

# THE SECOND MOMENT OF THE NUMBER OF INTEGRAL POINTS ON ELLIPTIC CURVES IS BOUNDED

LEVENT ALPOGE AND WEI HO

**ABSTRACT.** In this paper, we show that the second moment of the number of integral points on elliptic curves over  $\mathbb{Q}$  is bounded. In particular, we prove that, for any  $0 < s < \log_2 5 = 2.3219\dots$ , the  $s$ -th moment of the number of integral points is bounded for many families of elliptic curves — e.g., for the family of all integral short Weierstrass curves ordered by naive height, for the family of only minimal such Weierstrass curves, for the family of semistable curves, or for subfamilies thereof defined by finitely many congruence conditions. For certain other families of elliptic curves, such as those with a marked point or a marked 2-torsion point, the same methods show that for  $0 < s < \log_2 3 = 1.5850\dots$ , the  $s$ -th moment of the number of integral points is bounded.

The main new ingredient in our proof is an upper bound on the number of integral points on an affine integral Weierstrass model of an elliptic curve depending only on the rank of the curve and the number of square divisors of the discriminant. We obtain the bound by studying a bijection first observed by Mordell between integral points on these curves and certain types of binary quartic forms. The theorems on moments then follow from Hölder's inequality, analytic techniques, and results on bounds on the average sizes of Selmer groups in the families.

## 1. INTRODUCTION

In this paper, we prove several theorems about the number of integral points on elliptic curves over  $\mathbb{Q}$ . We bound the number of integral points using only the rank of the elliptic curve and the higher order divisors of its discriminant, and, using this bound, we show that the second moment of the number of integral points on elliptic curves over  $\mathbb{Q}$  is bounded. To our knowledge, this is the first instance of an unconditional bound on higher moments for arithmetic data on elliptic curves.

We first give an explicit upper bound on the number of integral points on affine integral Weierstrass models of elliptic curves over  $\mathbb{Q}$ , depending only on the rank and the number of square divisors of the discriminant of the curve.

**Theorem 1.1.** *Let  $A, B \in \mathbb{Z}$  be such that  $\Delta_{A,B} := -16(4A^3 + 27B^2) \neq 0$ . Let  $E_{A,B}$  be the elliptic curve given by  $y^2 = x^3 + Ax + B$ . Then*

$$\#E_{A,B}(\mathbb{Z}) \ll 2^{\text{rank } E_{A,B}(\mathbb{Q})} \prod_{p^2 | \Delta_{A,B}} \left( 4 \left\lfloor \frac{v_p(\Delta_{A,B})}{2} \right\rfloor + 1 \right).$$

*In fact, the factor in the product may be taken to be the smaller of  $4 \left\lfloor \frac{v_p(\Delta_{A,B})}{2} \right\rfloor + 1$  and  $2 \cdot 10^7$ .*

Here  $v_p$  denotes the  $p$ -adic valuation for a prime  $p$ , and  $E_{A,B}(\mathbb{Z})$  is the set of integer solutions  $\{(x, y) \in \mathbb{Z}^2 : y^2 = x^3 + Ax + B\}$ . By  $f \ll g$  we mean that there is a positive absolute constant  $c > 0$  such that  $|f| \leq c \cdot |g|$ . Mordell was the first to prove the finiteness of the number of integral points on an elliptic curve (basically by the invariant-theoretic method we employ in this paper), a theorem generalized by Siegel to all curves of genus  $g \geq 1$ .

**Remark 1.2.** If  $v_p(\Delta_{A,B}) = 2$  or  $3$ , the factor in Theorem 1.1 for  $p$  may be improved to 4 (rather than 5). This results from a slightly more careful analysis of the  $p$ -adic argument at the end of Bombieri–Schmidt [BS87]; see Remark 3.3.

Previous upper bounds on the number of integral points on elliptic curves have similar shapes but are not strong enough for our theorems on moments. For example, Helfgott and Venkatesh [HV06] show that, for any  $E_{A,B}$ ,

$$\#E_{A,B}(\mathbb{Z}) \ll O(1)^{\omega(\Delta_{A,B})} \cdot (\log |\Delta_{A,B}|)^2 \cdot 1.33^{\text{rank } E_{A,B}(\mathbb{Q})}$$

where  $\omega(n)$  denotes the number of distinct prime factors of  $n$ . For minimal short Weierstrass curves  $E_{A,B}$  (i.e., such that there does not exist a prime  $p$  with  $p^4 \mid A$  and  $p^6 \mid B$ ), Silverman [Sil87] shows that

$$\#E_{A,B}(\mathbb{Z}) \ll O(1)^{\text{rank } E_{A,B}(\mathbb{Q}) + \omega_{\text{ss}}(\Delta_{A,B})}$$

where  $\omega_{\text{ss}}(\Delta_{A,B})$  denotes the number of primes of semistable bad reduction and the  $O(1)$  is on the order of  $10^{10}$ . (Since we have control on the average size of  $n$ -Selmer groups only for small  $n$ , and thus control on the average size of  $n^{\text{rank } E_{A,B}(\mathbb{Q})}$  only for small  $n$ , this bound is unsuitable for our application.) The strongest bound for minimal short Weierstrass curves is by Hindry and Silverman [HS88]:

$$\#E_{A,B}(\mathbb{Z}) \ll O(1)^{\text{rank } E_{A,B}(\mathbb{Q}) + \sigma_{A,B}} \quad (1)$$

where  $\sigma_{A,B} := \frac{\log |\Delta_{A,B}|}{\log N_{A,B}}$  is the Szpiro ratio of  $E_{A,B}$  (here  $N_{A,B}$  denotes the conductor of  $E_{A,B}$ ). Since the ABC conjecture implies that the Szpiro ratio is at most  $6 + o(1)$ , the Hindry–Silverman bound (1) implies that, conditional on ABC and uniform boundedness of ranks, the number of integral points is uniformly bounded. (In fact, all one needs is Lang’s conjecture that  $\hat{h}(P) \gg h(E_{A,B})$  for non-torsion  $P \in E_{A,B}(\mathbb{Q})$ .)<sup>1</sup>

We use Theorem 1.1 to prove that the second moment of the number of integral points on elliptic curves over  $\mathbb{Q}$  is bounded. In particular, we consider the family  $\mathcal{F}_{\text{univ}}$  of all integral Weierstrass models

$$y^2 = x^3 + Ax + B$$

of elliptic curves over  $\mathbb{Q}$ , where  $A, B \in \mathbb{Z}$  with  $\Delta_{A,B} \neq 0$ , and order this family by *naive height*

$$H(E_{A,B}) = H(A, B) := \max(4|A|^3, 27B^2).$$

Not only do we prove that the second moment of the number of integral points in this family is bounded, we obtain the following slightly stronger result:

**Theorem 1.3.** *If  $0 < s < \log_2 5 = 2.3219\dots$ , we have*

$$\text{Avg}(|E_{A,B}(\mathbb{Z})|^s) \ll_s 1 \quad (2)$$

where the average is taken over all elliptic curves in  $\mathcal{F}_{\text{univ}}$  ordered by naive height.

More precisely, let

$$\mathcal{F}_{\text{univ}}^{\leq T} := \{(A, B) : \Delta_{A,B} \neq 0, H(A, B) \leq T\}$$

parametrize all integral Weierstrass models  $E_{A,B}$  of elliptic curves with naive height up to  $T$ . Then there exists a constant  $C_s$ , depending only on  $s$ , such that

$$\limsup_{T \rightarrow \infty} \frac{\sum_{(A,B) \in \mathcal{F}_{\text{univ}}^{\leq T}} |E_{A,B}(\mathbb{Z})|^s}{|\mathcal{F}_{\text{univ}}^{\leq T}|} < C_s.$$

In fact, we may take  $C_s$  to be  $O(1)^{20(\log_2 5 - s)^{-1}}$ . One expects that elliptic curves should have no “unexpected points” on average, i.e., that all these moments should be 0 (since we are not counting the point at infinity). In [Alp14], it is proved that (2) holds for  $0 < s < \log_3 5 = 1.4649\dots$  (and

<sup>1</sup>Alternatively, Abramovich [Abr97] has shown that the Lang–Vojta conjecture for varieties of log general type implies uniform boundedness of the number of  $S$ -integral points on a stably minimal model of an elliptic curve.

thus by taking  $s = 1$ , that the average number of integral points is bounded, a result also proved by D. Kim [Kim15]).

**Remark 1.4.** A related but different question is to show that most elliptic curves have very few integral points; perhaps the strongest known result in this direction is that 80% of curves in  $\mathcal{F}_{\text{univ}}$  have at most 2 integral points (by combining the fact that 100% of rank 1 curves in  $\mathcal{F}_{\text{univ}}$  have at most 2 points [Alp14, Lemma 20] with Bhargava–Shankar’s result [BS13] that at least 80% of curves in  $\mathcal{F}_{\text{univ}}$  have rank 0 or 1). Note that these bounds do not imply that the average number of integral points is bounded, since it is a priori possible that there is some small exceptional subset in which the curves have an enormous number of points.

**Remark 1.5.** Theorem 1.3 gives a bound on  $\text{Avg}(|E_{A,B}(\mathbb{Z})|^s)$  for  $s < \log_2 5$  by using Bhargava–Shankar’s bound on the average size of 5-Selmer groups over this family [BS13]. If a bound on the average size of  $n$ -Selmer groups over this family were known, the same argument would yield a similar bound for all  $s < \log_2(n)$ .

**Method of proof.** Theorem 1.1 follows from studying a bijection first observed by Mordell between integral points on an integral Weierstrass model  $E_{A,B}$  of an elliptic curve and binary quartics of the form  $X^4 + 6cX^2Y^2 + 8dXY^3 + eY^4$  with  $c, d, e \in \mathbb{Z}$  and invariants  $I = -48A$  and  $J = -1728B$ . The natural map taking the integral point to an element of the 2-Selmer group of the elliptic curve translates precisely to taking the corresponding binary quartic to its  $\text{PGL}_2(\mathbb{Q})$ -equivalence class. By working explicitly (and using results of Bombieri–Schmidt [BS87] and Evertse [Eve97] on Thue equations), we bound the size of a fibre by

$$\ll \prod_{p^2 | \Delta_{A,B}} \min \left\{ 4 \left\lfloor \frac{v_p(\Delta_{A,B})}{2} \right\rfloor + 1, 2 \cdot 10^7 \right\}.$$

The image lies in  $E_{A,B}(\mathbb{Q})/2E_{A,B}(\mathbb{Q})$ , whose size is at most  $\leq 4 \cdot 2^{\text{rank } E_{A,B}(\mathbb{Q})}$ , giving the theorem.

Theorem 1.3 follows fairly straightforwardly from Theorem 1.1, Hölder’s inequality, standard analytic techniques, and knowledge of bounds on the average sizes of 5-Selmer groups in this family from Bhargava–Shankar [BS13]. We also obtain similar bounds on the second moments of the number of integral points on elliptic curves in so-called *large* families (see Definition 3.4), for which average sizes of 5-Selmer groups are bounded by [BS13]. For example, for the family  $\mathcal{F}_{\text{min}}$  of *minimal* Weierstrass models (the subset of  $\mathcal{F}_{\text{univ}}$  where there does not exist a prime  $p$  with  $p^4 \mid A$  and  $p^6 \mid B$ ), or for the family  $\mathcal{F}_{\text{ss}}$  of all semistable elliptic curves, we obtain the same result as in Theorem 1.3, when the curves are ordered by naive height.

For certain other families of elliptic curves, e.g., the family

$$\mathcal{F}_1 := \{y^2 + d_3y = x^3 + d_2x^2 + d_4x : d_2, d_3, d_4 \in \mathbb{Z}, \Delta \neq 0\}$$

of elliptic curves in Weierstrass form with a marked point at  $(0, 0)$ , ordered by an analogous notion of height, we also find that the average (and the  $s$ -moments for  $0 < s < \log_2 3$ ) of the number of integral points is bounded. These types of results follow from the same techniques as for Theorem 1.3 and bounds on the average 3-Selmer group size from [BH18].

**Acknowledgments.** We thank Manjul Bhargava, Arul Shankar, and Joe Silverman for helpful comments and conversations. LA and WH were supported by the NSF GRFP and NSF grant DMS-1701437, respectively.

## 2. BINARY QUARTIC FORMS AND INTEGRAL POINTS ON ELLIPTIC CURVES

**2.1. Preliminaries on binary quartic forms.** Given a binary quartic form

$$f(X, Y) = aX^4 + bX^3Y + cX^2Y^2 + dXY^3 + eY^4 \tag{3}$$

with coefficients in  $\mathbb{Q}$ , the group  $\mathrm{SL}_2(\mathbb{Q})$  naturally acts by linear substitutions of the variables, i.e., for  $g \in \mathrm{SL}_2(\mathbb{Q})$ , one has

$$g \cdot f(X, Y) = f((X, Y) \cdot g). \quad (4)$$

There exist degree 2 and 3 polynomial invariants  $I$  and  $J$  that generate the  $\mathrm{SL}_2(\mathbb{Q})$ -invariant ring as a polynomial ring. The standard normalizations of  $I$  and  $J$  are as follows:

$$\begin{aligned} I &= 12ae - 3bd + c^2, \\ J &= 72ace - 27ad^2 - 27b^2e + 9bcd - 2c^3. \end{aligned}$$

The discriminant  $\Delta(f) = \frac{1}{27}(4I^3 - J^2)$  of  $f$  is a polynomial invariant with integer coefficients. It is well known that if  $\Delta(f)$  is nonzero, then the double cover  $Z^2 = f(X, Y)$  of  $\mathbb{P}^1$  is a genus one curve with Jacobian isomorphic to the elliptic curve

$$E : y^2 = x^3 - \frac{I}{3}x - \frac{J}{27}.$$

Conversely, over any field  $K$ , a smooth genus one curve with a rational degree 2 divisor or line bundle (thereby giving a degree 2 map to  $\mathbb{P}^1$ ) has a model of the form  $Z^2 = f(X, Y)$  for a binary quartic form  $f$  over  $K$ .

We say that a binary quartic form (3) is *integral* if  $a, b, c, d, e \in \mathbb{Z}$  and *integer-matrix* if additionally 4 divides  $b$  and  $d$  and 6 divides  $c$ . Both conditions are preserved by the action of  $\mathrm{SL}_2(\mathbb{Z})$ . For an integer-matrix binary quartic form  $f$ , there are polynomial invariants  $I'(f), J'(f)$  with *integral* coefficients such that  $12I' = I$  and  $432J' = J$ , so the elliptic curve associated to  $f$  is isomorphic to

$$y^2 = x^3 - 4I'x - 16J'. \quad (5)$$

In the sequel, we will mostly work with binary quartics of a special type, so we name them as follows:

**Definition 2.1.** We say a binary quartic form (3) is *flattened* if it is integral and monic with no  $X^3Y$ -coefficient, i.e., if  $a = 0$ ,  $b = 1$ , and  $c, d, e \in \mathbb{Z}$ .

**2.2. Mordell's construction.** In [Mor69, Chapter 25], Mordell shows that, given an integral affine Weierstrass model of an elliptic curve  $y^2 = x^3 + Ax + B$  with an integral point, there exists an integer-matrix binary quartic form  $f(X, Y)$  and  $p, q \in \mathbb{Z}$  such that  $f(p, q) = 1$  and  $I'(f) = -4A$  and  $J'(f) = -4B$ ; however, his construction is not explicit. Conversely, given an integer-matrix binary quartic form  $f(X, Y)$  such that  $I'$  and  $J'$  are multiples of 4 and  $p, q \in \mathbb{Z}$  such that  $f(p, q) = 1$ , one may explicitly produce (using covariants of  $f$ ) an integral point on the elliptic curve (5). In the next two subsections, we give a geometric explanation of Mordell's construction, which yields an explicit construction of a monic integer-matrix binary quartic form associated to an integral point on an elliptic curve.

Let  $\mathcal{E}$  be an elliptic curve over  $\mathbb{Q}$  with affine integral Weierstrass model

$$E_{A,B} : y^2 = x^3 + Ax + B \quad (6)$$

with  $A, B \in \mathbb{Z}$ . Let  $O$  denote the point at infinity. Given a point  $P = (x_0, y_0) \in E_{A,B}(\mathbb{Q})$ , the degree 2 divisor  $O + P$  induces a map from  $\mathcal{E}$  to  $\mathbb{P}^1$  as a double cover ramified at four (not necessarily rational) points. In other words, we obtain a rational binary quartic form, which is easily computed [CFS10, BH18]:

$$f(X, Y) = X^4 - 6x_0X^2Y^2 + 8y_0XY^3 + (-4A - 3x_0^2)Y^4. \quad (7)$$

It is easy to check that  $I'(f) = -4A$  and  $J'(f) = -4B$ .

Conversely, given a binary quartic  $f(X, Y) = X^4 + 6cX^2Y^2 + 4dXY^3 + eY^4$ , we may easily solve for the coefficients of the elliptic curve and the integral point by equating the coefficients with (7). We obtain the elliptic curve

$$E : y^2 = x^3 - \frac{3c^2 + e}{4}x + \frac{c^3 + d^2 - ce}{4}$$

which contains the point  $P = (x_0, y_0) = (-c, d/2)$ . Assuming that  $I'(f) = 3c^2 + e$  and  $J'(f) = -c^3 - d^2 + ce$  are both divisible by 4, we immediately have that  $d$  must be even, in which case the elliptic curve  $E$  and the point  $P$  both have integral coefficients.

It is clear that these constructions are inverse to one another. We thus obtain the explicit maps for the bijection in the following theorem:

**Theorem 2.2** (Mordell). *The following two sets are in bijection:*

- *integral affine Weierstrass models  $y^2 = x^3 + Ax + B$  of elliptic curves with integral points  $(x_0, y_0)$ ,*
- *binary quartics  $X^4 + 6cX^2Y^2 + 8dXY^3 + eY^4$  with  $c, d, e \in \mathbb{Z}$  and  $e \equiv c^2 \pmod{4}$ .*

Note that the binary quartics in Theorem 2.2 are flattened.

**2.3. Binary quartics with representations of 1.** We now relate the sets in Theorem 2.2 with binary quartic forms with representations of 1, which is Mordell's original correspondence [Mor69, Chapter 25]. This subsection is not needed for proving the main theorems in this paper, but we include it to give a more modern interpretation of Mordell's work.

We show that integer-matrix binary quartic forms  $f(X, Y)$  with a representation of 1 (i.e., with  $p, q \in \mathbb{Z}$  with  $f(p, q) = 1$ ) may be transformed, under the standard action of  $\mathrm{SL}_2(\mathbb{Z})$ , to flattened integer-matrix binary quartics. An element  $g \in \mathrm{SL}_2(\mathbb{Z})$  acts on  $f(X, Y)$  by linear transformations as in (4) and on  $(p, q)$  satisfying  $f(p, q) = 1$  by  $(p, q) \cdot g$ .

**Lemma 2.3.** *There is a bijection between flattened integer-matrix binary quartics*

$$X^4 + 6cX^2Y^2 + 4dXY^3 + eY^4 \tag{8}$$

*with  $c, d, e \in \mathbb{Z}$  and  $\mathrm{SL}_2(\mathbb{Z})$ -equivalence classes of triples  $(f, p, q)$ , where  $f$  is an integer-matrix binary quartic form and  $p, q \in \mathbb{Z}$  with  $f(p, q) = 1$ .*

*Furthermore, restricting to flattened integer-matrix binary quartics where  $d$  is even and  $c^2 \equiv e \pmod{4}$  gives a bijection with triples  $(f, p, q)$  where  $I'(f)$  and  $J'(f)$  are divisible by 4.*

*Proof.* Given an integer-matrix binary quartic form  $f(X, Y)$  and  $p, q \in \mathbb{Z}$  with  $f(p, q) = 1$ , because  $p$  and  $q$  must be relatively prime, there exist integers  $\alpha$  and  $\beta$  with  $\alpha p + \beta q = 1$ . Since the action of  $\begin{pmatrix} \alpha & \beta \\ -q & p \end{pmatrix}$  takes  $(p, q)$  to  $(1, 0)$ , there exists an  $\mathrm{SL}_2(\mathbb{Z})$ -transformation taking  $f$  to a monic integer-matrix binary quartic form. Then ‘‘completing the quartic’’ (which is possible over  $\mathbb{Z}$  because of the coefficients of 4 and 6) shows that there exists a unique  $\mathrm{SL}_2(\mathbb{Z})$ -transformation of  $f$  giving a binary quartic of the form (8).

Given two binary quartics  $f$  and  $f'$  of the above form, each with the representation  $(p, q) = (1, 0)$  of 1, it is easy to check that there is no nontrivial element of  $\mathrm{SL}_2(\mathbb{Z})$  taking  $(f, 1, 0)$  to  $(f', 1, 0)$ .

The last statement follows trivially since for the binary quartic (8), we compute  $I' = 3c^2 + e$  and  $J' = -c^3 - d^2 + ce$ .  $\square$

Combining Lemma 2.3 with Theorem 2.2, we have the following:

**Corollary 2.4.** *The following sets are in bijection:*

- (i) *integral affine Weierstrass models  $y^2 = x^3 + Ax + B$  of elliptic curves with integral points  $(x_0, y_0)$ ,*
- (ii) *binary quartics  $X^4 + 6cX^2Y^2 + 8dXY^3 + eY^4$  with  $c, d, e \in \mathbb{Z}$  and  $e \equiv c^2 \pmod{4}$ ,*
- (iii)  *$\mathrm{SL}_2(\mathbb{Z})$ -equivalence classes of triples  $(f, p, q)$ , where  $f(X, Y)$  is an integer-matrix binary quartic form with  $4 \mid I'(f)$  and  $4 \mid J'(f)$  and  $p, q \in \mathbb{Z}$  with  $f(p, q) = 1$ .*

## 3. COUNTING INTEGRAL POINTS ON ELLIPTIC CURVES

**3.1. Integral points and Selmer elements.** Given an elliptic curve  $\mathcal{E}$  over  $\mathbb{Q}$  with an integral affine Weierstrass model  $E_{A,B}$  of the form (6), we consider the sequence of maps

$$\Psi: E_{A,B}(\mathbb{Z}) \hookrightarrow \mathcal{E}(\mathbb{Q}) \rightarrow \mathcal{E}(\mathbb{Q})/2\mathcal{E}(\mathbb{Q}) \xrightarrow{\xi} \text{Sel}_2(\mathcal{E}) \quad (9)$$

where  $E_{A,B}(\mathbb{Z})$  denotes the integral points on  $E_{A,B}$  and  $\text{Sel}_2(\mathcal{E})$  is the 2-Selmer group of  $\mathcal{E}$ .

It is well known that elements of  $\text{Sel}_2(\mathcal{E})$  may be represented as binary quartic forms  $f(X, Y)$  over  $\mathbb{Q}$  such that the Jacobian of the associated genus one curve  $C(f): Z^2 = f(X, Y)$  is isomorphic to  $\mathcal{E}$  and  $C$  is locally soluble. More precisely, elements of  $\text{Sel}_2(\mathcal{E})$  are in bijection with  $\text{PGL}_2(\mathbb{Q})$ -equivalence classes of such binary quartic forms (see, e.g., [BSD63, BS15, BH16]). The  $\text{PGL}_2(\mathbb{Q})$ -action on binary quartic forms is induced from the following twisted action of  $\text{GL}_2(\mathbb{Q})$  on binary quartics: for  $g \in \text{GL}_2(\mathbb{Q})$  and a binary quartic  $f(X, Y)$ , we have  $g \cdot f(X, Y) = (\det g)^{-2} f((X, Y) \cdot g)$ . The ring of  $\text{PGL}_2(\mathbb{Q})$ -invariants is still the polynomial ring generated by  $I$  and  $J$ .

The map  $\xi: \mathcal{E}(\mathbb{Q})/2\mathcal{E}(\mathbb{Q}) \hookrightarrow \text{Sel}_2(\mathcal{E})$  sends a rational point  $P \in \mathcal{E}(\mathbb{Q})$  to the rational binary quartic form arising from the degree 2 map  $\mathcal{E} \rightarrow \mathbb{P}^1$  given by the divisor  $O + P$  (as described in §2.2). The composition  $\Psi$  of the maps in (9) is thus given by one direction of the bijection in Theorem 2.2, from an integral point  $P = (x_0, y_0) \in E_{A,B}(\mathbb{Z})$  to the  $\text{PGL}_2(\mathbb{Q})$ -equivalence class of the corresponding binary quartic form  $f_P(X, Y) := X^4 - 6x_0X^2Y^2 + 8y_0XY^3 + (-4A - 3x_0^2)Y^4$ . Note that the genus one curve  $C(f_P)$  associated to such a form (in fact, any monic binary quartic form) is automatically globally soluble; indeed,  $f_P(1, 0) = 1$  gives a rational solution. This is not surprising since, by construction, the image of  $P$  in  $\text{Sel}_2(\mathcal{E})$  lies in the subset of globally soluble forms, namely the image of  $\mathcal{E}(\mathbb{Q})/2\mathcal{E}(\mathbb{Q})$ .

Writing  $\mathcal{E}(\mathbb{Q}) \cong \mathbb{Z}^{\text{rk} \mathcal{E}(\mathbb{Q})} \oplus \mathcal{E}(\mathbb{Q})_{\text{tors}}$  and noting that  $|\mathcal{E}(\mathbb{Q})_{\text{tors}}| \ll 1$  by Mazur's theorem [Maz77], we see that  $|\mathcal{E}(\mathbb{Q})/2\mathcal{E}(\mathbb{Q})| \ll 2^{\text{rk} \mathcal{E}(\mathbb{Q})}$  (in fact,  $\leq 4 \cdot 2^{\text{rk} \mathcal{E}(\mathbb{Q})}$ ). Hence the image of  $\xi$ , and thus the image of the composition map  $\Psi$ , is of size at most

$$\ll 2^{\text{rk}(\mathcal{E}(\mathbb{Q}))}.$$

Therefore, to prove Theorem 1.1, it suffices to show that the size of each fibre of the map  $\Psi$  is bounded as follows:

**Proposition 3.1.** *Let  $f(X, Y) = X^4 + a_2X^2Y^2 + a_3XY^3 + a_4Y^4 \in \mathbb{Z}[X, Y]$  be a flattened binary quartic form. The number of elements  $\gamma \in \text{PGL}_2(\mathbb{Q})$  such that  $\gamma \cdot f$  is flattened is*

$$\ll \prod_{p^2 | \Delta(f)} \min \left\{ 4 \left\lfloor \frac{v_p(\Delta(f))}{2} \right\rfloor + 1, 2 \cdot 10^7 \right\}.$$

To prove Proposition 3.1, we first establish properties of any  $\gamma \in \text{PGL}_2(\mathbb{Q})$  that sends a flattened binary quartic form  $f$  to another flattened form. We then show that each such  $\gamma$  gives rise to a solution of a Thue equation, and invoke the fact that the number of such solutions is bounded. For example, in the simplest case when the discriminant  $\Delta(f)$  is as squarefree as possible<sup>2</sup>, then the size of the fiber of  $\Psi$  is bounded by  $4 \lfloor \frac{12}{2} \rfloor + 1 = 25$  times the number of solutions to  $f(x, y) = 1$ , which is uniformly bounded by  $2 \cdot 10^7$  for all  $f$  [Eve97]. (Once  $\Delta(f) \gg 1$ , Akhtari [Akh12] gives a much better bound of 26 for the number of solutions.)

**Lemma 3.2.** *Let  $f(X, Y) = X^4 + a_2X^2Y^2 + a_3XY^3 + a_4Y^4 \in \mathbb{Z}[X, Y]$  be a flattened binary quartic form. For any  $\gamma \in \text{PGL}_2(\mathbb{Q})$  such that  $\gamma \cdot f$  is flattened, if we write  $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  with  $a, b, c, d \in \mathbb{Z}$  and  $\gcd(a, b, c, d) = 1$ , then we have*

- (i)  $\gcd(a, b) = 1$ , and

<sup>2</sup>If  $f$  is a binary quartic form associated to an elliptic curve  $E_{A,B}$ , then  $\Delta(f) = 2^8 \Delta_{A,B} = -2^{12}(4A^3 + 27B^2)$ , so here we mean  $2^{-12}\Delta(f)$  is squarefree and not divisible by 2.



(ii)  $f(a, b) = (\det \gamma)^2$  divides the discriminant  $\Delta(f)$  of  $f$ .

*Proof.* Let  $\gamma \in \mathrm{PGL}_2(\mathbb{Q})$  be such that  $\gamma \cdot f$  is flattened, so  $(\gamma \cdot f)(1, 0) = f(a, b)(\det \gamma)^{-2} = 1$  and  $\gamma \cdot f$  has zero subleading coefficient. We may represent any  $\gamma \in \mathrm{PGL}_2(\mathbb{Q})$  by  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  with  $a, b, c, d \in \mathbb{Z}$  and  $\gcd(a, b, c, d) = 1$ . Let  $g := \gcd(a, b)$ , and write  $a = g\alpha$  and  $b = g\beta$  with  $\alpha, \beta \in \mathbb{Z}$ , so  $\gcd(\alpha, \beta) = 1$ . Then there exist integers  $\tilde{\alpha}, \tilde{\beta}$  such that  $\alpha\tilde{\alpha} - \beta\tilde{\beta} = 1$ . Let  $\tilde{\gamma} := \begin{pmatrix} \alpha & \beta \\ \tilde{\beta} & \tilde{\alpha} \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ , and let  $\eta := c\tilde{\alpha} - d\tilde{\beta} \in \mathbb{Z}$ . Define

$$U := \gamma\tilde{\gamma}^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} \tilde{\alpha} & -\beta \\ -\tilde{\beta} & \alpha \end{pmatrix} = \begin{pmatrix} g & 0 \\ \eta & \frac{\det \gamma}{g} \end{pmatrix},$$

implying that  $\gamma = U\tilde{\gamma}$ .

We now show that  $g$  divides all the entries of  $U$ , namely  $g^2 \mid \det \gamma$  and  $g \mid \eta$ . Let  $\tilde{f} := \tilde{\gamma} \cdot f \in \mathbb{Z}[X, Y]$ , with no twisted action necessary since  $\tilde{\gamma} \in \mathrm{SL}_2(\mathbb{Z})$ . Write

$$\tilde{f}(X, Y) =: \tilde{a}_0 X^4 + \tilde{a}_1 X^3 Y + \tilde{a}_2 X^2 Y^2 + \tilde{a}_3 X Y^3 + \tilde{a}_4 Y^4 \in \mathbb{Z}[X, Y].$$

Then  $(\gamma \cdot f)(X, Y) = (\det \gamma)^{-2}(U \cdot \tilde{f})(X, Y) = (\det \gamma)^{-2} \tilde{f}\left(gX + \eta Y, \frac{\det \gamma}{g} Y\right)$ . Expanding, we compute that the  $X^4$ -coefficient in  $\gamma \cdot f$  is

$$(\gamma \cdot f)(1, 0) = f(a, b) = g^4 f(\alpha, \beta) = \frac{g^4 \tilde{a}_0}{(\det \gamma)^2}.$$

Since it is also 1 by hypothesis, we find that  $\frac{(\det \gamma)^2}{g^4} = \tilde{a}_0 \in \mathbb{Z}$ . Thus  $g^4$  divides  $(\det \gamma)^2$ , so  $g^2$  divides  $\det \gamma$ . Now the  $X^3 Y$ -coefficient of  $\gamma \cdot f$  is

$$\frac{4g^3 \eta \cdot \tilde{a}_0 + g^2 (\det \gamma) \cdot \tilde{a}_1}{(\det \gamma)^2} = 0.$$

Substituting for  $\tilde{a}_0$ , we find that  $\tilde{a}_1 = -4 \cdot \frac{(\det \gamma) \cdot \eta}{g^3} \in \mathbb{Z}$ . Finally, the  $X^2 Y^2$ -coefficient of  $\gamma \cdot f$  is

$$\frac{6g^2 \eta^2 \cdot \tilde{a}_0 + 3g\eta (\det \gamma) \cdot \tilde{a}_1 + (\det \gamma)^2 \cdot \tilde{a}_2}{(\det \gamma)^2} = -\frac{6\eta^2}{g^2} + \tilde{a}_2$$

after substituting for  $\tilde{a}_0$  and  $\tilde{a}_1$ . Since  $\tilde{a}_2 \in \mathbb{Z}$  and this coefficient is integral as well, we deduce that  $g^2$  divides  $6\eta^2$ , so  $g$  divides  $\eta$ .

Since  $g$  divides all the entries of  $U$ , we see that  $g$  divides all the entries of  $U \cdot \tilde{\gamma} = \gamma$ , implying that  $g$  divides  $\gcd(a, b, c, d) = 1$  and thus  $g = 1$  as desired.

Substituting  $g = 1$  shows that  $\tilde{a}_1 = -4\eta \det \gamma$ , so  $\det \gamma$  divides  $\tilde{a}_1$ , and of course  $(\det \gamma)^2$  divides (since it is in fact equal to)  $\tilde{a}_0$ . We thus find that  $(\det \gamma)^2$  divides  $\Delta(\tilde{f}) = \Delta(f)$  (since every term of  $\Delta(\tilde{f})$  is a multiple of either  $\tilde{a}_0$  or  $\tilde{a}_1^2$ ).  $\square$

*Proof of Proposition 3.1.* Lemma 3.2 shows that for any  $\gamma \in \mathrm{PGL}_2(\mathbb{Q})$  (represented by  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  with  $a, b, c, d \in \mathbb{Z}$  and  $\gcd(a, b, c, d) = 1$ ) such that  $\gamma \cdot f$  is flattened, we have that  $\gcd(a, b) = 1$  and  $f(a, b) = (\det \gamma)^2$  is a square dividing  $\Delta(f)$ . We now claim that the map

$$\Phi: \{\gamma \in \mathrm{PGL}_2(\mathbb{Q}) : \gamma \cdot f \text{ is flattened}\} \rightarrow \{(a, b) \in \mathbb{Z}^2 : \gcd(a, b) = 1, f(a, b) = \square, f(a, b) \mid \Delta(f)\},$$

taking  $\gamma$  as above to  $(a, b)$ , is injective. Indeed, if  $\gamma, \gamma' \in \mathrm{PGL}_2(\mathbb{Q})$  map to the same  $(a, b) \in \mathbb{Z}^2$ , write  $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  and  $\gamma' = \begin{pmatrix} a & b \\ c' & d' \end{pmatrix}$ , and note that

$$\gamma' \gamma^{-1} = \begin{pmatrix} 1 & 0 \\ \frac{c'd - cd'}{\det \gamma} & 1 \end{pmatrix}.$$

Let  $\lambda := \frac{c'd - cd'}{\det \gamma} \in \mathbb{Q}$ . Since

$$(\gamma' \cdot f)(X, Y) = ((\gamma' \gamma^{-1}) \cdot (\gamma \cdot f))(X, Y) = (\gamma \cdot f)(X + \lambda Y, Y)$$

and both  $\gamma \cdot f$  and  $\gamma' \cdot f$  are flattened by hypothesis, it follows that  $\lambda = 0$  and so  $\gamma = \gamma'$ , as desired.

Thus, for the map  $\Phi$ , the size of the domain is bounded by the size of the codomain, which is simply

$$\sum_{\delta^2 | \Delta(f)} \#\{(a, b) \in \mathbb{Z}^2 : \gcd(a, b) = 1, f(a, b) = \delta^2\}. \quad (10)$$

We divide the set of primes  $p^2 | \Delta(f)$  into two sets to obtain a hybrid bound. Let  $S$  be the set<sup>3</sup> of primes such that  $2 \cdot 10^7 \leq 4 \left\lfloor \frac{v_p(\Delta(f))}{2} \right\rfloor + 1$ , and let

$$D := \prod_{\substack{p^2 | \Delta(f) \\ p \notin S}} p^{v_p(\Delta(f))}.$$

Given  $\delta$  such that  $\delta^2 | \Delta(f)$ , set  $\nu := \gcd(\delta, D)$  and  $\mu := \frac{\delta}{\nu}$ .

The argument of Bombieri–Schmidt in [BS87, Section VI], specifically the second-to-last paragraph, produces  $\leq 4^{\omega(\nu)}$  many quartic forms  $f_{\nu,i}$ , depending only on  $f$  and  $\nu$ , such that the number of relatively prime solutions to  $f(a, b) = \delta^2 = \nu^2 \mu^2$  is bounded above by the sum of the numbers of relatively prime solutions of  $f_{\nu,i}(a, b) = \mu^2$ . Rewriting (10) in terms of  $\mu$  and  $\nu$  gives

$$\begin{aligned} & \sum_{\delta^2 | \Delta(f)} \#\{(a, b) \in \mathbb{Z}^2 : \gcd(a, b) = 1, f(a, b) = \delta^2\} \\ &= \sum_{\nu^2 | D} \sum_{\mu^2 | \frac{\Delta(f)}{D}} \#\{(a, b) \in \mathbb{Z}^2 : \gcd(a, b) = 1, f(a, b) = \mu^2 \nu^2\} \\ &\leq \sum_{\nu^2 | D} \sum_{i=1}^{4^{\omega(\nu)}} \sum_{\mu^2 | \frac{\Delta(f)}{D}} \#\{(a, b) \in \mathbb{Z}^2 : \gcd(a, b) = 1, f_{\nu,i}(a, b) = \mu^2\}. \end{aligned} \quad (11)$$

To bound the number of solutions to  $f_{\nu,i}(a, b) = \mu^2$ , we use the following theorem of Evertse [Eve97, Theorem 1]: for a number field  $K$  with a finite set of places  $T$  containing all infinite places and a homogeneous polynomial  $g \in \mathcal{O}_K[x, y]$  of degree at least 3, the inclusion  $f(a, b) \in \mathcal{O}_{K,T}^\times$  has at most  $(5 \cdot 10^6 \deg(g))^{\#T}$  many solutions, modulo the action of  $\mathcal{O}_{K,T}^\times$ . Applying this theorem here with  $K = \mathbb{Q}$ ,  $T = S \cup \{\infty\}$ , and  $g = f_{\nu,i}$ , we see that the innermost sum of (11) is at most

$$\#\{(a, b) \in \mathbb{Z}^2 : \gcd(a, b) = 1, f_{\nu,i}(a, b) \in \mathbb{Z}[S^{-1}]^\times = \mathcal{O}_{\mathbb{Q},S}^\times\} \ll (2 \cdot 10^7)^{\#S}.$$

(Note that the condition  $\gcd(a, b) = 1$  implies that only  $(a, b)$  and  $(-a, -b)$  are solutions in the same coset under the action of  $\mathcal{O}_{K,T}^\times$ .)

Hence the size (10) of the codomain of  $\Phi$  is

$$\begin{aligned} \sum_{\delta^2 | \Delta(f)} \#\{(a, b) \in \mathbb{Z}^2 : \gcd(a, b) = 1, f(a, b) = \delta^2\} &\ll \sum_{\nu^2 | D} \sum_{i=1}^{4^{\omega(\nu)}} (2 \cdot 10^7)^{\#S} \\ &\ll (2 \cdot 10^7)^{\#S} \sum_{\nu^2 | D} 4^{\omega(\nu)} \\ &= (2 \cdot 10^7)^{\#S} \prod_{p^2 | D} \left( 4 \left\lfloor \frac{v_p(\Delta(f))}{2} \right\rfloor + 1 \right), \end{aligned}$$

completing the argument. □

<sup>3</sup>Taking  $S$  to be the empty set in this argument gives a weaker but simpler upper bound for (10), and thus for Proposition 3.1, of  $\prod_{p^2 | \Delta(f)} \left( 4 \left\lfloor \frac{v_p(\Delta(f))}{2} \right\rfloor + 1 \right)$ .



**Remark 3.3.** When  $2 \leq v_p(\Delta(f)) < 4$ , evidently  $p \notin S$ , and either  $p \nmid \nu$  (in which case there is no factor corresponding to  $p$ ) or  $v_p(\nu) = 2$ . By simply enumerating cases of  $f$  over  $\mathbb{Q}_p$  one finds that the number of disks required for [BS87, Lemma 7] is in fact at most 3, because at least two roots lie in the same residue disk modulo  $p$ . This translates into a factor of 3 corresponding to  $p$ , rather than 4, and thus gives the claim of Remark 1.2 (after applying this improved bound in the following proof of Theorem 1.1).

Having proved Proposition 3.1, Theorem 1.1 follows easily.

*Proof of Theorem 1.1.* We showed that the map  $\Psi: E_{A,B}(\mathbb{Z}) \rightarrow \mathcal{E}(\mathbb{Q})/2\mathcal{E}(\mathbb{Q}) \subseteq \text{Sel}_2(\mathcal{E})$ , taking an integral point of  $E_{A,B}$  to the  $\text{PGL}_2(\mathbb{Q})$ -equivalence class of its corresponding binary quartic form  $f$  (by Theorem 2.2), has image of size  $\ll 2^{\text{rk } \mathcal{E}(\mathbb{Q})}$ . Recall that the binary quartic  $f$  has invariants  $I = -48A$  and  $J = -1728B$ , so  $\Delta(f) = 2^8 \Delta_{A,B}$ . Thus, applying Proposition 3.1, we see that the size of a fibre of the map  $\Psi$  is bounded by

$$\ll \prod_{p^2 | \Delta_{A,B}} \min \left\{ 4 \left\lfloor \frac{v_p(\Delta_{A,B})}{2} \right\rfloor + 1, 2 \cdot 10^7 \right\}.$$

Combining the two estimates gives the theorem.  $\square$

**3.2. Bounding moments of the number of integral points on elliptic curves.** Theorem 1.3 follows from “averaging” the bound in Theorem 1.1 and analytic techniques; the additional crucial input is Bhargava–Shankar’s result that the average size of the 5-Selmer group of elliptic curves in  $\mathcal{F}_{\text{univ}}$ , ordered by naive height, is bounded [BS13]. In fact, because the average size of the 5-Selmer group is bounded in any “large” family [BS13, Theorem 31], we may prove the same result for such families as well.

**Definition 3.4.** We say a subfamily  $\mathcal{F} \subseteq \mathcal{F}_{\text{univ}}$  of elliptic curves over  $\mathbb{Q}$  is *defined by congruence conditions* if, for all primes  $p$ , there exists a closed subset  $\Sigma_p$  of  $\{(A, B) \in \mathbb{Z}_p^2 : \Delta_{A,B} \neq 0\}$  whose boundary has measure 0 such that  $E_{A,B} \in \mathcal{F}$  for  $(A, B) \in \Sigma_p$ .

For such a family  $\mathcal{F}$ , let  $\text{Inv}(\mathcal{F}) := \{(A, B) \in \mathbb{Z} \times \mathbb{Z} : E_{A,B} \in \mathcal{F}\}$  and let  $\text{Inv}_p(\mathcal{F})$  be the  $p$ -adic closure of  $\text{Inv}(\mathcal{F})$  in  $\mathbb{Z}_p^2 \setminus \{\Delta_{A,B} = 0\}$ . The set  $\text{Inv}_\infty(\mathcal{F})$  is defined to be  $\{(A, B) \in \mathbb{R}^2 : \Delta_{A,B} \bowtie 0\}$ , where  $\bowtie$  is  $>$ ,  $<$ , or  $\neq 0$  depending on whether  $\mathcal{F}$  contains only curves of positive discriminant, negative discriminant, or both, respectively. Then  $\mathcal{F}$  is a *large family* if, for all but finitely many primes  $p$ , the set  $\text{Inv}_p(\mathcal{F})$  contains all  $(A, B) \in \mathbb{Z}_p^2$  such that  $p^2 \nmid \Delta_{A,B}$ .

Examples of large families include  $\mathcal{F}_{\text{univ}}$  itself, the family  $\mathcal{F}_{\text{min}}$  of minimal Weierstrass curves, the family  $\mathcal{F}_{\text{ss}}$  of semistable elliptic curves, and any family defined by finitely many congruence conditions. We now prove the following stronger version of Theorem 1.3:

**Theorem 3.5.** *For any large family  $\mathcal{F}$  of elliptic curves and any  $0 < s < \log_2 5 = 2.3219\dots$ , we have*

$$\text{Avg}_{\mathcal{F}}(|E(\mathbb{Z})|^s) \ll_s 1$$

where the average is taken over all elliptic curves  $E$  in  $\mathcal{F}$  ordered by naive height.

*Proof of Theorem 3.5.* Let  $\mathcal{F}^{\leq T} := \{(A, B) \in \text{Inv}(\mathcal{F}) : \Delta_{A,B} \neq 0, H(A, B) \leq T\}$  represent the curves in  $\mathcal{F}$  of naive height at most  $T$ . By simply applying the bound of Theorem 1.1 and noting that  $4e + 1 \leq (e + 1)^3$  for any positive integer  $e$ , for any  $(A, B) \in \mathcal{F}^{\leq T}$  we have

$$|E_{A,B}(\mathbb{Z})|^s \ll (2^s)^{\text{rk } E_{A,B}(\mathbb{Q})} \cdot \prod_{p^2 | \Delta_{A,B}} \left( \left\lfloor \frac{v_p(\Delta_{A,B})}{2} \right\rfloor + 1 \right)^{3s}.$$

Summing over all  $(A, B)$  in  $\mathcal{F}^{\leq T}$  and then applying Hölder's inequality with dual exponent pair  $(1 + \varepsilon, 1 + \varepsilon^{-1})$  yields

$$\begin{aligned} \sum_{(A,B) \in \mathcal{F}^{\leq T}} |E_{A,B}(\mathbb{Z})|^s &\ll \sum_{(A,B) \in \mathcal{F}^{\leq T}} (2^s)^{\text{rk } E_{A,B}(\mathbb{Q})} \cdot \prod_{p^2 | \Delta_{A,B}} \left( \left\lfloor \frac{v_p(\Delta_{A,B})}{2} \right\rfloor + 1 \right)^{3s} \\ &\leq \left( \sum_{(A,B) \in \mathcal{F}^{\leq T}} \left( 2^{(1+\varepsilon)s} \right)^{\text{rk } E_{A,B}(\mathbb{Q})} \right)^{\frac{1}{1+\varepsilon}} \left( \sum_{(A,B) \in \mathcal{F}^{\leq T}} \prod_{p^2 | \Delta_{A,B}} \left( \left\lfloor \frac{v_p(\Delta_{A,B})}{2} \right\rfloor + 1 \right)^{3(1+\varepsilon^{-1})s} \right)^{\frac{\varepsilon}{1+\varepsilon}}. \end{aligned} \quad (12)$$

We bound the first term in (12) as follows. Since  $0 < s < \log_2 5$ , we may choose  $\varepsilon > 0$  such that  $(1 + \varepsilon) \cdot s < \log_2 5$ , or, equivalently,  $0 < \varepsilon < \frac{\log_2 5}{s} - 1$ . Then  $2^{(1+\varepsilon)s} \leq 5$ , so

$$\left( 2^{(1+\varepsilon) \cdot s} \right)^{\text{rk } E_{A,B}(\mathbb{Q})} \leq 5^{\text{rk } E_{A,B}(\mathbb{Q})} \leq |E_{A,B}(\mathbb{Q})/5E_{A,B}(\mathbb{Q})| \leq |\text{Sel}_5(E_{A,B})|. \quad (13)$$

Bhargava and Shankar [BS13, Theorem 31] show that  $\text{Avg}_{\mathcal{F}} |\text{Sel}_5(E_{A,B})| = 6$  for any large family  $\mathcal{F}$  ordered by naive height, and combining this average with (13) implies that

$$\sum_{(A,B) \in \mathcal{F}^{\leq T}} \left( 2^{(1+\varepsilon) \cdot s} \right)^{\text{rk } E_{A,B}(\mathbb{Q})} \leq (6 + o(1)) \cdot |\mathcal{F}^{\leq T}|. \quad (14)$$

In Lemma 3.6, we will bound the second term in (12) by showing that for all  $t \geq 0$ ,

$$\sum_{(A,B) \in \mathcal{F}^{\leq T}} \prod_{p^2 | \Delta_{A,B}} \left( \left\lfloor \frac{v_p(\Delta_{A,B})}{2} \right\rfloor + 1 \right)^t \ll_t |\mathcal{F}^{\leq T}|. \quad (15)$$

(By the Selberg–Delange method, we find that the implicit constant is  $\ll O(1)^{16^t}$ , but we need not be so precise.) The desired theorem follows from the two bounds (14) and (15).  $\square$

**Lemma 3.6.** *For any large family  $\mathcal{F}$  and all  $t \geq 0$ , we have*

$$\sum_{(A,B) \in \mathcal{F}^{\leq T}} \prod_{p^2 | \Delta_{A,B}} \left( \left\lfloor \frac{v_p(\Delta_{A,B})}{2} \right\rfloor + 1 \right)^t \ll_t |\mathcal{F}^{\leq T}|.$$

*Proof.* Let  $0 < \delta < 1$  be a parameter, to be taken to be small but  $\gg 1$  (in fact,  $\delta = \frac{1}{100}$  suffices, though one may optimize the argument and take  $\delta$  even larger). We may use the standard trick of decomposing a product over all primes into products of small and large primes:

$$\prod_{p^2 | n} \left( \left\lfloor \frac{v_p(n)}{2} \right\rfloor + 1 \right) = \left( \prod_{\substack{p^2 | n \\ p < n^\delta}} \left( \left\lfloor \frac{v_p(n)}{2} \right\rfloor + 1 \right) \right) \cdot \left( \prod_{\substack{p^2 | n \\ p \geq n^\delta}} \left( \left\lfloor \frac{v_p(n)}{2} \right\rfloor + 1 \right) \right).$$

Now note that the number of  $p \geq n^\delta$  for which  $p^2 | n$  is trivially at most  $\frac{1}{2\delta}$ , and, for each such  $p$ , we have that  $v_p(n) \leq \frac{1}{\delta}$  as well. It follows that

$$\prod_{p^2 | n, p \geq n^\delta} \left( \left\lfloor \frac{v_p(n)}{2} \right\rfloor + 1 \right) \leq \left( \frac{1}{2\delta} + 1 \right)^{\frac{1}{2\delta}} \leq \delta^{-\delta^{-1}} \ll 1. \quad (16)$$

Using the large prime bound (16) and the fact that  $|\Delta_{A,B}| \ll T$ , the sum in (15) may thus be bounded by

$$\sum_{(A,B) \in \mathcal{F}^{\leq T}} \prod_{p^2 | \Delta_{A,B}} \left( \left\lfloor \frac{v_p(\Delta_{A,B})}{2} \right\rfloor + 1 \right)^t \ll_t \sum_{(A,B) \in \mathcal{F}^{\leq T}} \prod_{\substack{p^2 | \Delta_{A,B} \\ p \ll T^\delta}} \left( \left\lfloor \frac{v_p(\Delta_{A,B})}{2} \right\rfloor + 1 \right)^t.$$

We now bound the product over small primes, first for  $\mathcal{F} = \mathcal{F}_{\text{univ}}$ . Let  $m \ll T^{\frac{1}{3}}$ . Each fibre of the natural reduction map  $\mathcal{F}_{\text{univ}}^{\leq T} \rightarrow (\mathbb{Z}/m\mathbb{Z})^2$  is of size<sup>4</sup>

$$\ll \left( \frac{T^{\frac{1}{3}}}{m} + 1 \right) \cdot \left( \frac{T^{\frac{1}{2}}}{m} + 1 \right) \ll \frac{|\mathcal{F}_{\text{univ}}^{\leq T}|}{m^2}. \quad (17)$$

Moreover, the number of  $(\bar{A}, \bar{B}) \in (\mathbb{Z}/m\mathbb{Z})^2$  such that  $\Delta_{\bar{A}, \bar{B}} = -16(4\bar{A}^3 + 27\bar{B}^2) \equiv 0 \pmod{m}$  is

$$\ll m \cdot O(1)^{\omega(m)}.$$

Indeed, in  $\mathbb{Z}/p^{v_p(m)}\mathbb{Z}$ , by Hensel's lemma<sup>5</sup> one has  $\ll 1$  solutions for  $B$  for any fixed  $A$ , and the Chinese remainder theorem then implies that, for fixed  $A \in \mathbb{Z}/m\mathbb{Z}$ , one has  $\ll O(1)^{\omega(m)}$  solutions for  $B \in \mathbb{Z}/m\mathbb{Z}$ . Therefore, for  $m \ll T^{\frac{1}{3}}$ , the number of  $(A, B) \in \mathcal{F}_{\text{univ}}^{\leq T}$  with  $m \mid \Delta_{A,B}$  is

$$\ll \frac{|\mathcal{F}_{\text{univ}}^{\leq T}|}{m^2} \cdot (m \cdot O(1)^{\omega(m)}) = |\mathcal{F}_{\text{univ}}^{\leq T}| \frac{O(1)^{\omega(m)}}{m}. \quad (18)$$

Similarly, for any large  $\mathcal{F}$ , we obtain the same bounds. Indeed, from [BS15, Proposition 3.16], for any prime  $p$ , the number of  $(A, B) \in \mathcal{F}_{\text{univ}}^{\leq T}$  with  $p^2$  dividing  $\Delta_{A,B}$  is  $O(T^{\frac{5}{6}} p^{-\frac{3}{2}})$ . Then [BS15, Theorem 3.17] implies that the number of elliptic curves with height  $\leq T$  in any large family  $\mathcal{F}$  is the product of local densities with bounded error:

$$\int_{\substack{(A,B) \in \text{Inv}_\infty(\mathcal{F}) \\ H(A,B) < T}} dA dB = \prod_p \int_{(A,B) \in \text{Inv}_p(\mathcal{F})} dA dB + o(T^{\frac{5}{6}}).$$

Since  $\mathcal{F}$  is defined by congruence conditions and contains all curves with discriminant not a multiple of  $p^2$  for almost all  $p$ , we have that  $\lim_{T \rightarrow \infty} \frac{|\mathcal{F}^{\leq T}|}{|\mathcal{F}_{\text{univ}}^{\leq T}|} \geq c_{\mathcal{F}}$  for some constant  $0 < c_{\mathcal{F}} \leq 1$  depending only on  $\mathcal{F}$ . In other words, any large family makes up a positive proportion of  $\mathcal{F}_{\text{univ}}$ . Thus, the fibers of the reduction map  $\mathcal{F}^{\leq T} \rightarrow (\mathbb{Z}/m\mathbb{Z})^2$  are also of size  $\ll |\mathcal{F}^{\leq T}| m^{-2}$  (analogously to (17)). And just as in (18), the number of  $(A, B) \in \mathcal{F}^{\leq T}$  with  $m \mid \Delta_{A,B}$  is therefore

$$\ll \frac{|\mathcal{F}^{\leq T}|}{m^2} \cdot (m \cdot O(1)^{\omega(m)}) = |\mathcal{F}^{\leq T}| \frac{O(1)^{\omega(m)}}{m}. \quad (19)$$

We will apply (19) to the case of  $m = d^2$  where the prime factors of  $d$  satisfy  $p \ll T^\delta$ , i.e.,  $d$  is  $T^\delta$ -smooth. For  $(A, B) \in \mathcal{F}^{\leq T}$ , let

$$n_{A,B} := \prod_{\substack{p^2 | \Delta_{A,B} \\ p \ll T^\delta}} p^{\left\lfloor \frac{v_p(\Delta_{A,B})}{2} \right\rfloor}.$$

<sup>4</sup>This bound is unsurprising since we expect equidistribution.

<sup>5</sup>For  $p = 2$  or  $3$ , either the valuation  $v_p(m)$  is large enough (e.g.,  $\geq 10$ ), in which case Hensel's lemma applies, or else there are  $O(1)$  many elements of  $\mathbb{Z}/p^{v_p(m)}\mathbb{Z}$  anyway.

We claim there always exists a divisor  $d_{A,B}$  of  $n_{A,B}$  (and hence  $d_{A,B}^2 \mid \Delta_{A,B}$ ) such that

$$\min(n_{A,B}, T^{\frac{1}{6}-\delta}) \ll d_{A,B} \ll T^{\frac{1}{6}} \quad (20)$$

$$\text{and} \quad \tau(n_{A,B}) \ll \tau(d_{A,B})^4, \quad (21)$$

where  $\tau$  denotes the standard number-of-divisors function.

To see this, we first note that, for any  $m \in \mathbb{Z}^+$ ,  $x \geq 1$ , and  $1 \leq y \leq m$  such that  $p \leq x$  for all primes  $p$  dividing  $m$ , there is always a divisor of  $m$  in the interval  $[y, xy]$ . Indeed, either  $m$  already lies in this interval, or we may, by removing one prime of  $m$  at a time, reduce the size of the divisor by at most a factor of  $x$  at each step. Since we eventually reach  $1 \leq y$ , we must eventually cross this interval. We will call this the ‘greedy argument’.

Now we apply this to  $n_{A,B}$  and its divisors. First, if  $n_{A,B}$  already satisfies the inequality (20) we are done and may take  $d_{A,B} := n_{A,B}$ . Otherwise, there is at least one divisor  $d \mid n_{A,B}$  with  $T^{\frac{1}{6}-\delta} \ll d \ll T^{\frac{1}{6}}$  by using the greedy argument with  $m = n_{A,B}$ ,  $x \asymp T^\delta$ , and  $y \asymp T^{\frac{1}{6}-\delta}$ . Thus we may take  $d_{A,B}$  to be a divisor  $d$  in that interval maximizing  $\tau(d)$  among all divisors in the interval:

$$d_{A,B} := \operatorname{argmax} \left\{ \tau(d) : d \mid n_{A,B}, T^{\frac{1}{6}-\delta} \ll d \ll T^{\frac{1}{6}} \right\}.$$

In other words, we have  $d_{A,B}$  divides  $n_{A,B}$  and  $T^{\frac{1}{6}-\delta} \ll d_{A,B} \ll T^{\frac{1}{6}}$ , and if  $d \mid n_{A,B}$  and  $T^{\frac{1}{6}-\delta} \ll d \ll T^{\frac{1}{6}}$ , then  $\tau(d_{A,B}) \geq \tau(d)$ .

For any  $d \mid n_{A,B}$  with  $d \ll T^{\frac{1}{6}}$ , there exists an integer  $d'$  with  $d \mid d' \mid n_{A,B}$  and  $T^{\frac{1}{6}-\delta} \ll d' \ll T^{\frac{1}{6}}$ . Indeed, if  $d \gg T^{\frac{1}{6}-\delta}$ , we take  $d' = d$ ; otherwise, if  $d \ll T^{\frac{1}{6}-\delta}$ , apply the greedy argument with  $m = d$ ,  $x \asymp T^\delta$ ,  $y \asymp \frac{T^{\frac{1}{6}-\delta}}{d}$  to find a divisor  $e \mid \frac{n}{d}$  with  $\frac{T^{\frac{1}{6}-\delta}}{d} \ll e \ll \frac{T^{\frac{1}{6}}}{d}$ , and take  $d' := de$ . Note that  $\tau(d) \leq \tau(d') \leq \tau(d_{A,B})$ .

Since  $n_{A,B}^2 \mid \Delta_{A,B}$  by definition and  $|\Delta_{A,B}| \ll T$ , we have that  $n_{A,B} \ll T^{\frac{1}{2}}$ . Moreover, applying the greedy argument at most two times, we may write  $n_{A,B} = d_{A,B} d_1 d_2 d_3$  with  $d_i \ll T^{\frac{1}{6}}$  for  $i = 1, 2, 3$ . Then  $\tau(d_i) \leq \tau(d_{A,B})$ , so  $\tau(n_{A,B}) = \tau(d_{A,B} d_1 d_2 d_3) \leq \tau(d_{A,B}) \tau(d_1) \tau(d_2) \tau(d_3) \leq \tau(d_{A,B})^4$ , yielding (21) and the claim.

Choosing one such  $d_{A,B}$  for each  $(A, B) \in \mathcal{F}^{\leq T}$ , we conclude that

$$\begin{aligned} \sum_{(A,B) \in \mathcal{F}^{\leq T}} \prod_{\substack{p^2 \mid \Delta_{A,B} \\ p \ll T^\delta}} \left( \left\lfloor \frac{v_p(\Delta_{A,B})}{2} \right\rfloor + 1 \right)^t &= \sum_{(A,B) \in \mathcal{F}^{\leq T}} \tau(n_{A,B})^t \\ &\ll_t \sum_{\substack{d \ll T^{\frac{1}{6}} \\ d \text{ } T^\delta\text{-smooth}}} \tau(d)^t \sum_{\substack{(A,B) \in \mathcal{F}^{\leq T} \\ d^2 \mid \Delta_{A,B}}} 1 \\ &\ll |\mathcal{F}^{\leq T}| \cdot \sum_{\substack{d \ll T^{\frac{1}{6}} \\ d \text{ } T^\delta\text{-smooth}}} \frac{O(1)^{\omega(d)} \tau(d)^t}{d^2} \quad \text{by (19)} \\ &\ll_t |\mathcal{F}^{\leq T}| \end{aligned}$$

as desired.  $\square$

**3.3. Other families.** The arguments in Section 3.2 may be modified appropriately to give averages or moments on the number of integral points on elliptic curves over  $\mathbb{Q}$  in some other families where we have finite upper bounds on the average  $d$ -Selmer group size, for some  $d > 2$ . These families

include the following:

$$\begin{aligned}\mathcal{F}_1 &:= \{y^2 + d_3y = x^3 + d_2x^2 + d_4x : d_2, d_3, d_4 \in \mathbb{Z}, \Delta \neq 0\} \\ \mathcal{F}_1(2) &:= \{y^2 = x^3 + d_2x^2 + d_4x : d_2, d_4 \in \mathbb{Z}, \Delta \neq 0\}\end{aligned}$$

The family  $\mathcal{F}_1$  (resp.,  $\mathcal{F}_1(2)$ ) has a marked point (resp., marked 2-torsion point) at  $(0, 0)$ , and the height  $H(E)$  of a curve  $E$  in either family is again a measure of the size of the coefficients, defined as  $\max |d_i|^{\frac{12}{i}}$ . By [BH18], the average size of the 3-Selmer group in each of these families, ordered by height, is bounded. We claim that the average number of integral points on the curves in these families is bounded, and in fact, a stronger statement holds:

**Theorem 3.7.** *For  $\mathcal{F} = \mathcal{F}_1$  or  $\mathcal{F}_1(2)$  and any  $0 < s < \log_2 3 = 1.5850\dots$ , we have*

$$\text{Avg}_{\mathcal{F}}(|E(\mathbb{Z})|^s) \ll_s 1$$

where the average is taken over all elliptic curves  $E$  in  $\mathcal{F}$  ordered by height.

*Proof.* The proof follows the same outline as that of Theorem 3.5. Let  $\mathcal{F}^{\leq T} := \{E \in \mathcal{F} : \Delta_E \neq 0, H(A, B) \leq T\}$  represent the curves in  $\mathcal{F}$  of height at most  $T$ . The bound of Theorem 1.1 and Hölder's inequality give an inequality analogous to (12):

$$\sum_{E \in \mathcal{F}^{\leq T}} |E(\mathbb{Z})|^s \ll \left( \sum_{E \in \mathcal{F}^{\leq T}} (2^{(1+\varepsilon) \cdot s})^{\text{rk } E(\mathbb{Q})} \right)^{\frac{1}{1+\varepsilon}} \left( \sum_{E \in \mathcal{F}^{\leq T}} \prod_{p^2 | \Delta_E} \left( \left\lfloor \frac{v_p(\Delta_E)}{2} \right\rfloor + 1 \right)^{3(1+\varepsilon^{-1})s} \right)^{\frac{\varepsilon}{1+\varepsilon}} \quad (22)$$

The first term is bounded as before, by choosing  $0 < \varepsilon < \frac{\log_2 3}{s} - 1$  so that

$$\left( 2^{(1+\varepsilon) \cdot s} \right)^{\text{rk } E(\mathbb{Q})} \leq 3^{\text{rk } E(\mathbb{Q})} \leq |E(\mathbb{Q})/3E(\mathbb{Q})| \leq |\text{Sel}_3(E)|.$$

The bounds on the average 3-Selmer size from [BH18] imply that

$$\sum_{E \in \mathcal{F}^{\leq T}} \left( 2^{(1+\varepsilon) \cdot s} \right)^{\text{rk } E(\mathbb{Q})} \ll |\mathcal{F}^{\leq T}|.$$

To bound the second term in (22), we imitate Lemma 3.6. We claim that for all  $t \geq 0$ ,

$$\sum_{E \in \mathcal{F}^{\leq T}} \prod_{p^2 | \Delta_E} \left( \left\lfloor \frac{v_p(\Delta_E)}{2} \right\rfloor + 1 \right)^t \ll_t |\mathcal{F}^{\leq T}|.$$

Almost all of the proof of Lemma 3.6 still works, including the large prime bound; we need only show that the number of  $E \in \mathcal{F}^{\leq T}$  with  $m | \Delta_E$  is

$$\ll |\mathcal{F}^{\leq T}| O(1)^{\omega(m)} m^{-1} \quad (23)$$

(as in (19)). For  $\mathcal{F} = \mathcal{F}_1$ , each fiber of the natural reduction map  $\mathcal{F}^{\leq T} \rightarrow (\mathbb{Z}/m\mathbb{Z})^3$  sending  $E \in \mathcal{F}_1$  to  $(d_2, d_3, d_4)$  modulo  $m$  is now of size  $\ll |\mathcal{F}^{\leq T}| \cdot m^{-3}$  (since there are now 3 parameters instead of 2 for each curve in  $\mathcal{F}_1$ ). But Hensel's lemma and the Chinese remainder theorem, in this case, show that there are  $\ll m^2 \cdot O(1)^{\omega(m)}$  solutions  $(d_2, d_3, d_4)$  modulo  $m$  to the discriminant vanishing modulo  $m$ . We thus still obtain the bound (23). The argument for  $\mathcal{F}_1(2)$  is almost identical.  $\square$

## REFERENCES

- [Abr97] Dan Abramovich, *Uniformity of stably integral points on elliptic curves*, Invent. Math. **127** (1997), no. 2, 307–317. MR 1427620
- [Akh12] Shabnam Akhtari, *Upper bounds for the number of solutions to quartic Thue equations*, Int. J. Number Theory **8** (2012), no. 2, 335–360. MR 2890483
- [Alp14] Levent Alpoge, *The average number of integral points on elliptic curves is bounded*, 2014, <https://arxiv.org/abs/1412.1047>.

- [BH16] Manjul Bhargava and Wei Ho, *Coregular spaces and genus one curves*, Cambridge J. Math. **4** (2016), no. 1, 1–119.
- [BH18] ———, *On the average sizes of Selmer groups in families of elliptic curves*, 2018, preprint, 46 pages.
- [BS87] E. Bombieri and W. M. Schmidt, *On Thue’s equation*, Invent. Math. **88** (1987), no. 1, 69–81.
- [BS13] Manjul Bhargava and Arul Shankar, *The average size of the 5-Selmer group of elliptic curves is 6, and the average rank is less than 1*, 2013, <http://arxiv.org/abs/1312.7859>.
- [BS15] ———, *Binary quartic forms having bounded invariants, and the boundedness of the average rank of elliptic curves*, Ann. of Math. (2) **181** (2015), no. 1, 191–242.
- [BSD63] B. J. Birch and H. P. F. Swinnerton-Dyer, *Notes on elliptic curves. I*, J. Reine Angew. Math. **212** (1963), 7–25.
- [CFS10] John E. Cremona, Tom A. Fisher, and Michael Stoll, *Minimisation and reduction of 2-, 3- and 4-coverings of elliptic curves*, Algebra Number Theory **4** (2010), no. 6, 763–820.
- [Eve97] Jan-Hendrik Evertse, *The number of solutions of the Thue-Mahler equation*, J. Reine Angew. Math. **482** (1997), 121–149. MR 1427659
- [HS88] M. Hindry and J. H. Silverman, *The canonical height and integral points on elliptic curves*, Invent. Math. **93** (1988), no. 2, 419–450.
- [HV06] H. A. Helfgott and A. Venkatesh, *Integral points on elliptic curves and 3-torsion in class groups*, J. Amer. Math. Soc. **19** (2006), no. 3, 527–550.
- [Kim15] Dohyeong Kim, *Descent for the punctured universal elliptic curve, and the average number of integral points on elliptic curves*, 2015, <https://arxiv.org/abs/1502.04923>.
- [Maz77] B. Mazur, *Modular curves and the Eisenstein ideal*, Inst. Hautes Études Sci. Publ. Math. (1977), no. 47, 33–186.
- [Mor69] L. J. Mordell, *Diophantine equations*, Pure and Applied Mathematics, Vol. 30, Academic Press, London-New York, 1969. MR 0249355
- [Sil87] Joseph H. Silverman, *A quantitative version of Siegel’s theorem: integral points on elliptic curves and Catalan curves*, J. Reine Angew. Math. **378** (1987), 60–100. MR 895285