b_j(x), \tag{8}$$

and this will be the main expansion that we will be using for the rest of the paper.

Now we want to define the infinitesimal variation of a quantity under a conformal change. Fix $w \in \mathcal{P}$ and for a given quantity F_θ depending on θ denote $\delta F_\theta := \frac{d}{dr}_{|r=0} F_{e^{rw}\theta}$. Next, we define the zeta function of A_θ by

$$\zeta_{A_\theta}(s) := \sum_{j=1}^{\infty} \frac{1}{\lambda_j^s},$$

where $0 < \lambda_1 \leq \lambda_2 \leq \cdots$ is the spectrum of the operator $A_\theta : W^{2,2}(M) \cap \mathcal{P} \to \mathcal{P}$. In what follows $\mathrm{TR}[A]$ is to be understood as the trace of the operator A in $\mathcal{P}$. Then we have the following proposition.

Proposition 2.1. *With the notations above, we have*

$$\zeta_{A_\theta}(0)=\int_M a_4(x)dx-1.$$

Moreover,

$$\delta\zeta_{A_\theta}(0)=0 \quad \textit{and} \quad \delta\zeta'_{A_\theta}(0)=2\int_M w\Big(a_4(x)-\frac{1}{V}\Big)\,d\nu,$$

where $V=\int_M d\nu_\theta$ is the volume of M.

Proof. Most of the computations in this part are relatively standard and they can be found in [Branson and Ørsted 1986; 1991a; 1991b] in the Riemannian setting. First we use the Mellin transform and (8) to write

$$\begin{aligned}
\zeta_{A_\theta}(s)&=\frac{1}{\Gamma(s)}\int_0^\infty t^{s-1}(\mathrm{TR}[e^{-tA_\theta}]-1)\,dt\\
&=\frac{1}{\Gamma(s)}\Big(-\frac{1}{s}+\int_0^1 t^{s-1}\sum_{j=0}^N t^{\frac14(j-4)}\int_M a_j(x)\,d\nu\,dt\\
&\qquad+\int_M t^{s-1}O(t^{\frac14(N+1-4)})\,dt+\sum_{j=1}^N\int_0^1 t^{j+s-1}\ln(t)\int_M b_j\,d\nu\,dt\\
&\qquad\qquad+\int_0^1 t^{s-1}O(t^{N+1}\ln(t))\,dt+\int_1^\infty t^{s-1}\sum_{j=1}^\infty e^{-\lambda_j t}\,dt\Big)\\
&=\frac{1}{\Gamma(s)}\Big(\frac{-1}{s}+\sum_{j=0}^N\frac{1}{s+\frac14(j-4)}\int_M a_j(x)\,d\nu+\int_0^1 t^{s-1}O(t^{\frac14(N+1-4)})\,dt\\
&\qquad+\sum_{j=1}^N\frac{1}{(s+j)^2}\int_M b_j\,d\nu+\int_0^1 t^{s-1}O(t^{N+1}\ln(t))\,dt\int_1^\infty t^{s-1}\sum_{j=1}^\infty e^{-\lambda_j t}\,dt\Big).
\end{aligned}$$

Since, Γ has a simple pole at $s=0$ with residue 1, we see that by taking $s\to 0$, there are only two terms that survive, leading to

$$\zeta_{A_\theta}(0)=\int_M a_4(x)\,d\nu-1.$$

Next we move to the study of the variation of ζ_{A_θ}. Let $f\in C^\infty(M)$ and $v\in\mathcal{P}$. Then

$$\int_M \tau_{rw}(f)v\,d\nu_{rw}=\int_M f v e^{2rw}\,d\nu.$$

Differentiating with respect to r and evaluating at 0 yields

$$\int_M(\delta\tau(f)+2w\tau(f)-2wF)v\,d\nu=0.$$

Hence,

$$\delta\tau(f) = 2\tau(wf - w\tau(f)).$$

If we let M_w be the multiplication by w, then

$$\delta\tau = 2(\tau M_w - \tau M_w \tau).$$

In particular, if $f \in \mathcal{P}$, then $\delta\tau(f) = 0$.

Next we want to evaluate δA_θ. Recall that $A_{e^{rw}\theta} = \tau_{e^{rw}\theta} e^{-2rw} A_\theta$. Therefore,

$$\delta A_\theta = \delta\tau A_\theta - 2\tau M_w A_\theta = -2\tau M_w A_\theta.$$

Thus,

$$\begin{aligned}\delta\,\mathrm{TR}[e^{-tA_\theta}] &= -t\,\mathrm{TR}[\delta A_\theta e^{-tA_\theta}]\\ &= 2t\,\mathrm{TR}[\tau M_w A_\theta e^{-tA_\theta}]\\ &= 2t\,\mathrm{TR}[M_w A_\theta e^{-tA_\theta}].\end{aligned}$$

The last equality follows from $\mathrm{TR}[AB] = \mathrm{TR}[BA]$ and $e^{-tA_\theta}\tau = e^{-tA_\theta}$. But

$$-tM_w A_\theta e^{-tA_\theta} = \frac{d}{d\varepsilon}_{|\varepsilon=0} M_w e^{-t(1+\varepsilon)A_\theta}.$$

Using the expansion (8), we have

$$K(t(1+\varepsilon), x, x) \sim \sum_{j=0}^{\infty} (1+\varepsilon)^{\frac{1}{4}(j-4)} t^{\frac{1}{4}(j-4)} a_j(x) + H,$$

where H is the logarithmic part. Hence, comparing the terms in the expansion after integration, we get

$$\delta \int_M a_j\,dv = \frac{4-j}{2} \int_M w a_j\,dv.$$

In particular, we have $\delta \int_M a_4\,dv = 0$.

Similarly,

$$\Gamma(s)\zeta_{A_\theta}(s) = \Gamma(s)\big(\zeta_{A_\theta}(0) + s\zeta'_{A_\theta}(0) + O(s^2)\big).$$

Hence, since $\delta\zeta_{A_\theta}(0) = 0$, and $s\Gamma(s) \sim 0$ when $s \to 0$, we have

$$\delta\zeta'_{A_\theta}(0) = [\Gamma(s)\delta\zeta_{A_\theta}(s)]_{s=0}.$$

But,

$$\Gamma(s)\delta\zeta_{A_\theta}(s) = \int_0^\infty 2t^s \operatorname{TR}[wA_\theta e^{-tA_\theta}]\,dt \tag{9}$$
$$= -\int_0^\infty 2t^s \frac{d}{dt}\operatorname{TR}[we^{-tA_\theta}]\,dt$$
$$= \int_0^\infty 2st^{s-1}\operatorname{TR}\Big[w\Big(e^{-tA_\theta}-\frac{1}{V}\Big)\Big]\,dt.$$

Using again the expansion (8) and a similar computation as in the previous case yields

$$\delta\zeta'_{A_\theta}(0) = 2\int_M w\Big(a_4 - \frac{1}{V}\Big)\,dv.$$ □

Proposition 2.2. *There exists* $c \neq 0$ *such that*

$$\zeta_{A_\theta}(0) = c\int_M Q'_\theta\,dv - 1.$$

Moreover $c = -\dfrac{1}{24\pi^2}$.

Proof. First we notice that a_4 is a pseudo-Hermitian invariant of order -2, that is

$$a_{4,e^r\theta} = e^{-2r}a_{4,\theta}$$

for all $r \in \mathbb{R}$. So from [Hirachi 2014], we have the existence of $c_1, c_2, c_3, c_4, c_5 \in \mathbb{R}$ such that

$$a_4 = c_1 Q'_\theta + c_2\Delta_b R + c_3 R_{,0} + c_4 R^2 + c_5 Q_\theta,$$

where $Q'_\theta = 2\Delta_b R - 4|A|^2 + R^2$ and $Q_\theta = -\frac{2}{3}\Delta_b R + 2\operatorname{Im}(A_{11,\bar{1}\bar{1}})$. Since we are in a pseudo-Einstein manifold and $w \in \mathcal{P}$ we can assume that $Q_\theta = 0$. So after integration, we have

$$\int_M a_4\,dv = c_1\int_M Q'_\theta\,dv + c_4\int_M R^2\,dv.$$

Since $\int_M a_4\,dv$ is invariant under the conformal change $e^w\theta$, it is easy to see that $c_4 = 0$. Hence,

$$\int_M a_4\,dv = c_1\int_M Q'_\theta\,dv.$$

Next we want to calculate c_1 (compare to [Stanton 1989], where the invariant k_2 is always 0). We take the case of the sphere S^3. Based on the computations in [Branson et al. 2013], we have

$$\zeta_{A_\theta}(s) = 2\sum_{j=1}^\infty \frac{j+1}{(j(j+1))^s} = 2\sum_{j=2}^\infty \frac{1}{j^{2s-1}}\left(\frac{1}{1-\frac{1}{j}}\right)^s.$$

Using the expansion of

$$\left(1-\frac{1}{j}\right)^{-s}=1+\frac{s}{j}+\frac{s(s+1)}{2j^2}+sO\left(\frac{1}{j^3}\right),$$

we see that

$$\zeta_{A_\theta}(s)=2\left(\zeta_R(2s-1)-1+s(\zeta_R(2s)-1)-\frac{s(s+1)}{2}(\zeta_R(2s+1)-1)\right)+sH(s),$$

with $H(s)$ holomorphic near $s=0$ and ζ_R the classical Riemann Zeta function. Now we recall that ζ_R is regular at $s=0$ and $s=-1$ but has a simple pole at $s=1$ with residue equal to 1. Hence

$$\zeta_{A_\theta}(0)=2\left(-\frac{1}{12}-1+\frac{1}{4}\right)=-\frac{5}{3}\neq 0.$$

Knowing that $\int_{S^3} Q'_\theta\, dv=16\pi^2$, we have

$$16\pi^2 c-1=-\frac{5}{3}.$$

Thus,

$$c=-\frac{1}{24\pi^2}.$$

□

3. The expression for the determinant

Recall that in the previous section, we found that $a_4=c_1Q'+c_2\Delta_b R+c_3R_{,0}$. In particular,

$$\begin{aligned}\delta\zeta'_{A_\theta}(0)&=\int_M 2w\left(a_4(x)-\frac{1}{V}\right)dv\\&=c_1\int_M 2w\left(Q'_\theta-\frac{1}{c_1V}\right)dv+c_2\int_M 2R\Delta_b w\,dv-c_3\int_M 2w_{,0}R\,dv\\&=c_1A_1+c_2A_2+c_3A_3.\end{aligned}$$

We will calculate the change of each term under conformal change of θ. The easiest term to handle is the first one. Indeed, recall that if $\tilde\theta=e^w\theta$ then

$$\begin{aligned}\widetilde{Q'_\theta}e^{2w}&=P'_\theta w+Q'_\theta \mod \mathcal{P}^\perp,\\ \widetilde{R}&=[R-|\nabla_b w|^2-2\Delta_b w]e^{-w},\\ \widetilde{\Delta}_b f&=e^{-w}[\Delta_b f+\nabla_b f\cdot\nabla_b w].\end{aligned}$$

So if $\hat\theta_u=e^{uw}\theta$, we have

$$\int_M 2w\left[\widehat{Q'_\theta}-\frac{1}{c_1}\frac{1}{\widehat{V}}\right]d\hat v=\int_M 2uwP'_\theta w+2Q'w-\frac{1}{c_1}\frac{2we^{2uw}}{\int_M e^{2uw}\,dv}\,dv.$$

Integrating u in [0, 1] yields

$$\widetilde{A}_1(w) = \int_M w A_\theta w + Q'_\theta w \, d\nu - \frac{1}{c_1} \ln\left(\fint_M e^{2w} \, d\nu \right).$$

For the second term, we have

$$\begin{aligned}\int_M \widehat{R}\hat{\Delta}_b w \, d\hat{\nu} &= \int_M \left[R - u^2 |\nabla_b w|^2 - 2u \Delta_b w \right] \left[\Delta_b w + u |\nabla_b w|^2 \right] d\nu \\ &= \int_M R \Delta_b w - u^2 |\nabla_b w|^2 \Delta_b w - 2u (\Delta_b w)^2 + Ru |\nabla_b w|^2 \\ &\qquad\qquad - u^3 |\nabla_b w|^4 - 2u^2 |\nabla_b w|^2 \Delta_b w \, d\nu.\end{aligned}$$

In particular after integrating over u between 0 and 1, we get

$$\begin{aligned}\widetilde{A}_2(w) &= 2 \int_M R \Delta_b w - |\nabla_b w|^2 \Delta_b w - (\Delta_b w)^2 + \frac{1}{2} R |\nabla_b w|^2 - \frac{1}{4} |\nabla_b w|^4 \, d\nu \\ &= 2 \int_M R \Delta_b w - \left(\Delta_b w + \frac{1}{2} |\nabla_b w|^2 \right)^2 + \frac{R}{2} |\nabla_b w|^2 \, d\nu.\end{aligned}$$

Next we compute

$$\int_M \widehat{T} w \widehat{R} d\hat{\nu} = \int_M \left[w_{,0} R - u^2 w_{,0} |\nabla_b w|^2 - 2u w_{,0} \Delta_b w \right] d\nu,$$

where T is the characteristic vector field of θ and we are adopting the notation $Tf = f_{,0}$. Integrating as above yields

$$\widetilde{A}_3(w) = 2 \int_M w_{,0} R - \frac{1}{3} w_{,0} |\nabla_b w|^2 - w_{.0} \Delta_b w \, d\nu.$$

Therefore, one has

$$\zeta'_{\widetilde{A}_\theta}(0) - \zeta'_{A_\theta}(0) = c_1 \widetilde{A}_1(w) + c_2 \widetilde{A}_2(w) - c_3 \widetilde{A}_3(w)$$

or equivalently

$$\ln\left(\frac{\det(A_\theta)}{\det(A_{\tilde{\theta}})} \right) = c_1 \widetilde{A}_1(w) + c_2 \widetilde{A}_2(w) - c_3 \widetilde{A}_3(w). \qquad \square$$

4. Scaling invariant functional

We focus now on the study of the functional F defined by

$$F(w) = c_1 II(w) + c_2 III(w) + c_3 IV(w),$$

where $c_1 = -1/(24\pi^2)$. For the sake of notation, we will keep using c_1 instead of its numerical value. We will also be using constants C_k depending on M and θ.

We recall first that there exists a constant C such that

$$\frac{1}{16\pi^2}\int_M wA_\theta w+2Q'w\,d\nu-\ln\Bigl(\fint e^{2w}\,d\nu\Bigr)\geq C. \tag{10}$$

In fact this follows from the CR Beckner–Onofri inequality proved in [Branson et al. 2013] and also treated in [Case and Yang 2013]. Since the functional F is scaling invariant (that is, $F(w+c)=F(w)$), we can assume without loss of generality that $\bar{w}=\int_M w\,d\nu=0$. Also, we recall that $a_4=c_1Q'_\theta+c_2\Delta_bR+c_3R_{,0}$. Therefore,

$$\int_M a_4w\,d\nu=\int_M c_1Q'_\theta w+c_2R\Delta_bw-c_3Rw_{,0}\,d\nu.$$

Hence, we can write F as

$$\begin{aligned}F(w)=2\int_M a_4w\,d\nu+c_1\Bigl(\int_M wA_\theta w\,d\nu-\int_M Q'_\theta\,d\nu\ln\Bigl(\fint e^{2w}\Bigr)\Bigr)\\ -2c_2\int_M\Bigl(\Delta_bw+\frac{1}{2}|\nabla_bw|^2\Bigr)^2d\nu+c_2\int_M R|\nabla_bw|^2\,d\nu\\ +2c_3\int_M w_{,0}\Bigl(\frac{1}{3}|\nabla_bw|^2+\Delta_bw\Bigr)\,d\nu.\end{aligned}$$

Using (10), we have

$$\int_M Q'_\theta\,d\nu\ln\Bigl(\fint e^{2w}\,d\nu\Bigr)\leq\frac{\int_M Q'_\theta\,d\nu}{16\pi^2}\Bigl[\int_M wA_\theta w+2Q'_\theta w\,d\nu\Bigr]-C.$$

Therefore, for

$$a=\frac{\int_M Q'_\theta\,d\nu}{16\pi^2}$$

and since $c_1<0$, we have

$$\begin{aligned}&c_1\Bigl(\int_M wA_\theta w\,d\nu-\int_M Q'_\theta\,d\nu\ln\Bigl(\fint e^{2w}\Bigr)\Bigr)\\ &\leq c_1\Bigl[(1-a)\int_M wA_\theta w\,d\nu+2a\int Q'_\theta w\,d\nu\Bigr]+C_1\\ &\leq 4c_1(1-a)\int_M(\Delta_bw)^2\,d\nu+C_2\int_M|\nabla_bw|^2\,d\nu+2ac_1\int Q'_\theta w\,d\nu+C_3,\end{aligned} \tag{11}$$

where in the second line we used the expression (7). On the other hand, for the mixed term of $III(w)$, we have for every $\alpha>0$,

$$2\int_M\Delta_bw|\nabla_bw|^2\,d\nu\leq\alpha\int_M(\Delta_bw)^2+\frac{1}{\alpha}\int_M|\nabla_bw|^4\,d\nu. \tag{12}$$

Next, we let λ denote the best constant appearing in the estimate

$$\|f_{,0}\|_{L^2}\leq\lambda\|\Delta_bf\|_{L^2}.$$

Then we have

$$
\begin{aligned}
(13)\quad 2\int_M w_{,0}\Big(\frac{1}{3}|\nabla_b w|^2+\Delta_b w\Big)\,dv &\le 2\Big(\lambda\|\Delta_b w\|_{L^2}\Big\|\frac{1}{3}|\nabla_b w|^2+\Delta_b w\Big\|_{L^2}\Big)\\
&\le 2\lambda\Big(\|\Delta_b w\|_{L^2}^2+\frac{1}{3}\|\Delta_b w\|_{L^2}\big\||\nabla_b w|^2\big\|_{L^2}\Big)\\
&\le \lambda\Big(\Big(2+\frac{\alpha}{3}\Big)\|\Delta_b w\|_{L^2}^2+\frac{1}{3\alpha}\int_M|\nabla_b w|^4\,dv\Big).
\end{aligned}
$$

Hence, combining (11), (12) and (13) and assuming that $c_2>0$ and $c_3\ge 0$, we get

$$
\begin{aligned}
(14)\quad & c_1 II(w)+c_2 III(w)+c_3 IV(w)\\
&\le \Big(4c_1(1-a)+c_2(\alpha-2)+c_3\Big(2+\frac{\alpha}{3}\Big)\lambda\Big)\int_M(\Delta_b w)^2\,dv\\
&\quad+\Big(c_2\Big(\frac{1}{\alpha}-\frac{1}{2}\Big)+c_3\frac{\lambda}{3\alpha}\Big)\int_M|\nabla_b w|^4\,dv+C_4\int_M|\nabla_b w|^2\,dv\\
&\quad+2\int_M a_4 w\,dv+2ac_1\int_M Q'_\theta w\,dv+C_5.
\end{aligned}
$$

Now we need to choose α in a way that the coefficients of $\int_M(\Delta_b w)^2\,dv$ and $\int_M|\nabla_b w|^4\,dv$ are both negative. For this to happen, one needs that

$$
\begin{cases}
4c_1(1-a)-2c_2+2\lambda c_3<0,\\
\alpha<\dfrac{2c_2-2\lambda c_3-4c_1(1-a)}{c_2+\frac{1}{3}\lambda c_3},\\
\dfrac{1}{\alpha}<\dfrac{c_2}{2c_2+\frac{2}{3}\lambda c_3}.
\end{cases}
$$

Hence, we need

$$
\frac{2c_2+\frac{2}{3}\lambda c_3}{c_2}<\frac{2c_2-2\lambda c_3-4c_1(1-a)}{c_2+\frac{1}{3}\lambda c_3}.
$$

This is possible if condition (6) is satisfied for $\mu=3/(2\lambda)$. This yields in particular that if $w_k\in W_H^{2,2}(M)\cap\mathcal{P}$ is a maximizing sequence for F, then there exists $C>0$ such that

$$
\int_M(\Delta_b w_k)^2\,dv+\int_M|\nabla_b w_k|^4\,dv\le C.
$$

Indeed, we have

$$
-\varepsilon<F(0)\le F(w_k).
$$

Therefore, there exists $c>0$ such that

$$-\varepsilon \leq -c\left(\int_M (\Delta_b w_k)^2\,dv + \int_M |\nabla_b w_k|^4\,dv\right) + C_4\int_M |\nabla_b w_k|^2\,dv \\ +2\int_M a_4 w_k\,dv + 2ac_1\int_M Q'_\theta w_k\,dv + C_5. \tag{15}$$

Thus

$$\int_M (\Delta_b w_k)^2\,dv+\int_M |\nabla_b w_k|^4\,dv \leq C_6\left(\int_M |\nabla_b w_k|^4\,dv\right)^{\frac{1}{2}} + C_7\left(\int_M |\nabla_b w_k|^4\,dv\right)^{\frac{1}{4}} + C_8,$$

where here we used the fact that $\bar{w}=0$ and

$$\int_M f w\,dv \leq \|f\|_{L^2}\|w\|_{L^2} \leq \|f\|_{L^2}\|\nabla_b w\|_{L^2} \leq \|f\|_{L^2} V^{\frac{1}{4}}\|\nabla_b w\|_{L^4},$$

in order to bound the terms of $\int_M a_4 w_k\,dv$ and $\int_M Q'_\theta w_k\,dv$ and Hölder's inequality to bound the term $\int_M |\nabla_b w_k|^2\,dv$.

Hence, $(w_k)_k$ is bounded in $W_H^{2,2}(M)$ and has a weakly convergent subsequence that converges to w_∞. Passing to the lim sup in $F(w_k)$, we get that the weak limit w_∞ is in fact a maximizer of F.

Finally, based on the remark below (5), we see that the critical points of F under the constraint $\int_M e^{2w}dv = 1$ satisfy the equation

$$\tilde{\tau}\left[c_1\widetilde{Q'_\theta} + c_2\widetilde{\Delta_b}\widetilde{R} + c_3 e^{-2w}\mathcal{H}\right] = cte.$$

Acknowledgement

The author wants to acknowledge that the problem treated in this paper was initiated by a question of Carlo Morpurgo. The author is grateful for the fruitful discussions that led to the finalized version of this paper. The author also wants to thank the referee for a careful review of the paper and suggestions that led to the improvement of this manuscript.

References

[Beckner 1993] W. Beckner, "Sharp Sobolev inequalities on the sphere and the Moser–Trudinger inequality", *Ann. of Math.* (2) **138**:1 (1993), 213–242. MR Zbl

[Branson and Ørsted 1986] T. P. Branson and B. Ørsted, "Conformal indices of Riemannian manifolds", *Compositio Math.* **60**:3 (1986), 261–293. MR Zbl

[Branson and Ørsted 1991a] T. P. Branson and B. Ørsted, "Conformal geometry and global invariants", *Differential Geom. Appl.* **1**:3 (1991), 279–308. MR Zbl

[Branson and Ørsted 1991b] T. P. Branson and B. Ørsted, "Explicit functional determinants in four dimensions", *Proc. Amer. Math. Soc.* **113**:3 (1991), 669–682. MR Zbl

[Branson et al. 2013] T. P. Branson, L. Fontana, and C. Morpurgo, "Moser–Trudinger and Beckner-Onofri's inequalities on the CR sphere", *Ann. of Math.* (2) **177**:1 (2013), 1–52. MR Zbl

[Burns and Epstein 1988] D. M. Burns, Jr. and C. L. Epstein, "A global invariant for three-dimensional CR-manifolds", *Invent. Math.* **92**:2 (1988), 333–348. MR Zbl

[Case and Yang 2013] J. S. Case and P. Yang, "A Paneitz-type operator for CR pluriharmonic functions", *Bull. Inst. Math. Acad. Sin. (N.S.)* **8**:3 (2013), 285–322. MR Zbl

[Case et al. 2016] J. S. Case, C.-Y. Hsiao, and P. Yang, "Extremal metrics for the Q'-curvature in three dimensions", *C. R. Math. Acad. Sci. Paris* **354**:4 (2016), 407–410. MR

[Chang and Yang 1995] S.-Y. A. Chang and P. C. Yang, "Extremal metrics of zeta function determinants on 4-manifolds", *Ann. of Math.* (2) **142**:1 (1995), 171–212. MR Zbl

[Chêng and Lee 1990] J. H. Chêng and J. M. Lee, "The Burns–Epstein invariant and deformation of CR structures", *Duke Math. J.* **60**:1 (1990), 221–254. MR Zbl

[Esposito and Malchiodi 2019] P. Esposito and A. Malchiodi, "Critical metrics for Log-determinant functionals in conformal geometry", 2019. arXiv

[Fefferman and Hirachi 2003] C. Fefferman and K. Hirachi, "Ambient metric construction of Q-curvature in conformal and CR geometries", *Math. Res. Lett.* **10**:5-6 (2003), 819–831. MR Zbl

[Gover and Graham 2005] A. R. Gover and C. R. Graham, "CR invariant powers of the sub-Laplacian", *J. Reine Angew. Math.* **583** (2005), 1–27. MR Zbl

[Gursky and Malchiodi 2012] M. Gursky and A. Malchiodi, "Non-uniqueness results for critical metrics of regularized determinants in four dimensions", *Comm. Math. Phys.* **315**:1 (2012), 1–37. MR Zbl

[Hirachi 2014] K. Hirachi, "Q-prime curvature on CR manifolds", *Differential Geom. Appl.* **33**:suppl. (2014), 213–245. MR Zbl

[Ho 2019] P. T. Ho, "Prescribing the Q'-curvature in three dimension", *Discrete Contin. Dyn. Syst.* **39**:4 (2019), 2285–2294. MR Zbl

[Maalaoui 2019a] A. Maalaoui, "Logarithmic Hardy–Littlewood–Sobolev Inequality on Pseudo-Einstein 3-manifolds and the Logarithmic Robin Mass", 2019. arXiv

[Maalaoui 2019b] A. Maalaoui, "Prescribing the $\bar{Q}'$-Curvature on Pseudo-Einstein CR 3-Manifolds", 2019. arXiv

[Boutet de Monvel and Guillemin 1981] L. Boutet de Monvel and V. Guillemin, *The spectral theory of Toeplitz operators*, Annals of Mathematics Studies **99**, Princeton University Press, 1981. MR

[Morpurgo 1996] C. Morpurgo, "The logarithmic Hardy–Littlewood–Sobolev inequality and extremals of zeta functions on S^n", *Geom. Funct. Anal.* **6**:1 (1996), 146–171. MR Zbl

[Morpurgo 2002] C. Morpurgo, "Sharp inequalities for functional integrals and traces of conformally invariant operators", *Duke Math. J.* **114**:3 (2002), 477–553. MR Zbl

[Okikiolu 2001] K. Okikiolu, "Critical metrics for the determinant of the Laplacian in odd dimensions", *Ann. of Math.* (2) **153**:2 (2001), 471–531. MR Zbl

[Okikiolu 2008] K. Okikiolu, "Extremals for logarithmic Hardy–Littlewood–Sobolev inequalities on compact manifolds", *Geom. Funct. Anal.* **17**:5 (2008), 1655–1684. MR Zbl

[Polyakov 1981a] A. M. Polyakov, "Quantum geometry of bosonic strings", *Phys. Lett. B* **103**:3 (1981), 207–210. MR

[Polyakov 1981b] A. M. Polyakov, "Quantum geometry of fermionic strings", *Phys. Lett. B* **103**:3 (1981), 211–213. MR

[Ponge 2007] R. Ponge, "Noncommutative residue for Heisenberg manifolds. Applications in CR and contact geometry", *J. Funct. Anal.* **252**:2 (2007), 399–463. MR Zbl

[Stanton 1989] N. K. Stanton, "Spectral invariants of CR manifolds", *Michigan Math. J.* **36**:2 (1989), 267–288. MR Zbl

Received August 16, 2020. Revised October 6, 2020.

ALI MAALAOUI
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
CLARK UNIVERSITY
WORCESTER, MA
UNITED STATES

and

DEPARTMENT OF MATHEMATICS AND NATURAL SCIENCES
AMERICAN UNIVERSITY OF RAS AL KHAIMAH
RAS AL KHAIMAH
UNITED ARAB EMIRATES

amaalaoui@clarku.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 2, 2020
https://doi.org/10.2140/pjm.2020.309.437

DISTRIBUTION OF DISTANCES IN POSITIVE CHARACTERISTIC

THANG PHAM AND LÊ ANH VINH

Let $\mathbb{F}_q$ be an arbitrary finite field, and $\mathcal{E}$ be a point set in $\mathbb{F}_q^d$. Let $\Delta(\mathcal{E})$ be the set of distances determined by pairs of points in $\mathcal{E}$. Using Kloosterman sums, Iosevich and Rudnev (2007) proved that if $|\mathcal{E}| \geq 4q^{(d+1)/2}$ then $\Delta(\mathcal{E}) = \mathbb{F}_q$. In general, this result is sharp in odd-dimensional spaces over arbitrary finite fields. We use the point-plane incidence bound due to Rudnev to prove that if $\mathcal{E}$ has Cartesian product structure in vector spaces over prime fields, then we can break the exponent $(d+1)/2$ and still cover all distances. We also show that the number of pairs of points in $\mathcal{E}$ of any given distance is close to its expected value.

1. Introduction

Let $\mathcal{E}$ be a finite subset of $\mathbb{R}^d$ ($d \geq 2$), and $\Delta(\mathcal{E})$ be the distance set determined by $\mathcal{E}$. The Erdős distinct distances problem is to find the best lower bound of the size of the distance set $\Delta(\mathcal{E})$ in terms of the size of the point set $\mathcal{E}$.

In the plane case, Erdős [1946] conjectured that $|\Delta(\mathcal{E})| \gg |\mathcal{E}|/\sqrt{\log|\mathcal{E}|}$. This conjecture was proved up to logarithmic factor by Guth and Katz [2015] in 2010. More precisely, they showed that $|\Delta(\mathcal{E})| \gg |\mathcal{E}|/\log|\mathcal{E}|$. In higher dimension cases, Erdős [1946] also conjectured that $|\Delta(\mathcal{E})| \gg |\mathcal{E}|^{2/d}$. Interested readers are referred to [Solymosi and Vu 2008] for results on Erdős distinct distances problem in three and higher dimensions.

In this paper, we use the following notation: $X \ll Y$ means that there exists some absolute constant $C_1 > 0$ such that $X \leq C_1 Y$, $X \sim Y$ means $Y \ll X \ll Y$, $X \gtrsim Y$ means $X \gg (\log Y)^{-C_2} Y$ for some absolute constant $C_2 > 0$, and $X \gtrsim_d Y$ means $X \geq C_3 (\log Y)^{-C_4} Y$ for some positive constants C_3, C_4 depending on d.

As a continuous analogue of the Erdős distinct distances problem, Falconer [1985] asked how large the Hausdorff dimension of $\mathcal{E} \subset \mathbb{R}^d$ needs to be to ensure that the Lebesgue measure of $\Delta(\mathcal{E})$ is positive. He conjectured that for any subset $\mathcal{E} \subset \mathbb{R}^d$ of the Hausdorff dimension greater than $d/2$, $\mathcal{E}$ determines a distance set of a positive Lebesgue measure. This conjecture is still open in all dimensions.

MSC2020: 14N10, 51A45, 52C10.
Keywords: distances, finite fields, incidence, Rudnev's point-plane incidence bound.

We refer readers to [Du et al. 2018; Guth et al. 2020] for recent updates on this conjecture.

Let $\mathbb{F}_q$ be the finite field of order q, where q is an odd prime power. Given two points $\boldsymbol{x} = (x_1, \dots, x_d)$ and $\boldsymbol{y} = (y_1, \dots, y_d)$ in $\mathbb{F}_q^d$, we denote the distance between $\boldsymbol{x}$ and $\boldsymbol{y}$ by

$$\|\boldsymbol{x} - \boldsymbol{y}\| := (x_1 - y_1)^2 + \cdots + (x_d - y_d)^2.$$

Note that the distance function defined here is not a metric but it is invariant under translations and actions of the orthogonal group.

For a subset $\mathcal{E} \subset \mathbb{F}_q^d$, we denote the set of all distances determined by $\mathcal{E}$ by

$$\Delta(\mathcal{E}) := \{\|\boldsymbol{x} - \boldsymbol{y}\| : \boldsymbol{x}, \boldsymbol{y} \in \mathcal{E}\}.$$

The finite field analogue of the Erdős distinct distances problem was first studied by Bourgain, Katz, and Tao [Bourgain et al. 2004]. More precisely, they proved that in the prime field $\mathbb{F}_p$ with $p \equiv 3 \mod 4$, $|\Delta(\mathcal{E})| \gg |\mathcal{E}|^{1/2+\epsilon}$ for some $\epsilon = \epsilon(\alpha) > 0$, for any subset $\mathcal{E} \subset \mathbb{F}_p^2$ of the cardinality $|\mathcal{E}| = p^\alpha$, $0 < \alpha < 2$,

Note that the condition $p \equiv 3 \mod 4$ in Bourgain, Katz, and Tao's result is necessary, since if $p \equiv 1 \mod 4$, then there exists $i \in \mathbb{F}_p$ such that $i^2 = -1$. By taking $\mathcal{E} = \{(x, ix) : x \in \mathbb{F}_p\}$, we have $|\mathcal{E}| = p$ and $\Delta(\mathcal{E}) = \{0\}$.

In the setting of arbitrary finite fields $\mathbb{F}_q$, Iosevich and Rudnev [2007] showed that Bourgain, Katz, and Tao's result does not hold. For example, assuming that $q = p^2$, one can take $\mathcal{E} = \mathbb{F}_p^2$, then $\Delta(\mathcal{E}) = \mathbb{F}_p$ or $|\Delta(\mathcal{E})| = |\mathcal{E}|^{1/2}$. Thus, Iosevich and Rudnev reformulated the problem in the spirit of the Falconer distance conjecture over the Euclidean spaces. More precisely, they asked, for a subset $\mathcal{E} \subset \mathbb{F}_q^d$, how large $|\mathcal{E}|$ needs to be to ensure that $\Delta(\mathcal{E})$ covers the whole field or at least a positive proportion of all elements of the field?

Using Fourier analytic methods, Iosevich and Rudnev [2007] proved that $\Delta(\mathcal{E}) = \mathbb{F}_q$ for any point set $\mathcal{E} \subset \mathbb{F}_q^d$ with the cardinality $|\mathcal{E}| \ge 4q^{(d+1)/2}$. Hart, Iosevich, Koh, and Rudnev [Hart et al. 2011] showed that, in general, the exponent $(d+1)/2$ cannot be improved when d is odd, even if we only want to cover a positive proportion of all the distances. In even-dimensional cases, it has been conjectured that the exponent $(d+1)/2$ can be improved to $d/2$, which is in line with the Falconer distance conjecture in the Euclidean space.

In the plane case, Bennett, Hart, Iosevich, Pakianathan, and Rudnev [Bennett et al. 2017] proved that if $\mathcal{E}$ is a subset of $\mathbb{F}_q^2$ of cardinality $|\mathcal{E}| \ge q^{4/3}$, then $\Delta(\mathcal{E})$ covers a positive proportion of all distances. Murphy and Petridis [2019] showed that there are infinite subsets of $\mathbb{F}_q^2$ of size $q^{4/3}$ whose distance sets do not cover the whole field $\mathbb{F}_q$. It is not known whether there exist a small $c > 0$ and a set $\mathcal{E} \subset \mathbb{F}_q^2$ with $|\mathcal{E}| \ge cq^{3/2}$ such that $\Delta(\mathcal{E}) \ne \mathbb{F}_q$. We refer the interested reader to [Hart et al. 2011, Theorem 2.7] for a construction in odd-dimensional spaces.

Chapman et al. [2012] broke the exponent $(d+1)/2$ to $d^2/(2d-1)$ under the additional assumption that the set $\mathcal{E}$ has Cartesian product structure. However, in this case, they can cover only a positive proportion of all distances. In the setting of prime fields, it has been proved in [Pham et al. 2019] that for $A \subset \mathbb{F}_p$, we have $|\Delta(A^d)| \geq \frac{1}{c} \cdot \min\{|A|^{2-1/2^{d-2}}, p\}$ with $c = 2^{(2^{d-1}-1)/2^{d-2}}$. Therefore, $|\Delta(A^d)| \geq p/c$ under the condition $|A| \geq p^{2^{d-2}/(2^{d-1}-1)}$. However, this result again only gives us a positive proportion of all distances, and does not tell us the number of pairs of any given distance.

We will show that if $\mathcal{E} \subset \mathbb{F}_p^d$ has Cartesian product structure, we can break the exponent $(d+1)/2$ due to Iosevich and Rudnev [2007] and still cover all possible distances. Our main tool is the point-plane incidence bound due to Rudnev [2018].

Our first result is for odd-dimensional cases.

Theorem 1.1. *Let $\mathbb{F}_p$ be a prime field, and A be a set in $\mathbb{F}_p$. For an integer $d \geq 3$, suppose the set $A^{2d+1} \subset \mathbb{F}_p^{2d+1}$ satisfies*

$$|A^{2d+1}| \gtrsim_d p^{\frac{2d+2}{2} - \frac{3 \cdot 2^{d-2} - d - 1}{3 \cdot 2^{d-1} - 1}};$$

then:

- *The distance set covers all elements in $\mathbb{F}_p$, namely,*
$$\Delta(A^{2d+1}) = \underbrace{(A-A)^2 + \cdots + (A-A)^2}_{2d+1 \text{ terms}} = \mathbb{F}_p.$$
- *The number of pairs $(\boldsymbol{x}, \boldsymbol{y}) \in A^{2d+1} \times A^{2d+1}$ satisfying $\|\boldsymbol{x} - \boldsymbol{y}\| = \lambda$ is $\sim p^{-1}|A|^{4d+2}$ for any $\lambda \in \mathbb{F}_p$.*

Corollary 1.2. *For $A \subset \mathbb{F}_p$, suppose that $|A| \gtrsim p^{6/11}$; then we have*

$$\Delta(A^7) = (A-A)^2 + (A-A)^2 + (A-A)^2 + (A-A)^2 + (A-A)^2 + (A-A)^2 + (A-A)^2 = \mathbb{F}_p.$$

Our second result is for even-dimensional cases.

Theorem 1.3. *Let $\mathbb{F}_p$ be a prime field, and A be a set in $\mathbb{F}_p$. For an integer $d \geq 3$, suppose the set $A^{2d} \subset \mathbb{F}_p^{2d}$ satisfies*

$$|A^{2d}| \gtrsim_d p^{\frac{2d+1}{2} - \frac{2^d - 2d - 1}{2^{d+1} - 2}};$$

then:

- *The distance set covers all elements in $\mathbb{F}_p$, namely,*
$$\Delta(A^{2d}) = \underbrace{(A-A)^2 + \cdots + (A-A)^2}_{2d \text{ terms}} = \mathbb{F}_p.$$
- *The number of pairs $(\boldsymbol{x}, \boldsymbol{y}) \in A^{2d} \times A^{2d}$ satisfying $\|\boldsymbol{x} - \boldsymbol{y}\| = \lambda$ is $\sim p^{-1}|A|^{4d}$ for any $\lambda \in \mathbb{F}_p$.*

Corollary 1.4. *For $A \subset \mathbb{F}_p$, suppose that $|A| \gtrsim p^{4/7}$; then we have*

$$\Delta(A^6) = (A-A)^2 + (A-A)^2 + (A-A)^2 + (A-A)^2 + (A-A)^2 + (A-A)^2 = \mathbb{F}_p.$$

Remark. In the setting of arbitrary finite fields $\mathbb{F}_q$, it is not possible to break the exponent $(d+1)/2$ and still cover all distances with the method in this paper and the distance energy in [Pham et al. 2017, Lemma 3.1]. More precisely, for $A \subset \mathbb{F}_q$, one can follow the proofs of Theorems 1.1 and 1.3 to get the conditions $|A^{2d+1}| \gg q^{(2d+2)/2+1/(4d)}$ and $|A^{2d}| \gg q^{(2d+1)/2+1/(4d-2)}$ for odd and even dimensions, respectively.

Remark. The Cauchy–Davenport theorem states that for subsets X and Y of $\mathbb{F}_p$, we have $|X+Y| \geq \min\{p, |X|+|Y|-1\}$. It is not hard to check that $\Delta(A^{2d}) = \Delta(A^d) + \Delta(A^d)$. The result of [Chapman et al. 2012] tells us that $|\Delta(A^d)| \geq p/2$ whenever $|A| \gg p^{d/(2d-1)}$. Therefore, one can apply the Cauchy–Davenport theorem to show that $|\Delta(A^{2d})| \geq p-1$ under the condition $|A| \geq p^d/(2d-1)$. However, our set A^{2d} lies on the $2d$-dimensional space $\mathbb{F}_p^{2d}$; thus the exponent $d/(2d-1)$ is worse than the Iosevich–Rudnev's exponent $(2d+1)/(4d)$. The same happens for odd-dimensional spaces. Note that the bound $|\Delta(A^d)| \geq 1/c \cdot \min\{|A|^{2-1/2^{d-2}}, p\}$ with $c = 2^{(2^{d-1}-1)/2^{d-2}}$ in [Pham et al. 2019] is not suitable for this approach since the constant factor $1/c$ is too small.

Let $\mathbb{F}_q$ be an arbitrary finite field, and $\mathcal{E} \subset \mathbb{F}_q^d$. The product set of $\mathcal{E}$, denoted by $\Pi(\mathcal{E})$, is defined as follows:

$$\Pi(\mathcal{E}) := \{\boldsymbol{x} \cdot \boldsymbol{y} : \boldsymbol{x}, \boldsymbol{y} \in \mathcal{E}\}.$$

Using Fourier analysis, Hart and Iosevich [2008] proved that if $|\mathcal{E}| \gg q^{(d+1)/2}$, then $\Pi(\mathcal{E}) \supseteq \mathbb{F}_q \setminus \{0\}$. Moreover, under the same condition on the size of $\mathcal{E}$, we have that the number of pairs $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{E} \times \mathcal{E}$ satisfying $\boldsymbol{x} \cdot \boldsymbol{y} = \lambda$ is $\sim q^{-1}|\mathcal{E}|^2$ for any $\lambda \neq 0$. If $\mathcal{E}$ has Cartesian product structure, i.e., $\mathcal{E} = A^d$ for some $A \subset \mathbb{F}_q$, then the condition $|\mathcal{E}| \gg q^{(d+1)/2}$ is equivalent with $|A| \gg q^{1/2+1/(2d)}$.

In the setting of prime fields $\mathbb{F}_p$, if $d = 8$, Glibichuk and Konyagin [2007] proved that for $A, B \subset \mathbb{F}_p$, if $|A|\lceil |B|/2 \rceil \geq p$, then we have $8A \cdot B = \mathbb{F}_p$. This result has been extended to arbitrary finite fields in [Glibichuk and Rudnev 2009].

In this paper, using the techniques in the proof of Theorem 1.3, we are able to obtain the following theorem.

Theorem 1.5. *For $A \subset \mathbb{F}_p$, suppose that $|A| \gtrsim p^{4/7}$; then*:

- $6A \cdot A = A \cdot A + A \cdot A + A \cdot A + A \cdot A + A \cdot A + A \cdot A = \mathbb{F}_p$.
- *For any $\lambda \in \mathbb{F}_p$, the number of pairs $(\boldsymbol{x}, \boldsymbol{y}) \in A^6 \times A^6$ such that $\boldsymbol{x} \cdot \boldsymbol{y} = \lambda$ is $\sim p^{-1}|A|^{12}$.*

Note that our exponent $4/7$ improves the exponent $7/12$ of Hart and Iosevich [2008] in the case $d = 6$. The following is the conjecture due to Iosevich.

Conjecture 1.6. *Let A be a set in $\mathbb{F}_p$ and suppose that $|A| \gg p^{\frac{1}{2}+\epsilon}$ for any $\epsilon > 0$. Then,*

$$A \cdot A + A \cdot A = \mathbb{F}_p, \quad (A-A)^2 + (A-A)^2 = \mathbb{F}_p.$$

In the spirit of sum-product problems, it has been proved in [Pham et al. 2019] that for $A \subset \mathbb{F}_p$, if $|A| \ll p^{1/2+1/(5\cdot 2^{d-1}-2)}$, $d \geq 2$, then we have

$$\max\{|\Delta(A^d)|, |\Pi(A^d)|\} \gg |A|^{2-\frac{1}{5\cdot 2^d-3}}.$$

Using our energies (Lemmas 2.2 and 2.4 below), and the prime field analogue of Balog–Wooley decomposition energy due to Rudnev, Shkredov, and Stevens [Rudnev et al. 2020], we are able to give the energy variant of this result.

Theorem 1.7. *Let $d \geq 2$ be an integer and A a set in $\mathbb{F}_p$ with $|A| \ll p^{1/2+1/(5\cdot 2^{d-1}-2)}$. There exist two disjoint subsets B and C of A such that $A = B \sqcup C$ and*

$$\max\{E_d((B-B)^2), E_d(C\cdot C)\} \ll d^4 (\log|A|)^4 |A|^{4d-2+\frac{1}{5\cdot 2^{d-3}}},$$

where $E_d((B-B)^2)$ is the number of $4d$-tuples $\{(a_i, b_i, c_i, e_i)\}_{i=1}^d$ with a_i, c_i, b_i, e_i in B such that $(a_1-b_1)^2 + \cdots + (a_d-b_d)^2 = (c_1-e_1)^2 + \cdots + (c_d-e_d)^2$, and $E_d(C\cdot C)$ be the number of $4d$-tuples $\{(a_i, b_i, c_i, e_i)\}_{i=1}^d$ with $a_i, c_i, b_i, e_i \in C$ such that $a_1b_1 + \cdots + a_db_d = c_1e_1 + \cdots + c_de_d$.

2. Preliminaries

Let E and F be multisets in $\mathbb{F}_p^2$. We denote by $\bar{E}$ and $\bar{F}$ the sets of distinct elements in E and F, respectively. For any multiset X, we use the notation $|X|$ to denote the size of X. For $\lambda \in \mathbb{F}_p$, let $N(E, F, \lambda)$ be the number of pairs $((e_1, e_2), (f_1, f_2)) \in E \times F$ such that $e_1 f_1 + e_2 + f_2 = \lambda$. In the following lemma, we provide an upper bound and a lower bound of $N(E, F, \lambda)$ for any $\lambda \in \mathbb{F}_p$. Note that, this lemma is essentially the weighted version of the point-line incidences of [Vinh 2011] in the plane $\mathbb{F}_q^2$ (see also [Hanson et al. 2016, Lemma 14]).

Lemma 2.1. *Let E, F be multisets in $\mathbb{F}_p^2$. For any $\lambda \in \mathbb{F}_p$, we have*

$$\left| N(E, F, \lambda) - \frac{|E||F|}{p} \right| \leq p^{\frac{1}{2}} \Bigg(\sum_{(e_1,e_2)\in\bar{E}} m_E((e_1, e_2))^2 \sum_{(f_1,f_2)\in\bar{F}} m_F((f_1, f_2))^2 \Bigg)^{1/2},$$

where $m_X((a, b))$ is the multiplicity of (a, b) in X with $X \in \{E, F\}$.

Proof. Let χ be a nontrivial additive character on $\mathbb{F}_p$. We have

$$N(E, F, \lambda) = \sum_{(e_1,e_2)\in\bar{E},(f_1,f_2)\in\bar{F}} \frac{1}{p} m_E((e_1, e_2)) m_F((f_1, f_2)) \sum_{s\in\mathbb{F}_p} \chi(s\cdot(e_1 f_1 + e_2 + f_2 - \lambda)).$$

This gives us

$$N(E, F, \lambda) = \frac{|E||F|}{p} + L,$$

where

$$L = \sum_{(e_1,e_2)\in\bar{E},(f_1,f_2)\in\bar{F}} m_E((e_1, e_2))m_F((f_1, f_2))\frac{1}{p}\sum_{s\neq 0}\chi(s\cdot(e_1 f_1 + e_2 + f_2 - \lambda)).$$

If we view L as a sum in $(e_1, e_2) \in \bar{E}$, then we can apply the Cauchy–Schwarz inequality to derive the following:

$$\begin{aligned}
L^2 \leq& \sum_{(e_1,e_2)\in\bar{E}} m_E((e_1, e_2))^2 \\
&\sum_{(e_1,e_2)\in\mathbb{F}_p^2} \frac{1}{p^2}\sum_{s,s'\neq 0}\sum_{(f_1,f_2),(f_1',f_2')\in\bar{F}} m_F((f_1, f_2))m_F((f_1', f_2')) \\
&\times\chi(s\cdot(e_1 f_1 + e_2 + f_2 - \lambda))\chi(s'\cdot(-e_1 f_1' - e_2 - f_2' + \lambda)) \\
=& \sum_{(e_1,e_2)\in\bar{E}} m_E((e_1, e_2))^2\frac{1}{p^2} \\
&\sum_{\substack{(e_1,e_2)\in\mathbb{F}_p^2,\ (f_1,f_2)\in\bar{F}\\ (f_1',f_2')\in\bar{F},\ s,s'\neq 0}} m_F((f_1, f_2))m_F((f_1', f_2'))\chi(e_1(sf_1 - s' f_1'))\chi(e_2(s - s')) \\
&\times\chi(s(f_2 - \lambda) - s'(f_2' - \lambda)) \\
=& \sum_{(e_1,e_2)\in\bar{E}} m_E((e_1, e_2))^2 \\
&\sum_{\substack{s\neq 0,\ (f_1,f_2)\in\bar{F}\\ (f_1',f_2')\in\bar{F},\ f_1=f_1'}} m_F((f_1, f_2))m_F((f_1', f_2'))\chi(s\cdot(f_2 - f_2')) = I + II,
\end{aligned}$$

where I is the sum over all pairs $((f_1, f_2), (f_1, f_2'))$ with $f_2 = f_2'$, and II is the sum over all pairs $((f_1, f_2), (f_1', f_2'))$ with $f_2 \neq f_2'$.

It is not hard to check that if $f_2 \neq f_2'$, then

$$\sum_{s\neq 0}\chi(s\cdot(f_2 - f_2')) = -1,$$

so $II < 0$. Note that $|II| \leq I$ since $L^2 \geq 0$.

On the other hand, if $f_2 = f_2'$, then

$$\sum_{s\neq 0}\chi(s\cdot(f_2 - f_2')) = p - 1.$$

In other words,

$$I \le p \sum_{(e_1,e_2)\in\overline{E}} m_E((e_1,e_2))^2 \sum_{(f_1,f_2)\in\overline{F}} m_F((f_1,f_2))^2,$$

which implies that

$$|L| \le \sqrt{I+II} \le p^{\frac{1}{2}}\left(\sum_{(e_1,e_2)\in\overline{E}} m_E((e_1,e_2))^2 \sum_{(f_1,f_2)\in\overline{F}} m_F((f_1,f_2))^2\right)^{1/2}.$$

This completes the proof of the lemma. □

For $A \subset \mathbb{F}_p$, let $E_d((A-A)^2)$ be the number of $4d$-tuples $\{(a_i,b_i,c_i,e_i)\}_{i=1}^d$ with $a_i, c_i, b_i, e_i \in A$ such that

$$(a_1-b_1)^2+\cdots+(a_d-b_d)^2=(c_1-e_1)^2+\cdots+(c_d-e_d)^2.$$

Similarly, let $E_d(A\cdot A)$ be the number of $4d$-tuples $\{(a_i,b_i,c_i,e_i)\}_{i=1}^d$ with $a_i, c_i, b_i, e_i \in A$ such that

$$a_1b_1+\cdots+a_db_d=c_1e_1+\cdots+c_de_d.$$

In our next lemmas, we give recursive formulas for $E_d((A-A)^2)$ and $E_d(A\cdot A)$.

Lemma 2.2. *For $A \subset \mathbb{F}_p$, we have*

$$E_d((A-A)^2) \le Cd^2(\log|A|)^2\left(\frac{|A|^{4d}}{p}+|A|^{2d+1}\sqrt{E_{d-1}((A-A)^2)}\right),$$

for some positive constant C.

The proof of this lemma will be given in the next section. The following result is a direct consequence, which tells us an upper bound of $E_d((A-A)^2)$.

Corollary 2.3. *Let A be a set in $\mathbb{F}_p$. For $d\ge 2$, suppose that $|A| \gg (d\log|A|)p^{1/2}$; then we have*

$$E_d((A-A)^2) \ll d^2(\log|A|)^2\frac{|A|^{4d}}{p}+d^4(\log|A|)^4|A|^{4d-2+\frac{1}{2^{d-1}}}.$$

Proof. We prove by induction on d that

$$E_d((A-A)^2) \le 2C^2d^2(\log|A|)^2\frac{|A|^{4d}}{p}+2C^2d^4(\log|A|)^4|A|^{4d-2+\frac{1}{2^{d-1}}},$$

whenever $|A| \ge \sqrt{2}C(d\log|A|)p^{1/2}$, where the constant C comes from Lemma 2.2.

The base case $d=2$ follows directly from Lemma 2.2 by using the trivial upper bound $|A|^3$ of $E_1((A-A)^2)$.

Suppose the statement holds for any $d-1\geq 2$; we now prove that it also holds for d. Indeed, by induction hypothesis, we have

$$\begin{aligned}(1)\quad & E_{d-1}((A-A)^2)\\ &\leq 2C^2(d-1)^2(\log|A|)^2\frac{|A|^{4(d-1)}}{p}+2C^2(d-1)^4(\log|A|)^4|A|^{4d-6+\frac{1}{2^{d-2}}}\\ &\leq 2C^2d^2(\log|A|)^2\frac{|A|^{4(d-1)}}{p}+2C^2d^4(\log|A|)^4|A|^{4d-6+\frac{1}{2^{d-2}}}.\end{aligned}$$

On the other hand, it follows from Lemma 2.2 that

$$(2)\qquad E_d((A-A)^2)\leq Cd^2(\log|A|)^2\left(\frac{|A|^{4d}}{p}+|A|^{2d+1}\sqrt{E_{d-1}((A-A)^2)}\right).$$

Putting (1) and (2) together, we obtain

$$\begin{aligned}E_d((A-A)^2)\leq Cd^2(\log|A|)^2\frac{|A|^{4d}}{p}&+\sqrt{2}C^2d^2(\log|A|)^2|A|^{2d+1}\\ &\times\left(d\log|A|\frac{|A|^{2(d-1)}}{p^{1/2}}+d^2(\log|A|)^2|A|^{2d-3+\frac{1}{2^{d-1}}}\right).\end{aligned}$$

Since $\sqrt{2}C(d\log|A|)p^{1/2}\leq|A|$, we have

$$\sqrt{2}C^2d\log|A|\frac{|A|^{2(d-1)}}{p^{1/2}}\leq C\frac{|A|^{2d-1}}{p}.$$

This implies that

$$E_d((A-A)^2)\leq 2Cd^2(\log|A|)^2\frac{|A|^{4d}}{p}+2C^2d^4(\log|A|)^4|A|^{4d-2+\frac{1}{2^{d-1}}}.\qquad\square$$

Similarly, for the case of product sets, we have:

Lemma 2.4. *For $A\subset\mathbb{F}_p$,*

$$E_d(A\cdot A)\leq Cd^2(\log|A|)^2\left(\frac{|A|^{4d}}{p}+|A|^{2d+1}\sqrt{E_{d-1}(A\cdot A)}\right),$$

for some positive constant C.

Proof. The proof of this lemma is almost identical to that of Lemma 2.2, so we omit it. □

Corollary 2.5. *Let A be a set in $\mathbb{F}_p$. For $d\geq 2$, suppose that $|A|\gg(d\log|A|)p^{1/2}$; then we have*

$$E_d(A\cdot A)\ll d^2(\log|A|)^2\frac{|A|^{4d}}{p}+d^4(\log|A|)^4|A|^{4d-2+\frac{1}{2^{d-1}}}.$$

Proof. The proof of Corollary 2.5 is identical to that of Corollary 2.3 with Lemma 2.4 in the place of Lemma 2.2; thus we omit it. □

2.1. *Proof of Lemma 2.2.* In the proof of Lemma 2.2, we will use a point-plane incidence bound due to Rudnev [2018] and an argument in [Shkredov 2018, Theorem 32].

Let us first recall that if $\mathcal{R}$ is a set of points in $\mathbb{F}_p^3$ and $\mathcal{S}$ is a set of planes in $\mathbb{F}_p^3$, then the number of incidences between $\mathcal{R}$ and $\mathcal{S}$, denoted by $\mathcal{I}(\mathcal{R}, \mathcal{S})$, is the cardinality of the set $\{(r, s) \in \mathcal{R} \times \mathcal{S} : r \in s\}$. The following is a version of Rudnev's point-plane incidence bound, which can be found in [de Zeeuw 2016].

Theorem 2.6 [Rudnev 2018; de Zeeuw 2016]. *Let $\mathcal{R}$ be a set of points in $\mathbb{F}_p^3$ and $\mathcal{S}$ be a set of planes in $\mathbb{F}_p^3$, with $|\mathcal{R}| \le |\mathcal{S}|$. Suppose that there is no line that contains k points of $\mathcal{R}$ and is contained in k planes of $\mathcal{S}$. Then*

$$\mathcal{I}(\mathcal{R}, \mathcal{S}) \ll \frac{|\mathcal{R}||\mathcal{S}|}{p} + |\mathcal{R}|^{1/2}|\mathcal{S}| + k|\mathcal{S}|.$$

Proof of Lemma 2.2. We first have

$$E_d((A-A)^2) = \sum_{t_1,t_2} r_{(d-1)(A-A)^2}(t_1) r_{(d-1)(A-A)^2}(t_2) f(t_1, t_2),$$

where $r_{(d-1)(A-A)^2}(t)$ is the number of $2(d-1)$ tuples $(a_1, \dots, a_{d-1}, b_1, \dots, b_{d-1})$ in A^{2d-2} such that $(a_1 - b_1)^2 + \cdots + (a_{d-1} - b_{d-1})^2 = t$, and $f(t_1, t_2)$ is the sum $\sum_s r_{(A-A)^2+t_1}(s) r_{(A-A)^2+t_2}(s)$. We now split the sum $E_d((A-A)^2)$ into intervals:

$$E_d((A-A)^2) \ll \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} \sum_{t_1,t_2} f(t_1, t_2) r^{(i)}_{(d-1)(A-A)^2}(t_1) r^{(j)}_{(d-1)(A-A)^2}(t_2),$$

where $L_1 \le \log(|A|^{2d-2})$, $L_2 \le \log(|A|^{2d-2})$, and $r^{(i)}_{(d-1)(A-A)^2}(t_1)$ is the restriction of the function $r_{(d-1)(A-A)^2}(x)$ on the set $P_i := \{t : 2^i \le r_{(d-1)(A-A)^2}(t) < 2^{i+1}\}$.

Using the pigeon-hole principle two times, there exist sets P_{i_0} and P_{j_0} for some i_0 and j_0 such that

$$\begin{aligned} E_d((A-A)^2) &\le (2d-2)^2 (\log|A|)^2 \sum_{t_1,t_2} f(t_1, t_2) r^{(i_0)}_{(d-1)(A-A)^2}(t_1) r^{(j_0)}_{(d-1)(A-A)^2}(t_2) \\ &\ll d^2 (\log|A|)^2 2^{i_0} 2^{j_0} \sum_{t_1,t_2} f(t_1, t_2) P_{i_0}(t_1) P_{j_0}(t_2). \end{aligned}$$

One can check that the sum $\sum_{t_1,t_2} f(t_1, t_2) P_{i_0}(t_1) P_{j_0}(t_2)$ is equal to the number of incidences between the point set $\mathcal{R}$ of points $(-2a, e, t_1 + a^2 - e^2) \in \mathbb{F}_p^3$ with $a \in A$, $e \in A$, $t_1 \in P_{i_0}$, and the plane set $\mathcal{S}$ of planes in $\mathbb{F}_p^3$ defined by

$$bX + 2cY + Z = t_2 - b^2 + c^2,$$

where $b \in A$, $c \in A$ and $t_2 \in P_{j_0}$. Without loss of generality, we can assume that $|P_{i_0}| \le |P_{j_0}|$.

To apply Theorem 2.6, we need to bound the maximal number of collinear points in $\mathcal{R}$. The projection of $\mathcal{R}$ into the plane of the first two coordinates is the set $-2A \times A$; thus if a line is not vertical, then it contains at most $|A|$ points from $\mathcal{R}$. If a line is vertical, then it contains at most $|P_{i_0}|$ points from $\mathcal{R}$, but that line is not contained in any plane in $\mathcal{S}$. In other words, we can apply Theorem 2.6 with $k = |A|$, and obtain

$$\sum_{t_1,t_2} f(t_1,t_2)P_{i_0}(t_1)P_{j_0}(t_2) \ll \frac{|A|^4|P_{i_0}||P_{j_0}|}{p} + |A|^3|P_{i_0}|^{1/2}|P_{j_0}| + |A|^3|P_{j_0}|$$
$$\ll \frac{|A|^4|P_{i_0}||P_{j_0}|}{p} + |A|^3|P_{i_0}|^{1/2}|P_{j_0}|.$$

We now fall into the following cases:

Case 1: If the first term dominates, we have

$$\sum_{t_1,t_2} f(t_1,t_2)P_{i_0}(t_1)P_{j_0}(t_2) \ll \frac{|A|^4|P_{i_0}||P_{j_0}|}{p}.$$

Case 2: If the second term dominates, we have

$$\sum_{t_1,t_2} f(t_1,t_2)P_{i_0}(t_1)P_{j_0}(t_2) \ll |A|^3|P_{i_0}|^{1/2}|P_{j_0}|.$$

Therefore,

$$E_d((A-A)^2) \ll d^2(\log|A|)^2 2^{i_0}2^{j_0}\left(\frac{|A|^4|P_{i_0}||P_{j_0}|}{p} + |A|^3|P_{i_0}|^{1/2}|P_{j_0}|\right)$$
$$\ll d^2(\log|A|)^2\left(\frac{|A|^{4d}}{p} + |A|^{2d+1}\sqrt{E_{d-1}((A-A)^2)}\right),$$

where we have used the facts that

- $2^{i_0}|P_{i_0}|^{1/2} \ll \sqrt{E_{d-1}((A-A)^2)}$.
- $2^{j_0}|P_{j_0}| \ll |A|^{2d-2}$.
- $2^{i_0}|P_{i_0}| \ll |A|^{2d-2}$.

This completes the proof of the lemma. □

3. Proof of Theorem 1.1

Proof of Theorem 1.1. Let λ be an arbitrary element in $\mathbb{F}_p$. Let E be the multiset of points $(2x, x^2+(y_1-z_1)^2+\cdots+(y_d-z_d)^2) \in \mathbb{F}_p^2$ with $x, y_i, z_i \in A$, and F be the multiset of points $(-t, t^2+(u_1-v_1)^2+\cdots+(u_d-v_d)^2) \in \mathbb{F}_p^2$ with $t, u_i, v_i \in A$. We have $|E| = |F| = |A|^{2d+1}$.

It follows from Lemma 2.1 that

$$(3)\quad \left|N(E,F,\lambda)-\frac{|E||F|}{p}\right|\le p^{\frac{1}{2}}\Bigg(\sum_{(e_1,e_2)\in\overline{E}} m_E((e_1,e_2))^2\sum_{(f_1,f_2)\in\overline{F}} m_F((f_1,f_2))^2\Bigg)^{1/2}.$$

We observe that $N(E,F,\lambda)$ equals the number of pairs $(\boldsymbol{x},\boldsymbol{y})\in A^{2d+1}\times A^{2d+1}$ such that $\|\boldsymbol{x}-\boldsymbol{y}\|=\lambda$.

From the setting of E and F, it is not hard to see that

$$(4)\quad \begin{aligned}\sum_{(e_1,e_2)\in\overline{E}} m_E((e_1,e_2))^2 &= |A|E_d((A-A)^2),\\ \sum_{(f_1,f_2)\in\overline{F}} m_F((f_1,f_2))^2 &= |A|E_d((A-A)^2).\end{aligned}$$

Putting (3) and (4) together, we have

$$(5)\quad \left|N(E,F,\lambda)-\frac{|A|^{4d+2}}{p}\right|\le p^{\frac{1}{2}}|A|E_d((A-A)^2).$$

On the other hand, Corollary 2.3 gives us

$$(6)\quad E_d((A-A)^2)\ll d^2(\log|A|)^2\frac{|A|^{4d}}{p}+d^4(\log|A|)^4|A|^{4d-2+\frac{1}{2^{d-1}}}.$$

Substituting (6) into (5), we obtain $N(E,F,\lambda)\sim|A|^{4d+2}p^{-1}$ whenever

$$|A^{2d+1}|\gtrsim_d p^{\frac{2d+2}{2}-\frac{3\cdot 2^{d-2}-d-1}{3\cdot 2^{d-1}-1}}.$$

Since λ is arbitrary in $\mathbb{F}_p$, the theorem follows. □

4. Proofs of Theorems 1.3 and 1.5

The proof of Theorem 1.3 is similar to that of Theorem 1.1, but we need a higher-dimensional version of Lemma 2.1.

Let E and F be multisets in $\mathbb{F}_p^3$. For $\lambda\in\mathbb{F}_p$, let $N(E,F,\lambda)$ be the number of pairs $((e_1,e_2,e_3),(f_1,f_2,f_3))\in E\times F$ such that $e_1f_1+e_2f_2+e_3+f_3=\lambda$. One can follow step by step the proof of Lemma 2.1 to obtain the following.

Lemma 4.1. *Let E, F be multisets in $\mathbb{F}_p^3$. For any $\lambda\in\mathbb{F}_p$, we have*

$$\left|N(E,F,\lambda)-\frac{|E||F|}{p}\right| \le p\Bigg(\sum_{(e_1,e_2,e_3)\in\overline{E}} m_E((e_1,e_2,e_3))^2\sum_{(f_1,f_2,f_3)\in\overline{F}} m_F((f_1,f_2,f_3))^2\Bigg)^{1/2}.$$

We are now ready to prove Theorem 1.3.

Proof of Theorem 1.3. Let λ be an arbitrary element in $\mathbb{F}_p$. Let E be the multiset of points $(2x_1, 2x_2, x_1^2+x_2^2+(y_1-z_1)^2+\cdots+(y_{d-1}-z_{d-1})^2) \in \mathbb{F}_p^3$ with $x_i, y_i, z_i \in A$, and F the multiset of points $(-t_1, -t_2, t_1^2+t_2^2+(u_1-v_1)^2+\cdots+(u_{d-1}-v_{d-1})^2) \in \mathbb{F}_p^3$ with $t_i, u_i, v_i \in A$. We have $|E| = |A|^{2d}$ and $|F| = |A|^{2d}$.

It follows from Lemma 4.1 that

$$\left|N(E, F, \lambda) - \frac{|E||F|}{p}\right| \le p\Bigg(\sum_{(e_1,e_2,e_3)\in\overline{E}} m_E((e_1, e_{2,3}))^2 \sum_{(f_1,f_2,f_3)\in\overline{F}} m_F((f_1, f_2, f_3))^2\Bigg)^{1/2}. \tag{7}$$

We observe that $N(E, F, \lambda)$ is equal to the number of pairs $(\boldsymbol{x}, \boldsymbol{y}) \in A^{2d} \times A^{2d}$ such that $\|\boldsymbol{x} - \boldsymbol{y}\| = \lambda$.

From the setting of E and F, it is not hard to see that

$$\begin{aligned} \sum_{(e_1,e_2,e_3)\in\overline{E}} m_E((e_1, e_2, e_3))^2 &= |A|^2 E_{d-1}((A-A)^2), \\ \sum_{(f_1,f_2,f_3)\in\overline{F}} m_F((f_1, f_2, f_3))^2 &= |A|^2 E_{d-1}((A-A)^2). \end{aligned} \tag{8}$$

Putting (7) and (8) together, we have

$$\left|N(E, F, \lambda) - \frac{|A|^{4d}}{p}\right| \le p|A|^2 E_{d-1}((A-A)^2). \tag{9}$$

On the other hand, Corollary 2.3 gives us

$$E_{d-1}((A-A)^2) \ll d^2(\log|A|)^2 \frac{|A|^{4d-4}}{p} + d^4(\log|A|)^4 |A|^{4d-6+\frac{1}{2^{d-2}}}. \tag{10}$$

Substituting (10) into (9), we obtain $N(E, F, \lambda) \sim |A|^{4d} p^{-1}$ whenever

$$|A^{2d}| \gtrsim_d p^{\frac{2d+1}{2} - \frac{2^d-2d-1}{2^{d+1}-2}}.$$

Since λ is arbitrary in $\mathbb{F}_p$, the theorem follows. □

Proof of Theorem 1.5. The proof of Theorem 1.5 is similar to that of Theorem 1.3 with Corollary 2.5 in the place of Corollary 2.3. □

5. Proof of Theorem 1.7

Let us first recall the prime field analogue of Balog–Wooley decomposition energy due to Rudnev, Shkredov and Stevens [Rudnev et al. 2020].

Theorem 5.1 [Rudnev et al. 2020]. *Let A be a set in $\mathbb{F}_p$ with $|A| \le p^{5/8}$. There exist two disjoint subsets B and C of A such that $A = B \sqcup C$ and*

$$\max\{E^+(B), E^\times(C)\} \lesssim |A|^{14/5},$$

where

$$E^+(B) = |\{(a,b,c,d) \in B^4 : a+b = c+d\}|, \quad E^\times(C) = |\{(a,b,c,d) \in C^4 : ab = cd\}|.$$

We refer the interested reader to [Balog and Wooley 2017] for the original result over $\mathbb{R}$. The most up to date bound for this result over $\mathbb{R}$ is due to Shakan [2019].

The following is another corollary of Lemma 2.2.

Corollary 5.2. *Let A be a set in $\mathbb{F}_p$, and B be a subset of A. For an integer $d \ge 2$, suppose that $|A| \ll p^{1/2+1/(5\cdot 2^{d-1}-2)}$ and $E^+(B) \lesssim |A|^{14/5}$; then we have*

$$E_d((B-B)^2) \ll d^4 (\log|A|)^4 |A|^{4d-2+\frac{1}{5\cdot 2^{d-3}}}.$$

Proof. We prove by induction on d that

$$E_d((B-B)^2) \le 4C^2 d^4 (\log|A|)^4 |A|^{4d-2+\frac{1}{5\cdot 2^{d-3}}},$$

whenever $|A| \le (Cp)^{1/2+1/(5\cdot 2^{d-1}-2)}$, where the constant C comes from Lemma 2.2.

The base case $d = 2$ follows from Lemma 2.2 and the facts that $E_1((B-B)^2) \ll E^+(B)$ and $|B| \le |A|$.

Suppose the corollary holds for $d-1 \ge 2$. We now show that it also holds for general d. Indeed, it follows from Lemma 2.2 that

$$E_d((B-B)^2) \le Cd^2 (\log|A|)^2 \left(\frac{|B|^{4d}}{p} + |B|^{2d+1} \sqrt{E_{d-1}((B-B)^2)} \right).$$

On the other hand, by the inductions hypothesis, we have

$$\begin{aligned} E_{d-1}((B-B)^2) &\le 4C^2 (d-1)^4 (\log|A|)^4 |A|^{4d-6+\frac{1}{5\cdot 2^{d-4}}} \\ &\le 4C^2 d^4 (\log|A|)^4 |A|^{4d-6+\frac{1}{5\cdot 2^{d-4}}}. \end{aligned}$$

Thus, using the fact that $|B| \le |A|$, we obtain

$$\begin{aligned} E_d((B-B)^2) &\le d^2 (\log|A|)^2 \left(C\frac{|A|^{4d}}{p} + 2C^2 d^2 (\log|A|)^2 |A|^{4d-2+\frac{1}{5\cdot 2^{d-3}}} \right) \\ &\le 4C^2 d^4 (\log|A|)^4 |A|^{4d-2+\frac{1}{5\cdot 2^{d-3}}}, \end{aligned}$$

whenever $|A| \le (Cp)^{1/2+1/(5\cdot 2^{d-1}-2)}$. □

Using the same argument, we also have another corollary of Lemma 2.4.

Corollary 5.3. *Let A be a set in $\mathbb{F}_p$, and C be a subset of A. For an integer $d \geq 2$, suppose that $|A| \ll p^{1/2+1/(5\cdot 2^{d-1}-2)}$ and $E^\times(C) \lesssim |A|^{14/5}$; then we have*

$$E_d(C \cdot C) \ll d^4(\log|A|)^4|A|^{4d-2+\frac{1}{5\cdot 2^{d-3}}}.$$

We are now ready to prove Theorem 1.7.

Proof of Theorem 1.7. It follows from Theorem 5.1 that there exist two disjoint subsets B and C of A such that $A = B \sqcup C$ and $\max\{E^+(B), E^\times(C)\} \lesssim |A|^{14/5}$. One now can apply Corollaries 5.2 and 5.3 to derive

$$\max\{E_d((B-B)^2), E_d(C \cdot C)\} \ll d^4(\log|A|)^4|A|^{4d-2+\frac{1}{5\cdot 2^{d-3}}}. \qquad \square$$

Acknowledgments:

The authors are grateful to the referee for useful comments and corrections. Pham was supported by Swiss National Science Foundation grant P400P2-183916. Vinh was supported by the National Foundation for Science and Technology Development Project. 101.99-2019.318.

References

[Balog and Wooley 2017] A. Balog and T. D. Wooley, "A low-energy decomposition theorem", *Q. J. Math.* **68**:1 (2017), 207–226. MR

[Bennett et al. 2017] M. Bennett, D. Hart, A. Iosevich, J. Pakianathan, and M. Rudnev, "Group actions and geometric combinatorics in $\mathbb{F}_q^d$", *Forum Math.* **29**:1 (2017), 91–110. MR

[Bourgain et al. 2004] J. Bourgain, N. Katz, and T. Tao, "A sum-product estimate in finite fields, and applications", *Geom. Funct. Anal.* **14**:1 (2004), 27–57. MR

[Chapman et al. 2012] J. Chapman, M. B. Erdoğan, D. Hart, A. Iosevich, and D. Koh, "Pinned distance sets, k-simplices, Wolff's exponent in finite fields and sum-product estimates", *Math. Z.* **271**:1-2 (2012), 63–93. MR

[Du et al. 2018] X. Du, L. Guth, Y. Ou, H. Wang, B. Wilson, and R. Zhang, "Weighted restriction estimates and application to Falconer distance set problem", 2018. arXiv

[Erdös 1946] P. Erdös, "On sets of distances of n points", *Amer. Math. Monthly* **53** (1946), 248–250. MR

[Falconer 1985] K. J. Falconer, "On the Hausdorff dimensions of distance sets", *Mathematika* **32**:2 (1985), 206–212. MR

[Glibichuk and Konyagin 2007] A. A. Glibichuk and S. V. Konyagin, "Additive properties of product sets in fields of prime order", pp. 279–286 in *Additive combinatorics*, edited by M. B. N. Andrew Granville and J. Solymos, CRM Proc. Lecture Notes **43**, Amer. Math. Soc., Providence, RI, 2007. MR

[Glibichuk and Rudnev 2009] A. Glibichuk and M. Rudnev, "On additive properties of product sets in an arbitrary finite field", *J. Anal. Math.* **108** (2009), 159–170. MR

[Guth and Katz 2015] L. Guth and N. H. Katz, "On the Erdős distinct distances problem in the plane", *Ann. of Math.* (2) **181**:1 (2015), 155–190. MR

[Guth et al. 2020] L. Guth, A. Iosevich, Y. Ou, and H. Wang, "On Falconer's distance set problem in the plane", *Invent. Math.* **219**:3 (2020), 779–830. MR

[Hanson et al. 2016] B. Hanson, B. Lund, and O. Roche-Newton, "On distinct perpendicular bisectors and pinned distances in finite fields", *Finite Fields Appl.* **37** (2016), 240–264. MR

[Hart and Iosevich 2008] D. Hart and A. Iosevich, "Sums and products in finite fields: an integral geometric viewpoint", pp. 129–135 in *Radon transforms, geometry, and wavelets*, edited by G. Ólafsson et al., Contemp. Math. **464**, Amer. Math. Soc., Providence, RI, 2008. MR

[Hart et al. 2011] D. Hart, A. Iosevich, D. Koh, and M. Rudnev, "Averages over hyperplanes, sum-product theory in vector spaces over finite fields and the Erdős–Falconer distance conjecture", *Trans. Amer. Math. Soc.* **363**:6 (2011), 3255–3275. MR

[Iosevich and Rudnev 2007] A. Iosevich and M. Rudnev, "Erdős distance problem in vector spaces over finite fields", *Trans. Amer. Math. Soc.* **359**:12 (2007), 6127–6142. MR

[Murphy and Petridis 2019] B. Murphy and G. Petridis, "An example related to the Erdős–Falconer question over arbitrary finite fields", *Bull. Hellenic Math. Soc.* **63** (2019), 38–39. MR

[Pham et al. 2017] D. H. Pham, T. Pham, and L. A. Vinh, "An improvement on the number of simplices in $\mathbb{F}_q^d$", *Discrete Appl. Math.* **221** (2017), 95–105. MR

[Pham et al. 2019] T. Pham, L. A. Vinh, and F. de Zeeuw, "Three-variable expanding polynomials and higher-dimensional distinct distances", *Combinatorica* **39**:2 (2019), 411–426. MR

[Rudnev 2018] M. Rudnev, "On the number of incidences between points and planes in three dimensions", *Combinatorica* **38**:1 (2018), 219–254. MR

[Rudnev et al. 2020] M. Rudnev, I. Shkredov, and S. Stevens, "On the energy variant of the sum-product conjecture", *Rev. Mat. Iberoam.* **36**:1 (2020), 207–232. MR

[Shakan 2019] G. Shakan, "On higher energy decompositions and the sum-product phenomenon", *Math. Proc. Cambridge Philos. Soc.* **167**:3 (2019), 599–617. MR

[Shkredov 2018] I. D. Shkredov, "On asymptotic formulae in some sum-product questions", *Trans. Moscow Math. Soc.* **79** (2018), 231–281. MR

[Solymosi and Vu 2008] J. Solymosi and V. H. Vu, "Near optimal bounds for the Erdős distinct distances problem in high dimensions", *Combinatorica* **28**:1 (2008), 113–125. MR

[Vinh 2011] L. A. Vinh, "The Szemerédi–Trotter type theorem and the sum-product estimate in finite fields", *European J. Combin.* **32**:8 (2011), 1177–1181. MR

[de Zeeuw 2016] F. de Zeeuw, "A short proof of Rudnev's point-plane incidence bound", 2016. arXiv

Received June 18, 2020. Revised July 22, 2020.

THANG PHAM
DEPARTMENT OF MATHEMATICS
ETH, ZURICH
ZURICH
SWITZERLAND
phamanhthang.vnu@gmail.com

LÊ ANH VINH
VIETNAM INSTITUTE OF EDUCATIONAL SCIENCES
HA NOI
VIETNAM
vinhle@vnies.edu.vn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 2, 2020
https://doi.org/10.2140/pjm.2020.309.453

ELLIPTIC GRADIENT ESTIMATES FOR A PARABOLIC EQUATION WITH V-LAPLACIAN AND APPLICATIONS

JIAN-HONG WANG AND YU ZHENG

In this paper, we establish a local elliptic gradient estimate for positive bounded solutions to a parabolic equation concerning the V-Laplacian

$$\big(\Delta_V - \partial_t - q(x,t)\big)u(x,t) = F(u(x,t))$$

on an n-dimensional complete Riemannian manifold with the Bakry–Émery Ricci curvature Ric_V bounded below, which is weaker than the m-Bakry–Émery Ricci curvature Ric_V^m bounded below considered by Chen and Zhao (2018). As applications, we obtain the local elliptic gradient estimates for the cases that $F(u) = au \ln u$ and au^γ. Moreover, we prove parabolic Liouville theorems for the solutions satisfying some growth restriction near infinity and study the problem about conformal deformation of the scalar curvature. In the end, we also derive a global Bernstein-type gradient estimate for the above equation with $F(u) = 0$.

1. Introduction and main results

In this paper, we will study local and global elliptic gradient estimates for positive smooth bounded solutions $u(x,t)$ to a parabolic equation

$$\big(\Delta_V - \partial_t - q(x,t)\big)u(x,t) = F(u(x,t)) \tag{1.1}$$

on an n-dimensional complete Riemannian manifold (M^n, g), where $q(x,t)$ is a function which is C^2 in the x-variable and C^1 in the t-variable, and $F(u)$ is a C^2 function of u.

The Equation (1.1) is an important extension of the Schrödinger equation. The V-Laplacian is defined by

$$\Delta_V := \Delta - V \cdot \nabla,$$

where V is a smooth vector field.

Yu Zheng is supported by the NSFC (No.11671141).
MSC2010: primary 58J35; secondary 35B53, 35K05.
Keywords: gradient estimate, Liouville theorem, V-Laplacian, Bakry–Émery Ricci curvature, parabolic equation.

As in [Chen et al. 2012], we define the m-Bakry–Émery Ricci curvature

$$\mathrm{Ric}_V^m := \mathrm{Ric} + \tfrac{1}{2}\mathcal{L}_V g - \frac{1}{m} V^* \otimes V^*$$

for any number $m \geq 0$, where Ric is the Ricci tensor, $\mathcal{L}_V$ is the Lie derivative in the direction of V, and V^* is the metric-dual of V. When $m = 0$, it means that $V \equiv 0$ and Ric_V^m returns to the usual Ricci tensor. The (∞)-Bakry–Émery Ricci curvature is

$$\mathrm{Ric}_V := \mathrm{Ric} + \tfrac{1}{2}\mathcal{L}_V g.$$

It is easy to see that $\mathrm{Ric}_V^m \geq c$ implies $\mathrm{Ric}_V \geq c$, but not vice versa.

If $\mathrm{Ric}_V = \lambda g$ for some real constant λ, then (M^n, g) is a Ricci soliton, which is a natural extension of Einstein metric. A Ricci soliton is called shrinking, steady or expanding, if $\lambda > 0$, $\lambda = 0$ or $\lambda < 0$, respectively. In particular, when $V = \nabla f$ for some function $f \in C^\infty(M)$, since $\mathcal{L}_{\nabla f} g = 2\,\mathrm{Hess}\, f$ (Hess is the Hessian with respect to the metric g), a Ricci soliton becomes a gradient Ricci soliton. The gradient Ricci soliton plays an important role in the formation of singularities of the Ricci flow, and has been studied by many authors; see [Cao 2010; Hamilton 1995] for nice surveys.

Relating to the V-Laplacian, we have, for $\mathrm{Ric}_V^m (0 < m < \infty)$, the following Bochner formula:

$$\begin{aligned} \tfrac{1}{2}\Delta_V |\nabla u|^2 &= |\nabla\nabla u|^2 + \langle \nabla u, \nabla \Delta_V u\rangle + \mathrm{Ric}_V^m(\nabla u, \nabla u) + \frac{1}{m}|\langle V, \nabla u\rangle|^2 \\ &\geq \frac{(\Delta_V u)^2}{m+n} + \langle \nabla u, \nabla \Delta_V u\rangle + \mathrm{Ric}_V^m(\nabla u, \nabla u). \end{aligned} \tag{1.2}$$

When $m = \infty$, we have

$$\tfrac{1}{2}\Delta_V |\nabla u|^2 = |\nabla\nabla u|^2 + \langle \nabla u, \nabla \Delta_V u\rangle + \mathrm{Ric}_V(\nabla u, \nabla u). \tag{1.3}$$

The formula (1.2) looks like the classical Bochner formula on an $(m+n)$-dimensional manifold with Ricci tensor, therefore many geometric results for the Laplacian on n-dimensional manifolds with Ric bounded below can be possibly extended to the V-Laplacian on $(m+n)$-dimensional manifolds with Ric_V^m bounded below, such as the mean curvature comparison theorem, the volume comparison theorem, etc. However, for Ric_V, due to lack of the term $\frac{1}{m}|\langle V, \nabla u\rangle|^2$, there seems essential obstacles to obtaining some important conclusions when Ric_V is only bounded below.

To the best of our knowledge, the gradient estimate technique was originated by S.-T. Yau [1975] in the 1970s, who first proved a gradient estimate for the harmonic function on manifolds. In the 1980s, this technique was developed by Li and Yau [1986] for the heat equation on manifolds, and yielded a parabolic gradient estimate (sometimes called Li–Yau's gradient estimate). More precisely,

Theorem A [Li and Yau 1986]. *Let M be a complete manifold with dimension $n \geq 2$, $\operatorname{Ric}(M) \geq -k$, $k \geq 0$. Suppose u is any positive solution to the heat equation in $B(x_0, R) \times [t_0 - T, t_0]$. Then*

$$\frac{|\nabla u|^2}{u^2} - \frac{u_t}{u} \leq \frac{c_n}{R^2} + \frac{c_n}{T} + c_n k$$

in $B\big(x_0, \frac{R}{2}\big) \times \big[t_0 - \frac{T}{2}, t_0\big]$. Here c_n depends only on n.

Li and Yau [1986] also proved a parabolic gradient estimate for the Schrödinger equation

$$\big(\Delta - \partial_t - q(x,t)\big)u(x,t) = 0,$$

which can be seen as the special case of (1.1) (see [Li and Yau 1986, Theorem 1.2]).

In the 1990s, R. Hamilton [1993] proved a global elliptic gradient estimate (sometimes called Hamilton's gradient estimate) for the heat equation on closed manifolds.

Theorem B [Hamilton 1993]. *Let M be an n-dimensional closed manifold with $\operatorname{Ric} \geq -K$ for nonnegative constant K, and let u be a positive solution of the heat equation*

$$\frac{\partial u}{\partial t} = \Delta u$$

with $u \leq A$ for all time. Then

$$t|\nabla u|^2 \leq (1 + 2Kt)u^2 \ln \frac{A}{u}.$$

Hamilton's gradient estimate requires that the equation be defined on closed manifolds. Later, Souplet and Zhang [2006] proved a local elliptic gradient estimate (sometimes called Souplet–Zhang's gradient estimate) for the heat equation on noncompact manifolds by inserting a necessary logarithmic correction term.

Theorem C [Souplet and Zhang 2006]. *Let M be a Riemannian manifold with dimension $n \geq 2$ and $\operatorname{Ric} \geq -k$, $k \geq 0$. Assume u is any positive solution to the heat equation in $Q_{R,T} = B_{x_0}(R) \times [t_0 - T, t_0] \subset M \times (-\infty, \infty)$ with $u \leq M$. Then there exists a dimensional constant c such that*

$$|\nabla \ln u| \leq c\left(\frac{1}{R} + \frac{1}{\sqrt{T}} + \sqrt{k}\right)\left(1 + \ln \frac{M}{u}\right)$$

in $Q_{R/2,\,T/2}$.

Moreover, if M has nonnegative Ricci curvature and u is any positive solution of the heat equation on $M \times (0, \infty)$, then there exist dimensional constants c_1, c_2 such that

$$|\nabla \ln u| \leq c_1 \frac{1}{\sqrt{t}}\left(c_2 + \ln \frac{u(x, 2t)}{u(x,t)}\right)$$

for all $x \in M$ and $t > 0$.

Apart from the above theorems, Li–Yau's, Hamilton's and Souplet–Zhang's gradient estimates have been generalized to other linear and nonlinear equations on Riemannian manifolds, see, e.g., [Brighton 2013; Chow and Hamilton 1997; Chen and Qiu 2016; Cao and Zhang 2011; Huang and Ma 2016; Li and Xu 2011; Li 1991; 2012; 2015; Ma 2006; Ruan 2007; Wu 2015; Yang 2008; Zhu 2016].

We now give the main theorems, a local elliptic (Souplet–Zhang's) gradient estimate for positive smooth solutions to (1.1), which is based on the arguments of Souplet and Zhang [2006] for the heat equation, Brighton [2013] for the f-harmonic function and J.-Y. Wu [2015] for the f-heat equation. It is important that our gradient estimate does not depend on any assumption on V.

Theorem 1.1. *Let (M^n, g) be an n-dimensional complete Riemannian manifold, and let $B_{x_0}(R)$ be a geodesic ball of radius R around x_0 and $R \geq 2$. Assume* $\mathrm{Ric}_V \geq -k$ *in $B_{x_0}(R)$ for some constant $k \geq 0$. Let u be a positive solution of (1.1) in $Q_{R,T} = B_{x_0}(R) \times [t_0 - T, t_0] \subset M^n \times (-\infty, \infty)$ with $u \leq M$ for some positive constant M, where $t_0 \in \mathbb{R}$ and $T > 0$. Then there exists a dimensional constant $C(n)$ such that*

$$\begin{aligned} (1.4)\quad & |\nabla \ln u| \\ & \leq C(n)\left(\sqrt{\frac{1+|\delta|}{R}} + \frac{1}{\sqrt{t-t_0+T}} + \sqrt{k+\lambda_1+\lambda_2+\lambda_3+\lambda_4}\right)\left(1+\ln\frac{M}{u}\right) \end{aligned}$$

in $Q_{R/2,\,T}$ with $t \neq t_0 - T$. Here

$$\lambda_1 = -\min\left\{0, \min_{Q_{R,T}} \frac{F(u)}{u}\right\}, \quad \lambda_2 = -\min\left\{0, \min_{Q_{R,T}}\left(F'(u) - \frac{F(u)}{u}\right)\right\},$$

$$\lambda_3 = 2\max_{Q_{R,T}}\{q^-\} \quad (q^- = \max\{-q, 0\} \textit{ is the negative part of } q),$$

$$\lambda_4 = \max_{Q_{R,T}}\left|\nabla\sqrt{|q|}\right|,$$

which are nonnegative constants, and $\delta = \max_{\{x \mid d(x,x_0)=1\}} \Delta_V r(x)$.

Remark 1.2. Theorem 1.1 describes local elliptic gradient estimate under only Ric_V bounded below, whose assumption on Ric_V is obviously weaker than the assumption on $\mathrm{Ric}_V^m (m < \infty)$ which was considered by Chen and Zhao [2018].

On one hand, we apply Theorem 1.1 to analyze the existence of solutions to the special case of (1.1). Moreover, we study the problem about conformal deformation of the scalar curvature on complete noncompact manifolds; see Corollary 2.7 in Section 2.

Theorem 1.3. *Let (M^n, g) be an n-dimensional complete manifold with* $\mathrm{Ric}_V \geq 0$. *Consider the equation*

$$(1.5)\qquad \big(\Delta_V - \partial_t - q(x)\big)u(x,t) = au^{\gamma}$$

for some constants $a \geq 0$ *and* $\gamma > 1$. *Suppose that* $q(x) \neq 0$ *and*

$$q^- = o(R^{-1}), \quad \left|\nabla\sqrt{|q|}\right| = o(R^{-1}) \quad as \;\; R \to \infty.$$

Then there does not exist any positive ancient solution (*that is*, *a solution defined in all space and negative time*) *to* (1.5) *such that* $u(x,t) = o(r(x)^{1/2} + |t|^{1/2})$ *near infinity. In particular*, *if* $V \equiv 0$, *we only assume* $u(x,t) = o(r(x) + |t|^{1/2})$ *near infinity.*

Remark 1.4. If $q(x)$ is a positive constant, it naturally satisfies the growth conditions of $q(x)$ in Theorem 1.3. There also exist many nontrivial functions $q(x)$ satisfying these growth conditions, such as $q(x) = e^{-x}$ in R^1.

On the other hand, we apply Theorem 1.1 to prove the parabolic Liouville theorem for the V-heat equation under certain growth conditions of solutions. This result is similar to the cases of the heat equation and the f-heat equation, obtained by Souplet and Zhang [2006] and Wu [2015], respectively.

Theorem 1.5. *Let* (M^n, g) *be an* n*-dimensional complete Riemannian manifold with* $\mathrm{Ric}_V \geq 0$. *Let* $u(x,t)$ *be an eternal solution* (*that is*, *a solution defined in all space and time*) *to*

$$(\Delta_V - \partial_t)u = 0. \tag{1.6}$$

Then the following conclusions hold.

(i) *If* $u(x,t) = e^{o(r^{1/2}(x)+|t|^{1/2})}$ *near infinity and* $u > 0$, *then* u *is a constant.*

(ii) *If* $u(x,t) = o(r^{1/2}(x) + |t|^{1/2})$ *near infinity*, *then* u *is a constant.*

Remark 1.6. The growth condition of u is necessary. For example, let $u = e^{x+2t}$, $V = \nabla f$, $f = -x$ in $\mathbb{R}^1$. Then u is a nonconstant positive eternal solution to (1.6). Any complete shrinking or steady Ricci solitons satisfy $\mathrm{Ric}_V \geq 0$, hence Theorem 1.5 also holds on shrinking or steady Ricci solitons.

In the end, we derive a global Bernstein-type gradient estimate for positive bounded solution to (1.1) with $F(u) = 0$ on complete Riemannian manifolds with Ric_V bounded below, which is inspired by the works of Kotschwar [2007] for the heat equation and Wu [2015] for the f-heat equation.

Theorem 1.7. *Let* (M^n, g) *be an* n*-dimensional complete Riemannian manifold with* $\mathrm{Ric}_V \geq -k$ *for some constant* $k \geq 0$. *Let* u *be a solution to*

$$\big(\Delta_V - \partial_t - q(x,t)\big)u(x,t) = 0 \tag{1.7}$$

in $M^n \times [0, T]$ *with* $0 < T < \infty$. *Suppose that*

$$0 < u \leq M, \quad q^-(x,t) \leq \alpha \quad and \quad \left|\nabla\sqrt{|q|}\right| \leq \beta,$$

where M, α, β are positive constants. Then there exists an absolute constant C such that

$$t|\nabla u|^2 \leq CM^2\big(1+(k+\alpha+\beta)T\big)$$

in $M^n \times [0, T]$.

The rest of this paper is organized as follows. In Section 2, we give a useful lemma and a cut-off function to prove Theorem 1.1 via the maximum principle and V-Laplacian comparison theorem. As applications of Theorem 1.1, we prove Theorems 1.3 and 1.5. Moreover, we apply Theorem 1.3 to discuss Yamabe type problems and obtain Corollary 2.7. In Section 3, we prove Theorem 1.7 by using another local elliptic gradient estimate for (1.7).

2. Local elliptic gradient estimate

In this section, we first follow the techniques of [Souplet and Zhang 2006; Brighton 2013; Wu 2015] to prove Theorem 1.1. Notice that $0 < u \leq M$ is a solution of (1.1). Define a smooth function

$$f = \ln \frac{u}{M} \quad \text{in } \; Q_{R,T}.$$

Obviously, $f \leq 0$. By (1.1), we have

$$(\Delta_V - \partial_t) f + |\nabla f|^2 - q(x,t) = A(f), \quad \text{where } \; A(f) = \frac{F(Me^f)}{Me^f} = \frac{F(u)}{u}. \tag{2.1}$$

Let

$$g := |\nabla \ln(1-f)|^2 = \frac{|\nabla f|^2}{(1-f)^2},$$

we have the following lemma.

Lemma 2.1. *Let (M^n, g) be an n-dimensional complete Riemannian manifold with* $\mathrm{Ric}_V \geq -k$ *for some constant $k \geq 0$. Then g satisfies*

$$(\Delta_V - \partial_t) g \geq \frac{2f}{1-f}\langle \nabla f, \nabla g\rangle + 2(1-f)g^2 - 2(k+\lambda_1+\lambda_2+\lambda_3)g - 2\lambda_4^2, \tag{2.2}$$

where $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are the same as in Theorem 1.1.

Proof. Let $h = \ln(1-f)$, i.e., $g = |\nabla h|^2$. By the Bochner formula (1.3) and $\mathrm{Ric}_V \geq -k$, we have

$$\begin{aligned} \Delta_V g &= 2\big(|\nabla\nabla h|^2 + \langle \nabla h, \nabla \Delta_V h\rangle + \mathrm{Ric}_V(\nabla h, \nabla h)\big) \\ &\geq 2\big(\langle \nabla h, \nabla \Delta_V h\rangle - k|\nabla h|^2\big). \end{aligned} \tag{2.3}$$

Since $\nabla h = -\frac{\nabla f}{1-f}$ and

$$\Delta_V h = \Delta h - \langle V, \nabla h\rangle = -\frac{(1-f)\Delta f + |\nabla f|^2}{(1-f)^2} + \left\langle V, \frac{\nabla f}{1-f}\right\rangle = -\frac{\Delta_V f}{1-f} - \frac{|\nabla f|^2}{(1-f)^2}.$$

By a direct computation, we obtain

$$\langle \nabla h, \nabla \Delta_V h\rangle = \frac{2\nabla\nabla f(\nabla f, \nabla f)}{(1-f)^3} + \frac{2|\nabla f|^4}{(1-f)^4} + \frac{\langle \nabla f, \nabla \Delta_V f\rangle}{(1-f)^2} + \frac{|\nabla f|^2 \Delta_V f}{(1-f)^3}.$$

Hence, (2.3) becomes

$$\Delta_V g \geq \frac{4\nabla\nabla f(\nabla f, \nabla f)}{(1-f)^3} + \frac{4|\nabla f|^4}{(1-f)^4} + \frac{2\langle \nabla f, \nabla \Delta_V f\rangle}{(1-f)^2} + \frac{2|\nabla f|^2 \Delta_V f}{(1-f)^3} - 2k\frac{|\nabla f|^2}{(1-f)^2}. \tag{2.4}$$

By using (2.1), we obtain

$$\begin{aligned}\partial_t g &= \frac{2\langle \nabla f, \nabla f_t\rangle}{(1-f)^2} + \frac{2|\nabla f|^2 f_t}{(1-f)^3} \\ &= \frac{2\langle \nabla f, \nabla \Delta_V f\rangle}{(1-f)^2} + \frac{4\nabla\nabla f(\nabla f, \nabla f)}{(1-f)^2} + \frac{2|\nabla f|^2 \Delta_V f}{(1-f)^3} + \frac{2|\nabla f|^4}{(1-f)^3} \\ &\quad - \frac{2\langle \nabla f, \nabla(q+A)\rangle}{(1-f)^2} - \frac{2(q+A)|\nabla f|^2}{(1-f)^3}.\end{aligned} \tag{2.5}$$

Combining (2.4) and (2.5), we have

$$\begin{aligned}(\Delta_V - \partial_t)g \geq{}& \frac{4\nabla\nabla f(\nabla f, \nabla f)}{(1-f)^3} + \frac{4|\nabla f|^4}{(1-f)^4} - \frac{4\nabla\nabla f(\nabla f, \nabla f)}{(1-f)^2} - \frac{2|\nabla f|^4}{(1-f)^3} \\ &- 2k\frac{|\nabla f|^2}{(1-f)^2} + \frac{2\langle \nabla f, \nabla(q+A)\rangle}{(1-f)^2} + \frac{2(q+A)|\nabla f|^2}{(1-f)^3}.\end{aligned} \tag{2.6}$$

Since $g = \frac{|\nabla f|^2}{(1-f)^2}$, then

$$\langle \nabla g, \nabla f\rangle = \frac{2\nabla\nabla f(\nabla f, \nabla f)}{(1-f)^2} + \frac{2|\nabla f|^4}{(1-f)^3},$$

which implies

$$\begin{aligned}0 ={}& -2\langle \nabla g, \nabla f\rangle + \frac{4\nabla\nabla f(\nabla f, \nabla f)}{(1-f)^2} + \frac{4|\nabla f|^4}{(1-f)^3} \\ &+ \frac{1}{1-f}\left(2\langle \nabla g, \nabla f\rangle - \frac{4|\nabla f|^4}{(1-f)^3}\right) - \frac{4\nabla\nabla f(\nabla f, \nabla f)}{(1-f)^3}.\end{aligned} \tag{2.7}$$

From (2.7), we know that

$$\frac{4\nabla\nabla f(\nabla f,\nabla f)}{(1-f)^3}+\frac{4|\nabla f|^4}{(1-f)^4}-\frac{4\nabla\nabla f(\nabla f,\nabla f)}{(1-f)^2}-\frac{2|\nabla f|^4}{(1-f)^3} = \frac{2f}{1-f}\langle\nabla g,\nabla f\rangle+\frac{2|\nabla f|^4}{(1-f)^3}.$$

Using the above equality, (2.6) becomes

$$(\Delta_V-\partial_t)g\ge\frac{2f}{1-f}\langle\nabla f,\nabla g\rangle+2(1-f)g^2-2\Big(k-A'(f)-\frac{A}{1-f}\Big)g - \frac{2}{1-f}|\nabla q|\sqrt{g}+\frac{2q}{1-f}g. \tag{2.8}$$

Since $0<\frac{1}{1-f}\le 1$ and

$$2|\nabla q|\sqrt{g}\le 2|q|g+\frac{|\nabla q|^2}{2|q|}=2|q|g+2\big|\nabla\sqrt{|q|}\big|^2. \tag{2.9}$$

Noticing that the inequality (2.9) is trivial when $q=0$. Hence,

$$-\frac{2|\nabla q|\sqrt{g}}{1-f}\ge-\frac{2|q|g}{1-f}-\frac{2\big|\nabla\sqrt{|q|}\big|^2}{1-f}\ge-\frac{2|q|g}{1-f}-2\big|\nabla\sqrt{|q|}\big|^2.$$

Using this inequality, (2.8) can be rewritten as

$$(\Delta_V-\partial_t)g\ge\frac{2f}{1-f}\langle\nabla f,\nabla g\rangle+2(1-f)g^2-2\bigg(k-A'(f)-\frac{A}{1-f}+\frac{2q^-}{1-f}\bigg)g -2\big|\nabla\sqrt{|q|}\big|^2.$$

By the definitions of $\lambda_1,\lambda_2,\lambda_3,\lambda_4$, we have

$$-\frac{A}{1-f}\le\frac{A^-}{1-f}\le A^-=-\min\{0,\frac{F(u)}{u}\}\le\lambda_1,$$
$$-A'(f)=-\Big(F'(u)-\frac{F(u)}{u}\Big)\le\lambda_2,\qquad \frac{2q^-}{1-f}\le\lambda_3,\qquad \big|\nabla\sqrt{|q|}\big|\le\lambda_4.$$

Hence, (2.2) immediately follows. □

In order to prove Theorem 1.1, we introduce a useful cut-off function which originated with Li and Yau [1986] (see also [Bailesteanu et al. 2010; Souplet and Zhang 2006]).

Lemma 2.2. *Fix $t_0\in\mathbb{R}$ and $T>0$. For any given $\tau\in(t_0-T,t_0]$, there exists a smooth function $\bar\eta:[0,\infty)\times[t_0-T,t_0]\to\mathbb{R}$ satisfying following propositions*:

(1) $0\le\bar\eta(r,t)\le 1$ *in* $[0,R]\times[t_0-T,t_0]$, *and it is supported in a subset of* $[0,R]\times[t_0-T,t_0]$.

(2) $\bar{\eta}(r,t)=1$ *and* $\partial_r\bar{\eta}(r,t)=0$ *in* $\left[0,\frac{R}{2}\right]\times[\tau,t_0]$ *and* $\left[0,\frac{R}{2}\right]\times[t_0-T,t_0]$, *respectively.*

(3) $|\partial_t\bar{\eta}|\le C\bar{\eta}^{1/2}/(\tau-t_0+T)$ *in* $[0,\infty)\times[t_0-T,t_0]$ *for some constant* $C>0$, *and* $\bar{\eta}(r,t_0-T)=0$ *for all* $r\in[0,\infty)$.

(4) $-(C_\epsilon/R)\bar{\eta}^\epsilon\le\partial_r\bar{\eta}\le 0$ *and* $|\partial_r^2\bar{\eta}|\le C_\epsilon\bar{\eta}^\epsilon/R^2$ *in* $[0,\infty)\times[t_0-T,t_0]$ *for every* $\epsilon\in(0,1)$ *with some constant* C_ϵ *depending on* ϵ.

Now, we apply Lemmas 2.1 and 2.2 to prove Theorem 1.1 via the maximum principle and the V-Laplacian comparison theorem [Wu 2018, Theorem 2.1] in a local space-time supported set.

Proof of Theorem 1.1. Fix any number $\tau\in(t_0-T,t_0]$, we will show that (1.4) holds at (x,τ) for all $x\in B_{x_0}\left(\frac{R}{2}\right)$. The assertion of theorem will immediately follows due to τ is arbitrary.

Choose a cut-off function $\bar{\eta}(r,t)$ satisfying the propositions of Lemma 2.2. Let $\eta:M\times[t_0-T,t_0]\to\mathbb{R}$ such that $\eta(x,t)=\bar{\eta}(r(x),t)$, where $r(x)=d(x,x_0)$. It is easy to see that $\eta(x,t)$ is supported in $Q_{R,T}$. Our aim is to calculate $(\Delta_V-\partial_t)(\eta g)$ and estimate each term at a space-time point where ηg attains its maximum.

From Lemma 2.1, we conclude

$$(\Delta_V-\partial_t)(\eta g)-\left(\frac{2f}{1-f}\nabla f+2\frac{\nabla\eta}{\eta}\right)\nabla(\eta g)\\ \ge 2(1-f)\eta g^2-\left(\frac{2f}{1-f}\langle\nabla f,\nabla\eta\rangle\right)g-2\frac{|\nabla\eta|^2}{\eta}g+(\Delta_V\eta)g-\eta_t g\\ -2(k+\lambda_1+\lambda_2+\lambda_3)\eta g-2\lambda_4^2\eta. \tag{2.10}$$

Assume

$$(\eta g)(x_1,t_1)=\max_{B_{x_0}(R)\times[t_0-T,\tau]}(\eta g).$$

We may assume $(\eta g)(x_1,t_1)>0$, otherwise, $g(x,\tau)\le 0$ and (1.4) naturally holds at (x,τ) whenever $d(x,x_0)<\frac{R}{2}$. Notice that $t_1\ne t_0-T$ due to $(\eta g)(x_1,t_1)>0$. We may also assume that $\eta(x,t)$ is smooth at (x_1,t_1) by the standard Calabi argument [1958]. Using the maximum principle, at (x_1,t_1), we have

$$\Delta_V(\eta g)\le 0,\quad (\eta g)_t\ge 0\quad\text{and}\quad \nabla(\eta g)=0.$$

Hence, (2.10) can be simplified as

$$2(1-f)\eta g^2\le\left(\frac{2f}{1-f}\langle\nabla f,\nabla\eta\rangle+2\frac{|\nabla\eta|^2}{\eta}\right)g-(\Delta_V\eta)g+\eta_t g\\ +2(k+\lambda_1+\lambda_2+\lambda_3)\eta g+2\lambda_4^2\eta. \tag{2.11}$$

at (x_1,t_1). In the following, we will estimate each term on the right hand side of (2.11) and obtain the desired gradient estimate in Theorem 1.1. We will get it by two steps.

Case I. Assume $x_1 \notin B_{x_0}(1)$. Since $\mathrm{Ric}_V \geq -k$ and $d(x_1, x_0) \geq 1$ in $B_{x_0}(R)$, $R \geq 2$, by the V-Laplacian comparison theorem, we have

$$\Delta_V r(x_1) \leq \delta + k(R-1),$$

where $\delta = \max_{\{x|d(x,x_0)=1\}} \Delta_V r(x)$.

Below the Young's inequality and Lemma 2.2 will be repeatedly used in the following estimate. Let c be a constant depending only on n whose value may change from line to line. Then we have the following inequalities:

$$\begin{aligned} \frac{2f}{1-f}\langle \nabla f, \nabla \eta\rangle g &\leq 2|f||\nabla \eta| g^{3/2} \\ &= 2\big[\eta(1-f)g^2\big]^{3/4} \cdot \frac{|f||\nabla \eta|}{\big[\eta(1-f)\big]^{3/4}} \\ &\leq \eta(1-f)g^2 + c\frac{f^4|\nabla\eta|^4}{(1-f)^3\eta^3} \\ &\leq \eta(1-f)g^2 + c\frac{f^4}{R^4(1-f)^3}. \end{aligned} \tag{2.12}$$

$$2\frac{|\nabla\eta|^2}{\eta} g \leq \tfrac{1}{8}\eta g^2 + 8\frac{|\nabla\eta|^4}{\eta^3} \leq \tfrac{1}{8}\eta g^2 + \frac{c}{R^4}. \tag{2.13}$$

$$\begin{aligned} -(\Delta_V \eta) g &= -\big(\partial_r^2\bar\eta + \partial_r\bar\eta \Delta_V r\big) g \\ &\leq \Big(|\partial_r^2\bar\eta| + \big(|\delta| + k(R-1)\big)|\partial_r\bar\eta|\Big) g \\ &= \eta^{1/2} g \frac{|\partial_r^2\bar\eta|}{\bar\eta^{1/2}} + \big(|\delta| + k(R-1)\big)\eta^{1/2} g \frac{|\partial_r\bar\eta|}{\bar\eta^{1/2}} \\ &\leq \tfrac{1}{8}\eta g^2 + c\left(\frac{|\partial_r^2\bar\eta|^2}{\bar\eta} + \delta^2\frac{|\partial_r\bar\eta|^2}{\bar\eta} + k^2(R-1)^2\frac{|\partial_r\bar\eta|^2}{\bar\eta}\right) \\ &\leq \tfrac{1}{8}\eta g^2 + \frac{c}{R^4} + \frac{c\delta^2}{R^2} + ck^2\frac{(R-1)^2}{R^2} \\ &\leq \tfrac{1}{8}\eta g^2 + \frac{c}{R^4} + \frac{c\delta^2}{R^2} + ck^2. \end{aligned} \tag{2.14}$$

$$|\eta_t| g = \eta^{1/2} g \frac{|\bar\eta_t|}{\bar\eta^{1/2}} \leq \tfrac{1}{8}\eta g^2 + 8\frac{|\bar\eta_t|^2}{\bar\eta} \leq \tfrac{1}{8}\eta g^2 + \frac{c}{(\tau - t_0 + T)^2}. \tag{2.15}$$

$$2(k + \lambda_1 + \lambda_2 + \lambda_3)\eta g \leq \tfrac{1}{8}\eta g^2 + 8(k + \lambda_1 + \lambda_2 + \lambda_3)^2. \tag{2.16}$$

$$2\lambda_4^2 \eta \leq 2\lambda_4^2. \tag{2.17}$$

Substituting (2.12)–(2.17) into the right hand side of (2.11), at (x_1, t_1), we have

$$2\eta(1-f)g^2 \le \eta(1-f)g^2 + \tfrac{1}{2}\eta g^2 + c\frac{f^4}{R^4(1-f)^3} + \frac{c}{R^4} + \frac{c\delta^2}{R^2} + ck^2 + \frac{c}{(\tau - t_0 + T)^2} + 8(k+\lambda_1+\lambda_2+\lambda_3)^2 + 2\lambda_4^2.$$

Since $1-f \ge 1$, the above estimate implies

$$\begin{aligned}(\eta g^2)(x_1,t_1) &\le \frac{1}{1-f}\left(\tfrac{1}{2}\eta g^2 + c\frac{f^4}{R^4(1-f)^3} + \frac{c}{R^4} + \frac{c\delta^2}{R^2} + ck^2 + \frac{c}{(\tau - t_0 + T)^2}\right. \\ &\qquad \left. + 8(k+\lambda_1+\lambda_2+\lambda_3)^2 + 2\lambda_4^2\right) \\ &\le \tfrac{1}{2}\eta g^2 + c\frac{f^4}{R^4(1-f)^4} + \frac{c}{R^4} + \frac{c\delta^2}{R^2} + ck^2 + \frac{c}{(\tau - t_0 + T)^2} \\ &\qquad + 8(k+\lambda_1+\lambda_2+\lambda_3)^2 + 2\lambda_4^2 \\ &\le \tfrac{1}{2}\eta g^2 + \frac{c}{R^4} + \frac{c\delta^2}{R^2} + \frac{c}{(\tau - t_0 + T)^2} + c(k+\lambda_1+\lambda_2+\lambda_3+\lambda_4)^2.\end{aligned}$$

Since $0 \le \eta \le 1$, then

$$\begin{aligned}(\eta g)^2(x_1,t_1) &\le (\eta g^2)(x_1,t_1) \\ &\le \frac{c}{R^4} + \frac{c\delta^2}{R^2} + \frac{c}{(\tau - t_0 + T)^2} + c(k+\lambda_1+\lambda_2+\lambda_3+\lambda_4)^2.\end{aligned}$$

Since $\eta(x,\tau) = 1$ when $x \in B_{x_0}\left(\frac{R}{2}\right)$ by the proposition (2) in Lemma 2.2 and $R \ge 2$, we obtain

$$\begin{aligned}g(x,\tau) = (\eta g)(x,\tau) &\le (\eta g)(x_1,t_1) \\ &\le \frac{c}{R^2} + \frac{c|\delta|}{R} + \frac{c}{\tau - t_0 + T} + c(k+\lambda_1+\lambda_2+\lambda_3+\lambda_4) \\ &\le \frac{c(1+|\delta|)}{R} + \frac{c}{\tau - t_0 + T} + c(k+\lambda_1+\lambda_2+\lambda_3+\lambda_4)\end{aligned}$$

for all $x \in B_{x_0}\left(\frac{R}{2}\right)$.

Case II. Assume $x_1 \in B_{x_0}(1) \subset B_{x_0}\left(\frac{R}{2}\right)$ when $R \ge 2$. In this case, η is a constant in space direction in $Q_{R/2,T}$. Hence (2.11) can be simplified as

$$2(1-f)\eta g^2 \le \eta_t g + 2(k+\lambda_1+\lambda_2+\lambda_3)\eta g + 2\lambda_4^2\eta$$

at (x_1, t_1). Since $1-f \ge 1$, this implies

$$2\eta g^2 \le \eta_t g + 2(k+\lambda_1+\lambda_2+\lambda_3)\eta g + 2\lambda_4^2\eta$$

at (x_1, t_1). Substituting (2.15)–(2.17) into the right hand side of the above inequality, we have

$$(\eta g^2)(x_1,t_1) \le c\Big(\frac{1}{(\tau-t_0+T)^2} + (k+\lambda_1+\lambda_2+\lambda_3)^2+\lambda_4^2\Big).$$

Since $0 \le \eta \le 1$, we have

$$(\eta g)(x_1,t_1) \le c\Big(\frac{1}{\tau-t_0+T} + k+\lambda_1+\lambda_2+\lambda_3+\lambda_4\Big).$$

Since $\eta(x,\tau)=1$ whenever $x \in B_{x_0}\big(\frac{R}{2}\big)$,

$$\begin{aligned} g(x,\tau) &= (\eta g)(x,\tau) \le (\eta g)(x_1,t_1) \\ &\le c\Big(\frac{1}{\tau-t_0+T} + k+\lambda_1+\lambda_2+\lambda_3+\lambda_4\Big) \end{aligned}$$

for all $x \in B_{x_0}\big(\frac{R}{2}\big)$.

Combining the above two cases, by the definition of g and the fact that $\tau \in (t_0-T, t_0]$ was chosen arbitrarily, we obtain

$$\frac{|\nabla f|}{1-f}(x,t) \le C(n)\bigg(\sqrt{\frac{1+|\delta|}{R}} + \frac{1}{\sqrt{t-t_0+T}} + \sqrt{k+\lambda_1+\lambda_2+\lambda_3+\lambda_4}\bigg)$$

for any $(x,t) \in Q_{R/2,T}$ with $t \ne t_0 - T$. Substituting $f = \ln\frac{u}{M}$ into the above estimate completes the proof of theorem. □

Remark 2.3. If $q(x,t) = F(u) = 0$, $V = \nabla f$ for some function f, then Theorem 1.1 is the same as Theorem 1.1 in [Wu 2018] for the weighted heat equation. From the proof of Theorem 1.1, we know that the term $\sqrt{(1+|\delta|)/R}$ in (1.4) can be changed into $\frac{1}{R}$ whenever $V \equiv 0$.

When V is bounded, we can prove another gradient estimate of (1.1) in any geodesic ball. Its proof is similar to that of Theorem 1.1 except that the V-Laplacian comparison theorem in Theorem 1.1 is replaced by another V-Laplacian comparison theorem [Wu 2018, Theorem 2.2]. We only provide the conclusion and omit the proof.

Theorem 2.4. *Let (M^n, g) be an n-dimensional complete Riemannian manifold. Assume $\mathrm{Ric}_V \ge -k$ and $|V| \le a$ in $B_{x_0}(R)$ for some nonnegative constants k and a. Let $0 < u \le M$ be a solution of (1.1) in $Q_{R,T}$. Then there exists a dimensional constant $C(n)$ such that*

$$|\nabla \ln u| \le C(n)\bigg(\sqrt{\frac{1+a}{R}} + \frac{1}{\sqrt{t-t_0+T}} + \sqrt{k+\lambda_1+\lambda_2+\lambda_3+\lambda_4}\bigg)\Big(1+\ln\frac{M}{u}\Big) \tag{2.18}$$

in $Q_{R/2,T}$ with $t \ne t_0 - T$, where $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are the same as in Theorem 1.1.

As applications of Theorem 1.1, we can derive some corollaries by considering the special cases of (1.1). More precisely,

$$\big(\Delta_V - \partial_t - q(x,t)\big)u = au\ln u, \tag{2.19}$$

$$\big(\Delta_V - \partial_t - q(x,t)\big)u = au^{\gamma}, \tag{2.20}$$

where a and γ are constants.

When $V \equiv 0$, the elliptic version of (2.19) is closely related to the gradient Ricci soliton; (see [Ma 2006]). In fact, consider the gradient Ricci soliton

$$\mathrm{Ric} + \nabla\nabla f + \lambda g = 0,$$

where λ is a constant. Taking the trace of the above equality, we have

$$R + \Delta f + n\lambda = 0.$$

Using the contracted Bianchi identity and Ricci identity, then

$$|\nabla f|^2 + R - 2\lambda f = c$$

for some constant c. Hence

$$|\nabla f|^2 - \Delta f - 2\lambda f = n\lambda + c.$$

Setting $u = e^{-f}$, we obtain

$$\Delta u - (c + n\lambda)u = -2\lambda u\ln u.$$

When $V \equiv 0$, the elliptic version of (2.20) is related to conformal deformation of the scalar curvature on manifolds. In fact, for any n-dimensional ($n \geq 3$) manifold, consider a conformal metric $\tilde{g} = u^{4/(n-2)}g$ for some positive function u. Then the scalar curvature $\tilde{s}$ of metric $\tilde{g}$ related to the scalar curvature s of metric g is given by

$$\Delta u - \frac{n-2}{4(n-1)}su + \frac{n-2}{4(n-1)}\tilde{s}u^{(n+2)/(n-2)} = 0. \tag{2.21}$$

We have known that if M is compact and $\tilde{s}$ is a constant, the existence of u is the well-known Yamabe problem which has been solved by R. Schoen [1984] (see also [Lee and Parker 1987; Mastrolia et al. 2012] for more details).

Corollary 2.5. *Let (M^n, g) be an n-dimensional complete Riemannian manifold with $\mathrm{Ric}_V \geq -k$ for some constant $k \geq 0$. Let $0 < u \leq M$ be a smooth solution of (2.19) with $a \leq 0$ in $M^n \times [t_0 - T, t_0]$. Suppose that $q^- \leq c_1$ and $\big|\nabla\sqrt{|q|}\big| \leq c_2$ for some constants c_1, c_2. Then there exists a dimensional constant $c(n)$ such that*

$$|\nabla \ln u| \leq c(n)\left(\frac{1}{\sqrt{t-t_0+T}} + \sqrt{k - a\ln(\max\{M,1\}) - a + 2c_1 + c_2}\right) \cdot \left(1 + \ln\frac{M}{u}\right) \tag{2.22}$$

in $M^n \times (t_0 - T, t_0]$.

Proof. Since $F(u) = au\ln u$ $(a \le 0)$ and $0 < u \le M$, by the definitions of λ_i $(i = 1, 2, 3, 4)$, it is easy to obtain

$$\lambda_1 = -a\ln(\max\{M, 1\}), \quad \lambda_2 = -a, \quad \lambda_3 = 2c_1, \quad \lambda_4 = c_2.$$

Applying Theorem 1.1 and setting $R \to \infty$, (2.22) immediately follows. □

Corollary 2.6. *Let* (M^n, g), q^- *and* $\left|\nabla\sqrt{|q|}\right|$ *be the same as in Corollary 2.5. Let* $0 < u \le M$ *be a smooth solution of (2.20) with* $\gamma > 1$ *in* $M^n \times [t_0-T, t_0]$. *Then there exists a dimensional constant* $c(n)$ *such that*

$$(2.23) \quad |\nabla \ln u| \le c(n)\left(\frac{1}{\sqrt{t-t_0+T}} + \sqrt{k + \frac{\operatorname{sgn} a - 1}{2}(a\gamma M^{\gamma-1}) + 2c_1 + c_2}\right) \cdot \left(1 + \ln\frac{M}{u}\right)$$

in $M^n \times (t_0-T, t_0]$, *where*

$$\operatorname{sgn} a = \begin{cases} 1 & \text{if } a > 0, \\ 0 & \text{if } a = 0, \\ -1 & \text{if } a < 0. \end{cases}$$

Proof. Since $F(u) = au^\gamma$ $(\gamma > 1)$ and $0 < u \le M$, by the direct calculations, we have

(i) If $a \ge 0$, then $\lambda_1 = \lambda_2 = 0$.

(ii) If $a < 0$, then $\lambda_1 = -aM^{\gamma-1}$, $\lambda_2 = -a(\gamma-1)M^{\gamma-1}$.

Using (1.4) and letting $R \to \infty$, the desired result (2.23) follows. □

Next, we apply Theorem 1.1 to prove Theorem 1.3 which analyze the existence of solutions to the parabolic equation (1.5) when the coefficient $q(x)$ and solutions $u(x, t)$ satisfying some growth conditions. Furthermore, we can use Theorem 1.3 to study the problem about conformal deformation of the scalar curvature on complete manifolds.

Proof of Theorem 1.3. From the proof of Theorem 1.1 and Corollary 2.6, since $a \ge 0$, then $\lambda_1 = \lambda_2 = 0$, we get a local elliptic gradient estimate for (1.5):

$$(2.24) \qquad |\nabla \ln u| \le c(n)\left(\sqrt{\frac{1+|\delta|}{R}} + \frac{1}{\sqrt{t-t_0+T}} + \sqrt{\lambda_3 + \lambda_4}\right)\left(1 + \ln\frac{M}{u}\right)$$

for any $(x, t) \in Q_{\frac{R}{2},T}$ with $t \ne t_0 - T$, where λ_3, λ_4 be defined in Theorem 1.1.

For any fixed space-time point (x_0, t_0), by the growth assumptions of $u(x, t)$ and $q(x)$, applying (2.24) to $u(x_0, t_0)$ in the space-time set $Q_{R,R} = B_{x_0}(R) \times [t_0-R, t_0]$, then

$$(2.25) \quad |\nabla \ln u(x_0, t_0)| \le c(n)\left(\sqrt{\frac{1+|\delta|}{R}} + o(R^{-1/2})\right)\left(1 + o(\ln\sqrt{R}) - \ln u(x_0, t_0)\right)$$

for sufficiently large $R \geq 2$.

Notice that $\ln u(x_0, t_0)$ is a fixed value, which implies

$$|\nabla u(x_0, t_0)| = 0 \quad \text{as } R \to \infty.$$

Since (x_0, t_0) was chosen arbitrarily, then $u(x, t) \equiv u(t)$, and Equation (1.5) becomes

$$u'(t) = -q(x)u(t) - au^{\gamma}(t). \tag{2.26}$$

Case I. $a = 0$.

In this case, $u'(t) = -q(x)u(t)$. Since $q(x) \neq 0$, we solve this equation and obtain

$$u(t) = Ce^{-q(x)t}, \tag{2.27}$$

where C is an arbitrary constant.

From (2.27), we know $q(x) = c$ for some constant $c > 0$ due to the growth assumption of q^-. Then

$$u(t) = u(0)e^{-ct} = u(0)e^{c|t|},$$

which contradicts the assumption that $u(x, t) = o(r(x)^{1/2} + |t|^{1/2})$ near infinity.

Case II. $a > 0$.

In this case, (2.26) can be regraded as a one-order linear ordinary equation which has a general solution

$$u^{1-\gamma}(t) = Ce^{(\gamma-1)q(x)t} - \frac{a}{q(x)}, \tag{2.28}$$

where C is an arbitrary constant.

By the same way, we know $q(x) = c$ for some constant $c > 0$, then

$$u^{1-\gamma}(t) = \Big(u^{1-\gamma}(0) + \frac{a}{c}\Big)e^{(\gamma-1)ct} - \frac{a}{c}.$$

Since $a, c, \gamma - 1$ and $u(0)$ are positive constants, which imply

$$u^{1-\gamma}(t) \to -\frac{a}{c} < 0 \quad \text{as } t \to -\infty,$$

this is impossible since $u > 0$.

As for the case $V \equiv 0$, the term $\sqrt{(1 + |\delta|)/R}$ in (2.24) can be changed into $\frac{1}{R}$. Since $u(x, t) = o(r(x) + |t|^{1/2})$ near infinity, we apply (2.24) to $u(x_0, t_0)$ in $Q_{R,R^2} = B_{x_0}(R) \times [t_0 - R^2, t_0]$ and the proof is almost the same as before except that (2.25) is replaced by

$$|\nabla \ln u(x_0, t_0)| \leq c(n)\Big(\frac{1}{R} + o(R^{-1/2})\Big)\big(1 + o(\ln R) - \ln u(x_0, t_0)\big). \qquad \square$$

As an application of Theorem 1.3, we discuss the Yamabe type problem of complete Riemannian manifolds and immediately obtain the following corollary.

Corollary 2.7. *Let (M^n, g) be an n-dimensional $(n \geq 3)$ complete Riemannian manifold with* $\mathrm{Ric} \geq 0$ *and the scalar curvature s of g satisfying*

$$\sup_{B_{x_0}(R)} |\nabla \sqrt{s}| = o(R^{-1})$$

as $R \to \infty$. Then there does not exist complete metric

$$\tilde{g} \in \{u^{4/(n-2)} g \mid 0 < u \in C^\infty(M) \text{ and } u(x) = o(r(x)^{1/2})\},$$

such that the scalar curvature $\tilde{s}$ of $\tilde{g}$ is some nonpositive constant.

Proof. It is equivalent to prove that if $\tilde{s}$ is some nonpositive constant, then there does not exist any positive solution to (2.21) satisfying $u(x) = o(r(x)^{1/2})$. In Theorem 1.3, let

$$u(x,t) = u(x), \quad V = 0, \quad q(x) = \frac{n-2}{4(n-1)} s \geq 0,$$
$$a = -\frac{n-2}{4(n-1)} \tilde{s} \geq 0, \quad \gamma = \frac{n+2}{n-2} > 1,$$

we know that $q(x)$ satisfies the growth conditions in Theorem 1.3 due to the assumptions on s, hence the conclusion follows. □

We also apply Theorem 1.1 to derive the parabolic Liouville theorem for the V-heat equation which extends some known results.

Proof of Theorem 1.5. Since $F(u) = q = 0$, using Theorem 1.1, we have

$$|\nabla \ln u| \leq C(n)\left(\sqrt{\frac{1+|\delta|}{R}} + \frac{1}{\sqrt{t-t_0+T}}\right)\left(1 + \ln \frac{M}{u}\right). \tag{2.29}$$

(i) By the assumption of $u(x,t)$, we have

$$\ln u = o(r^{1/2}(x) + |t|^{1/2})$$

near infinity. For any space-time point (x_0, t_0), we apply (2.29) to $u(x_0, t_0)$ in the space-time set $Q_{R,R} = B_{x_0}(R) \times [t_0 - R, t_0]$, then

$$\frac{|\nabla u(x_0,t_0)|}{u(x_0,t_0)} \leq \frac{C(n,\delta)}{\sqrt{R}}\big(1 + o(\sqrt{R}) - \ln u(x_0,t_0)\big)$$

for sufficiently large $R \geq 2$.

For the fixed value $\ln u(x_0, t_0)$, setting $R \to \infty$ in the above inequality, we get

$$|\nabla u(x_0,t_0)| = 0.$$

Then u is only a time-dependent function due to (x_0, t_0) being arbitrary. Moreover, u is a constant by using (1.6).

(ii) Let $M_R = \sup_{Q_{\sqrt{R},\sqrt{R}}} |u|$. Considering the function $U = u + 2M_{2R}$, then

$$M_{2R} \le U(x,t) \le 3M_{2R}$$

whenever $(x,t) \in Q_{2\sqrt{R},\,2\sqrt{R}}$. For any fixed point (x_0, t_0), applying (2.29) to U, we have

$$\frac{|\nabla u(x_0,t_0)|}{u(x_0,t_0) + 2M_{2R}} \le \frac{C(n,\delta)}{\sqrt{R}}$$

for sufficiently large $R \ge 2$. By the assumption of $u(x,t)$, we have $M_{2R} = o(R^{1/4})$. The conclusion immediately follows by taking $R \to \infty$. □

3. Global elliptic gradient estimate

In this section, we follow the arguments of Kotschwar [2007] and Wu [2015] to prove Theorem 1.7. The key is to derive a local elliptic gradient estimate which is different from Souplet–Zhang's gradient estimate. Our proof is based on the technique of Shi [1989] from the estimation of derivatives of curvature under the Ricci flow. Firstly, we give the following lemma.

Lemma 3.1. *Define*

$$G(x,t) := (4M^2 + u^2)|\nabla u|^2.$$

Under the same assumptions as in Theorem 1.7, we have

$$(\partial_t - \Delta_V)G \le \left(\tfrac{5}{2}k + \tfrac{7}{2}\alpha\right)G - \frac{2}{125M^4}G^2 + 50\beta^2 M^4. \tag{3.1}$$

Proof. By straightforward calculations, we obtain

$$(\partial_t - \Delta_V)u^2 = -2|\nabla u|^2 - 2qu^2.$$
$$(\partial_t - \Delta_V)|\nabla u|^2 \le -2|\nabla\nabla u|^2 - 2u\langle\nabla u, \nabla q\rangle - 2q|\nabla u|^2 + 2k|\nabla u|^2.$$

Then,

$$\begin{aligned}(\partial_t - \Delta_V)G &= (\partial_t - \Delta_V)u^2 \cdot |\nabla u|^2 + (4M^2+u^2)(\partial_t - \Delta_V)|\nabla u|^2 \\ &\qquad - 2\langle\nabla u^2, \nabla|\nabla u|^2\rangle \\ &\le (-2|\nabla u|^2 - 2qu^2)|\nabla u|^2 \\ &\quad + (4M^2+u^2)\big(-2|\nabla\nabla u|^2 - 2u\langle\nabla u, \nabla q\rangle - 2q|\nabla u|^2 + 2k|\nabla u|^2\big) \\ &\qquad - 8u\nabla\nabla u(\nabla u, \nabla u).\end{aligned} \tag{3.2}$$

Since

$$-8u\nabla\nabla u(\nabla u,\nabla u) \le 10u^2|\nabla\nabla u|^2 + \tfrac{8}{5}|\nabla u|^4, \quad 5u^2 \le 4M^2 + u^2 \le 5M^2,$$

the inequality (3.2) can be simplified as

$$(3.3)\quad (\partial_t-\Delta_V)G \le 10kM^2|\nabla u|^2-\tfrac{2}{5}|\nabla u|^4-2qu^2|\nabla u|^2 \\ -2(4M^2+u^2)u\langle\nabla u,\nabla q\rangle-2(4M^2+u^2)q|\nabla u|^2.$$

Using the Young's inequality, then

$$\begin{aligned}-2(4M^2+u^2)u\langle\nabla u,\nabla q\rangle &\le 10M^2u|\nabla u||\nabla q|\\ &\le 2|q|u^2|\nabla u|^2+\tfrac{25}{2}\frac{|\nabla q|^2}{|q|}M^4\\ &=2|q|u^2|\nabla u|^2+50\big|\nabla\sqrt{|q|}\big|^2M^4.\end{aligned}$$

Notice that

$$-2(4M^2+u^2)q|\nabla u|^2\le 10M^2q^-|\nabla u|^2,$$

and

$$|q|-q=2q^-.$$

Hence, (3.3) can be written as

$$\begin{aligned}(\partial_t-\Delta_V)G &\le 10kM^2|\nabla u|^2-\tfrac{2}{5}|\nabla u|^4+4q^-u^2|\nabla u|^2+10M^2q^-|\nabla u|^2\\ &\qquad\qquad +50\big|\nabla\sqrt{|q|}\big|^2M^4\\ &\le -\tfrac{2}{5}|\nabla u|^4+(10k+14\alpha)M^2|\nabla u|^2+50\beta^2M^4,\end{aligned}$$

where we used the assumptions $q^-\le\alpha$ and $\big|\nabla\sqrt{|q|}\big|\le\beta$.

By the definition of G, we know

$$4M^2|\nabla u|^2\le G\le 5M^2|\nabla u|^2.$$

Hence, the inequality (3.1) follows. □

Now, applying Lemma 3.1, we give a proof of Theorem 1.7.

Proof of Theorem 1.7. As in [Li and Yau 1986], we take a cut-off function $\bar\phi(s)$ which is defined in $[0,\infty)$ such that $0\le\bar\phi(s)\le 1$ and

$$\bar\phi(s)=1\quad\text{for } s\in\big[0,\tfrac{1}{2}\big],\qquad \bar\phi(s)=0\quad\text{for } s\in[1,\infty).$$

$\bar\phi(s)$ also satisfies

$$-c_1\le\frac{\bar\phi'(s)}{\bar\phi^{1/2}(s)}\le 0,\quad \bar\phi''(s)\ge -c_2$$

for positive absolute constants c_1 and c_2.

Let $\phi(x)=\bar\phi\big(\frac{r(x)}{R}\big)$ for $R\ge 2$, where $r(x)$ denotes the distance from the fixed point x_0 to x. Using the argument of Calabi [1958], we may assume $\phi(x)\in C^2(M)$

with support in $B_{x_0}(R)$. By direct calculations, we have

$$\frac{|\nabla\phi|^2}{\phi} \le \frac{c_3}{R^2}, \quad \Delta_V\phi = \frac{\bar{\phi}'\Delta_V r}{R} + \frac{\bar{\phi}''}{R^2} \tag{3.4}$$

for some positive absolute constant c_3.

Considering $t\phi G$ in $B_{x_0}(R)\times[0,T]$, using Lemma 3.1, we get

$$\begin{aligned}(\partial_t - \Delta_V)(t\phi G) \le \phi G + t\phi\Big(\big(\tfrac{5}{2}k + \tfrac{7}{2}\alpha\big)G - \frac{2}{125M^4}G^2 + 50\beta^2M^4\Big)\\ - tG\Delta_V\phi - 2t\langle\nabla\phi, \nabla G\rangle\end{aligned} \tag{3.5}$$

Assume

$$(t\phi G)(x_1, t_1) = \max_{B_{x_0}(R)\times(0,T]}(t\phi G).$$

If $t\phi G$ is not identically zero (i.e., u is not a constant in $\operatorname{supp}\phi$), then

$$(t\phi G)(x_1, t_1) > 0.$$

By the maximum principle, at (x_1, t_1),

$$\nabla(t\phi G) = 0, \quad (\partial_t - \Delta_V)(t\phi G) \ge 0.$$

In the following, we will estimate the each term on the right hand side of (3.5) at (x_1, t_1).

Case I. Assume $x_1 \in B_{x_0}(1) \subset B_{x_0}\big(\frac{R}{2}\big)$ because of $R \ge 2$.

In this case, $\phi \equiv 1$ implies $\nabla\phi = \Delta_V\phi = 0$. The inequality (3.5) can be simplified as

$$0 \le G + t\Big(\big(\tfrac{5}{2}k + \tfrac{7}{2}\alpha\big)G - \frac{2G^2}{125M^4} + 50\beta^2M^4\Big).$$

It is equivalent to

$$\frac{2}{125M^4}tG^2 - \Big(\big(\tfrac{5}{2}k + \tfrac{7}{2}\alpha\big)t + 1\Big)G - 50\beta^2M^4t \le 0$$

at (x_1, t_1). Since $0 < t \le T$, we have

$$\frac{2}{125M^4}(tG)^2 - \Big(\big(\tfrac{5}{2}k + \tfrac{7}{2}\alpha\big)T + 1\Big)tG - 50\beta^2M^4T^2 \le 0.$$

At this time, (x_1, t_1) is also the maximum point of tG in $B_{x_0}\big(\frac{R}{2}\big)\times(0,T]$. Hence, we obtain

$$(tG)(x,t) \le (tG)(x_1, t_1) \le CM^4\big(1 + (k + \alpha + \beta)T\big)$$

in $B_{x_0}\big(\frac{R}{2}\big)\times[0,T]$.

Case II. Assume $x_1 \notin B_{x_0}(1)$.

In this case, since $\mathrm{Ric}_V \geq -k$, $d(x_1, x_0) \geq 1$ and $R \geq 2$, by the V-Laplacian comparison theorem, we have

$$\Delta_V r(x_1) \leq \delta + k(R-1) \leq |\delta| + k(R-1),$$

where $\delta = \max_{\{x \mid d(x,x_0)=1\}} \Delta_V r(x)$. Hence,

$$\Delta_V \phi \geq -\frac{c_1}{R}\big(|\delta| + k(R-1)\big) - \frac{c_2}{R^2}. \tag{3.6}$$

By using Lemma 3.1, (3.4) and (3.6), at (x_1, t_1), we have

$$\begin{aligned}
0 &\leq (\partial_t - \Delta_V)(t\phi G) \qquad (3.7)\\
&= \phi G + t\phi(\partial_t - \Delta_V)G - tG\Delta_V\phi - 2t\langle \nabla\phi, \nabla G\rangle \\
&= \left(\phi + 2t\frac{|\nabla\phi|^2}{\phi} - t\Delta_V\phi\right)G + t\phi(\partial_t - \Delta_V)G - 2\left\langle \nabla(t\phi G), \frac{\nabla\phi}{\phi}\right\rangle \\
&\leq \left(1 + \frac{c_4 t}{R^2} + \frac{c_1 t}{R}\big(|\delta| + k(R-1)\big)\right)G \\
&\qquad\qquad + t\phi\left(\left(\tfrac{5}{2}k + \tfrac{7}{2}\alpha\right)G - \frac{2G^2}{125M^4} + 50\beta^2 M^4\right) \\
&\leq -\frac{c_5}{M^4}t\phi G^2 + \left(1 + c_6\left(\frac{1+|\delta|}{R} + k + \alpha\right)T\right)G + c_7\beta^2 M^4 T.
\end{aligned}$$

Multiplying both sides of (3.7) by $t\phi$ and using $t\phi \leq T$, we have

$$\frac{c_5}{M^4}(t\phi G)^2 - \left(1 + c_6\left(\frac{1+|\delta|}{R} + k + \alpha\right)T\right)(t\phi G) - c_7\beta^2 M^4 T^2 \leq 0.$$

We solve this inequality and obtain that

$$(t\phi G)(x_1, t_1) \leq c_8 M^4\left(1 + \left(\frac{1+|\delta|}{R} + k + \alpha + \beta\right)T\right).$$

Notice that the above constants $c_i\,(i = 1, 2, \ldots, 8)$ are all absolute positive constants. Consequently,

$$\begin{aligned}
(tG)(x,t) = (t\phi G)(x,t) &\leq (t\phi G)(x_1, t_1) \\
&\leq c_8 M^4\left(1 + \left(\frac{1+|\delta|}{R} + k + \alpha + \beta\right)T\right)
\end{aligned}$$

for any $(x,t) \in B_{x_0}\left(\frac{R}{2}\right) \times [0, T]$.

Combining the above two cases, we obtain

$$(tG)(x,t) \leq CM^4\left(1 + \left(\frac{1+|\delta|}{R} + k + \alpha + \beta\right)T\right)$$

for any $(x,t) \in B_{x_0}\left(\frac{R}{2}\right) \times [0, T]$.

Since $G \geq 4M^2|\nabla u|^2$, the theorem follows by taking $R \to \infty$. □

Acknowledgement

The authors would like to thank the referee for helpful comments and suggestions to improve this article.

References

[Bailesteanu et al. 2010] M. Bailesteanu, X. Cao, and A. Pulemotov, "Gradient estimates for the heat equation under the Ricci flow", *J. Funct. Anal.* **258**:10 (2010), 3517–3542. MR Zbl

[Brighton 2013] K. Brighton, "A Liouville-type theorem for smooth metric measure spaces", *J. Geom. Anal.* **23**:2 (2013), 562–570. MR Zbl

[Calabi 1958] E. Calabi, "An extension of E. Hopf's maximum principle with an application to Riemannian geometry", *Duke Math. J.* **25** (1958), 45–56. MR Zbl

[Cao 2010] H.-D. Cao, "Recent progress on Ricci solitons", pp. 1–38 in *Recent advances in geometric analysis*, edited by Y.-I. Lee et al., Adv. Lect. Math. (ALM) **11**, International Press, Somerville, MA, 2010. MR Zbl

[Cao and Zhang 2011] X. Cao and Z. Zhang, "Differential Harnack estimates for parabolic equations", pp. 87–98 in *Complex and differential geometry*, edited by W. Ebeling et al., Springer Proc. Math. **8**, Springer, 2011. MR Zbl

[Chen and Qiu 2016] Q. Chen and H. Qiu, "Gradient estimates and Harnack inequalities of a nonlinear parabolic equation for the V-Laplacian", *Ann. Global Anal. Geom.* **50**:1 (2016), 47–64. MR Zbl

[Chen and Zhao 2018] Q. Chen and G. Zhao, "Li–Yau type and Souplet–Zhang type gradient estimates of a parabolic equation for the V-Laplacian", *J. Math. Anal. Appl.* **463**:2 (2018), 744–759. MR Zbl

[Chen et al. 2012] Q. Chen, J. Jost, and H. Qiu, "Existence and Liouville theorems for V-harmonic maps from complete manifolds", *Ann. Global Anal. Geom.* **42**:4 (2012), 565–584. MR Zbl

[Chow and Hamilton 1997] B. Chow and R. S. Hamilton, "Constrained and linear Harnack inequalities for parabolic equations", *Invent. Math.* **129**:2 (1997), 213–238. MR Zbl

[Hamilton 1993] R. S. Hamilton, "A matrix Harnack estimate for the heat equation", *Comm. Anal. Geom.* **1**:1 (1993), 113–126. MR Zbl

[Hamilton 1995] R. S. Hamilton, "The formation of singularities in the Ricci flow", pp. 7–136 in *Surveys in differential geometry* (Cambridge, MA, 1993), vol. II, edited by S.-T. Yau, International Press, Cambridge, MA, 1995. MR Zbl

[Huang and Ma 2016] G. Huang and B. Ma, "Hamilton–Souplet–Zhang's gradient estimates for two types of nonlinear parabolic equations under the Ricci flow", *J. Funct. Spaces* (2016), art. id. 2894207, 7 pp. MR Zbl

[Kotschwar 2007] B. L. Kotschwar, "Hamilton's gradient estimate for the heat kernel on complete manifolds", *Proc. Amer. Math. Soc.* **135**:9 (2007), 3013–3019. MR Zbl

[Lee and Parker 1987] J. M. Lee and T. H. Parker, "The Yamabe problem", *Bull. Amer. Math. Soc.* (*N.S.*) **17**:1 (1987), 37–91. MR Zbl

[Li 1991] J. Li, "Gradient estimates and Harnack inequalities for nonlinear parabolic and nonlinear elliptic equations on Riemannian manifolds", *J. Funct. Anal.* **100**:2 (1991), 233–256. MR Zbl

[Li 2012] P. Li, *Geometric analysis*, Cambridge Studies in Advanced Mathematics **134**, Cambridge University Press, 2012. MR Zbl

[Li 2015] Y. Li, "Li–Yau–Hamilton estimates and Bakry–Emery–Ricci curvature", *Nonlinear Anal.* **113** (2015), 1–32. MR Zbl

[Li and Xu 2011] J. Li and X. Xu, "Differential Harnack inequalities on Riemannian manifolds I: Linear heat equation", *Adv. Math.* **226**:5 (2011), 4456–4491. MR Zbl

[Li and Yau 1986] P. Li and S.-T. Yau, "On the parabolic kernel of the Schrödinger operator", *Acta Math.* **156**:3-4 (1986), 153–201. MR Zbl

[Ma 2006] L. Ma, "Gradient estimates for a simple elliptic equation on complete non-compact Riemannian manifolds", *J. Funct. Anal.* **241**:1 (2006), 374–382. MR Zbl

[Mastrolia et al. 2012] P. Mastrolia, M. Rigoli, and A. G. Setti, *Yamabe-type equations on complete, noncompact manifolds*, Progress in Mathematics **302**, Springer, 2012. MR Zbl

[Ruan 2007] Q. Ruan, "Elliptic-type gradient estimate for Schrödinger equations on noncompact manifolds", *Bull. Lond. Math. Soc.* **39**:6 (2007), 982–988. MR Zbl

[Schoen 1984] R. Schoen, "Conformal deformation of a Riemannian metric to constant scalar curvature", *J. Differential Geom.* **20**:2 (1984), 479–495. MR Zbl

[Shi 1989] W.-X. Shi, "Deforming the metric on complete Riemannian manifolds", *J. Differential Geom.* **30**:1 (1989), 223–301. MR Zbl

[Souplet and Zhang 2006] P. Souplet and Q. S. Zhang, "Sharp gradient estimate and Yau's Liouville theorem for the heat equation on noncompact manifolds", *Bull. London Math. Soc.* **38**:6 (2006), 1045–1053. MR Zbl

[Wu 2015] J.-Y. Wu, "Elliptic gradient estimates for a weighted heat equation and applications", *Math. Z.* **280**:1-2 (2015), 451–468. MR Zbl

[Wu 2018] J.-Y. Wu, "Myers' type theorem with the Bakry–Émery Ricci tensor", *Ann. Global Anal. Geom.* **54**:4 (2018), 541–549. MR Zbl

[Yang 2008] Y. Yang, "Gradient estimates for a nonlinear parabolic equation on Riemannian manifolds", *Proc. Amer. Math. Soc.* **136**:11 (2008), 4095–4102. MR Zbl

[Yau 1975] S. T. Yau, "Harmonic functions on complete Riemannian manifolds", *Comm. Pure Appl. Math.* **28** (1975), 201–228. MR Zbl

[Zhu 2016] X. Zhu, "Gradient estimates and Liouville theorems for linear and nonlinear parabolic equations on Riemannian manifolds", *Acta Math. Sci. Ser. B* (*Engl. Ed.*) **36**:2 (2016), 514–526. MR Zbl

Received October 12, 2018. Revised December 12, 2019.

JIAN-HONG WANG
SCHOOL OF STATISTICS AND MATHEMATICS
SHANGHAI LIXIN UNIVERSITY OF ACCOUNTING AND FINANCE
SHANGHAI
CHINA
jianhongwang@163.com

YU ZHENG
SCHOOL OF MATHEMATICAL SCIENCES
EAST CHINA NORMAL UNIVERSITY
SHANGHAI
CHINA
zhyu@math.ecnu.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 2, 2020
https://doi.org/10.2140/pjm.2020.309.475

OPTIMAL L^2 EXTENSION OF SECTIONS FROM SUBVARIETIES IN WEAKLY PSEUDOCONVEX MANIFOLDS

XIANGYU ZHOU AND LANGFENG ZHU

We obtain optimal L^2 extension of holomorphic sections of a holomorphic vector bundle from subvarieties in weakly pseudoconvex Kähler manifolds. Moreover, in the case of a line bundle the Hermitian metric is allowed to be singular.

1. Introduction and main results

The L^2 extension problem is an important topic in several complex variables and complex geometry. Many generalizations and applications (see [Manivel 1993; Ohsawa 1995; 2001; Siu 1996; 1998; Berndtsson 1996; Demailly 2000; 2012a; Siu 2002; McNeal and Varolin 2007; Berndtsson and Păun 2008], etc.) have been obtained since the original work of Ohsawa and Takegoshi [1987]. Recent progress concerns the optimal L^2 extension and its applications (see [Guan et al. 2011; Zhu et al. 2012; Guan and Zhou 2012; 2015a; 2015c; Błocki 2013; Zhou 2015; Ohsawa 2015; Berndtsson and Lempert 2016; Cao 2017; Zhou and Zhu 2018], etc.).

Most recently, several general L^2 extension theorems with optimal estimates were proved in [Guan and Zhou 2015c] for holomorphic sections defined on subvarieties in Stein or projective manifolds. In [Demailly 2016], several L^2 extension theorems were obtained for holomorphic sections defined on subvarieties in weakly pseudoconvex Kähler manifolds.

In this paper, we prove an optimal L^2 extension theorem, which generalizes the main theorems in [Guan and Zhou 2015c] to weakly pseudoconvex Kähler manifolds and a main theorem in [Zhou and Zhu 2018], and optimizes a main theorem in [Demailly 2016] (cf. Theorem 2.8 and Remark 2.9 in [Demailly 2016]).

Let us recall some usual notions and some definitions in [Demailly 2016].

Zhou was partially supported by the National Natural Science Foundation of China (No. 11688101 and No. 11431013). Zhu was partially supported by the National Natural Science Foundation of China (No. 11201347, No. 11671306 and No. 12022110) and the China Scholarship Council. Zhu is the corresponding author.

MSC2010: primary 32D15, 32J25, 32Q15, 32U05, 32W05; secondary 14F18, 32L10.

Keywords: optimal L^2 extension, plurisubharmonic function, multiplier ideal sheaf, strong openness, weakly pseudoconvex manifold, Kähler manifold.

Definition 1.1. A function $\psi : X \to [-\infty, +\infty)$ on a complex manifold X is said to be quasi-plurisubharmonic if ψ is locally the sum of a plurisubharmonic function and a smooth function. In addition, we say that ψ has neat analytic singularities if every point $x \in X$ possesses an open neighborhood U on which ψ can be written as

$$\psi = c \log \sum_{1 \le j \le j_0} |g_j|^2 + u,$$

where c is a nonnegative number, $g_j \in \mathcal{O}_X(U)$ and $u \in C^\infty(U)$.

Definition 1.2. If ψ is a quasi-plurisubharmonic function on a complex manifold X, the multiplier ideal sheaf $\mathcal{I}(\psi)$ is the coherent analytic subsheaf of $\mathcal{O}_X$ defined by

$$\mathcal{I}(\psi)_x = \left\{ f \in \mathcal{O}_{X,x} : \exists U \ni x, \int_U |f|^2 e^{-\psi} \, d\lambda < +\infty \right\},$$

where U is an open coordinate neighborhood of x, and $d\lambda$ is the Lebesgue measure in the corresponding open chart of $\mathbb{C}^n$. We say that the singularities of ψ are log canonical along the zero variety $Y = V(\mathcal{I}(\psi))$ if $\mathcal{I}((1-\varepsilon)\psi)\big|_Y = \mathcal{O}_X\big|_Y$ for every $\varepsilon > 0$.

If ω is a Kähler metric on X, we let $dV_{X,\omega} := \omega^n/n!$ be the corresponding Kähler volume element, where $n = \dim X$. In case ψ has log canonical singularities along $Y = V(\mathcal{I}(\psi))$, one can associate in a natural way a measure $dV_{X,\omega}[\psi]$ on the set $Y^0 = Y_{\text{reg}}$ of regular points of Y as follows.

Definition 1.3. If $g \in C_c(Y^0)$ is a compactly supported nonnegative continuous function on Y^0 and $\tilde{g}$ is a compactly supported nonnegative continuous extension of g to X such that $(\operatorname{supp} \tilde{g}) \cap Y \subset Y^0$, then we set

$$\int_{Y^0} g \, dV_{X,\omega}[\psi] = \varlimsup_{t \to -\infty} \int_{\{x \in X : t < \psi(x) < t+1\}} \tilde{g} e^{-\psi} \, dV_{X,\omega}.$$

Remark 1.4. By Hironaka's desingularization theorem (Theorem 2.8), it is not hard to see that the limit in the above definition does not depend on the extension $\tilde{g}$ and then $dV_{X,\omega}[\psi]$ is well defined on Y^0 (see Proposition 4.5 in [Demailly 2016] for a proof).

Remark 1.5. The definition of $dV_{X,\omega}[\psi]$ here has a slight difference with the one in [Guan and Zhou 2015c]. In fact, if we denote the measure in [Guan and Zhou 2015c] by $d\widehat{V}_{X,\omega}[\psi]$, the integral $\int_{Y^0} g \, dV_{X,\omega}[\psi]$ here is equal to

$$\sum_{1 \le j \le n} \frac{\pi^j}{j!} \int_{Y_{n-j}} g \, d\widehat{V}_{X,\omega}[\psi],$$

where Y_{n-j} is the $(n-j)$-dimensional component of Y_{reg}.

We will define a class of functions before the statement of our main theorem.

Definition 1.6. Let $\alpha_0 \in (-\infty, +\infty]$ and $\alpha_1 \in [0, +\infty)$. When $\alpha_0 \neq +\infty$, let $\mathfrak{R}_{\alpha_0,\alpha_1}$ be the class of functions defined by

$$\Big\{ R \in C^\infty(-\infty, \alpha_0] : R > 0,\ R \text{ is decreasing near } -\infty,$$

$$\overline{\lim_{t\to-\infty}}\, e^t R(t) < +\infty, \quad C_R := \int_{-\infty}^{\alpha_0} \frac{1}{R(t)}\,dt < +\infty,$$

$$\int_t^{\alpha_0}\left(\frac{\alpha_1}{R(\alpha_0)} + \int_{t_2}^{\alpha_0}\frac{dt_1}{R(t_1)}\right)dt_2 + \frac{(\alpha_1)^2}{R(\alpha_0)} < R(t)\left(\frac{\alpha_1}{R(\alpha_0)} + \int_t^{\alpha_0}\frac{dt_1}{R(t_1)}\right)^2 \tag{1-1}$$

$$\text{for all } t \in (-\infty, \alpha_0)\Big\}.$$

When $\alpha_0 = +\infty$, we replace $R \in C^\infty(-\infty, \alpha_0]$ with $R \in C^\infty(-\infty, +\infty)$ and $R(+\infty) := \underline{\lim}_{t\to+\infty} R(t) \in (0, +\infty]$ in the above definition of $\mathfrak{R}_{\alpha_0,\alpha_1}$.

Remark 1.7. The numbers α_0, α_1 and the function $R(t)$ are equal to the numbers A, $1/\delta$ and the function $1/(c_A(-t)e^t)$ which are defined just before the main theorems in [Guan and Zhou 2015c]. If $\alpha_0 \neq +\infty$ and R is decreasing on $(-\infty, \alpha_0]$, then (1-1) holds for all $t \in (-\infty, \alpha_0)$. If $\alpha_0 = +\infty$, then (1-1) implies that $\int_t^{+\infty}(\alpha_1/R(+\infty))\,dt_2 < +\infty$ for all $t \in (-\infty, +\infty)$. Therefore, $\alpha_1/R(+\infty) = 0$, i.e., $\alpha_1 = 0$ or $R(+\infty) = +\infty$.

Theorem 1.8 (The main theorem). *Let $R \in \mathfrak{R}_{\alpha_0,\alpha_1}$. Let (X, ω) be a weakly pseudoconvex complex n-dimensional manifold possessing a Kähler metric ω, and ψ be a quasi-plurisubharmonic function on X with neat analytic singularities. Let Y be the analytic subvariety of X defined by $Y = V(\mathcal{I}(\psi))$ and assume that ψ has log canonical singularities along Y. Let L (resp. E) be a holomorphic line bundle (resp. a holomorphic vector bundle) over X equipped with a singular Hermitian metric $h = h_L$ (resp. a smooth Hermitian metric $h = h_E$), which is written locally as $e^{-\phi_L}$ for some quasi-plurisubharmonic function ϕ_L with respect to a local holomorphic frame of L. Assume that*

(i) $\sqrt{-1}\Theta_h + \sqrt{-1}\partial\bar\partial\psi$ *is semipositive on $X \setminus \{\psi = -\infty\}$ in the sense of currents* (*resp. in the sense of Nakano*),

and that there is a continuous function $\alpha < \alpha_0$ on X such that the following two assumptions hold:

(ii) $\sqrt{-1}\Theta_h + \sqrt{-1}\partial\bar\partial\psi + (1/\tilde\chi(\alpha))\sqrt{-1}\partial\bar\partial\psi$ *is semipositive on $X \setminus \{\psi = -\infty\}$ in the sense of currents* (*resp. in the sense of Nakano*),

(iii) $\psi \le \alpha$,

where $\widetilde{\chi}(t)$ is the function

$$\frac{\displaystyle\int_t^{\alpha_0}\left(\frac{\alpha_1}{R(\alpha_0)}+\int_{t_2}^{\alpha_0}\frac{dt_1}{R(t_1)}\right)dt_2+\frac{(\alpha_1)^2}{R(\alpha_0)}}{\displaystyle\frac{\alpha_1}{R(\alpha_0)}+\int_t^{\alpha_0}\frac{dt_1}{R(t_1)}}. \tag{1-2}$$

Then for every section $f \in H^0(Y^0, (K_X \otimes L)|_{Y^0})$ (resp. $f \in H^0(Y^0, (K_X \otimes E)|_{Y^0})$) on $Y^0 = Y_{\mathrm{reg}}$ such that

$$\int_{Y^0} |f|^2_{\omega,h}\, dV_{X,\omega}[\psi] < +\infty, \tag{1-3}$$

there exists a section $F \in H^0(X, K_X \otimes L)$ (resp. $F \in H^0(X, K_X \otimes E)$) such that $F = f$ on Y^0 and

$$\int_X \frac{|F|^2_{\omega,h}}{e^{\psi}R(\psi)}\, dV_{X,\omega} \le \left(\frac{\alpha_1}{R(\alpha_0)} + C_R\right)\int_{Y^0} |f|^2_{\omega,h}\, dV_{X,\omega}[\psi]. \tag{1-4}$$

Remark 1.9. The case of Theorem 1.8 when X is Stein or projective was proved in [Guan and Zhou 2015c] (see also Proposition 4.1 in [Zhou and Zhu 2018] for a simplified version). Hence Theorem 1.8 can be regarded as a generalization of the main theorems in [Guan and Zhou 2015c] to weakly pseudoconvex Kähler manifolds. Then it is easy to see from Remark 1.5 and the main theorems in [Guan and Zhou 2015c] that the constant $\alpha_1/R(\alpha_0) + C_R$ in (1-4) is optimal. Hence Theorem 1.8 gives an optimal version of a main theorem in [Demailly 2016] (cf. Theorem 2.8 and Remark 2.9 in [Demailly 2016]).

Remark 1.10. In [Zhou and Zhu 2018], Theorem 1.8 was proved for L in the special case when $\psi = m \log |s|^2$, $\alpha_0 = \alpha_1 = 0$ and R is decreasing on $(-\infty, 0]$, where s is a global holomorphic section of some holomorphic vector bundle of rank m over X equipped with a smooth Hermitian metric, and s is transverse to the zero section. Similarly as in [Zhou and Zhu 2018], a global plurisubharmonic negligible weight can be added to Theorem 1.8 by adding another regularization process to Step 2 in Section 4.

Remark 1.11. In order to deal with the singular metric h_L on the weakly pseudoconvex Kähler manifold X, not only the regularization Theorem 2.2 and the error term method of solving $\bar{\partial}$ equations (Lemma 2.1) are needed, but also a limit problem about L^2 integrals with singular weights needs to be solved. We solve the limit problem in Proposition 3.2. Then by using Propositions 3.1, 3.2 and the strong openness property of multiplier ideal sheaves (Theorem 2.7) as the key tools, we construct a family of smooth extensions of f satisfying some uniform estimates, and overcome the difficulty in dealing with the singular metric (see also [Zhou and Zhu 2018] for the special case).

The rest of this paper are organized as follows. First, we recall some results used in the proof of Theorem 1.8 in Section 2. Then, we prove two key propositions in Section 3 which will be used to deal with the singular metric h_L. Finally, we prove Theorem 1.8 in Section 4 by using the results in Sections 2 and 3.

2. Some results used in the proof of Theorem 1.8

In this section, we recall some results which will be used in the proof of Theorem 1.8.

Lemma 2.1 [Demailly 2000; 2016]. *Let (X, ω) be a complete Kähler manifold equipped with a (not necessarily complete) Kähler metric ω, and let (Q, h) be a holomorphic vector bundle over X equipped with a smooth Hermitian metric h. Assume that τ and A are smooth and bounded positive functions on X and let*

$$\mathrm{B} := \big[\tau\sqrt{-1}\Theta_{Q,h} - \sqrt{-1}\partial\bar{\partial}\tau - \sqrt{-1}A^{-1}\partial\tau \wedge \bar{\partial}\tau, \Lambda\big].$$

Assume that $\delta \geq 0$ is a nonnegative number such that $\mathrm{B} + \delta\mathrm{I}$ is semipositive definite everywhere on $\wedge^{n,q}T_X^* \otimes Q$ for some $q \geq 1$. Then given a form $g \in L^2(X, \wedge^{n,q}T_X^* \otimes Q)$ such that $\mathrm{D}''g = 0$ and

$$\int_X \langle(\mathrm{B} + \delta\mathrm{I})^{-1}g, g\rangle_{\omega,h}\, dV_{X,\omega} < +\infty,$$

there exists an approximate solution $u \in L^2(X, \wedge^{n,q-1}T_X^* \otimes Q)$ and a correcting term $v \in L^2(X, \wedge^{n,q}T_X^* \otimes Q)$ such that $\mathrm{D}''u + \sqrt{\delta}v = g$ and

$$\int_X \frac{|u|^2_{\omega,h}}{\tau + A}\, dV_{X,\omega} + \int_X |v|^2_{\omega,h}\, dV_{X,\omega} \leq \int_X \langle(\mathrm{B} + \delta\mathrm{I})^{-1}g, g\rangle_{\omega,h}\, dV_{X,\omega}.$$

The following regularization theorem due to Demailly plays an important role in the present paper.

Theorem 2.2 [Demailly 1994, Theorem 6.1]. *Let (X, ω) be a complex manifold equipped with a Hermitian metric ω, and $\Omega \Subset X$ be an open subset. Assume that $T = \widetilde{T} + (\sqrt{-1}/\pi)\partial\bar{\partial}\varphi$ is a closed $(1, 1)$-current on X, where $\widetilde{T}$ is a smooth real $(1, 1)$-form and φ is a quasi-plurisubharmonic function. Let γ be a continuous real $(1, 1)$-form such that $T \geq \gamma$. Suppose that the Chern curvature tensor of T_X satisfies*

$$(\sqrt{-1}\Theta_{T_X} + \varpi \otimes \mathrm{Id}_{T_X})(\kappa_1 \otimes \kappa_2, \kappa_1 \otimes \kappa_2) \geq 0 \quad (\forall \kappa_1, \kappa_2 \in T_X \text{ with } \langle\kappa_1, \kappa_2\rangle = 0)$$

for some continuous nonnegative $(1, 1)$-form ϖ on X. Then there is a family of closed $(1, 1)$-currents $T_{\varsigma,\rho} = \widetilde{T} + (\sqrt{-1}/\pi)\partial\bar{\partial}\varphi_{\varsigma,\rho}$ defined on a neighborhood of $\overline{\Omega}$ ($\varsigma \in (0, +\infty)$ and $\rho \in (0, \rho_1)$ for some positive number ρ_1) independent of γ, such that

(i) *$\varphi_{\varsigma,\rho}$ is quasi-plurisubharmonic on a neighborhood of $\overline{\Omega}$, smooth on $\Omega \setminus E_\varsigma(T)$, increasing with respect to ς and ρ on Ω, and converges to φ on Ω as $\rho \to 0$,*

(ii) $T_{\varsigma,\rho} \geq \gamma - \varsigma\varpi - \delta_\rho\omega$ *on* Ω,

where $E_\varsigma(T) := \{x \in X : \nu(T, x) \geq \varsigma\}$ $(\varsigma > 0)$ *is the* ς*-upperlevel set of Lelong numbers*, *and* $\{\delta_\rho\}$ *is an increasing family of positive numbers such that* $\lim_{\rho\to 0} \delta_\rho = 0$.

Remark 2.3. Although Theorem 2.2 is stated in [Demailly 1994] in the case X is compact, almost the same proof as in [Demailly 1994] shows that Theorem 2.2 holds in the noncompact case while uniform estimates are obtained only on the relatively compact subset Ω.

Lemma 2.4 [Demailly 1982, Theorem 1.5]. *Let* X *be a Kähler manifold*, *and* Z *be an analytic subset of* X. *Assume that* Ω *is a relatively compact open subset of* X *possessing a complete Kähler metric. Then* $\Omega \setminus Z$ *carries a complete Kähler metric.*

Lemma 2.5 [Hörmander 1990, Theorem 4.4.2]. *Let* Ω *be a pseudoconvex open set in* $\mathbb{C}^n$, *and* φ *be a plurisubharmonic function on* Ω. *For every* $w \in L^2_{(p,q+1)}(\Omega, e^{-\varphi})$ *with* $\bar\partial w = 0$ *there is a solution* $s \in L^2_{(p,q)}(\Omega, \mathrm{loc})$ *of the equation* $\bar\partial s = w$ *such that*

$$\int_\Omega \frac{|s|^2}{(1+|z|^2)^2} e^{-\varphi}\, d\lambda \leq \int_\Omega |w|^2 e^{-\varphi}\, d\lambda,$$

where $d\lambda$ *is the* $2n$*-dimensional Lebesgue measure on* $\mathbb{C}^n$.

Lemma 2.6 [Demailly 1982, Lemma 6.9]. *Let* Ω *be an open subset of* $\mathbb{C}^n$ *and* Z *be a complex analytic subset of* Ω. *Assume that* u *is a* $(p, q-1)$*-form with* L^2_{loc} *coefficients and* g *is a* (p, q)*-form with* L^1_{loc} *coefficients such that* $\bar\partial u = g$ *on* $\Omega \setminus Z$ (*in the sense of currents*). *Then* $\bar\partial u = g$ *on* Ω.

Multiplier ideal sheaves have the following new property which was conjectured by Demailly.

Theorem 2.7 (strong openness property of multiplier ideal sheaves [Guan and Zhou 2015b]). *Let* φ *be a negative plurisubharmonic function on the unit polydisk* $\Delta^n \subset \mathbb{C}^n$. *Assume that* F *is a holomorphic function on* Δ^n *satisfying*

$$\int_{\Delta^n} |F|^2 e^{-\varphi}\, d\lambda < +\infty.$$

Then there exists $r \in (0, 1)$ *and* $\beta \in (0, +\infty)$ *such that*

$$\int_{\Delta^n_r} |F|^2 e^{-(1+\beta)\varphi}\, d\lambda < +\infty,$$

where $\Delta^n_r := \{(z_1, \ldots, z_n) \in \mathbb{C}^n : |z_k| < r, 1 \leq k \leq n\}$.

We'll also need the following desingularization theorem due to Hironaka.

Theorem 2.8 (Hironaka's desingularization theorem [Hironaka 1964; Bierstone and Milman 1991]). *Let X be a complex manifold, and M be an analytic subvariety in X. Then there is a local finite sequence of blow-ups $\mu_j : X_{j+1} \to X_j$ $(X_1 := X,\ j = 1, 2, \dots)$ with smooth centers S_j such that:*

(1) *Each component of S_j lies either in $(M_j)_{\text{sing}}$ or in $M_j \cap E_j$, where $M_1 := M$, M_{j+1} denotes the strict transform of M_j by μ_j, $(M_j)_{\text{sing}}$ denotes the singular set of M_j, and E_{j+1} denotes the exceptional divisor $\mu_j^{-1}(S_j \cup E_j)$.*

(2) *Let M' and E' denote the final strict transform of M and the exceptional divisor respectively. Then:*
   (a) *The underlying point-set $|M'|$ is smooth.*
   (b) *$|M'|$ and E' simultaneously have only normal crossings.*

Remark 2.9. We say that $|M'|$ and E' simultaneously have only normal crossings if, locally, there is a coordinate system in which E' is a union of coordinate hyperplanes, and $|M'|$ is a coordinate subspace.

3. Key propositions used to deal with the singular metric h_L

In order to deal with the singular metric h_L, we prove two key propositions in this section, which are generalizations of the key propositions in [Zhou and Zhu 2018].

Proposition 3.1. *Let R be a positive continuous function defined on $(-\infty, 0]$ such that $\beta_R := \sup_{t\le 0}(e^t R(t)) < +\infty$ and $\hat{\beta}_R := \inf_{t\le 0} R(t) > 0$. Let $\Omega \subset \mathbb{C}^n$ be a bounded pseudoconvex domain, ϕ be a plurisubharmonic function on Ω, and Υ be a quasi-plurisubharmonic function defined on a neighborhood on $\overline{\Omega}$. Assume that Υ has neat analytic singularities and the singularities of Υ are log canonical along the zero variety $Y = V(\mathcal{I}(\Upsilon))$. Set*

$$U = \{x \in \Omega : \Upsilon(x) < 0\}.$$

Furthermore, assume that

$$\sqrt{-1}\partial\bar{\partial}\Upsilon \ge -\gamma\sqrt{-1}\partial\bar{\partial}|z|^2$$

on Ω for some nonnegative number γ, where $z := (z_1, \dots, z_n)$ is the coordinate vector in $\mathbb{C}^n$. Then for every $\beta_1 \in (0, 1)$ and every holomorphic n-form f on U satisfying

$$\int_U \frac{|f|^2 e^{-\phi}}{e^{\Upsilon} R(\Upsilon)}\, d\lambda < +\infty,$$

there exists a holomorphic n-form F on Ω satisfying $F = f$ on Y,

$$\int_U \frac{|F|^2 e^{-\phi}\, d\lambda}{e^{\Upsilon} R(\Upsilon)} \le e^{2\gamma \sup_\Omega |z|^2}\left(2 + \frac{72\beta_R}{\beta_1 \hat{\beta}_R}\right)\int_U \frac{|f|^2 e^{-\phi}\, d\lambda}{e^{\Upsilon} R(\Upsilon)}, \tag{3-1}$$

and

$$\int_{\Omega}\frac{|F|^2e^{-\phi}\,d\lambda}{(1+e^{\Upsilon})^{1+\beta_1}}\le e^{2\gamma\sup_{\Omega}|z|^2}\left(\beta_R+\frac{36\beta_R}{\beta_1 2^{\beta_1}}\right)\int_U\frac{|f|^2e^{-\phi}\,d\lambda}{e^{\Upsilon}R(\Upsilon)}. \tag{3-2}$$

Proof. This proposition is a modification of a theorem in [Demailly 2012b].

Since Ω is a pseudoconvex domain, there is a sequence of pseudoconvex subdomains $\Omega_k \Subset \Omega$ $(k=1,2,\dots)$ such that $\bigcup_{k=1}^{+\infty}\Omega_k=\Omega$. Then for fixed k, by convolution we can get a decreasing family of smooth plurisubharmonic functions $\{\phi_j\}_{j=1}^{+\infty}$ defined on a neighborhood of $\overline{\Omega_k}$ such that $\lim_{j\to+\infty}\phi_j=\phi$.

Let $\theta:\mathbb{R}\to[0,1]$ be a smooth function such that $\theta=1$ on $\left(-\infty,\frac{1}{4}\right)$, $\theta=0$ on $\left(\frac{3}{4},+\infty\right)$ and $|\theta'|\le 3$ on $\mathbb{R}$.

Fix k and j. Set $\hat f=\theta(e^{\Upsilon})f$. Then the construction of $\hat f$ implies that $\hat f$ is smooth on Ω and $\hat f=f$ on $Y\cap\Omega$.

Set $g=\bar\partial\hat f$. Then $g=\theta'(e^{\Upsilon})e^{\Upsilon}\bar\partial\Upsilon\wedge f$ on Ω.

Let $\Sigma:=\{\Upsilon=-\infty\}$. Lemma 2.4 implies that $\Omega_k\setminus\Sigma$ is a complete Kähler manifold. Let $\Omega_k\setminus\Sigma$ be endowed with the Euclidean metric and let Q be the trivial line bundle on $\Omega_k\setminus\Sigma$ equipped with the metric

$$h:=e^{-\phi_j-\Upsilon-\beta_1\log(1+e^{\Upsilon})-2\gamma|z|^2}.$$

Then we want to solve a $\bar\partial$ equation on $\Omega_k\setminus\Sigma$ by applying Lemma 2.1 to the case $\tau=1$, $A=0$ and $\delta=0$ (in fact, the case $\tau=1$ and $A=0$ is the nontwisted version of Lemma 2.1). The key step in applying Lemma 2.1 is to estimate the term

$$\int_{\Omega_k\setminus\Sigma}\langle \mathrm{B}^{-1}g,g\rangle_h\,d\lambda,$$

where $\mathrm{B}:=[\sqrt{-1}\Theta_h,\Lambda]$.

Set $\nu=\partial\Upsilon$. Then $g=\theta'(e^{\Upsilon})e^{\Upsilon}\bar\nu\wedge f$ on Ω.

Since

$$\begin{aligned}
&\sqrt{-1}\Theta_h\big|_{\Omega_k\setminus\Sigma}\\
&\quad=\sqrt{-1}\partial\bar\partial\phi_j+\sqrt{-1}\partial\bar\partial\Upsilon+\beta_1\sqrt{-1}\partial\bar\partial\log(1+e^{\Upsilon})+2\gamma\sqrt{-1}\partial\bar\partial|z|^2\\
&\quad=\sqrt{-1}\partial\bar\partial\phi_j+\left(1+\frac{\beta_1e^{\Upsilon}}{1+e^{\Upsilon}}\right)\sqrt{-1}\partial\bar\partial\Upsilon+2\gamma\sqrt{-1}\partial\bar\partial|z|^2+\frac{\beta_1e^{\Upsilon}\sqrt{-1}\partial\Upsilon\wedge\bar\partial\Upsilon}{(1+e^{\Upsilon})^2}\\
&\quad\ge\frac{\beta_1e^{\Upsilon}\sqrt{-1}\nu\wedge\bar\nu}{(1+e^{\Upsilon})^2},
\end{aligned}$$

we get

$$\mathrm{B}\ge\frac{\beta_1e^{\Upsilon}}{(1+e^{\Upsilon})^2}\mathrm{T}_{\bar\nu}\mathrm{T}_{\bar\nu}^*$$

on $\Omega_k\setminus\Sigma$, where $T_{\bar\nu}$ denotes the operator $\bar\nu\wedge\bullet$ and $\mathrm{T}_{\bar\nu}^*$ is its Hilbert adjoint operator.

Then we get $\langle \mathrm{B}^{-1}g, g\rangle_h\big|_{\Omega_k\setminus U}=0$ and

$$\begin{aligned}\langle \mathrm{B}^{-1}g, g\rangle_h\big|_{(U\cap\Omega_k)\setminus\Sigma} &= \langle \mathrm{B}^{-1}(\theta'(e^{\Upsilon})e^{\Upsilon}\bar{\nu}\wedge f), \theta'(e^{\Upsilon})e^{\Upsilon}\bar{\nu}\wedge f\rangle_h \\ &\le \frac{(1+e^{\Upsilon})^2}{\beta_1 e^{\Upsilon}}|\theta'(e^{\Upsilon})e^{\Upsilon}f|^2 e^{-\phi_j-\Upsilon-\beta_1\log(1+e^{\Upsilon})-2\gamma|z|^2} \\ &= \frac{(1+e^{\Upsilon})^{2-\beta_1}}{\beta_1}|\theta'(e^{\Upsilon})f|^2 e^{-\phi_j-2\gamma|z|^2} \le \frac{36}{\beta_1 2^{\beta_1}}|f|^2 e^{-\phi_j-2\gamma|z|^2}.\end{aligned}$$

Hence it follows from Lemma 2.1 that there exists $u_{k,j}\in L^2(\Omega_k\setminus\Sigma,\ K_\Omega\otimes Q,\ h)$ such that $\bar\partial u_{k,j}=g=\bar\partial\hat{f}$ on $\Omega_k\setminus\Sigma$ and

$$\int_{\Omega_k\setminus\Sigma}|u_{k,j}|_h^2\,d\lambda\le\int_{\Omega_k\setminus\Sigma}\langle \mathrm{B}^{-1}g, g\rangle_h\,d\lambda.$$

Thus

$$\begin{aligned}\text{(3-3)}\qquad \int_{\Omega_k\setminus\Sigma}\frac{|u_{k,j}|^2e^{-\phi_j-2\gamma|z|^2}}{e^{\Upsilon}(1+e^{\Upsilon})^{\beta_1}}\,d\lambda &\le \frac{36}{\beta_1 2^{\beta_1}}\int_{U\cap\Omega_k}|f|^2e^{-\phi_j-2\gamma|z|^2}\,d\lambda \\ &\le \frac{36\beta_R}{\beta_1 2^{\beta_1}}\int_U\frac{|f|^2e^{-\phi-2\gamma|z|^2}}{e^{\Upsilon}R(\Upsilon)}\,d\lambda.\end{aligned}$$

Hence we have $u_{k,j}\in L^2(\Omega_k\setminus\Sigma,\ K_\Omega)$. Since $g\in C^\infty(\Omega_k,\ \wedge^{n,1}T^*_\Omega)$, Lemma 2.6 implies that $\bar\partial u_{k,j}=g$ holds on Ω_k.

Let $F_{k,j}:=\hat{f}-u_{k,j}$. Then $\bar\partial F_{k,j}=0$ on Ω_k. Thus $F_{k,j}$ is holomorphic on Ω_k. Hence $u_{k,j}$ is smooth on Ω_k. Then the nonintegrability of $e^{-\Upsilon}$ along Y implies that $u_{k,j}=0$ on $Y\cap\Omega_k$. Therefore, $F_{k,j}=f$ on $Y\cap\Omega_k$.

It follows from (3-3) that

$$\begin{aligned}\int_{U\cap\Omega_k}\frac{|u_{k,j}|^2e^{-\phi_j-2\gamma|z|^2}}{e^{\Upsilon}R(\Upsilon)}\,d\lambda &\le \frac{2^{\beta_1}}{\hat\beta_R}\int_{U\cap\Omega_k}\frac{|u_{k,j}|^2e^{-\phi_j-2\gamma|z|^2}}{e^{\Upsilon}(1+e^{\Upsilon})^{\beta_1}}\,d\lambda \\ &\le \frac{36\beta_R}{\beta_1\hat\beta_R}\int_U\frac{|f|^2e^{-\phi-2\gamma|z|^2}}{e^{\Upsilon}R(\Upsilon)}\,d\lambda.\end{aligned}$$

Since

$$|F_{k,j}|^2\big|_{U\cap\Omega_k}\le 2|\hat f|^2+2|u_{k,j}|^2\le 2|f|^2+2|u_{k,j}|^2,$$

we get

$$\begin{aligned}\text{(3-4)}\qquad \int_{U\cap\Omega_k}\frac{|F_{k,j}|^2e^{-\phi_j-2\gamma|z|^2}}{e^{\Upsilon}R(\Upsilon)}\,d\lambda &\le 2\int_{U\cap\Omega_k}\frac{(|f|^2+|u_{k,j}|^2)e^{-\phi_j-2\gamma|z|^2}}{e^{\Upsilon}R(\Upsilon)}\,d\lambda \\ &\le \left(2+\frac{72\beta_R}{\beta_1\hat\beta_R}\right)\int_U\frac{|f|^2e^{-\phi-2\gamma|z|^2}}{e^{\Upsilon}R(\Upsilon)}\,d\lambda.\end{aligned}$$

Since

$$\langle\kappa_1+\kappa_2,\ \kappa_1+\kappa_2\rangle\le\langle\kappa_1,\kappa_1\rangle+\langle\kappa_2,\kappa_2\rangle+c\langle\kappa_1,\kappa_1\rangle+\frac{1}{c}\langle\kappa_2,\kappa_2\rangle \tag{3-5}$$

for any inner product space (H, $\langle\bullet,\bullet\rangle$), where $\kappa_1,\kappa_2\in$ H, we get

$$|F_{k,j}|^2\big|_{U\cap\Omega_k}\le(|f|+|u_{k,j}|)^2\le(1+e^{\Upsilon})|f|^2+\Big(1+\frac{1}{e^{\Upsilon}}\Big)|u_{k,j}|^2.$$

Then

$$\frac{|F_{k,j}|^2}{(1+e^{\Upsilon})^{1+\beta_1}}\bigg|_{U\cap\Omega_k}\le|f|^2+\frac{|u_{k,j}|^2}{e^{\Upsilon}(1+e^{\Upsilon})^{\beta_1}}.$$

Since $|F_{k,j}|^2\big|_{\Omega_k\setminus U}=|u_{k,j}|^2$, we get

$$\frac{|F_{k,j}|^2}{(1+e^{\Upsilon})^{1+\beta_1}}\bigg|_{\Omega_k\setminus U}\le\frac{|u_{k,j}|^2}{e^{\Upsilon}(1+e^{\Upsilon})^{\beta_1}}.$$

Hence it follows from the two inequalities above and (3-3) that

$$\begin{aligned}\int_{\Omega_k}\frac{|F_{k,j}|^2e^{-\phi_j-2\gamma|z|^2}}{(1+e^{\Upsilon})^{1+\beta_1}}\,d\lambda&\le\int_U|f|^2e^{-\phi-2\gamma|z|^2}\,d\lambda+\int_{\Omega_k}\frac{|u_{k,j}|^2e^{-\phi_j-2\gamma|z|^2}}{e^{\Upsilon}(1+e^{\Upsilon})^{\beta_1}}\,d\lambda\\&\le\Big(\beta_R+\frac{36\beta_R}{\beta_1 2^{\beta_1}}\Big)\int_U\frac{|f|^2e^{-\phi-2\gamma|z|^2}}{e^{\Upsilon}R(\Upsilon)}\,d\lambda.\end{aligned}\tag{3-6}$$

Since $e^{-2\gamma\sup_\Omega|z|^2}\le e^{-2\gamma|z|^2}\le 1$ on Ω, the desired holomorphic n-form F on Ω and the L^2 estimates (3-1) and (3-2) can be obtained from (3-4) and (3-6) by applying Montel's theorem and extracting weak limits of $\{F_{k,j}\}_{k,j}$, first as $j\to+\infty$ and then as $k\to+\infty$. □

Proposition 3.2. *Let X, ψ, Y and Y^0 be as in Theorem 1.8. Let $U\Subset V\Subset\Omega$ be three local coordinate balls in X, ϕ be a plurisubharmonic function on Ω such that* $\sup_\Omega\phi<+\infty$, *and υ be a nonnegative continuous function on Ω with* $\operatorname{supp}\upsilon\subset U$. *Let C, β, c_1 and c_2 be positive numbers, and let β_1 be a small enough positive number. Assume that f is a holomorphic function on $\Omega\cap Y$ satisfying*

$$\int_{\Omega\cap Y^0}|f|^2e^{-\phi}\,d\lambda[\psi]<+\infty, \tag{3-7}$$

and that $f_t\in\mathcal{O}(\Omega)$ $\big(t\in(-\infty,0)\big)$ are a family of holomorphic functions such that for all $t\in(-\infty,0)$, $f_t=f$ on $\Omega\cap Y$,

$$\sup_V|f_t|^2\le Ce^{-\beta_1 t}, \tag{3-8}$$

$$\frac{1}{e^t}\int_{\Omega\cap\{\psi<t+c_2\}}|f_t|^2e^{-(1+\beta)\phi}\,d\lambda\le C. \tag{3-9}$$

Then

$$\varlimsup_{t\to-\infty}\int_{U\cap\{t-c_1<\psi<t+c_2\}}\frac{e^t\upsilon|f_t|^2e^{-\phi}}{(e^{\psi}+e^t)^2}\,d\lambda\le\int_{U\cap Y^0}\upsilon|f|^2e^{-\phi}\,d\lambda[\psi]. \tag{3-10}$$

Remark 3.3. One of the key points in the proof of Proposition 3.2 is to verify that the upper limit in (3-10) produces the zero measure on the singular set of Y, i.e., we have (3-16). Then the key uniform estimates in Step 2 of the proof are obtained.

In order to prove Proposition 3.2, we prove the following lemma at first.

Lemma 3.4. *Let r_1, r_2 and γ be positive numbers such that $r_1 < r_2 < \gamma$. Let φ be a bounded negative subharmonic function on Δ_γ, where $\Delta_\gamma := \{w \in \mathbb{C} : |w| < \gamma\}$. Assume that $\{v_t\}_{t\in(-\infty,0)}$ are nonnegative continuous functions defined on Δ_γ such that*

$$\lim_{t\to-\infty} \sup_{\{w\in\mathbb{C}\,:\, e^t(r_1)^{2\alpha}<|w|^{2\alpha}<e^t(r_2)^{2\alpha}\}} |v_t(w) - v_0| = 0, \tag{3-11}$$

where $\alpha \in [1, +\infty)$ and $v_0 \in [0, +\infty)$. Let

$$P_t := \int_{\{w\in\mathbb{C}\,:\, e^t(r_1)^{2\alpha}<|w|^{2\alpha}<e^t(r_2)^{2\alpha}\}} \frac{e^t|w|^{2\alpha-2}v_t(w)e^{-\varphi(w)}}{(|w|^{2\alpha}+e^t)^2}\, d\lambda(w).$$

Then

$$\overline{\lim_{t\to-\infty}}\, P_t \le \frac{\pi v_0 e^{-\varphi(0)}}{\alpha}. \tag{3-12}$$

Proof. Put

$$S_{\delta,t} = \{z \in \Delta_\gamma : \varphi(e^{t/(2\alpha)}z) < (1+\delta)\varphi(0)\}, \quad \delta \in (0, +\infty), \quad t \in (-\infty, 0).$$

Denote by $\lambda(S_{\delta,t})$ the 2-dimensional Lebesgue measure of $S_{\delta,t}$.

Since $\varphi(w)$ is a negative upper semicontinuous function on Δ_γ and $\varphi(0) > -\infty$, we have that for every $\varepsilon \in (0, 1)$, there exists $t_\varepsilon \in (-\infty, 0)$ such that

$$\varphi(e^{t/(2\alpha)}z) \le (1-\varepsilon)\varphi(0)$$

for all $z \in \Delta_\gamma$ when $t \in (-\infty, t_\varepsilon)$.

Since $\varphi(e^{t/(2\alpha)}z)$ is subharmonic on Δ_γ with respect to z for any $t \in (-\infty, t_\varepsilon)$, it follows from the mean value inequality that, for all $t \in (-\infty, t_\varepsilon)$,

$$\begin{aligned}
\varphi(0) &\le \frac{1}{\pi\gamma^2}\int_{z\in\Delta_\gamma} \varphi(e^{t/(2\alpha)}z)\, d\lambda(z) \\
&= \frac{1}{\pi\gamma^2}\int_{z\in\Delta_\gamma\setminus S_{\delta,t}} \varphi(e^{t/(2\alpha)}z)\, d\lambda(z) + \frac{1}{\pi\gamma^2}\int_{z\in S_{\delta,t}} \varphi(e^{t/(2\alpha)}z)\, d\lambda(z) \\
&\le \frac{(1-\varepsilon)\varphi(0)\big(\pi\gamma^2 - \lambda(S_{\delta,t})\big)}{\pi\gamma^2} + \frac{(1+\delta)\varphi(0)\lambda(S_{\delta,t})}{\pi\gamma^2} \\
&= \varphi(0)\bigg(1 - \varepsilon + \frac{(\delta+\varepsilon)\lambda(S_{\delta,t})}{\pi\gamma^2}\bigg).
\end{aligned}$$

Then $\varphi(0) < 0$ implies that

$$\lambda(S_{\delta,t}) \le \frac{\pi\gamma^2\varepsilon}{\delta+\varepsilon} \le \frac{\pi\gamma^2}{\delta}\varepsilon$$

when $t \in (-\infty, t_\varepsilon)$. Hence

$$\lim_{t\to-\infty} \lambda(S_{\delta,t}) = 0 \qquad \text{for all } \delta \in (0,+\infty). \tag{3-13}$$

Since φ is bounded, we have

$$-\varphi \le C_1$$

for some positive number C_1.

Equation (3-11) implies that

$$\sup_{\{w\in\mathbb{C}\,:\, e^t(r_1)^{2\alpha}<|w|^{2\alpha}<e^t(r_2)^{2\alpha}\}} v_t(w) \le C_2$$

for some positive number C_2 independent of t when t is small enough.

Then by the change of variables $w = e^{t/(2\alpha)}z$, we have

$$\begin{aligned}
P_t &= \int_{\{z\in\mathbb{C}\,:\, r_1<|z|<r_2\}} \frac{|z|^{2\alpha-2} v_t(e^{t/(2\alpha)}z)e^{-\varphi(e^{t/(2\alpha)}z)}}{(|z|^{2\alpha}+1)^2}\,d\lambda(z)\\
&= \int_{\{r_1<|z|<r_2\}\cap S_{\delta,t}} \frac{|z|^{2\alpha-2} v_t(e^{t/(2\alpha)}z)e^{-\varphi(e^{t/(2\alpha)}z)}}{(|z|^{2\alpha}+1)^2}\,d\lambda(z)\\
&\qquad\qquad + \int_{\{r_1<|z|<r_2\}\setminus S_{\delta,t}} \frac{|z|^{2\alpha-2} v_t(e^{t/(2\alpha)}z)e^{-\varphi(e^{t/(2\alpha)}z)}}{(|z|^{2\alpha}+1)^2}\,d\lambda(z)\\
&\le \frac{(r_2)^{2\alpha-2}C_2e^{C_1}}{\big((r_1)^{2\alpha}+1\big)^2}\cdot\lambda(S_{\delta,t})\\
&\qquad + \Big(\sup_{r_1<|z|<r_2} v_t(e^{t/(2\alpha)}z)\Big)e^{-(1+\delta)\varphi(0)} \int_{\{r_1<|z|<r_2\}} \frac{|z|^{2\alpha-2}}{(|z|^{2\alpha}+1)^2}\,d\lambda(z).
\end{aligned}$$

Since

$$\int_{\{r_1<|z|<r_2\}} \frac{|z|^{2\alpha-2}}{(|z|^{2\alpha}+1)^2}\,d\lambda(z) \le \frac{\pi}{\alpha},$$

we obtain from (3-11), (3-13) that

$$\overline{\lim_{t\to-\infty}}\, P_t \le \frac{\pi v_0 e^{-(1+\delta)\varphi(0)}}{\alpha}.$$

Since δ is an arbitrary positive number, we get (3-12). □

Now we begin to prove Proposition 3.2.

Proof. Let $\beta_v := \sup_U v$.

Without loss of generality, we may suppose that ϕ is negative on Ω.

We will use Hironaka's desingularization theorem (Theorem 2.8) to deal with the measure $d\lambda[\psi]$. This idea comes from [Demailly 2016].

At first we use Theorem 2.8 on X to resolve the singularities of Y and we denote the corresponding proper modification by μ_1. Next, we make a blow-up μ_2 along $|Y'|$. Then we use Theorem 2.8 again to resolve the singularities of Σ and we denote the corresponding proper holomorphic modification by μ_3, where Σ denote the strict transform of $\{\psi = -\infty\}$ by $\mu_1 \circ \mu_2$. Finally, we make a blow-up μ_4 along $|\Sigma'|$. Thus we can get a proper holomorphic map $\mu : \widetilde{X} \to X$, which is locally a finite composition of blow-ups with smooth centers and is equal to $\mu_1 \circ \mu_2 \circ \mu_3 \circ \mu_4$. Moreover, $\widetilde{Y}$ and the divisor $\mu^{-1}(\{\psi = -\infty\}) \setminus \widetilde{Y}$ simultaneously have only normal crossings in $\widetilde{X}$, where $\widetilde{Y}$ denotes the strict transform of $\mu_2^{-1}(|Y'|)$ by $\mu_3 \circ \mu_4$.

Step 1: Representing the measure $|f|^2 d\lambda[\psi]$ on $Y^0 \cap U$ explicitly as an integral on $\widetilde{Y}$ (see (3-15)).

For any $\tilde{x} \in \overline{\mu^{-1}(U)} \cap \mu^{-1}(\{\psi = -\infty\})$, there exists a relatively compact coordinate ball $(W; w_1, \dots, w_n)$ contained in $\mu^{-1}(V)$ centered at $\tilde{x}$ such that $w^b = 0$ is the zero divisor of the Jacobian J_μ, and $\psi \circ \mu$ can be written on W as

$$\psi \circ \mu(w) = c \log |w^a|^2 + \tilde{u}(w),$$

where c is a positive number, $w := (w_1, \dots, w_n)$, $\tilde{u} \in C^\infty(\overline{W})$, $w^a := \prod_{p=1}^n w_p^{a_p}$ and $w^b := \prod_{p=1}^n w_p^{b_p}$ for some nonnegative integers a_p and b_p.

Let $D_p := \{w_p = 0\}$. Then as proved in [Demailly 2016], the multiplier ideal sheaf $\mathcal{I}(\psi)$ is given by the direct image formula

$$\mathcal{I}(\psi) = \mu_* \mathcal{O}_{\widetilde{X}}\Big(-\sum_{p=1}^n \lfloor c a_p - b_p \rfloor_+ D_p \Big),$$

where $\lfloor c a_p - b_p \rfloor_+$ denotes the minimal nonnegative integer bigger than $c a_p - b_p - 1$. Since ψ has log canonical singularities, by the construction of μ and Theorem 2.8, one of the following cases is true on W:

(A) $\widetilde{Y}$ is given on W precisely by D_{p_0} (if W is small enough) for some p_0 satisfying $c a_{p_0} - b_{p_0} = 1$, and $c a_p - b_p \le 1$ for $p \ne p_0$;

(B) $\widetilde{Y} \cap W = \varnothing$, and $c a_p - b_p \le 1$.

By definition, the measure $|f|^2 d\lambda[\psi]$ can be defined as

$$g \mapsto \varlimsup_{t \to -\infty} \int_{\{t < c \log |w^a|^2 + \tilde{u}(w) < t+1\}} \frac{|\tilde{f} \circ \mu|^2 (\tilde{g} \circ \mu) \xi e^{-\tilde{u}}}{|w^{ca-b}|^2} \, d\lambda(w), \tag{3-14}$$

where $d\lambda(w) :=$ the Lebesgue measure with respect to the coordinate vector w, $\tilde{f}$ is a holomorphic extension of f to Ω, g and $\tilde{g}$ are defined as in Definition 1.3, and ξ is the smooth positive function $|J_\mu|^2/|w^b|^2$ (as stated in [Demailly 2016], one would still have to take into account a partition of unity on the various coordinate charts covering the fibers of μ, but we will avoid this technicality for the simplicity of notation).

In Case (A), let us denote $w=(w', w_{p_0})\in\mathbb{C}^{n-1}\times\mathbb{C}$, $a=(a', a_{p_0})$, $b=(b', b_{p_0})$ and $d\lambda(w)=d\lambda(w')d\lambda(w_{p_0})$. Then (3-14) becomes

$$g\mapsto \varlimsup_{t\to-\infty}\int_{\{t<c\log|w^a|^2+\tilde{u}(w)<t+1\}}\frac{|\tilde{f}\circ\mu|^2}{|(w')^{ca'-b'}|^2}\cdot\frac{(\tilde{g}\circ\mu)\xi e^{-\tilde{u}}}{|w_{p_0}|^2}\,d\lambda(w).$$

Since the domain of integration can be written as

$$\big\{e^{t-\tilde{u}(w)}|(w')^{a'}|^{-2c}<|w_{p_0}|^{2ca_{p_0}}<e^{t+1-\tilde{u}(w)}|(w')^{a'}|^{-2c}\big\},$$

the mapping (3-14) becomes

$$g\mapsto\frac{\pi}{ca_{p_0}}\int_{w'\in D_{p_0}}\frac{|f\circ\mu|^2}{|(w')^{ca'-b'}|^2}\cdot(g\circ\mu)\xi e^{-\tilde{u}}\,d\lambda(w'). \tag{3-15}$$

Set $\kappa=\{p: ca_p-b_p=1\}$.

If $p\in\kappa\setminus\{p_0\}$, then Theorem 2.8 and the construction of μ imply that an image of D_p under a finite sequence of blow-ups in the desingularization process must be contained in a smooth center contained in Y or $\mu_2^{-1}(|Y'|)$. Hence the images of D_p and $D_p\cap D_{p_0}$ coincide under the composition of these blow-ups.

Since it is implied from (3-7) and (3-15) that $f\circ\mu\big|_{D_p\cap D_{p_0}}=0$, we obtain that

$$f\circ\mu\big|_{D_p}=0 \tag{3-16}$$

holds for all $p\in\kappa\setminus\{p_0\}$ in Case (A).

Similarly, we can get that (3-16) holds for all $p\in\kappa$ in Case (B). Then (3-14) is the zero measure in Case (B).

Therefore, we represent the measure $|f|^2d\lambda[\psi]$ on $Y^0\cap U$ explicitly as in (3-15).

Step 2: Obtaining some uniform estimates for $f_t\circ\mu$.

By Cauchy's inequality for holomorphic functions, it follows from (3-8) that

$$\sup_{U_1}|\partial^\gamma f_t|^2\le C_1\sup_V|f_t|^2\le C_1Ce^{-\beta_1 t} \tag{3-17}$$

for any $t\in(-\infty,0)$ and any multi-index γ satisfying $|\gamma|\le n$, where $U_1\Subset V$ is a neighborhood of $\overline{U}$, and C_1 is a positive number independent of t and γ.

Let $W_t:=W\cap\mu^{-1}(U)\cap\{\psi\circ\mu<t+c_2\}$.

In Case (A), by applying the mean value theorem to $f_t \circ \mu$ successively along the directions in κ, we get from (3-17) and (3-16) that for any $w = (w', w_{p_0}) \in W_t$,

$$|f_t \circ \mu(w', w_{p_0}) - f_t \circ \mu(w', 0)|^2 \le C_2 \prod_{p\in\kappa} |w_p|^2 \sup_{|\gamma|\le|\kappa|} \sup_{\mu^{-1}(U_1)} |\partial^\gamma f_t|^2 \le C_3 e^{-\beta_1 t} \prod_{p\in\kappa} |w_p|^2 \tag{3-18}$$

and

$$|f_t \circ \mu(w', 0)|^2 = |f \circ \mu(w', 0)|^2 \le C_4 \prod_{p\in\kappa\setminus\{p_0\}} |w_p|^2 \tag{3-19}$$

when t is small enough, where C_2, C_3 and C_4 are positive numbers independent of t.

In Case (B), if $\kappa \neq \varnothing$, take $p_1 \in \kappa$ and denote $w = (w'', w_{p_1})$. Since $f_t \circ \mu(w'', 0) = f \circ \mu(w'', 0) = 0$, by a similar method we have that

$$|f_t \circ \mu(w'', w_{p_1})|^2 \le C_5 e^{-\beta_1 t} \prod_{p\in\kappa} |w_p|^2 \tag{3-20}$$

for any $w = (w'', w_{p_1}) \in W_t$ when t is small enough, where C_5 is a positive number independent of t. If $\kappa = \varnothing$, (3-8) implies that

$$|f_t \circ \mu(w)|^2 \le C e^{-\beta_1 t} \tag{3-21}$$

for any $w \in W_t$.

Step 3: The proof of (3-10).

Let j be a positive integer. Then (3-9) implies that

$$\frac{1}{e^t} \int_{\{\phi\le -j\}\cap U\cap\{\psi<t+c_2\}} |f_t|^2 e^{-\phi}\, d\lambda \le \frac{1}{e^t} \int_{\{\phi\le -j\}\cap U\cap\{\psi<t+c_2\}} |f_t|^2 e^{-(1+\beta)\phi-\beta j}\, d\lambda \le C e^{-\beta j}$$

for any $t \in (-\infty, 0)$.

Therefore, for every $\varepsilon \in (0, 1)$, there exists a positive integer j_ε such that

$$\int_{\{\phi\le -j_\varepsilon\}\cap U\cap\{t-c_1<\psi<t+c_2\}} \frac{e^t v |f_t|^2 e^{-\phi}}{(e^\psi + e^t)^2}\, d\lambda \le \frac{1}{(e^{-c_1}+1)^2 e^t} \int_{\{\phi\le -j_\varepsilon\}\cap U\cap\{\psi<t+c_2\}} v|f_t|^2 e^{-\phi}\, d\lambda \le \frac{\beta_v C e^{-\beta j_\varepsilon}}{(e^{-c_1}+1)^2} < \frac{\varepsilon}{2} \tag{3-22}$$

for any $t \in (-\infty, 0)$.

Set $\phi_\varepsilon = \max\{\phi, -j_\varepsilon\}$. We want to prove

$$\overline{\lim_{t\to-\infty}} \int_{U\cap\{t-c_1<\psi<t+c_2\}} \frac{e^t v |f_t|^2 e^{-\phi_\varepsilon}}{(e^\psi + e^t)^2}\, d\lambda \le \int_{U\cap Y^0} v|f|^2 e^{-\phi_\varepsilon}\, d\lambda[\psi]. \tag{3-23}$$

Set

$$I_0=\overline{\lim_{t\to-\infty}}\int_{W\cap\mu^{-1}(U)\cap\{t-c_1<\psi\circ\mu<t+c_2\}}\frac{e^t(v\circ\mu)|f_t\circ\mu|^2e^{-\phi_\varepsilon\circ\mu}|J_\mu|^2}{(e^{\psi\circ\mu}+e^t)^2}\,d\lambda.$$

Then by Step 1, it suffices to prove that

$$I_0\le\frac{\pi}{ca_{p_0}}\int_{W\cap\mu^{-1}(U)\cap D_{p_0}}\frac{(v\circ\mu)|f\circ\mu|^2\xi e^{-\tilde{u}-\phi_\varepsilon\circ\mu}}{|(w')^{ca'-b'}|^2}\,d\lambda(w') \tag{3-24}$$

in Case (A) and $I_0=0$ in Case (B), where ξ is the smooth positive function $|J_\mu|^2/|w^b|^2$ defined in Step 1.

In Case (A), let

$$\Phi_t(w'):=\int_{W_{t,w'}}\frac{e^t(v\circ\mu)|f_t\circ\mu|^2e^{-\phi_\varepsilon\circ\mu}|J_\mu|^2}{(e^{\psi\circ\mu}+e^t)^2}\,d\lambda(w_{p_0})$$

and

$$\Phi(w'):=\frac{\pi}{ca_{p_0}}\cdot\frac{v\circ\mu(w',0)|f\circ\mu(w',0)|^2\xi(w',0)e^{-\tilde{u}(w',0)-\phi_\varepsilon\circ\mu(w',0)}}{|(w')^{ca'-b'}|^2},$$

where $W_{t,w'}$ is the 1-dimensional open set

$$\left\{e^{t-c_1-\tilde{u}(w',w_{p_0})}|(w')^{a'}|^{-2c}<|w_{p_0}|^{2ca_{p_0}}<e^{t+c_2-\tilde{u}(w',w_{p_0})}|(w')^{a'}|^{-2c}\right\}\cap W\cap\mu^{-1}(U)$$

for every fixed t and w' $\left(w'\in D_{p_0}\setminus\bigcup_{p\neq p_0}D_p\right)$. Then

$$I_0=\overline{\lim_{t\to-\infty}}\int_{W\cap\mu^{-1}(U)\cap D_{p_0}}\Phi_t(w')\,d\lambda(w'). \tag{3-25}$$

Since $-c_1<\psi\circ\mu-t<c_2$ holds on $W_{t,w'}$, we obtain from (3-18) and (3-19) that

$$\begin{aligned}\Phi_t(w')&\le C_6\int_{W_{t,w'}}\frac{(v\circ\mu)|f_t\circ\mu|^2e^{-\phi_\varepsilon\circ\mu}|J_\mu|^2}{e^{\psi\circ\mu}}\,d\lambda(w_{p_0})\\&\le C_7\int_{W_{t,w'}}\frac{|f_t\circ\mu|^2}{|w^{ca-b}|^2}\,d\lambda(w_{p_0})\\&\le C_8\int_{W_{t,w'}}\frac{\prod_{p\in\kappa}|w_p|^2}{|w^{(1+\beta_1)ca-b}|^2}\,d\lambda(w_{p_0})+C_8\int_{W_{t,w'}}\frac{\prod_{p\in\kappa\setminus\{p_0\}}|w_p|^2}{|w^{ca-b}|^2}\,d\lambda(w_{p_0}),\end{aligned}$$

where C_7 and C_8 are positive numbers independent of t.

Since it is easy to prove that the right-hand side of the above inequality is dominated by a function of w' which is independent of t and which belongs to $L^1(W \cap \mu^{-1}(U) \cap D_{p_0})$ when

$$\beta_1 < \min_{\{p\,:\,a_p \neq 0\}} \frac{1-(ca_p - b_p) + \lfloor ca_p - b_p \rfloor_+}{ca_p},$$

it follows from (3-25) and Fatou's lemma that

$$I_0 \leq \int_{W\cap\mu^{-1}(U)\cap D_{p_0}} \overline{\lim_{t\to-\infty}} \, \Phi_t(w')\, d\lambda(w'). \tag{3-26}$$

Since (3-18) implies that

$$\lim_{t\to-\infty} \sup_{w_{p_0}\in W_{t,w'}} |f_t \circ \mu(w', w_{p_0}) - f \circ \mu(w', 0)| = 0$$

for every fixed $w' \in \big(W \cap \mu^{-1}(U) \cap D_{p_0}\big) \setminus \bigcup_{p \neq p_0} (D_{p_0} \cap D_p)$ when $\beta_1 < 1/ca_{p_0}$, it follows from Lemma 3.4 that

$$\overline{\lim_{t\to-\infty}} \, \Phi_t(w') \leq \Phi(w') \quad \text{for all } w' \in \big(W \cap \mu^{-1}(U) \cap D_{p_0}\big) \setminus \bigcup_{p\neq p_0} (D_{p_0} \cap D_p).$$

Hence (3-24) follows from (3-26). Similarly, we can obtain from (3-20) and (3-21) that $I_0 = 0$ in Case (B) when

$$\beta_1 < \min_{\{p\,:\,a_p \neq 0\}} \frac{1-(ca_p - b_p) + \lfloor ca_p - b_p \rfloor_+}{ca_p}.$$

Thus we get (3-23).

It is easy to see that (3-10) follows from (3-22) and (3-23). Thus we finish the proof of Proposition 3.2. □

4. Proof of Theorem 1.8

Without loss of generality, we can suppose that f is not 0 identically.

Let h_0 be any fixed smooth metric of L on X. Then $h = h_0 e^{-\phi}$ for some global function ϕ on X, which is quasi-plurisubharmonic by the assumption in the theorem.

Since X is weakly pseudoconvex, there exists a smooth plurisubharmonic exhaustion function P on X. Let $X_k := \{P < k\}$ ($k = 1, 2, \ldots$, we choose P such that $X_1 \neq \varnothing$).

Our proof consists of several steps. We will discuss for fixed k until the end of Step 5.

We will give the proof for the line bundle L in the first five steps, and we will give the proof for the vector bundle E in Step 6.

Step 1: Construction of a family of special smooth extensions $\tilde{f}_t$ of f to a neighborhood of $\overline{X_k} \cap Y$ in X.

In order to deal with singular metrics of holomorphic line bundles on weakly pseudoconvex Kähler manifolds, we construct in this step a family of smooth extensions $\tilde{f}_t$ of f satisfying some special estimates by using the results in Section 3.

Let $\varepsilon \in \left(0, \frac{1}{2}\right)$.

For the sake of clarity, we divide this step into four parts.

Part I: Construction of local coordinate patches $\{\Omega_i\}_{i=1}^N$, $\{U_i\}_{i=1}^N$ and a partition of unity $\{\xi_i\}_{i=1}^{N+1}$.

For any point $x \in Y$, we can find a local coordinate ball Ω'_x in X centered at x such that there exists a local holomorphic frame of L on Ω'_x and such that ϕ can be written as a sum of a smooth function and a plurisubharmonic function on Ω'_x. Moreover, we assume that ψ can be written on Ω'_x as

$$\psi = c_x \log \sum_{1 \le j \le j_0} |g_{x,j}|^2 + u_x, \tag{4-1}$$

where c_x is a positive number, $g_{x,j} \in \mathcal{O}_X(\Omega'_x)$ and $u_x \in C^\infty(\Omega'_x)$.

Let $U_x \Subset V_x \Subset \Omega_x \Subset \Omega'_x$ be three small coordinate balls.

Since $\overline{X_k} \cap Y$ is compact, there exist points $x_1, x_2, \dots, x_N \in \overline{X_k} \cap Y$ such that $\overline{X_k} \cap Y \subset \bigcup_{i=1}^N U_{x_i}$.

For simplicity, we denote Ω'_{x_i}, Ω_{x_i}, U_{x_i}, V_{x_i} and u_{x_i} by Ω'_i, Ω_i, U_i, V_i and u_i, respectively. We denote the local expression (4-1) on Ω'_i by

$$\psi = \Upsilon_i + u_i.$$

Choose an open set U_{N+1} in X such that $\overline{X_k} \cap Y \subset X \setminus \overline{U_{N+1}} \Subset \bigcup_{i=1}^N U_i$. Set $U = X \setminus \overline{U_{N+1}}$.

Let $\{\xi_i\}_{i=1}^{N+1}$ be a partition of unity subordinate to the cover $\{U_i\}_{i=1}^{N+1}$ of X. Then $\operatorname{supp} \xi_i \Subset U_i$ for $i = 1, \dots, N$ and $\sum_{i=1}^N \xi_i = 1$ on U.

Part II: Construction of local holomorphic extensions $\hat{f}_{i,t}$ $(1 \le i \le N)$ of f to $\Omega_i \cap \{\psi < t + c_2\}$, where c_2 will be defined in this part.

By Remark 1.9, f has local L^2 extensions to local coordinate balls around every point in Y. Hence f is indeed a holomorphic section well defined on Y (not only on Y^0). By Step 1 (see (3-15)) in the proof of Proposition 3.2, (1-3) is equivalent to

$$\int_{D_{p_0}} \frac{|f \circ \mu|^2_{\omega,h_0} \xi e^{-\tilde{u} - \phi \circ \mu}}{|(w')^{ca' - b'}|^2} \, d\lambda(w') < +\infty.$$

Hence by Theorem 2.7, there exists a positive number $\beta \in (0, 1)$ such that

$$\int_{\Omega_i \cap Y^0} |f|^2_{\omega,h_0} e^{-(1+\beta)\phi}\, dV_{X,\omega}[\psi] < +\infty \quad (1 \le i \le N). \tag{4-2}$$

Let $\tilde{\alpha}_0 < \alpha_0$ be a fixed number such that R is decreasing on $(-\infty, \tilde{\alpha}_0]$. Then set $R_0(t) = R(\tilde{\alpha}_0)e^{-\beta_2(t-\tilde{\alpha}_0)}$, $t \in (-\infty, \tilde{\alpha}_0]$, where β_2 is a positive number which will be determined later in Step 4. Let

$$R_1(t) := \min\{R_0(t+\tilde{\alpha}_0), R(t+\tilde{\alpha}_0)\}, \quad t \in (-\infty, 0].$$

Then R_1 is decreasing and thereby satisfies all the requirements for the functions in $\mathfrak{R}_{0,\alpha_1}$ except that R_1 is only continuous.

Let $c_1 = c_2 := \log((2-\varepsilon)/\varepsilon)$, $m_i := \inf_{\Omega_i} u_i$ and $M_i := \sup_{\Omega_i} u_i$.

For each fixed $t \in (-\infty, 0)$, by Remark 1.9, we apply Theorem 1.8 to the Stein manifold $\Omega_i \cap \{\Upsilon_i < t + c_2 - m_i\}$, to the negative plurisubharmonic function $\Upsilon_i - t - c_2 + m_i$, to the holomorphic section f on $\Omega_i \cap Y^0$ with the L^2 condition (4-2) and to the function R_1 (R_1 is only needed to be continuous by the remark after Theorem 2.1 in [Guan and Zhou 2015c]), and then we obtain L^2 extensions of f from $\Omega_i \cap Y^0$ to

$$\Omega_i \cap \{\Upsilon_i < t + c_2 - m_i\},$$

where we equip the line bundle L with the singular metric $h_0 e^{-(1+\beta)\phi}$. More precisely, there exists a uniform positive number C_1 (independent of t) and holomorphic extensions $\hat{f}_{i,t}$ $(1 \le i \le N)$ of f from $\Omega_i \cap Y^0$ to $\Omega_i \cap \{\Upsilon_i < t + c_2 - m_i\}$ such that

$$\begin{aligned} \int_{\Omega_i \cap \{\Upsilon_i < t + c_2 - m_i\}} & \frac{|\hat{f}_{i,t}|^2_{\omega,h_0} e^{-(1+\beta)\phi}}{e^{\Upsilon_i - t - c_2 + m_i} R_1(\Upsilon_i - t - c_2 + m_i)}\, dV_{X,\omega} \\ &\le C_1 \int_{\Omega_i \cap Y^0} |f|^2_{\omega,h_0} e^{-(1+\beta)\phi}\, dV_{X,\omega}[\Upsilon_i - t - c_2 + m_i] \\ &\le C_2 e^t \int_{\Omega_i \cap Y^0} |f|^2_{\omega,h_0} e^{-(1+\beta)\phi}\, dV_{X,\omega}[\psi], \end{aligned} \tag{4-3}$$

where C_2 is a positive number independent of t. Furthermore, we get that f is in fact holomorphic on $\Omega_i \cap Y$ and $\hat{f}_{i,t} = f$ on $\Omega_i \cap Y$.

Part III: Construction of local holomorphic extensions $\tilde{f}_{i,t}$ $(1 \le i \le N)$ of f to Ω_i.

For each fixed t, applying Proposition 3.1 to the local extensions $\hat{f}_{i,t}$ $(1 \le i \le N)$ with the weight $(1+\beta)\phi$ and to the case $\Upsilon = \Upsilon_i - t - c_2 + m_i$, $\Omega = \Omega_i$ and some small positive number β_1 which will be determined later in Step 4, we obtain from (4-3) holomorphic sections $\tilde{f}_{i,t}$ $(1 \le i \le N)$ on Ω_i satisfying $\tilde{f}_{i,t} = \hat{f}_{i,t} = f$ on $\Omega_i \cap Y^0$,

$$\int_{\Omega_i \cap \{\Upsilon_i < t + c_2 - m_i\}} \frac{|\tilde{f}_{i,t}|^2_{\omega,h_0} e^{-(1+\beta)\phi}}{e^{\Upsilon_i - t - c_2 + m_i} R_1(\Upsilon_i - t - c_2 + m_i)} \, dV_{X,\omega} \le C_3 e^t, \tag{4-4}$$

and

$$\int_{\Omega_i} \frac{|\tilde{f}_{i,t}|^2_{\omega,h_0} e^{-(1+\beta)\phi}}{(1 + e^{\Upsilon_i - t - c_2 + m_i})^{1+\beta_1}} \, dV_{X,\omega} \le C_3 e^t \tag{4-5}$$

for some positive number C_3 independent of t.

Since $\sup_{t \le 0}(e^t R_1(t)) < +\infty$, it follows from (4-4) that

$$\int_{\Omega_i \cap \{\psi < t + c_2\}} |\tilde{f}_{i,t}|^2_{\omega,h_0} e^{-(1+\beta)\phi} \, dV_{X,\omega} \le C_4 e^t \tag{4-6}$$

for any t, where C_4 is a positive number independent of t.

Since Υ_i is bounded above on Ω_i, it follows from (4-5) that

$$\int_{\Omega_i} |\tilde{f}_{i,t}|^2_{\omega,h_0} e^{-(1+\beta)\phi} \, dV_{X,\omega} \le C_5 e^{-\beta_1 t} \tag{4-7}$$

for any t, where C_5 is a positive number independent of t.

Since $|\tilde{f}_{i,t}|^2$ is plurisubharmonic on Ω_i, by mean value inequality, we get from (4-7) that

$$\sup_{V_i} |\tilde{f}_{i,t}|^2_{\omega,h_0} \le C_6 e^{-\beta_1 t} \tag{4-8}$$

for any t, where C_6 is a positive number independent of t.

Since (4-6) and (4-8) imply that the assumptions in Proposition 3.2 hold for $\tilde{f}_{i,t}$, we apply Proposition 3.2 to $\tilde{f}_{i,t}$ $(1 \le i \le N)$ and get

$$\varlimsup_{t \to -\infty} \int_{U_i \cap \{t - c_1 < \psi < t + c_2\}} \frac{e^t \xi_i |\tilde{f}_{i,t}|^2_{\omega,h_0} e^{-\phi}}{(e^\psi + e^t)^2} \, dV_{X,\omega} \le \int_{U_i \cap Y^0} \xi_i |f|^2_{\omega,h_0} e^{-\phi} \, dV_{X,\omega}[\psi], \tag{4-9}$$

which will be used in Step 4.

Part IV: Construction of a family of smooth extensions $\tilde{f}_t$ of f to a neighborhood of $\overline{X_k} \cap Y$ in X.

Define $\tilde{f}_t = \sum_{i=1}^N \xi_i \tilde{f}_{i,t}$ for all t.

Since

$$\tilde{f}_t|_{U_j} = \sum_{i=1}^{N} \xi_i \tilde{f}_{j,t} + \sum_{i=1}^{N} \xi_i (\tilde{f}_{i,t} - \tilde{f}_{j,t}) = \tilde{f}_{j,t} + \sum_{i=1}^{N} \xi_i (\tilde{f}_{i,t} - \tilde{f}_{j,t})$$

for any $j = 1, \dots, N$, we have

$$|\mathrm{D}'' \tilde{f}_t|_{\omega,h_0}\big|_{U_j} = \Big|\sum_{i=1}^{N} \bar{\partial}\xi_i \wedge (\tilde{f}_{i,t} - \tilde{f}_{j,t})\Big|_{\omega,h_0} \quad \text{for all } t. \tag{4-10}$$

Let μ and W be as in the beginning of the proof of Proposition 3.2 (here W is centered at a point $\tilde{x} \in \overline{\mu^{-1}(U_i \cap U_j)} \cap \{\psi = -\infty\}$). For similar reasons as in (3-18), (3-20) and (3-21), we get from (4-8) that

$$|\tilde{f}_{i,t} \circ \mu - \tilde{f}_{j,t} \circ \mu|^2_{\omega,h_0}\big|_{W_{i,j,t}} \le C_7 e^{-\beta_1 t} \prod_{p \in \kappa} |w_p|^2 \tag{4-11}$$

when $\kappa \neq \varnothing$ and t is small enough, and that

$$|\tilde{f}_{i,t} \circ \mu - \tilde{f}_{j,t} \circ \mu|^2_{\omega,h_0}\big|_{W_{i,j,t}} \le C_7 e^{-\beta_1 t} \tag{4-12}$$

when $\kappa = \varnothing$ and t is small enough, where

$$W_{i,j,t} := W \cap \mu^{-1}(U_i \cap U_j) \cap \{\psi \circ \mu < t + c_2\}$$

and C_7 is a positive number independent of t.

Step 2: Singularity attenuation process for the current $\sqrt{-1}\partial\bar{\partial}\phi$.

Since the singularities of $\sqrt{-1}\partial\bar{\partial}\psi$ obstruct the application of Theorem 2.2, we work on $\widetilde{X}$ first and then go back to X. Some ideas in this step come from [Yi 2012].

Let $\mu : \widetilde{X} \to X$ be as in the beginning of the proof of Proposition 3.2. Let $\widetilde{X}_{k+1} := \mu^{-1}(X_{k+1})$, $\widetilde{X}_k := \mu^{-1}(X_k)$ and $\widetilde{\Sigma_0} := \mu^{-1}(\Sigma_0)$, where $\Sigma_0 := \{\psi = -\infty\}$. Then

$$\gamma_1 := \sqrt{-1}\partial\bar{\partial}(\psi \circ \mu) - \sum_j q_j [D_j]$$

is a smooth real $(1,1)$-form for some positive numbers q_j, where (D_j) are the irreducible components of $\widetilde{\Sigma_0}$. It is not hard to prove the following lemma and we won't give its proof.

Lemma 4.1. *There exists a positive number $\tilde{n}_k$ such that*

$$\widetilde{\omega}_{k+1} := \tilde{n}_k \mu^* \omega + \sqrt{-1}\partial\bar{\partial}(\psi \circ \mu) - \sum_j q_j [D_j]$$

is a Kähler metric on $\widetilde{X}_{k+1}$.

Since $\mu:\widetilde{X}\setminus\widetilde{\Sigma_0}\to X\setminus\Sigma_0$ is biholomorphic and $\sum_j q_j[D_j]\big|_{\widetilde{X}\setminus\widetilde{\Sigma_0}}=0$, the curvature assumptions (i) and (ii) in Theorem 1.8 imply that

$$\sqrt{-1}\partial\bar\partial(\phi\circ\mu)\big|_{\widetilde{X}\setminus\widetilde{\Sigma_0}}+\gamma_2\big|_{\widetilde{X}\setminus\widetilde{\Sigma_0}}\geq 0 \quad\text{and}\quad \sqrt{-1}\partial\bar\partial(\phi\circ\mu)\big|_{\widetilde{X}\setminus\widetilde{\Sigma_0}}+\gamma_3\big|_{\widetilde{X}\setminus\widetilde{\Sigma_0}}\geq 0$$

hold on $\widetilde{X}\setminus\widetilde{\Sigma_0}$, where

$$\gamma_2:=\sqrt{-1}\mu^*\Theta_{L,h_0}+\gamma_1,\quad \gamma_3:=\sqrt{-1}\mu^*\Theta_{L,h_0}+\left(1+\frac{1}{\widetilde{\chi}(\alpha\circ\mu)}\right)\gamma_1.$$

Since γ_2 and γ_3 are continuous on $\widetilde{X}$, and $\phi\circ\mu$ is quasi-plurisubharmonic on $\widetilde{X}$, we get that

$$\sqrt{-1}\partial\bar\partial(\phi\circ\mu)+\gamma_2\geq 0 \tag{4-13}$$

and

$$\sqrt{-1}\partial\bar\partial(\phi\circ\mu)+\gamma_3\geq 0 \tag{4-14}$$

hold on $\widetilde{X}$. Since there must exist a continuous nonnegative (1, 1)-form ϖ_{k+1} on the Kähler manifold $(\widetilde{X}_{k+1},\widetilde{\omega}_{k+1})$ such that

$$(\sqrt{-1}\Theta_{T_{\widetilde{X}_{k+1}}}+\varpi_{k+1}\otimes \mathrm{Id}_{T_{\widetilde{X}_{k+1}}})(\kappa_1\otimes\kappa_2,\kappa_1\otimes\kappa_2)\geq 0 \quad (\text{for all } \kappa_1,\kappa_2\in T_{\widetilde{X}_{k+1}})$$

holds on $\widetilde{X}_{k+1}$, by Theorem 2.2, we obtain from (4-13) and (4-14) a family of functions $\{\widetilde{\phi}_{\varsigma,\rho}\}_{\varsigma>0,\rho\in(0,\rho_1)}$ on a neighborhood of the closure of $\widetilde{X}_k$ such that

(i) $\widetilde{\phi}_{\varsigma,\rho}$ is quasi-plurisubharmonic on a neighborhood of the closure of $\widetilde{X}_k$, smooth on $\widetilde{X}_k\setminus E_\varsigma(\phi\circ\mu)$, increasing with respect to ς and ρ on $\widetilde{X}_k$, and converges to $\phi\circ\mu$ on $\widetilde{X}_k$ as $\rho\to 0$,

(ii) $\dfrac{\sqrt{-1}}{\pi}\partial\bar\partial\widetilde{\phi}_{\varsigma,\rho}\geq -\dfrac{\gamma_2}{\pi}-\varsigma\varpi_{k+1}-\delta_\rho\widetilde{\omega}_{k+1}$ on $\widetilde{X}_k$,

(iii) $\dfrac{\sqrt{-1}}{\pi}\partial\bar\partial\widetilde{\phi}_{\varsigma,\rho}\geq -\dfrac{\gamma_3}{\pi}-\varsigma\varpi_{k+1}-\delta_\rho\widetilde{\omega}_{k+1}$ on $\widetilde{X}_k$,

where $E_\varsigma(\phi\circ\mu):=\{x\in\widetilde{X}:\nu(\phi\circ\mu,x)\geq\varsigma\}$ $(\varsigma>0)$ is the ς-upperlevel set of Lelong numbers of $\phi\circ\mu$, and $\{\delta_\rho\}$ is an increasing family of positive numbers such that $\lim_{\rho\to 0}\delta_\rho=0$.

Since $\widetilde{\omega}_{k+1}$ is a Kähler metric on $\widetilde{X}_{k+1}$ by Lemma 4.1 and $\widetilde{X}_k$ is relatively compact in $\widetilde{X}_{k+1}$, there exists a positive number $n_k>1$ such that $n_k\widetilde{\omega}_{k+1}\geq\varpi_{k+1}$ holds on $\widetilde{X}_k$. Take $\varsigma=\delta_\rho$ and denote $\widetilde{\phi}_{\delta_\rho,\rho}$ simply by $\widetilde{\phi}_\rho$. Then $\widetilde{\phi}_\rho$ is quasi-plurisubharmonic on a neighborhood of the closure of $\widetilde{X}_k$, smooth on $\widetilde{X}_k\setminus E_{\delta_\rho}(\phi\circ\mu)$, increasing with respect to ρ on $\widetilde{X}_k$, and converges to $\phi\circ\mu$ on $\widetilde{X}_k$ as $\rho\to 0$. Furthermore,

$$\sqrt{-1}\partial\bar\partial\widetilde{\phi}_\rho+\gamma_2+2\pi n_k\delta_\rho\widetilde{\omega}_{k+1}\geq 0 \quad\text{and}\quad \sqrt{-1}\partial\bar\partial\widetilde{\phi}_\rho+\gamma_3+2\pi n_k\delta_\rho\widetilde{\omega}_{k+1}\geq 0$$

hold on $\widetilde{X}_k$. Since $\mu : \widetilde{X}_k \setminus \widetilde{\Sigma_0} \to X_k \setminus \Sigma_0$ is biholomorphic, we get that

$$\sqrt{-1}\partial\bar{\partial}(\widetilde{\phi}_\rho \circ \mu^{-1}) + (\mu^{-1})^*\gamma_2 + 2\pi n_k \delta_\rho (\mu^{-1})^* \widetilde{\omega}_{k+1} \geq 0$$

and

$$\sqrt{-1}\partial\bar{\partial}(\widetilde{\phi}_\rho \circ \mu^{-1}) + (\mu^{-1})^*\gamma_3 + 2\pi n_k \delta_\rho (\mu^{-1})^* \widetilde{\omega}_{k+1} \geq 0$$

hold on $X_k \setminus \Sigma_0$. Then, replacing γ_2, γ_3 and $\widetilde{\omega}_{k+1}$ with their definitions, we obtain that

$$\sqrt{-1}\partial\bar{\partial}(\widetilde{\phi}_\rho \circ \mu^{-1}) + \sqrt{-1}\Theta_{L,h_0} + (1+2\pi n_k \delta_\rho)\sqrt{-1}\partial\bar{\partial}\psi \geq -2\pi n_k \tilde{n}_k \delta_\rho \omega \tag{4-15}$$

and

$$\sqrt{-1}\partial\bar{\partial}(\widetilde{\phi}_\rho \circ \mu^{-1}) + \sqrt{-1}\Theta_{L,h_0} + \left(1 + 2\pi n_k \delta_\rho + \frac{1}{\widetilde{\chi}(\alpha)}\right)\sqrt{-1}\partial\bar{\partial}\psi \geq -2\pi n_k \tilde{n}_k \delta_\rho \omega \tag{4-16}$$

hold on $X_k \setminus \Sigma_0$.

Since $E_{\delta_\rho}(\phi \circ \mu)$ is an analytic set in $\widetilde{X}$, Remmert's proper mapping theorem implies that

$$\Sigma_\rho := \mu\big(E_{\delta_\rho}(\phi \circ \mu)\big)$$

is an analytic set in X. By Lemma 2.4, $X_k \setminus (\Sigma_0 \cup \Sigma_\rho)$ is a complete Kähler manifold.

It follows from the properties of $\widetilde{\phi}_\rho$ that $\widetilde{\phi}_\rho \circ \mu^{-1}$ is smooth on $X_k \setminus (\Sigma_0 \cup \Sigma_\rho)$, increasing with respect to ρ on $X_k \setminus \Sigma_0$, uniformly bounded above on $X_k \setminus \Sigma_0$ with respect to ρ, and converges to ϕ on $X_k \setminus \Sigma_0$ as $\rho \to 0$.

In Step 3, we will use $\widetilde{\phi}_\rho \circ \mu^{-1}$ to construct a smooth metric of L on $X_k \setminus (\Sigma_0 \cup \Sigma_\rho)$.

Step 3: Construction of additional weights and twist factors.

Let ζ, χ and η be the solution to the following system of ODEs defined on $(-\infty, \alpha_0)$:

$$\chi(t)\zeta'(t) - \chi'(t) = 1, \tag{4-17}$$

$$\big(\chi(t) + \eta(t)\big)e^{\zeta(t)} = \left(\frac{\alpha_1}{R(\alpha_0)} + C_R\right)R(t), \tag{4-18}$$

$$\frac{(\chi'(t))^2}{\chi(t)\zeta''(t) - \chi''(t)} = \eta(t), \tag{4-19}$$

where we assume that ζ, χ and η are smooth on $(-\infty, \alpha_0)$, and that $\inf_{t<\alpha_0} \zeta(t) = 0$, $\inf_{t<\alpha_0} \chi(t) = \alpha_1$, $\eta > 0$, $\zeta' > 0$ and $\chi' < 0$ on $(-\infty, \alpha_0)$. If $\alpha_0 = +\infty$, we replace the assumption $\inf_{t<\alpha_0} \chi(t) = \alpha_1$ by $\chi > 0$. By a similar calculation as in [Guan and Zhou 2015c] or [Zhou and Zhu 2018], we can solve the system of ODEs and the solution is

$$\begin{aligned}
\chi(t) &= \widetilde{\chi}(t),\\
\zeta(t) &= \log\Big(\frac{\alpha_1}{R(\alpha_0)}+C_R\Big)-\log\Big(\frac{\alpha_1}{R(\alpha_0)}+\int_t^{\alpha_0}\frac{dt_1}{R(t_1)}\Big),\\
\eta(t) &= R(t)\Big(\frac{\alpha_1}{R(\alpha_0)}+\int_t^{\alpha_0}\frac{dt_1}{R(t_1)}\Big)-\widetilde{\chi}(t),
\end{aligned}$$

where $\widetilde{\chi}(t)$ is defined by (1-2).

Let $\varepsilon\in\big(0,\frac{1}{2}\big)$ be as in Step 1 and put $\sigma_t=\log(e^{\psi}+e^{t})-\varepsilon$. Then there exists a negative number t_ε such that $\sigma_t\le\alpha-\frac{\varepsilon}{2}$ on $\overline{X_k}$ for any $t\in(-\infty,t_\varepsilon)$.

Let $h_{\rho,t}$ be the new metric on the line bundle L over $X_k\setminus(\Sigma_0\cup\Sigma_\rho)$ defined by

$$h_{\rho,t}:=h_0e^{-\widetilde{\phi}_\rho\circ\mu^{-1}-(1+2\pi n_k\delta_\rho)\psi-\zeta(\sigma_t)}.$$

Let $\tau_t:=\chi(\sigma_t)$ and $A_t:=\eta(\sigma_t)$. Set $\mathrm{B}_{\rho,t}=[\Theta_{\rho,t},\ \Lambda]$ on $X_k\setminus(\Sigma_0\cup\Sigma_\rho)$, where

$$\Theta_{\rho,t}:=\tau_t\sqrt{-1}\Theta_{L,h_{\rho,t}}-\sqrt{-1}\partial\bar\partial\tau_t-\sqrt{-1}\frac{\partial\tau_t\wedge\bar\partial\tau_t}{A_t}.$$

Set $\nu_t=\partial\sigma_t$. We want to prove

$$\Theta_{\rho,t}\big|_{X_k\setminus(\Sigma_0\cup\Sigma_\rho)}\ge\frac{e^{t}}{e^{\psi}}\sqrt{-1}\nu_t\wedge\bar\nu_t-2\pi n_k\widetilde{n}_k\chi(\sigma_t)\delta_\rho\omega. \tag{4-20}$$

It follows from (4-17) and (4-19) that

$$\begin{aligned}
&\Theta_{\rho,t}\big|_{X_k\setminus(\Sigma_0\cup\Sigma_\rho)}\\
&=\chi(\sigma_t)\big(\sqrt{-1}\Theta_{L,h_0}+\sqrt{-1}\partial\bar\partial(\widetilde{\phi}_\rho\circ\mu^{-1})+(1+2\pi n_k\delta_\rho)\sqrt{-1}\partial\bar\partial\psi\big)\\
&\qquad+\big(\chi(\sigma_t)\zeta'(\sigma_t)-\chi'(\sigma_t)\big)\sqrt{-1}\partial\bar\partial\sigma_t\\
&\qquad+\Big(\chi(\sigma_t)\zeta''(\sigma_t)-\chi''(\sigma_t)-\frac{(\chi'(\sigma_t))^2}{\eta(\sigma_t)}\Big)\sqrt{-1}\partial\sigma_t\wedge\bar\partial\sigma_t\\
&=\chi(\sigma_t)\big(\sqrt{-1}\Theta_{L,h_0}+\sqrt{-1}\partial\bar\partial(\widetilde{\phi}_\rho\circ\mu^{-1})+(1+2\pi n_k\delta_\rho)\sqrt{-1}\partial\bar\partial\psi\big)+\sqrt{-1}\partial\bar\partial\sigma_t\\
&=\chi(\sigma_t)\big(\sqrt{-1}\Theta_{L,h_0}+\sqrt{-1}\partial\bar\partial(\widetilde{\phi}_\rho\circ\mu^{-1})+(1+2\pi n_k\delta_\rho)\sqrt{-1}\partial\bar\partial\psi\big)\\
&\qquad+\frac{e^{t}}{e^{\psi}}\sqrt{-1}\nu_t\wedge\bar\nu_t+\frac{e^{\psi}}{e^{\psi}+e^{t}}\sqrt{-1}\partial\bar\partial\psi.
\end{aligned}$$

Since χ is decreasing and $\chi=\widetilde{\chi}$, it follows from (4-15) and (4-16) that

$$\begin{aligned}
&\chi(\sigma_t)\big(\sqrt{-1}\Theta_{L,h_0}+\sqrt{-1}\partial\bar\partial(\widetilde{\phi}_\rho\circ\mu^{-1})+(1+2\pi n_k\delta_\rho)\sqrt{-1}\partial\bar\partial\psi\big)+\frac{e^{\psi}}{e^{\psi}+e^{t}}\sqrt{-1}\partial\bar\partial\psi\\
&=\chi(\sigma_t)\big(\sqrt{-1}\Theta_{L,h_0}+\sqrt{-1}\partial\bar\partial(\widetilde{\phi}_\rho\circ\mu^{-1})+(1+2\pi n_k\delta_\rho)\sqrt{-1}\partial\bar\partial\psi+2\pi n_k\tilde n_k\delta_\rho\omega\big)\\
&\qquad-2\pi n_k\tilde n_k\chi(\sigma_t)\delta_\rho\omega+\frac{\chi(\alpha)e^{\psi}}{e^{\psi}+e^{t}}\cdot\frac{\sqrt{-1}\partial\bar\partial\psi}{\chi(\alpha)}
\end{aligned}$$

$$\geq \frac{\chi(\alpha)e^{\psi}}{e^{\psi}+e^{t}}\Big(\sqrt{-1}\Theta_{L,h_0}+\sqrt{-1}\partial\bar\partial(\widetilde{\phi}_\rho\circ\mu^{-1})+(1+2\pi n_k\delta_\rho)\sqrt{-1}\partial\bar\partial\psi$$
$$+2\pi n_k\tilde n_k\delta_\rho\omega+\frac{\sqrt{-1}\partial\bar\partial\psi}{\chi(\alpha)}\Big)-2\pi n_k\tilde n_k\chi(\sigma_t)\delta_\rho\omega$$
$$\geq -2\pi n_k\tilde n_k\chi(\sigma_t)\delta_\rho\omega$$

on $X_k\setminus(\Sigma_0\cup\Sigma_\rho)$. Hence we get (4-20) as desired.

Let β be as in Step 1. Let β_0 and β_3 be two positive numbers which will be determined later in Step 4. We choose an increasing family of positive numbers $\{\rho_t\}_{t\in(-\infty,t_\varepsilon)}$ such that $\lim_{t\to-\infty}\rho_t=0$ and for any t,

$$2\pi n_k\tilde n_k\chi(t-1)\delta_{\rho_t}<e^{\beta_0 t}, \tag{4-21}$$
$$2\pi n_k\delta_{\rho_t}<\beta_3, \tag{4-22}$$
$$\Big(\frac{\varepsilon}{2-\varepsilon}e^t\Big)^{2\pi n_k\delta_{\rho_t}}>\frac{1}{1+\varepsilon}. \tag{4-23}$$

Since $\sigma_t\geq t-1$ on X_k and χ is decreasing, we have $\chi(\sigma_t)\leq\chi(t-1)$ on X_k. Then it follows from (4-20) and (4-21) that

$$\Theta_{\rho_t,t}\big|_{X_k\setminus(\Sigma_0\cup\Sigma_{\rho_t})}\geq\frac{e^t}{e^\psi}\sqrt{-1}\nu_t\wedge\bar\nu_t-e^{\beta_0 t}\omega.$$

Hence

$$\mathrm{B}_{\rho_t,t}+e^{\beta_0 t}\mathrm{I}\geq\Big[\frac{e^t}{e^\psi}\sqrt{-1}\nu_t\wedge\bar\nu_t,\ \Lambda\Big]=\frac{e^t}{e^\psi}\mathrm{T}_{\bar\nu_t}\mathrm{T}^*_{\bar\nu_t}\geq 0 \tag{4-24}$$

on $X_k\setminus(\Sigma_0\cup\Sigma_{\rho_t})$ as an operator on $(n,1)$-forms, where $\mathrm{T}_{\bar\nu_t}$ denotes the operator $\bar\nu_t\wedge\bullet$ and $\mathrm{T}^*_{\bar\nu_t}$ is its Hilbert adjoint operator.

Step 4: Construction of suitably truncated forms and solving $\bar\partial$ globally with L^2 estimates.

In this step and Step 5, we denote $\mathrm{B}_{\rho_t,t}$ and $h_{\rho_t,t}$ simply by B_t and h_t respectively.

Let $\varepsilon\in\big(0,\frac{1}{2}\big)$ be as in Step 1. It is easy to construct a smooth function $\theta:\mathbb{R}\to[0,1]$ such that $\theta=0$ on $\big(-\infty,\frac{\varepsilon}{2}\big]$, $\theta=1$ on $\big[1-\frac{\varepsilon}{2},+\infty\big)$ and $|\theta'|\leq(1+\varepsilon)/(1-\varepsilon)$ on $\mathbb{R}$.

Define $g_t=\mathrm{D}''\big(\theta(e^t/(e^\psi+e^t))\tilde f_t\big)$, where $\tilde f_t$ is constructed in Step 1. Then $\mathrm{D}''g_t=0$ and

$$g_t=-\theta'\Big(\frac{e^t}{e^\psi+e^t}\Big)\frac{e^{\psi+t}}{(e^\psi+e^t)^2}\bar\partial\psi\wedge\tilde f_t+\theta\Big(\frac{e^t}{e^\psi+e^t}\Big)\mathrm{D}''\tilde f_t=g_{1,t}+g_{2,t},$$

where $g_{1,t}$ and $g_{2,t}$ denote

$$-\bar\nu_t\wedge\theta'\Big(\frac{e^t}{e^\psi+e^t}\Big)\frac{e^t}{e^\psi+e^t}\tilde f_t\quad\text{and}\quad\theta\Big(\frac{e^t}{e^\psi+e^t}\Big)\mathrm{D}''\tilde f_t,$$

respectively.

Then

$$\operatorname{supp} g_{1,t} \subset \{t-c_1<\psi<t+c_2\} \quad \text{and} \quad \operatorname{supp} g_{2,t} \subset \{\psi<t+c_2\},$$

where c_1 and c_2 are defined as in Step 1.

It follows from (3-5) and (4-24) that

$$\begin{aligned}
(4\text{-}25)\quad & \langle (\mathrm{B}_t+2e^{\beta_0 t}\mathrm{I})^{-1} g_t, g_t\rangle_{\omega,h_t}\big|_{X_k\setminus(\Sigma_0\cup\Sigma_{\rho_t})} \\
&\le (1+\varepsilon)\langle (\mathrm{B}_t+2e^{\beta_0 t}\mathrm{I})^{-1} g_{1,t}, g_{1,t}\rangle_{\omega,h_t} + \frac{1+\varepsilon}{\varepsilon}\langle (\mathrm{B}_t+2e^{\beta_0 t}\mathrm{I})^{-1} g_{2,t}, g_{2,t}\rangle_{\omega,h_t} \\
&\le (1+\varepsilon)\langle (\mathrm{B}_t+e^{\beta_0 t}\mathrm{I})^{-1} g_{1,t}, g_{1,t}\rangle_{\omega,h_t} + \frac{1+\varepsilon}{\varepsilon}\Big\langle \frac{1}{e^{\beta_0 t}} g_{2,t}, g_{2,t}\Big\rangle_{\omega,h_t}.
\end{aligned}$$

By (4-24), we have

$$\langle (\mathrm{B}_t+e^{\beta_0 t}\mathrm{I})^{-1} g_{1,t}, g_{1,t}\rangle_{\omega,h_t}\big|_{X_k\setminus(\Sigma_0\cup\Sigma_{\rho_t})} \le \frac{e^{\psi}}{e^t}\Big|\theta'\Big(\frac{e^t}{e^{\psi}+e^t}\Big)\frac{e^t}{e^{\psi}+e^t}\tilde{f}_t\Big|^2_{\omega,h_t}.$$

Then $\zeta>0$ implies that

$$\begin{aligned}
I_{1,t} &:= \int_{X_k\setminus(\Sigma_0\cup\Sigma_{\rho_t})} \langle (\mathrm{B}_t+e^{\beta_0 t}\mathrm{I})^{-1} g_{1,t}, g_{1,t}\rangle_{\omega,h_t}\, dV_{X,\omega} \\
&\le \frac{(1+\varepsilon)^2}{(1-\varepsilon)^2}\int_{X_k\cap\{t-c_1<\psi<t+c_2\}} \frac{e^t|\tilde{f}_t|^2_{\omega,h_0} e^{-\widetilde{\phi}_{\rho_t}\circ\mu^{-1}}}{(e^{\psi}+e^t)^2 e^{2\pi n_k\delta_{\rho_t}\psi}}\, dV_{X,\omega}.
\end{aligned}$$

Since $\widetilde{\phi}_{\rho_t}\circ\mu^{-1}\ge\phi$ on $X_k\setminus\Sigma_0$, it follows from (4-23) that

$$\begin{aligned}
I_{1,t} &\le \frac{(1+\varepsilon)^2}{(1-\varepsilon)^2}\int_{X_k\cap\{t-c_1<\psi<t+c_2\}} \frac{e^t|\tilde{f}_t|^2_{\omega,h_0} e^{-\phi}\, dV_{X,\omega}}{(e^{\psi}+e^t)^2\big(\frac{\varepsilon}{2-\varepsilon}e^t\big)^{2\pi n_k\delta_{\rho_t}}} \\
&\le \frac{(1+\varepsilon)^3}{(1-\varepsilon)^2}\int_{X_k\cap\{t-c_1<\psi<t+c_2\}} \frac{e^t|\tilde{f}_t|^2_{\omega,h_0} e^{-\phi}\, dV_{X,\omega}}{(e^{\psi}+e^t)^2}.
\end{aligned}$$

Since

$$|\tilde{f}_t|^2_{\omega,h_0}\big|_U = \Big|\sum_{i=1}^N \sqrt{\xi_i}\cdot\sqrt{\xi_i}\,\tilde{f}_{i,t}\Big|^2_{\omega,h_0} \le \Big(\sum_{i=1}^N \xi_i\Big)\Big(\sum_{i=1}^N \xi_i|\tilde{f}_{i,t}|^2_{\omega,h_0}\Big) = \sum_{i=1}^N \xi_i|\tilde{f}_{i,t}|^2_{\omega,h_0}$$

by the Cauchy–Schwarz inequality, we have

$$I_{1,t} \le \frac{(1+\varepsilon)^3}{(1-\varepsilon)^2}\sum_{i=1}^N\int_{X_k\cap\{t-c_1<\psi<t+c_2\}} \frac{e^t\xi_i|\tilde{f}_{i,t}|^2_{\omega,h_0} e^{-\phi}\, dV_{X,\omega}}{(e^{\psi}+e^t)^2}.$$

Then it follows from (4-9) that

$$\begin{aligned}\varlimsup_{t\to-\infty} I_{1,t} &\le \sum_{i=1}^{N} \varlimsup_{t\to-\infty}\left(\frac{(1+\varepsilon)^3}{(1-\varepsilon)^2}\int_{X_k\cap\{t-c_1<\psi<t+c_2\}} \frac{e^t\xi_i|\tilde{f}_{i,t}|^2_{\omega,h_0}e^{-\phi}dV_{X,\omega}}{(e^{\psi}+e^t)^2}\right)\\ &\le \sum_{i=1}^{N}\frac{(1+\varepsilon)^3}{(1-\varepsilon)^2}\int_{U_i\cap Y^0}\xi_i|f|^2_{\omega,h_0}e^{-\phi}\,dV_{X,\omega}[\psi]\\ &\le \frac{(1+\varepsilon)^3}{(1-\varepsilon)^2}\int_{Y^0}|f|^2_{\omega,h_0}e^{-\phi}\,dV_{X,\omega}[\psi].\end{aligned}$$

Then

$$I_{1,t} \le \frac{(1+\varepsilon)^4}{(1-\varepsilon)^2}\int_{Y^0}|f|^2_{\omega,h_0}e^{-\phi}\,dV_{X,\omega}[\psi] \tag{4-26}$$

when t is small enough.

Since $\zeta(\sigma_t)>0$ and $\widetilde{\phi}_{\rho_t}\circ\mu^{-1}\ge\phi$ on $X_k\setminus\Sigma_0$, by (4-22), we have

$$\begin{aligned} I_{2,t} &:= \int_{X_k\setminus(\Sigma_0\cup\Sigma_{\rho_t})}\left\langle\frac{1}{e^{\beta_0 t}}g_{2,t},\,g_{2,t}\right\rangle_{\omega,h_t} dV_{X,\omega}\\ &\le \frac{1}{e^{\beta_0 t}}\int_{X_k\cap\{\psi<t+c_2\}}\frac{|\mathrm{D}''\tilde{f}_t|^2_{\omega,h_0}e^{-\widetilde{\phi}_{\rho_t}\circ\mu^{-1}}}{e^{(1+2\pi n_k\delta_{\rho_t})\psi}}\,dV_{X,\omega}\\ &\le \frac{1}{e^{\beta_0 t}}\int_{X_k\cap\{\psi<t+c_2\}}\frac{|\mathrm{D}''\tilde{f}_t|^2_{\omega,h_0}e^{-\phi}}{e^{(1+\beta_3)\psi}}\,dV_{X,\omega}.\end{aligned}$$

Then it follows from (4-10) and the Cauchy–Schwarz inequality that $I_{2,t}$ is bounded by the sum of the terms

$$\frac{C_8}{e^{\beta_0 t}}\int_{U_i\cap U_j\cap\{\psi<t+c_2\}}\frac{|\tilde{f}_{i,t}-\tilde{f}_{j,t}|^2_{\omega,h_0}e^{-\phi}}{e^{(1+\beta_3)\psi}}\,dV_{X,\omega}\qquad(1\le i,\,j\le N),$$

where C_8 is some positive number independent of t.

By the definition of R_1 (see Part II in Step 1), (4-4) implies that for $i=1,\dots,N$,

$$\int_{\Omega_i\cap\{\psi<t+c_2\}}\frac{|\tilde{f}_{i,t}|^2_{\omega,h_0}e^{-(1+\beta)\phi}}{e^{\psi}R_0(\psi)}\,dV_{X,\omega}\le C_9 \tag{4-27}$$

for some positive number C_9 independent of t when t is small enough. Then by the Hölder inequality, we get

$$\begin{aligned}
&\int_{U_i\cap U_j\cap\{\psi<t+c_2\}} \frac{|\tilde f_{i,t}-\tilde f_{j,t}|^2_{\omega,h_0}e^{-\phi}}{e^{(1+\beta_3)\psi}}\,dV_{X,\omega} \\
&\quad\le \biggl(\int_{U_i\cap U_j\cap\{\psi<t+c_2\}} \frac{|\tilde f_{i,t}-\tilde f_{j,t}|^2_{\omega,h_0}e^{-(1+\beta)\phi}}{e^{\psi}R_0(\psi)}\,dV_{X,\omega}\biggr)^{\frac{1}{1+\beta}} \\
&\qquad\qquad\times\biggl(\int_{U_i\cap U_j\cap\{\psi<t+c_2\}} \frac{|\tilde f_{i,t}-\tilde f_{j,t}|^2_{\omega,h_0}(R_0(\psi))^{\frac{1}{\beta}}}{e^{\left(1+\beta_3\cdot\frac{1+\beta}{\beta}\right)\psi}}\,dV_{X,\omega}\biggr)^{\frac{\beta}{1+\beta}} \\
&\quad\le C_{10}\biggl(\int_{U_i\cap U_j\cap\{\psi<t+c_2\}} \frac{|\tilde f_{i,t}-\tilde f_{j,t}|^2_{\omega,h_0}}{e^{\left(1+\beta_3\cdot\frac{1+\beta}{\beta}+\beta_2\cdot\frac{1}{\beta}\right)\psi}}\,dV_{X,\omega}\biggr)^{\frac{\beta}{1+\beta}}
\end{aligned}$$

when t is small enough, where C_{10} is a positive number independent of t.

We will estimate the last integral above by estimating its pull-back under μ. We cover $\mu^{-1}(U_i\cap U_j)\cap\{\psi\circ\mu<t+c_2\}$ by a finite number of coordinate balls such as W in Step 1 in the proof of Proposition 3.2. It follows from (4-11) and (4-12) that for each W,

$$\int_{W_{i,j,t}} \frac{|\tilde f_{i,t}\circ\mu-\tilde f_{j,t}\circ\mu|^2_{\omega,h_0}|J_\mu|^2}{e^{\left(1+\beta_3\cdot\frac{1+\beta}{\beta}+\beta_2\cdot\frac{1}{\beta}\right)\psi\circ\mu}}\,d\lambda(w)\le C_{11}\int_{W_{i,j,t}}\frac{1}{\prod_{p=1}^n|w_p|^{2\beta_{5,p}}}\,d\lambda(w),$$

where

$$\begin{aligned}
\beta_{5,p}&:=\beta_4ca_p+(ca_p-b_p)-\lfloor ca_p-b_p\rfloor_+,\\
\beta_4&:=\beta_3\cdot\frac{1+\beta}{\beta}+\beta_2\cdot\frac{1}{\beta}+\beta_1,
\end{aligned}$$

and C_{11} is a positive number independent of t.

Since

$$\bigl(W\cap\{\psi\circ\mu<t+c_2\}\bigr)\subset\bigcup_{p=1}^n\bigl(\bigl\{|w_p|<e^{(t+c_2-m)/(2c|a|)}\bigr\}\cap W\bigr),$$

where $m:=\inf_W\tilde u(w)$, we obtain

$$\begin{aligned}
\int_{W_{i,j,t}}\frac{1}{\prod_{p=1}^n|w_p|^{2\beta_{5,p}}}\,d\lambda(w)&\le\sum_{p=1}^n\int_{\{|w_p|<e^{(t+c_2-m)/(2c|a|)}\}\cap W}\frac{1}{\prod_{p=1}^n|w_p|^{2\beta_{5,p}}}\,d\lambda(w)\\
&\le C_{12}\sum_{p=1}^n e^{((1-\beta_{5,p})/c|a|)t}
\end{aligned}$$

when $\max_{1\le p\le n}\beta_{5,p}<1$, where C_{12} is a positive number independent of t.

Let β_1 be a positive number such that

$$\beta_1 < \min_{\{p\,:\,a_p\neq 0\}} \frac{1-(ca_p-b_p)+\lfloor ca_p-b_p\rfloor_+}{3ca_p}. \tag{4-28}$$

Take $\beta_2=\beta_1\beta$, $\beta_3=\beta_1\beta/(1+\beta)$. Then $\beta_4=3\beta_1$ and $\max_{1\le p\le n}\beta_{5,p}<1$. Let β_0 be a positive number such that

$$\beta_0 < \min_{1\le p\le n} \frac{\beta(1-\beta_{5,p})}{2(1+\beta)c|a|}$$

for every W. Then we have

$$I_{2,t} \le C_{13}\cdot e^{\beta_0 t}, \tag{4-29}$$

where C_{13} is a positive number independent of t.

Therefore, it follows from (4-25), (4-26) and (4-29) that

$$\int_{X_k\setminus(\Sigma_0\cup\Sigma_{\rho_t})} \langle(\mathrm{B}_t+2e^{\beta_0 t}\mathrm{I})^{-1}g_t,\,g_t\rangle_{\omega,h_t}\,dV_{X,\omega} \le (1+\varepsilon)I_{1,t}+\frac{1+\varepsilon}{\varepsilon}I_{2,t}\le C(t),$$

where

$$C(t):=\frac{(1+\varepsilon)^5}{(1-\varepsilon)^2}\int_{Y^0}|f|^2_{\omega,h_0}e^{-\phi}\,dV_{X,\omega}[\psi]+\frac{1+\varepsilon}{\varepsilon}C_{13}\cdot e^{\beta_0 t}.$$

Then by Lemma 2.1, there exists $u_{k,\varepsilon,t}\in L^2(X_k\setminus(\Sigma_0\cup\Sigma_{\rho_t}),\ K_X\otimes L,\ h_t)$ and $v_{k,\varepsilon,t}\in L^2(X_k\setminus(\Sigma_0\cup\Sigma_{\rho_t}),\ \wedge^{n,1}T^*_X\otimes L,\ h_t)$ such that

$$\mathrm{D}''u_{k,\varepsilon,t}+\sqrt{2e^{\beta_0 t}}\,v_{k,\varepsilon,t}=g_t \tag{4-30}$$

on $X_k\setminus(\Sigma_0\cup\Sigma_{\rho_t})$ and

$$\int_{X_k\setminus(\Sigma_0\cup\Sigma_{\rho_t})}\frac{|u_{k,\varepsilon,t}|^2_{\omega,h_0}e^{-\widetilde{\phi}_{\rho_t}\circ\mu^{-1}-(1+2\pi n_k\delta_{\rho_t})\psi-\zeta(\sigma_t)}}{\tau_t+A_t}dV_{X,\omega} + \int_{X_k\setminus(\Sigma_0\cup\Sigma_{\rho_t})}|v_{k,\varepsilon,t}|^2_{\omega,h_0}e^{-\widetilde{\phi}_{\rho_t}\circ\mu^{-1}-(1+2\pi n_k\delta_{\rho_t})\psi-\zeta(\sigma_t)}\,dV_{X,\omega}\le C(t). \tag{4-31}$$

Since $\{\widetilde{\phi}_{\rho_t}\circ\mu^{-1}\}$ are uniformly bounded above on $X_k\setminus\Sigma_0$ with respect to t as obtained in Step 2, we have

$$e^{-\widetilde{\phi}_{\rho_t}\circ\mu^{-1}}\ge C_{14} \tag{4-32}$$

on $X_k\setminus\Sigma_0$ for any t, where C_{14} is a positive number independent of t. Since $t-\varepsilon\le\sigma_t\le\alpha-\frac{\varepsilon}{2}$ on $\overline{X_k}$ and ψ is upper semicontinuous on X, we have that ψ, $\zeta(\sigma_t)$ and τ_t+A_t are all bounded above on $\overline{X_k}$ for each fixed t. Then it follows from (4-31) that $u_{k,\varepsilon,t} \in L^2$ and $v_{k,\varepsilon,t} \in L^2$. Hence it follows from (4-30) and Lemma 2.6 that

$$\mathrm{D}'' u_{k,\varepsilon,t} + \sqrt{2e^{\beta_0 t}}\, v_{k,\varepsilon,t} = \mathrm{D}''\left(\theta\left(\frac{e^t}{e^\psi + e^t}\right)\tilde{f}_t\right) \tag{4-33}$$

holds on X_k. Furthermore, (4-31) and (4-18) imply that

$$\int_{X_k} \frac{|u_{k,\varepsilon,t}|^2_{\omega,h_0} e^{-\widetilde{\phi}_{\rho_t}\circ\mu^{-1}}}{\left(\frac{\alpha_1}{R(\alpha_0)} + C_R\right) e^\psi R(\sigma_t)}\, dV_{X,\omega} + \int_{X_k} |v_{k,\varepsilon,t}|^2_{\omega,h_0} e^{-\widetilde{\phi}_{\rho_t}\circ\mu^{-1} - \psi - \zeta(\sigma_t)}\, dV_{X,\omega} \le e^{2\pi n_k \delta_{\rho_t} M_\psi} C(t), \tag{4-34}$$

where $M_\psi := \sup_{X_k} \psi$.

Define

$$F_{k,\varepsilon,t} = -u_{k,\varepsilon,t} + \theta\left(\frac{e^t}{e^\psi + e^t}\right)\tilde{f}_t.$$

Then (4-33) implies that $\mathrm{D}'' F_{k,\varepsilon,t} = \sqrt{2e^{\beta_0 t}}\, v_{k,\varepsilon,t}$ on X_k. Since $\widetilde{\phi}_{\rho_t}\circ\mu^{-1} \ge \phi$ on $X_k \setminus \Sigma_0$, it follows from (3-5) and (4-34) that

$$\begin{aligned} &\int_{X_k} \frac{|F_{k,\varepsilon,t}|^2_{\omega,h_0} e^{-\widetilde{\phi}_{\rho_t}\circ\mu^{-1}}}{e^\psi \max\{R(\psi-\varepsilon), R(\sigma_t)\}}\, dV_{X,\omega} \\ &\quad \le (1+\varepsilon)\int_{X_k} \frac{|u_{k,\varepsilon,t}|^2_{\omega,h_0} e^{-\widetilde{\phi}_{\rho_t}\circ\mu^{-1}}}{e^\psi R(\sigma_t)}\, dV_{X,\omega} \\ &\qquad + \frac{1+\varepsilon}{\varepsilon}\int_{X_k} \frac{\left|\theta\left(\frac{e^t}{e^\psi+e^t}\right)\tilde{f}_t\right|^2_{\omega,h_0} e^{-\widetilde{\phi}_{\rho_t}\circ\mu^{-1}}}{e^\psi R(\psi-\varepsilon)}\, dV_{X,\omega} \\ &\quad \le (1+\varepsilon) e^{2\pi n_k \delta_{\rho_t} M_\psi}\left(\frac{\alpha_1}{R(\alpha_0)} + C_R\right) C(t) + \widetilde{C}(t) \end{aligned} \tag{4-35}$$

when t is small enough, where

$$\widetilde{C}(t) := \frac{1+\varepsilon}{\varepsilon}\int_{X_k\cap\{\psi<t+c_2\}} \frac{|\tilde{f}_t|^2_{\omega,h_0} e^{-\phi}}{e^\psi R(\psi-\varepsilon)}\, dV_{X,\omega}.$$

Now we want to prove

$$\lim_{t\to-\infty} \widetilde{C}(t) = 0. \tag{4-36}$$

As in (4-27), we can obtain from (4-4) that for $i = 1, \dots, N$,

$$\int_{\Omega_i\cap\{\psi<t+c_2\}} \frac{|\tilde{f}_{i,t}|^2_{\omega,h_0} e^{-(1+\beta)\phi}}{e^\psi R(\psi-\varepsilon)}\, dV_{X,\omega} \le C_{15}$$

for some positive number C_{15} independent of t when t is small enough. Then by the Hölder inequality, we have that

$$\begin{aligned}\int_{U_i\cap X_k\cap\{\psi<t+c_2\}} &\frac{|\tilde{f}_{i,t}|^2_{\omega,h_0}e^{-\phi}}{e^{\psi}R(\psi-\varepsilon)}\,dV_{X,\omega}\\ &\le\left(\int_{U_i\cap\{\psi<t+c_2\}}\frac{|\tilde{f}_{i,t}|^2_{\omega,h_0}e^{-(1+\beta)\phi}}{e^{\psi}R(\psi-\varepsilon)}\,dV_{X,\omega}\right)^{\frac{1}{1+\beta}}\\ &\qquad\times\left(\int_{U_i\cap\{\psi<t+c_2\}}\frac{|\tilde{f}_{i,t}|^2_{\omega,h_0}}{e^{\psi}R(\psi-\varepsilon)}\,dV_{X,\omega}\right)^{\frac{\beta}{1+\beta}}\\ &\le C_{15}^{\frac{1}{1+\beta}}\left(\int_{U_i\cap\{\psi<t+c_2\}}\frac{|\tilde{f}_{i,t}|^2_{\omega,h_0}}{e^{\psi}R(\psi-\varepsilon)}\,dV_{X,\omega}\right)^{\frac{\beta}{1+\beta}}\end{aligned}$$

when t is small enough.

We cover $\mu^{-1}(U_i)\cap\{\psi\circ\mu<t+c_2\}$ by a finite number of coordinate balls such as W in Step 1 in the proof of Proposition 3.2. Then, in order to prove $\lim_{t\to-\infty}\widetilde{C}(t)=0$, it suffices to prove

$$\lim_{t\to-\infty}\int_{W_{i,t}}\frac{|\tilde{f}_{i,t}\circ\mu|^2_{\omega,h_0}|J_\mu|^2}{e^{\psi\circ\mu}R(\psi\circ\mu-\varepsilon)}\,d\lambda(w)=0,$$

where

$$W_{i,t}:=W\cap\mu^{-1}(U_i)\cap\{\psi\circ\mu<t+c_2\}.$$

Then by (3-18)–(3-21), it suffices to prove

$$\lim_{t\to-\infty}\int_{W_{i,t}}\frac{d\lambda(w)}{R(\psi\circ\mu-\varepsilon)|w_{p_0}|^2\prod_{1\le p\le n,p\ne p_0}|w_p|^{2(ca_p-b_p)-2\lfloor ca_p-b_p\rfloor_+}}=0 \tag{4-37}$$

in Case (A) and

$$\lim_{t\to-\infty}\int_{W_{i,t}}\frac{d\lambda(w)}{R(\psi\circ\mu-\varepsilon)\prod_{p=1}^n|w_p|^{2\beta_1ca_p+2(ca_p-b_p)-2\lfloor ca_p-b_p\rfloor_+}}=0 \tag{4-38}$$

in Case (A) and Case (B).

Applying Fubini's theorem with respect to (w',w_{p_0}) and then using change of variables, we can obtain that

$$\begin{aligned}\lim_{t\to-\infty}\int_{W_{i,t}}&\frac{d\lambda(w)}{R(\psi\circ\mu-\varepsilon)|w_{p_0}|^2\prod_{1\le p\le n,p\ne p_0}|w_p|^{2(ca_p-b_p)-2\lfloor ca_p-b_p\rfloor_+}}\\ &\le C_{16}\lim_{t\to-\infty}\int_{-\infty}^{t+c_2-m}\frac{ds}{R(s+M-\varepsilon)}=0,\end{aligned}$$

where $M:=\sup_W\tilde{u}(w)$, $m:=\inf_W\tilde{u}(w)$ and C_{16} is a positive number independent of t. Hence we get (4-37).

Similarly, it is easy to see that (4-28) implies that (4-38).

Therefore, we obtain (4-36).

Let $\widehat{\alpha}_k := \sup_{X_k} \alpha$. Then

$$e^{\psi} \max\{R(\psi-\varepsilon), R(\sigma_t)\} \le e^{\varepsilon} \sup_{t\le\widehat{\alpha}_k}\big(e^t R(t)\big).$$

Hence it follows from (4-32) and (4-35) that

$$\int_{X_k} |F_{k,\varepsilon,t}|^2_{\omega,h_0}\, dV_{X,\omega} \le C_{17} \tag{4-39}$$

for some positive number C_{17} independent of t when t is small enough.

Since the positive continuous function R is decreasing near $-\infty$, it is easy to see that $\max\{R(\psi-\varepsilon), R(\sigma_t)\}$ is equal to $R(\psi-\varepsilon)$ near $\{\psi=-\infty\}$ and converges uniformly to $R(\psi-\varepsilon)$ on $\overline{X_k}$ as $t\to-\infty$.

Since $\widetilde{\phi}_{\rho_t}\circ\mu^{-1}$ is increasing with respect to t and converges to ϕ on $X_k\setminus\Sigma_0$ as $t\to-\infty$, by extracting weak limits of $\{F_{k,\varepsilon,t}\}$ as $t\to-\infty$, we get from (4-39) and (4-35) a sequence $\{t_j\}_{j=1}^{+\infty}$ and $F_{k,\varepsilon}\in L^2$ such that $\lim_{j\to+\infty} t_j=-\infty$, $F_{k,\varepsilon,t_j}\rightharpoonup F_{k,\varepsilon}$ weakly in L^2 as $j\to+\infty$ and

$$\int_{X_k} \frac{|F_{k,\varepsilon}|^2_{\omega,h_0} e^{-\phi}}{e^{\psi} R(\psi-\varepsilon)}\, dV_{X,\omega} \le \frac{(1+\varepsilon)^6}{(1-\varepsilon)^2}\left(\frac{\alpha_1}{R(\alpha_0)}+C_R\right)\int_{Y^0} |f|^2_{\omega,h_0} e^{-\phi}\, dV_{X,\omega}[\psi]. \tag{4-40}$$

Since $\sigma_t\le\alpha-\frac{\varepsilon}{2}$ on X_k, $\widehat{\alpha}_k := \sup_{X_k}\alpha$ and ζ is increasing, we get

$$e^{-\zeta(\sigma_t)} \ge e^{-\zeta\left(\widehat{\alpha}_k-\frac{\varepsilon}{2}\right)} \tag{4-41}$$

on X_k. Then (4-34), (4-32) and (4-41) imply that

$$\int_{X_k} |v_{k,\varepsilon,t}|^2_{\omega,h_0}\, dV_{X,\omega} \le e^{\zeta\left(\widehat{\alpha}_k-\frac{\varepsilon}{2}\right)+(1+2\pi n_k\delta_{\rho_t})M_\psi} C_{14}^{-1} C(t).$$

Hence $\sqrt{2e^{\beta_0 t_j}}\, v_{k,\varepsilon,t_j}\to 0$ in L^2 as $j\to+\infty$. Since $\mathrm{D}'' F_{k,\varepsilon,t}=\sqrt{2e^{\beta_0 t}}\, v_{k,\varepsilon,t}$ on X_k, we get $\mathrm{D}''F_{k,\varepsilon}=0$ on X_k. Then $F_{k,\varepsilon}$ is a holomorphic section of $K_X\otimes L$ on X_k. In Step 5, we will prove that $F_{k,\varepsilon}=f$ on $X_k\cap Y^0$ by solving $\bar\partial$ locally.

Step 5: Solving $\bar\partial$ locally with L^2 estimates and the end of the proof for the line bundle L.

For any $x\in X_k\cap Y$, let Ω_x be as in Step 1. Let

$$\widehat{\Omega}_x \Subset (X_k\cap\Omega_x)$$

be a coordinate ball with center x. Since the bundle L is trivial on Ω_x, $u_{k,\varepsilon,t}$ and $v_{k,\varepsilon,t}$ can be regarded as forms on Ω_x with values in $\mathbb{C}$ and the metric h_0 of L on Ω_x can be regarded as a positive smooth function.

It is easy to see that $C(t) \leq C_{18}$ for some positive number C_{18} independent of t when t is small enough. Then it follows from (4-34), (4-41) and (4-32) that

$$\int_{\widehat{\Omega}_x} |v_{k,\varepsilon,t}|^2 e^{-\psi}\, d\lambda \leq C_{19} C_{18}$$

for some positive number C_{19} independent of t when t is small enough.

Since $\bar{\partial} v_{k,\varepsilon,t} = 0$ on $\widehat{\Omega}_x$ by (4-33), applying Lemma 2.5 to the $(n, 1)$-form

$$\sqrt{2e^{\beta_0 t}}\, v_{k,\varepsilon,t} \in L^2_{(n,1)}(\widehat{\Omega}_x,\ e^{-\psi}),$$

we get an $(n, 0)$-form $s_{k,\varepsilon,t} \in L^2_{(n,0)}(\widehat{\Omega}_x,\ e^{-\psi})$ such that

$$\bar{\partial} s_{k,\varepsilon,t} = \sqrt{2e^{\beta_0 t}}\, v_{k,\varepsilon,t}$$

on $\widehat{\Omega}_x$ and

$$\int_{\widehat{\Omega}_x} |s_{k,\varepsilon,t}|^2 e^{-\psi}\, d\lambda \leq C_{20} \int_{\widehat{\Omega}_x} |\sqrt{2e^{\beta_0 t}}\, v_{k,\varepsilon,t}|^2 e^{-\psi}\, d\lambda \leq 2C_{20} C_{19} C_{18} e^{\beta_0 t} \tag{4-42}$$

for some positive number C_{20} independent of t. Hence

$$\int_{\widehat{\Omega}_x} |s_{k,\varepsilon,t}|^2\, d\lambda \leq C_{21} e^{\beta_0 t} \tag{4-43}$$

for some positive number C_{21} independent of t.

Now define

$$G_{k,\varepsilon,t} = -u_{k,\varepsilon,t} - s_{k,\varepsilon,t} + \theta\left(\frac{e^t}{e^{\psi} + e^t}\right) \tilde{f}_t$$

on $\widehat{\Omega}_x$. Then $G_{k,\varepsilon,t} = F_{k,\varepsilon,t} - s_{k,\varepsilon,t}$ and $\bar{\partial} G_{k,\varepsilon,t} = 0$. Hence $G_{k,\varepsilon,t}$ is holomorphic in $\widehat{\Omega}_x$. Therefore, $u_{k,\varepsilon,t} + s_{k,\varepsilon,t}$ is smooth in $\widehat{\Omega}_x$. Furthermore, we get from (4-39) and (4-43) that

$$\int_{\widehat{\Omega}_x} |G_{k,\varepsilon,t}|^2\, d\lambda \leq 2 \int_{\widehat{\Omega}_x} |F_{k,\varepsilon,t}|^2\, d\lambda + 2 \int_{\widehat{\Omega}_x} |s_{k,\varepsilon,t}|^2\, d\lambda \leq C_{22} \tag{4-44}$$

for some positive number C_{22} independent of t when t is small enough.

We get from (4-32) and (4-34) that

$$\int_{\widehat{\Omega}_x} \frac{|u_{k,\varepsilon,t}|^2 e^{-\psi}}{R(\sigma_t)}\, d\lambda \leq C_{23} C(t) \leq C_{23} C_{18}$$

for some positive number C_{23} independent of t when t is small enough. Since $R(\sigma_t) \leq R(t - \varepsilon)$ on $\widehat{\Omega}_x$ when t is small enough, we have that

$$\int_{\widehat{\Omega}_x} |u_{k,\varepsilon,t}|^2 e^{-\psi}\, d\lambda \leq C_{23} C_{18} R(t - \varepsilon).$$

Therefore, combining the last inequality and (4-42), we obtain that

$$\int_{\widehat{\Omega}_x} |u_{k,\varepsilon,t}+s_{k,\varepsilon,t}|^2 e^{-\psi}\,d\lambda \le 2C_{23}C_{18}R(t-\varepsilon)+4C_{20}C_{19}C_{18}e^{\beta_0 t}.$$

Then the nonintegrability of $e^{-\psi}$ along $\widehat{\Omega}_x \cap Y$ and the smoothness of $u_{k,\varepsilon,t}+s_{k,\varepsilon,t}$ in $\widehat{\Omega}_x$ show that $u_{k,\varepsilon,t}+s_{k,\varepsilon,t}=0$ on $\widehat{\Omega}_x \cap Y$ for any t. Hence $G_{k,\varepsilon,t}=f$ on $\widehat{\Omega}_x \cap Y^0$ for any t.

Since $s_{k,\varepsilon,t_j} \to 0$ in $L^2_{(n,0)}(\widehat{\Omega}_x)$ by (4-43) and $F_{k,\varepsilon,t_j} \rightharpoonup F_{k,\varepsilon}$ weakly in $L^2_{(n,0)}(\widehat{\Omega}_x)$ as $j \to +\infty$, $G_{k,\varepsilon,t_j} \rightharpoonup F_{k,\varepsilon}$ weakly in $L^2_{(n,0)}(\widehat{\Omega}_x)$ as $j \to +\infty$. Hence it follows from (4-44) and routine arguments applying Montel's theorem that a subsequence of $\{G_{k,\varepsilon,t_j}\}_{j=1}^{+\infty}$ converges to $F_{k,\varepsilon}$ uniformly on compact subsets of $\widehat{\Omega}_x$. Then $F_{k,\varepsilon}=f$ on $\widehat{\Omega}_x \cap Y^0$ and thereby on $X_k \cap Y^0$.

Since the positive continuous function R is decreasing near $-\infty$, $e^t R(t)$ is bounded above near $-\infty$ and ϕ is locally bounded above, applying Montel's theorem and extracting weak limits of $\{F_{k,\varepsilon}\}_{k,\varepsilon}$, first as $\varepsilon \to 0$, and then as $k \to +\infty$, we get from (4-40) a holomorphic section F on X with values in $K_X \otimes L$ such that $F=f$ on Y^0 and

$$\int_X \frac{|F|^2_{\omega,h}}{e^{\psi}R(\psi)}\,dV_{X,\omega} \le \left(\frac{\alpha_1}{R(\alpha_0)}+C_R\right)\int_{Y^0} |f|^2_{\omega,h}\,dV_{X,\omega}[\psi].$$

Theorem 1.8 is thus proved for the line bundle L.

Step 6: The proof for the vector bundle E.

The proof for E is similar but simpler. We only point out the main modifications by examining the proof for L.

In Step 1, we don't need to construct a family of special smooth extensions $\tilde{f}_t$ of f since h_E is smooth. Hence the strong openness property and the key propositions are not needed. Delete Parts II and III in Step 1 and replace the family of sections $\tilde{f}_{i,t}$ with a fixed local holomorphic extension $\tilde{f}_i$. Then $\tilde{f}_t$ becomes a fixed smooth extension $\tilde{f}=\sum_{i=1}^N \xi_i \tilde{f}_i$. Then it is easy to see that (4-9)–(4-12) hold for $\tilde{f}_{i,t}=\tilde{f}_i$, $\tilde{f}_t=\tilde{f}$ and $\beta_1=0$.

Step 2 is not needed since h_E is already smooth.

In Step 3, the negative term will not appear on the right-hand side of (4-20) since $\delta_\rho=0$.

In Step 4, it is easy to prove the estimate (4-26) for $I_{1,t}$ by the modified (4-9). It is also not hard to prove the estimate (4-29) for $I_{2,t}$ by the modified (4-10)–(4-12). Equation (4-36) can be easily obtained since h_E is smooth.

Step 5 for E is almost the same and Theorem 1.8 is thus proved for the vector bundle E.

References

[Berndtsson 1996] B. Berndtsson, "The extension theorem of Ohsawa–Takegoshi and the theorem of Donnelly-Fefferman", *Ann. Inst. Fourier* (*Grenoble*) **46**:4 (1996), 1083–1094. MR Zbl

[Berndtsson and Lempert 2016] B. Berndtsson and L. Lempert, "A proof of the Ohsawa–Takegoshi theorem with sharp estimates", *J. Math. Soc. Japan* **68**:4 (2016), 1461–1472. MR Zbl

[Berndtsson and Păun 2008] B. Berndtsson and M. Păun, "Bergman kernels and the pseudoeffectivity of relative canonical bundles", *Duke Math. J.* **145**:2 (2008), 341–378. MR Zbl

[Bierstone and Milman 1991] E. Bierstone and P. D. Milman, "A simple constructive proof of canonical resolution of singularities", pp. 11–30 in *Effective methods in algebraic geometry* (Castiglioncello, 1990), edited by T. Mora and C. Traverso, Progr. Math. **94**, Birkhäuser, Boston, 1991. MR Zbl

[Błocki 2013] Z. Błocki, "Suita conjecture and the Ohsawa–Takegoshi extension theorem", *Invent. Math.* **193**:1 (2013), 149–158. MR Zbl

[Cao 2017] J. Cao, "Ohsawa–Takegoshi extension theorem for compact Kähler manifolds and applications", pp. 19–38 in *Complex and symplectic geometry*, edited by D. Angella et al., Springer INdAM Ser. **21**, Springer, 2017. MR Zbl

[Demailly 1982] J.-P. Demailly, "Estimations L^2 pour l'opérateur $\bar{\partial}$ d'un fibré vectoriel holomorphe semi-positif au-dessus d'une variété kählérienne complète", *Ann. Sci. École Norm. Sup.* (4) **15**:3 (1982), 457–511. MR Zbl

[Demailly 1994] J.-P. Demailly, "Regularization of closed positive currents of type (1, 1) by the flow of a Chern connection", pp. 105–126 in *Contributions to complex analysis and analytic geometry*, edited by H. Skoda and J.-M. Trépreau, Aspects Math. **26**, Friedr. Vieweg, Braunschweig, 1994. MR Zbl

[Demailly 2000] J.-P. Demailly, "On the Ohsawa–Takegoshi-Manivel L^2 extension theorem", pp. 47–82 in *Complex analysis and geometry* (Paris, 1997), edited by P. Dolbeault et al., Progr. Math. **188**, Birkhäuser, Basel, 2000. MR Zbl

[Demailly 2012a] J.-P. Demailly, *Analytic methods in algebraic geometry*, Surveys of Modern Mathematics **1**, International Press, Somerville, MA, 2012. MR Zbl

[Demailly 2012b] J.-P. Demailly, *Complex analytic and differential geometry*, 2012.

[Demailly 2016] J.-P. Demailly, "Extension of holomorphic functions defined on non reduced analytic subvarieties", pp. 191–222 in *The legacy of Bernhard Riemann after one hundred and fifty years, Vol. I*, edited by L. Ji et al., Adv. Lect. Math. (ALM) **35**, International Press, Somerville, MA, 2016. MR Zbl

[Guan and Zhou 2012] Q. Guan and X. Zhou, "Optimal constant problem in the L^2 extension theorem", *C. R. Math. Acad. Sci. Paris* **350**:15-16 (2012), 753–756. MR Zbl

[Guan and Zhou 2015a] Q. Guan and X. Zhou, "Optimal constant in an L^2 extension problem and a proof of a conjecture of Ohsawa", *Sci. China Math.* **58**:1 (2015), 35–59. MR Zbl

[Guan and Zhou 2015b] Q. Guan and X. Zhou, "A proof of Demailly's strong openness conjecture", *Ann. of Math.* (2) **182**:2 (2015), 605–616. MR Zbl

[Guan and Zhou 2015c] Q. Guan and X. Zhou, "A solution of an L^2 extension problem with an optimal estimate and applications", *Ann. of Math.* (2) **181**:3 (2015), 1139–1208. MR Zbl

[Guan et al. 2011] Q. Guan, X. Zhou, and L. Zhu, "On the Ohsawa–Takegoshi L^2 extension theorem and the twisted Bochner–Kodaira identity", *C. R. Math. Acad. Sci. Paris* **349**:13-14 (2011), 797–800. MR Zbl

[Hironaka 1964] H. Hironaka, "Resolution of singularities of an algebraic variety over a field of characteristic zero: I, II", *Ann. of Math.* (2) **79** (1964), 109–203, 205–326. MR Zbl

[Hörmander 1990] L. Hörmander, *An introduction to complex analysis in several variables*, 3rd ed., North-Holland Mathematical Library **7**, North-Holland, Amsterdam, 1990. MR

[Manivel 1993] L. Manivel, "Un théorème de prolongement L^2 de sections holomorphes d'un fibré hermitien", *Math. Z.* **212**:1 (1993), 107–122. MR Zbl

[McNeal and Varolin 2007] J. D. McNeal and D. Varolin, "Analytic inversion of adjunction: L^2 extension theorems with gain", *Ann. Inst. Fourier* (*Grenoble*) **57**:3 (2007), 703–718. MR Zbl

[Ohsawa 1995] T. Ohsawa, "On the extension of L^2 holomorphic functions, III: Negligible weights", *Math. Z.* **219**:2 (1995), 215–225. MR Zbl

[Ohsawa 2001] T. Ohsawa, "On the extension of L^2 holomorphic functions, V: Effects of generalization", *Nagoya Math. J.* **161** (2001), 1–21. MR Zbl

[Ohsawa 2015] T. Ohsawa, *L^2 approaches in several complex variables*, Springer, 2015. MR Zbl

[Ohsawa and Takegoshi 1987] T. Ohsawa and K. Takegoshi, "On the extension of L^2 holomorphic functions", *Math. Z.* **195**:2 (1987), 197–204. MR Zbl

[Siu 1996] Y.-T. Siu, "The Fujita conjecture and the extension theorem of Ohsawa–Takegoshi", pp. 577–592 in *Geometric complex analysis* (Hayama, 1995), edited by J. Noguchi et al., World Sci. Publ., River Edge, NJ, 1996. MR Zbl

[Siu 1998] Y.-T. Siu, "Invariance of plurigenera", *Invent. Math.* **134**:3 (1998), 661–673. MR Zbl

[Siu 2002] Y.-T. Siu, "Extension of twisted pluricanonical sections with plurisubharmonic weight and invariance of semipositively twisted plurigenera for manifolds not necessarily of general type", pp. 223–277 in *Complex geometry* (Göttingen, 2000), edited by I. Bauer et al., Springer, 2002. MR Zbl

[Yi 2012] L. Yi, *Théorèmes d'extension et métriques de Kähler-Einstein Généralisées*, Ph.D. Thesis, Université de Lorraine, Nancy, 2012, available at http://www.theses.fr/2012LORR0151.

[Zhou 2015] X. Zhou, "A survey on L^2 extension problem", pp. 291–309 in *Complex geometry and dynamics*, edited by J. E. Fornæss et al., Abel Symp. **10**, Springer, 2015. MR Zbl

[Zhou and Zhu 2018] X. Zhou and L. Zhu, "An optimal L^2 extension theorem on weakly pseudoconvex Kähler manifolds", *J. Differential Geom.* **110**:1 (2018), 135–186. MR Zbl

[Zhu et al. 2012] L. Zhu, Q. Guan, and X. Zhou, "On the Ohsawa–Takegoshi L^2 extension theorem and the Bochner–Kodaira identity with non-smooth twist factor", *J. Math. Pures Appl.* (9) **97**:6 (2012), 579–601. MR Zbl

Received September 27, 2019. Revised June 5, 2020.

XIANGYU ZHOU
INSTITUTE OF MATHEMATICS, AMSS
AND
HUA LOO-KENG KEY LABORATORY OF MATHEMATICS
CHINESE ACADEMY OF SCIENCES
BEIJING
CHINA
xyzhou@math.ac.cn

LANGFENG ZHU
SCHOOL OF MATHEMATICS AND STATISTICS
WUHAN UNIVERSITY
WUHAN
CHINA
zhulangfeng@amss.ac.cn

CONTENTS

Volume 309, no. 1 and no. 2

Guidelines for Authors

Authors may submit articles at msp.org/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 309 No. 2 December 2020