

# On Vojta's proof of the Mordell conjecture

Xinyi Yuan

December 4, 2025

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Vojta's inequality</b>	<b>2</b>
2.1	Vojta's inequality . . . . .	2
2.2	Mumford's equality . . . . .	4
2.3	A precise statement . . . . .	7
2.4	The Arakelov-geometric setting . . . . .	8
2.5	Plan of proof . . . . .	11
<b>3</b>	<b>Step 1: existence of small section</b>	<b>12</b>
3.1	Hilbert–Samuel type of estimates . . . . .	12
3.2	Estimate of the arithmetic Euler characteristic . . . . .	13
<b>4</b>	<b>Step 2: lower bound of the height</b>	<b>15</b>
4.1	The hermitian line bundle by restriction . . . . .	16
4.2	The section of the line bundle . . . . .	17
4.3	The degree of the line bundle . . . . .	18
4.4	The uniform choice . . . . .	20
<b>5</b>	<b>Step 3: upper bound of the index</b>	<b>21</b>
5.1	First reduction: remove fibers . . . . .	22
5.2	Second reduction: trivialize weight . . . . .	23
5.3	Proof of the essential case . . . . .	25
<b>6</b>	<b>Completion of the proof</b>	<b>27</b>

## 1 Introduction

The celebrated Mordell conjecture, proved by Faltings, is the following theorem.

**Theorem 1.1** (Mordell conjecture). *Let  $C$  be a smooth, projective, and geometrically connected curve of genus  $g > 1$  over a number field  $K$ . Then  $C(K)$  is finite.*

The conjecture was raised by Mordell when he proved the Mordell–Weil theorem for elliptic curves over  $\mathbb{Q}$  in 1922. Faltings’ proof of the conjecture in [Fal83] signified a milestone in the history of Diophantine geometry. By an analogue to the Thue–Siegel–Roth theorem in Diophantine approximation, Vojta [Voj91] gave a new proof of the Mordell conjecture in 1991. In terms of  $p$ -adic Hodge theory, Lawrence–Venkatesh [LV20] gave a third proof of the Mordell conjecture.

Each of these three proofs has its own valuable feature. For example, Vojta’s proof was extended by Faltings [Fal91] to prove the Bombieri–Lang conjecture for subvarieties of abelian varieties. Moreover, Vojta essentially proved a deep inequality concerning distribution of rational points of large heights. Combining with the uniform Bogomolov conjecture recently proved by Dimitrov–Gao–Habegger [DGH21] and Kühne [Kuh21], Vojta’s inequality implies a uniform bound on the number of rational points in terms of the genus and the Mordell–Weil rank.

Vojta’s proof depends heavily on high-dimensional Arakelov geometry developed by Arakelov [Ara74], Deligne [Del87], and Gillet–Soulé [GS90, GS92]. Bombieri [Bom90] (based on [Fal91]) wrote a relatively elementary variant of Vojta’s proof, which replaced the use of Arakelov geometry by those of Siegel’s lemma and classical height theory, at the cost of involving lots of less important extra objects and estimates. We refer to the textbooks [HS00, BG06, IKM22] for Bombieri’s version of the proof.

The goal of this paper is to introduce a new variant of Vojta’s original proof. The most technical part of Vojta’s original proof is the application of the arithmetic Riemann–Roch theorem of Gillet–Soulé [GS92] to construct a small section, which also involved an estimate of some analytic torsion. Our treatment here replaces the application of the arithmetic Riemann–Roch theorem by a quick application of the arithmetic Siu inequality of Yuan [Yua08]. So our treatment is still based on Arakelov geometry.

Note that the arithmetic Siu inequality in [Yua08] is based on Gillet–Soulé’s arithmetic Hilbert–Samuel formula and Zhang’s arithmetic Nakai–Moishezon criterion in [Zha92, Zha95a]. The arithmetic Hilbert–Samuel formula is originally proved by Gillet–Soulé [GS92] as a consequence of their arithmetic Riemann–Roch theorem, but it also has a direct proof by Abbes–Bouche [AB95]. In this sense, our new proof avoids using the arithmetic Riemann–Roch theorem.

The proofs in [Voj91, Bom90] emphasize the role of the bounds of the index, which is analogous to the proof of the Thue–Siegel–Roth theorem. However, in our exposition, we will emphasize more on the role of the bounds of the height and its relation with small sections and base points, which is closer to the spirit of Arakelov geometry.

Most parts of our treatment are essentially duplicated from Vojta’s original proof, but we still include a complete exposition in this paper for convenience of readers.

## Acknowledgments

The author would like to thank Zheng Xiao and Chunhui Liu, who taught a course on Roth's theorem and Vojta's inequality in Peking University in the fall of 2022. Motivated by their lectures, the author obtained the variant of Vojta's proof in the current paper.

The author would also like to thank the support of the China–Russia Mathematics Center. The author is supported by grants NO. 12250004 and NO. 12321001 from the National Science Foundation of China, and by the Xplorer Prize from the New Cornerstone Science Foundation.

## 2 Vojta's inequality

The goal of this section is to introduce Vojta's inequality, Mumford's equality, and another version of Vojta's inequality using Arakelov geometry. In the end, we sketch the plan of our proof of Vojta's inequality.

### 2.1 Vojta's inequality

Let  $K$  be a number field. Let  $C$  be a smooth, projective, and geometrically integral curve of genus  $g > 1$  over  $K$ . Denote by  $J$  the Jacobian variety of  $C$  over  $K$ . Denote by  $[m] : J \rightarrow J$  the morphism given by multiplication by  $m \in \mathbb{Z}$ .

First we recall some basic facts on the Néron–Tate heights on  $C$  and  $J$ . Since

$$[2g - 2] : \text{Pic}^0(C_{\overline{K}}) \longrightarrow \text{Pic}^0(C_{\overline{K}})$$

is surjective, there is a line bundle  $\alpha$  on  $C_{\overline{K}}$  such that  $(2g - 2)\alpha$  is isomorphic to the canonical sheaf  $\omega_C$ . Replacing  $K$  by a finite extension if necessary, we can assume that  $\alpha$  is actually a line bundle on  $C$ .

Consider the Abel–Jacobi embedding

$$i_\alpha : C \longrightarrow J, \quad x \longmapsto (x) - \alpha.$$

Recall that the theta divisor on  $J$  is given by

$$\theta_\alpha = \underbrace{i_\alpha(C) + \cdots + i_\alpha(C)}_{g-1 \text{ copies}}.$$

It is well-known that  $\theta_\alpha$  is ample and gives a principal polarization of  $J$ . By [Ser89, p. 74, eq. (1)],  $\theta_\alpha$  is symmetric in the sense that  $[-1]\theta_\alpha$  is linearly equivalent to  $\theta_\alpha$ .

Denote  $J(\overline{K})_{\mathbb{R}} = J(\overline{K}) \otimes_{\mathbb{Z}} \mathbb{R}$ . It is well-known that the Néron–Tate height

$$\hat{h}_{\theta_\alpha} : J(\overline{K}) \longrightarrow \mathbb{R}$$

is quadratic and induces a positive definite quadratic form

$$\hat{h}_{\theta_\alpha} : J(\overline{K})_{\mathbb{R}} \longrightarrow \mathbb{R}.$$

See [Ser89, §3.3, §3.8] for example, and we normalize Weil heights over number fields as in [Ser89, §2.2]. Denote the Néron–Tate height pairing

$$\langle P, Q \rangle : J(\overline{K}) \times J(\overline{K}) \longrightarrow \mathbb{R}$$

by

$$\langle P, Q \rangle = \frac{1}{2}(\hat{h}_{\theta_\alpha}(P + Q) - \hat{h}_{\theta_\alpha}(P) - \hat{h}_{\theta_\alpha}(Q)), \quad P, Q \in J(\overline{K}).$$

To do Euclidean geometry, we define

$$|P| := \sqrt{\hat{h}_{\theta_\alpha}(P)}, \quad P \in J(\overline{K}).$$

We have  $|P| = \sqrt{\langle P, P \rangle}$  by definition. These definitions extend to  $P, Q \in J(\overline{K})_{\mathbb{R}}$  naturally. We will apply  $|\cdot|$  and  $\langle \cdot, \cdot \rangle$  to  $C(\overline{K})$  via the embedding  $i_\alpha : C \rightarrow J$ .

Now we are ready to introduce Vojta’s inequality, which is essentially due to Vojta [Voj91].

**Theorem 2.1** (Vojta’s inequality). *For any real number  $\mu > \frac{1}{\sqrt{g}}$ , there exist positive constants  $A_1, A_2$  such that for all points  $P_1, P_2 \in C(\overline{K})$  satisfying*

$$|P_1| > A_1, \quad |P_2| > A_2|P_1|,$$

*we have*

$$\langle P_1, P_2 \rangle < \mu \cdot |P_1| \cdot |P_2|.$$

Combined with the Mordell–Weil theorem for the Jacobian variety, Vojta’s inequality implies the Mordell conjecture without much effort. In fact, the finite-dimensional real vector space  $V = J(K)_{\mathbb{R}}$ , endowed with the inner product  $\langle \cdot, \cdot \rangle$ , is isometric to a Euclidean space  $\mathbb{R}^r$ . For any two vectors  $x, y \in V$ , denote by  $\angle(x, y)$  the angle between these two vectors, which is set to be 0 if  $x = 0$  or  $y = 0$ . Take a real number  $\mu$  in the open interval  $(1/\sqrt{g}, 1)$ , and denote  $\theta = \arccos \mu$ . The conclusion

$$\langle P_1, P_2 \rangle < \mu \cdot |P_1| \cdot |P_2|$$

implies

$$\cos \angle(P_1, P_2) < \mu, \quad \angle(P_1, P_2) > \theta.$$

For every nonzero vector  $x \in V$ , consider the cone

$$\text{cone}(x) := \{y \in V : \angle(x, y) \leq \theta/2\}.$$

Then  $V$  is covered by finitely many cones  $\text{cone}(x_1), \dots, \text{cone}(x_n)$ , as a consequence of the compactness of the unit sphere of  $V$ . For every  $i$ , denote by  $\Sigma_i$  the set of points of  $C(K)$  whose images in  $V$  lie in  $\text{cone}(x_i)$ . It suffices to prove that every  $\Sigma_i$  is finite. Note that any two points of  $\Sigma_i$  have an angle at most  $\theta$ . Apply Theorem 2.1 to this situation. If every  $P \in \Sigma_i$  satisfies  $|P| \leq A_1$ , then  $\Sigma_i$  is finite by the Northcott property. If there exists  $P \in \Sigma_i$  with  $|P| > A_1$ , then Vojta’s inequality implies  $|P'| \leq A_2|P|$  for any  $P' \in \Sigma_i$ , and thus  $\Sigma_i$  is still finite by the Northcott property.

## 2.2 Mumford's equality

To prove Vojta's inequality, we first apply Mumford's equality to convert the inequality to a lower bound of Weil heights of some relevant points.

Let  $K$  be a number field. Let  $C$  be a smooth projective and geometrically connected curve of genus  $g > 1$  over  $K$ . As above, let  $\alpha$  be a line bundle on  $C$  with  $(2g - 2)\alpha \cong \omega_{C/K}$ .

Let  $X = C^2$ , and let  $\Delta$  be the diagonal of  $C^2$ . Denote by  $p_1, p_2 : X \rightarrow C$  the two projections. Define the *Vojta line bundle* to be

$$L_0 := (\delta_1 - 1)p_1^*\alpha + (\delta_2 - 1)p_2^*\alpha + \mathcal{O}(\Delta),$$

where  $\delta_1, \delta_2$  are positive rational numbers. The Vojta line bundle  $L_0$  is a  $\mathbb{Q}$ -line bundle on  $X$ , i.e. an element of  $\text{Pic}(X)_{\mathbb{Q}} = \text{Pic}(X) \otimes_{\mathbb{Z}} \mathbb{Q}$ . Here we write tensor products of line bundles additively.

We will see that it is often more convenient to write

$$L_0 = \delta_1 p_1^*\alpha + \delta_2 p_2^*\alpha + \mathcal{O}(\Delta)^\circ,$$

where

$$\mathcal{O}(\Delta)^\circ := \mathcal{O}(\Delta) - p_1^*\alpha - p_2^*\alpha.$$

is the the modified diagonal. For example, it gives convenience in computing intersection numbers by

$$\mathcal{O}(\Delta)^\circ \cdot p_1^*\alpha = \mathcal{O}(\Delta)^\circ \cdot p_2^*\alpha = 0.$$

Let  $h_{p_1^*\alpha}, h_{p_2^*\alpha}, h_\Delta$  be Weil heights on  $X(\overline{K})$  with respect to the line bundles  $p_1^*\alpha, p_2^*\alpha, \mathcal{O}(\Delta)$ . Take a Weil height function

$$h_{L_0} : X(\overline{K}) \longrightarrow \mathbb{R}$$

with respect to the line bundle  $L_0$  on  $X$  by

$$h_{L_0} = (\delta_1 - 1)h_{p_1^*\alpha} + (\delta_2 - 1)h_{p_2^*\alpha} + h_\Delta.$$

Note that  $h_{p_1^*\alpha}, h_{p_2^*\alpha}$ , and  $h_\Delta$  are only determined up to constants, but we choose them once for all.

**Theorem 2.2** (Mumford's equality). *For any  $P_1, P_2 \in C(\overline{K})$ ,*

$$h_{L_0}(P_1, P_2) = \frac{1}{g}\delta_1|P_1|^2 + \frac{1}{g}\delta_2|P_2|^2 - 2\langle P_1, P_2 \rangle + O(\delta_1 + \delta_2 + 1).$$

*Here the coefficient in the error term has a bound independent of  $(P_1, P_2, \delta_1, \delta_2)$ .*

*Proof.* Let us start with the pull-back formulas. Denote by

$$j_\alpha = (i_\alpha, i_\alpha) : C \times C \longrightarrow J \times J$$

the natural morphism. Denote by  $p'_1, p'_2 : J \times J \rightarrow J$  the projections. Denote by  $P$  the Poincaré line bundle on  $J \times J$ . Then the following are true:

(1)  $i_\alpha^* \mathcal{O}(\theta_\alpha) = g\alpha$  in  $\text{Pic}(C)$ .

(2)  $j_\alpha^* P = \mathcal{O}(\Delta)^\circ$  in  $\text{Pic}(C \times C)$ .

The first identity is [Ser89, p. 75, eq. (2)], the second identity is [Ser89, p. 76, eq. (4)]. All these require the condition  $(2g - 2)\alpha = \omega_{C/K}$ .

Apply the functoriality of Weil heights for pull-backs of line bundles as in [Ser89, p. 23] to the above identities. We have for  $k = 1, 2$ ,

$$h_{p_k^* \alpha}(P_1, P_2) = h_\alpha(P_k) + O(1) = \frac{1}{g} h_{\theta_\alpha}(i_\alpha(P_k)) + O(1) = \frac{1}{g} |P_k|^2 + O(1),$$

and

$$h_{\mathcal{O}(\Delta)^\circ}(P_1, P_2) = h_P(i_\alpha(P_1), i_\alpha(P_2)) + O(1) = 2\langle P_1, P_2 \rangle + O(1).$$

Here the last inequality follows from [Ser89, p. 39, Theorem], and the factor 2 comes from different normalization of the pairing.  $\square$

*Remark 2.3.* We can have a precise equality of the above theorem without the error  $O(1)$  in terms of admissible pairings of Zhang [Zha93] (or the reformulation of [Yua21, Appendix A] in terms of adelic line bundles of Zhang [Zha95b]).

In the case  $\delta_1 = \delta_2 = 1$ , Mumford's equality gives

$$h_\Delta(P_1, P_2) = \frac{1}{g} |P_1|^2 + \frac{1}{g} |P_2|^2 - 2\langle P_1, P_2 \rangle + O(1).$$

As  $\Delta$  is effective, its Weil height function satisfies

$$h_\Delta(P_1, P_2) \geq O(1), \quad P_1 \neq P_2.$$

This implies an inequality

$$\frac{1}{g} |P_1|^2 + \frac{1}{g} |P_2|^2 - 2\langle P_1, P_2 \rangle \geq O(1), \quad P_1 \neq P_2.$$

This inequality is due to Mumford [Mum65, §3, Cor. 1]. It gives a non-trivial bound on the angle between  $P_1$  and  $P_2$  if  $|P_1|$  and  $|P_2|$  are large with a quotient close to 1. See also [Ser89, §5.7].

To illustrate Vojta's idea to use Mumford's equality to prove Vojta's inequality, let us assume some situations which are too good to be true but might eventually be modified to be true. Let  $\delta_1, \delta_2$  be positive rational numbers. A basic application of the Riemann–Roch theorem on  $X$  proves that Vojta's line bundle

$$L_0 = \delta_1 p_1^* \alpha + \delta_2 p_2^* \alpha + \mathcal{O}(\Delta)^\circ$$

has volume

$$\text{vol}(L_0) \geq L_0^2 = 2(\delta_1 \delta_2 - g).$$

In particular,  $L_0$  is big if  $\delta_1 \delta_2 > g$ . See Theorem 3.2 for this result. We will assume that  $\delta_1 \delta_2 > g$  throughout this paper. It follows that  $h_{L_0}(P)$  is bounded

below for  $P = (P_1, P_2) \in X(\overline{K})$  lying outside of the base locus of positive multiples of  $L_0$ . See [Ser89, §2.10, Theorem]. By Mumford's equality, this gives the relation

$$\frac{1}{g}\delta_1|P_1|^2 + \frac{1}{g}\delta_2|P_2|^2 - 2\langle P_1, P_2 \rangle \geq O(\delta_1 + \delta_2 + 1)$$

for  $P \in X(\overline{K})$  outside the base locus.

For the sake of demonstrating ideas, we ignore the base locus and uniformity of the error, so we simply assume that under the condition  $\delta_1\delta_2 > g$ , we have

$$\frac{1}{g}\delta_1|P_1|^2 + \frac{1}{g}\delta_2|P_2|^2 - 2\langle P_1, P_2 \rangle \geq 0$$

for all  $P_1, P_2 \in C(\overline{K})$ . For fixed  $(P_1, P_2)$ , we want to choose  $\delta_1, \delta_2$  so that the inequality is as optimal as possible. We also fix  $\lambda = \delta_1\delta_2 - g > 0$ . The most optimal case amounts to taking the minimal value of the left-hand side of the inequality under the constraint  $\delta_1\delta_2 = g + \lambda$ . By the inequality between the arithmetic mean and the geometric mean, the left-hand side of the inequality is minimal if

$$\delta_1 = \sqrt{g + \lambda} \frac{|P_2|}{|P_1|}, \quad \delta_2 = \sqrt{g + \lambda} \frac{|P_1|}{|P_2|}.$$

In this case, the inequality becomes

$$\frac{\sqrt{g + \lambda}}{g} |P_1| \cdot |P_2| \geq \langle P_1, P_2 \rangle.$$

This would imply Vojta's inequality.

However, this ideal situation is not practical by many obstacles. First, the point  $(P_1, P_2)$  might lie in the base locus of the line bundle  $L_0$ . Second, we need uniformity of the error term when varying  $(P_1, P_2)$  and  $(\delta_1, \delta_2)$ . In the following, we write a precise statement of the inequality, and then introduce more ideas to overcome these two issues.

### 2.3 A precise statement

Let  $\delta_1, \delta_2$  be positive rational numbers. As above, denote

$$L_0 = (\delta_1 - 1)p_1^*\alpha + (\delta_2 - 1)p_2^*\alpha + \mathcal{O}(\Delta).$$

This gives a Weil height

$$h_{L_0} := (\delta_1 - 1)h_{p_1^*\alpha} + (\delta_2 - 1)h_{p_2^*\alpha} + h_\Delta$$

on  $X = C^2$  as above. The estimate we can achieve is as follows.

**Theorem 2.4.** *Let  $\delta_1, \delta_2$  be positive rational numbers satisfying*

$$\delta_1 > 2g\delta_2, \quad g < \delta_1\delta_2 < g + \frac{1}{4}.$$

Let  $P_1, P_2 \in C(\overline{K})$  be points. Then

$$h_{L_0}(P_1, P_2) \geq -2c(\delta_1, \delta_2) \cdot (\delta_1|P_1|^2 + \delta_2|P_2|^2) + O(\delta_1/(\delta_1\delta_2 - g)).$$

Here the coefficient in the error term is independent of  $(P_1, P_2, \delta_1, \delta_2)$ , and

$$c(\delta_1, \delta_2) = \sqrt{\frac{2}{g}(\delta_1\delta_2 - g) + (2g - 1)\frac{\delta_2}{\delta_1}}.$$

By Mumford's equality, we can prove that Theorem 2.4 implies Theorem 2.1. In fact, by Theorem 2.2, Theorem 2.4 gives

$$\left(\frac{1}{g} + 2c(\delta_1, \delta_2)\right) (\delta_1|P_1|^2 + \delta_2|P_2|^2) - 2\langle P_1, P_2 \rangle + \frac{c_0\delta_1}{\delta_1\delta_2 - g} \geq 0.$$

Here  $c_0$  is a non-negative constant independent of  $(P_1, P_2, \delta_1, \delta_2)$ . By approximation, the inequality also holds for real numbers  $\delta_1, \delta_2$  satisfying the condition of Theorem 2.4.

Let  $\lambda$  be a real number with  $0 < \lambda < 1/4$ . Set

$$\delta_1 = \sqrt{g + \lambda} \frac{|P_2|}{|P_1|}, \quad \delta_2 = \sqrt{g + \lambda} \frac{|P_1|}{|P_2|}.$$

Then  $\delta_1\delta_2 - g = \lambda > 0$ . The requirement  $\delta_1 > 2g\delta_2$  transfers to  $|P_2| > \sqrt{2g}|P_1|$ , which we assume. We further assume  $|P_1| > 0$ . The inequality becomes

$$2\sqrt{g + \lambda} \left(\frac{1}{g} + 2c(\delta_1, \delta_2)\right) |P_1| \cdot |P_2| + \sqrt{g + \lambda} \frac{c_0|P_2|}{\lambda|P_1|} \geq 2\langle P_1, P_2 \rangle,$$

or equivalently

$$\frac{\langle P_1, P_2 \rangle}{|P_1| \cdot |P_2|} \leq \sqrt{g + \lambda} \left(\frac{1}{g} + 2c(\delta_1, \delta_2) + \frac{c_0}{2\lambda|P_1|^2}\right).$$

Here

$$c(\delta_1, \delta_2) = \sqrt{\frac{2}{g}\lambda + (2g - 1)\frac{|P_1|^2}{|P_2|^2}}.$$

For any fixed  $\lambda > 0$ , if  $|P_1|^2 > \lambda^{-2}c_0$  and  $|P_2|^2 > \lambda^{-1}|P_1|^2$ , then the inequality implies

$$\frac{\langle P_1, P_2 \rangle}{|P_1| \cdot |P_2|} \leq \sqrt{g + \lambda} \left(\frac{1}{g} + 2\sqrt{\frac{2}{g}\lambda + (2g - 1)\lambda} + \frac{\lambda}{2}\right).$$

As  $\lambda \rightarrow 0$ , the right-hand side converges to  $1/\sqrt{g}$ . It follows that the right-hand side can be taken to be smaller than any given  $\mu > 1/\sqrt{g}$ . This proves Theorem 2.1 by Theorem 2.4.

## 2.4 The Arakelov-geometric setting

The Weil heights in Theorem 2.4 are only well-defined up to bounded functions, which is inconvenient for estimation when varying the line bundles. So we are going to realize them as height functions associated to hermitian line bundles.

### 2.4.1 Basic Arakelov geometry

Let us start with some terminology of Arakelov geometry of Arakelov [Ara74] and Gillet–Soulé [GS90].

Let  $\mathcal{X}$  be an arithmetic variety, i.e. a projective and flat integral scheme over  $\text{Spec } \mathbb{Z}$ . Assume that the generic fiber  $\mathcal{X}_{\mathbb{Q}} = \mathcal{X} \times_{\mathbb{Z}} \mathbb{Q}$  of  $\mathcal{X}$  is smooth. Let  $\overline{\mathcal{L}} = (\mathcal{L}, \|\cdot\|)$  be a *hermitian line bundle* on  $\mathcal{X}$ , i.e.,  $\mathcal{L}$  is a line bundle on  $\mathcal{X}$ , and  $\|\cdot\|$  is a (smooth) hermitian metric of  $\mathcal{L}(\mathbb{C})$  on the complex manifold  $\mathcal{X}(\mathbb{C})$  invariant under the complex conjugation.

For any section  $s \in \Gamma(\mathcal{X}, \mathcal{L})_{\mathbb{R}} = \Gamma(\mathcal{X}, \mathcal{L}) \otimes_{\mathbb{Z}} \mathbb{R} \subset \Gamma(\mathcal{X}_{\mathbb{C}}, \mathcal{L}_{\mathbb{C}})$ , we have the *supremum norm*

$$\|s\|_{\text{sup}} = \sup_{z \in \mathcal{X}(\mathbb{C})} \|s(z)\|.$$

Picking any Haar measure on  $\Gamma(\mathcal{X}, \mathcal{L})_{\mathbb{R}}$ , we have the *arithmetic Euler characteristic*

$$\hat{\chi}(\overline{\mathcal{L}}) = \log \frac{\text{vol}(B(\overline{\mathcal{L}}))}{\text{vol}(\Gamma(\mathcal{X}, \mathcal{L})_{\mathbb{R}}/\Gamma(\mathcal{X}, \mathcal{L}))},$$

where

$$B(\overline{\mathcal{L}}) = \{s \in \Gamma(\mathcal{X}, \mathcal{L})_{\mathbb{R}} : \|s\|_{\text{sup}} \leq 1\}$$

is the corresponding unit ball. It is easy to see that this definition is independent of the choice of the Haar measure.

A hermitian line bundle  $\overline{\mathcal{L}}$  on  $\mathcal{X}$  is called *nef* if the following conditions are satisfied:

- (1)  $\overline{\mathcal{L}}$  is *relatively semipositive*; i.e., the curvature of  $\overline{\mathcal{L}}$  is semipositive and  $\deg(\mathcal{L}|_{\mathcal{Y}}) \geq 0$  for any closed integral curve  $\mathcal{Y}$  on any special fibre of  $\mathcal{X}$  over  $\text{Spec}(\mathbb{Z})$ ;
- (2)  $\overline{\mathcal{L}}$  is *horizontally semipositive*; i.e., the arithmetic degree  $\widehat{\deg}(\overline{\mathcal{L}}|_{\mathcal{Y}}) \geq 0$  for any one-dimensional horizontal closed integral subscheme  $\mathcal{Y}$  of  $\mathcal{X}$ .

Note that the second condition in (2) means that  $\mathcal{L}$  is nef on any special fiber of  $\mathcal{X}$  over  $\mathbb{Z}$  in the classical sense. By an arithmetic version of Kleiman’s theorem (cf. [YZ21, Cor. A.4.3]), if  $\overline{\mathcal{L}}$  is nef, then  $\overline{\mathcal{L}}^{\dim \mathcal{Y}} \cdot \mathcal{Y} \geq 0$  for any closed integral subscheme  $\mathcal{Y}$  of  $\mathcal{X}$ .

As a convention of this paper, we write tensor products of (hermitian) line bundles additively, so  $n(\overline{\mathcal{L}} - \overline{\mathcal{M}})$  means  $(\overline{\mathcal{L}} \otimes \overline{\mathcal{M}}^{\vee})^{\otimes n}$ .

Now we have the following arithmetic Siu inequality of Yuan [Yua08], whose proof is based on the arithmetic Hilbert–Samuel formula of Gillet–Soulé [GS92] and the arithmetic Nakai–Moishezon criterion of Zhang [Zha95a].

**Theorem 2.5.** *Let  $\overline{\mathcal{L}}$  and  $\overline{\mathcal{M}}$  be nef hermitian line bundles on an arithmetic variety  $\mathcal{X}$  of dimension  $r$ . Then*

$$\hat{\chi}(n(\overline{\mathcal{L}} - \overline{\mathcal{M}})) \geq \frac{n^r}{r!} (\overline{\mathcal{L}}^r - r \overline{\mathcal{L}}^{r-1} \cdot \overline{\mathcal{M}}) + o(n^r).$$

*Proof.* If  $\overline{\mathcal{L}}$  and  $\overline{\mathcal{M}}$  are ample, the theorem follows from [Yua08, Thm. 2.2]. If  $\overline{\mathcal{L}}$  and  $\overline{\mathcal{M}}$  are nef, take an ample hermitian line bundle  $\overline{\mathcal{A}}$  on  $\mathcal{X}$ . Apply the theorem to  $(m\overline{\mathcal{L}} + \overline{\mathcal{A}}, m\overline{\mathcal{M}} + \overline{\mathcal{A}}, i(\overline{\mathcal{L}} - \overline{\mathcal{M}}))$  for positive integers  $m$  and  $i = 0, \dots, m - 1$ . Set  $m \rightarrow \infty$ .  $\square$

## 2.4.2 Arithmetic models

We fix the following arithmetic datum:

- (1) Let  $\mathcal{C}$  be a regular (projective and flat) integral model of  $C$  over  $O_K$ .
- (2) Let  $\mathcal{X}$  be a (projective and flat) integral model of  $X$  over  $O_K$  such that the identification  $\mathcal{X}_K \rightarrow C^2$  extends to a morphism  $\psi : \mathcal{X} \rightarrow \mathcal{C} \times_{O_K} \mathcal{C}$  which induces an isomorphism from  $\psi^{-1}(\mathcal{C}^{\text{sm}} \times_{O_K} \mathcal{C}^{\text{sm}})$  to  $\mathcal{C}^{\text{sm}} \times_{O_K} \mathcal{C}^{\text{sm}}$ . Here  $\mathcal{C}^{\text{sm}}$  denotes the smooth locus of  $\mathcal{C}$  over  $O_K$ .
- (3) Denote by  $\tilde{p}_1, \tilde{p}_2 : \mathcal{X} \rightarrow \mathcal{C}$  the morphism extending the projections  $p_1, p_2 : X \rightarrow C$ , which are also the compositions of  $\psi : \mathcal{X} \rightarrow \mathcal{C} \times_{O_K} \mathcal{C}$  with the projections  $p_1, p_2 : \mathcal{C} \times_{O_K} \mathcal{C} \rightarrow \mathcal{C}$ .
- (4) Let  $\overline{\alpha} = (\tilde{\alpha}, \|\cdot\|)$  be a nef hermitian  $\mathbb{Q}$ -line bundle on  $\mathcal{C}$  extending  $\alpha$ .
- (5) Let  $\overline{\Delta}$  be an arithmetic divisor on  $\mathcal{X}$  extending  $\Delta$ .

There are various methods to construct  $(\mathcal{C}, \mathcal{X}, \overline{\alpha}, \overline{\Delta})$ , since our requirements are very weak. For example, assume that  $C$  has semistable reduction over  $O_K$ , which can be achieved by a finite base change. Then let  $\mathcal{C}$  be the minimal regular model of  $C$  over  $O_K$ , and let  $\overline{\omega}_{\mathcal{C}/O_K}$  be the relative dualizing sheaf with Arakelov's metric at archimedean places. By [Fal84, §5, Thm. 5],  $\overline{\omega}_{\mathcal{C}/O_K}$  is nef. Note that this result is not essential for our purpose, since we could always change the hermitian metric by a small constant multiple to make  $\overline{\omega}_{\mathcal{C}/O_K}$  nef. Take  $\overline{\alpha}$  to be the unique extension of  $\alpha$  such that  $(2g-2)\overline{\alpha} = \overline{\omega}_{\mathcal{C}/O_K}$ . Then we set  $\mathcal{X}$  to be the blowing-up of  $\mathcal{C} \times_{O_K} \mathcal{C}$  along its (0-dimensional) non-smooth locus, which is known to be regular by an easy explicit computation (cf. [Zha10, §3.1]). Set  $\overline{\Delta}$  to be the Zariski closure of  $\Delta$  in  $\mathcal{X}$  with an arbitrary Green function.

Finally, we define Vojta's hermitian  $\mathbb{Q}$ -line bundle  $\overline{\mathcal{L}}_0$  on  $\mathcal{X}$  by

$$\overline{\mathcal{L}}_0 := (\delta_1 - 1)\tilde{p}_1^*\overline{\alpha} + (\delta_2 - 1)\tilde{p}_2^*\overline{\alpha} + \mathcal{O}(\overline{\Delta}).$$

Recall the height function

$$h_{\overline{\mathcal{L}}_0} : X(\overline{K}) \rightarrow \mathbb{R}$$

by

$$h_{\overline{\mathcal{L}}_0}(x) = \frac{1}{\deg(x)} \widehat{\deg}(\overline{\mathcal{L}}_0|_{\bar{x}}),$$

where  $\bar{x}$  is the multi-section of  $\mathcal{X}$  over  $\text{Spec } O_K$  corresponding to  $x$ , and  $\deg(x)$  is the degree over  $\bar{x}$  over  $\text{Spec } O_K$ . For convenience, we normalize the height function by

$$h_{\overline{\mathcal{L}}_0}(x)_{\mathbb{Q}} = \frac{1}{[K : \mathbb{Q}]} h_{\overline{\mathcal{L}}_0}(x), \quad x \in X(\overline{K}).$$

Then the following theorem is equivalent to Theorem 2.4.

**Theorem 2.6.** *Let  $\delta_1, \delta_2$  be positive rational numbers satisfying*

$$\delta_1 > 2g\delta_2, \quad g < \delta_1\delta_2 < g + \frac{1}{4}.$$

*Let  $P_1, P_2 \in C(\overline{K})$  be points. Then*

$$h_{\overline{\mathcal{L}}_0}(P_1, P_2)_{\mathbb{Q}} \geq -2c(\delta_1, \delta_2) \cdot (\delta_1|P_1|^2 + \delta_2|P_2|^2) + O(\delta_1/(\delta_1\delta_2 - g)).$$

*Here the coefficient in the error term is independent of  $(P_1, P_2, \delta_1, \delta_2)$ , and*

$$c(\delta_1, \delta_2) = \sqrt{\frac{2}{g}(\delta_1\delta_2 - g) + (2g - 1)\frac{\delta_2}{\delta_1}}.$$

Note that we can assume that  $P_1, P_2 \in C(K)$  by taking finite base changes, and track that the error term in our proof is uniform in this sense. Hence, from now on, we always assume  $P_1, P_2 \in C(K)$ , and omit the part of checking the uniformity under base change.

## 2.5 Plan of proof

Recall that the Mordell conjecture is reduced to Theorem 2.6. The proof of the theorem will take up the rest of this paper. Let us sketch the idea here by three major steps.

The first step is to construct a small section  $s$  of  $\overline{\mathcal{L}} = d\overline{\mathcal{L}}_0$ , i.e. a global section  $s$  of  $d\mathcal{L}_0$  with small supremum norm  $\|s\|_{\text{sup}}$  over  $X(\mathbb{C})$ . Here  $d$  is any sufficiently large integer such that  $d\overline{\mathcal{L}}_0$  is an integral hermitian line bundle. To achieve this, it suffices to have a suitable lower bound of  $\hat{\chi}(d\overline{\mathcal{L}}_0)$ . We apply Yuan's arithmetic Siu inequality (cf. Theorem 2.5) to achieve this, while Vojta applied the arithmetic Riemann–Roch theorem.

The second step is to obtain a suitable lower bound of the height  $h_{\overline{\mathcal{L}}}(P)$  in terms of the small section  $s$ . This is easy and standard if  $s$  does not vanish at  $P$ . However, it is not practical to expect the non-vanishing, and thus we have to treat the case that  $s$  vanishes at  $P$ . The index of  $s$  at  $P$  (with a proper weight) is a geometric term for “the vanishing order” of  $s$  at  $P$ . Then we give a lower bound of  $h_{\overline{\mathcal{L}}}(P)$  in terms of  $-\log \|s\|_{\text{sup}}$  and a contribution from the index. The smaller the index is, the better the lower bound is. The proof here is mostly to unravel definitions in Arakelov geometry.

The third step is to obtain an upper bound of the index of  $s$  at  $P$ , which will be used to simplify the lower bound of the height  $h_{\overline{\mathcal{L}}}(P)$  to prove Theorem 2.6. In history, there are two approaches to such an upper bound. Vojta used Dyson's lemma, which is geometric in nature; [Bom90] used Roth lemma, which is arithmetic in nature. We follow Vojta's approach by Dyson's lemma. The key idea to prove Dyson's lemma is that by blowing-up  $X$  at  $P$ , the index of  $s$  at  $P$  (with the trivial weight) becomes the multiplicity of the exceptional divisor in the pull-back of  $\text{div}(s)$ , and then it can be bounded by intersection numbers of the line bundles involved.

### Convention for error terms

Throughout this paper, our estimates are sensitive to  $(P_1, P_2, \delta_1, \delta_2, d)$ . As a convention, by an inequality of the form

$$F \geq G + O(H)$$

for real number  $F, G, H$  with  $H > 0$ , we mean

$$F \geq G + cH$$

for a (possibly negative) constant  $c$  independent of  $(P_1, P_2, \delta_1, \delta_2, d)$ . To emphasize this situation, we may also say that the coefficient in the error term is independent of  $(P_1, P_2, \delta_1, \delta_2, d)$ . The convention for  $F \leq G + O(H)$  and  $F = G + O(H)$  are similar.

## 3 Step 1: existence of small section

Recall that

$$\bar{\mathcal{L}}_0 := (\delta_1 - 1)\tilde{p}_1^*\bar{\alpha} + (\delta_2 - 1)\tilde{p}_2^*\bar{\alpha} + \mathcal{O}(\bar{\Delta}).$$

Here  $(\mathcal{C}, \bar{\alpha}, \mathcal{X}, \bar{\Delta})$  is the arithmetic model of  $(C, \alpha, X, \Delta)$  chosen above. In particular,  $\bar{\alpha}$  is nef on  $\mathcal{C}$  by the choice. The goal of this step is to prove the following theorem, which is a variant of [Voj91, Thm. 2.2].

**Theorem 3.1** (small section). *Let  $\delta_1, \delta_2$  be positive rational numbers satisfying  $\delta_1 > \delta_2$  and  $\delta_1\delta_2 > g$ . Then there exists a positive integers  $d_0$  depending on  $\bar{\mathcal{L}}_0$ , such that for any integer  $d \geq d_0$  making  $d\bar{\mathcal{L}}_0$  an (integral) hermitian line bundle on  $\mathcal{X}$ , there exists a nonzero element  $s \in H^0(\mathcal{X}, d\mathcal{L}_0)$  satisfying*

$$-\log \|s\|_{\text{sup}} \geq O\left(\frac{d\delta_1^2\delta_2}{\delta_1\delta_2 - g}\right).$$

### 3.1 Hilbert–Samuel type of estimates

Denote

$$\alpha_1 = p_1^*\alpha, \quad \alpha_2 = p_2^*\alpha, \quad \bar{\alpha}_1 = \tilde{p}_1^*\bar{\alpha}, \quad \bar{\alpha}_2 = \tilde{p}_2^*\bar{\alpha}.$$

Recall

$$L_0 = \delta_1 \alpha_1 + \delta_2 \alpha_2 + (\mathcal{O}(\Delta) - \alpha_1 - \alpha_2)$$

and

$$\bar{\mathcal{L}}_0 = \delta_1 \bar{\alpha}_1 + \delta_2 \bar{\alpha}_2 + (\mathcal{O}(\bar{\Delta}) - \bar{\alpha}_1 - \bar{\alpha}_2).$$

We first have the following geometric estimates, which implies that  $L_0$  is big on  $X$ . Note that we cannot count sections of the  $\mathbb{Q}$ -line bundle  $L_0$  directly, so take multiples of  $L_0$  which are (integral) line bundles on  $X$ .

**Theorem 3.2.** *For positive integers  $d$  such that  $dL_0$  is a line bundle on  $X$ , we have*

$$h^0(dL_0) \geq (\delta_1\delta_2 - g)d^2 + O(d).$$

*Proof.* By the Riemann–Roch theorem, for any positive integer  $d$  such that  $dL_0$  is an integral line bundle on  $X$ ,

$$h^0(dL_0) - h^1(dL_0) + h^2(dL_0) = \frac{1}{2}d^2L_0^2 - \frac{1}{2}dL_0 \cdot \omega_X + \chi(\mathcal{O}_X).$$

By the Serre duality,  $h^2(dL_0) = h^0(\omega_X - dL_0)$ . The line bundle  $M = \alpha_1 + \alpha_2$  is ample on  $X$ . An easy calculation gives  $L_0 \cdot M = \delta_1 + \delta_2$ . It follows that for sufficiently large  $d$ , the intersection number  $M \cdot (\omega_X - dL_0) < 0$ , and thus  $h^0(\omega_X - dL_0) = 0$ . It follows that

$$h^0(dL_0) \geq \frac{1}{2}d^2L_0^2 - \frac{1}{2}dL_0 \cdot \omega_X + \chi(\mathcal{O}_X).$$

□

*Remark 3.3.* One can also prove that  $h^0(dL_0) \leq (\delta_1\delta_2)d^2 + O(d)$ , but we do not need this fact here.

The following key estimate can be viewed as an arithmetic counterpart of Theorem 3.2. We will prove it by Yuan’s arithmetic Siu inequality instead of Gillet–Soulé’s arithmetic Riemann–Roch theorem.

**Theorem 3.4.** *There is a positive constant  $A_{\delta_1, \delta_2}$  depending on  $\delta_1, \delta_2$  such that for positive integers  $d > A_{\delta_1, \delta_2}$  making  $d\bar{\mathcal{L}}_0$  is a hermitian line bundle on  $\mathcal{X}$ , we have*

$$\hat{\chi}(d\bar{\mathcal{L}}_0) \geq \frac{1}{6}d^3(\bar{\mathcal{L}}_0^3) + O(\delta_1 d^3) = O(\delta_1^2 \delta_2 d^3).$$

Let us first see how Theorem 3.2 and Theorem 3.4 imply Theorem 3.1 by applying Minkowski’s theorem. In fact, denote  $\epsilon = (\delta_1\delta_2 - g)/2$ . By the theorems, for sufficiently large and sufficiently divisible  $d$  (relatively to  $(\delta_1, \delta_2)$ ), we have

$$h^0(dL_0) \geq \frac{1}{2}d^2 \text{vol}(L_0) - \epsilon d^2 \geq \frac{1}{2}d^2(\delta_1\delta_2 - g),$$

and

$$\hat{\chi}(d\bar{\mathcal{L}}_0) \geq O(\delta_1^2 \delta_2 d^3).$$

By Minkowski’s theorem, there exists a nonzero section  $s \in \Gamma(\mathcal{X}, d\mathcal{L}_0)$  such that

$$\begin{aligned} -\log \|s\|_{\text{sup}} &\geq \frac{1}{[K : \mathbb{Q}]} \frac{\hat{\chi}(d\bar{\mathcal{L}}_0)}{h^0(dL_0)} - \log 2 \\ &\geq \frac{O(d^3 \delta_1^2 \delta_2)}{d^2(\delta_1\delta_2 - g)} - \log 2 = O\left(\frac{d\delta_1^2 \delta_2}{\delta_1\delta_2 - g}\right). \end{aligned}$$

Here the second inequality uses the lower bound of  $h^0(dL_0)$  (instead of an upper bound) due to the extreme possibility that  $O(d^3 \delta_1^2 \delta_2)$  might represent a negative number. This proves Theorem 3.1.

### 3.2 Estimate of the arithmetic Euler characteristic

Now we prove Theorem 3.4. Recall

$$\bar{\mathcal{L}}_0 = \delta_1 \bar{\alpha}_1 + \delta_2 \bar{\alpha}_2 + (\mathcal{O}(\bar{\Delta}) - \bar{\alpha}_1 - \bar{\alpha}_2).$$

Write

$$\mathcal{O}(\bar{\Delta}) - \bar{\alpha}_1 - \bar{\alpha}_2 = \bar{\mathcal{A}} - \bar{\mathcal{B}}$$

for nef hermitian line bundles  $\bar{\mathcal{A}}, \bar{\mathcal{B}}$  on  $\mathcal{X}$ . It follows that

$$\bar{\mathcal{L}}_0 = \bar{\mathcal{A}}' - \bar{\mathcal{B}}$$

for

$$\bar{\mathcal{A}}' = \delta_1 \bar{\alpha}_1 + \delta_2 \bar{\alpha}_2 + \bar{\mathcal{A}}.$$

By the construction,  $\bar{\alpha}$  is nef on  $\mathcal{C}$ , so  $\bar{\mathcal{A}}'$  is also nef on  $\mathcal{X}$ . Then we are ready to apply Theorem 2.5. It gives

$$\hat{\chi}(d\bar{\mathcal{L}}_0) \geq \frac{d^3}{6}(\bar{\mathcal{A}}'^3 - 3\bar{\mathcal{A}}'^2 \cdot \bar{\mathcal{B}}) + o_{\delta_1, \delta_2}(d^3) = \frac{d^3}{6}(\bar{\mathcal{L}}_0^3 - 3\bar{\mathcal{A}}' \cdot \bar{\mathcal{B}}^2 + \bar{\mathcal{B}}^3) + o_{\delta_1, \delta_2}(d^3).$$

Here the coefficient of the error term  $o_{\delta_1, \delta_2}(d^3)$  depends on  $\delta_1, \delta_2$ .

By the expressions defining  $\bar{\mathcal{L}}_0, \bar{\mathcal{A}}'$  and  $\bar{\mathcal{B}}$ , we see that the intersection numbers on the right-hand side are polynomials of  $\delta_1, \delta_2$  of degree at most 3, whose coefficients are given by intersection numbers of  $\bar{\alpha}_1, \bar{\alpha}_2, \bar{\Delta}, \bar{\mathcal{A}}, \bar{\mathcal{B}}$  on  $\mathcal{X}$ . The polynomials appearing in  $\bar{\mathcal{A}} \cdot \bar{\mathcal{B}}^2$  and  $\bar{\mathcal{B}}^3$  have degrees at most 1. As  $\delta_1 > \delta_2$ , we obtain

$$-3\bar{\mathcal{A}}' \cdot \bar{\mathcal{B}}^2 + \bar{\mathcal{B}}^3 = O(\delta_1).$$

Note that  $\delta_1^2 \delta_2 > g\delta_1$  by assumption. Then it remains to prove

$$\bar{\mathcal{L}}_0^3 = O(\delta_1^2 \delta_2).$$

The coefficients of  $\delta_1^3$  and  $\delta_1^2$  in  $\bar{\mathcal{L}}_0^3$  are  $\bar{\alpha}_1^3$  and  $3\bar{\alpha}_1^2 \cdot (\bar{\Delta} - \bar{\alpha}_1 - \bar{\alpha}_2)$  respectively. We claim that  $\bar{\alpha}_1^3 = \bar{\alpha}_1^2 \cdot (\bar{\Delta} - \bar{\alpha}_1 - \bar{\alpha}_2) = 0$ . With the claim, the next largest monomial of  $\delta_1, \delta_2$  is  $\delta_1^2 \delta_2$  since  $\delta_1 > \delta_2$  and  $\delta_1 \delta_2 > g$ . This proves the theorem.

It remains to prove the claim that  $\bar{\alpha}_1^3 = \bar{\alpha}_1^2 \cdot (\bar{\Delta} - \bar{\alpha}_1 - \bar{\alpha}_2) = 0$ . This is just a special case of the following basic result for the map  $\tilde{p}_1 : \mathcal{X} \rightarrow \mathcal{C}$ .

**Lemma 3.5.** *Let  $f : \mathcal{X} \rightarrow \mathcal{Y}$  be a morphism of arithmetic varieties over  $\mathbb{Z}$ . Let  $r = \dim \mathcal{Y}$  and  $r + n = \dim \mathcal{X}$ . Let  $\eta \in \mathcal{Y}$  be the generic point and  $\mathcal{X}_\eta$  be the generic fiber of  $f$ . Let  $\bar{\mathcal{L}}_1, \dots, \bar{\mathcal{L}}_n$  be hermitian line bundles on  $\mathcal{X}$  and  $\bar{\mathcal{M}}_1, \dots, \bar{\mathcal{M}}_r$  be hermitian line bundles on  $\mathcal{Y}$ . Then*

$$\bar{\mathcal{L}}_1 \cdots \bar{\mathcal{L}}_n \cdot f^* \bar{\mathcal{M}}_1 \cdots f^* \bar{\mathcal{M}}_r = (\mathcal{L}_{1, \eta} \cdots \mathcal{L}_{n, \eta})(\bar{\mathcal{M}}_1 \cdots \bar{\mathcal{M}}_r).$$

Here the expressions in the three brackets are intersection numbers on  $\mathcal{X}, \mathcal{X}_\eta, \mathcal{Y}$  respectively. If  $f$  is not surjective, then  $\mathcal{X}_\eta$  is empty and the intersection number on it is 0 by convention.

*Proof.* We use induction on the relative dimension  $n$ . The case  $n = 0$  follows from the projection formula. Assume that the result holds for  $n - 1$ , and then consider the case  $n$ . Take a global section  $s$  of  $\mathcal{L}_n$ , and write  $\text{div}(s) = \sum_{i=1}^m a_i \mathcal{X}_i$  for prime divisors  $\mathcal{X}_i$  on  $\mathcal{X}$ . Denote by  $f_i : \mathcal{X}_i \rightarrow \mathcal{Y}$  the restriction of  $f$  to  $\mathcal{X}_i$ . We have the induction formula

$$\begin{aligned} \bar{\mathcal{L}}_1 \cdots \bar{\mathcal{L}}_n \cdot f^* \bar{\mathcal{M}}_1 \cdots f^* \bar{\mathcal{M}}_r &= \sum_{i=1}^m a_i \bar{\mathcal{L}}_1|_{\mathcal{X}_i} \cdots \bar{\mathcal{L}}_{n-1}|_{\mathcal{X}_i} \cdot f_i^* \bar{\mathcal{M}}_1 \cdots f_i^* \bar{\mathcal{M}}_r \\ &\quad - \sum_{i=1}^m a_i \int_{\mathcal{X}_i(\mathbb{C})} \log \|s\| c_1(\bar{\mathcal{L}}_1) \cdots c_1(\bar{\mathcal{L}}_n) f_i^* c_1(\bar{\mathcal{M}}_1) \cdots f_i^* c_1(\bar{\mathcal{M}}_r). \end{aligned}$$

The integral equals 0 since  $c_1(\bar{\mathcal{M}}_1) \cdots c_1(\bar{\mathcal{M}}_r) = 0$  on  $\mathcal{Y}(\mathbb{C})$  by dimension reason. It follows that

$$\bar{\mathcal{L}}_1 \cdots \bar{\mathcal{L}}_n \cdot f^* \bar{\mathcal{M}}_1 \cdots f^* \bar{\mathcal{M}}_r = \sum_{i=1}^m a_i \bar{\mathcal{L}}_1|_{\mathcal{X}_i} \cdots \bar{\mathcal{L}}_{n-1}|_{\mathcal{X}_i} \cdot f_i^* \bar{\mathcal{M}}_1 \cdots f_i^* \bar{\mathcal{M}}_r.$$

Let us first simplify every term of the right-hand side.

If  $\mathcal{X}_i$  is vertical in the sense that the image of  $\mathcal{X}_i \rightarrow \text{Spec } \mathbb{Z}$  is a closed point represented by a prime number  $p$ , then

$$\bar{\mathcal{L}}_1|_{\mathcal{X}_i} \cdots \bar{\mathcal{L}}_{n-1}|_{\mathcal{X}_i} \cdot f_i^* \bar{\mathcal{M}}_1 \cdots f_i^* \bar{\mathcal{M}}_r = \mathcal{L}_1|_{\mathcal{X}_i} \cdots \mathcal{L}_{n-1}|_{\mathcal{X}_i} \cdot f_i^*(\mathcal{M}_1|_{\mathcal{Y}_p}) \cdots f_i^*(\mathcal{M}_r|_{\mathcal{Y}_p}).$$

is 0 as  $r > \dim \mathcal{Y}_p$ . If  $\mathcal{X}_i$  is horizontal in that  $\mathcal{X}_i \rightarrow \text{Spec } \mathbb{Z}$  is surjective, then it is an arithmetic variety, and the induction hypothesis gives

$$\bar{\mathcal{L}}_1|_{\mathcal{X}_i} \cdots \bar{\mathcal{L}}_{n-1}|_{\mathcal{X}_i} \cdot f_i^* \bar{\mathcal{M}}_1 \cdots f_i^* \bar{\mathcal{M}}_r = (\mathcal{L}_{1,\eta} \cdots \mathcal{L}_{n-1,\eta} \cdot \mathcal{X}_{i,\eta})(\bar{\mathcal{M}}_1 \cdots \bar{\mathcal{M}}_r).$$

Putting these into the formula, we get the result for  $n$ .  $\square$

## 4 Step 2: lower bound of the height

Recall that Theorem 3.1 constructs a small section  $s$  of  $\bar{\mathcal{L}} = d\bar{\mathcal{L}}_0$ . Write

$$\bar{\mathcal{L}} = (d_1 - d)\tilde{p}_1^* \bar{\alpha} + (d_2 - d)\tilde{p}_2^* \bar{\alpha} + \mathcal{O}(d\bar{\Delta}),$$

where

$$d_1 = \delta_1 d, \quad d_2 = \delta_2 d.$$

The goal here is to give a lower bound of  $h_{\bar{\mathcal{L}}}(P)$  using the section  $s$ .

Denote by  $\mathcal{L}$  the underlying line bundle of  $\bar{\mathcal{L}}$ , and by  $L = \mathcal{L}_K$  the generic fiber as before. Recall that  $\mathcal{C}$  and  $\mathcal{X}$  are respectively integral models of  $C$  and  $X = C^2$  over  $O_K$ . Recall that  $P = (P_1, P_2)$  is a  $K$ -point of  $X$ . Denote by  $\mathcal{P}$  (resp.  $\mathcal{P}_1, \mathcal{P}_2$ ) the Zariski closure of  $P$  (resp.  $P_1, P_2$ ) in  $\mathcal{X}$  (resp.  $\mathcal{C}, \mathcal{C}$ ).

If  $s(P) \neq 0$ , we would simply have

$$h_{\bar{\mathcal{L}}}(P) = \widehat{\text{deg}}(\bar{\mathcal{L}}|_{\mathcal{P}}) \geq - \sum_{\sigma: K \rightarrow \mathbb{C}} \log \|s(P)\|_{\sigma} \geq - \sum_{\sigma: K \rightarrow \mathbb{C}} \log \|s\|_{\sigma, \text{sup}}.$$

This would imply Theorem 2.4 by the bound in Theorem 3.1. However, in general, we cannot avoid the case  $s(P) = 0$  where the above argument does not work. We will solve this problem by carefully analyzing of the effect of the index (which we will define later) of  $s$  at  $P$  on the height.

For  $i = 1, 2$ , choose a local coordinate  $t_i$  of  $C_i$  at  $P_i$ , i.e.  $t_i$  is a generator of the maximal ideal of the local ring of  $C_i$  at  $P_i$ . Let  $\mathcal{O}_X$  denotes the structure sheaf of  $X$ . For every local regular function  $f$  at  $P$  (i.e.  $f \in \mathcal{O}_{X,P}$ ), let  $a_{i_1, i_2}$  be the coefficient of  $t_1^{i_1} t_2^{i_2}$  in the power series expansion of  $f$  in terms of variables  $t_1, t_2$ . Although  $a_{i_1, i_2}$  may depends on the choice of  $(s_0, t_1, t_2)$ , the pairs  $(i_1, i_2)$  such that  $a_{i_1, i_2} \neq 0$  does not depend on the choice of  $(s_0, t_1, t_2)$ . Therefore, we can define *the index of  $f$  at  $P$  of weight  $(d_1, d_2)$*  by

$$\text{Ind}_P(f, (d_1, d_2)) = \min \left\{ \frac{i_1}{d_1} + \frac{i_2}{d_2} : a_{i_1, i_2} \neq 0 \right\}. \quad (1)$$

Now consider the line bundle  $L$  and its section  $s$ . Choose a local section  $s_0$  of  $L$  at  $P$  such that  $s_0(P) \neq 0$ . Then  $s/s_0$  is a local regular function at  $P$ , and we define *the index of  $s$  at  $P$  of weight  $(d_1, d_2)$*  by

$$\text{Ind}_P(s, (d_1, d_2)) = \text{Ind}_P \left( \frac{s}{s_0}, (d_1, d_2) \right).$$

Note that  $\text{Ind}_P(s, (d_1, d_2))$  is independent of the choice of  $s_0$ .

Assume that the index  $\text{Ind}_P(s, (d_1, d_2))$  is obtained at the monomial  $t_1^{e_1} t_2^{e_2}$  of the power series expansion of  $s$  at  $P$ . The following is the main result of this section, whose current form does not require the previous assumptions on  $(\delta_1, \delta_2, d, P_1, P_2)$  and  $s$ .

**Theorem 4.1.** *Let  $s$  be a nonzero global section of  $\mathcal{L}$  on  $\mathcal{X}$ . Then*

$$h_{\bar{\mathcal{L}}}(P) \geq - \sum_{\sigma: K \rightarrow \mathbb{C}} \log \|s\|_{\sigma, \text{sup}} - \frac{2g-2}{g} [K : \mathbb{Q}] (e_1 |P_1|^2 + e_2 |P_2|^2) + O(d_1 + d_2 + d + e_1 + e_2).$$

Here the coefficients in the error terms are independent of  $(\delta_1, \delta_2, d, P_1, P_2, s)$ .

## 4.1 The hermitian line bundle by restriction

For  $i = 1, 2$ , denote by  $\bar{\mathcal{P}}_i = (\mathcal{P}_i, g_{\mathcal{P}_i})$  the arithmetic divisor on  $\mathcal{X}$  with underlying divisor  $\mathcal{P}_i$  and Arakelov's Green function  $g_{\mathcal{P}_i}$ . This induces hermitian line bundles  $\mathcal{O}(\bar{\mathcal{P}}_i)$  on  $\mathcal{X}$ .

Define a hermitian line bundle  $\bar{\mathcal{M}}$  on  $\text{Spec } O_K$  by

$$\bar{\mathcal{M}} := (\bar{\mathcal{L}} - e_1 \tilde{p}_1^* \mathcal{O}(\bar{\mathcal{P}}_1) - e_2 \tilde{p}_2^* \mathcal{O}(\bar{\mathcal{P}}_2))|_{\mathcal{P}}.$$

Here  $\tilde{p}_i : \mathcal{X} \rightarrow \mathcal{C}$  is the morphism extending the projection  $p_i : X \rightarrow C$ .

Denote by  $\bar{\omega}_{\mathcal{C}/O_K}$  the relative dualizing sheaf of  $\mathcal{C}$  over  $O_K$  endowed with the Arakelov metric  $\|\cdot\|_{\text{Ar}}$  as in [Ara74, §4]. Then we obtain the following result immediately.

**Lemma 4.2.**

$$\begin{aligned} h_{\bar{\mathcal{L}}}(P) &= \widehat{\deg}(\bar{\mathcal{M}}) - e_1 h_{\bar{\omega}_{\mathcal{C}/O_K}}(P_1) - e_2 h_{\bar{\omega}_{\mathcal{C}/O_K}}(P_2) \\ &= \widehat{\deg}(\bar{\mathcal{M}}) - \frac{2g-2}{g} [K : \mathbb{Q}] (e_1 |P_1|^2 + e_2 |P_2|^2) + O(e_1 + e_2). \end{aligned}$$

*Proof.* By definition,

$$\widehat{\deg}(\bar{\mathcal{M}}) = \bar{\mathcal{L}} \cdot \mathcal{P} - e_1 \tilde{p}_1^* \mathcal{O}(\bar{\mathcal{P}}_1) \cdot \mathcal{P} - e_2 \tilde{p}_2^* \mathcal{O}(\bar{\mathcal{P}}_2) \cdot \mathcal{P}.$$

Note that

$$\tilde{p}_i^* \mathcal{O}(\bar{\mathcal{P}}_i) \cdot \mathcal{P} = \mathcal{O}(\bar{\mathcal{P}}_i) \cdot \mathcal{P}_i = -\bar{\omega}_{\mathcal{C}/O_K} \cdot \mathcal{P}_i = -[K : \mathbb{Q}] \frac{2g-2}{g} |P_i|^2 + O(1),$$

where the first equality follows from the projection formula for  $\tilde{p}_i : \mathcal{X} \rightarrow \mathcal{C}$ , the second equality follows from the arithmetic adjunction formula, and the third equality is proved in Theorem 2.2.  $\square$

## 4.2 The section of the line bundle

The underlying line bundle  $\mathcal{M}$  of  $\bar{\mathcal{M}}$  on  $\text{Spec } O_K$  is given by

$$\mathcal{M} = (\mathcal{L} - e_1 \tilde{p}_1^* \mathcal{O}(\mathcal{P}_1) - e_2 \tilde{p}_2^* \mathcal{O}(\mathcal{P}_2))|_{\mathcal{P}}.$$

It can be identified with a locally free  $O_K$ -module of rank one. We will define a nonzero section  $s^\circ$  of  $\mathcal{M}$  from the section  $s$  of  $\mathcal{L}$ .

Start with the generic fiber

$$M = (L - e_1 p_1^* \mathcal{O}(P_1) - e_2 p_2^* \mathcal{O}(P_2))|_P.$$

As in the definition of the index, let  $s_0$  be a local section of  $L$  at  $P$  with  $s_0(P) \neq 0$ , and for  $i = 1, 2$  let  $t_i$  be a local coordinate of  $C_i$  at  $P_i$ . Then we have a power series expansion

$$s = s_0(a_{e_1, e_2} t_1^{e_1} t_2^{e_2} + \dots).$$

Via the canonical injection  $\mathcal{O}(-P_i) \rightarrow \mathcal{O}_C$ , the local section  $t_i$  of  $\mathcal{O}_C$  corresponds to a local section  $t_i^\dagger$  of the line bundle  $\mathcal{O}(-P_i)$  with  $t_i^\dagger(P_i) \neq 0$ . Finally, we obtain a nonzero element  $s^\circ$  of  $M$  by setting

$$s^\circ = (a_{e_1, e_2} s_0) \otimes (t_1^\dagger)^{\otimes e_1} \otimes (t_2^\dagger)^{\otimes e_2}$$

according to the decomposition

$$M = (L + e_1 p_1^* \mathcal{O}(-P_1) + e_2 p_2^* \mathcal{O}(-P_2))|_P.$$

**Lemma 4.3.** *The element  $s^\circ \in M$  is independent of the choice of  $(s_0, t_1, t_2)$ , and lies in the lattice  $\mathcal{M}$  of  $M$ .*

*Proof.* We first prove the first statement. If we change  $s_0$  to  $s'_0 = us_0$  for an invertible local regular function  $u$  at  $P$ , then the expansion

$$s = s_0(a_{e_1, e_2} t_1^{e_1} t_2^{e_2} + \dots)$$

becomes

$$s = s'_0(a'_{e_1, e_2} t_1^{e_1} t_2^{e_2} + \dots)$$

with  $a'_{e_1, e_2} = u(0)^{-1} a_{e_1, e_2}$ . Thus  $s^\circ$  does not change.

Similarly, if we change  $(t_1, t_2)$  to  $(t'_1, t'_2) = (u_1 t_1, u_2 t_2)$  for invertible local regular functions  $u_1, u_2$  at  $P$ , then the process does not change  $s^\circ$  either. This proves that  $s^\circ \in M$  is independent of the choice of  $(s_0, t_1, t_2)$ .

To prove the second statement, we first review some geometric properties. Since  $\mathcal{C}$  is regular, the sections  $\mathcal{P}_1, \mathcal{P}_2$  lie in the smooth locus  $\mathcal{C}^{\text{sm}}$  of  $\mathcal{C}$  above  $O_K$ . See [Liu02, §9.1, Cor. 1.32] for this well-known fact. By assumption, the natural map  $\phi : \mathcal{X} \rightarrow \mathcal{C} \times_{O_K} \mathcal{C}$  is an isomorphism above  $\mathcal{C}^{\text{sm}} \times_{O_K} \mathcal{C}^{\text{sm}}$ , so it is an isomorphism from a neighborhood of  $\mathcal{P}$  in  $\mathcal{X}$  to a neighborhood of  $\mathcal{P}_1 \times_{O_K} \mathcal{P}_2$  in  $\mathcal{C} \times_{O_K} \mathcal{C}$ .

Now we prove the second statement that  $s^\circ$  lies in  $\mathcal{M}$ . It suffices to prove that  $s^\circ$  lies in every localization

$$\mathcal{M}_{(\wp)} = \mathcal{M} \otimes_{O_K} O_{K,(\wp)},$$

where  $O_{K,(\wp)}$  denotes the local ring of  $O_K$  at a prime ideal  $\wp$ . This will be done by an integral version of the definition of  $s^\circ$ . For  $i = 1, 2$ , let  $\mathcal{P}_{i, \wp}$  (resp.  $\mathcal{P}_\wp$ ) be a closed point of  $\mathcal{P}_i$  (resp.  $\mathcal{P}$ ) above  $\wp$ . Since  $\mathcal{P}_i$  lies in the smooth locus of  $\mathcal{C}$ , the maximal ideal of the local ring  $\mathcal{O}_{\mathcal{C}, \mathcal{P}_{i, \wp}}$  can be generated by a single element, which we denote by  $t_i$ . Moreover, the completion of  $\mathcal{O}_{\mathcal{C}, \mathcal{P}_{i, \wp}}$  is isomorphic to the formal power series ring  $O_{K,(\wp)}[[t_i]]$ . Via the morphism  $\phi$ , the completion of  $\mathcal{O}_{\mathcal{X}, \mathcal{P}_\wp}$  is isomorphic to the formal power series ring  $O_{K,(\wp)}[[t_1, t_2]]$ .

Let  $s_0$  be a generator of the stalk of  $\mathcal{L}$  at  $\mathcal{P}_\wp$ . Then  $s/s_0$  lies in  $\mathcal{O}_{\mathcal{X}, \mathcal{P}_\wp}$ , and thus has a power series expansion

$$s/s_0 = \sum_{i_1, i_2} a_{i_1, i_2} t_1^{i_1} t_2^{i_2}, \quad a_{i_1, i_2} \in O_{K,(\wp)}.$$

In particular, we have  $a_{e_1, e_2} \in O_{K,(\wp)}$ . It follows that

$$s^\circ = (a_{e_1, e_2} s_0) \otimes (t_1^\dagger)^{\otimes e_1} \otimes (t_2^\dagger)^{\otimes e_2}$$

lies in

$$\mathcal{M} = (\mathcal{L} + e_1 \tilde{p}_1^* \mathcal{O}(-\mathcal{P}_1) + e_2 \tilde{p}_2^* \mathcal{O}(-\mathcal{P}_2))|_{\mathcal{P}}.$$

This completes the proof.  $\square$

### 4.3 The degree of the line bundle

In the above paragraphs, we have constructed a nonzero section

$$s^\circ = (a_{e_1, e_2} s_0) \otimes (t_1^\dagger)^{\otimes e_1} \otimes (t_2^\dagger)^{\otimes e_2}$$

of the underlying line bundle  $\mathcal{M}$  of

$$\overline{\mathcal{M}} = (\overline{\mathcal{L}} - e_1 \widehat{p}_1^* \mathcal{O}(\overline{\mathcal{P}}_1) - e_2 \widehat{p}_2^* \mathcal{O}(\overline{\mathcal{P}}_2))|_{\mathcal{P}}.$$

By definition, for every embedding  $\sigma : K \rightarrow \mathbb{C}$ , the metric

$$\|s^\circ\|_\sigma = |a_{e_1, e_2}|_\sigma \cdot \|s_0(P)\|_\sigma \cdot \|t_1^\dagger(P_1)\|_\sigma^{e_1} \cdot \|t_2^\dagger(P_2)\|_\sigma^{e_2}.$$

Here the metric

$$\|t_i^\dagger(P_i)\|_\sigma = \|t_i \cdot 1^\dagger(P_i)\|_\sigma = \lim_{Q \rightarrow P_i} (|t_i(Q)|_\sigma \cdot \|1^\dagger(Q)\|_\sigma) = \lim_{Q \rightarrow P_i} (|t_i(Q)|_\sigma \cdot e^{g_{\mathcal{P}_i, \sigma}(Q)}),$$

where  $Q$  is a point on  $C_\sigma(\mathbb{C})$  converging to  $P_i$ . By the definition of the Arakelov metric  $\|\cdot\|_{\text{Ar}}$  on  $\overline{\omega}_{C/O_K}$ , we have exactly

$$\|t_i^\dagger(P_i)\|_\sigma = \|(dt_i)(P_i)\|_{\text{Ar}, \sigma}.$$

It follows that

$$\|s^\circ\|_\sigma = |a_{e_1, e_2}|_\sigma \cdot \|s_0(P)\|_\sigma \cdot \|(dt_1)(P_1)\|_{\text{Ar}, \sigma}^{e_1} \cdot \|(dt_2)(P_2)\|_{\text{Ar}, \sigma}^{e_2}.$$

Now we are ready to estimate the degree of  $\overline{\mathcal{M}}$ . By definition,

$$\widehat{\deg}(\overline{\mathcal{M}}) = \log \#(\mathcal{M}/sO_K) - \sum_{\sigma: K \rightarrow \mathbb{C}} \log \|s^\circ\|_\sigma \geq - \sum_{\sigma: K \rightarrow \mathbb{C}} \log \|s^\circ\|_\sigma.$$

It follows that

$$\begin{aligned} \widehat{\deg}(\overline{\mathcal{M}}) \geq - \sum_{\sigma: K \rightarrow \mathbb{C}} \left( \log |a_{e_1, e_2}|_\sigma + \log \|s_0(P)\|_\sigma \right. \\ \left. + e_1 \log \|(dt_1)(P_1)\|_{\text{Ar}, \sigma} + e_2 \log \|(dt_2)(P_2)\|_{\text{Ar}, \sigma} \right). \quad (2) \end{aligned}$$

Denote by

$$D_{r_i} = \{Q \in C_\sigma(\mathbb{C}) : |t_i(Q)| \leq r_i\}$$

the closed disc of radius  $r_i$  in  $C_\sigma(\mathbb{C})$  with center  $P_i$ . We can choose  $r_1, r_2$  small enough so that  $s_0$  is holomorphic and everywhere non-vanishing at a neighborhood of the poly-disc

$$D = D_{r_1, r_2} = D_{r_1} \times D_{r_2}$$

in  $X_\sigma(\mathbb{C})$ . Then  $f = s/s_0$  is a holomorphic function on  $D$ . By Cauchy's integration formula,

$$a_{e_1, e_2} = \frac{1}{(2\pi i)^2} \int_{\partial D_{r_1}} \int_{\partial D_{r_2}} \frac{f(z_1, z_2)}{z_1^{e_1+1} z_2^{e_2+1}} dz_2 dz_1.$$

It follows that

$$|a_{e_1, e_2}|_\sigma \leq r_1^{-e_1} r_2^{-e_2} |f|_{D, \text{sup}}.$$

Here  $|f|_{D,\text{sup}}$  denotes the maximal value of  $|f|$  on  $D$ , and we will use other similar notation. For example, we have

$$|f|_{D,\text{sup}} = \|s/s_0\|_{D,\text{sup}} \leq \|s\|_{D,\text{sup}}/\|s_0\|_{D,\text{inf}} \leq \|s\|_{\sigma,\text{sup}}/\|s_0\|_{D,\text{inf}}.$$

Therefore, (2) becomes

$$\widehat{\text{deg}}(\overline{\mathcal{M}}) \geq - \sum_{\sigma:K \rightarrow \mathbb{C}} \left( \log \|s\|_{\sigma,\text{sup}} + \log \frac{\|s_0(P)\|_{\sigma}}{\|s_0\|_{D,\text{inf}}} - e_1 \log r_1 - e_2 \log r_2 \right. \\ \left. + e_1 \log \|(dt_1)(P_1)\|_{\text{Ar},\sigma} + e_2 \log \|(dt_2)(P_2)\|_{\text{Ar},\sigma} \right).$$

#### 4.4 The uniform choice

Now we are going to choose  $(s_0, (t_1, t_2))$  carefully to get a uniform bound with respect to  $P$ . Consider the set of quadruples  $(U_1 \times U_2, V, (z_1, z_2), s_0)$ , where

- (1)  $U_1, U_2$  are open subsets of  $C_{\sigma}(\mathbb{C})$  under the complex topology;
- (2)  $z_1: U_1 \rightarrow \{w \in \mathbb{C} : |w| < 3\}$  and  $z_2: U_2 \rightarrow \{w \in \mathbb{C} : |w| < 3\}$  are biholomorphic maps;
- (3)  $V = \{Q \in U_1 \times U_2 : |z_1(Q)| < 1, |z_2(Q)| < 1\}$  is a poly-disc in  $U$ ;
- (4)  $s_0$  is a holomorphic section of  $L$  on  $U$  which is everywhere non-vanishing on  $U$ .

By compactness, we can choose finitely many quadruples  $(U_1 \times U_2, V, (z_1, z_2), s_0)$  such that these poly-discs  $V$  cover  $X_{\sigma}(\mathbb{C})$ . For every point  $P = (P_1, P_2) \in X_{\sigma}(\mathbb{C})$ , there exists some chosen quadruple  $(U_1 \times U_2, V, (z_1, z_2), s_0)$  such that  $P \in V$ . Then we set  $(t_1, t_2) = (z_1 - z_1(P_1), z_2 - z_2(P_2))$ . Take the poly-disc

$$D = \{Q \in U : |t_1(Q)| \leq 1, |t_2(Q)| \leq 1\},$$

which is contained in

$$\{Q \in U_1 \times U_2 : |z_1(Q)| < 2, |z_2(Q)| < 2\}.$$

By the currently chosen  $(s_0, (t_1, t_2))$  for  $P$ , we have the bound

$$\widehat{\text{deg}}(\overline{\mathcal{M}}) \geq - \sum_{\sigma:K \rightarrow \mathbb{C}} \left( \log \|s\|_{\sigma,\text{sup}} + \log \frac{\|s_0(P)\|_{\sigma}}{\|s_0\|_{D,\text{inf}}} - e_1 \log r_1 - e_2 \log r_2 \right. \\ \left. + e_1 \log \|(dt_1)(P_1)\|_{\text{Ar},\sigma} + e_2 \log \|(dt_2)(P_2)\|_{\text{Ar},\sigma} \right).$$

We see that  $r_1 = r_2 = 1$ , and

$$\log \|s_0(P)\|_{\sigma}, \quad \log \|s_0\|_{D,\text{inf}}, \quad \log \|(dt_1)(P_1)\|_{\text{Ar},\sigma}, \quad \log \|(dt_2)(P_2)\|_{\text{Ar},\sigma}$$

are all bounded from both above and below uniformly with respect to  $P$ . Moreover, as  $(d_1, d_2, d)$  varies, we can choose the section  $s_0$  of

$$L = (d_1 - d)p_1^*\alpha + (d_2 - d)p_2^*\alpha + \mathcal{O}(d\Delta)$$

on every  $U$  as tensor powers of fixed sections of the individual line bundles  $p_1^*\alpha, p_2^*\alpha, \mathcal{O}(d\Delta)$ . This makes  $\log \|s_0(P)\|_\sigma$  and  $\log \|s_0\|_{D, \text{inf}}$  grows as  $O(d_1 + d_2 + d)$ . As a consequence, we have

$$\widehat{\text{deg}}(\overline{\mathcal{M}}) \geq - \sum_{\sigma: K \rightarrow \mathbb{C}} \log \|s\|_{\sigma, \text{sup}} + O(d_1 + d_2 + d) + O(e_1 + e_2).$$

## 5 Step 3: upper bound of the index

The goal of this section is to provide an upper bound of the index used in Theorem 4.1. This lies in the framework of the classical Dyson's lemma. Our exposition follows from [Voj89], which in turn is inspired by [EV84, Vio85]. Let us first introduce the framework of this section, which is slightly different from that of the previous sections.

Let  $K$  be an algebraically closed field of characteristic 0. Let  $C_1, C_2$  be smooth projective curves of genera  $g_1, g_2$  over  $K$  respectively. Let  $X = C_1 \times C_2$  be the product, and  $p_1: X \rightarrow C_1, p_2: X \rightarrow C_2$  be the projection morphisms. Denote by  $F_1$  (resp.  $F_2$ ) a closed fiber of  $p_1$  (resp.  $p_2$ ), which will be used for computing intersection numbers. Let  $D$  be a nonzero effective divisor on  $X$ . Denote by  $e(D)$  the maximum of  $\text{ord}_Y D$  over all prime divisors  $Y$  of  $X$  which are not fibers of  $p_1$  or  $p_2$ .

Let  $b_1, b_2$  be positive integers. For every point  $P = (P_1, P_2) \in X(K)$ , we define the *index of  $D$  at  $P$  of weight  $(b_1, b_2)$*  to be

$$\text{Ind}_P(D, (b_1, b_2)) := \text{Ind}_P(f, (b_1, b_2)),$$

where  $f$  is a local generator of the ideal sheaf  $\mathcal{I}_D$  at  $P$ , i.e. the stalk  $\mathcal{I}_{D, P}$ , and the index on the right is defined by (1) (here we view  $f$  as a local regular function in  $\mathcal{O}_{X, P}$  via the inclusion  $\mathcal{I}_D \subset \mathcal{O}_X$ ). Note that  $\text{Ind}_P(D, (b_1, b_2))$  does not depend on the choice of the local generator  $f$ .

Let  $Q_1, \dots, Q_m$  be distinct  $K$ -points on  $X = C_1 \times C_2$ . For  $i = 1, 2$ , denote by

$$m_i = m_i(Q_1, \dots, Q_m)$$

the number of different points in  $p_i(Q_1), \dots, p_i(Q_m)$ . For  $k = 1, \dots, m$ , let

$$I_k := \text{Ind}_{Q_k}(D, (b_1, b_2))$$

denote the index.

Finally, denote a function  $V: [0, \infty) \rightarrow \mathbb{R}$  by

$$V(a) := \text{vol}\left(\{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y \leq 1, x + y \leq a\}\right).$$

Here the right-hand side is the area in the Euclidean plane.

Finally, the main theorem of this section is the following theorem from [Voj89, Thm. 0.4].

**Theorem 5.1** (generalized Dyson's lemma). *Assume that  $m_1 = m_2 = m$ . Assume that  $b_1 \leq D \cdot F_2$  and  $b_2 \leq D \cdot F_1$  are positive real numbers. Then*

$$\sum_{k=1}^m V(I_k) \leq \frac{D \cdot D}{2b_1 b_2} + e(D) \frac{D \cdot F_1}{2b_1 b_2} \max\{2g_1 - 2 + m_1, 0\}.$$

The intuition for the inequality is as follows. Write  $D = \text{div}(s)$  for a nonzero section  $s$  of a line bundle  $L$  on  $X$ . Fix  $X, L, Q_1, \dots, Q_m, b_1, b_2$ , but vary  $s$  in  $H^0(X, L)$ . Fix positive real numbers  $I'_1, \dots, I'_m$  in the interval  $(0, 1)$ . Consider conditions for the existence of a section  $s$  satisfying  $\text{Ind}_{Q_k}(\text{div}(s), (b_1, b_2)) \geq I'_k$ . These conditions become linear equations on the section  $s$  at  $Q_1, \dots, Q_m$ , and the number of linear equations is roughly  $\sum_{k=1}^m b_1 b_2 V(I'_k)$ . If these linear equations are linearly independent, then the existence of  $s$  is equivalent to

$$\sum_{k=1}^m b_1 b_2 V(I'_k) < h^0(X, L) + \text{error}.$$

If  $L$  is sufficiently ample, we expect

$$h^0(X, L) = \frac{1}{2} L^2 + \text{error}.$$

The combination of these two relations explains the inequality in Dyson's lemma.

## 5.1 First reduction: remove fibers

We first reduce Theorem 5.1 to the following case where  $D$  contains no fibers of  $p_1$  or  $p_2$  and the assumptions  $m_1 = m_2 = m$ ,  $b_1 \leq D \cdot F_2$ ,  $b_2 \leq D \cdot F_1$  are removed.

**Theorem 5.2.** *Assume that  $D$  does not contain any fiber of  $p_1$  or  $p_2$ . Then*

$$\sum_{k=1}^m \frac{1}{2} I_k^2 \leq \frac{D \cdot D}{2b_1 b_2} + e(D) \frac{D \cdot F_1}{2b_1 b_2} \max\{2g_1 - 2 + m_1, 0\}.$$

Let us prove Theorem 5.1 by Theorem 5.2. Let  $D$  be as in Theorem 5.1. If  $D$  contains a fiber of  $p_1$  or  $p_2$  which does not pass through any of  $Q_1, \dots, Q_m$ , then removing this fiber from  $D$  does not change the left-hand side of the inequality of the theorem while decrease the right-hand side of the inequality. Therefore, we can assume that  $D$  does not contain such fibers.

Write

$$\begin{aligned} D &= D' + \sum_{k=1}^m (x_k (p_1^{-1}(p_1(Q_k))) + y_k (p_2^{-1}(p_2(Q_k)))) \\ &= D' + \sum_{k=1}^m (x_k p_1(Q_k) \times C_2 + y_k C_1 \times p_2(Q_k)), \end{aligned}$$

where the effective divisor  $D'$  on  $X$  does not contain any fiber of  $p_1$  or  $p_2$ , and the multiplicities  $x_1, \dots, x_m, y_1, \dots, y_m$  are non-negative integers. Since  $m_1 = m_2 = m$ , i.e.  $p_1(Q_1), \dots, p_1(Q_m)$  are distinct and  $p_2(Q_1), \dots, p_2(Q_m)$  are also distinct, it is easy to see that the indices  $I_k = \text{Ind}_{Q_k}(D, (b_1, b_2))$  and  $I'_k = \text{Ind}_{Q_k}(D', (b_1, b_2))$  satisfy

$$I_k = I'_k + \frac{x_k}{b_1} + \frac{y_k}{b_2}.$$

By Theorem 5.2,

$$\sum_{k=1}^m \frac{1}{2} I_k^2 \leq \frac{D' \cdot D'}{2b_1 b_2} + e(D') \frac{D' \cdot F_1}{2b_1 b_2} \max\{2g_1 - 2 + m_1, 0\}.$$

Note  $e(D) = e(D')$ , and  $V(a) \leq a^2/2$  for all  $a \geq 0$ . By taking difference, it suffices to prove

$$\begin{aligned} \sum_{k=1}^m (V(I_k) - V(I'_k)) &\leq \frac{D \cdot D - D' \cdot D'}{2b_1 b_2} \\ &\quad + e(D) \frac{(D - D') \cdot F_1}{2b_1 b_2} \max\{2g_1 - 2 + m_1, 0\}. \end{aligned}$$

By the expression of  $D$  in terms of  $D'$ , we can view both sides of the inequality as functions in non-negative variables  $x_1, \dots, x_m, y_1, \dots, y_m$ . When the variables are 0, the inequality is an equality. Therefore, it suffices to prove the corresponding inequality for the partial derivatives of both sides. Note that the derivative  $V'(a)$  exists and  $V'(a) \leq 1$  for all  $a \geq 0$ . Then

$$\frac{\partial}{\partial x_k} \sum_{i=1}^m (V(I_i) - V(I'_i)) = \frac{d}{dx_k} V(I_k) = \frac{1}{b_1} V'(I_k) \leq \frac{1}{b_1}.$$

On the other hand,

$$\frac{\partial}{\partial x_k} \left( \frac{D \cdot D - D' \cdot D'}{2b_1 b_2} \right) = \frac{2D \cdot F_1}{2b_1 b_2} \geq \frac{1}{b_1}.$$

Here we have used the assumption  $D \cdot F_1 \geq b_2$ . This proves the inequality for the partial derivative with respect to  $x_k$ , and that for  $y_k$  is similar. This completes the first reduction process.

## 5.2 Second reduction: trivialize weight

Now we reduce Theorem 5.2 to the case  $b_1 = b_2 = 1$  by the covering trick of [Voj89, §3]. The proof of [Voj89] applies the Kodaira–Parshin construction to construct ramified covering of  $C_i$ . Here we use a topological method instead, which works best over  $\mathbb{C}$ .

**Lemma 5.3.** *Let  $C$  be a smooth complex projective curve of genus  $g$ . Let  $n > 1$  be a positive integer. Let  $S$  be a finite set of points of  $C$ . In the case  $g = 0$ , we further assume that  $|S| \geq 2$ . Then there is a smooth projective complex curve  $C'$  with a finite and flat morphism  $f : C' \rightarrow C$  such that any point  $P' \in C'$  is ramified over  $C$  if and only if  $f(P') \in S$ , and in that case, the ramification index of  $P'$  is exactly  $n$ .*

*Proof.* If  $S = \emptyset$ , then  $g > 0$ , and the result is elementary in algebraic topology. Roughly speaking, an unramified cyclic covering of  $C$  of order  $n$  is given by a surjection  $\pi_1(C) \rightarrow \mathbb{Z}/n\mathbb{Z}$ , or equivalently a surjection  $H_1(C, \mathbb{Z}) \rightarrow \mathbb{Z}/n\mathbb{Z}$  by taking abelianization. There are plenty of such maps as  $H_1(C, \mathbb{Z}) \cong \mathbb{Z}^{2g}$ .

Now we assume that  $|S| \geq 1$ . By taking an unramified covering  $h : C_2 \rightarrow C$  of degree 2, and replacing  $(C, S)$  by  $(C_2, h^{-1}(S))$ , we can assume further that  $|S| \geq 2$ . By factoring  $n$ , we can assume that  $n$  is a prime number.

For every  $P \in S$ , take a closed disc  $D_P$  in  $C$  with center  $P$  of positive radius such that these discs are disjoint. There is a basic exact sequence

$$0 \longrightarrow \mathbb{Z} \xrightarrow{\iota_1} \bigoplus_{P \in S} H_1(D_P \setminus P, \mathbb{Z}) \xrightarrow{\iota_2} H_1(C \setminus S, \mathbb{Z}) \longrightarrow H_1(C, \mathbb{Z}) \longrightarrow 0.$$

Here  $H_1(D_P \setminus P, \mathbb{Z}) \cong \mathbb{Z}$  is generated by the boundary of  $D_P$ , and the arrow  $\iota_1$  is the diagonal map under this identification. Then  $H_1(C \setminus S, \mathbb{Z}) \cong \mathbb{Z}^{2g+|S|-1}$ .

Let  $H_1(C \setminus S, \mathbb{Z}) \rightarrow \mathbb{Z}/n\mathbb{Z}$  be a surjection whose kernel does not contain the image of  $H_1(D_P \setminus P, \mathbb{Z})$  for any  $P \in S$ . The surjection induces an unramified cyclic covering  $f^\circ : U' \rightarrow C \setminus S$  of degree  $n$ . Let  $C'$  be the unique smooth compactification of  $U'$ . Then we obtain a possibly ramified covering  $f : C' \rightarrow C$ . We will prove that this covering satisfies the requirement.

By the choice, the surjection  $H_1(C \setminus S, \mathbb{Z}) \rightarrow \mathbb{Z}/n\mathbb{Z}$  does not factor through  $H_1(C \setminus (S \setminus P), \mathbb{Z}) \rightarrow \mathbb{Z}/n\mathbb{Z}$  for any  $P \in S$ . In other words, the map  $f : C' \rightarrow C$  is ramified above  $P$ . As  $f^\circ$  is Galois, the Galois action extends to  $C'$ . Then the ramification indices of all points in  $f^{-1}(P)$  are equal and divides  $n$ , so they must be equal to  $n$  by the assumption that  $n$  is prime. This finishes the proof.  $\square$

Now we are ready to reduce Theorem 5.2 to the case  $b_1 = b_2 = 1$ . It suffices to reduce it to the case  $b_1 = b_2$ , since in the later case both sides of the inequality are quadratic in  $b_1^{-1}$ .

If  $g_i = 0$  and  $m_i = 1$  for some  $i = 1, 2$ , then we can just add a point  $Q_{m+1}$  of  $X \setminus |D|$  to the set  $\{Q_1, \dots, Q_m\}$  such that  $m_1 \geq 2$  and  $m_2 \geq 2$ . This does not change both sides of the inequality. Therefore, we can assume that for every  $i = 1, 2$ , we have either  $g_i > 0$  or  $m_i \geq 2$ .

By the Lefschetz principle, we can assume that  $(C, D, Q_1, \dots, Q_m)$  are defined over a finitely generated fields of characteristic 0, and by base change we can assume that they are actually defined over  $\mathbb{C}$ . This reduce the problem to the case  $K = \mathbb{C}$ .

Now we are ready to apply Lemma 5.3. Set  $(n_1, n_2) = (b_2, b_1)$ . By the lemma, for  $i = 1, 2$ , there is a covering  $f_i : C'_i \rightarrow C_i$  which is unramified outside  $\{p_i(Q_1), \dots, p_i(Q_m)\}$  and the ramification index of points above  $\{p_i(Q_1), \dots, p_i(Q_m)\}$

are equal to  $n_i$ . Let  $X = C'_1 \times C'_2$ , and let  $f: X' \rightarrow X$  the induced map. Denote by  $Q'_1, \dots, Q'_{m'}$  the points of  $f^{-1}(Q_1, \dots, Q_m)$ .

Now we claim that Theorem 5.2 for  $(X, (Q'_1, \dots, Q'_{m'}), f^*D, (b_1n_1, b_2n_2))$  implies that for  $(X, (Q_1, \dots, Q_m), D, (b_1, b_2))$ . In fact, we have the relation

$$\text{Ind}_{Q'}(D', (b_1n_1, b_2n_2)) = \text{Ind}_{Q_k}(D, (b_1, b_2))$$

for any  $Q' \in f^{-1}(Q_k)$ . The Riemann–Hurwitz formula gives the relation

$$2g'_1 - 2 + m'_1 = \deg(f_1)(2g_1 - 2 + m_1),$$

where  $g'_1$  denotes the genus of  $C'_1$ ,  $m'_1$  denotes the number of distinct points in  $p'_1(Q'_1, \dots, Q'_{m'})$ , and  $p'_1: X' \rightarrow C'_1$  denotes the projection. By these, we can check that the left-hand side (resp. right-hand side) of the inequality in Theorem 5.2 for  $(X, (Q'_1, \dots, Q'_{m'}), f^*D, (b_1n_1, b_2n_2))$  is equal to  $\deg(f_1) \deg(f_2)/(n_1, n_2)$  times that for  $(X, (Q_1, \dots, Q_m), D, (b_1, b_2))$ . This finishes the reduction by  $b_1n_1 = b_2n_2$ .

### 5.3 Proof of the essential case

In the above paragraphs, we reduced Theorem 5.2 to the case  $b_1 = b_2 = 1$ . Now we prove this case.

Denote by  $\psi: X' \rightarrow X$  the blowing-up of  $X$  along  $\{Q_1, \dots, Q_m\}$ . Denote by  $E_1, \dots, E_m$  the exceptional divisors above  $Q_1, \dots, Q_m$  respectively. Write

$$D = a_1D_1 + \dots + a_rD_r$$

in terms of distinct prime divisors  $D_1, \dots, D_r$  of  $X$ . Here the coefficients  $a_1, \dots, a_r$  are strictly positive. For  $j = 1, \dots, r$ , denote by  $D'_j$  the strict transform of  $D_j$ . Then the strict transform of  $D$  is just

$$D' = a_1D'_1 + \dots + a_rD'_r.$$

By a direct calculation, the index  $I_k = \text{Ind}_{Q_k}(D, (1, 1))$  is the multiplicity of  $E_k$  in the pull-back  $\psi^*D$ . In other words, we have

$$\psi^*D = D' + I_1E_1 + \dots + I_mE_m.$$

By the projection formula for  $\psi$ , the intersection number  $\psi^*D \cdot E_k = 0$ . It follows that

$$D'^2 = (\psi^*D - I_1E_1 + \dots + I_mE_m)^2 = D^2 - I_1^2 - \dots - I_m^2.$$

Then

$$I_1^2 + \dots + I_m^2 = D^2 - D'^2.$$

Therefore, Theorem 5.2 (for  $b_1 = b_2 = 1$ ) is equivalent to

$$D'^2 + e(D)(D \cdot F_1) \max\{2g_1 - 2 + m_1, 0\} \geq 0.$$

This is further equivalent to

$$D'^2 + e(D')(D' \cdot \psi^* F_1) \max\{2g_1 - 2 + m_1, 0\} \geq 0.$$

Here  $e(D') = e(D) = \max\{a_1, \dots, a_r\}$ .

We first prove the inequality holds for a single irreducible component. Namely, we have the following stronger statement, which is the reason for all other generalization.

**Theorem 5.4.** *Let  $Y$  be a integral closed curve of  $X'$  which is not contained in fibers of  $p'_1 : X' \rightarrow C_1$  or  $p'_2 : X' \rightarrow C_2$ . Then*

$$Y^2 + (Y \cdot \psi^* F_1)(2g_1 - 2 + m_1) \geq 0. \quad (3)$$

*Proof.* As in [Voj89], the key of the proof is the adjunction formula and the Riemann–Hurwitz formula.

Recall the adjunction formula

$$2h^1(\mathcal{O}_Y) - 2 = Y^2 + Y \cdot \omega_{X'}, \quad (4)$$

which is well-known if  $Y$  is smooth. The singular case is a combination of the expression

$$\chi(\mathcal{O}_Y) = \chi(\mathcal{O}_{X'}) - \chi(\mathcal{O}_{X'}(-Y)),$$

which comes from the additivity of the Euler characteristic  $\chi$ , and the Riemann–Roch theorem for the line bundle  $\mathcal{O}_{X'}(-Y)$  on  $X'$ .

Let  $f: \tilde{Y} \rightarrow Y$  be the normalization of  $Y$ . Then we have the following exact sequence

$$0 \longrightarrow \mathcal{O}_Y \longrightarrow f_* \mathcal{O}_{\tilde{Y}} \longrightarrow \mathcal{Q} \longrightarrow 0.$$

Since the quotient sheaf  $\mathcal{Q} = f_* \mathcal{O}_{\tilde{Y}} / \mathcal{O}_Y$  is supported on the singular locus of  $Y$ ,  $H^1(Y, \mathcal{Q}) = 0$ . Thus we see that

$$h^1(\mathcal{O}_Y) \geq h^1(\mathcal{O}_{\tilde{Y}}). \quad (5)$$

By the Riemann–Hurwitz formula,

$$2h^1(\mathcal{O}_{\tilde{Y}}) - 2 \geq \deg(\tilde{Y}/C_2)(2g_2 - 2) = (Y \cdot \psi^* F_2)(2g_2 - 2) = Y \cdot p_2'^* \omega_{C_2}. \quad (6)$$

Then (4), (5), and (6) imply that

$$Y^2 + Y \cdot \omega_{X'} \geq Y \cdot p_2'^* \omega_{C_2}. \quad (7)$$

Moreover, the canonical sheaf of blowing-up is given by

$$\omega_{X'} = \psi^* \omega_X + \mathcal{O}(E_1 + \dots + E_m) = p_1'^* \omega_{C_1} + p_2'^* \omega_{C_2} + \mathcal{O}(E_1 + \dots + E_m).$$

Thus (7) becomes

$$Y^2 + Y \cdot p_1'^* \omega_{C_1} + Y \cdot E_1 + \dots + Y \cdot E_m \geq 0.$$

In order to prove (3), it remains to prove that

$$Y \cdot E_1 + \cdots + Y \cdot E_m \leq (Y \cdot \psi^* F_1) m_1.$$

Finally, denote by  $\{P_1, \dots, P_{m_1}\}$  the distinct points of  $\{p_1(Q_1), \dots, p_1(Q_m)\}$ . Then the divisor  $F = p_1^{-1}(P_1) + \cdots + p_1^{-1}(P_{m_1})$  passing through all points of  $\{Q_1, \dots, Q_m\}$ . By the blowing-up process,  $\psi^* F - E_1 - \cdots - E_m$  is an effective divisor intersecting  $Y$  properly. This gives

$$Y \cdot (\psi^* F - E_1 - \cdots - E_m) \geq 0.$$

This completes the proof by the fact that  $F$  is numerically equivalent to  $m_1 F_1$ .  $\square$

Now we are ready to complete the proof of Theorem 5.2. Recall that right before Theorem 5.4, we are left to prove that

$$D'^2 + e(D')(D' \cdot \psi^* F_1) \max\{2g_1 - 2 + m_1, 0\} \geq 0.$$

Here

$$D' = a_1 D'_1 + \cdots + a_r D'_r, \quad e(D') = \max\{a_1, \dots, a_r\}.$$

As the proper intersection  $D'_j \cdot D'_{j'} \geq 0$  for  $j \neq j'$ , we have

$$D'^2 \geq \sum_{j=1}^r a_j^2 D_j'^2.$$

Since  $F_1$  is nef, we have

$$e(D')(D' \cdot \psi^* F_1) \geq \sum_{j=1}^r a_j^2 (D'_j \cdot \psi^* F_1).$$

Thus it suffices to prove that

$$\sum_{j=1}^r a_j^2 D_j'^2 + \sum_{j=1}^r a_j^2 (D'_j \cdot \psi^* F_1) \max\{2g_1 - 2 + m_1, 0\} \geq 0. \quad (8)$$

By Theorem 5.4,

$$D_j'^2 + (D'_j \cdot \psi^* F_1)(2g_1 - 2 + m_1) \geq 0.$$

Sum over  $j$ , and then (8) is proved.

*Remark 5.5.* The proof only works for base fields of characteristic 0 due to the application of the Riemann–Hurwitz formula. In fact, Dyson’s lemma does not hold for positive characteristics by the example in [Voj89, §4].

## 6 Completion of the proof

Now we are prepared to prove Theorem 2.6. Assume the setting of this theorem. Recall that

$$\bar{\mathcal{L}} = d\bar{\mathcal{L}}_0 = (d_1 - d)\tilde{p}_1^*\bar{\alpha} + (d_2 - d)\tilde{p}_2^*\bar{\alpha} + \mathcal{O}(d\bar{\Delta}),$$

where

$$d_1 = \delta_1 d, \quad d_2 = \delta_2 d.$$

By Theorem 3.1, for sufficiently large and sufficiently divisible integer  $d$ , there exists a nonzero element  $s \in \Gamma(\mathcal{X}, \mathcal{L})$  satisfying

$$-\log \|s\|_{\sigma, \text{sup}} \geq O\left(\frac{d\delta_1^2\delta_2}{\delta_1\delta_2 - g}\right) = O\left(\frac{d\delta_1}{\delta_1\delta_2 - g}\right)$$

for every embedding  $\sigma: K \rightarrow \mathbb{C}$ . Here we assume that  $\text{Ind}_P(s, (d_1, d_2))$  is attained at exponent  $(e_1, e_2)$ , and we use  $g < \delta_1\delta_2 < g + 1/4$  to simplify the error term.

Taking this section  $s$  in Theorem 4.1, we have

$$h_{\bar{\mathcal{L}}}(P)_{\mathbb{Q}} \geq -\frac{2g-2}{g}(e_1|P_1|^2 + e_2|P_2|^2) + O\left(\frac{d\delta_1}{\delta_1\delta_2 - g}\right) + O(e_1 + e_2).$$

We need the following bound of the index terms.

**Lemma 6.1.** *Let  $\delta_1, \delta_2$  be positive rational numbers satisfying*

$$\delta_1 > 2g\delta_2, \quad g < \delta_1\delta_2 < g + \frac{1}{4}.$$

*Then the exponent  $(e_1, e_2)$  satisfies*

$$e_1 + e_2 < c(\delta_1, \delta_2)\delta_1 d,$$

and

$$e_1|P_1|^2 + e_2|P_2|^2 \leq c(\delta_1, \delta_2) \cdot (\delta_1|P_1|^2 + \delta_2|P_2|^2) \cdot d.$$

Here

$$c(\delta_1, \delta_2) = \sqrt{\frac{2}{g}(\delta_1\delta_2 - g) + (2g-1)\frac{\delta_2}{\delta_1}} < 1.$$

By the lemma, the bound of the height becomes

$$h_{\bar{\mathcal{L}}}(P)_{\mathbb{Q}} \geq -\frac{2g-2}{g}c(\delta_1, \delta_2) \cdot (\delta_1|P_1|^2 + \delta_2|P_2|^2) \cdot d + O\left(\frac{d\delta_1}{\delta_1\delta_2 - g}\right).$$

It follows that

$$h_{\bar{\mathcal{L}}_0}(P)_{\mathbb{Q}} \geq -\frac{2g-2}{g}c(\delta_1, \delta_2) \cdot (\delta_1|P_1|^2 + \delta_2|P_2|^2) + O\left(\frac{\delta_1}{\delta_1\delta_2 - g}\right),$$

which proves Theorem 2.6.

Finally, we deduce Lemma 6.1 from Theorem 5.1 as follows.

*Proof of Lemma 6.1.* Set  $C_1 = C_2 = C$  and  $(b_1, b_2) = (d_1, d_2)$  to compute the index

$$I = \text{Ind}_P(s, (d_1, d_2)) = \text{Ind}_P(\text{div}(s), (d_1, d_2)).$$

Then Theorem 5.1 gives

$$V(I) \leq \frac{d_1 d_2 - g d^2}{d_1 d_2} + e(\text{div}(s)) \frac{1}{2d_1} (2g - 1).$$

Here  $e(\text{div}(s))$  is the maximal multiplicity of irreducible components  $Y$  of  $D$  which are not fibers of  $p_1$  or  $p_2$ . It follows that

$$d_2 = L \cdot F_1 \geq e(\text{div}(s)) D \cdot F_1 \geq e(\text{div}(s)).$$

As a consequence, we have

$$V(I) \leq \frac{d_1 d_2 - g d^2}{d_1 d_2} + \frac{(2g - 1)d_2}{2d_1} = \frac{\delta_1 \delta_2 - g}{\delta_1 \delta_2} + \frac{(2g - 1)\delta_2}{2\delta_1}.$$

Since  $\delta_1 > 2g\delta_2$  and  $g < \delta_1 \delta_2 < g + 1/4$ ,

$$\frac{\delta_1 \delta_2 - g}{\delta_1 \delta_2} + \frac{(2g - 1)\delta_2}{2\delta_1} < \frac{1}{4g} + \frac{2g - 1}{2 \cdot 2g} = \frac{1}{2}.$$

It follows that  $V(I) < 1/2$ , and thus

$$V(I) = \frac{1}{2} I^2 = \frac{1}{2} \left( \frac{e_1}{d_1} + \frac{e_2}{d_2} \right)^2.$$

As a consequence,

$$\left( \frac{e_1}{d_1} + \frac{e_2}{d_2} \right)^2 \leq \frac{2(\delta_1 \delta_2 - g)}{\delta_1 \delta_2} + (2g - 1) \frac{\delta_2}{\delta_1}.$$

By  $\delta_1 \delta_2 > g$ , we have

$$\frac{e_1}{d_1} + \frac{e_2}{d_2} < c(\delta_1, \delta_2),$$

and thus

$$e_1 \delta_2 + e_2 \delta_1 < c(\delta_1, \delta_2) \delta_1 \delta_2 d.$$

As  $\delta_1 > \delta_2$ , this implies

$$e_1 + e_2 < c(\delta_1, \delta_2) \delta_1 d.$$

On the other hand, the trivial bound

$$e_1 |P_1|^2 + e_2 |P_2|^2 \leq (e_1 \delta_2 + e_2 \delta_1) \cdot (\delta_1 |P_1|^2 + \delta_2 |P_2|^2) \cdot (\delta_1 \delta_2)^{-1}$$

gives

$$e_1 |P_1|^2 + e_2 |P_2|^2 \leq c(\delta_1, \delta_2) d \cdot (\delta_1 |P_1|^2 + \delta_2 |P_2|^2).$$

□

## References

- [AB95] A. Abbes and T. Bouche. Théorème de Hilbert-Samuel “arithmétique”. *Ann. Inst. Fourier (Grenoble)*, 45(2):375–401, 1995.
- [Ara74] S. Ju. Arakelov. An intersection theory for divisors on an arithmetic surface. *Izv. Akad. Nauk SSSR Ser. Mat.*, 38:1179–1192, 1974.
- [BG06] Enrico Bombieri and Walter Gubler. *Heights in Diophantine geometry*, volume 4 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2006.
- [Bom90] Enrico Bombieri. The Mordell conjecture revisited. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 17(4):615–640, 1990.
- [DGH21] V. Dimitrov, Z. Gao, P. Habegger, Uniformity in Mordell-Lang for curves, *Ann. of Math. (2)* 194 (2021), no. 1, 237–298.
- [Del87] P. Deligne. Le déterminant de la cohomologie. In *Current trends in arithmetical algebraic geometry (Arcata, Calif., 1985)*, volume 67 of *Contemp. Math.*, pages 93–177. Amer. Math. Soc., Providence, RI, 1987.
- [EV84] Hélène Esnault and Eckart Viehweg. Dyson’s lemma for polynomials in several variables (and the theorem of Roth). *Invent. Math.*, 78(3):445–490, 1984.
- [Fal83] Gerd Faltings. Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. *Invent. Math.*, 73(3):349–366, 1983.
- [Fal84] Gerd Faltings. Calculus on arithmetic surfaces. *Ann. of Math. (2)*, 119(2):387–424, 1984.
- [Fal91] Gerd Faltings. Diophantine approximation on abelian varieties. *Ann. of Math. (2)*, 133(3):549–576, 1991.
- [GS90] Henri Gillet and Christophe Soulé. Arithmetic intersection theory. *Inst. Hautes Études Sci. Publ. Math.*, (72):93–174 (1991), 1990.
- [GS92] Henri Gillet and Christophe Soulé. An arithmetic Riemann-Roch theorem. *Invent. Math.*, 110(3):473–543, 1992.
- [HS00] Marc Hindry and Joseph H. Silverman. *Diophantine geometry*, volume 201 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000.
- [IKM22] Hideaki Ikoma, Shu Kawaguchi, and Atsushi Moriwaki. *The Mordell conjecture—a complete proof from Diophantine geometry*, volume 226 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2022.
- [Kuh21] L. Kühne, *Equidistribution in Families of Abelian Varieties and Uniformity*. arXiv:2101.10272v2.

- [Liu02] Qing Liu. *Algebraic geometry and arithmetic curves*, volume 6 of *Oxford Graduate Texts in Mathematics*. Oxford University Press, Oxford, 2002. Translated from the French by Reinie Ern e, Oxford Science Publications.
- [LV20] Brian Lawrence and Akshay Venkatesh. Diophantine problems and  $p$ -adic period mappings. *Invent. Math.*, 221(3):893–999, 2020.
- [Mum65] D. Mumford. A remark on Mordell’s conjecture. *Amer. J. Math.* 87 (1965), 1007–1016.
- [Ser89] J-P. Serre. *Lectures on the Mordell-Weil theorem*. Friedr. Vieweg & Sohn, Braunschweig, 1989. Translated from the French and edited by Martin Brown from notes by Michel Waldschmidt.
- [Vio85] Carlo Viola. On Dyson’s lemma. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 12(1):105–135, 1985.
- [Voj89] Paul Vojta. Dyson’s lemma for products of two curves of arbitrary genus. *Invent. Math.*, 98(1):107–113, 1989.
- [Voj91] Paul Vojta. Siegel’s theorem in the compact case. *Ann. of Math. (2)*, 133(3):509–548, 1991.
- [Yua08] Xinyi Yuan. Big line bundles over arithmetic varieties. *Invent. Math.*, 173(3):603–649, 2008.
- [Yua21] Xinyi Yuan. Arithmetic bigness and a uniform Bogomolov-type result, arXiv:2108.05625, to appear in the *Annals of Mathematics*.
- [YZ21] X. Yuan, S. Zhang, Adelic line bundles on quasi-projective varieties, arXiv:2105.13587v6, to appear in the *Annals of Mathematics Studies*.
- [Zha92] Shou-Wu Zhang. Positive line bundles on arithmetic surfaces. *Ann. of Math. (2)*, 136(3):569–587, 1992.
- [Zha93] S. Zhang, Admissible pairing on a curve. *Invent. Math.* 112 (1993), 171–193.
- [Zha95a] Shou-Wu Zhang. Positive line bundles on arithmetic varieties. *J. Amer. Math. Soc.*, 8(1):187–221, 1995.
- [Zha95b] Shou-Wu Zhang. Small points and adelic metrics. *J. Algebraic Geom.*, 4(2):281–300, 1995.
- [Zha10] Shou-Wu Zhang. Gross-Schoen cycles and dualising sheaves. *Invent. Math.*, 179(1):1–73, 2010.

Xinyi Yuan  
 Address: BICMR, Peking University, Haidian District, Beijing 100871, China  
 Email: [yxy@bicmr.pku.edu.cn](mailto:yxy@bicmr.pku.edu.cn)