

MORDELL–LANG PLUS GEOMETRIC BOGOMOLOV

LONG LIU

ABSTRACT. We prove a geometric version of Poonen's "Mordell–Lang plus Bogomolov" theorem for semi-abelian varieties in characteristic zero. This is a generalization of the Mordell–Lang conjecture and the geometric Bogomolov conjecture.

CONTENTS

1. Introduction	1
2. Heights and approximation of a point	6
3. Approximation with bounded heights	9
4. A Vojta inequality over function fields	13
5. Proof of the main theorem	35
References	37

1. INTRODUCTION

We prove a geometric version of Poonen's "Mordell–Lang plus Bogomolov" theorem for semi-abelian varieties in characteristic zero. This is a generalization of the Mordell–Lang conjecture and the geometric Bogomolov conjecture. More precisely, we study subvarieties of a semi-abelian variety over a function field that contain a Zariski dense set of points arbitrarily close in canonical height to a finite-rank division group. We characterize them, after passage to the stabilizer quotient, as images of constant subvarieties translated by points of the group. A geometric Vojta inequality supplies the height bound needed for this classification.

1.1. Semi-abelian varieties and heights. Let k be an algebraically closed field of characteristic zero, and let K/k be a finitely generated regular extension with $\text{trdeg}(K/k) \geq 1$. Fix an algebraic closure $\overline{K}$. Fix a polarization $(\mathcal{B}, \mathcal{H})$ of K/k , where $\mathcal{B}$ is a normal projective k -variety with function field K and $\mathcal{H}$ is an ample line bundle on $\mathcal{B}$. Heights are normalized by degrees over K , so they are unchanged when the field of definition of a point is enlarged.

Let G be a semi-abelian variety over K , given by an exact sequence

$$1 \longrightarrow T \longrightarrow G \xrightarrow{\pi} A \longrightarrow 0,$$

Date: 29 September 2026.

2020 Mathematics Subject Classification. Primary 14G40; Secondary 11G50, 11J25, 14K12.

Key words and phrases. Mordell–Lang conjecture, geometric Bogomolov conjecture, semi-abelian varieties, canonical heights, function fields, Vojta inequality.

where T is a torus and A is an abelian variety. We use additive notation for the group law and write $[n]$ for multiplication by $n \in \mathbb{Z}$.

We use the height conventions of Luo–Yu [LY25, §1.1 and §2.5]. Put $t = \dim T$. After a finite extension of K , choose a splitting $T \simeq \mathbb{G}_m^t$ and the associated standard equivariant compactification $\bar{G}$. The projection extends to $\bar{\pi} : \bar{G} \rightarrow A$. Let D be the sum of the zero and infinity divisors in the projective-line factors of $\bar{G}$, and put $M = \mathcal{O}_{\bar{G}}(D)$. Choose a symmetric ample line bundle N on A , and set

$$L = M + \bar{\pi}^* N.$$

Here and below, tensor products of line bundles are written additively. If $t = 0$, take $\bar{G} = A$ and $M = 0$. The canonical height associated with these data is

$$\hat{h}_L(x) = \hat{h}_M(x) + \hat{h}_N(\pi(x)), \quad x \in G(\bar{K}),$$

where $\hat{h}_M$ is the canonical height associated with the boundary line bundle M and $\hat{h}_N$ is the Néron–Tate height associated with N on A . Both summands are nonnegative. Under multiplication by a positive integer, $\hat{h}_M$ is homogeneous of degree one and $\hat{h}_N$ is homogeneous of degree two.

We fix these data throughout the paper. For $\epsilon > 0$, put

$$B_\epsilon(G) = \{x \in G(\bar{K}) : \hat{h}_L(x) < \epsilon\}.$$

The set $B_\epsilon(G)$ depends on L . Heights from two standard choices are comparable by positive multiplicative constants; consequently, the condition below involving every $\epsilon > 0$ is independent of the choice. The comparison is recalled in Section 2. We say that a subvariety contains a dense set of small points if its intersection with $B_\epsilon(G)$ is dense for every $\epsilon > 0$.

Let $\Gamma_0 \subset G(\bar{K})$ be a finitely generated subgroup, and put

$$\Gamma = \Gamma_0^{\text{div}} := \{x \in G(\bar{K}) : [n]x \in \Gamma_0 \text{ for some } n \geq 1\}, \quad \Gamma_\epsilon = \Gamma + B_\epsilon(G).$$

Following Poonen [Poo99, §1], we call Γ the division group of Γ_0 and may visualize Γ_ϵ as a fattening of Γ , a "slab" in the topology on $G(\bar{K})$. For every $n \geq 1$, it satisfies $[n]^{-1}\Gamma = \Gamma$, where the inverse image is taken in $G(\bar{K})$. In particular, Γ contains all torsion points. A subgroup Λ has finite rank if $\dim_{\mathbb{Q}}(\Lambda \otimes_{\mathbb{Z}} \mathbb{Q}) < \infty$. The groups Γ_0^{div} are exactly the finite-rank subgroups Λ satisfying $[n]^{-1}\Lambda = \Lambda$ for every $n \geq 1$.

1.2. The main theorem. Poonen [Poo99, Theorem 8 and Corollary 9] proved Mordell–Lang plus Bogomolov over number fields for division groups on almost split semi-abelian varieties, that is, varieties isogenous to a product of an abelian variety and a torus. Zhang [Zha00, Theorem 1.2] independently proved the abelian case by equidistribution of almost division points. Rémond [Rem03, Theorems 1.1 and 1.2] treated arbitrary semi-abelian varieties over $\bar{\mathbb{Q}}$ and finite-rank subgroups.

We formulate and prove the geometric analogue. Since geometric canonical heights vanish on constant points and their images under homomorphisms, constant subvarieties must occur in the conclusion even when they are not cosets.

Let $X \subset G_{\bar{K}}$ be a closed integral subvariety. Let $\text{Stab}_{G_{\bar{K}}}(X)$ be its stabilizer, and set

$$S = \text{Stab}_{G_{\bar{K}}}(X)^0, \quad q : G_{\bar{K}} \longrightarrow \tilde{G} = G_{\bar{K}}/S.$$

Here S is the connected stabilizer of X ; quotienting by S makes the stabilizer of $q(X)$ finite. We say that X is Γ -special if there exist $\gamma \in \Gamma$, a semi-abelian variety G_0 over k , a homomorphism $h : (G_0)_{\bar{k}} \rightarrow \tilde{G}$ with finite kernel, and a closed integral k -subvariety $Y \subset G_0$ such that

$$(1) \quad q(X) = q(\gamma) + h(Y_{\bar{k}}).$$

This is the notion used in Luo–Yu’s geometric Bogomolov theorem [LY25, Theorem 1.1], with the torsion translation replaced by an element of Γ .

There is an equivalent formulation using the semi-abelian Chow trace. Let

$$(T_X, \tau_X) = \text{Tr}_{\bar{k}/k}(\tilde{G}).$$

The group T_X is defined over k , and $\tau_X : (T_X)_{\bar{k}} \rightarrow \tilde{G}$ is universal for homomorphisms from constant semi-abelian varieties; see [Liu24, Corollary 2.4.4(2)] and [Liu26]. Then X is Γ -special if and only if there exist $\gamma \in \Gamma$ and a closed integral k -subvariety $Y' \subset T_X$ such that

$$(2) \quad q(X) = q(\gamma) + \tau_X(Y'_{\bar{k}}).$$

Section 2 proves the equivalence. In both formulations we first pass to the stabilizer quotient.

Theorem 1.1. *With the notation above, the following conditions are equivalent.*

(1) *For every $\epsilon > 0$, the set*

$$X(\bar{k}) \cap \Gamma_{\epsilon}$$

is Zariski dense in X .

(2) *X is Γ -special, equivalently it satisfies (1), or equivalently (2).*

When $\Gamma_0 = \{0\}$, the group Γ is the torsion subgroup, and Theorem 1.1 recovers Luo–Yu’s geometric Bogomolov theorem: translation by torsion preserves canonical heights. For a group of positive rank, the approximating elements of Γ can have unbounded height. Theorem 4.1 supplies the estimate needed to handle them. Even after this estimate, bounded height does not imply finiteness, since every constant point has height zero; Proposition 3.1 instead classifies the subvarieties on which points of bounded height approach Γ arbitrarily closely.

Induction on dimension gives a single height threshold and a finite family of special subvarieties, as in [Poo99, Introduction, Conjectures 1 and 2].

Corollary 1.2. *With the notation above, there are $\epsilon_0 > 0$ and a finite, possibly empty, family of integral Γ -special subvarieties $Z_1, \dots, Z_m \subset X$ such that*

$$X(\bar{k}) \cap \Gamma_{\epsilon_0} \subseteq \bigcup_{i=1}^m Z_i(\bar{k}).$$

If X is not Γ -special, all the Z_i may be taken to be proper subvarieties of X .

The method in this paper does not work in positive characteristic; see §5.3.

1.3. Related results. Mordell conjectured that a curve over $\mathbb{Q}$ of genus at least two has finitely many rational points. Faltings proved the number-field form [Fal83, Satz 7]: for a smooth projective geometrically integral curve C of genus at least two over a number field F , the set $C(F)$ is finite. A rational base point, when one exists, embeds C in its Jacobian. The Mordell–Weil theorem makes the Jacobian’s group of F -rational points finitely generated, placing Mordell’s conjecture in a more general intersection problem.

Lang [Lan60, p. 29] formulated this problem for tori and abelian varieties and asked for its extension to semi-abelian varieties. Faltings [Fal91, Fal94] proved the abelian case, Vojta [Voj96, Theorem 0.2 and pp. 133–134] the semi-abelian case for finitely generated groups, and McQuillan [McQ95, Introduction and §3] the finite-rank case. Thus the intersection of a subvariety of a complex semi-abelian variety with a finite-rank subgroup is a finite union of intersections of that subgroup with semi-abelian cosets contained in the subvariety; each coset can be translated by a point of the subgroup. This gives Mordell’s conjecture because a curve of genus at least two in its Jacobian contains no positive-dimensional abelian translate. The torsion case is the Manin–Mumford theorem. Han–Luo–Yu [HLY26, Theorem 1.1] give a bound for the number of cosets in terms of dimension, degree, and rank in characteristic zero. These are statements about exact intersections.

Over function fields of characteristic zero, Manin [Man63] and Grauert [Gra65, p. 132] proved the geometric Mordell theorem: a smooth projective geometrically integral non-isotrivial curve of genus at least two has finitely many K -rational points. Here a curve is non-isotrivial if it does not become defined over k after any finite extension of K . Constant curves have Zariski dense constant points, which explains the non-isotriviality assumption. Hrushovski [Hru96, Theorem 1.1] proved the relative function-field Mordell–Lang theorem in all characteristics; see [BBP18, §4] for its semi-abelian formulation. In characteristic zero, the subgroup may have finite rational rank. In characteristic $p > 0$, the cited formulation requires containment in the prime-to- p divisible hull of a finitely generated group: every point must have a multiple in that group with multiplier prime to p . Finite rational rank alone does not imply this condition.

Over number fields, Ullmo [Ull98] proved Bogomolov’s conjecture for curves in their Jacobians, and Zhang [Zha98] proved it for subvarieties of abelian varieties. The conclusion is that a subvariety with dense small points is a torsion translate of an abelian subvariety.

For geometric heights, constant subvarieties also occur. Gubler [Gub07, Theorem 1.1] proved the theorem when the abelian variety is totally degenerate at one place, using tropical equidistribution. Yamaki [Yam18, Theorem 1.5] reduced the general abelian conjecture to nowhere-degenerate abelian varieties with trivial trace; here nowhere degenerate means potentially good reduction at every divisorial place. Gao–Habegger [GH19, Theorem 1.1 and Appendix A] proved the abelian geometric Bogomolov theorem in characteristic zero and transcendence degree one. Cantat–Gao–Habegger–Xie [CGHX21, Theorem A] treated arbitrary transcendence degree in characteristic zero, and Xie–Yuan [XY22, Theorem 1.1] proved the theorem in arbitrary characteristic. Luo–Yu [LY25, Theorem 1.1] extended it to semi-abelian varieties. Their Example 1.2 explains why one must allow positive heights tending to zero: a special subvariety of a nonconstant extension need not have dense height-zero points.

The closest arithmetic estimate to our argument is Rémond’s Vojta inequality [Rem03, Theorem 4.1]. His approximation theorem [Rem03, Theorem 1.2] permits errors z satisfying

$$\widehat{h}_L(z) \leq \epsilon(1 + \widehat{h}_L(\gamma))$$

for the corresponding arithmetic canonical height, and gives finiteness outside positive-dimensional cosets. In the abelian case, Ge [Ge24, Theorems 1.2 and 1.2’] chooses the height threshold in terms of the dimension of the ambient variety and the degree of the subvariety, and bounds the number of cosets needed to cover the intersection exponentially in the rank of the group.

Moriwaki [Mor01, Theorem A] treats abelian varieties over finitely generated extensions of $\mathbb{Q}$ using big arithmetic polarizations; those heights include arithmetic contributions. Hultberg [Hul26, Theorems 2 and 4] proves Bogomolov for semi-abelian varieties over globally valued fields of characteristic zero and a gap principle over non-archimedean globally valued fields for small points on subvarieties with finite stabilizer whose differences generate the ambient group. The Vojta inequality below concerns geometric heights over the fixed algebraically closed constant field. For fixed geometric data, its constants are independent of the finite extension of K over which the points are defined, and the auxiliary comparison is uniform in the integer weights.

After completing the proof of Theorem 1.1 in August 2026, the author learned in September 2026 of an independent proof of a similar result by Ningjun Jiang. The proof presented here was developed independently.

1.4. Plan of proof. The implication from specialness to the density of approximations follows from geometric Bogomolov. For the converse, pass to the connected stabilizer quotient; the image of Γ remains a finite-rank division group. We are then reduced to a subvariety X with finite stabilizer.

First suppose that the approximating points on X have bounded height. Sections 2 and 3 use the seminorm

$$\rho(x) = \hat{h}_M(x) + \sqrt{\hat{h}_N(\pi(x))},$$

whose two summands have the same degree of homogeneity. Finite rank gives a finite covering of the relevant bounded part of Γ by sets of small ρ -diameter. One part contains approximations to a dense subset of X . Taking differences cancels its center. Thus, for each integer $m \geq 2$, the image of

$$X^m \longrightarrow G^{m-1}, \quad (x_1, \dots, x_m) \longmapsto (x_2 - x_1, \dots, x_m - x_1)$$

has dense small points. For sufficiently large m , the image has finite stabilizer. Geometric Bogomolov makes it constant up to torsion translation, and a fiber of the next difference map gives $X = a + Y_{\overline{K}}$ with Y defined over k .

Determining the translate requires more than geometric Bogomolov. A generically finite difference map on Y , defined over k , gives a height inequality without an additive error. It follows that

$$\inf_{\gamma \in \Gamma} \rho(a - \gamma) = 0.$$

The positive gap for nonzero heights over each fixed finite extension of K implies that the real kernel of ρ on a finite-dimensional rational subspace is defined over $\mathbb{Q}$. The identities $[n]^{-1}\Gamma = \Gamma$ then give $a = \gamma + z$ with $\gamma \in \Gamma$ and $\hat{h}_L(z) = 0$, as required.

To remove the height bound, let Z_X be the union of the positive-dimensional semi-abelian cosets contained in X . The Vojta inequality, Theorem 4.1, excludes tuples in $X \setminus Z_X$ whose heights grow rapidly and whose normalized differences are small. The two height summands require different normalizations. Covering the corresponding normalized images of Γ by finitely many small sets gives the height bound of Proposition 5.1. Since X has finite stabilizer, Z_X is a proper closed subset; removing it preserves density and reduces the theorem to the bounded-height case.

The proof of the Vojta inequality occupies Section 4. We first work over a curve. The geometric product estimate retains the Samuel multiplicities in the derivative intersections

and bounds the degree and height of each product factor. The comparison obtained from the auxiliary section then descends to products of smaller dimension with leading constants independent of the integer weights and the finite extensions of K over which the points are defined. Fixing those weights and a finite extension of K containing the points before letting the auxiliary degree tend to infinity removes the genus term in Siegel's lemma. Finally, one generic complete-intersection curve preserves normalized heights, up to a fixed factor, over every finite extension of K and gives the result for the original polarization.

1.5. Notation and terminology. A variety is an integral separated scheme of finite type over a field, and a subvariety is closed and integral. A subscript denotes extension of scalars. A constant variety over $\bar{K}$ is the base change of a variety over k . Products and stabilizers are taken over $\bar{K}$ unless another field is specified, and density means Zariski density. We use multiplicative notation only in the torus examples.

1.6. Acknowledgement and AI usage. The author is grateful to Professor Huayi Chen for his encouragement and thoughtful advice during the preparation of this manuscript.

The ideas are due to the author's brain. The author used ChatGPT to polish the paper and verify the mathematical correctness. The author thanks OpenAI's ChatGPT for assistance with the presentation of the manuscript.

2. HEIGHTS AND APPROXIMATION OF A POINT

We prove that a point approximated arbitrarily closely by Γ_0^{div} , with Γ_0 finitely generated, belongs to Γ_0^{div} modulo a point of height zero. The proof uses a seminorm associated with the height and a positive lower bound for nonzero heights over each fixed finite extension of K .

2.1. Canonical heights. Keep the line bundles M , N , and L chosen in the introduction. If h_M and h_N are corresponding Weil heights, Tate's limiting process gives, for $x \in G(\bar{K})$ and $a \in A(\bar{K})$,

$$\hat{h}_M(x) = \lim_{n \rightarrow \infty} \frac{h_M([n]x)}{n}, \quad \hat{h}_N(a) = \lim_{n \rightarrow \infty} \frac{h_N([n]a)}{n^2}.$$

Let $\bar{M}$ and $\bar{N}$ be the resulting canonically metrized adelic line bundles. Following Luo–Yu [LY25, §2.4–§2.5], put $\bar{L} = \bar{M} + \bar{\pi}^* \bar{N}$, so that $\hat{h}_L = \hat{h}_{\bar{L}}$. These heights are the geometric counterparts of those in [Poo99, §3 and Lemma 4]. For every integer n , they satisfy

$$(3) \quad \hat{h}_M([n]x) = |n| \hat{h}_M(x), \quad \hat{h}_N([n]a) = n^2 \hat{h}_N(a) \quad (n \in \mathbb{Z}).$$

To give the two summands the same scaling under multiplication, put

$$(4) \quad \rho(x) = \hat{h}_M(x) + \sqrt{\hat{h}_N(\pi(x))}.$$

Both summands are homogeneous of degree one and satisfy the triangle inequality. For $\hat{h}_M$, apply the geometric Weil-height inequality obtained from the effective-divisor inequality in [Rem03, Corollary 3.1] to $[n]x$ and $[n]y$, divide by n , and let n tend to infinity. For $\sqrt{\hat{h}_N}$, use the nonnegativity of the quadratic form $\hat{h}_N$. Hence

$$\rho(x+y) \leq \rho(x) + \rho(y), \quad \rho([n]x) = |n| \rho(x).$$

Each of $\hat{h}_M$, $\sqrt{\hat{h}_N \circ \pi}$, and ρ vanishes on torsion and induces a seminorm on $G(\bar{K}) \otimes_{\mathbb{Z}} \mathbb{Q}$. These seminorms extend continuously to finite-dimensional real scalar extensions: if $e_1, \dots, e_r$ is a

rational basis, the bound $\rho(\sum_i a_i e_i) \leq \sum_i |a_i| \rho(e_i)$ gives a unique continuous extension, and the same argument applies to the other two seminorms.

The elementary bounds

$$(5) \quad \rho(x) \leq \widehat{h}_L(x) + \sqrt{\widehat{h}_L(x)}, \quad \widehat{h}_L(x) \leq \rho(x) + \rho(x)^2$$

allow us to replace $\widehat{h}_L$ by ρ when considering boundedness or convergence to zero. For a product with chosen line bundles $L_1, \dots, L_r$ defining the heights, we use the external tensor product, whose height is

$$\widehat{h}_{L_1 \boxtimes \dots \boxtimes L_r}(x_1, \dots, x_r) = \sum_{i=1}^r \widehat{h}_{L_i}(x_i).$$

Let G' be another semi-abelian variety, with abelian quotient A' . Choose its standard compactification, boundary line bundle M' , and symmetric ample line bundle N' on A' , and put $L' = M' + \overline{\pi}'^* N'$, where $\overline{\pi}'$ extends $G' \rightarrow A'$. For a homomorphism $f : G \rightarrow G'$, [LY25, Lemma 3.3, proof of (1)] gives $\widehat{h}_{L'}(f(x)) \leq C \widehat{h}_L(x)$, where $C > 0$ depends on f and the height data. If f has finite kernel, the reverse comparison also holds. This is part (2) of the same lemma for isogenies. In general, choose character bases so that the map on the tori has the form

$$(t_1, \dots, t_r) \mapsto (t_1^{d_1}, \dots, t_r^{d_r}, 1, \dots, 1), \quad d_i \geq 1.$$

The target boundary pulls back to $\sum_i d_i D_i$, where D_i is the sum of the zero and infinity divisors in the i th source factor. Nonnegativity of their canonical heights gives the reverse comparison for $\widehat{h}_M$. The map on abelian quotients is finite onto its image. Indeed, the inverse image of a positive-dimensional connected kernel would map with finite kernel to the target torus; it would therefore be affine, contradicting its positive-dimensional abelian quotient. The pullback of N' is consequently ample and gives the reverse comparison for $\widehat{h}_N \circ \pi$. Applying the upper bound to an isomorphism and its inverse shows that changing the torus coordinates preserves these comparisons.

We next identify the points of height zero. Let $(G^{\overline{K}/k}, \text{tr})$ be the geometric $\overline{K}/k$ -trace of $G_{\overline{K}}$; see [Liu24, Corollary 2.4.4(2)] and [LY25, §3.2]. Thus every homomorphism from a constant semi-abelian variety to $G_{\overline{K}}$ factors uniquely through tr . Its connected reduced kernel descends to k by [Liu24, Theorem 2.3.12]. The quotient by this kernel is constant and maps to $G_{\overline{K}}$, so universality gives a map from the quotient back to $G^{\overline{K}/k}$. Uniqueness in the universal property makes the composite with the quotient map the identity. The connected reduced kernel is therefore zero, and tr has finite kernel. In characteristic zero, every finite subgroup of a constant semi-abelian variety is defined over k . Its quotient, and hence the image of tr , is constant. Write H for this image with its induced k -structure.

The universal property also identifies the two descriptions of specialness in the introduction. For the quotient $\widetilde{G}$, it factors the homomorphism in (1) as $h = \tau_X \circ \nu_{\overline{K}}$, where $\nu : G_0 \rightarrow T_X$ is defined over k . The kernel of ν is finite because that of h is finite. Thus ν is finite onto its image, and $Y' = \nu(Y)$ is an integral closed k -subvariety of T_X . This gives (2). Conversely, (2) gives (1) with $G_0 = T_X$ and $h = \tau_X$.

By [LY25, Theorem 1.1] applied to a point,

$$(6) \quad \{x \in G(\overline{K}) : \widehat{h}_L(x) = 0\} = G(\overline{K})_{\text{tors}} + H(k).$$

2.2. A lower bound for positive heights. Lang–Néron and positive definiteness modulo the trace bound the positive values of $\widehat{h}_N$ from below over each fixed finite extension of K ; see [Con06, Theorems 7.1 and 9.15]. Above points of quadratic height zero, we express $\widehat{h}_M$ as an integral boundary degree divided by a fixed denominator. This proves the following analogue of [Poo99, Lemma 10(5)].

Lemma 2.1. *For every finite extension F/K , there is $a_F > 0$ such that*

$$P \in G(F), \quad \widehat{h}_L(P) < a_F \implies \widehat{h}_L(P) = 0.$$

Proof. A lower bound after a finite extension of F also applies to $G(F)$. We may thus enlarge F so that the splitting, compactification, and line bundles defining the heights are defined over it. Let $(A^{F/k}, \text{tr}_{F/k})$ be the F/k -trace of A_F . The group

$$A(F)/\text{tr}_{F/k}(A^{F/k}(k))$$

is finitely generated by Lang–Néron, and its Néron–Tate height is positive definite after real scalar extension; see [Con06, Theorems 7.1 and 9.15]. Modulo torsion, this group is a lattice, so its positive height values have a positive lower bound. Fix an integer $e \geq 1$ killing the torsion in the displayed quotient. If $\widehat{h}_L(P)$ is below the positive lower bound, then $\widehat{h}_N(\pi(P)) = 0$ and

$$[e]\pi(P) = \text{tr}_{F/k}(a) \quad \text{for some } a \in A^{F/k}(k).$$

The integer e is independent of P . We now bound the positive linear heights above the trace image by $[F:K]^{-1}$.

Put $A_0 = A^{F/k}$ and pull $G_F \rightarrow A_F$ back to $(A_0)_F$. The resulting extension by the split torus is determined by rigidified algebraically trivial line bundles, hence by finitely many points $\eta_j \in A_0^\vee(F)$. Let $\mathcal{B}_F$ be the normalization of $\mathcal{B}$ in F , and let $\mathcal{H}_F$ be the pullback of $\mathcal{H}$. Properness of $A_0^\vee$ makes every rational map $\mathcal{B}_F \dashrightarrow A_0^\vee$ defined by an η_j regular at the codimension-one points. All these maps are therefore morphisms on one open subset $U \subset \mathcal{B}_F$ whose complement has codimension at least two.

The rigidified Poincaré bundles give the pulled-back extension and its standard compactification over $A_0 \times U$. Let D be the sum of the zero and infinity divisors in its projective-line factors. The multiplication morphism satisfies

$$(7) \quad [n]^* D = nD \quad (n \geq 1).$$

This follows from the rigidified identity $[n]^* Q \simeq Q^{\otimes n}$ for each algebraically trivial extension bundle Q : in the associated projective-line bundle, the fiber coordinates are raised to their n th powers.

A point P' of the pulled-back extension above $a \in A_0(k)$ gives a rational section over U . It extends at codimension-one points by properness of the compactification. Its contact with D is an effective integral Weil divisor $D_{P'}$ on $\mathcal{B}_F$. Its model height is

$$(8) \quad \frac{\deg_{\mathcal{H}_F}(D_{P'})}{[F:K]}, \quad \deg_{\mathcal{H}_F}(D_{P'}) = \mathcal{H}_F^{\dim \mathcal{B}-1} \cdot D_{P'}.$$

The generic-fiber boundary is the pullback of the chosen boundary on G , so its model height differs from the corresponding Weil height by a bounded function. By (7), the model height scales exactly by n under $[n]$. Dividing by n and taking the limit therefore identifies (8) with $\widehat{h}_M$ at the image of P' in $G(F)$. A nonzero effective integral divisor has positive integral $\mathcal{H}_F$ -degree. Hence this height is either zero or at least $[F:K]^{-1}$.

Apply this to the lift of $[e]P$ above a . Since $\hat{h}_M([e]P) = e\hat{h}_M(P)$, a point with $\hat{h}_N(\pi(P)) = 0$ has either zero linear height or linear height at least $(e[F:K])^{-1}$. Choose a_F smaller than this bound and the positive gap for $\hat{h}_N \circ \pi$. Then $\hat{h}_L(P) < a_F$ forces both summands to vanish. $\square$

The height gap implies that the real kernel of ρ is rational on every finitely generated subgroup. The lattice argument is the seminorm analogue of [Con06, Lemma 10.1].

Lemma 2.2. *Let $\Lambda \subset G(\bar{K})$ be a finitely generated subgroup. The kernel of ρ on $\Lambda \otimes_{\mathbb{Z}} \mathbb{R}$ is defined over $\mathbb{Q}$ with respect to $\Lambda \otimes_{\mathbb{Z}} \mathbb{Q}$.*

Proof. Choose a finite extension F/K over which generators of Λ are defined. Lemma 2.1 and (5) give a positive lower bound for the nonzero ρ -values on $\Lambda/\Lambda_{\text{tors}}$. Its image in $(\Lambda \otimes_{\mathbb{Z}} \mathbb{R})/\ker \rho$ is therefore discrete. Since the image spans the quotient, it is a lattice of rank equal to the dimension of the quotient. The kernel of the map from $\Lambda/\Lambda_{\text{tors}}$ to this lattice consequently has rank $\dim_{\mathbb{R}} \ker \rho$ and spans $\ker \rho$. $\square$

Applying this rationality statement to a point and a finite set of generators gives the following analogue of [Poo99, Lemma 14]. Here the term of height zero includes constant points of the trace, as in (6).

Corollary 2.3. *Let $\Gamma \subset G(\bar{K})$ be a subgroup of finite rank satisfying $[n]^{-1}\Gamma = \Gamma$ for every $n \geq 1$. If $a \in G(\bar{K})$ satisfies*

$$\inf_{\gamma \in \Gamma} \rho(a - \gamma) = 0,$$

then $a = \gamma + z$ for some $\gamma \in \Gamma$ and some point z with $\hat{h}_L(z) = 0$.

Proof. Choose $g_1, \dots, g_r \in \Gamma$ whose images span $\Gamma \otimes_{\mathbb{Z}} \mathbb{Q}$, and apply Lemma 2.2 to the group generated by $a, g_1, \dots, g_r$. Quotient its real scalar extension by the kernel of ρ . The approximation hypothesis places the image of a in the closure of the span of the images of the g_i . This span is finite-dimensional, hence closed. Both it and the kernel of ρ are rational. Solving the corresponding linear equations over $\mathbb{Q}$ gives rational numbers c_i such that

$$\rho\left(a - \sum_i c_i g_i\right) = 0$$

in the rational scalar extension. Choose an integer $n \geq 1$ with $nc_i \in \mathbb{Z}$. Multiplication by n is surjective on $G(\bar{K})$, so there is $\gamma \in G(\bar{K})$ with $[n]\gamma = \sum_i nc_i g_i$. Since $[n]\gamma \in \Gamma$ and $[n]^{-1}\Gamma = \Gamma$, we have $\gamma \in \Gamma$. Homogeneity of ρ now yields $\rho(a - \gamma) = 0$, and (5) finishes the proof. $\square$

3. APPROXIMATION WITH BOUNDED HEIGHTS

We characterize approximation by a finite-rank subgroup when the approximating points have bounded height. Over $\bar{\mathbb{Q}}$, Rémond [Rem03, §5, pp. 209–211] combines a finite covering argument with Bogomolov's theorem. Over a function field, the conclusion involves a constant subvariety of the trace. Retain H and ρ from Section 2, and put $\Lambda_\epsilon = \Lambda + B_\epsilon(G)$ for a subgroup $\Lambda \subset G(\bar{K})$.

Proposition 3.1. *Let $\Lambda \subset G(\bar{K})$ be a subgroup of finite rank, and let $X \subset G_{\bar{K}}$ have finite stabilizer. The following are equivalent.*

- (1) *There is $R \geq 0$ such that, for every $\epsilon > 0$, the set*

$$\{x \in X(\bar{K}) \cap \Lambda_\epsilon : \hat{h}_L(x) \leq R\}$$

is dense in X .

(2) There are $a \in G(\overline{K})$ and a closed k -subvariety $Y \subset H$ such that

$$X = a + Y_{\overline{K}}, \quad \inf_{\lambda \in \Lambda} \rho(a - \lambda) = 0.$$

If $[n]^{-1}\Lambda = \Lambda$ for every $n \geq 1$, these conditions imply that X is Λ -special.

3.1. Difference maps. Under condition (1), we first show that a translate of X is defined over k . The finite covering argument of [Rem03, Lemma 5.1(3)] produces a dense subset close to one element of Λ . Taking differences cancels that element and gives small points on the image of the difference morphism.

Lemma 3.2. *Assume condition (1) of Proposition 3.1. There are points $\lambda_n \in \Lambda$, positive numbers $b_n \rightarrow 0$, and dense subsets $E_n \subset X(\overline{K})$ such that*

$$\rho(x - \lambda_n) \leq b_n \quad (x \in E_n).$$

Proof. Choose $\epsilon_n > 0$ with $\epsilon_n + \sqrt{\epsilon_n} \leq 1/n$. If $x \in X(\overline{K}) \cap \Lambda_{\epsilon_n}$ and $\hat{h}_L(x) \leq R$, choose $\lambda \in \Lambda$ with $\hat{h}_L(x - \lambda) < \epsilon_n$. Then

$$\rho(x - \lambda) \leq 1/n, \quad \rho(\lambda) \leq R + \sqrt{R} + 1.$$

The points λ satisfying the second inequality have bounded image in the finite-dimensional normed space $(\Lambda \otimes_{\mathbb{Z}} \mathbb{R}) / \ker \rho$. Cover this bounded image by finitely many balls of radius $1/n$ with centers in the image, and lift the centers to Λ . The dense set under consideration is then covered by finitely many sets of the form $\rho(x - \lambda) \leq 2/n$. Irreducibility of X forces one of these sets to be dense. Take it for E_n , its center for λ_n , and put $b_n = 2/n$. $\square$

For an integer $m \geq 2$, define the difference morphism

$$\alpha_m : X^m \longrightarrow G^{m-1}, \quad (x_1, \dots, x_m) \longmapsto (x_2 - x_1, \dots, x_m - x_1),$$

and let W_m be the closure of its image. Zhang [Zha98, Lemma 3.1] proves generic finiteness for sufficiently many factors in an abelian variety when the stabilizer is trivial. The following argument applies to semi-abelian varieties with finite stabilizer and also controls the stabilizer of W_m .

Lemma 3.3. *Let $m \geq 2$ satisfy $m > \dim G$, and suppose that X has finite stabilizer. Then*

- (1) α_m is generically finite onto W_m ;
- (2) W_m has finite stabilizer.

Proof. If X is a point, both assertions are immediate. Assume $d = \dim X > 0$, and put $g = \dim G$ and $F = \text{Stab}(X)(\overline{K})$. For $t \notin F$, the distinct irreducible varieties X and $X - t$ satisfy

$$\dim(X \cap (X - t)) \leq d - 1.$$

The incidence of tuples $(x_1, \dots, x_m, t)$ with $t \notin F$ and $x_i + t \in X$ for every i has dimension at most

$$g + m(d - 1) < md.$$

It cannot dominate X^m . Two tuples in a fiber of α_m differ by one common translation. A general tuple thus admits only translations in the finite group F , proving (1).

Let $v = (v_2, \dots, v_m)$ stabilize W_m , and put $v_1 = 0$. The constructible image of α_m contains a dense open subset of W_m . Therefore, for a general tuple $(x_i) \in X^m$, the point $\alpha_m((x_i)) + v$ has a preimage. Equivalently, there is $t \in G(\bar{K})$ such that

$$x_i + t + v_i \in X \quad (1 \leq i \leq m).$$

Suppose that $v \notin F^{m-1}$. Outside the finite set $\bigcup_i (F - v_i)$, every intersection $X \cap (X - t - v_i)$ has dimension at most $d - 1$, so the corresponding incidence again has dimension at most $g + m(d - 1)$. For a point t in that finite set, at least one $t + v_i$ lies outside F : otherwise $t \in F$ because $v_1 = 0$, and then all $v_i \in F$. The incidence over each such t consequently has dimension at most $md - 1$. Neither part dominates X^m , a contradiction. Thus $\text{Stab}(W_m)(\bar{K}) \subseteq F^{m-1}$, proving (2). $\square$

For sufficiently large m , finite stabilizer and geometric Bogomolov [LY25, Theorem 1.1] make W_m constant when it contains a dense set of small points. A fiber of $W_{m+1} \rightarrow W_m$ over a constant point then recovers a translate of X .

Lemma 3.4. *Suppose that X has finite stabilizer and every W_m contains a dense set of small points. Then $X = a + Y_{\bar{K}}$ for some $a \in G(\bar{K})$ and some closed k -subvariety $Y \subset H$.*

Proof. Choose $m > \dim G$ with $m \geq 2$. Lemma 3.3 gives a finite stabilizer for W_m . Since the trace of a product is the product of the traces, [LY25, Theorem 1.1] makes W_m a torsion translate of a constant subvariety of H^{m-1} . Diagonal tuples show that $0 \in W_m$, so the translating torsion point lies in H^{m-1} . It is constant in characteristic zero. Hence W_m is defined over k . The same holds for W_{m+1} .

The coordinate projections of W_{m+1} contain all differences $x - x'$ with $x, x' \in X(\bar{K})$. Hence $X - X \subset H$. Consider the projection

$$p: W_{m+1} \longrightarrow W_m$$

which forgets the last difference. Put $d = \dim X$. By Lemma 3.3, its source and target have dimensions $(m + 1)d$ and md . There is a nonempty open subset of W_m contained in the image of α_m , over which the fibers of α_m are finite and the fibers of p have dimension d .

The set $W_m(k)$ remains dense after extension to $\bar{K}$, since k is algebraically closed. Thus this open subset contains a point $w \in W_m(k)$ even though it need not be defined over k . Choose a preimage under α_m and denote its first coordinate by a .

Varying the last point in X^{m+1} shows that $X - a$ is contained in the fiber $p^{-1}(w)$, viewed in its last coordinate in H . It is closed, irreducible, and of dimension d , so it is an irreducible component of that fiber. The fiber is defined over the algebraically closed field k , and each of its geometric irreducible components is defined over k . This proves the assertion. $\square$

3.2. The translating point. To locate the translating point, we must deduce small heights on Y from small heights of differences. The usual inequality for a generically finite map has an additive error; see [Sil11, Theorem 1]. Since the heights here tend to zero, we need to remove that error. The argument using a nonzero section in [Sil11, §1] does so on models defined over k .

Lemma 3.5. *Let H be a semi-abelian variety over k , and let $Y \subset H$ have finite stabilizer. Choose a standard compactification $\bar{H}$ of $H_{\bar{K}}$ and a line bundle $L_H = M_H + \pi_H^* N_H$, where π_H extends the map from $H_{\bar{K}}$ to its abelian quotient, M_H is the boundary line bundle, and N_H is symmetric*

and ample on that quotient. There are an integer $r \geq 2$, a nonempty open subset $U \subset Y^r$ defined over k , and $C > 0$ such that

$$(9) \quad \sum_{i=1}^r \hat{h}_{L_H}(y_i) \leq C \sum_{i=2}^r \hat{h}_{L_H}(y_i - y_1) \quad ((y_1, \dots, y_r) \in U(\bar{K})).$$

Proof. Height comparison reduces the assertion to line bundles defined over k . Choose $r > \dim H$ with $r \geq 2$. Compactify Y^r and the image of its difference morphism over k , and let Z be the closure of the graph of the difference map. Let p and f be the projections to these compactifications, respectively. Let $\mathcal{L}$ and $\mathcal{N}$ be the restrictions of the external products of the ample line bundles defining the heights on the corresponding powers of a standard compactification of H defined over k . On the open parts, their geometric model heights are the sums in (9).

By Lemma 3.3, f is generically finite, so $f^*\mathcal{N}$ is big. The characterization of big line bundles in [Laz04, Corollary 2.2.7] gives positive integers C, b and a nonzero section of

$$(f^*\mathcal{N})^{\otimes bC} \otimes (p^*\mathcal{L})^{\otimes (-b)}$$

over k . Remove the zero divisor of this section and the loci where the graph projections do not represent the original morphism. This gives a nonempty open subset $U \subset Y^r$ defined over k .

Geometric model heights on constant models are exactly additive and functorial. They are nonnegative outside an effective constant divisor: pull its section back along the rational map from a normal model of a finite extension of K defining the point. Properness extends this map at codimension-one points, where the resulting divisor is effective and has nonnegative polarized degree. Applying this to the section above gives

$$bC h_{\mathcal{N}}(f(z)) - b h_{\mathcal{L}}(p(z)) \geq 0$$

for points above U . The constant model heights of the standard bundles equal their canonical heights: the multiplication identities already hold on those models. Dividing by b yields (9). $\square$

Proof of Proposition 3.1. Assume (1), and use the sets E_n , centers λ_n , and numbers b_n of Lemma 3.2. For every fixed m , the set E_n^m is dense in X^m , and its differences satisfy

$$\rho(x_i - x_1) \leq 2b_n.$$

Their image is dense in W_m . By (5), the product height of each image point is at most $(m-1)(2b_n + 4b_n^2)$. Consequently W_m contains a dense set of small points. Lemma 3.4 gives $X = a + Y_{\bar{K}}$ with $Y \subset H$ defined over k .

The stabilizer of Y is finite. Choose the line bundle L_H defining the height on H and choose r, U, C as in Lemma 3.5. For each n , density of E_n^r gives a tuple $(x_1, \dots, x_r)$ whose translate by $(-a, \dots, -a)$ belongs to U . Write $x_i = a + y_i$ and $z_i = x_i - \lambda_n$; these points depend on n . Then

$$y_i - y_1 = z_i - z_1, \quad \rho(z_i - z_1) \leq 2b_n.$$

By (5), these differences have $\hat{h}_L$ -height at most $2b_n + 4b_n^2$. Height comparison for $H \hookrightarrow G$ gives $\hat{h}_{L_H}(y_i - y_1) \rightarrow 0$. Since the constants in (9) are fixed, we obtain $\hat{h}_{L_H}(y_1) \rightarrow 0$. Since

$$a - \lambda_n = z_1 - y_1,$$

the triangle inequality, the bound $\rho(z_1) \leq b_n$, and height comparison give $\rho(a - \lambda_n) \rightarrow 0$. This proves (2).

Conversely, assume (2). The set $a + Y(k)$ is dense in X . Every $y \in Y(k)$ has height zero, so $\rho(a + y) = \rho(a)$, by applying the triangle inequality in both directions. Given $\epsilon > 0$, choose $\lambda \in \Lambda$ with $\rho(a - \lambda)$ small enough that $\rho(a - \lambda) + \rho(a - \lambda)^2 < \epsilon$. Then, for every $y \in Y(k)$,

$$\widehat{h}_L(a + y - \lambda) \leq \rho(a - \lambda) + \rho(a - \lambda)^2 < \epsilon.$$

Also $\widehat{h}_L(a + y) \leq \rho(a) + \rho(a)^2$, so (1) holds with $R = \rho(a) + \rho(a)^2$.

Finally, if $[n]^{-1}\Lambda = \Lambda$ for every $n \geq 1$, Corollary 2.3 gives $a = \lambda + z$ with $\lambda \in \Lambda$ and $\widehat{h}_L(z) = 0$. By (6), write $z = \tau + c$ with τ torsion and $c \in H(k)$. Absorb τ into λ , since Λ contains all torsion, and replace Y by $c + Y$. This gives the required special form. $\square$

4. A VOJTA INEQUALITY OVER FUNCTION FIELDS

For a closed subvariety $V \subset G_{\overline{K}}$, let Z_V be the union of the translates of positive-dimensional semi-abelian subvarieties contained in V . This is the *Mordell exceptional locus* of [Abr94, §0]. By [Abr94, Theorems 1 and 2], it is closed, and $Z_V = V$ implies $\dim \text{Stab}(V) > 0$.

The following theorem is the function-field analogue of [Rem03, Theorem 4.1]. It excludes tuples outside Z_V whose heights grow rapidly and whose normalized differences are small. The two normalizations reflect the linear growth of the boundary height and the quadratic growth of the abelian height under multiplication.

Theorem 4.1. *Suppose that $\dim V \geq 1$, and put $m = \dim V + 1$. There are positive constants c_1, c_2, c_3 , depending on G, V , the line bundles M, N , and the fixed polarization, such that no tuple $(x_1, \dots, x_m) \in (V \setminus Z_V)(\overline{K})^m$ satisfies all of*

- (1) $\widehat{h}_L(x_1) \geq c_3$ and $\widehat{h}_L(x_{i+1}) \geq c_2^2 \widehat{h}_L(x_i)$ for $1 \leq i < m$;
- (2) for $1 \leq i < m$,

$$\widehat{h}_M\left(\frac{x_i}{\widehat{h}_L(x_i)} - \frac{x_{i+1}}{\widehat{h}_L(x_{i+1})}\right) \leq c_1^{-1},$$

$$\widehat{h}_N\left(\frac{\pi(x_i)}{\sqrt{\widehat{h}_L(x_i)}} - \frac{\pi(x_{i+1})}{\sqrt{\widehat{h}_L(x_{i+1})}}\right) \leq c_1^{-2}.$$

The expressions in (2) are taken in the real scalar extension of the group generated by the points; the heights extend there by homogeneity and continuity. The constants are independent of the finite extensions of K containing the points.

We first work over the function field of a curve. Proposition 4.2, applied to the semi-abelian difference maps, gives an inequality with integer weights. Approximating inverse square roots of the point heights by ratios of positive integers then gives the two normalized differences in the theorem; the boundary differences use the squares of these integers. The comparison must therefore have constants independent of the weights and the fields containing the points. We prove it by auxiliary sections, differentiation, and an index estimate, as in [Yua25, §§3–5], with degree and height bounds for the smaller product factors produced by the index estimate. Sections 4.7 and 4.8 give the semi-abelian application and the passage to higher transcendence degree.

4.1. Heights over a curve. Suppose $K = k(C)$, where C is a smooth projective curve over k . For a finite extension F/K , let C_F be the smooth projective curve with function field F , equipped with its finite morphism to C . The absolute values $|a|_v = \exp(-\text{ord}_v(a))$, indexed by the closed points v of C_F , are trivial on k^* and satisfy the product formula. For $z = (z_0 : \cdots : z_N) \in \mathbb{P}^N(F)$, put

$$h(z) = \frac{1}{[F : K]} \sum_v \log \max_j |z_j|_v.$$

This is the height $h_{\mathcal{O}(1)}$ for the constant model of the hyperplane line bundle. The product formula makes it independent of the homogeneous coordinates, and the normalization makes it invariant under finite extension of F .

For a nonzero polynomial, h denotes the height of its coefficient vector. For a finite family of polynomials, it denotes the height of the joint coefficient vector, whose v -adic norm is the largest absolute value of any coefficient in the family. For a projective variety Y , let $h(Y)$ be its Chow height, divided by $[F : K]$ but not by $\deg Y$. Equivalently, this is the corresponding hyperplane intersection on the closure over C_F . These projective heights are distinct from the canonical height $\hat{h}_L$ on G .

4.2. A uniform height comparison. The following proposition compares the height of an auxiliary line bundle with a weighted sum of projective heights. It is the function-field analogue of [Rem05, Theorem 1.2] with $\omega = 0$; [Dil20, Theorem 1.1] treats different projective factors over $\overline{\mathbb{Q}}$. All heights use the normalization of Section 4.1.

Proposition 4.2. *Let $X \subset \mathbb{P}^N$ be an integral projective variety of positive dimension, and let $L_X = \mathcal{O}_X(1)$. Fix integers $m \geq 2$, $t_1, t_2, r_0, \theta \geq 1$ and a real number $\delta \geq 1$. For a tuple $a = (a_1, \dots, a_m)$ of positive integers, put $|a| = \sum_i a_i$. Let $\pi : \mathcal{X} \rightarrow X^m$ be projective and birational, let $\mathcal{P}$ be very ample on $\mathcal{X}$, and put $\mathcal{N}_a = \pi^*(\boxtimes_i L_X^{a_i})$. Let $\mathcal{M}$ be a line bundle on $\mathcal{X}$, and let Σ be a family of r_0 global sections of $\mathcal{P} \otimes \mathcal{M}^{-1}$. Suppose there are injections*

$$\mathcal{P} \longrightarrow \mathcal{N}_a^{t_1}, \quad \mathcal{P} \otimes \mathcal{M}^{-1} \longrightarrow \mathcal{N}_a^{t_2}.$$

Let U be an open set on which π and both injections are isomorphisms. Identify U with its image under π , and fix $x = (x_1, \dots, x_m) \in U$. Assume the following.

- (1) *The line bundle $\mathcal{M}$ is nef, and Σ generates $\mathcal{P} \otimes \mathcal{M}^{-1}$.*
- (2) *In an embedding given by $\mathcal{P}$, the first injection sends the coordinate sections to monomials with coefficient 1. The second sends Σ to polynomials $P_1, \dots, P_{r_0}$ of multidegree $t_2 a$, whose joint coefficient vector satisfies*

$$h(P_1, \dots, P_{r_0}) \leq \delta |a|.$$

- (3) *For every product $Y = \prod_i Y_i \subset X^m$ containing x , with each Y_i integral, its strict transform $\tilde{Y}$ satisfies*

$$(10) \quad (\mathcal{M}^{\dim Y} \cdot \tilde{Y}) \geq \theta^{-1} \prod_i a_i^{\dim Y_i}.$$

Define

$$h_{\mathcal{M}}(x) = h_{\mathcal{P}}(x) - h(\Sigma(x)),$$

where $h_{\mathcal{P}}$ uses the stated coordinate sections. There are constants $c_1, c_2, c_3 > 0$, depending only on $X, N, m, t_1, t_2, r_0, \theta, \delta$, such that

$$(11) \quad \sum_i a_i h(x_i) \leq c_1 h_{\mathcal{M}}(x)$$

whenever $a_i/a_{i+1} \geq c_2$ for $1 \leq i < m$ and $h(x_i) \geq c_3$ for all i . The constants are uniform in $\mathcal{X}$, the weights, the point, and their fields of definition.

The auxiliary section will belong to $\mathcal{M}^d \otimes \mathcal{N}_a^{-d\varepsilon}$, with $\varepsilon > 0$ rational and d a large integer. If its first nonzero derivative at x has small weighted order, the product formula gives the height comparison. Otherwise, the index estimate produces a hypersurface through one of the x_i , with controlled degree and height. We choose a product through x of least dimension subject to bounds preserved by this construction, so the second alternative is impossible. Section 4.6 chooses these bounds and completes the proof.

4.2.1. Finite projections. Fix the data of Proposition 4.2. We work on a product $Y = \prod_i Y_i \subset X^m$ containing x , with each Y_i integral, so that the estimates will also apply after replacing a factor by a proper subvariety. Put

$$u_i = \dim Y_i, \quad D_i = \deg Y_i, \quad u = \sum_i u_i, \quad D = \prod_i D_i.$$

We use this notation until Section 4.6. Let $\tilde{Y}$ be the strict transform of Y in $\mathcal{X}$, and let W_{ij} be the original homogeneous coordinates on the i th factor. We retain these coordinates because the coordinate sections of $\mathcal{P}$ map to monomials in them. Omit coordinates identically zero on Y_i , together with the resulting zero projective coordinates and generating sections. If a remaining coordinate vanishes at x_i , it already gives a proper hyperplane section through x_i with coefficient height zero.

Finite linear projections provide local parameters for differentiation wherever they are étale. Their equations also control the coefficients of the auxiliary section. The following construction is the geometric-height counterpart of [Rem05, Proposition 2.2 and Lemma 2.3]; constant coefficients have absolute value at most 1 at every place.

Lemma 4.3. *For every i , there are constant linear forms $V_{i0}, \dots, V_{iu_i}$ defining a finite projection $\rho_i : Y_i \rightarrow \mathbb{P}^{u_i}$ of degree D_i , with $V_{i0}(x_i) \neq 0$, having the following properties.*

- (1) *Each remaining W_{ij} satisfies a monic homogeneous equation $Q_{ij}(V_{i0}, \dots, V_{iu_i}, W_{ij}) = 0$ of degree D_i . If $u_i < N$, the same holds for an additional constant linear form whose ratio to V_{i0} generates the function-field extension. Let B_i be the joint coefficient vector of these equations, including their common leading coefficient 1. Then*

$$(12) \quad h(B_i) \leq h(Y_i).$$

- (2) *Write $Q_{ij} = Q_{ij,1}^{b_{ij}}$, where $Q_{ij,1}$ is monic and irreducible. The discriminants of these irreducible equations and of the equation for the additional linear form are nonzero homogeneous polynomials Δ in the projection coordinates satisfying*

$$(13) \quad \deg \Delta \leq 2D_i^2, \quad h(\Delta) \leq 2D_i h(B_i).$$

If these discriminants do not vanish at $\rho_i(x_i)$, then ρ_i is étale at x_i and all coordinate derivative denominators are nonzero there. For $Y_i = \mathbb{P}^N$, one may take the identity projection and linear coordinate equations.

Proof. Generic projection and the primitive element theorem give a nonempty open set of choices for the projection and the additional linear form. We can also require $V_{i0}(x_i) \neq 0$. The coefficients can be chosen in k , since constant points are Zariski dense in the parameter space.

Specialize the Chow form of Y_i to the u_i hyperplanes $V_{i\ell} - T_\ell V_{i0}$ and the hyperplane $W_{ij} - ZV_{i0}$. The coefficient of Z^{D_i} is obtained by taking V_{i0} as the last hyperplane; it is a nonzero scalar independent of j . Divide by this common scalar and homogenize. This gives Q_{ij} , and the same construction applies to the additional linear form. The coefficients are constant linear combinations of the Chow coefficients, up to one common scalar. The ultrametric inequality bounds their local norms, and the product formula removes the scalar, proving (12).

Multiplicativity of the Gauss norm gives $|Q_{ij,1}|_v = |Q_{ij}|_v^{1/b_{ij}} \leq |B_i|_v$, since B_i contains 1. Discriminants are resultants of total coefficient degree at most $2D_i$, which proves (13) for the coordinate equations and the equation of the additional linear form. If the discriminant of the latter equation does not vanish at $\rho_i(x_i)$, the localized algebra generated by the ratio of that form to V_{i0} is étale over the base, hence normal. The coordinate algebra of Y_i is integral over this algebra and has the same fraction field, so the two algebras agree there. Thus ρ_i is étale at x_i . Nonvanishing of the other discriminants gives the asserted derivative denominators. $\square$

Choose these projections over a finite extension F/K defining x , Y , the line bundles, the sections, and the projection equations. Enlarge F to contain the finitely many roots introduced in Section 4.4. The field F and the weights a_i remain fixed while the auxiliary degree tends to infinity. Hypersurface intersections may later require further extensions, under which the normalized heights remain unchanged.

4.3. Existence of an auxiliary section. Assume $u > 0$ and fix an upper bound $D_* \in \mathbb{Z}$ for D . Put

$$\varepsilon = \frac{1}{4u\theta(t_1 m)^u D_*}, \quad \mathcal{Q}_d = \mathcal{M}^d \otimes \mathcal{N}_a^{-d\varepsilon},$$

where d runs through positive integers with $d\varepsilon \in \mathbb{Z}$. All line bundles below are restricted to $\tilde{Y}$. The intersection hypothesis ensures that subtracting $d\varepsilon \mathcal{N}_a$ from $d\mathcal{M}$ leaves a space of sections of order d^u . We use the dimension argument of [Rem05, Proposition 4.1], with the nef hypothesis required in [Dil20, proof of Proposition 3.1].

Lemma 4.4. *As d tends to infinity through positive integers with $d\varepsilon \in \mathbb{Z}$,*

$$(14) \quad h^0(\tilde{Y}, \mathcal{Q}_d) \geq \frac{d^u}{4\theta u!} \prod_i a_i^{u_i} + o(d^u).$$

Proof. Choose an effective Cartier divisor $E \in |\mathcal{N}_a|$ on $\tilde{Y}$ and put $q = d\varepsilon$. Successively restricting to E gives

$$h^0(d\mathcal{M} - q\mathcal{N}_a) \geq h^0(d\mathcal{M}) - q h^0(E, d\mathcal{M}|_E).$$

Here each $(d\mathcal{M} - j\mathcal{N}_a)|_E$ injects into $d\mathcal{M}|_E$: multiply by a section of $j\mathcal{N}_a|_E$ that avoids the associated points of E . Such a section exists because $\mathcal{N}_a$ is globally generated and the ground field is infinite. By the nef asymptotic Riemann–Roch formula [Laz04, Theorem 1.4.40 and Corollary 1.4.41], the two terms on the right have leading coefficients $\mathcal{M}^u/u!$ and

$\mathcal{M}^{u-1}E/(u-1)!$. On the possibly nonreduced divisor E , use the coherent-sheaf form of the same estimate and Riemann–Roch on its fundamental cycle; lower-dimensional terms are $o(d^{u-1})$. Hence

$$\liminf_{d \rightarrow \infty} \frac{u!}{d^u} h^0(\tilde{Y}, \mathcal{M}^d \mathcal{N}_a^{-d\varepsilon}) \geq \mathcal{M}^u - u\varepsilon \mathcal{M}^{u-1} \mathcal{N}_a.$$

For $u = 1$, the restriction term is simply the length of E . The injection into $\mathcal{N}_a^{t_1}$ is nonzero on $\tilde{Y}$, because it is an isomorphism at x . Thus $t_1 \mathcal{N}_a - \mathcal{P}$ is effective there. The class $\mathcal{P} - \mathcal{M}$ is globally generated. Intersecting these classes with the nef classes $\mathcal{M}$ and $\mathcal{N}_a$ gives

$$\mathcal{M}^{u-1} \mathcal{N}_a \leq t_1^{u-1} \mathcal{N}_a^u \leq (t_1 m)^u D \prod_i a_i^{u_i}.$$

Together with (10) and the choice of ε , this bounds the liminf below by $3(4\theta)^{-1} \prod_i a_i^{u_i}$. In particular, (14) holds. $\square$

As in [Rem05, Proposition 4.2], we express a section of $\mathcal{Q}_d$ by polynomials whose coefficients satisfy linear equations. Set $r = r_0(N+1)^m$ and let s_λ denote the r sections

$$\sigma^d \prod_i W_{i\ell_i}^{d\varepsilon a_i}, \quad \sigma \in \Sigma, \quad 0 \leq \ell_i \leq N,$$

which generate $\mathcal{P}^d \mathcal{M}^{-d} \mathcal{N}_a^{d\varepsilon}$. For large d , Serre vanishing allows us to choose a basis of $H^0(\tilde{Y}, \mathcal{P}^d)$ consisting of monomials in the projective coordinates. Write a family $\mathbf{F} = (F_\lambda)_{1 \leq \lambda \leq r}$ in this basis. It represents a section of $\mathcal{Q}_d$ precisely when the local quotients F_λ / s_λ agree, or equivalently when

$$F_\lambda s_\mu = F_\mu s_\lambda \quad (1 \leq \lambda, \mu \leq r).$$

Let N_d be the number of coefficient unknowns and q_d the dimension of the kernel of this linear system. We allow zero generating sections, so omitting coordinates does not change r .

Lemma 4.5. *The compatibility system has the following properties.*

(1) *Its kernel is naturally $H^0(\tilde{Y}, \mathcal{Q}_d)$, and*

$$(15) \quad \frac{N_d}{q_d} \leq 4\theta r (t_1 m)^u D + o(1).$$

(2) *It can be written as scalar linear equations over F whose joint coefficient vector $\mathcal{A}_d$ satisfies*

$$(16) \quad h(\mathcal{A}_d) \leq d \left(\delta |a| + (t_1 + t_2 + \varepsilon) \sum_i a_i h(B_i) \right) + O_{Y,a}(1).$$

Proof. Since the s_λ generate, every compatible family determines a section of $\mathcal{Q}_d$ by gluing its local quotients. Multiplication by the s_λ gives the inverse map. Thus $q_d = h^0(\tilde{Y}, \mathcal{Q}_d)$. The injection into $\mathcal{N}_a^{d\varepsilon}$ gives

$$N_d \leq r \frac{d^u}{u!} (t_1 m)^u D \prod_i a_i^{u_i} + o(d^u),$$

and Lemma 4.4 proves (15).

Since the coordinate sections of $\mathcal{P}$ map to monomials, each coefficient in a compatibility equation is a polynomial of the form $\pm W^\alpha P_j^d$ of multidegree $\gamma = d(t_1 + t_2 + \varepsilon)a$. Reduce this polynomial using the monic equations Q_{ij} . The result is a sum of monomials W^k with

$k_{ij} < D_i$, whose coefficients are polynomials in the projection coordinates. Each substitution decreases the total degree in the non-projection coordinates, so along any branch there are at most γ_i substitutions in factor i . Since $|B_i|_v \geq 1$, the ultrametric inequality bounds the norm of the entire coefficient family by

$$|(P_1, \dots, P_{r_0})|_v^d \prod_i |B_i|_v^{\gamma_i}.$$

There are only finitely many possible monomials W^k , independently of d . Express them in the function-field basis given by powers of the primitive elements, with exponent less than D_i in factor i . A single nonzero polynomial R_Y in the projection variables clears the denominators of all these expressions; write $R_Y W^k = S_k$ in this basis. Multiplying the compatibility equations by R_Y preserves their kernel. Substitution, followed by taking coefficients in the projection variables and powers of the primitive elements, gives scalar equations over F . The polynomials S_k are fixed as d varies, so their contribution to the height is bounded independently of d . Summing the local bounds proves (16). $\square$

Thunder's function-field Siegel lemma gives a nonzero solution in terms of the row heights. We use [Fuk10, Theorem 1.1 and (14)] in place of the number-field estimate [Rem05, Lemma 2.6]. Let g_F be the genus of C_F . A system with N unknowns and $s < N$ independent nonzero rows has a nonzero kernel vector z satisfying

$$(17) \quad h(z) \leq \frac{\sum_i h(\text{row } i)}{N - s} + \frac{2g_F + 1}{[F : K]}.$$

For this normalization, choose $O \in C_F(k)$ and, by Riemann–Roch, a nonconstant $t \in F$ whose only pole is at O , of order $e \leq g_F + 1$. Then $[F : k(t)] = e$. Since k is algebraically closed, the constant $m(F, k(t))$ in [Fuk10, (14)] equals e . With heights normalized by e , the cited basis bound gives a vector of logarithmic height at most the kernel's Plücker height divided by $N - s$, plus $(g_F - 1 + e)/e$. By Plücker duality [Fuk10, (21)] and the ultrametric determinant bound, the kernel height is at most the sum of the row heights. Multiplication by $e/[F : K]$ gives an additive term at most $2g_F/[F : K]$, and hence (17). If $s = 0$, a coordinate vector has height zero.

The genus term depends on F but not on d . Dividing by d with F fixed gives the following estimate.

Lemma 4.6. *For all sufficiently large positive integers d with $d\varepsilon \in \mathbb{Z}$, there is a nonzero section $s_d \in H^0(\tilde{Y}, \mathcal{Q}_d)$ represented by a family F_d of homogeneous polynomials in the chosen monomial basis such that*

$$(18) \quad \limsup_{d \rightarrow \infty} \frac{h(F_d)}{d} \leq 4\theta r(t_1 m)^u D \left(\delta|a| + (t_1 + t_2 + \varepsilon) \sum_i a_i h(B_i) \right).$$

The expression on the right is independent of F and linear in the weights. The lower bound on d may depend on the fixed data, including F and a .

Proof. Apply (17) to independent rows of the compatibility system. Each row has height at most $h(\mathcal{A}_d)$, so a nonzero kernel vector can be chosen with

$$h(F_d) \leq \frac{N_d}{q_d} h(\mathcal{A}_d) + \frac{2g_F + 1}{[F : K]}.$$

Divide by d and use (15) and (16). All data, including F and a , are fixed while d tends to infinity, and hence

$$(19) \quad \lim_{d \rightarrow \infty} \frac{2g_F + 1}{d[F:K]} = 0.$$

Neither this term nor $O_{Y,a}(1)/d$ contributes to the limit, which proves (18). $\square$

4.4. A height bound from differentiation. We next estimate the contribution of differentiation to the height inequality. The local input is [Rem05, Lemma 2.7, pp. 473–474]: its proof at a simple root, on pp. 474–475, uses only the Leibniz rule and the ultrametric inequality and applies at every place of F . All derivatives are taken over F .

Let $T = (T_1, \dots, T_u)$ be affine coordinates, and let $y(T)$ be a local algebraic function satisfying a polynomial relation

$$Q(T, y) = 0, \quad Q_Y(0, y_0) \neq 0,$$

where $y_0 = y(0)$ and $Q_Y = \partial Q / \partial Y$. At a fixed place v , rescale the coordinate so that $|y_0|_v \leq 1$, and let $|Q|_v$ be the maximum of the absolute values of its coefficients after this rescaling. For a multi-index α , write $\partial^\alpha = (\alpha!)^{-1} \partial^{|\alpha|} / \partial T^\alpha$. Put

$$c_v = \frac{|Q|_v}{|Q_Y(0, y_0)|_v} \geq 1.$$

The local argument at a simple root in the cited proof uses no irreducibility hypothesis on Q and gives

$$(20) \quad |\partial^\alpha y(0)|_v \leq c_v^{2|\alpha|-1} \quad (|\alpha| > 0).$$

Return to the projections of Lemma 4.3. Assume that their discriminants and the remaining original coordinates do not vanish at x . The projections are then étale at x . On a neighborhood of x_i in the chart $V_{i0} \neq 0$, the remaining affine coordinate functions are invertible. Write

$$w_{ij} = W_{ij}/V_{i0}, \quad z_{ij} = w_{ij}(x_i), \quad \mathbf{v}_i = (V_{i1}/V_{i0}, \dots, V_{iu_i}/V_{i0}).$$

The entries of $\mathbf{v}_i - \mathbf{v}_i(x_i)$ are local coordinates at x_i . To bound the coefficients of the coordinate equations after translation to x_i , put, for each place v of F ,

$$M_{iv} = \max_j |z_{ij}|_v, \quad A_{iv} = |B_i|_v M_{iv}^{D_i}.$$

Since $1 = V_{i0}/V_{i0}$ is a constant linear combination of the w_{ij} , we have $M_{iv} \geq 1$. Every affine projection coordinate has norm at most M_{iv} at x_i , and $|B_i|_v \geq 1$. With $Q_{ij} = Q_{ij,1}^{b_{ij}}$ as above, define

$$R_{ij} = \left(\frac{1}{b_{ij}!} \partial_Y^{b_{ij}} Q_{ij} \right) (1, \mathbf{v}_i(x_i), z_{ij}), \quad c_i = \prod_j (z_{ij}^{b_{ij}} R_{ij})^{1/b_{ij}}.$$

The discriminant condition gives $R_{ij} \neq 0$. The product c_i collects the derivative denominators in factor i , which will disappear from the global estimate by the product formula. The roots in its definition lie in the fixed field F and do not depend on d . Derivatives below are divided derivatives in the local projection coordinates. The next lemma is the non-archimedean form of [Rem05, Corollary 5.1 and Lemma 5.2]; see also the correction in [Dil20, Lemma 5.1].

Lemma 4.7. *For every Laurent monomial J in the coordinates w_{ij} , every multi-index $\kappa = (\kappa_i)$, and every place v of F ,*

$$(21) \quad |\partial^\kappa J(x)|_v \leq |J(x)|_v \prod_i A_{iv}^{2(N+1)|\kappa_i|} |c_i|_v^{-2|\kappa_i|}.$$

Proof. Translate the projection coordinates to x_i and replace Y by $z_{ij}Y$ in the coordinate equation. Its coefficient norm is at most A_{iv} , and the branch w_{ij}/z_{ij} takes the value 1 at the origin. Since $Q_{ij} = Q_{ij,1}^{b_{ij}}$, the Gauss norm of the transformed irreducible equation is at most $A_{iv}^{1/b_{ij}}$. Its first derivative in Y at this root, raised to the power b_{ij} , equals $z_{ij}^{b_{ij}} R_{ij}$. Applying (20) to this simple root gives, for $|\alpha| > 0$ and $\eta = 1$,

$$|\partial^\alpha w_{ij}^\eta(x_i)|_v \leq |z_{ij}^\eta|_v \left(\frac{A_{iv}^{1/b_{ij}}}{|(z_{ij}^{b_{ij}} R_{ij})^{1/b_{ij}}|_v} \right)^{2|\alpha|}.$$

The ratio is at least 1, so we may increase the exponent from $2|\alpha| - 1$ to $2|\alpha|$. Apply the same argument to the reciprocal polynomial $Y^{D_i} Q_{ij}(1, \mathbf{v}_i(x_i) + T, z_{ij}/Y)$. Its coefficient norm is unchanged, and its relevant derivative at $Y = 1$ differs only by a sign. The displayed estimate therefore also holds for $\eta = -1$.

The product of these ratios over j is at most $A_{iv}^{N+1}/|c_i|_v$. Write J as a product of coordinates and inverse coordinates and distribute the divided derivatives among its factors. Each term is bounded by the right-hand side of (21); the ultrametric inequality gives the same bound for their sum. $\square$

The relevant order of vanishing weights differentiation in factor i by $(dt_1 a_i)^{-1}$. For a nonzero section s_d of $\mathcal{Q}_d$, let σ_d be the least value of

$$\sum_i \frac{|\kappa_i|}{dt_1 a_i}$$

for which the divided derivative of a local representative of s_d at x is nonzero. This is the weighted index of s_d at x . Multiplication by a unit does not change the least weighted degree of a nonzero term, so the definition is independent of the local generator. Let

$$H_x = \sum_i a_i h(x_i), \quad H_B = \sum_i a_i h(B_i).$$

We apply the product-formula argument of [Rem05, proof of Proposition 5.3]. A small weighted index allows the differentiation terms involving H_x to be absorbed into the left side of the height inequality.

Lemma 4.8. *Let s_d be represented by the polynomial family F_d , and let κ attain its weighted index σ_d . Then*

(1) *one has*

$$(22) \quad d\varepsilon H_x - dh_{\mathcal{M}}(x) \leq h(F_d) + 2(N+1) \sum_i |\kappa_i| (h(B_i) + D_i h(x_i));$$

(2) *if $\sigma_d \leq \varepsilon/[4t_1(N+1)\max_i D_i]$, then*

$$(23) \quad \frac{\varepsilon}{2} H_x \leq h_{\mathcal{M}}(x) + \frac{h(F_d)}{d} + \frac{\varepsilon}{2} H_B.$$

Proof. At each place choose a projective coordinate Ξ_{j_v} of $\mathcal{P}$ of maximal norm, a section σ_{k_v} in Σ of maximal norm, and original coordinates $W_{i\ell_{vi}}$ of maximal norm. They define a nonvanishing local generator

$$s_v = \Xi_{j_v}^d \sigma_{k_v}^{-d} \prod_i W_{i\ell_{vi}}^{-d\epsilon a_i}$$

of $\mathcal{Q}_d$ near x . Fix one rational local generator s_0 and write $s_v = \alpha_v s_0$ and $s_d = \alpha s_0$. The coordinate heights give

$$dh_{\mathcal{M}}(x) - d\epsilon H_x = \frac{1}{[F:K]} \sum_v \log |\alpha_v(x)|_v.$$

The representing family gives $\alpha = \alpha_v F_{k_v, \ell_v}(\Xi/\Xi_{j_v})$. All derivatives of smaller weighted order vanish, so

$$\partial^K \alpha(x) = \alpha_v(x) \partial^K (F_{k_v, \ell_v}(\Xi/\Xi_{j_v}))(x).$$

Apply the product formula to the nonzero left side. Each monomial on the right is a Laurent monomial in the original w_{ij} , since the injection for $\mathcal{P}$ is monomial. Its value at x has norm at most 1 by the choice of Ξ_{j_v} . Lemma 4.7 bounds its derivative. Sum over the places and apply the product formula to each fixed nonzero c_i . Since

$$\frac{1}{[F:K]} \sum_v \log M_{iv} = h(x_i),$$

this gives (22).

Under the bound in (2), we have $|\kappa_i| \leq d\epsilon a_i / [4(N+1)D_i]$. Substitution in (22), followed by division by d , proves (23). $\square$

4.5. The geometric product estimate. We use a large weighted index to cut one factor by a proper hypersurface whose degree and height bounds are independent of the auxiliary degree. The algebraic part of Rémond's product estimate gives the product structure. We supply the geometric height bound by intersection theory over the base curve.

4.5.1. Intersections with multiplicities. Let $s \geq 1$ and $n_1, \dots, n_s \geq 0$ be integers. Set $\mathcal{P} = C_F \times \prod_{i=1}^s \mathbb{P}^{n_i}$, let pr_{C_F} be its projection to C_F , and let H_i be the hyperplane class from its i th projective factor. Write $\mathcal{P}_F$ for the generic fiber, $n = (n_1, \dots, n_s)$, $|\alpha| = \sum_i \alpha_i$, $\alpha! = \prod_i \alpha_i!$, and $H^\alpha = \prod_i H_i^{\alpha_i}$. For a subvariety V of codimension c in the generic fiber, let $\mathcal{V}$ be its closure and put

$$d_\alpha(V) = \mathcal{V} \cdot [\text{fiber}] \cdot H^{n-\alpha} \quad (|\alpha| = c), \quad h_\alpha(V) = \frac{\mathcal{V} \cdot H^{n-\alpha}}{[F:K]} \quad (|\alpha| = c-1).$$

We use only indices with $0 \leq \alpha_i \leq n_i$. The numbers $d_\alpha(V)$ are the multidegrees of V , and the numbers $h_\alpha(V)$ are the intersection numbers over C_F used to define its heights. All are nonnegative. Multiprojective Chow heights are fixed linear combinations of the $h_\alpha(V)$.

The geometric Bézout estimate corresponding to [Rem01, Proposition 3.2] follows directly from intersection theory. Let P be a nonzero multihomogeneous polynomial of multidegree $(\delta_1, \dots, \delta_s)$ that does not vanish identically on V . Normalizing one coefficient to 1 and clearing poles gives a divisor of class

$$\text{pr}_{C_F}^* D_P + \sum_i \delta_i H_i, \quad \deg D_P = [F:K] h(P).$$

The closure of the generic intersection cycle is bounded by this effective intersection cycle. Let e_i denote the i th coordinate vector. Intersecting with the hyperplane classes gives, for $|\alpha| = c$,

$$(24) \quad h_\alpha(V \cdot V(P)) \leq \sum_i \delta_i h_{\alpha - e_i}(V) + h(P) d_\alpha(V),$$

where $h_{\alpha - e_i}(V) = 0$ if $\alpha_i = 0$. Iterating this estimate bounds successive proper intersections with their intersection multiplicities.

The index gives a lower bound for the Samuel multiplicity of a component of the ideal generated by the derivatives. We must therefore retain that multiplicity in both estimates. The following is the geometric counterpart of [Rem01, Proposition 3.2]. Write $\delta^\alpha = \prod_i \delta_i^{\alpha_i}$.

Lemma 4.9. *Let $\mathcal{I}$ be a nonzero ideal sheaf on $\mathcal{P}_F$ generated by multihomogeneous polynomials of multidegrees at most $\delta = (\delta_i)$, with coefficient poles bounded by one effective divisor E on C_F . Let T be a component of its zero locus of codimension $c \geq 1$, and put $\mu = e(\mathcal{I}_{\eta_T}, \mathcal{O}_{\mathcal{P}_F, \eta_T})$, the Samuel multiplicity at the generic point η_T of T . Then*

(1) for $|\alpha| = c$,

$$(25) \quad \mu d_\alpha(T) \leq \frac{c!}{\alpha!} \delta^\alpha;$$

(2) for $|\alpha| = c - 1$,

$$(26) \quad \mu h_\alpha(T) \leq \frac{c!}{\alpha!} \delta^\alpha \frac{\deg E}{[F : K]}.$$

Proof. Let $\mathcal{T}$ be the closure of T . Multiplying each generator by all monomials that raise its multidegree to δ preserves the ideal sheaf, since locally one such monomial is a unit. The resulting sections belong to the nef line bundle

$$\mathcal{L} = \text{pr}_{C_F}^* \mathcal{O}_{C_F}(E) \otimes \mathcal{O}(\delta_1, \dots, \delta_s).$$

Let B be the common zero locus of these sections and put $R_0 = [\mathcal{P}]$. Inductively, cut R_{j-1} by a k -linear combination of these sections that is nonzero on each of its components. Such a choice exists because no component of R_{j-1} lies in B and k is infinite. Write the intersection cycle as $A_j + R_j$, where A_j consists of the components contained in B , with their multiplicities. The identities

$$c_1(\mathcal{L})R_{j-1} = A_j + R_j$$

give

$$c_1(\mathcal{L})^c [\mathcal{P}] = R_c + \sum_{j=1}^c c_1(\mathcal{L})^{c-j} A_j.$$

No component discarded before the c th intersection contains η_T , since T is a component of the generic base locus. The first c sections therefore form a system of parameters in the regular local ring $\mathcal{O}_{\mathcal{P}_F, \eta_T}$. Their local intersection multiplicity is their quotient length, which is at least μ by [Rem01, Lemma 3.1]. Hence A_c contains $\mu \mathcal{T}$.

Intersect the last cycle identity with respectively $[\text{fiber}] H^{n-\alpha}$ and $H^{n-\alpha}$. All the other terms have nonnegative degree, because the cycles are effective and $\mathcal{L}$, the H_i , and the fiber class are nef. Finally,

$$c_1(\mathcal{L})^c = \left(\sum_i \delta_i H_i \right)^c + c \text{pr}_{C_F}^*[E] \left(\sum_i \delta_i H_i \right)^{c-1},$$

since $[E]^2 = 0$. The coefficients of these two terms give (25) and (26). $\square$

4.5.2. Components of the index loci. Let P be a nonzero multihomogeneous polynomial on $\prod_{i=1}^s \mathbb{P}^{n_i}$ with positive multidegrees δ_i . Its *weighted index* at a point is the minimum of

$$\sum_{i=1}^s \frac{|\alpha_i|}{\delta_i}$$

over the divided derivatives whose values at the point are nonzero, after dehomogenization in an affine chart containing that point. Here α_i is the multi-index in the i th factor, and derivatives are divided as in Section 4.4. This agrees with the local-section definition of weighted index used there. Let $Z_\epsilon(P)$ be the closed locus where this index exceeds ϵ . By the Euler identities for a multihomogeneous polynomial, it is also cut out by the homogeneous divided derivatives of weighted order at most ϵ , as in [Rem01, §1, p. 289].

A component occurring at two distinct index levels has a large Samuel multiplicity. If the successive polynomial degrees are sufficiently separated, the resulting multidegree inequalities force the component to be a product. We use the multiplicity estimate in [Rem01, Proposition 3.1] and the multidegree argument in [Rem01, proof of Proposition 2.1, pp. 294–296]; Lemma 4.9 supplies the geometric height estimates. For $1 \leq c \leq \sum_i n_i$, put

$$\phi(c) = \max_{|a|=c} \frac{c!}{a!},$$

where the maximum is over the s -component nonnegative multi-indices.

Lemma 4.10. *Let T be a common geometric irreducible component of $Z_\sigma(P)$ and $Z_{\sigma+\tau}(P)$, where $\sigma \geq 0$ and $\tau > 0$, and put $c = \text{codim } T$. Suppose that $\delta_i/\delta_{i+1} > 1$ for $1 \leq i < s$. Then*

- (1) $\phi(c)/\tau^c \geq 1$;
- (2) *if $\delta_i/\delta_{i+1} > \phi(c)/\tau^c$ for $1 \leq i < s$, then $T = T_1 \times \cdots \times T_s$ for geometrically integral subvarieties $T_i \subset \mathbb{P}^{n_i}$. Writing $a_i = \text{codim } T_i$, we have*

$$(27) \quad \prod_i \deg T_i \leq \frac{\phi(c)}{\tau^c}, \quad \delta_i h(T_i) \prod_{j \neq i} \deg T_j \leq a_i \frac{\phi(c)}{\tau^c} h(P) \quad (a_i > 0).$$

Proof. Enlarge F so that T and the factors that occur below are geometrically integral over F . Normalized heights are unchanged.

Normalize one coefficient of P to 1, and let E be the common pole divisor of its coefficients. Then $\deg E = [F : K]h(P)$. Divided differentiation multiplies coefficients by integers, whose valuations are nonnegative. Thus E also bounds the poles of the derivative family, to which Lemma 4.9 applies.

For $1 \leq i \leq s$, let t_i be the codimension of the projection of T to the factors $i, \dots, s$; put $t_{s+1} = 0$ and $b_i = t_i - t_{i+1}$. Let $\mathcal{I}_\sigma$ be the ideal generated by the derivatives of weighted order at most σ , and let μ be its Samuel multiplicity at η_T . Rémond's algebraic multiplicity estimate [Rem01, Proposition 3.1] gives

$$(28) \quad \mu \geq \tau^c \prod_i \delta_i^{b_i}.$$

The derivative hypothesis in that proposition follows from the Leibniz rule: derivatives of $\mathcal{I}_\sigma$ of weighted order at most τ belong to the ideal defining $Z_{\sigma+\tau}(P)$, and hence vanish at η_T . The proposition and its proof in [Rem01, §4, pp. 296–297] are algebraic over a characteristic-zero field, so they apply over F .

To obtain the product structure, we follow the multidegree calculation in [Rem01, proof of Proposition 2.1, pp. 294–296]. Let p_i be the codimension of the projection of T to the first i factors, put $p_0 = 0$, and set $a_i = p_i - p_{i-1}$. For this multi-index a , we have $d_a(T) > 0$, as in the cited proof. Moreover $p_i + t_{i+1} \leq c$, because T lies in the product of its two projections. Combining (25) and (28) gives

$$(29) \quad 1 \leq d_a(T) \leq \frac{c!}{a! \tau^c} \prod_{i < s} \left(\frac{\delta_i}{\delta_{i+1}} \right)^{p_i + t_{i+1} - c}.$$

Since $p_i + t_{i+1} - c \leq 0$ and the successive degree ratios exceed 1, (29) gives $1 \leq \phi(c)/\tau^c$. If every successive degree ratio exceeds $\phi(c)/\tau^c$, then (29) forces $p_i + t_{i+1} = c$ for every i . The inclusion of T in the product of its two projections is then an equality, by irreducibility and equality of dimensions. Applying this to all cuts gives $T = \prod_i T_i$.

For this product, $a_i = \text{codim } T_i = b_i$ and

$$d_a(T) = \prod_i \deg T_i, \quad h_{a-e_i}(T) = h(T_i) \prod_{j \neq i} \deg T_j \quad (a_i > 0).$$

For the second identity, each factor closure is integral and hence flat over the smooth curve C_F . Their fiber product is flat with integral generic fiber. Its local algebras inject into their generic fibers, so it is integral and equals the closure of T . The identity follows from the projection formula. Equations (25)–(28) give (27). $\square$

We now choose index levels so that some two successive loci have a common component. This gives the following geometric version of [Rem01, Corollary 1.2], with bounds independent of the field of definition.

Lemma 4.11. *Suppose $s \geq 2$ and $n_i \geq 1$, and put $u = \sum_i n_i$. For $\epsilon > 0$, set $B_\epsilon = (s/\epsilon)^u$. If*

$$(30) \quad 0 < \epsilon < s \left(\frac{\log(u+1)}{2u^2} \right)^u, \quad \frac{\delta_i}{\delta_{i+1}} > B_\epsilon \quad (1 \leq i < s),$$

every geometric irreducible component of $Z_\epsilon(P)$ is contained in a proper product $Z_1 \times \cdots \times Z_s$ of geometrically integral projective subvarieties satisfying

- (1) $\prod_i \deg Z_i \leq B_\epsilon$;
- (2) *for each i with $Z_i \subsetneq \mathbb{P}^{n_i}$,*

$$(31) \quad \delta_i h(Z_i) \leq n_i B_\epsilon h(P).$$

The factors may be defined over a finite extension of K ; their heights use the normalization above.

Proof. The numerical calculation in [Rem01, §2, pp. 292–293] gives, under (30),

$$\sum_{j=1}^u \phi(j)^{1/j} B_\epsilon^{-1/j} < \epsilon.$$

Choose $\tau_j > \phi(j)^{1/j} B_\epsilon^{-1/j}$ with $\sum_{j=1}^u \tau_j < \epsilon$ and put $\sigma_j = \sum_{r \leq j} \tau_r$, $\sigma_0 = 0$. Given a component W of $Z_\epsilon(P)$, choose components W_j of $Z_{\sigma_j}(P)$ such that $W \subset W_u \subset \cdots \subset W_0$. Here W_0 has codimension one. Choose the first j such that $\text{codim } W_j \leq j$; this exists because $\text{codim } W_u \leq u$. By minimality of j and the inclusion $W_j \subset W_{j-1}$,

$$j \leq \text{codim } W_{j-1} \leq \text{codim } W_j \leq j.$$

Thus $W_j = W_{j-1}$ and its codimension is exactly $c = j$. Then $\phi(c)/\tau_j^c < B_\epsilon$. Lemma 4.10 applies to $W_{j-1} = W_j$, since the degree ratios exceed $B_\epsilon > 1$. Writing this common component as $\prod_i T_i$, we obtain

$$\prod_i \deg T_i \leq B_\epsilon, \quad \delta_i h(T_i) \leq n_i B_\epsilon h(P).$$

The product is proper because it is a component of an index locus of the nonzero polynomial P . $\square$

For one factor $\mathbb{P}^n$, the same choice of a common component applies to a polynomial of degree δ with thresholds $0, \beta/n, \dots, \beta$, where $n \geq 1$ and $0 < \beta < 1$. A component of $Z_\beta(P)$ is contained in a common component T of two successive loci, of some codimension $1 \leq c \leq n$. In the local ring at the generic point of T , the ideal generated by the derivatives of weighted order at most the lower of the two successive thresholds has Samuel multiplicity at least $(\beta\delta/n)^c$, by (28). Lemma 4.9 therefore gives

$$\deg T \leq (n/\beta)^c \leq (n/\beta)^n, \quad \delta h(T) \leq c(n/\beta)^c h(P) \leq n(n/\beta)^n h(P).$$

We use this form after omitting zero-dimensional factors.

4.5.3. From a large index to a hypersurface. We return to Proposition 4.2, retaining Y , the finite projections ρ_i , and D, u, H_B from Sections 4.2–4.4. To apply the product estimate to an auxiliary section, we take a norm to the product of projective spaces. Fix a positive integer d with $d\epsilon \in \mathbb{Z}$, a nonzero section s_d of $\mathcal{Q}_d$, and its representing family F_d . Let σ_d be its weighted index. Assume that the discriminants and remaining original coordinates are nonzero at x , and put $\rho = \prod_i \rho_i$.

Lemma 4.12. *Let $s_d \neq 0$ be represented by F_d . Choose a component of F_d whose corresponding generating section s_λ is nonzero at x , and take its norm under ρ . This gives a nonzero polynomial G_d on $\prod_i \mathbb{P}^{u_i}$ with the following properties.*

- (1) *Its multidegree is $dt_1 Da$, and its index at $\rho(x)$ for these degrees is at least σ_d/D .*
- (2) *Its coefficient height satisfies*

$$(32) \quad h(G_d) \leq D(h(F_d) + dt_1 H_B).$$

Proof. Substitute the monomial expressions for the projective coordinates of $\mathcal{P}$ into the chosen component. It is a nonzero section of multidegree $dt_1 a$ on Y . Its norm is regular because the projection is finite and the section is integral on every affine projection chart; the pole bound at infinity gives multidegree $dt_1 Da$.

Extend the Gauss norm in the affine projection variables to a splitting field. The monic equations and $|B_i|_v \geq 1$ bound every conjugate of W_{ij}/V_{i0} by $|B_i|_v$. Each conjugate of the chosen component therefore has norm at most $|F_d|_v \prod_i |B_i|_v^{dt_1 a_i}$. Taking the product of the D conjugates and summing over the places gives (32).

Let α be the chosen component in a local trivialization. Because $s_\lambda(x) \neq 0$, its index for the weights $dt_1 a$ equals that of s_d . The quotient $N(\alpha)/\alpha$ is a product of integral conjugates and is integral over the base local ring. It lies in the function field of Y , and the local ring $\mathcal{O}_{Y,x}$ is normal because ρ is étale at x . The quotient is therefore regular at x . Its weighted order is nonnegative. The étale projection identifies completed local rings, while the weights for G_d are D times the original weights. This gives the index bound. $\square$

The degree of the norm contains a factor d , so the product estimate replaces $h(F_d)$ by $h(F_d)/d$. For the auxiliary sections already constructed, this quotient is bounded as d tends to infinity. Put $n = m \dim X$, and let s be the number of positive u_i . We omit the zero-dimensional factors when applying Lemma 4.11.

Lemma 4.13. *Suppose $0 < \beta < 1$, $\sigma_d/D > \beta$, and $a_i/a_{i+1} > (m/\beta)^u$ for $1 \leq i < m$. If $s \geq 2$, suppose also that*

$$\beta < s \left(\frac{\log(u+1)}{2u^2} \right)^u.$$

Set $P = ((m+n)/\beta)^n$. Then there is an index i with $u_i > 0$ and a nonzero homogeneous polynomial U on $\mathbb{P}^{u_i}$ vanishing at $\rho_i(x_i)$ such that

$$(33) \quad \deg U \leq P, \quad a_i h(U) \leq \frac{u_i P}{t_1} \left(\frac{h(F_d)}{d} + t_1 H_B \right).$$

Its pullback to Y_i defines a proper hypersurface section through x_i .

Proof. Apply Lemma 4.11 to G_d after omitting the zero-dimensional factors. The ratios between the remaining successive weights still exceed $(m/\beta)^u$. If only one factor remains, use the one-factor estimate following that lemma. In both cases the bound P covers the required degree bound, including $(u/\beta)^u$ in the one-factor case. We obtain a proper factor $Z_i \subsetneq \mathbb{P}^{u_i}$ through $\rho_i(x_i)$ with degree at most P and

$$dt_1 D a_i h(Z_i) \leq u_i P h(G_d).$$

Choose a constant linear projection of Z_i to a hypersurface in $\mathbb{P}^{\dim Z_i + 1}$. Its equation is a constant specialization of the Chow form of Z_i . Pulling it back gives a nonzero homogeneous polynomial U vanishing on Z_i , with $\deg U \leq \deg Z_i$ and $h(U) \leq h(Z_i)$. Combine these bounds with Lemma 4.12 to obtain (33). Since ρ_i is finite and surjective, the nonzero equation U does not vanish identically on Y_i . $\square$

4.6. Uniform constants and completion of the comparison. We prove Proposition 4.2 by choosing a product through x of least dimension among those satisfying prescribed degree and height bounds. At dimension u , denote these bounds by D_u and $B_u|a|$. The recursion below allows every proper section supplied by the preceding lemmas. Start with

$$n = m \dim X, \quad D_n = (\deg X)^m, \quad B_n = \max\{1, h(X), \delta\}.$$

At dimension u , use e_u as the value of ε in $\mathcal{Q}_d$, $A_u|a|$ as a bound for $h(F_d)/d$, and s_u as the threshold for the weighted index. Define these and the constants for hypersurface sections recursively, for $u = n, n-1, \dots, 1$, by

$$\begin{aligned} e_u &= \frac{1}{4u\theta(t_1 m)^u D_u}, & R_u &= 8\theta r_0 (N+1)^m (t_1 m)^u D_u, \\ A_u &= R_u (\delta + (t_1 + t_2 + 1) B_u) + 1, & s_u &= \frac{e_u}{8t_1 (N+1) D_u}, \\ \beta_u &= \min\{s_u / (2D_u), (4n^2)^{-n}\}, & P_u &= \lceil ((m+n)/\beta_u)^n \rceil, \\ L_u &= \max \left\{ \frac{n P_u}{t_1} (A_u + t_1 B_u) + 1, 2D_u B_u, 1 \right\}, & T_u &= \max\{P_u, 2D_u^2\}, \\ D_{u-1} &= D_u T_u, & B_{u-1} &= (T_u + 1) B_u + D_u L_u. \end{aligned}$$

Lemma 4.6 gives $h(F_d)/d \leq A_u|a|$ for large d when the exponent is e_u . For index at most s_u , Lemma 4.8 gives a height bound. For larger index, β_u permits Lemma 4.13, and $T_u, L_u|a|$ bound the degree and weighted height of its equation. The same bounds include the discriminants of Lemma 4.3. Choose

$$c_2 > \max_{1 \leq u \leq n} P_u, \quad c_3 > \max \left\{ B_0, \max_{1 \leq u \leq n} \frac{4(A_u + e_u B_u/2)}{e_u} \right\}, \quad c_1 = \max_{1 \leq u \leq n} \frac{4}{e_u}.$$

These constants depend only on the data listed in Proposition 4.2.

The recursion for D_{u-1} and B_{u-1} preserves the degree and height bounds when one factor is replaced by an irreducible component of a proper hypersurface section.

Lemma 4.14. *Suppose a product $Y = \prod_i Y_i$ through x of dimension $u > 0$ satisfies*

$$\prod_i \deg Y_i \leq D_u, \quad \sum_i a_i h(Y_i) \leq B_u|a|.$$

Suppose an equation U on the original coordinates of one factor, or on its projection coordinates, has $\deg U \leq T_u$ and $a_i h(U) \leq L_u|a|$, vanishes at x_i , and does not vanish identically on Y_i . Replacing Y_i by an irreducible component through x_i of the corresponding hypersurface section gives a product Y' of dimension $u-1$ satisfying

$$\prod_j \deg Y'_j \leq D_{u-1}, \quad \sum_j a_j h(Y'_j) < B_{u-1}|a|.$$

Proof. Substitution of constant projection forms does not increase the coefficient height or degree of U . Bézout's degree inequality and (24) give

$$\prod_j \deg Y'_j \leq D_u T_u = D_{u-1}, \quad \sum_j a_j h(Y'_j) \leq T_u B_u|a| + D_u L_u|a| < B_{u-1}|a|.$$

The section is proper, so each of its irreducible components has dimension $u_i - 1$. Finite extension to define the chosen component preserves the normalized heights. $\square$

Proof of Proposition 4.2. Choose a product $Y = \prod_i Y_i$ through x of least dimension subject to

$$\prod_i \deg Y_i \leq D_u, \quad \sum_i a_i h(Y_i) \leq B_u|a|, \quad u = \sum_i \dim Y_i.$$

Such a product exists because X^m satisfies the bounds. Its dimension is positive: if $u = 0$, then $Y_i = \{x_i\}$ for every i , and $c_3|a| \leq \sum_i a_i h(x_i) \leq B_0|a|$, contrary to $c_3 > B_0$.

Choose the projections of Lemma 4.3. If an original coordinate vanishes at x_i but not identically on Y_i , its equation has degree 1 and height 0. If a discriminant vanishes at $\rho_i(x_i)$, its equation has degree at most T_u and weighted height at most $L_u|a|$, by (13). Each equation cuts Y_i properly, the latter because ρ_i is finite and surjective. Lemma 4.14 contradicts the choice of Y in either case. Thus the projections are étale at x and all remaining coordinates and derivative denominators are nonzero there.

Fix a finite extension F/K defining the point, product, projections, and auxiliary data, as in Section 4.2. Apply Lemma 4.6 with $\varepsilon = e_u$ and $D_* = D_u$. For every sufficiently large positive integer d with $de_u \in \mathbb{Z}$, its section satisfies $h(F_d)/d \leq A_u|a|$. If $\sigma_d > s_u$, then

$$\sigma_d/D > s_u/D_u \geq 2\beta_u, \quad a_i/a_{i+1} \geq c_2 > P_u.$$

Since $P_u \geq (m/\beta_u)^u$, these inequalities give the weight condition in Lemma 4.13. The choice $\beta_u \leq (4n^2)^{-n}$ also gives its smallness condition: for $2 \leq s \leq u \leq n$, we have $(4n^2)^{-n} < s(2u^2)^{-u} < s(\log(u+1)/(2u^2))^u$. This also applies after zero-dimensional factors are omitted.

That lemma gives an equation of degree at most T_u and weighted height at most $L_u|a|$. Lemma 4.14 again contradicts minimality.

Hence $\sigma_d \leq s_u$ for every sufficiently large positive integer d with $de_u \in \mathbb{Z}$. Apply Lemma 4.8 and let d tend to infinity with F and a fixed. Since $H_B \leq B_u|a|$, we obtain

$$\frac{e_u}{2} H_x \leq h_{\mathcal{M}}(x) + (A_u + e_u B_u/2)|a|.$$

The inequality $H_x \geq c_3|a|$ and the choice of c_3 give $(e_u/4)H_x \leq h_{\mathcal{M}}(x)$, proving (11). Only the lower bound on d depends on the field and the weights; the constants c_1, c_2, c_3 do not. $\square$

4.7. The semi-abelian difference maps. We apply Proposition 4.2 to differences on G and its abelian quotient. Since the boundary height has degree one and the abelian height has degree two under multiplication, we use weights a_i^2 for the former and a_i for the latter. Both contributions then have degree two in the weights.

Let t be the torus rank of G . Choose $n \geq 1$ so that $N_0 = N^{\otimes n}$ gives the projectively normal embedding of [Rem03, §2], and put $c = (t+1)n$. Set $L^* = M + c\bar{\pi}^*N = M + (t+1)\bar{\pi}^*N_0$. We first prove Theorem 4.1 for its canonical height

$$\hat{h}_{L^*} = \hat{h}_M + c(\hat{h}_N \circ \pi).$$

In Theorem 4.1(2), we accordingly use $\hat{h}_M$ and $c\hat{h}_N$ for its two components.

Let $a = (a_1, \dots, a_m)$ be a tuple of positive integers. Each endpoint occurs in one adjacent difference and each interior point in two. We record these multiplicities by η_i , and define the exponents b_i used in the line bundles below and their sum H_a by

$$\eta_1 = \eta_m = 1, \quad \eta_i = 2 \quad (1 < i < m), \quad b_i = \eta_i a_i^2, \quad H_a = |b| = \sum_i b_i.$$

For $x = (x_1, \dots, x_m) \in (V \setminus Z_V)(\bar{K})^m$, write

$$D_a(x) = \sum_{i < m} (\hat{h}_M(a_i^2 x_i - a_{i+1}^2 x_{i+1}) + c\hat{h}_N(\pi(a_i x_i - a_{i+1} x_{i+1}))).$$

We first prove the following estimate for positive integer weights and then use it to deduce Theorem 4.1. We prove it on a projective graph, where the weighted differences define line bundles. Compare [Rem03, Lemma 4.2 and proof of Theorem 4.1].

Proposition 4.15. *There are positive constants κ, h_0, C_4, C_5 , depending only on G, V , the chosen embeddings and heights, such that*

$$(34) \quad \sum_i \eta_i a_i^2 \hat{h}_{L^*}(x_i) \leq C_4 D_a(x) + C_5 H_a$$

for all $x_i \in (V \setminus Z_V)(\bar{K})$ and positive integers a_i satisfying

- (1) $\hat{h}_{L^*}(x_i) \geq h_0$ for every i ;
- (2) $a_i / a_{i+1} \geq \kappa$ for $1 \leq i < m$.

The constants are independent of the weights and the fields of definition of the points.

Fix $x \in (V \setminus Z_V)(\bar{K})^m$ and positive integer weights a . We use the graph construction of [Rem03, §3]. To apply Proposition 4.2, we need an intersection bound and a coefficient bound with constants independent of a . We then compare the Weil height of the line bundle defined by these differences with $D_a(x)$.

4.7.1. *Line bundles and intersections.* Write the chosen standard compactification $\overline{G}$ as the fiber product over A of the bundles $\mathbb{P}(\mathcal{O}_A \oplus Q_j)$, where $Q_1, \dots, Q_t \in \text{Pic}^0(A)$; for $t = 0$, it is A . Let $\overline{V}$ be the closure of V in $\overline{G}$. The boundary line bundle is the fixed M . For positive integers $a_1, \dots, a_m$, the graph compactification in [Rem03, §3] extends the map

$$(x_i) \mapsto (a_i^2 x_i - a_{i+1}^2 x_{i+1})_{i < m}$$

to a morphism $\overline{\beta}_a$. The graph projection to $\overline{V}^m$ is an isomorphism over V^m . On the same graph, let r_a be the morphism to A^{m-1} with components $a_i \pi(x_i) - a_{i+1} \pi(x_{i+1})$. Set

$$\mathcal{M}_a = \overline{\beta}_a^*(M^{\boxtimes(m-1)}) \otimes r_a^*(N_0^{\boxtimes(m-1)})^{\otimes(t+1)}.$$

The bundle $\mathcal{M}_a$ is nef. To see this, pull the tautological quotient on $\mathbb{P}(\mathcal{O}_A \oplus Q_j)$ back to a normalized projective curve. It receives a nonzero map from one of two degree-zero line bundles, so it has nonnegative degree. The boundary bundle $\mathcal{O}(2) \otimes p^* Q_j^{-1}$ is therefore nef, where p denotes projection to A . Tensor products and pullbacks preserve nefness, and N_0 is ample.

The line bundle $L^* = M + (t+1)\overline{\pi}^* N_0$ is very ample for this choice of n . Pull $\mathcal{N}_b = \boxtimes_i (L^*)^{\otimes b_i}$ back to the graph. The section maps of [Rem03, Lemmas 3.2–3.3 and pp. 204–205] give a very ample bundle $\mathcal{P}_a$ and injections

$$\mathcal{P}_a \longrightarrow \mathcal{N}_b^{\otimes(t+2)}, \quad \mathcal{P}_a \otimes \mathcal{M}_a^{-1} \longrightarrow \mathcal{N}_b^{\otimes 4(t+1)}.$$

The first sends coordinate sections to monomials with coefficient 1. The bundle $\mathcal{P}_a \otimes \mathcal{M}_a^{-1}$ is generated by $r_0 = (N_2 + 1)^m (N_3 + 1)$ sections, where N_2, N_3 are the fixed projective dimensions in the boundary and abelian graph embeddings of [Rem03, §3]. For the second injection, choose a coordinate section nonzero at x . Both injections are then isomorphisms near x . These constructions use the theorem of the cube and the projective embeddings, so they apply over our characteristic-zero field. Thus the parameters in Proposition 4.2 are

$$(35) \quad t_1 = t + 2, \quad t_2 = 4(t + 1), \quad \theta = 2^{m^2}.$$

The coefficient parameter δ will be supplied by Lemma 4.18.

Let $Y = \prod_i Y_i$ be a product of integral subvarieties with $x_i \in Y_i \subset \overline{V}$, and put $u_i = \dim Y_i$ and $u = \sum_i u_i$. Write $\tilde{Y}_a$ for its strict transform on the graph and $\mathbf{1} = (1, \dots, 1)$. By Vojta's homogeneity theorem [Voj96, Theorem 5.5], it suffices to prove positivity when all weights are one. We use the adjacent-pair form in [Rem03, pp. 205–206].

Lemma 4.16. *For every such product Y , we have*

$$(36) \quad (\mathcal{M}_a^u \cdot \tilde{Y}_a) \geq \prod_i a_i^{2u_i} \geq 2^{-m^2} \prod_i b_i^{u_i}.$$

Proof. The cited homogeneity theorem gives

$$(37) \quad (\mathcal{M}_a^u \cdot \tilde{Y}_a) = \left(\prod_i a_i^{2u_i} \right) (\mathcal{M}_1^u \cdot \tilde{Y}_1).$$

To pass from the cited proof over $\mathbb{C}$ to our field, descend the graph, line bundles, and product to a finitely generated characteristic-zero subfield and embed it into $\mathbb{C}$. Intersection numbers are invariant under field extension. It remains to prove positivity for $a_1 = \dots = a_m = 1$.

Set $U_i = Y_i \cap G$. Since $x_i \in U_i \setminus Z_V$, no U_i is contained in Z_V . The difference map

$$(y_1, \dots, y_m) \mapsto (y_1 - y_2, \dots, y_{m-1} - y_m)$$

is therefore generically finite on $\prod_i U_i$ by [Rem03, p. 206], which applies the algebraic argument of [Rem00, Lemma 2.1] with $m = \dim V + 1$. When $a_1 = \cdots = a_m = 1$, $\mathcal{M}_1$ is the pullback of the ample bundle $(L^*)^{\boxtimes(m-1)}$ by the extension of this map. The projection formula gives $(\mathcal{M}_1^u \cdot \tilde{Y}_1) \geq 1$, including when $u = 0$. Since $u \leq m(m-1) < m^2$ and $\eta_i \leq 2$, this proves (36). $\square$

4.7.2. Polynomials representing multiplication. We use [Rem00, Proposition 5.2, pp. 126–128] and [Rem03, Lemma 3.3] for the coefficient estimate. We give the estimate for geometric heights, keeping track of the common scalar in each coefficient family. Put $n_0 = h^0(A, N_0) - 1$. Let $\xi_0, \dots, \xi_{n_0}$ be the coordinate sections of N_0 , and let $Z_0, \dots, Z_{n_0}$ be the corresponding polynomial variables. By the theorem of the cube, projective normality, and the Künneth formula, there are bihomogeneous polynomials $R_{\ell, \ell'}$ of bidegree $(2, 2)$ representing the pullbacks of $\xi_\ell \boxtimes \xi_{\ell'}$ by $(y, z) \mapsto (y + z, y - z)$. Fix these polynomials and let $H_{\text{add}} \geq 1$ bound the height of their joint coefficient vector.

Multiplying the sections by a power of one coordinate permits a recursion by addition and subtraction. The exponents below make all the required coordinate powers nonnegative, so this recursion gives polynomials whose coefficient heights have quadratic bounds. Put

$$f(r) = \left\lfloor \frac{r^2 - 1}{8} \right\rfloor \quad (r \geq 1).$$

Lemma 4.17. *For every $r \geq 1$ and fixed coordinate index j , there is a family $\mathbf{Q}_{r,j} = (Q_{r,i,j})_{0 \leq i \leq n_0}$ of homogeneous polynomials such that*

- (1) *each polynomial has degree $r^2 + f(r)$ and represents $\xi_j^{f(r)} [r]^* \xi_i$ under $[r]^* N_0 \simeq N_0^{\otimes r^2}$;*
- (2) *the height $\mathcal{H}_{r,j}$ of the family's joint coefficient vector satisfies $\mathcal{H}_{r,j} \leq (r^2 - 1) H_{\text{add}}$.*

The estimate is uniform in j .

Proof. Start with $Q_{1,i,j} = Z_i$. For $r > 1$, put $r_- = \lfloor r/2 \rfloor$ and $r_+ = \lceil r/2 \rceil$. The addition and subtraction of $[r_+]y$ and $[r_-]y$ give $[r]y$ and $[r_+ - r_-]y$. If r is odd, the substituted polynomial $R_{i,j}(\mathbf{Q}_{r_+,j}, \mathbf{Q}_{r_-,j})$ therefore represents

$$\xi_j^{2f(r_-)+2f(r_+)+1} [r]^* \xi_i.$$

Multiply it by $Z_j^{f(r)-2f(r_-)-2f(r_+)-1}$. If r is even, choose once and for all an index j_0 such that $\xi_{j_0}(0) \neq 0$. Substitute into R_{i,j_0} , divide by the resulting common nonzero scalar, and multiply by $Z_j^{f(r)-4f(r_-)}$. These exponents are nonnegative because

$$f(2s) \geq 4f(s), \quad f(2s+1) \geq 2f(s) + 2f(s+1) + 1.$$

The two constructions have the required degree and represent the required sections. Multiplication by a monomial and division by a common nonzero scalar preserve the height of the coefficient vector. The ultrametric inequality gives

$$\mathcal{H}_{1,j} = 0, \quad \mathcal{H}_{r,j} \leq H_{\text{add}} + 2\mathcal{H}_{r_-,j} + 2\mathcal{H}_{r_+,j} \leq (r^2 - 1) H_{\text{add}},$$

where the last inequality follows by induction from $1 + 2(r_-^2 - 1) + 2(r_+^2 - 1) \leq r^2 - 1$. This bound holds for each fixed j , uniformly in the choice of j . $\square$

4.7.3. *Coefficients of the section maps.* We apply these polynomials to adjacent pairs. The formula in [Rem03, Lemma 3.3, p. 202] then gives the coefficient bound for the second injection.

Lemma 4.18. *The second injection and its generating family can be chosen so that their images are represented by a polynomial family $\mathbf{P}_a$ of multidegree $4(t+1)b$ satisfying*

$$(38) \quad h(\mathbf{P}_a) \leq \delta H_a, \quad \delta = 2(t+1)H_{\text{add}}.$$

The constant is independent of the weights and the fields of definition of the points.

Proof. For positive integers a, b and fixed coordinate indices j, k , the family

$$P_{a,b;j,k;\ell,\ell'} = R_{\ell,\ell'}(\mathbf{Q}_{a,j}, \mathbf{Q}_{b,k})$$

has bidegree $(2(a^2 + f(a)), 2(b^2 + f(b)))$. Its restriction represents the pullbacks of $\xi_\ell \boxtimes \xi_{\ell'}$ by $(y, z) \mapsto (ay + bz, ay - bz)$, multiplied by $p_1^* \xi_j^{2f(a)} p_2^* \xi_k^{2f(b)}$, where $p_1, p_2 : A \times A \rightarrow A$ are the projections. Its joint coefficient height is at most

$$H_{\text{add}} + 2H_{a,j} + 2H_{b,k} \leq 2(a^2 + b^2)H_{\text{add}}.$$

Choose j_i such that $\xi_{j_i}(\pi(x_i)) \neq 0$ for each i , and write $Z_j^{(i)}$ for the coordinate variables on the i th factor. The formula in [Rem03, Lemma 3.3, p. 202] takes the product of these pair polynomials over adjacent factors, raises each to the power d , and multiplies by

$$\prod_i (Z_{j_i}^{(i)})^{2d\eta_i(a_i^2 - f(a_i))}.$$

The resulting family has exactly multidegree $4d\eta_i a_i^2$ in factor i : the adjacent-pair product contributes $2d\eta_i(a_i^2 + f(a_i))$. Its joint coefficient height is at most

$$2dH_{\text{add}} \sum_{i < m} (a_i^2 + a_{i+1}^2) = 2dH_{\text{add}}H_a.$$

For the second injection above, take $d = t+1$. The additional boundary factors are monomials with coefficient 1. This gives (38). The fixed addition family and invariance of normalized heights under finite extension make δ independent of the weights and the field of definition of the points. $\square$

4.7.4. *Comparison with canonical heights.* Let $h_{\mathcal{M}_a}$ be the Weil height defined by the graph coordinates. We next compare it with the canonical heights, uniformly in a . The following is the geometric form of the section calculation in [Rem03, Lemma 4.2].

Lemma 4.19. *There is a constant $C > 0$, depending only on the fixed embeddings and height data, such that*

$$(39) \quad h_{\mathcal{M}_a}(x) \leq \sum_{i < m} (\widehat{h}_M(a_i^2 x_i - a_{i+1}^2 x_{i+1}) + c\widehat{h}_N(\pi(a_i x_i - a_{i+1} x_{i+1}))) + CH_a$$

for all the weights and points under consideration, independently of their fields of definition.

Proof. After a fixed finite extension, choose $P_j \in A(\overline{K})$ with $Q_j \simeq \tau_{P_j}^* N_0 \otimes N_0^{-1}$, where τ_{P_j} denotes translation by P_j . For $P \in \{0, \pm P_j\}$, the identity

$$[2]^* \tau_P^* N_0 \simeq (\tau_P^* N_0)^{\otimes 2} \otimes N_0^{\otimes 2}$$

compares two fixed generating families of sections. Their local norm ratios are bounded on a projective model and contribute zero outside finitely many places. Normalization preserves these bounds after finite extension. The graph construction uses the families with total tensor exponent $O(H_a)$, so the resulting errors are $O(H_a)$ with a fixed implied constant.

Let $\mathcal{E}_a = \overline{\beta}_a^*(M^{\boxtimes(m-1)})$ and choose its Weil height by

$$h_{\mathcal{E}_a}(x) = h_{\mathcal{M}_a}(x) - (t+1) \sum_{i < m} \widehat{h}_{N_0}(a_i \pi(x_i) - a_{i+1} \pi(x_{i+1})),$$

and let $B_a(x)$ be the projective height of the boundary coordinates in [Rem03, Definition 3.2]. The section formulas and the parallelogram law give

$$\begin{aligned} h_{\mathcal{E}_a}(x) &= B_a(x) - t \sum_i \eta_i a_i^2 \widehat{h}_{N_0}(\pi(x_i)) + O(H_a), \\ 2B_a(x) + 2t \sum_i \eta_i a_i^2 \widehat{h}_{N_0}(\pi(x_i)) &\leq B_a([2]x) + O(H_a). \end{aligned}$$

Since $\widehat{h}_{N_0}([2]y) = 4\widehat{h}_{N_0}(y)$, these imply $2h_{\mathcal{E}_a}(x) \leq h_{\mathcal{E}_a}([2]x) + CH_a$. Iteration yields

$$h_{\mathcal{E}_a}(x) \leq 2^{-r} h_{\mathcal{E}_a}([2]^r x) + (1 - 2^{-r})CH_a.$$

For fixed a , $h_{\mathcal{E}_a}$ differs by a bounded function from the pullback of a Weil height for $M^{\boxtimes(m-1)}$. Its normalized limit is the sum of linear heights in (39); the bounded difference vanishes after multiplication by 2^{-r} . This proves the estimate with C independent of a . $\square$

Proof of Proposition 4.15. For the embedding defined by L^* , its Weil height satisfies $h_{L^*} = \widehat{h}_{L^*} + O(1)$ uniformly on algebraic points; see [Rem03, §2, p. 195]. Apply Proposition 4.2 with the parameters (35), the intersection bound (36), and the coefficient bound (38). Since $b_i/b_{i+1} \geq a_i^2/(2a_{i+1}^2)$, a fixed lower bound on a_i/a_{i+1} supplies the weight hypothesis. The bounded difference between h_{L^*} and $\widehat{h}_{L^*}$ supplies the height hypothesis after increasing h_0 . Combining the conclusion with Lemma 4.19 proves (34); the bounded height differences contribute a constant multiple of H_a . All constants depend only on the stated fixed data. $\square$

4.8. Completion of the Vojta inequality.

4.8.1. Integer weights and normalized differences.

Proof of Theorem 4.1 over a curve. We approximate the normalized points in Theorem 4.1 by integer multiples and apply Proposition 4.15, as in [Rem03, Lemma 4.1 and proof of Theorem 4.1]. Choose $C_1 > 0$, $C_2 > \max\{\kappa, 1\}$, and $C_3 \geq h_0$ so that

$$C_4(C_1^{-1} + C_1^{-2}) < 1, \quad 2C_5/C_3 < 1.$$

Suppose a tuple outside Z_V satisfies the hypotheses of Theorem 4.1 for $\widehat{h}_{L^*}$ with these constants. Put $H_i = \widehat{h}_{L^*}(x_i)$. For each positive integer v , choose $a_i(v)$ nearest to $v/\sqrt{H_i}$. These integers are positive for large v , and

$$\frac{a_i(v)}{a_{i+1}(v)} \longrightarrow \sqrt{\frac{H_{i+1}}{H_i}} \geq C_2 > \kappa, \quad \frac{a_i(v)^2}{v^2} \longrightarrow \frac{1}{H_i}.$$

Thus (34) applies. Divide it by v^2 and let v tend to infinity. Homogeneity and continuity on the real span of the fixed tuple identify the limit of $D_{a(v)}(x)/v^2$ with

$$\sum_{i < m} \left(\hat{h}_M \left(\frac{x_i}{H_i} - \frac{x_{i+1}}{H_{i+1}} \right) + c \hat{h}_N \left(\frac{\pi(x_i)}{\sqrt{H_i}} - \frac{\pi(x_{i+1})}{\sqrt{H_{i+1}}} \right) \right).$$

The inequalities in Theorem 4.1(2), applied to $\hat{h}_{L^*}$, bound this by $(m-1)(C_1^{-1} + C_1^{-2})$. Since $\sum_i \eta_i = 2(m-1)$ and $H_i \geq C_3$, the limiting inequality gives

$$2 \leq C_4(C_1^{-1} + C_1^{-2}) + \frac{2C_5}{C_3} < 2,$$

a contradiction. Here the auxiliary degree has already tended to infinity with the field and the weights fixed. Thus (34) has no genus term before the weights vary with v .

It remains to replace $\hat{h}_{L^*}$ by $\hat{h}_L$. Suppose that x, y have positive height and the two expressions in Theorem 4.1(2), after taking the square root of the second, are at most δ . Set

$$r_x = \frac{\hat{h}_{L^*}(x)}{\hat{h}_L(x)} \in [1, c], \quad r_y = \frac{\hat{h}_{L^*}(y)}{\hat{h}_L(y)} \in [1, c].$$

The identity $r_x = c - (c-1)\hat{h}_M(x)/\hat{h}_L(x)$ and the reverse triangle inequality give $|r_x - r_y| \leq (c-1)\delta$. On $[1, c]$, the functions s^{-1} and $s^{-1/2}$ have Lipschitz constants 1 and 1/2. For $\hat{h}_{L^*}$, whose quadratic part is $c(\hat{h}_N \circ \pi)$, the corresponding first expression is at most $c\delta$ and the square root of the second is at most $\sqrt{c}(1 + (c-1)/2)\delta \leq c^{3/2}\delta$. Moreover,

$$\hat{h}_{L^*}(x) \geq \hat{h}_L(x), \quad \frac{\hat{h}_{L^*}(y)}{\hat{h}_{L^*}(x)} \geq \frac{\hat{h}_L(y)}{c\hat{h}_L(x)}.$$

If C_1, C_2, C_3 are the constants for $\hat{h}_{L^*}$, choose $c_1 \geq c^{3/2}C_1$, $c_2 \geq \sqrt{c}C_2$, and $c_3 \geq C_3$. The hypotheses for $\hat{h}_L$ would imply those already excluded for $\hat{h}_{L^*}$. This proves Theorem 4.1 over a curve. $\square$

4.8.2. Higher transcendence degree. To pass to higher transcendence degree, we choose a generic curve on the fixed polarized model. We need a single curve that preserves normalized heights for every finite extension of K . Suppose $b = \dim \mathcal{B} > 1$. Choose $e \geq 1$ such that $\mathcal{H}^{\otimes e}$ is very ample, and let k' be an algebraic closure of the function field of the full parameter space of $(b-1)$ -tuples of hyperplanes in this embedding. Let C/k' be their generic complete intersection in $\mathcal{B}_{k'}$. Since $\mathcal{B}$ is normal, its singular locus has codimension at least two and is avoided by this generic curve. Bertini's theorem in characteristic zero then makes C smooth; it is projective as a closed subvariety of $\mathcal{B}_{k'}$. The incidence variety dominates $\mathcal{B}$, so it gives an injection

$$K \hookrightarrow k'(C).$$

We fix an extension of this injection from $\bar{K}$ to an algebraic closure of $k'(C)$.

Put $\lambda = e^{b-1}$, and denote the base changes of M, N, L by M_C, N_C, L_C . Heights over $k'(C)$ use the normalization of Section 4.1. The generic curve has the following properties.

Lemma 4.20. *The extension $K \hookrightarrow k'(C)$ has the following properties.*

- (1) *For every finite extension F/K , the algebra $F \otimes_K k'(C)$ is a field of degree $[F : K]$ over $k'(C)$.*
- (2) *The canonical heights satisfy $\hat{h}_{M_C} = \lambda \hat{h}_M$, $\hat{h}_{N_C} = \lambda \hat{h}_N$, and $\hat{h}_{L_C} = \lambda \hat{h}_L$ on the original algebraic points of G, A, G , respectively.*
- (3) *Every point of $(V \setminus Z_V)(\overline{K})$ remains outside the Mordell exceptional locus after extension to an algebraic closure of $k'(C)$.*

Proof. Let F/K be any finite extension and let $\mathcal{B}_F \rightarrow \mathcal{B}$ be the normalization. The normalization morphism is finite, so every nonempty fiber has dimension zero. Thus [PS22, Theorem 1.5] applies: while the image has dimension at least two, its bound makes the locus of hyperplanes whose inverse images are not geometrically irreducible a proper subset of the dual projective space. Successive generic hyperplanes therefore have geometrically irreducible inverse images. Generic reducedness follows from Bertini in characteristic zero, giving a geometrically integral curve over the parameter field. Thus, after extending the parameter field to k' , the induced curve has function field

$$F \otimes_K k'(C), \quad [F \otimes_K k'(C) : k'(C)] = [F : K].$$

The good open subset of the parameter space may depend on F , but its generic point belongs to every such open subset. Consequently the single field $k'(C)$ has this property for every finite F/K . The same generic complete intersection avoids every fixed codimension-two subset of $\mathcal{B}_F$ defined over k ; in particular, modifications needed to extend a given rational map do not change the induced curve.

We record the height comparison, including its normalization. For a projective point z over F , normalize one nonzero coordinate to 1, and let D_z be the common pole divisor of its coordinates on $\mathcal{B}_F$. Write $\mathcal{H}_F$ for the pullback of $\mathcal{H}$. The generic curve avoids the indeterminacy locus of the rational map defined by z , so restricting its coordinates and applying the projection formula gives

$$h_C(z) = \frac{D_z \cdot (e\mathcal{H}_F)^{b-1}}{[F : K]} = \lambda h(z).$$

The same calculation applies to a line bundle on a projective model: extend the point on the closure of its graph and restrict the pulled-back line bundle to the generic curve. It also applies to coefficient and Chow heights, while generic multidegrees are unchanged by extension of the field. These identities hold for every algebraic point and every multiple $[n]x$ with the same generic curve. Passing to the canonical limits gives

$$\hat{h}_{M_C} = \lambda \hat{h}_M, \quad \hat{h}_{N_C} = \lambda \hat{h}_N, \quad \hat{h}_{L_C} = \lambda \hat{h}_L.$$

Original points outside Z_V remain outside the Mordell exceptional locus after this extension. Indeed, any positive-dimensional semi-abelian subgroup over the enlarged algebraically closed field descends to $\overline{K}$ by [Liu24, Theorem 2.3.12]. If a coset of such a subgroup contains $x \in V(\overline{K})$, it is $x + B$ after base change, for a subgroup B over $\overline{K}$. Its containment in the extended variety V descends faithfully flatly, so $x + B \subset V$ and $x \in Z_V$. $\square$

Completion of the proof of Theorem 4.1. Both expressions in Theorem 4.1(2) are unchanged by the displayed scaling, by their respective linear and quadratic homogeneities. Height ratios are unchanged as well. If c'_1, c'_2, c'_3 are the constants over the curve, take $c_1 = c'_1$, $c_2 = c'_2$,

and $c_3 = c'_3/\lambda$. This proves the theorem for the original polarization. The trace and bounded-height arguments of Sections 2 and 3 continue to use the original extension K/k . $\square$

5. PROOF OF THE MAIN THEOREM

After passing to the stabilizer quotient, Theorem 4.1 bounds heights outside a proper closed subset. Proposition 3.1 will then give the main theorem.

5.1. Bounded height outside the cosets. Finite rank makes the quotients by the kernels of the seminorms $\hat{h}_M$ and $\sqrt{\hat{h}_N \circ \pi}$ finite-dimensional. A finite covering of their unit balls supplies the nearby normalized points required by Theorem 4.1. The following is the geometric counterpart of [Rem03, §5, Lemma 5.1(1) and p. 210]; see also [Ge24, Proposition 3.3] in the abelian arithmetic setting.

Proposition 5.1. *Let $V \subset G_{\bar{K}}$ be a closed subvariety, and let $\Lambda \subset G(\bar{K})$ be a subgroup of finite rank. There are $\eta > 0$ and $R \geq 0$ such that*

$$\hat{h}_L(x) \leq R \quad (x \in (V \setminus Z_V)(\bar{K}) \cap (\Lambda + B_\eta(G))).$$

Proof. If $V = \{v\}$, take $R = \hat{h}_L(v)$ and $\eta = 1$. Otherwise choose the constants in Theorem 4.1 with $c_1 \geq 1$ and $c_2 > 1$, and again set $\eta = 1$. We will obtain a finite partition in which pairs of sufficiently large points satisfy Theorem 4.1(2).

Write $x = \lambda + z$, where $\lambda \in \Lambda$ and $\hat{h}_L(z) < 1$. The triangle inequalities for $\hat{h}_M$ and $\sqrt{\hat{h}_N \circ \pi}$ give

$$|\hat{h}_L(x) - \hat{h}_L(\lambda)| \leq 1 + 2\sqrt{\hat{h}_L(\lambda)}.$$

The function $\sqrt{\hat{h}_M}$ is subadditive, as is $\sqrt{\hat{h}_N \circ \pi}$. Applying the Euclidean triangle inequality to this pair of functions proves the triangle inequality for $\sqrt{\hat{h}_L}$. In particular,

$$|\sqrt{\hat{h}_L(x)} - \sqrt{\hat{h}_L(\lambda)}| \leq 1.$$

If $\hat{h}_L(x) > 1$, then $\hat{h}_L(\lambda) > 0$, and we may normalize by both heights. Homogeneity and the triangle inequalities give

$$\begin{aligned} \hat{h}_M\left(\frac{x}{\hat{h}_L(x)} - \frac{\lambda}{\hat{h}_L(\lambda)}\right) &\leq \frac{1 + |\hat{h}_L(x) - \hat{h}_L(\lambda)|}{\hat{h}_L(x)} \leq \frac{4 + 2\sqrt{\hat{h}_L(\lambda)}}{\hat{h}_L(x)}, \\ \sqrt{\hat{h}_N\left(\frac{\pi(x)}{\sqrt{\hat{h}_L(x)}} - \frac{\pi(\lambda)}{\sqrt{\hat{h}_L(\lambda)}}\right)} &\leq \frac{1 + |\sqrt{\hat{h}_L(x)} - \sqrt{\hat{h}_L(\lambda)}|}{\sqrt{\hat{h}_L(x)}} \leq \frac{2}{\sqrt{\hat{h}_L(x)}}. \end{aligned}$$

Both bounds tend to zero as $\hat{h}_L(x) \rightarrow \infty$. Choose $R_0 > \max\{1, c_3\}$ so that they are at most $1/(4c_1)$ for $\hat{h}_L(x) \geq R_0$, independently of x and its decomposition.

Take the finite-dimensional normed quotients of $\Lambda \otimes_{\mathbb{Z}} \mathbb{R}$ by the kernels of $\hat{h}_M$ and $\sqrt{\hat{h}_N \circ \pi}$, respectively. The images of $\lambda/\hat{h}_L(\lambda)$ in the first quotient and of $\lambda/\sqrt{\hat{h}_L(\lambda)}$ in the second belong to the corresponding closed unit balls, since $\hat{h}_M(\lambda) \leq \hat{h}_L(\lambda)$ and $\hat{h}_N(\pi(\lambda)) \leq \hat{h}_L(\lambda)$. Compactness gives a finite $1/(4c_1)$ -net in each ball. For every x with $\hat{h}_L(x) \geq R_0$, fix a

decomposition $x = \lambda + z$ and assign it to a pair of net centers within $1/(4c_1)$ of the normalized images of λ . The pairs of centers define a finite partition.

If x and x' belong to the same part, the normalized subgroup images differ by at most $1/(2c_1)$ in each norm. Combining this with the preceding error bounds gives

$$\begin{aligned} \hat{h}_M \left(\frac{x}{\hat{h}_L(x)} - \frac{x'}{\hat{h}_L(x')} \right) &\leq c_1^{-1}, \\ \hat{h}_N \left(\frac{\pi(x)}{\sqrt{\hat{h}_L(x)}} - \frac{\pi(x')}{\sqrt{\hat{h}_L(x')}} \right) &\leq c_1^{-2}. \end{aligned}$$

If the heights in $(V \setminus Z_V)(\bar{K}) \cap (\Lambda + B_1(G))$ were unbounded, one part would have unbounded heights. Successively choose $m = \dim V + 1$ points in that part with first height at least R_0 and successive height ratios at least c_2^2 . They satisfy both conditions of Theorem 4.1, a contradiction. The asserted bound R follows. $\square$

5.2. The stabilizer quotient.

Proof of Theorem 1.1. Assume (1). Let $S = \text{Stab}(X)^0$ and $q : G_{\bar{K}} \rightarrow \tilde{G} = G_{\bar{K}}/S$ be the stabilizer quotient. Choose a standard compactification of $\tilde{G}$ and a line bundle $\tilde{L} = \tilde{M} + \tilde{\pi}^* \tilde{N}$, where $\tilde{M}$ is its boundary line bundle, $\tilde{N}$ is symmetric and ample on its abelian quotient, and $\tilde{\pi}$ extends the projection from $\tilde{G}$ to its abelian quotient. Define $B_\epsilon(\tilde{G})$ using $\hat{h}_{\tilde{L}}$. Height comparison for q shows that $q(X)(\bar{K}) \cap (q(\Gamma) + B_\epsilon(\tilde{G}))$ is dense in $q(X)$ for every $\epsilon > 0$.

The group $q(\Gamma)$ has finite rank and satisfies $[n]^{-1}q(\Gamma) = q(\Gamma)$ for every $n \geq 1$, with inverse images taken in $\tilde{G}(\bar{K})$. Indeed, if $[n]b = q(\gamma)$ for $\gamma \in \Gamma$, lift b to $x \in G(\bar{K})$. Multiplication by n is surjective on S , so some $s \in S(\bar{K})$ satisfies $[n]s = [n]x - \gamma$. Then $[n](x - s) = \gamma$, whence $x - s \in \Gamma$ and $b \in q(\Gamma)$. Applying q to the definition of Γ gives $q(\Gamma) \subset q(\Gamma_0)^{\text{div}}$, and the inverse-image identities give the reverse inclusion.

The quotient data descend to a finite extension of K , preserving normalized heights. Since $q^{-1}(q(X)) = X$, the image $q(X)$ is closed and has finite stabilizer $\text{Stab}(X)/\text{Stab}(X)^0$. Replace G, X, Γ_0, Γ by their images and M, N, L by the quotient line bundles $\tilde{M}, \tilde{N}, \tilde{L}$, and use this height to define Γ_ϵ . It is now enough to prove (2) when X has finite stabilizer.

Abramovich's theorem makes Z_X a proper closed subset when $\dim X > 0$; for a point, Z_X is empty. Proposition 5.1 gives $\eta > 0$ and $R \geq 0$ such that every point of $(X \setminus Z_X)(\bar{K}) \cap \Gamma_\eta$ has height at most R . For $0 < \epsilon \leq \eta$, the set $(X \setminus Z_X)(\bar{K}) \cap \Gamma_\epsilon$ remains dense and has height at most R . The same density of points of height at most R holds for $\epsilon > \eta$ by inclusion. Proposition 3.1 gives (2) for the quotient, which is exactly (1) for the original variety.

Conversely, assume (2) and choose γ in (1). The connected stabilizer of $X - \gamma$ is again S , and its quotient is the image of a constant subvariety under a homomorphism with finite kernel. Hence $X - \gamma$ is special in the sense of [LY25, Theorem 1.1], with torsion translation zero. That theorem gives a dense set of points of arbitrarily small height on $X - \gamma$. Translation by $\gamma \in \Gamma$ proves (1). $\square$

Proof of Corollary 1.2. Induct on $\dim X$. If X is Γ -special, take $Z_1 = X$. Otherwise Theorem 1.1 gives $\epsilon > 0$ such that the closure of $X(\bar{K}) \cap \Gamma_\epsilon$ is proper. An empty closure requires no exceptional subvarieties; this includes the dimension-zero case. Otherwise apply induction to its finitely many irreducible components and take the minimum of ϵ and their positive

thresholds. Every point in the resulting neighborhood lies in a component and hence in one of its exceptional subvarieties. $\square$

5.3. Positive characteristic. In this subsection, let k have characteristic $p > 0$, with the other field and height conventions unchanged. Stabilizers are given their reduced scheme structure, as in [LY25, §1.1]. The implication from Γ -specialness to the density of $X(\overline{K}) \cap \Gamma_\epsilon$ for every $\epsilon > 0$ still follows from [LY25, Theorem 1.1]. When Γ_0 has rank zero, this theorem also gives the converse: $\Gamma = \Gamma_0^{\text{div}}$ is the torsion subgroup, and adding torsion preserves canonical heights.

Lemma 2.1 remains valid with the F/k -trace, since the Poincaré model and integral boundary degrees do not require separability. Lemma 2.2 and Corollary 2.3 then apply unchanged. Together with geometric Bogomolov for a point, they prove Theorem 1.1 when $\dim X = 0$. The constant term in a point of height zero is expressed through the trace map, without identifying its image with a constant subgroup.

For a split torus, Proposition 3.1 also holds in positive characteristic. The torus is constant, the incidence estimates hold in every characteristic, and generic finiteness suffices for the argument using a big line bundle in Lemma 3.5. Homomorphisms of split tori are given by integer matrices on their character groups, so subtori and homomorphisms between them are defined over k . The arguments for descent and for the translating point therefore do not require the characteristic-zero identification of a general semi-abelian trace image.

Proposition 5.1 and Theorem 4.1 fail in positive characteristic, as the classical Frobenius solutions of $u + v = 1$ show; see [DM12, p. 1046, (1.3)–(1.4)].

Example 5.2. Let $K = k(t)$ and $G = \mathbb{G}_m^2$. On $\overline{G} = (\mathbb{P}^1)^2$, take $L = M = \mathcal{O}(2, 2)$ and $N = 0$ on the trivial abelian quotient. Thus $\widehat{h}_L(u, v) = 2h_{\mathcal{O}(1)}(u) + 2h_{\mathcal{O}(1)}(v)$, with the heights computed over $\mathbb{P}_k^1$ using $\mathcal{O}(1)$ as the base polarization. Set

$$X = \{(u, v) : u + v = 1\} \subset G, \quad P = (t, 1 - t), \quad \Gamma = \langle P \rangle^{\text{div}}.$$

A one-dimensional torus coset contained in X would equal X . Its coordinates would have the form az^r, bz^s for $a, b \in \overline{K}^*$ and integers r, s not both zero, whose sum cannot be identically 1. Hence Z_X is empty and the stabilizer is finite. Frobenius gives

$$[p^n]P = (t^{p^n}, 1 - t^{p^n}) \in X \cap \Gamma, \quad \widehat{h}_L([p^n]P) = p^n \widehat{h}_L(P) \longrightarrow \infty.$$

This disproves the positive-characteristic analogue of Proposition 5.1. It also disproves that of Theorem 4.1, since both expressions in Theorem 4.1(2) vanish for any two of these points. The curve is constant and therefore Γ -special, so it does not contradict the specialness conclusion.

The forward implication of Theorem 1.1 for general semi-abelian varieties in positive characteristic is not established by these arguments.

REFERENCES

- [Abr94] D. Abramovich. Subvarieties of semi-abelian varieties. *Compositio Math.*, 90(1):37–52, 1994.
- [BBP18] F. Benoist, E. Bouscaren, and A. Pillay. On function field Mordell–Lang: the semi-abelian case and the socle theorem. *Proc. Lond. Math. Soc.* (3), 116(1):182–208, 2018.
- [CGHX21] S. Cantat, Z. Gao, P. Habegger, and J. Xie. The geometric Bogomolov conjecture. *Duke Math. J.*, 170(2):247–277, 2021.

- [Con06] B. Conrad. Chow's K/k -image and K/k -trace, and the Lang–Néron theorem. *Enseign. Math. (2)*, 52(1–2):37–108, 2006.
- [DM12] H. Derksen and D. Masser. Linear equations over multiplicative groups, recurrences, and mixing I. *Proc. Lond. Math. Soc. (3)*, 104(5):1045–1083, 2012.
- [Dil20] G. A. Dill. Generalized Vojta–Rémond inequality. *Int. J. Number Theory*, 16(1):107–120, 2020.
- [Fal83] G. Faltings. Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. *Invent. Math.*, 73(3):349–366, 1983. Erratum: *Invent. Math.*, 75(2):381, 1984.
- [Fal91] G. Faltings. Diophantine approximation on abelian varieties. *Ann. of Math. (2)*, 133(3):549–576, 1991.
- [Fal94] G. Faltings. The general case of S. Lang’s conjecture. In *Barsotti Symposium in Algebraic Geometry* (Abano Terme, 1991), Perspectives in Mathematics, vol. 15, pp. 175–182. Academic Press, San Diego, CA, 1994.
- [Fuk10] L. Fukshansky. Algebraic points of small height missing a union of varieties. *J. Number Theory*, 130(10):2099–2118, 2010.
- [Ge24] T. Ge. Uniform Mordell–Lang plus Bogomolov. *Int. Math. Res. Not. IMRN*, 2024(9):7360–7378, 2024.
- [GH19] Z. Gao and P. Habegger. Heights in families of abelian varieties and the Geometric Bogomolov Conjecture. *Ann. of Math. (2)*, 189(2):527–604, 2019.
- [Gra65] H. Grauert. Mordells Vermutung über rationale Punkte auf algebraischen Kurven und Funktionenkörper. *Publ. Math. Inst. Hautes Études Sci.*, 25:131–149, 1965.
- [Gub07] W. Gubler. The Bogomolov conjecture for totally degenerate abelian varieties. *Invent. Math.*, 169(2):377–400, 2007.
- [HLY26] Z. Han, W. Luo, and J. Yu. Uniform Mordell–Lang conjecture for semi-abelian varieties. *Preprint*, [arXiv:2609.17233](#), 2026.
- [Hru96] E. Hrushovski. The Mordell–Lang conjecture for function fields. *J. Amer. Math. Soc.*, 9(3):667–690, 1996.
- [Hul26] N. Hultberg. New gap principle for semi-abelian varieties using globally valued fields. *Preprint*, [arXiv:2601.04972](#), 2026.
- [Lan60] S. Lang. Integral points on curves. *Publ. Math. Inst. Hautes Études Sci.*, 6:27–43, 1960.
- [Laz04] R. Lazarsfeld. Positivity in algebraic geometry I. *Ergebnisse der Mathematik und ihrer Grenzgebiete*, vol. 48. Springer-Verlag, Berlin, 2004.
- [Liu24] Long Liu. Chow trace of 1-motives and the Lang–Néron groups. *Preprint*, [arXiv:2404.13200](#), 2024. Accepted by *Algebra & Number Theory*.
- [Liu26] Long Liu. Representability of smooth-site direct images and Chow trace of semi-abelian varieties. *In preparation*.
- [LY25] Wenbin Luo and Jiawei Yu. Geometric Bogomolov conjecture for semiabelian varieties. *Preprint*, [arXiv:2505.07193](#), 2025.
- [Man63] Yu. I. Manin. Rational points on algebraic curves over function fields. *Izv. Akad. Nauk SSSR Ser. Mat.*, 27(6):1395–1440, 1963. Correction: *Letter to the Editor, Math. USSR-Izv.*, 34(2):465–466, 1990.
- [McQ95] M. McQuillan. Division points on semi-abelian varieties. *Invent. Math.*, 120(1):143–159, 1995.
- [Mor01] A. Moriwaki. A generalization of conjectures of Bogomolov and Lang over finitely generated fields. *Duke Math. J.*, 107(1):85–102, 2001.
- [Poo99] B. Poonen. Mordell–Lang plus Bogomolov. *Invent. Math.*, 137(2):413–425, 1999.
- [PS22] B. Poonen and K. Slavov. The exceptional locus in the Bertini irreducibility theorem for a morphism. *Int. Math. Res. Not. IMRN*, 2022(6):4503–4513, 2022.
- [Rem00] G. Rémond. Inégalité de Vojta en dimension supérieure. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 29(1):101–151, 2000.
- [Rem01] G. Rémond. Sur le théorème du produit. *J. Théor. Nombres Bordeaux*, 13(1):287–302, 2001.
- [Rem03] G. Rémond. Approximation diophantienne sur les variétés semi-abéliennes. *Ann. Sci. École Norm. Sup. (4)*, 36(2):191–212, 2003.
- [Rem05] G. Rémond. Inégalité de Vojta généralisée. *Bull. Soc. Math. France*, 133(4):459–495, 2005.
- [Sil11] J. H. Silverman. Height estimates for equidimensional dominant rational maps. *J. Ramanujan Math. Soc.*, 26(2):145–163, 2011.
- [Ull98] E. Ullmo. Positivité et discrétion des points algébriques des courbes. *Ann. of Math. (2)*, 147(1):167–179, 1998.
- [Voj96] P. Vojta. Integral points on subvarieties of semi-abelian varieties, I. *Invent. Math.*, 126(1):133–181, 1996.

- [XY22] J. Xie and X. Yuan. Geometric Bogomolov conjecture in arbitrary characteristics. *Invent. Math.*, 229(2):607–637, 2022.
- [Yam18] K. Yamaki. Trace of abelian varieties over function fields and the geometric Bogomolov conjecture. *J. Reine Angew. Math.*, 741:133–159, 2018.
- [Yua25] X. Yuan. On Vojta’s proof of the Mordell conjecture. *Preprint*, [arXiv:2508.11888](#), 2025.
- [Zha98] S.-W. Zhang. Equidistribution of small points on abelian varieties. *Ann. of Math. (2)*, 147(1):159–165, 1998.
- [Zha00] S.-W. Zhang. Distribution of almost division points. *Duke Math. J.*, 103(1):39–46, 2000.

INSTITUTE FOR THEORETICAL SCIENCES, WESTLAKE UNIVERSITY, HANGZHOU, ZHEJIANG, 310030, CHINA
Email address: liulong@westlake.edu.cn