Advances in Mathematics 457 (2024) 109946

Contents lists available at ScienceDirect

Advances in Mathematics

journal homepage: www.elsevier.com/locate/aim

The average number of integral points on the congruent number curves

Stephanie Chan [1]

Department of Mathematics, University of Michigan, 530 Church Street, Ann Arbor, MI 48109, USA

ARTICLE INFO

Article history:
Received 22 March 2023
Received in revised form 6 December 2023
Accepted 4 September 2024
Available online 11 September 2024
Communicated by George Andrews

Keywords:
Elliptic curve
Quadratic twist
Integral point

ABSTRACT

We show that the total number of non-torsion integral points on the elliptic curves $\mathcal{E}_D : y^2 = x^3 - D^2x$, where D ranges over positive squarefree integers less than N, is $O(N(\log N)^{-\frac{1}{4}+\epsilon})$. The proof involves a discriminant-lowering procedure on integral binary quartic forms and an application of Heath-Brown's method on estimating the average size of the 2-Selmer groups of the curves in this family.

1. Introduction

Given an elliptic curve over $\mathbb{Q}$ with short Weierstrass model

$$E : y^2 = x^3 + Ax + B, \; A, B \in \mathbb{Z}, \tag{1}$$

we study the quadratic twists of E, with the model

E-mail address: stephanie.chan@ist.ac.at.

[1] *Current address:* Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria.

https://doi.org/10.1016/j.aim.2024.109946

$$E_D : y^2 = x^3 + AD^2x + BD^3, \tag{2}$$

where D denotes a positive squarefree integer. Consider the set of integral points

$$E_D(\mathbb{Z}) \coloneqq \left\{(x, y) \in \mathbb{Z}^2 : y^2 = x^3 + AD^2x + BD^3\right\}.$$

It follows from a result of Mordell [17] that $\#E_D(\mathbb{Z})$ is always finite.

We are interested in the distribution of the number of integral points $\#E_D(\mathbb{Z})$ in a quadratic twist family, when E_D are ordered according to the size of D. If $E(\mathbb{Q})$ contains a 2-torsion point, this point must have the form $(a, 0)$ for some integer a under the model (1), and hence $(aD, 0) \in E_D(\mathbb{Z})$ for all squarefree integers D. Therefore we call an integral point *non-trivial* if it is not a 2-torsion point of $E_D(\mathbb{Q})$. Define the set of non-trivial integral points on E_D to be

$$E_D^*(\mathbb{Z}) \coloneqq \{(x, y) \in E_D(\mathbb{Z}) : y \neq 0\}.$$

Define

$$\mathcal{D} \coloneqq \{D \in \mathbb{Z} : D > 0 \text{ squarefree}\},$$
$$\mathcal{D}_N \coloneqq \{D \in \mathcal{D} : D \leq N\}.$$

Granville [9] conjectured that almost all curves within a quadratic twist family have no non-trivial integral point. We state the conjecture adapted to our model (2).

Conjecture 1.1 *(Granville [9]). Fix $A, B \in \mathbb{Z}$ such that $4A^3 + 27B^2 \neq 0$. Let $E_D : y^2 = x^3 + AD^2x + BD^3$, $D \in \mathcal{D}$. Then*

$$\#\{D \in \mathcal{D}_N : E_D^*(\mathbb{Z}) \neq \varnothing\} \sim C_{A,B} N^{\frac{1}{2}},$$

where $C_{A,B}$ is a constant that depends only on A, B.

We note that Granville's original conjecture considers a different model $Dy^2 = f(x)$, where $f \in \mathbb{Z}[x]$ and $\deg f = 3$. When $f(x) = x^3+Ax+B$, any point $(x, y) \in \mathbb{Z}^2$ satisfying $Dy^2 = f(x)$ corresponds to a point $(Dx, Dy) \in E_D(\mathbb{Z})$, so there are fewer integral points using the model $Dy^2 = f(x)$ when compared to our model (2). The exponent $\frac{1}{2}$ stated in Conjecture 1.1 replaces $\frac{1}{3}$ in the original conjecture because of this discrepancy. The exponent $\frac{1}{2}$ is suggested by some heuristics given in [5, p. 6677–6678] for the family $y^2 = x^3 - D^2x$.

In this direction, Matschke and Mudigonda [16] handled the case when $f(x)$ is reducible, assuming the *abc* conjecture.

Theorem 1.2 *(Matschke–Mudigonda [16]). Assume that the abc conjecture is true. Suppose $f(x) = x^3 + Ax + B$, $A, B \in \mathbb{Z}$, such that $4A^3 + 27B^2 \neq 0$ and $f(x)$ is reducible over $\mathbb{Q}$. Then*

$$\#\{D \in \mathcal{D}_N : Dy^2 = f(x) \text{ for some } x, y \in \mathbb{Z},\ y \neq 0\} \leq N^{\frac{2}{3}+o(1)}.$$

Our goal here is to gain progress towards Conjecture 1.1 on a specific quadratic twist family. We restrict our attention to the congruent number curve $\mathcal{E} : y^2 = x^3 - x$, and study its twists

$$\mathcal{E}_D : y^2 = x^3 - D^2 x.$$

It is well known that the torsion subgroup of $\mathcal{E}_D(\mathbb{Q})$ is $\{O,\ (0,0),\ (\pm D, 0)\} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ (see for example [13, Chapter I, Proposition 17]), where O denotes the point at infinity.

We now review some existing results concerning the family $\{\mathcal{E}_D : D \in \mathcal{D}\}$, and explain how it is implied that the moments of $\#\mathcal{E}_D(\mathbb{Z})$ are bounded. The 2-Selmer group of $\mathcal{E}_D$, which we denote by $\mathrm{Sel}_2(\mathcal{E}_D)$, is a finite abelian group with exponent 2 that is defined via local conditions and admits an injection $\mathcal{E}_D(\mathbb{Q})/2\mathcal{E}_D(\mathbb{Q}) \hookrightarrow \mathrm{Sel}_2(\mathcal{E}_D)$ (see for example [21, Chapter X]). In particular, the 2-Selmer rank provides an upper bound to the rank $\mathrm{rank}(\mathcal{E}_D(\mathbb{Q}))$ of the Mordell–Weil group of $\mathcal{E}_D$ over $\mathbb{Q}$. It is usually easier to compute the 2-Selmer groups of elliptic curves with a torsion subgroup $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ over $\mathbb{Q}$, since then most of the work can be done over $\mathbb{Q}$. Heath-Brown [11, Theorem 1] computed all the moments of the size of the 2-Selmer groups of $\mathcal{E}_D$. For any positive integer k, he showed that

$$\frac{1}{\#\mathcal{D}_N} \sum_{D \in \mathcal{D}_N} (\#\,\mathrm{Sel}_2(\mathcal{E}_D))^k = \prod_{j=1}^{k} (1 + 2^j) + o_k(1). \tag{3}$$

Since the 2-Selmer rank provides an upper bound to the rank of $\mathcal{E}_D$, the equation (3) implies that

$$\frac{1}{\#\mathcal{D}_N} \sum_{D \in \mathcal{D}_N} 2^{k \cdot \mathrm{rank}\, \mathcal{E}_D(\mathbb{Q})} \ll_k 1. \tag{4}$$

Lang [14, page 140] conjectured that the number of integral points on a quasi-minimal Weierstrass equation of an elliptic curve E should be bounded only in terms of $\mathrm{rank}\, E(\mathbb{Q})$. For the family $\{\mathcal{E}_D : D \in \mathcal{D}\}$, it follows from existing results in this direction by Silverman [20, Theorem A] and Hindry–Silverman [12, Theorem 0.7], that there exists some absolute constant C, such that

$$\#\mathcal{E}_D(\mathbb{Z}) \ll C^{\mathrm{rank}\, \mathcal{E}_D(\mathbb{Q})}. \tag{5}$$

In [5], we showed that C in (5) can be taken as small as 3.8. Combining the upper bounds in (5) and (4), we can bound the k-th moment of $\#\mathcal{E}_D(\mathbb{Z})$ by

$$\frac{1}{\#\mathcal{D}_N} \sum_{D \in \mathcal{D}_N} (\#\mathcal{E}_D(\mathbb{Z}))^k \ll_k 1. \tag{6}$$

We will show that in fact the moments of $\#\mathcal{E}_D^*(\mathbb{Z})$ tend to 0. The following is our main result.

Theorem 1.3. *For any $\epsilon > 0$ and any $k > 0$, we have*

$$\sum_{D\in\mathcal{D}_N} (\#\mathcal{E}_D^*(\mathbb{Z}))^k \ll_{\epsilon,k} N(\log N)^{-\frac{1}{4}+\epsilon}.$$

This shows that the k-th moment of $\#\mathcal{E}_D^*(\mathbb{Z})$ tends to 0 as N tends to infinity, since $\#\,\mathcal{D}_N \sim \frac{6}{\pi^2}N$.

To prove Theorem 1.3, it suffices to prove the following.

Theorem 1.4. *For any $\epsilon > 0$, we have*

$$\#\{D \in \mathcal{D}_N : \mathcal{E}_D^*(\mathbb{Z}) \neq \varnothing\} \ll_\epsilon N(\log N)^{-\frac{1}{4}+\epsilon}.$$

Indeed, by an application of Hölder's inequality, we have

$$\begin{aligned}\sum_{D\in\mathcal{D}_N} (\#\mathcal{E}_D^*(\mathbb{Z}))^k &\le \Big(\sum_{D\in\mathcal{D}_N} (\#\mathcal{E}_D^*(\mathbb{Z}))^{\frac{k}{\epsilon}}\Big)^{\epsilon} \left(\#\{D \in \mathcal{D}_N : \mathcal{E}_D^*(\mathbb{Z}) \neq \varnothing\}\right)^{1-\epsilon} \\ &\ll_{\epsilon,k} N(\log N)^{(-\frac{1}{4}+\epsilon)(1-\epsilon)},\end{aligned}$$

where we have inserted (6) and the estimate from Theorem 1.4. Rescaling ϵ gives Theorem 1.3.

We now give an outline of the proof of Theorem 1.4. In Section 2, for each integral point $(x, y) \in \mathcal{E}_D(\mathbb{Z})$, we use Mordell's correspondence [18, Chapter 25] to construct a corresponding integral binary quartic form f that represents 1 and has discriminant related to the discriminant of $\mathcal{E}_D$. Then in Section 3, we show that by picking an auxiliary prime $p \mid D/\gcd(x, D)$, we can transform f into an integral binary quartic form F that represents p and has discriminant lowered by a factor of p^6. In Section 4, we show that $\gcd(x, D)$ can be controlled by the image of (x, y) in the 2-Selmer group of $\mathcal{E}_D$ under the map

$$\mathcal{E}_D(\mathbb{Z}) \hookrightarrow \mathcal{E}_D(\mathbb{Q}) \twoheadrightarrow \mathcal{E}_D(\mathbb{Q})/2\mathcal{E}_D(\mathbb{Q}) \hookrightarrow \mathrm{Sel}_2(\mathcal{E}_D).$$

Then in Section 5 we extract some information about the distribution of 2-Selmer elements from work of Heath-Brown [10,11] to show that for almost all D, we are always able to pick the required auxiliary prime p of a suitable size. In particular, this p is not too small, so that there are $o(N)$ many discriminants for the discriminant-lowered quartic F to take. At the same time, by ensuring that p is not too large, we can deduce from upper bounds on the number of solutions to Thue inequalities, that each $\mathrm{SL}_2(\mathbb{Z})$-equivalence class of F can only be the image of finitely many integral points. In Section 6, we proceed to count the set of those quartics F that were discriminant-lowered by some suitable p.

We make use of the fact that every integral binary quartic form is $\mathrm{SL}_2(\mathbb{Z})$-equivalent to at least one reduced form with bounded seminvariants [6]. Applying the syzygy satisfied by the seminvariants returns a set of integral points on twists of $\mathcal{E}$ with bounded height. Then Theorem 1.4 follows from an application of an upper bound by Le Boudec [15].

2. Integer-matrix binary quartic forms

We say that a binary quartic form is *integer-matrix* if it has the form

$$f(X,Y) = a_0X^4 + 4a_1X^3Y + 6a_2X^2Y^2 + 4a_3XY^3 + a_4Y^4, \qquad a_i \in \mathbb{Z}.$$

Given any integral binary quartic form f and $(x_0, y_0) \in \mathbb{Z}^2$, define the action of

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z})$$

on the pair $(f, (x_0, y_0))$ by

$$\gamma \cdot (f(X,Y), (x_0, y_0)) = (f((X,Y) \cdot \gamma), (x_0, y_0) \cdot \gamma^{-1}),$$

where

$$(X,Y) \cdot \gamma = (aX + cY, bX + dY).$$

This action preserves the value of $f(x_0, y_0)$.

We recall some facts about the seminvariants of quartic forms [6, Section 4.1.1]. For our convenience, we choose to scale the seminvariants differently than in [6], since we will only be dealing with integer-matrix binary forms. The invariants of f are

$$\begin{aligned} I = I(f) &= a_0a_4 - 4a_1a_3 + 3a_2^2, \text{ and} \\ J = J(f) &= a_0a_2a_4 - a_0a_3^2 - a_1^2a_4 + 2a_1a_2a_3 - a_2^3. \end{aligned}$$

The discriminant of f is

$$\begin{aligned} \Delta(f) &:= I^3 - 27J^2 \\ &= a_0^3a_4^3 - 64a_1^3a_3^3 - 18a_0^2a_2^2a_4^2 - 12a_0^2a_1a_3a_4^2 - 6a_0a_1^2a_3^2a_4 \\ &\quad - 180a_0a_1a_2^2a_3a_4 + 81a_0a_2^4a_4 + 36a_1^2a_2^2a_3^2 - 27(a_0^2a_3^4 + a_1^4a_4^2) \\ &\quad + 54a_2(-a_2^2 + 2a_1a_3 + a_0a_4)(a_4a_1^2 + a_0a_3^2) \end{aligned}$$

The seminvariants attached to the form are I, J, $a = a(f) = a_0$,

$$H = H(f) = a_1^2 - a_0a_2, \text{ and } R = R(f) = 2a_1^3 + a_0^2a_3 - 3a_0a_1a_2.$$

Comparing to the formulas in [6, Section 4.1.1], here we have taken out a factor of -48 from their H, a factor of 32 from their R, a factor of 12 from their I, a factor 432 from their J, and a factor of $256 \cdot 27$ from their Δ. The seminvariants are related by the syzygy

$$H^3 - \frac{I}{4}a^2H - \frac{J}{4}a^3 = \left(\frac{R}{2}\right)^2. \tag{7}$$

Notice that when I and J are both divisible by 4, $(H, \frac{1}{2}R)$ defines an integral point on a twist of the elliptic curve $y^2 = x^3 - \frac{I}{4}x - \frac{J}{4}$.

2.1. Mordell's correspondence

For integers A, B such that $4A^3 + 27B^2 \neq 0$, define an elliptic curve over $\mathbb{Q}$ with the affine integral Weierstrass model

$$E_{A,B} : y^2 = x^3 + Ax + B.$$

The discriminant of $E_{A,B}$ is given by

$$\Delta_{E_{A,B}} = -16(4A^3 + 27B^2).$$

For integers $c, d, e \in \mathbb{Z}$, define an integer-matrix binary quartic form

$$f_{c,d,e}(X, Y) = X^4 + 6cX^2Y^2 + 8dXY^3 + eY^4.$$

Define

$$\mathcal{A} := \{(E_{A,B}, (x_0, y_0)) : A, B \in \mathbb{Z},\ 4A^3 + 27B^2 \neq 0,\ (x_0, y_0) \in E_{A,B}(\mathbb{Z})\},$$
$$\mathcal{B} := \{f_{c,d,e} : c, d, e \in \mathbb{Z},\ e \equiv c^2 \bmod 4,\ \Delta(f) \neq 0\}.$$

The following correspondence is given by Mordell [18, Chapter 25] (or see [3, Section 2.3] for a modern interpretation).

Theorem 2.1 *(Mordell). There is a bijection*

$$\mathcal{A} \to \mathcal{B}$$

given by

$$(E_{A,B}, (x_0, y_0)) \mapsto f,$$

where

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

Moreover, under this map, $\Delta(f) = \Delta_{E_{A,B}}$*,* $I(f) = -4A$ *and* $J(f) = -4B$*.*

The inverse map comes from the syzygy (7) satisfied by the seminvariants, but we will only make use of the map in the direction from $\mathcal{A}$ to $\mathcal{B}$ from Theorem 2.1.

3. Lowering the discriminant

We now fix an elliptic curve $E : y^2 = x^3 + Ax + B,\ A, B \in \mathbb{Z}$ and consider its quadratic twists $E_D : y^2 = x^3 + AD^2x + BD^3$, where $D \in \mathcal{D}$. For each $P = (c, d) \in E_D(\mathbb{Z})$, Theorem 2.1 gives the binary quartic form

$$f_P(X, Y) := X^4 - 6cX^2Y^2 + 8dXY^3 + (-4AD^2 - 3c^2)Y^4, \tag{8}$$

which satisfies $\Delta(f_P) = \Delta_E D^6$, $I(f_P) = -4AD^2$ and $J(f_P) = -4BD^3$.

Denote the space of integer-matrix binary quartic forms by V. Let $x(P)$ denote the x-coordinate of the point $P \in E_D(\mathbb{Z})$. Define a map

$$\Psi : \bigcup_{D\in\mathcal{D}} \left\{ (P, M) : \begin{matrix} P \in E_D(\mathbb{Z}),\ M \in \mathbb{Z},\ M > 0 \\ M \mid D,\ \gcd(2 \cdot x(P), M) = 1 \end{matrix} \right\} \to (V \times \mathbb{Z}^2)/\operatorname{SL}_2(\mathbb{Z}) \tag{9}$$

given by

$$(P, M) = ((c, d), M) \mapsto (F, (1, 0)),$$

where F is defined by taking k to be any integer such that $k \equiv dc^{-1} \bmod M$ and

$$F(X, Y) = \frac{1}{M^3} f_P(MX + kY, Y). \tag{10}$$

We will show that Ψ is well-defined and injective in Lemma 3.1 and Lemma 3.2.

In work of Bombieri and Schmidt [4], to bound the number of solutions to a Thue equation $F_1(X, Y) = h$, they transformed the equation to $F_2(X, Y) = 1$, where the discriminant of F_2 is raised by a factor of h^6 compared to that of F_1. Some applications of this idea can be found in [1,2]. Here we attempt to carry out the reverse process on the integral quartic forms f_P to lower their discriminants.

Lemma 3.1. *Take* $D \in \mathcal{D}$*. Let* $P = (c, d) \in E_D(\mathbb{Z})$ *and take* f_P *as defined in* (8)*. Fix a positive squarefree integer* M *dividing* D *that is coprime to* $2c$*. Then for any integer* k *such that* $k \equiv dc^{-1} \bmod M$*, we have that*

$$F(X, Y) := \frac{1}{M^3} \cdot f_P\left((X, Y) \cdot \begin{pmatrix} M & 0 \\ k & 1 \end{pmatrix}\right) = \frac{1}{M^3} \cdot f_P(MX + kY, Y)$$

is an integer-matrix binary quartic form. Moreover, we have

(i) $F(1,0) = M$,
(ii) $I(F) = -4A(D/M)^2$, $J(F) = -4B(D/M)^3$, *and*
(iii) $\Delta(F) = \Delta(f_P)/M^6 = -16(4A^3 + 27B^2)(D/M)^6$.

Proof. Since $(c,d) \in E_D(\mathbb{Z})$, we have $d^2 = c^3 + AD^2c + BD^3$. Taking any integer k such that $k \equiv dc^{-1} \bmod M$, we have $k^2 \equiv d^2c^{-2} \equiv c \bmod M$. Then by Hensel's lemma we can find a lift K of k such that $k \equiv K \bmod M$ and

$$c \equiv K^2 \bmod M^3. \tag{11}$$

It suffices to show that F is an integer-matrix binary quartic form with this choice of $k = K$, since otherwise $k = K + uM$ for some integer u, and we can consider $F(X - uY, Y)$ instead.

Next we put (11) into $d^2 = c^3 + AD^2c + BD^3$ and solve for $d \bmod M^3$. Since $d \equiv kc \equiv k^3 \equiv K^3 \bmod M$, we see from the two square roots of $(K^2)^3 + AD^2(K^2) + BD^3 \bmod M^3$, that

$$d \equiv K^3 + \frac{AD^2}{2K} \bmod M^3. \tag{12}$$

By (11) and (12), we see that the coefficients of

$$\begin{aligned} f_P(MX + KY, Y) = {} & M^4X^4 + 4M^3KX^3y + 6M^2(K^2 - c)X^2Y^2 \\ & + 4M(K^3 - 3cK + 2d)XY^3 + (K^4 - 6cK^2 + 8dK - 4AD^2 - 3c^2)Y^4 \end{aligned}$$

are all divisible by M^3. Therefore F is an integer-matrix binary quartic form. The remaining properties are then straightforward from the definition of F. □

Lemma 3.2. *The map Ψ is well-defined and injective.*

Proof. To show that Ψ is well-defined, by Lemma 3.1, it remains to show that the class $(F,(1,0))/\operatorname{SL}_2(\mathbb{Z})$ does not depend on the choice of k. Since k is determined up to modulo M by (c,d), if there are two choices of k, say k_1 and k_2, that gives two forms F_1 and F_2 via (10), they must satisfy $k_1 = k_2 + uM$ for some integer u. Then $F_2(X + uY, Y) = F_1(X,Y)$, and so $\begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix} \cdot (F_2,(1,0)) = (F_2,(1,0))$.

Next we check that Ψ is injective. The value of $F(1,0)$ determines M, and together with the discriminant of F, determines D. In the following, fix some $D \in \mathcal{D}$ and some $M \mid D$ such that $\gcd(2,M) = 1$. Suppose that $P, Q \in E_D(\mathbb{Z})$ satisfy $\gcd(x(P),M) = \gcd(x(Q),M) = 1$ and write $\Psi(Q,M) = (F_P,(1,0))$ and $\Psi(P,M) = (F_Q,(1,0))$. Suppose that $(F_P,(1,0))$ and $(F_Q,(1,0))$ are $\operatorname{SL}_2(\mathbb{Z})$-equivalent, so $\gamma \cdot (F_P,(1,0)) = (F_Q,(1,0))$ for some $\gamma \in \operatorname{SL}_2(\mathbb{Z})$. Then $(1,0) \cdot \gamma^{-1} = (1,0)$ implies that we can write $\gamma = \begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix}$ for some $u \in \mathbb{Z}$. Recall that

$$F_P(X,Y) = \frac{1}{M^3} \cdot f_P\left((X,Y) \cdot \begin{pmatrix} M & 0 \\ k_P & 1 \end{pmatrix}\right)$$

and

$$F_Q(X,Y) = \frac{1}{M^3} \cdot f_Q\left((X,Y) \cdot \begin{pmatrix} M & 0 \\ k_Q & 1 \end{pmatrix}\right)$$

for some integers k_P and k_Q which are determined up to modulo M. From $F_P((X,Y) \cdot \gamma) = F_Q(X,Y)$, we get

$$f_P\left((X,Y) \cdot \gamma \cdot \begin{pmatrix} M & 0 \\ k_P & 1 \end{pmatrix}\right) = f_Q\left((X,Y) \cdot \begin{pmatrix} M & 0 \\ k_Q & 1 \end{pmatrix}\right).$$

Then since

$$\begin{pmatrix} M & 0 \\ k_Q & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix} \begin{pmatrix} M & 0 \\ k_P & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ uM + k_P - k_Q & 1 \end{pmatrix},$$

we have

$$f_P\left((X,Y) \cdot \begin{pmatrix} 1 & 0 \\ uM + k_P - k_Q & 1 \end{pmatrix}\right) = f_Q(X,Y).$$

The X^3Y-coefficients of f_P and f_Q are both 0, so it must be that $uM + k_P - k_Q = 0$ and $f_P = f_Q$. Hence $P = Q$. $\square$

4. The 2-Selmer group of $y^2 = x^3 - D^2x$

In the following sections we will specialise in the case when $A = -1$ and $B = 0$, that is, the quadratic twist family containing the congruent number curves

$$\mathcal{E}_D : y^2 = x^3 - D^2x,$$

where $D \in \mathcal{D}$.

Heath-Brown [10,11] computed the moments of the size of the 2-Selmer groups of the congruent number curve family $\{\mathcal{E}_D : D \in \mathcal{D}\}$. We will extract some information about the 2-Selmer elements in this family from the argument in [10,11], in order to show that we can usually pick a suitable M to apply Lemma 3.1.

The 2-Selmer group of $\mathcal{E}_D$ is defined to be

$$\mathrm{Sel}_2(\mathcal{E}_D) := \ker\left(H^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), \mathcal{E}_D[2]) \to \prod_{p \text{ place of } \mathbb{Q}} H^1(\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p), \mathcal{E}_D)\right).$$

Since $\mathcal{E}_D$ has full 2-torsion, there is an isomorphism $H^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), \mathcal{E}_D[2]) \cong ((\mathbb{Q}^\times/(\mathbb{Q}^\times)^2)^2$, and it is possible to obtain explicit equations for the homogeneous spaces.

(See for example [21, Chapter X, Proposition 1.4].) For the curves $\mathcal{E}_D$, these equations were given as part of Heath-Brown's argument [10, Section 2]. As we will see, each 2-Selmer element of $\mathcal{E}_D(\mathbb{Q})$ corresponds to a system of two binary quadratic forms that is everywhere locally solvable. We will follow [10, Section 2] to recover the equations.

We begin by defining the set of tuples which we will use as representatives of 2-Selmer elements.

Definition 4.1. *For $D \in \mathcal{D}$, define $\mathcal{W}_D$ to be the set of all 4-tuples of positive squarefree integers (D_1, D_2, D_3, D_4) such that*

(1) *the system*

$$D_1X^2 + D_4W^2 = D_2Y^2,\ D_1X^2 - D_4W^2 = D_3Z^2, \tag{13}$$

is everywhere locally solvable, and

(2) $D_1D_2D_3D_4 = D$.

Consider the injective homomorphism

$$\theta : \mathcal{E}_D(\mathbb{Q})/2\mathcal{E}_D(\mathbb{Q}) \to \mathbb{Q}^\times/(\mathbb{Q}^\times)^2 \times \mathbb{Q}^\times/(\mathbb{Q}^\times)^2 \times \mathbb{Q}^\times/(\mathbb{Q}^\times)^2$$

given by

$$(x, y) \mapsto (x - D, x, x + D) \tag{14}$$

at non-torsion points. At torsion points, we have $\theta(O) = (1, 1, 1)$, $\theta((0, 0)) = (-D, -1, D)$, $\theta((D, 0)) = (2, D, 2D)$, $\theta((-D, 0)) = (-2D, -D, 1)$.

In the next lemma we establish the correspondence between $\mathcal{W}_D$ and $\mathrm{Sel}_2(\mathcal{E}_D)$.

Lemma 4.2. *The set $\mathcal{W}_D$ is in bijection with*

$$\begin{cases} \mathrm{Sel}_2(\mathcal{E}_D)/\theta(\{O, (0, 0), (\pm D, 0)\}) & \text{if } D \text{ is odd}, \\ \mathrm{Sel}_2(\mathcal{E}_D)/\theta(\{O, (0, 0)\}) & \text{if } D \text{ is even}, \end{cases}$$

where θ denotes the natural map

$$\theta : \mathcal{E}_D(\mathbb{Q}) \twoheadrightarrow \mathcal{E}_D(\mathbb{Q})/2\mathcal{E}_D(\mathbb{Q}) \hookrightarrow \mathrm{Sel}_2(\mathcal{E}_D).$$

More explicitly, identifying $\mathrm{Sel}_2(\mathcal{E}_D)$ as a subgroup of $(\mathbb{Q}^\times/(\mathbb{Q}^\times)^2)^3$ via (14), $(D_1, D_2, D_3, D_4) \in \mathcal{W}_D$ maps to

$$\begin{cases} \left\{ \begin{matrix} (D_1D_2, D_2D_3, D_3D_1),\ (-D_4D_3, -D_3D_2, D_2D_4), \\ (2D_2D_1, D_1D_4, 2D_4D_2),\ (-2D_3D_4, -D_4D_1, D_1D_3) \end{matrix} \right\} & \text{if } D \text{ is odd}, \\ \{(D_1D_2, D_2D_3, D_3D_1),\ (-D_4D_3, -D_3D_2, D_2D_4)\} & \text{if } D \text{ is even}. \end{cases}$$

Moreover, if the image of $(c, d) \in \mathcal{E}_D(\mathbb{Z})$ *under the map*

$$\mathcal{E}_D(\mathbb{Z}) \hookrightarrow \mathcal{E}_D(\mathbb{Q}) \twoheadrightarrow \mathcal{E}_D(\mathbb{Q})/2\mathcal{E}_D(\mathbb{Q}) \hookrightarrow \mathrm{Sel}_2(\mathcal{E}_D) \twoheadrightarrow \mathcal{W}_D \tag{15}$$

is (D_1, D_2, D_3, D_4)*, then*

$$\gcd(c, D) \in \{D_1D_2D_3,\ D_2D_3D_4,\ D_1D_2D_4,\ D_1D_3D_4\}. \tag{16}$$

Proof. Following [10, Section 2], we first show that there is a bijection between $\theta(\mathcal{E}_D(\mathbb{Q})) \cong \mathcal{E}_D(\mathbb{Q})/2\mathcal{E}_D(\mathbb{Q})$ and the set of tuples of squarefree integers (B_1, B_2, B_3, B_4) such that

$$B_1B_2B_3B_4 = \begin{cases} D \text{ or } 4D & \text{if } D \text{ is odd,} \\ D & \text{if } D \text{ is even,} \end{cases} \quad \gcd(B_1, B_2, B_3) = 1, \quad B_1B_2B_3 > 0, \tag{17}$$

and the system

$$B_1X^2 + B_4W^2 = B_2Y^2,\ B_1X^2 - B_4W^2 = B_3Z^2 \tag{18}$$

is solvable over $\mathbb{Q}$. Then in the same way, by working over $\mathbb{Q}_p$ over all places p of $\mathbb{Q}$ instead, $\mathrm{Sel}_2(\mathcal{E}_D)$ corresponds to the set of tuples (B_1, B_2, B_3, B_4) satisfying (17) and such that (18) is everywhere locally solvable. Note that (B_1, B_2, B_3, B_4) is not necessarily in $\mathcal{W}_D$ yet, but this will be adjusted in a later step.

We begin by constructing (B_1, B_2, B_3, B_4) from an arbitrary element of $(x, y) \in \mathcal{E}_D(\mathbb{Q})$. Suppose $(x, y) \in \mathcal{E}_D(\mathbb{Q})$, and write $x = r/s$ and $y = t/u$, where r, s, t, u are integers, $s, u > 0$, and $\gcd(r, s) = \gcd(t, u) =1$. Putting this into $y^2 = x^3 - D^2x$, we have

$$r(r + sD)(r - sD)u^2 = t^2s^3.$$

Then since $\gcd(t, u) = \gcd(r, s) = 1$, we must have $s^3 = u^2$, so $s = W^2$ for some integer W. Now write $\gcd(r, D) = B_0$, and $r = B_0r'$. From

$$r(r + sD)(r - sD) = t^2,$$

we see that $B_0^3 \mid t^2$, hence $B_0^2 \mid t$ since B_0 is squarefree. Then writing $B_4 = D/B_0$, we have $\gcd(r', sB_4) = 1$ by construction, and the equation becomes

$$r'(r' + sB_4)(r' - sB_4) = B_0(t/B_0^2)^2.$$

The factors on the left are pairwise coprime except possibly a common factor of 2 between $r' + sB_4$ and $r' - sB_4$, which only occurs when r' and sB_4 are both odd; in this case $r', (r' + sB_4)/2, (r' - sB_4)/2$ are pairwise coprime. Now we can write

$$r' = B_1X^2, \quad r' + sB_4 = B_2Y^2, \quad r' - sB_4 = B_3Z^2, \tag{19}$$

where B_1, B_2, B_3 are squarefree integers such that

$$B_1B_2B_3 = \begin{cases} B_0 & \text{if } B_1, B_2, B_3 \text{ are pairwise coprime,} \\ 4B_0 & \text{if } \gcd(B_2, B_3) = 2 \text{ and } B_1, B_4 \text{ are odd.} \end{cases}$$

In the first case $B_1B_2B_3B_4 = D$ and in the second case $B_1B_2B_3B_4 = 4D$ with B_1, B_4 odd and B_2, B_3 even. When D is even, the case $B_1B_2B_3B_4 = 4D$ is not possible since $8 \mid B_1B_2B_3B_4$ is not compatible with the parity conditions on the squarefree B_i. Putting $s = W^2$ into (19) and rearranging, we obtain a solution to the system (18). Identifying $\theta(\mathcal{E}_D(\mathbb{Q}))$ with a subgroup of $(\mathbb{Q}^\times/(\mathbb{Q}^\times)^2)^3$ as in (14), it is clear from the construction that $\theta((x, y)) = (B_1B_2, B_2B_3, B_3B_1)$.

Conversely, given (B_1, B_2, B_3, B_4) and a solution to (18), take B_0 to be the squarefree part of $B_1B_2B_3$, then $(x, y) = (B_0B_1X^2/W^2, B_0^2XYZ/W^3) \in \mathcal{E}_D(\mathbb{Q})$ and $\theta((x, y)) = (B_1B_2, B_2B_3, B_3B_1)$. For any element $w \in (\mathbb{Q}^\times/(\mathbb{Q}^\times)^2)^3$, we can check that there is at most one (B_1, B_2, B_3, B_4) satisfying (17) such that $w \equiv (B_1B_2, B_2B_3, B_3B_1)$. This shows that $\theta(\mathcal{E}_D(\mathbb{Q}))$ is in bijection with the set of (B_1, B_2, B_3, B_4) as claimed.

The tuple (B_1, B_2, B_3, B_4) constructed is not always in $\mathcal{W}_D$ because of the signs of B_1, B_2, B_3, B_4 and their valuations at 2. To obtain the bijection as required in the lemma, we add a suitable torsion point of $\mathcal{E}_D$ to (x, y). If D is odd, exactly one of $(x', y') \in \{(x, y),\ (x, y) + (0, 0),\ (x, y) + (D, 0),\ (x, y) + (-D, 0)\}$ satisfies $x' > 0$ and $v_2(x') \neq 0$. Take (D_1, D_2, D_3, D_4) to be the tuple corresponding to (x', y') and take this as the image of (x, y) in $\mathcal{W}_D$. By studying $\theta((x', y'))$, we see that the image of (x, y) in $\mathcal{W}_D$ relates to (B_1, B_2, B_3, B_4) as follows

$$(D_1, D_2, D_3, D_4) = \begin{cases} (B_1, B_2, B_3, B_4) & \text{if } v_2(x) \neq 0 \text{ and } x > 0, \\ (B_4, -B_3, B_2, -B_1) & \text{if } v_2(x) \neq 0 \text{ and } x < 0, \\ (B_2/2, B_1, B_4, B_3/2) & \text{if } v_2(x) = 0 \text{ and } x > 0, \\ (-B_3/2, B_4, -B_1, B_2/2) & \text{if } v_2(x) = 0 \text{ and } x < 0. \end{cases}$$

If D is even, then exactly one of $(x', y') \in \{(x, y), (x, y) + (0, 0)\}$ satisfies $x' > 0$ and $v_2(x') \equiv 1 \bmod 2$. We take the image of (x, y) in $\mathcal{W}_D$ to be

$$(D_1, D_2, D_3, D_4) = \begin{cases} (B_1, B_2, B_3, B_4) & \text{if and } x > 0, \\ (B_4, -B_3, B_2, -B_1) & \text{if and } x < 0. \end{cases}$$

For the final claim in the lemma, notice that if $(x, y) \in \mathcal{E}_D(\mathbb{Z})$, by construction $\gcd(x, D) = B_0$, which is the squarefree part of $B_1B_2B_3$. This gives (16) by rewriting $B_1B_2B_3$ in terms of D_1, D_2, D_3, D_4 in each of the above cases. $\square$

5. Generic 2-Selmer elements

In Lemma 4.2 we constructed a map from $\mathrm{Sel}_2(\mathcal{E}_D)$ to $\mathcal{W}_D$. We want to show that for almost all $P \in \mathcal{E}_D^*(\mathbb{Z})$, there exists a prime p of suitable size such that $p \mid D$ but $p \nmid x(P)$ in order to apply Lemma 3.1 with $M = p$. The observation in (16) suggests that it will be useful to show that D_1, D_2, D_3, D_4 all have prime factors in an expected range. To achieve this, we will follow Section 2 to Section 4 in work of Heath-Brown [10] closely with suitable modifications. (See also [11].)

Henceforth $0 < \epsilon < \frac{1}{4}$ will be a fixed constant. Let S be the interval

$$S := \left[\exp((\log N)^{2\epsilon}),\ \exp((\log N)^{1-2\epsilon})\right],$$

so that any $p \in S$ satisfies

$$2\epsilon \log\log N \leq \log\log p \leq (1-2\epsilon)\log\log N.$$

Define

$$\omega(n) := \#\{p \text{ prime} : p \mid n\},$$
$$\omega_S(n) := \#\{p \text{ prime} : p \mid n,\ p \in S\}.$$

Further define a parameter

$$N^{\ddagger} := \exp\left((\log N)^{\frac{1}{4}\epsilon}\right).$$

The goal of this section is to prove the following result.

Theorem 5.1. *Define two properties on $(D_1, D_2, D_3, D_4) \in \mathcal{W}_D$:*

(**S1**) *(D_1, D_2, D_3, D_4) comes from a torsion point on $\mathcal{E}_D(\mathbb{Q})$;*
(**S2**) *for each $i \in \{1,2,3,4\}$, we have $D_i > N^{\ddagger}$ and there exist some $p \mid D_i$ such that $p \in S$.*

Then

$$\#\left\{D \in \mathcal{D}_N : \begin{array}{l}(\textit{S1}) \text{ and } (\textit{S2}) \text{ both fail for} \\ \text{some } (D_1, D_2, D_3, D_4) \in \mathcal{W}_D\end{array}\right\} \ll_\epsilon N(\log N)^{-\frac{1}{4}+\epsilon}.$$

From Lemma 4.2, we can check that all torsion points map to $(1,1,1,D) \in \mathcal{W}_D$ if D is odd. If D is even, $\{O, (0,0)\}$ maps to $(1,1,1,D)$ and $\{(\pm D, 0)\}$ maps to $\left(2, 1, \frac{1}{2}D, 1\right)$ in $\mathcal{W}_D$. Therefore the condition (**S1**) is equivalent to

$$(D_1, D_2, D_3, D_4) = \begin{cases} (1,1,1,D) & \text{if } D \text{ is odd,} \\ (1,1,1,D) \text{ or } \left(2,1,\frac{1}{2}D,1\right) & \text{if } D \text{ is even.} \end{cases}$$

When D is even, exactly one of D_1, D_2, D_3, D_4 is even. For the subsequent character sum argument, it will be easier to first isolate the prime factor 2 by replacing the even D_i with $2d_i$ and consider instead tuples of odd integers. Define

$$\delta_i(\eta) := \begin{cases} 1 & \text{if } \eta = i, \\ 0 & \text{otherwise,} \end{cases}$$

so that if D_η is even, we have $D_i = 2^{\delta_i(\eta)} d_i$ for $i \in \{1,2,3,4\}$, and we take $\eta = 0$ when D is odd so that trivially $D_i = d_i = 2^{\delta_i(0)} d_i$ for $i \in \{1,2,3,4\}$. To prove Theorem 5.1, it suffices to bound the number of 4-tuples of positive odd integers (d_1, d_2, d_3, d_4) satisfying the following conditions for some $\eta \in \{0,1,2,3,4\}$:

(1) $\left(2^{\delta_1(\eta)} d_1, 2^{\delta_2(\eta)} d_2, 2^{\delta_3(\eta)} d_3, 2^{\delta_4(\eta)} d_4\right) \in \mathcal{W}_D$ for some $D \in \mathcal{D}_N$, and
(2) one of the conditions (**W1**) and (**W2**) listed below.

(**W1**) For some $i \in \{1,2,3,4\}$, we have $d_i \le N^\ddagger$, and

$$(d_1, d_2, d_3, d_4) \ne \begin{cases} (1,1,1,D) & \text{if } \eta = 0 \text{ or } 4, \\ (1,1,\frac{1}{2}D,1) & \text{if } \eta = 2. \end{cases} \tag{20}$$

(**W2**) We have $d_i > N^\ddagger$ for all $i \in \{1,2,3,4\}$, and there exists an i such that d_i has no prime factor in S.

In the above notation, $\eta = 0$ implies that $D = d_1 d_2 d_3 d_4$ is odd, and $\eta \in \{1,2,3,4\}$ implies that $D = 2d_1 d_2 d_3 d_4$ is even.

5.1. The indicator function

For $(D_1, D_2, D_3, D_4) = \left(2^{\delta_1(\eta)} d_1, 2^{\delta_2(\eta)} d_2, 2^{\delta_3(\eta)} d_3, 2^{\delta_4(\eta)} d_4\right)$ to lie in $\mathcal{W}_D$, the system (13) has to be everywhere locally solvable. Following the proof of [10, Lemma 3], we will package the local conditions as a sum of product of Jacobi symbols. The function we will obtain to detect the local solvability conditions is essentially the same as in [10, Lemma 3] for odd D. For simplicity, we shall only keep the conditions at odd primes dividing D. (Though we remark here that there are automatically real solutions because $D_i > 0$ and the conditions at 2 do not really contribute further, see [10, Lemma 2].) The condition that $(D_1, D_2, D_3, D_4) \in \mathcal{W}_D$ implies that

$$\begin{cases} \left(\frac{D_2D_4}{p}\right) = \left(\frac{-D_3D_4}{p}\right) = 1 & \text{if } p \mid d_1, \\ \left(\frac{-D_1D_4}{p}\right) = \left(\frac{2D_1D_3}{p}\right) = 1 & \text{if } p \mid d_2, \\ \left(\frac{2D_1D_2}{p}\right) = \left(\frac{D_1D_4}{p}\right) = 1 & \text{if } p \mid d_3, \\ \left(\frac{D_1D_2}{p}\right) = \left(\frac{D_1D_3}{p}\right) = 1 & \text{if } p \mid d_4. \end{cases} \tag{21}$$

Set

$$\Pi_1 := \prod_{p\mid d_1} \left(1 + \left(\frac{D_2D_4}{p}\right)\right)\left(1 + \left(\frac{-D_3D_4}{p}\right)\right),$$

$$\Pi_2 := \prod_{p\mid d_2} \left(1 + \left(\frac{-D_1D_4}{p}\right)\right)\left(1 + \left(\frac{2D_1D_3}{p}\right)\right),$$

$$\Pi_3 := \prod_{p\mid d_3} \left(1 + \left(\frac{2D_1D_2}{p}\right)\right)\left(1 + \left(\frac{D_1D_4}{p}\right)\right),$$

$$\Pi_4 := \prod_{p\mid d_4} \left(1 + \left(\frac{D_1D_2}{p}\right)\right)\left(1 + \left(\frac{D_1D_3}{p}\right)\right),$$

then

$$G_\eta(d_1, d_2, d_3, d_4) := 4^{-\omega(d_1d_2d_3d_4)}\Pi_1\Pi_2\Pi_3\Pi_4$$

takes the value 1 when η and (d_1, d_2, d_3, d_4) satisfy (21) and 0 otherwise. Since $(D_1, D_2, D_3, D_4) \in \mathcal{W}_D$ implies (21), we have

$$G_\eta(d_1, d_2, d_3, d_4) \geq \begin{cases} 1 & \text{if } \left(2^{\delta_1(\eta)}d_1, 2^{\delta_2(\eta)}d_2, 2^{\delta_3(\eta)}d_3, 2^{\delta_4(\eta)}d_4\right) \in \mathcal{W}_D, \\ 0 & \text{else.} \end{cases} \tag{22}$$

The next step is to expand $\Pi_1, \Pi_2, \Pi_3, \Pi_4$. Substituting $D_i = 2^{\delta_i(\eta)}d_i$, we get

$$\Pi_1 = \sum \left(\frac{2^{\delta_2(\eta)+\delta_4(\eta)}d_2d_4}{d_{13}}\right)\left(\frac{-2^{\delta_3(\eta)+\delta_4(\eta)}d_3d_4}{d_{12}}\right)\left(\frac{-2^{\delta_2(\eta)+\delta_3(\eta)}d_2d_3}{d_{14}}\right),$$

where the sum is over all factorisations $d_1 = d_{10}d_{12}d_{13}d_{14}$;

$$\Pi_2 = \sum \left(\frac{-2^{\delta_1(\eta)+\delta_4(\eta)}d_1d_4}{d_{23}}\right)\left(\frac{2^{1+\delta_1(\eta)+\delta_3(\eta)}d_1d_3}{d_{24}}\right)\left(\frac{-2^{1+\delta_3(\eta)+\delta_4(\eta)}d_3d_4}{d_{21}}\right),$$

where the sum is over all factorisations $d_2 = d_{20}d_{21}d_{23}d_{24}$;

$$\Pi_3 = \sum \left(\frac{2^{1+\delta_1(\eta)+\delta_2(\eta)}d_1d_2}{d_{34}}\right)\left(\frac{2^{\delta_1(\eta)+\delta_4(\eta)}d_1d_4}{d_{32}}\right)\left(\frac{2^{1+\delta_2(\eta)+\delta_4(\eta)}d_2d_4}{d_{31}}\right),$$

where the sum is over all factorisations $d_3 = d_{30}d_{31}d_{32}d_{34}$;

$$\Pi_4 = \sum \left(\frac{2^{\delta_1(\eta)+\delta_2(\eta)}d_1d_2}{d_{43}}\right)\left(\frac{2^{\delta_1(\eta)+\delta_3(\eta)}d_1d_3}{d_{42}}\right)\left(\frac{2^{\delta_2(\eta)+\delta_3(\eta)}d_2d_3}{d_{41}}\right),$$

where the sum is over all factorisations $d_4 = d_{40}d_{41}d_{42}d_{43}$.

Write $\mathbf{d} = (d_{ij})$ as the 16-tuple of positive odd integers that arise from the expansions above, where the indices (i, j) are in the range

$$1 \leq i \leq 4,\ 0 \leq j \leq 4,\ i \neq j.$$

For odd D, set

$$g_0(\mathbf{d}) := \left(\frac{-1}{\alpha}\right)\left(\frac{2}{\beta_0}\right)\prod_i 4^{-\omega(d_{i0})}\prod_{j\neq 0} 4^{-\omega(d_{ij})}\prod_{k\neq i,j}\prod_l \left(\frac{d_{kl}}{d_{ij}}\right),$$

where $\alpha = d_{12}d_{14}d_{23}d_{21}$ and $\beta_0 = d_{24}d_{21}d_{34}d_{31}$. Then

$$G_0(d_1, d_2, d_3, d_4) = \sum_{\substack{\mathbf{d} \\ \prod_{j\neq i} d_{ij}=d_i}} g_0(\mathbf{d}).$$

For even D, from the expansions of $\Pi_1, \Pi_2, \Pi_3, \Pi_4$, we see that the only difference from the odd case is with the terms $\left(\frac{2}{d_{ij}}\right)$ that appear in the sum. Set

$$g_\eta(\mathbf{d}) := \left(\frac{-1}{\alpha}\right)\left(\frac{2}{\beta_\eta}\right)\prod_i 4^{-\omega(d_{i0})}\prod_{j\neq 0} 4^{-\omega(d_{ij})}\prod_{k\neq i,j}\prod_l \left(\frac{d_{kl}}{d_{ij}}\right),$$

where

$$\beta_1 = d_{23}d_{21}d_{32}d_{31}d_{43}d_{42}, \qquad \beta_2 = d_{13}d_{14}d_{24}d_{21}d_{43}d_{41},$$
$$\beta_3 = d_{12}d_{14}d_{34}d_{31}d_{42}d_{41}, \qquad \beta_4 = d_{13}d_{12}d_{23}d_{24}d_{34}d_{32}.$$

Then

$$G_\eta(d_1, d_2, d_3, d_4) = \sum_{\substack{\mathbf{d} \\ \prod_{j\neq i} d_{ij}=d_i}} g_\eta(\mathbf{d}).$$

5.2. Setting up the sums

We now set up the sum which bounds the number of elements in $\mathcal{W}_D$ that satisfy (**W1**). For each $\eta \in \{0, 1, 2, 3, 4\}$, we want to estimate the sum

$$\sum_{\substack{(d_1,d_2,d_3,d_4) \\ (\mathbf{W1})}} G_\eta(d_1, d_2, d_3, d_4),$$

where the sum is taken over all positive odd integers d_1, d_2, d_3, d_4 that satisfy (**W1**) and such that $d_1d_2d_3d_4 \in \mathcal{D}_N$. Following [10, Section 3], dissect the sum according to the size of each d_{ij} in the factorisation. For each (i, j), take A_{ij} to run over powers of 2. Then for $\mathbf{A} = (A_{ij})$, define the restricted sum

$$S_\eta^{(k)}(\mathbf{A}) := \sum_{\substack{\mathbf{d} \\ A_{ij} < d_{ij} \leq 2A_{ij}}} g_\eta(\mathbf{d}),$$

where the sum is taken over all 16-tuples of odd positive integers $\mathbf{d} = (d_{ij})$ such that $\prod_{i,j} d_{ij} \in \mathcal{D}_N$ and $A_{ij} < d_{ij} \leq 2A_{ij}$ for every i, j, with the further condition that

$$\prod_j d_{kj} \leq N^\ddagger. \tag{23}$$

The property (23) is equivalent to $d_k \leq N^\ddagger$, which for any given $k \in \{1, 2, 3, 4\}$ is a subcase of (**W1**). Note that if $A_{ij} = \frac{1}{2}$, the interval $A_{ij} < d_{ij} \leq 2A_{ij}$ forces $d_{ij} = 1$. To capture the property (20) from (**W1**), we exclude $\mathbf{A}$ that satisfy

$$\begin{cases} A_{ij} = \frac{1}{2} \text{ for all } i \in \{1, 2, 3\} & \text{if } \eta = 0 \text{ or } 4, \\ A_{ij} = \frac{1}{2} \text{ for all } i \in \{1, 2, 4\} & \text{if } \eta = 2. \end{cases} \tag{24}$$

Then

$$\sum_{\substack{(d_1,d_2,d_3,d_4) \\ (\mathbf{W1})}} G_\eta(d_1, d_2, d_3, d_4) \leq \sum_{k=1}^{4} \sum_{\mathbf{A}} S_\eta^{(k)}(\mathbf{A}), \tag{25}$$

where the sum runs over all $\mathbf{A}$ except those that satisfy (24). We shall begin bounding (25) in Section 5.4.

We next set up the sum that treats the property (**W2**). For ease of notation, assume that it is $d_4 > N^\ddagger$ that has no prime factor in S. The cases with d_4 replaced by d_1, d_2, d_3 will turn out to be the same after relabelling. We want to bound

$$\sum_{\substack{d_1d_2d_3d_4 \in \mathcal{D}_N \\ d_1,d_2,d_3,d_4 > N^\ddagger \\ p \mid d_4 \Rightarrow p \notin S}} G_\eta(d_1, d_2, d_3, d_4).$$

Similar to the previous case, define the restricted sum

$$S'_\eta(\mathbf{A}) := \sum_{\substack{\mathbf{d} \\ A_{ij} < d_{ij} \leq 2A_{ij}}} g_\eta(\mathbf{d}),$$

where the sum is taken over all 16-tuples of odd positive integers $\mathbf{d} = (d_{ij})$ such that $\prod_{i,j} d_{ij} \in \mathcal{D}_N$ and $A_{ij} < d_{ij} \leq 2A_{ij}$ for every i, j, with the extra conditions that

$$(p \mid d_{40}d_{41}d_{42}d_{43} \Rightarrow p \notin S) \quad \text{and} \tag{26}$$

$$d_{10}d_{12}d_{13}d_{14},\ d_{20}d_{21}d_{23}d_{24},\ d_{30}d_{31}d_{32}d_{34},\ d_{40}d_{41}d_{42}d_{43} > N^{\ddagger}. \tag{27}$$

Then

$$\sum_{\substack{d_1d_2d_3d_4 \in \mathcal{D}_N \\ d_1,d_2,d_3,d_4 > N^{\ddagger} \\ p \mid d_4 \Rightarrow p \notin S}} G_\eta(d_1, d_2, d_3, d_4) = \sum_{\mathbf{A}} S'_\eta(\mathbf{A}). \tag{28}$$

5.3. Preliminaries

We collect some results used in [10] which we will utilise.

Lemma 5.2 *([19, Theorem 1]). Fix $0 < \epsilon < 1$ and some positive constant C. Let f be a multiplicative function such that $f(p^\ell) \leq C$ for all prime p and $\ell \geq 1$. Then*

$$\sum_{X-Y<n\leq X} f(n) \ll \frac{Y}{\log X} \exp\left(\sum_{p \leq X} \frac{f(p)}{p}\right)$$

uniformly for $2 \leq X^{1-\epsilon} \leq Y < X$.

The next result by Heath-Brown handles double oscillation of characters.

Lemma 5.3 *([10, Lemma 4]). Let a_m, b_n be complex numbers of modulus at most 1. Let $M, N, X \gg 1$. Then*

$$\sum_{m,n} a_m b_n \left(\frac{n}{m}\right) \ll MN \min\{M, N\}^{-\frac{1}{32}}$$

uniformly in X, where the sum is for squarefree m, n satisfying $M < m \leq 2M$, $N < n \leq 2N$, $mn \leq X$.

We will also use the following version of the Siegel–Walfisz theorem for character sums.

Lemma 5.4 *([10, Lemma 6]). Let $k > 0$ be given. Let $d(r)$ denote the number of divisors of r. Then for arbitrary positive integers q, r and any non-principal character χ mod q, we have*

$$\sum_{\substack{n \in \mathcal{D}_X \\ \gcd(n,r)=1}} 4^{-\omega(n)} \chi(n) \ll X \cdot d(r) \cdot \exp\left(-c\sqrt{\log X}\right),$$

with a positive constant c depending only on k, uniformly for $q \leq (\log X)^k$.

5.4. Bounding the subsums

We proceed to bound the subsums in (25) and (28) following [10, Section 3] closely. To study the indices $\mathbf{u}, \mathbf{v}$ of the symbols $\left(\frac{d_{\mathbf{u}}}{d_{\mathbf{v}}}\right)$ that can appear in the expression of g_η, we make the following definition as in [10].

Definition 5.5 *(linked indices). We call two indices $\mathbf{u} = (i, j)$ and $\mathbf{v} = (k, l)$ linked if*

$$i \neq k \text{ and precisely one of the conditions } \begin{cases} l \notin \{0, i\}, \\ j \notin \{0, k\} \end{cases} \text{ holds.}$$

If two indices $\mathbf{u}$ and $\mathbf{v}$ are linked, then exactly one of $\left(\frac{d_{\mathbf{u}}}{d_{\mathbf{v}}}\right)$ and $\left(\frac{d_{\mathbf{v}}}{d_{\mathbf{u}}}\right)$ appears in the sum g_η. When both $\left(\frac{d_{\mathbf{u}}}{d_{\mathbf{v}}}\right)$ and $\left(\frac{d_{\mathbf{v}}}{d_{\mathbf{u}}}\right)$ appear in the expression, which is a possibility if $\mathbf{u}$ and $\mathbf{v}$ are unlinked, we can apply quadratic reciprocity to get $\left(\frac{d_{\mathbf{u}}}{d_{\mathbf{v}}}\right)\left(\frac{d_{\mathbf{v}}}{d_{\mathbf{u}}}\right) = (-1)^{\frac{d_{\mathbf{u}}-1}{2} \cdot \frac{d_{\mathbf{v}}-1}{2}}$.

We first bound the contribution of $S_\eta^{(k)}(\mathbf{A})$ from $\mathbf{A}$ with less than 4 large indices to the sum (25).

Lemma 5.6. *We have*

$$\sum_{\substack{\mathbf{A} \\ \#\{\mathbf{u}: A_{\mathbf{u}} \geq N^\ddagger\} \leq 3}} |S_\eta^{(k)}(\mathbf{A})| \ll N(\log N)^{-\frac{1}{4}+\epsilon}.$$

Proof. Let $\mathcal{W} = \{\mathbf{u} : A_{\mathbf{u}} \geq N^\ddagger\}$. Bound $|g_\eta(\mathbf{d})|$ trivially by $\prod_{i,j} 4^{-\omega(d_{ij})}$. Write $m = \prod_{\mathbf{u} \notin \mathcal{W}} d_{\mathbf{u}}$ and $n = \prod_{\mathbf{u} \in \mathcal{W}} d_{\mathbf{u}}$. Then for any given prime factor of n there are $\#\mathcal{W} \leq 3$ ways to place it into one of $\{d_{\mathbf{u}} : \mathbf{u} \in \mathcal{W}\}$, and for any given prime factor of m there are at most 16 ways to place it into one of $\{d_{\mathbf{u}} : \mathbf{u} \notin \mathcal{W}\}$. Therefore

$$\sum_{\substack{\mathbf{A} \in \mathcal{F} \\ \#\mathcal{W} \leq 3}} |S_\eta^{(k)}(\mathbf{A})| \ll \sum_{m < (N^\ddagger)^{16}} \left(\frac{16}{4}\right)^{\omega(m)} \sum_{n \leq \frac{N}{m}} \left(\frac{3}{4}\right)^{\omega(n)}.$$

Applying Lemma 5.2 and Mertens theorem, the inner sum becomes

$$\sum_{n\le \frac{N}{m}} \left(\frac{3}{4}\right)^{\omega(n)} \ll \frac{N}{m}(\log N)^{-\frac{1}{4}}.$$

Then substituting this back gives

$$\sum_{\substack{\mathbf{A}\in\mathcal{F}\\ \#\mathcal{W}\le 3}} |S_\eta^{(k)}(\mathbf{A})| \ll N(\log N)^{-\frac{1}{4}} \sum_{m\le (N^\ddagger)^4} \frac{4^{\omega(m)}}{m} \ll N(\log N)^{-\frac{1}{4}+\epsilon}.$$

This gives the claimed upper bound. □

When there are two large variables with linked indices, we apply Lemma 5.3.

Lemma 5.7. *Suppose* $A_{\mathbf{u}}, A_{\mathbf{v}} \ge (\log N)^{544}$ *for some linked indices* $\mathbf{u}$ *and* $\mathbf{v}$*. Then*

$$S'_\eta(\mathbf{A}) \ll N(\log N)^{-17} \quad \text{and} \quad S_\eta^{(k)}(\mathbf{A}) \ll N(\log N)^{-17}.$$

Proof. We follow the proof of [10, Lemma 5]. Since $\mathbf{u}$ and $\mathbf{v}$ are linked, we can write

$$g_\eta(\mathbf{d}) = a(d_{\mathbf{u}})b(d_{\mathbf{v}})\left(\frac{d_{\mathbf{u}}}{d_{\mathbf{v}}}\right),$$

where the functions a and b depends on the other variables $(d_{\mathbf{w}})_{\mathbf{w}\ne\mathbf{u},\mathbf{v}}$ but is independent of $d_{\mathbf{v}}$ and $d_{\mathbf{u}}$. Moreover $|a(d_{\mathbf{u}})|, |b(d_{\mathbf{v}})| \le 1$.

For $S'_\eta(\mathbf{A})$, when $d_{\mathbf{u}}$ does not satisfy (26), we impose that $a(d_{\mathbf{u}}) = 0$. Similarly when $d_{\mathbf{v}}$ does not satisfy (26), we impose $b(d_{\mathbf{v}}) = 0$. We have

$$S'_\eta(\mathbf{A}) \ll \sum_{\substack{(d_{\mathbf{w}})_{\mathbf{w}\ne\mathbf{u},\mathbf{v}}\\ A_{\mathbf{w}}<d_{\mathbf{w}}\le 2A_{\mathbf{w}}}} \left| \sum_{\substack{d_{\mathbf{u}}\\ A_{\mathbf{u}}<d_{\mathbf{u}}\le 2A_{\mathbf{u}}}} \sum_{\substack{d_{\mathbf{v}}\\ A_{\mathbf{v}}<d_{\mathbf{v}}\le 2A_{\mathbf{v}}}} a(d_{\mathbf{u}})b(d_{\mathbf{v}})\left(\frac{d_{\mathbf{u}}}{d_{\mathbf{v}}}\right)\right|,$$

where the sum is further subject to $\prod_{i,j} d_{ij} \in \mathcal{D}_N$ and (27) being satisfied. Then an application of Lemma 5.3 implies that

$$S'_\eta(\mathbf{A}) \ll \sum_{\substack{(d_{\mathbf{w}})_{\mathbf{w}\ne\mathbf{u},\mathbf{v}}\\ A_{\mathbf{w}}<d_{\mathbf{w}}\le 2A_{\mathbf{w}}}} A_{\mathbf{u}}A_{\mathbf{v}}(\min\{A_{\mathbf{u}}, A_{\mathbf{v}}\})^{-\frac{1}{32}} \ll N(\log N)^{-17},$$

where we have substituted the lower bound for $A_{\mathbf{u}}, A_{\mathbf{v}}$.

The sum for $S_\eta^{(k)}(\mathbf{A})$ can be bounded similarly. □

If the set of indices $\mathcal{L}$ that are linked to $\mathbf{u}$ are such that $\prod_{\mathbf{v}\in\mathcal{L}} d_{\mathbf{v}} \ne 1$ is small and $d_{\mathbf{u}}$ is large, we apply Lemma 5.4 instead.

Lemma 5.8. *Fix an index* $\mathbf{u}$*. Let* $\mathcal{L}$ *be the set of indices that are linked to* $\mathbf{u}$*. Suppose* $A_{\mathbf{v}} < (\log N)^{544}$ *holds for every* $\mathbf{v} \in \mathcal{L}$*, and* $A_{\mathbf{v}} \neq \frac{1}{2}$ *for at least one* $\mathbf{v} \in \mathcal{L}$*. Then if* $A_{\mathbf{u}} \geq (N^{\ddagger})^{\frac{1}{4}}$*, we have*

$$S'_{\eta}(\mathbf{A}) \ll_{\epsilon} N(\log N)^{-17} \text{ whenever } \mathbf{u} \notin \{40, 41, 42, 43\},$$

and

$$S_{\eta}^{(k)}(\mathbf{A}) \ll_{\epsilon} N(\log N)^{-17} \text{ holds for any } \mathbf{u}.$$

Proof. We follow the proof of [10, Lemma 7]. Write d' for the product $\prod_{\mathbf{v}\in\mathcal{L}} d_{\mathbf{v}}$. We can put $g_{\eta}(\mathbf{d})$ into the form

$$4^{-\omega(d_{\mathbf{u}})}\left(\frac{d_{\mathbf{u}}}{d'}\right)\chi(d_{\mathbf{u}})C,$$

where χ is a character modulo 8, $|C| \leq 1$, and χ and C do not depend on $d_{\mathbf{u}}$. Then

$$S'_{\eta}(\mathbf{A}) \ll \sum_{(d_{\mathbf{w}})_{\mathbf{w}\neq\mathbf{u}}} \left|\sum_{d_{\mathbf{u}}} 4^{-\omega(d_{\mathbf{u}})}\left(\frac{d_{\mathbf{u}}}{d'}\right)\chi(d_{\mathbf{u}})\right|,$$

where the sum of $d_{\mathbf{u}}$ is subject to the conditions $A_{\mathbf{u}} < d_{\mathbf{u}} \leq 2A_{\mathbf{u}}$, $\prod_{i,j} d_{ij} \in \mathcal{D}_N$ and (27) (but not (26) because $\mathbf{u} \notin \{40, 41, 42, 43\}$ by assumption). Then $\left(\frac{\cdot}{d'}\right)\chi$ is a character mod $8d'$ and non-principal because $d' \neq 1$. Since $8d' \ll (\log N)^{544\cdot 15} \leq (\log(N^{\ddagger}))^{\frac{4}{\epsilon}\cdot 544\cdot 15}$, we can apply Lemma 5.4, then sum over all $(d_{\mathbf{w}})_{\mathbf{w}\neq\mathbf{u}}$, we get

$$S'_{\eta}(\mathbf{A}) \ll_{\epsilon} N(\log N)^{15}\exp\left(-c\sqrt{\log A_{\mathbf{u}}}\right),$$

where c is a constant depending only on ϵ. Inserting the lower bound $A_{\mathbf{u}} \geq (N^{\ddagger})^{\frac{1}{4}}$ yields the required estimate.

The proof for $S_{\eta}^{(k)}(\mathbf{A})$ is similar noting that imposing (23) does not require the assumption $\mathbf{u} \notin \{40, 41, 42, 43\}$. □

Combining Lemma 5.7 and Lemma 5.8 we have the following.

Lemma 5.9. *Suppose* $A_{\mathbf{u}} \geq (N^{\ddagger})^{\frac{1}{4}}$ *and* $A_{\mathbf{v}} \neq \frac{1}{2}$ *hold for some linked indices* $\mathbf{u}$ *and* $\mathbf{v}$*. Then*

$$S'_{\eta}(\mathbf{A}) \ll_{\epsilon} N(\log N)^{-17} \text{ whenever } \mathbf{u} \notin \{40, 41, 42, 43\},$$

and

$$S_{\eta}^{(k)}(\mathbf{A}) \ll_{\epsilon} N(\log N)^{-17} \text{ holds for any } \mathbf{u}.$$

The following is inferred by [10, Lemma 9].

Lemma 5.10. *If $\mathcal{U}$ is a set of pairwise unlinked indices, and $\#\mathcal{U} \geq 4$, then $\mathcal{U}$ takes one of the following form*

$$\begin{gathered}\{10,20,30,40\},\{i0,j0,ij,ji\},\{i0,ij,ik,il\},\\ \{i0,ji,ki,li\},\{ij,ik,lj,lk\},\{ij,ji,kl,lk\},\end{gathered} \tag{29}$$

where i,j,k,l denote different non-zero indices.

5.5. The case $d_i \leq N^{\ddagger}$

We now work towards bounding the contribution from (**W1**).

Lemma 5.11. *For each $\eta \in \{0,1,2,3,4\}$, we have*

$$\#\left\{\left(2^{\delta_1(\eta)}d_1, 2^{\delta_2(\eta)}d_2, 2^{\delta_3(\eta)}d_3, 2^{\delta_4(\eta)}d_4\right) \in \mathcal{W}_D : (\textbf{W1}) \text{ holds}\right\} \ll_\epsilon N(\log N)^{-\frac{1}{4}+\epsilon}.$$

We adapt the argument in [10, Lemma 9, Lemma 11] to prove Lemma 5.11.

Lemma 5.12. *For each $k \in \{1,2,3,4\}$,*

$$\sum_{\mathbf{A}} |S_\eta^{(k)}(\mathbf{A})| \ll_\epsilon N(\log N)^{-\frac{1}{4}+\epsilon},$$

where the sum is over all $\mathbf{A}$ other than those that satisfy

$$A_{\mathbf{u}} \geq N^{\ddagger} \text{ for all } \mathbf{u} \in \mathcal{U} \quad \text{and} \quad A_{\mathbf{u}} = \frac{1}{2} \text{ for all } \mathbf{u} \notin \mathcal{U} \tag{30}$$

for some $\mathcal{U}$ being one of

$$\begin{cases} \{10,20,30,40\},\{40,41,42,43\},\{20,12,32,42\},\{30,13,23,43\} & \text{if } \eta = 0,\\ \{10,20,30,40\},\{40,14,24,34\} & \text{if } \eta = 1,\\ \{10,20,30,40\},\{20,12,22,32\},\{30,31,32,34\} & \text{if } \eta = 2,\\ \{10,20,30,40\},\{30,13,23,43\} & \text{if } \eta = 3,\\ \{10,20,30,40\},\{10,11,21,31\},\{40,41,42,43\} & \text{if } \eta = 4. \end{cases} \tag{31}$$

Proof. By Lemma 5.6, we can assume that there exists a set of indices $\mathcal{U}$ of size at least 4, such that $A_{\mathbf{u}} \geq N^{\ddagger}$ for all $\mathbf{u} \in \mathcal{U}$. By Lemma 5.9, we can assume that the indices in $\mathcal{U}$ are pairwise unlinked.

Therefore it remains to show that for each $\eta \in \{0,1,2,3,4\}$, the bound

$$\sum_{\substack{\mathbf{A} \\ u\in\mathcal{U}\Rightarrow A_{\mathbf{u}}\geq N^{\ddagger} \\ u\notin\mathcal{U}\Rightarrow A_{\mathbf{u}}=\frac{1}{2}}} |S_{\eta}^{(k)}(\mathbf{A})| \ll_{\epsilon} N(\log N)^{\frac{1}{4}+\epsilon} \tag{32}$$

holds for every $\mathcal{U}$ in (29), unless $\mathcal{U}$ is one of the sets in (31). When D is odd, namely when $\eta = 0$, (32) essentially follows from [10, Lemma 11].

For even D, fix any $\eta \in \{1,2,3,4\}$, then consider $\mathbf{A}$ such that $A_{\mathbf{u}} \geq N^{\ddagger}$ for all $\mathbf{u} \in \mathcal{U}$ and $A_{\mathbf{u}} = \frac{1}{2}$ for all $\mathbf{u} \notin \mathcal{U}$. We see that for every possible $\mathcal{U}$, there exists $\mathbf{v} \in \mathcal{U}$ such that $d_{\mathbf{v}}$ is one of the variables in β_η. Now fix one such $\mathbf{v} \in \mathcal{U}$. Since the indices in $\mathcal{U}$ are unlinked, putting in $d_{\mathbf{u}} = 1$ for all $\mathbf{u} \notin \mathcal{U}$, we see that $g_\eta(\mathbf{d})$ has the form

$$g_\eta(\mathbf{d}) = \left(\frac{-1}{\alpha'}\right)\left(\frac{2}{\beta'_\eta}\right)\prod_{\mathbf{u}\in\mathcal{U}} 4^{-\omega(d_{\mathbf{u}})} \prod_{\{\mathbf{u},\mathbf{w}\}\subset\mathcal{U}} \varphi_{\mathbf{u},\mathbf{w}}(d_{\mathbf{u}}, d_{\mathbf{w}}),$$

where α' is the product of the variables dividing α with indices in $\mathcal{U}$, β'_η is the product of variables dividing β_η with indices in $\mathcal{U}$, and $\varphi_{\mathbf{u},\mathbf{w}}(d_{\mathbf{u}}, d_{\mathbf{w}})$ is either trivially 1 or $(-1)^{\frac{d_{\mathbf{u}}-1}{2}\cdot\frac{d_{\mathbf{w}}-1}{2}}$ depending on the indices $\mathbf{u}, \mathbf{w}$. Viewing $(d_{\mathbf{u}})_{\mathbf{u}\in\mathcal{U}\setminus\{\mathbf{v}\}}$ as fixed, we can write

$$g_\eta(\mathbf{d}) = 4^{-\omega(d_{\mathbf{v}})}\chi(d_{\mathbf{v}})C,$$

where C depends on $(d_{\mathbf{u}})_{u\in\mathcal{U}\setminus\{\mathbf{v}\}}$ but not $d_{\mathbf{v}}$ and satisfies $|C| \leq 1$, and the function $\chi(d_{\mathbf{v}})$ is $\left(\frac{2}{d_{\mathbf{v}}}\right)$ or $\left(\frac{-2}{d_{\mathbf{v}}}\right)$ depending on $(d_{\mathbf{u}})_{\mathbf{u}\in\mathcal{U}\setminus\{\mathbf{v}\}}$ and whether $\mathbf{v}$ is the index of a variable dividing α. Then we have

$$|S_\eta^{(k)}(\mathbf{A})| \ll \sum_{(d_{\mathbf{u}})_{\mathbf{u}\in\mathcal{U}\setminus\{\mathbf{v}\}}} \left|\sum_{d_{\mathbf{v}}} 4^{-\omega(d_{\mathbf{v}})}\chi(d_{\mathbf{v}})\right|,$$

where $(d_{\mathbf{u}})$ are restricted to satisfy $A_{\mathbf{u}} < d_{\mathbf{u}} \leq 2A_{\mathbf{u}}$, $\prod_{\mathbf{u}\in\mathcal{U}} d_{\mathbf{u}} \in \mathcal{D}_N$ and (23). Apply Lemma 5.4 to the inner sum we conclude that

$$|S_\eta^{(k)}(\mathbf{A})| \ll_\epsilon N(\log N)^{-17}$$

as in the proof of [10, Lemma 7]. Summing over all $O((\log N)^{16})$-many possible $\mathbf{A}$, the bound in (32) holds as required. □

We are ready to bound the contribution from (**W1**).

Proof of Lemma 5.11. By (22), G_η provides an upper bound to the indicator function of $\mathcal{W}_D$, so we can bound the number of elements in $\mathcal{W}_D$ satisfying (**W1**) by the sum in (25). The assumption (23) implies that $S_\eta^{(k)}(\mathbf{A})$ is an empty sum whenever $A_{kj} \geq N^{\ddagger}$

for some j. Checking the sets in (31), the only possibility that $S_\eta^{(k)}(\mathbf{A})$ is non-trivial and not covered by Lemma 5.12, is if $\mathbf{A}$ satisfies (30) with

$$\mathcal{U} = \begin{cases} \{40, 41, 42, 43\} & \text{when } \eta = 0, \\ \{30, 31, 32, 34\} & \text{when } \eta = 2, \\ \{40, 41, 42, 43\} & \text{when } \eta = 4, \end{cases}$$

but these are within the exclusions set out in (24). This completes the proof. □

5.6. Prime divisors of a large d_i

We now bound the contribution from (**W2**).

Lemma 5.13. *For each $\eta \in \{0, 1, 2, 3, 4\}$, we have*

$$\#\left\{\left(2^{\delta_1(\eta)}d_1, 2^{\delta_2(\eta)}d_2, 2^{\delta_3(\eta)}d_3, 2^{\delta_4(\eta)}d_4\right) \in \mathcal{W}_D : (\mathbf{W2}) \text{ holds}\right\} \ll_\epsilon N(\log N)^{-\frac{1}{4}+\epsilon}.$$

To prove Lemma 5.13, we again modify the estimates in [10, Section 3] to account for the extra restrictions in the sum.

Lemma 5.14.

$$\sum_{\mathbf{A}} S'_\eta(\mathbf{A}) \ll_\epsilon \sum_{\mathcal{U}} \sum_{\substack{\mathbf{A} \\ \mathbf{v}\notin\mathcal{U} \Rightarrow A_\mathbf{v} = \frac{1}{2}}} S'_\eta(\mathbf{A}) + N(\log N)^{-1},$$

where the sum over $\mathcal{U}$ is over all $\mathcal{U}$ of the form $\{1i, 2j, 3k, 4l\}$.

Proof. For each $\mathbf{A}$ such that $S'_\eta(\mathbf{A})$ is non-trivial, the condition (27) allows us to find a set of indices

$$\mathcal{U} = \{1i, 2j, 3k, 4l\},$$

where i, j, k, l are not necessarily distinct, such that d_{1i}, d_{2j}, d_{3k}, $d_{4l} > (N^\ddagger)^{\frac{1}{4}}$. Hence we may assume that $A_{1i}, A_{2j}, A_{3k}, A_{4l} \geq (N^\ddagger)^{\frac{1}{4}}$. By Lemma 5.7, we can further assume that the indices $1i, 2j, 3k, 4l$ are pairwise unlinked, so $\mathcal{U}$ must take one of the form in (29).

Now suppose $\mathbf{v} \notin \mathcal{U}$. If $\mathbf{v}$ is not linked to any one of $1i, 2j, 3k$, then $\{1i, 2j, 3k, \mathbf{v}\}$ is also a set of unlinked indices with size 4. Comparing against the list in (29), if $\{1i, 2j, 3k, 4l\}$ and $\{1i, 2j, 3k, \mathbf{v}\}$ are both sets of unlinked indices, they must be the same set, which contradicts the assumption that $\mathbf{v} \notin \mathcal{U}$. Therefore $\mathbf{v}$ must be linked to one of $\{1i, 2j, 3k\}$, and this allows us to apply Lemma 5.9. Hence we are left with the terms $S'_\eta(\mathbf{A})$ with $A_\mathbf{v} = \frac{1}{2}$ for all $\mathbf{v} \notin \mathcal{U}$. The sum of $S'_\eta(\mathbf{A})$ over those $\mathbf{A}$ treated by Lemma 5.7 and Lemma 5.9 contributes $O(N(\log N)^{-1})$ since there are $O((\log N)^{16})$ possible $\mathbf{A}$. □

Proof of Lemma 5.13. It suffices to bound (28). We further simplify the expression obtained in Lemma 5.14. Note that there are only finitely many possible $\mathcal{U} = \{1i, 2j, 3k, 4l\}$, then on setting $d_1 = d_{1i}$, $d_2 = d_{2j}$, $d_3 = d_{3k}$, $d_4 = d_{4l}$, we deduce that

$$\sum_{\mathcal{U}} \sum_{\substack{\mathbf{A} \\ \mathbf{v} \notin \mathcal{U} \Rightarrow A_{\mathbf{v}} = \frac{1}{2}}} S'_\eta(\mathbf{A}) \ll \sum_{\substack{d_1 d_2 d_3 d_4 \in \mathcal{D}_N \\ d_1, d_2, d_3, d_4 \geq N^{\ddagger} \\ p \mid d_4 \Rightarrow p \notin S}} 4^{-\omega(d_1 d_2 d_3 d_4)} \leq \sum_{D \in \mathcal{D}_N} 4^{-\omega(D)} \sum_{\substack{(d_1, d_2, d_3, d_4) \\ d_1 d_2 d_3 d_4 = D \\ p \mid d_4 \Rightarrow p \notin S}} 1$$

$$= \sum_{D \in \mathcal{D}_N} \left(\frac{3}{4}\right)^{\omega_S(D)} \ll_\epsilon N(\log N)^{-\frac{1}{4}+\epsilon},$$

where the final inequality follows from Lemma 5.2 and Mertens theorem. Therefore we conclude that

$$\sum_{\mathbf{A}} S'_\eta(\mathbf{A}) \ll_\epsilon N(\log N)^{-\frac{1}{4}+\epsilon},$$

which gives the required bound for (28) as desired. □

Combining Lemma 5.11 and Lemma 5.13 proves Theorem 5.1.

6. Counting generic points

The goal of this section is to prove Theorem 1.4. We begin by collecting the exceptional set of $D \in \mathcal{D}_N$ that will be disregarded in the subsequent argument. Take $\mathcal{G}_N$ to be the collection of $D \in \mathcal{D}_N$ that satisfy at least one of the following:

(**P1**) $\omega(D) \geq 2 \log \log N$,
(**P2**) $D < \exp(3(\log N)^{1-2\epsilon})$,
(**P3**) the conditions (**S1**) and (**S2**) both fail for some $(D_1, D_2, D_3, D_4) \in \mathcal{W}_D$.

Lemma 6.1. *We have*

$$\#\mathcal{G}_N \ll_\epsilon N(\log N)^{-\frac{1}{4}+\epsilon}.$$

Proof. By the Erdős-Kac theorem [7], the number of $D \in D_N$ satisfying (**P1**) is bounded by $O(N(\log N)^{-\frac{1}{2}})$. The number of $D \in \mathcal{D}_N$ that satisfy (**P2**) is trivially bounded by $\exp(3(\log N)^{1-2\epsilon})$. Theorem 5.1 allows us to bound the number of $D \in \mathcal{D}_N$ that satisfy (**P3**) by $O_\epsilon(N(\log N)^{-\frac{1}{4}+\epsilon})$. Collecting the upper bounds proves the lemma. □

Recall that any integral point in $\mathcal{E}_D(\mathbb{Z})$ maps to $\mathcal{W}_D$ under the map in (15), and

$$\mathcal{E}^*_D(\mathbb{Z}) = \mathcal{E}_D(\mathbb{Z}) \setminus \{(0,0), (\pm D, 0)\}.$$

For the non-trivial integral points that have image of the type (**S1**), we have the following bound from [5, Theorem 1.4] and the discussion after [5, Theorem 10.1].

Lemma 6.2. *We have*

$$\sum_{D\in\mathcal{D}_N}\sum_{T\in\{O,(0,0),(\pm D,0)\}}\#(\mathcal{E}_D^*(\mathbb{Z})\cap(T+2\mathcal{E}_D(\mathbb{Q})))\ll\sqrt{N}\log N.$$

Therefore it remains to handle the integral points on $\mathcal{E}_D$ with $D\in\mathcal{D}_N\setminus\mathcal{G}_N$ that have image satisfying (**S2**). Define

$$\mathcal{Z}_N:=\bigcup_{D\in\mathcal{D}_N\setminus\mathcal{G}_N}\{P\in\mathcal{E}_D(\mathbb{Z}):P\notin 2\mathcal{E}_D(\mathbb{Q})+\{O,(0,0),(\pm D,0)\}\}.$$

Then the image of $P=(x,y)\in\mathcal{Z}_N$ corresponds to $(D_1,D_2,D_3,D_4)\in\mathcal{W}_D$ of the type (**S2**). By Lemma 4.2, we have

$$\frac{D}{\gcd(x,D)}\in\{D_1,D_2,D_3,D_4\},$$

so the property (**S2**) allows us to pick a prime factor M_P of $D/\gcd(x,D)$ of size

$$\exp((\log N)^{2\epsilon})<M_P<\exp((\log N)^{1-2\epsilon}).\tag{33}$$

Now since D is squarefree, M_P divides D but does not divide x. Therefore we can apply the map Ψ defined in (9) to (P,M_P). Having fixed a choice of M_P for each $P\in\mathcal{Z}_N$, define

$$\Psi':\mathcal{Z}_N\to(V\times\mathbb{Z}^2)/\operatorname{SL}_2(\mathbb{Z})$$

by

$$P\mapsto(P,M_P)\overset{\Psi}{\longmapsto}(F,(1,0)).$$

Also define

$$\Phi:\mathcal{Z}_N\xrightarrow{\Psi'}(V\times\mathbb{Z}^2)/\operatorname{SL}_2(\mathbb{Z})\to V/\operatorname{SL}_2(\mathbb{Z})$$

by

$$P\overset{\Psi'}{\longmapsto}(F,(1,0))\mapsto F.$$

By Lemma 3.1(i), if $\Psi'(P)=(F,(1,0))$, then $F(1,0)=M_P$ and so (33) can be rewritten as

$$\exp((\log N)^{2\epsilon}) < F(1,0) < \exp((\log N)^{1-2\epsilon}). \tag{34}$$

For any $P \in \mathcal{Z}_N$, write

$$\tilde{D} = \frac{D}{M_P},$$

so $\Delta(F) = (2\tilde{D})^6$ if $F = \Phi(P)$ by Lemma 3.1(iii). Since $D \geq \exp(3(\log N)^{1-2\epsilon})$ by (**P2**) and M_P is in the range (33), we have

$$\exp(2(\log N)^{1-2\epsilon}) \leq D\exp(-(\log N)^{1-2\epsilon}) \leq \tilde{D} < D\exp(-(\log N)^{2\epsilon}). \tag{35}$$

6.1. Points lowered to the same quartic

We now show that each class in $\operatorname{im}\Phi$ cannot arise from too many integral points.

Lemma 6.3. *For any $F \in \operatorname{im}\Phi$, we have*

$$\#\Phi^{-1}(F) \ll 1,$$

where the implied constant is absolute.

Proof. From Lemma 3.2, we know that Ψ is injective, so Ψ' is also injective. Therefore we want to show that the size of the fibres of $\operatorname{im}\Psi' \to \operatorname{im}\Phi \subset V/\operatorname{SL}_2(\mathbb{Z})$ is bounded. Fix an arbitrary $F_0 \in \operatorname{im}\Phi$. Suppose $(F,(1,0)) \in \operatorname{im}\Psi'$ is such that F and F_0 are $\operatorname{SL}_2(\mathbb{Z})$-equivalent, so we can write

$$F_0(X,Y) = F((X,Y)\cdot\gamma)$$

for some $\gamma \in \operatorname{SL}_2(\mathbb{Z})$. Then

$$\gamma\cdot(F(X,Y),(1,0)) = (F((X,Y)\cdot\gamma),(1,0)\cdot\gamma^{-1}) = (F_0(X,Y),(1,0)\cdot\gamma^{-1}).$$

Setting $(x,y) = (1,0)\cdot\gamma^{-1}$, we see that $F_0(x,y) = F(1,0)$, then by (34), (x,y) gives a solution to the Thue inequality

$$1 \leq |F_0(X,Y)| \leq h, \tag{36}$$

where $h := \exp((\log N)^{1-2\epsilon})$. In particular this solution is primitive (i.e. x and y coprime), since $\gamma^{-1} \in \operatorname{SL}_2(\mathbb{Z})$ has determinant 1 and entries in $\mathbb{Z}$. Therefore to bound the number of possible $\operatorname{SL}_2(\mathbb{Z})$-equivalence classes of $(F,(1,0))$ that maps to F_0, it suffices to bound the number of primitive solutions to (33).

A result by Evertse [8, Theorem 6.4(ii)] implies that when $2^8\Delta(F_0) \geq (13h)^{10}$, the number of solutions to (36) is bounded by some absolute constant. Since $\Delta(F_0) =$

$(2\tilde{D})^6 \gg \exp(12(\log N)^{1-2\epsilon})$ from (35), and $h^{10} = \exp(10(\log N)^{1-2\epsilon})$, we conclude that the number of possible classes $(F,(1,0)) \in \operatorname{im}\Psi'$ that maps to each class of F_0 is absolutely bounded. □

6.2. Integral points with bounded height

The last piece of the argument is to bound the size of the image of Φ. In the remainder of this paper, our task is to prove the following estimate.

Lemma 6.4. *We have*

$$\#\operatorname{im}\Phi \ll_\epsilon N\exp(-(\log N)^\epsilon).$$

Every integral binary quartic form is $\mathrm{SL}_2(\mathbb{Z})$-equivalent to at least one reduced form in the sense of [6, Section 4.3]. The seminvariant a, H of the reduced form are bounded in terms of the seminvariants I and J. We restate a theorem in [6] in terms of our rescaled seminvariants. The scale factors of the seminvariants can be found in Section 2.

Theorem 6.5 *([6, Proposition 11]). Suppose $F_0(X,Y) \in \mathbb{Z}[X,Y]$ is a $\mathrm{SL}_2(\mathbb{Z})$-reduced quartic form, and $\Delta(F_0) > 0$, with leading coefficient $a = a(F_0)$ and seminvariant $H = H(F_0)$. Order the three real roots ϕ_1, ϕ_2, ϕ_3 of $X^3 - \frac{I}{4}X - \frac{J}{4}$ so that $a\phi_1 < a\phi_2 < a\phi_3$. Then (a, H) satisfies one of the following:*

(1) $|a| \le \frac{4}{3}|\phi_1 - \phi_3|$ *and* $\max\{a\phi_1,\ a\phi_3 - 4\phi_3^2 + \frac{1}{3}I\} \le H \le a\phi_2$*; or*
(2) $|a| \le \frac{4}{3}|\phi_1 - \phi_2|$ *and* $a\phi_3 \le H \le a\phi_2 - 4\phi_2^2 + \frac{1}{3}I$*.*

For $F \in \operatorname{im}\Phi$, recall from the properties in Lemma 3.1 that $\Delta(F) = (2\tilde{D})^6 > 0$, $I(F) = 4\tilde{D}^2$ and $J(F) = 0$, so in the notation of Theorem 6.5, we have $\{\phi_1, \phi_3\} = \{-\tilde{D}, \tilde{D}\}$ and $\phi_2 = 0$. Suppose F_0 is a reduced form of F with leading coefficient $a = a(F_0)$ and seminvariant $H = H(F_0)$. Then the two possible cases in Lemma 6.5 both lead to

$$|a| \le \frac{8}{3}\tilde{D} \qquad \text{and} \qquad |H| \le \frac{4}{3}\tilde{D}^2. \tag{37}$$

The syzygy in (7) for F_0 now takes the form

$$H^3 - (a\tilde{D})^2 H = \left(\frac{1}{2}R\right)^2. \tag{38}$$

Notice that this gives an integral point $(H, \frac{1}{2}R) \in \mathcal{E}_{|a\tilde{D}|}(\mathbb{Z})$ when $a \ne 0$. In the following, we show that the possibility that $a = 0$ does not happen to the forms in $\operatorname{im}\Phi$.

Lemma 6.6. *Suppose $F \in \operatorname{im}\Phi$. Then any form in the $\mathrm{SL}_2(\mathbb{Z})$-equivalence class of F has non-zero leading coefficient.*

Proof. Assume for contradiction that $\Phi(P) = F$ for some $P = (c, d) \in \mathcal{Z}_N$ and F is equivalent to some quartic form with leading coefficient 0. Then there is a non-trivial integral solution to $F(X, Y) = 0$. From $\Phi(P) = F$, we know that $F(X, Y) = \frac{1}{M^3} f_P(MX + kY, Y)$ for some $M, k \in \mathbb{Z}$, so $f_P(X, Y) = 0$ has a non-trivial rational solution, say (x_0, y_0). Then from the expression of f_P in (8),

$$f_P(x_0, y_0) = x_0^4 - 6cx_0^2y_0^2 + 8dx_0y_0^3 + (4D^2 - 3c^2)y_0^4 = 0.$$

We see that $y_0 \neq 0$ since the solution is non-trivial. The roots of $f_P(X, 1)$ are

$$\frac{x_0}{y_0} = -\sqrt{c} + \sqrt{c+D} + \sqrt{c-D}, \quad -\sqrt{c} - \sqrt{c+D} - \sqrt{c-D}, \\ \sqrt{c} + \sqrt{c+D} - \sqrt{c-D}, \quad \sqrt{c} - \sqrt{c+D} + \sqrt{c-D}.$$

For x_0/y_0 to be rational, it must be that $\theta(P) = (1, 1, 1)$, where θ is the 2-descent homomorphism defined in (14). This implies that $P \in 2\mathcal{E}_D(\mathbb{Q})$, but such points were excluded from $\mathcal{Z}_N$. □

Since the seminvariants a, H, R, I and J together determine the quartic form up to $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes, to bound $\#\operatorname{im}\Phi$, it suffices to count the number of tuples $(a, \tilde{D}, H, R)$ associated to reduced forms in $\operatorname{im}\Phi$. In light of Theorem 6.5 and Lemma 6.6, to prove Lemma 6.4, we can assume (35), (37), (38), and $a \neq 0$. We split into two cases according to whether $(H, \frac{1}{2}R)$ is a torsion point on $\mathcal{E}_{|a\tilde{D}|}(\mathbb{Z})$.

6.3. Torsion points

Here we bound the number of classes in $\operatorname{im}\Phi$ that contain a reduced form which produces a torsion point $(H, \frac{1}{2}R) \in \mathcal{E}_{|a\tilde{D}|}(\mathbb{Z})$ through the syzygy (38). Recall from (35) that

$$\tilde{D} \leq N \exp(-(\log N)^{2\epsilon}).$$

Let

$$\tilde{N} = N \exp(-(\log N)^{2\epsilon}),$$

so that $\tilde{D} \leq \tilde{N}$.

Lemma 6.7. *The total number of $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes that contain an integer-matrix binary form F that satisfies $a(F) \neq 0$, $H(F) \in \{-a(F)\tilde{D}, 0, a(F)\tilde{D}\}$, $I(F) = (2\tilde{D})^2$ and $J(F) = 0$ for some $\tilde{D} \in \mathcal{D}_{\tilde{N}}$, is bounded by*

$$\ll_\epsilon N \exp(-(\log N)^{\epsilon}).$$

Proof. Suppose that $F(X,Y) = a_0X^4 + 4a_1X^3Y + 6a_2X^2Y^2 + 4a_3XY^3 + a_4Y^4$, so $a(F) = a_0$. Since $H(F) = a_1^2 - a_0a_2$, and by assumption $a_0 \mid H(F)$, it must be that $a_0 \mid a_1^2$. Then $a_0 \mid \Delta(F) = (2\tilde{D})^6$ by the formula of the discriminant in Section 2. Therefore for each $\tilde{D}$, there can only be a maximum of $2 \cdot 7^{\omega(2\tilde{D})}$ possible a_0. Inserting the assumptions to (7), we see that $R(F) = 0$. Summing over $\tilde{D} \in \mathcal{D}_{\tilde{N}}$, then applying Lemma 5.2, the number of classes can be bounded by

$$\#\left\{(a,\tilde{D},H) \in \mathbb{Z}^3 : \begin{array}{l} a \neq 0,\ a \mid (2\tilde{D})^6,\ \tilde{D} \in \mathcal{D}_{\tilde{N}}, \\ H \in \{-a\tilde{D},\ 0,\ a\tilde{D}\} \end{array}\right\} \ll \sum_{\tilde{D} \leq \tilde{N}} 7^{\omega(\tilde{D})} \ll \tilde{N}(\log \tilde{N})^6.$$

Finally putting in $\tilde{N} = N\exp(-(\log N)^{2\epsilon})$ completes the proof. □

6.4. Non-torsion points

We now bound the number of classes in $\operatorname{im}\Phi$ that contain a reduced form which produces a non-torsion point $(H, \frac{1}{2}R) \in \mathcal{E}_{|a\tilde{D}|}(\mathbb{Z})$ through the syzygy (38) and satisfies (37). Since $D \in \mathcal{D}_N \setminus \mathcal{G}_N$, those $\mathcal{E}_D$ that satisfies (**P1**) have been removed, we can assume that $\omega(\tilde{D}) < \omega(D) < 2\log\log N$. Also by Lemma 6.6, $a \neq 0$.

Lemma 6.8. *We have*

$$\#\left\{(a,\tilde{D},H,R) \in \mathbb{Z}^4 : \begin{array}{l} 1 \leq |a| \leq \frac{8}{3}\tilde{N},\ \ \tilde{D} \in \mathcal{D}_{\tilde{N}}, \\ \omega(\tilde{D}) < 2\log\log N, \\ \left(H, \frac{1}{2}R\right) \in \mathcal{E}^*_{|a\tilde{D}|}(\mathbb{Z}),\ \ |H| \leq \frac{4}{3}\tilde{D}^2 \end{array}\right\} \ll_\epsilon N\exp(-(\log N)^\epsilon).$$

Proof. Write $n = |a\tilde{D}| \leq \frac{8}{3}\tilde{N}^2$. For each positive integer n, the number of positive squarefree divisor $\tilde{D}$ of n satisfying $\omega(\tilde{D}) < 2\log\log N$, is bounded by

$$\begin{aligned} \sum_{k \leq \min\{\omega(n), 2\log\log N\}} \binom{\omega(n)}{k} &< \sum_{k \leq 2\log\log N} (\omega(n))^k \\ &\ll \exp\left(2(\log\log N)^2 + O(\log\log N)\right), \end{aligned} \tag{39}$$

where we have used the fact that $\omega(n) \ll \log N$.

The number of integral points $P = (H, \frac{1}{2}R) \in \mathcal{E}_n(\mathbb{Z})$ we are counting are of bounded height $|x(P)| \leq \frac{4}{3}\tilde{N}^2$, so applying a result by Le Boudec [15, Theorem 2] we get

$$\sum_{n \geq 1} \#\left\{P \in \mathcal{E}^*_n(\mathbb{Z}) : |x(P)| \leq \frac{4}{3}\tilde{N}^2\right\} \ll \tilde{N}(\log \tilde{N})^6. \tag{40}$$

Now multiplying together the upper bounds in (39) and (40), then substituting $\tilde{N} = N\exp(-(\log N)^{2\epsilon})$, we get that the total number of $(a, \tilde{D}, H, R)$ is

$$\ll N \exp\left(-(\log N)^{2\epsilon} + 2(\log\log N)^2 + O(\log\log N)\right).$$

This proves the claim. □

Lemma 6.7 and Lemma 6.8 completes the proof of Lemma 6.4. Theorem 1.4 follows from Lemma 6.1, Lemma 6.2, Lemma 6.3 and Lemma 6.4.

References

[1] S. Akhtari, A positive proportion of locally soluble quartic Thue equations are globally insoluble, Math. Proc. Camb. Philos. Soc. 173 (2) (2022) 333–348.

[2] S. Akhtari, M. Bhargava, A positive proportion of Thue equations fail the integral Hasse principle, Am. J. Math. 141 (2) (2019) 283–307.

[3] L. Alpöge, W. Ho, The second moment of the number of integral points on elliptic curves is bounded, Preprint, arXiv:1807.03761, 2022.

[4] E. Bombieri, W.M. Schmidt, On Thue's equation, Invent. Math. 88 (1) (1987) 69–81.

[5] S. Chan, Integral points on the congruent number curve, Trans. Am. Math. Soc. 375 (9) (2022) 6675–6700.

[6] J.E. Cremona, Reduction of binary cubic and quartic forms, LMS J. Comput. Math. 2 (1999) 64–94.

[7] P. Erdös, M. Kac, The Gaussian law of errors in the theory of additive number theoretic functions, Am. J. Math. 62 (1940) 738–742.

[8] J.-H. Evertse, Upper Bounds for the Numbers of Solutions of Diophantine Equations, Mathematical Centre Tracts, vol. 168, Mathematisch Centrum, Amsterdam, 1983.

[9] A. Granville, Rational and integral points on quadratic twists of a given hyperelliptic curve, Int. Math. Res. Not. (8) (2007) 027.

[10] D.R. Heath-Brown, The size of Selmer groups for the congruent number problem, Invent. Math. 111 (1) (1993) 171–195.

[11] D.R. Heath-Brown, The size of Selmer groups for the congruent number problem. II, Invent. Math. 118 (2) (1994) 331–370, With an appendix by P. Monsky.

[12] M. Hindry, J.H. Silverman, The canonical height and integral points on elliptic curves, Invent. Math. 93 (2) (1988) 419–450.

[13] N. Koblitz, Introduction to Elliptic Curves and Modular Forms, Graduate Texts in Mathematics, vol. 97, Springer-Verlag, New York, 1984.

[14] S. Lang, Elliptic Curves: Diophantine Analysis, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 231, Springer-Verlag, Berlin-New York, 1978.

[15] P. Le Boudec, Linear growth for certain elliptic fibrations, Int. Math. Res. Not. 2015 (21) (2015) 10859–10871.

[16] B. Matschke, A.S. Mudigonda, Quadratic fields admitting elliptic curves with rational j-invariant and good reduction everywhere, J. Number Theory 247 (2023) 162–210.

[17] L.J. Mordell, On the integer solutions of the equation $ey^2 = ax^3 + bx^2 + cx + d$, Proc. Lond. Math. Soc. (2) 21 (1923) 415–419.

[18] L.J. Mordell, Diophantine Equations, Pure and Applied Mathematics, vol. 30, Academic Press, London-New York, 1969.

[19] P. Shiu, A Brun-Titchmarsh theorem for multiplicative functions, J. Reine Angew. Math. 313 (1980) 161–170.

[20] J.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.

[21] J.H. Silverman, The Arithmetic of Elliptic Curves, second edition, Graduate Texts in Mathematics, vol. 106, Springer, Dordrecht, 2009.