DISTRIBUTION OF SELMER GROUPS OF QUADRATIC TWISTS OF A FAMILY OF ELLIPTIC CURVES

MAOSHENG XIONG AND ALEXANDRU ZAHARESCU*

Abstract. We study the distribution of the size of the Selmer groups arising from a 2-isogeny and its dual 2-isogeny for quadratic twists of elliptic curves with full 2-torsion points in $\mathbb{Q}$. We show that one of these Selmer groups is almost always bounded, while the 2-rank of the other follows a Gaussian distribution. This provides us with a small Tate-Shafarevich group and a large Tate-Shararevich group. When combined with a result obtained by Yu ([32]), this shows that the mean value of the 2-rank of the large Tate-Shafarevich group for square-free positive integers n less than X is $\frac{1}{2}\log\log X + O(1)$, as $X \to \infty$.

1. Introduction

In [28] the authors have described the asymptotic behavior of the size of the Selmer groups arising from three 2-isogenies and their dual 2-isogenies for the elliptic curve $E_n : y^2 = x^3 - n^2x$, which is closely related with the congruent number problem. In this paper we would like to see to what extent such results hold true in general for quadratic twists of elliptic curves with full 2-torsion points in $\mathbb{Q}$. Namely, for any $a, b \in \mathbb{Z}$ with $ab(a-b) \neq 0$, we shall consider the elliptic curve $E = E(a, b)$ defined by the equation

$$E : y^2 = x(x+a)(x+b).$$

2000 *Mathematics Subject Classification.* 11G05, 14H52, 11L40, 11N45.

Key words and phrases. Elliptic curves, Selmer group, Tate-Shafarevich group, Erdös-Kac Theorem.

*The second author was supported by NSF grant number DMS-0456615, and by CNCSIS grant GR106/2007, code 1116, of the Romanian Ministry of Education and Research.

For a square-free integer n, the quadratic twist E_n is given by

$$E_n : y^2 = x(x+an)(x+bn). \tag{1}$$

Corresponding to the 2-torsion point $(0,0)$ one has the 2-isogeny $\phi : E_n \longrightarrow E_n'$ where

$$E_n' : Y^2 = X^3 - 2(a+b)nX^2 + (a-b)^2n^2X$$

and the 2-isogeny ϕ is given by (see pp 74, [27])

$$\phi(x,y) = \left(\frac{y^2}{x^2}, \frac{y(abn^2 - x^2)}{x^2}\right).$$

Let $\hat{\phi} : E_n' \to E_n$ be the dual isogeny of ϕ. For $X > 0$ and coprime integers C and h denote the set

$$S(X,h,C) = \{1 \le n \le X : n \equiv h \pmod{C} \text{ and } n \text{ is square-free}\}. \tag{2}$$

We will investigate the asymptotic behavior of the size of the Selmer groups $Sel^{(\phi)}(E_n/Q)$ and $Sel^{(\hat{\phi})}(E_n'/Q)$ for $n \in S(X,h,C)$ as $X \to \infty$.

Theorem 1. *Let $a, b \in \mathbb{Z}$ with $ab(a-b) \neq 0$ and ab not a square. Define*

$$C_0 = \prod_{p \,|\, ab(a-b)} p\,,$$

and let h and C be coprime integers such that $C_0 \,|\, C$. For $X > 0$ and $n \in S(X,h,C)$ denote

$$\#Sel^{(\phi)}(E_n/Q) = 2^{s(n,\phi)}\,, \quad \#Sel^{(\hat{\phi})}(E_n'/Q) = 2^{s(n,\hat{\phi})}\,,$$

where E_n is the elliptic curve given by (1). Then $s(n,\phi) \le \omega(a-b)+1$ for almost all $n \in S(X,h,C)$ as $X \to \infty$, where ω is the function counting the number of distinct prime divisors, and $s(n,\hat{\phi})$ follows a Gaussian distribution. More precisely, for any $\gamma \in \mathbb{R}$,

$$\lim_{X\to\infty} \frac{1}{\#S(X,h,C)} \# \left\{ n \in S(X,h,C) : \frac{s(n,\hat{\phi}) - \frac{1}{2}\log\log n}{\sqrt{\frac{1}{2}\log\log n}} \le \gamma \right\} = G(\gamma)\ ,$$

where the function G is defined by

$$G(\gamma) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\gamma} e^{\frac{-t^2}{2}} \, \mathrm{d}t \ .$$

A more careful analysis shows that if one further assumes the following four conditions on a, b and h:

- $\gcd(a, b) = 1$,
- $a + b \geq 0$ or $ab < 0$,
- $a - b \equiv 1 \pmod 2$,
- If $p \mid (a - b)$, then $(\frac{-bh}{p}) = -1$,

then $s(n, \phi) = 0$ for almost all $n \in S(X, h, C)$, as $X \to \infty$ in Theorem 1.

Combining the above result with a result obtained by Yu (Theorem 2, [32]), one can obtain information on the corresponding Tate-Shafarevich groups. We first describe several conditions, say (Ca), (Cb) and (Cc') as follows (where p denotes an odd prime):

$$\begin{aligned}
&(Ca): \text{ If } p|a \text{ and } \mathrm{ord}_p(a) \text{ is even, then } \left(\tfrac{bh}{p}\right) = -1. \\
&(Cb): \text{ If } p|b \text{ and } \mathrm{ord}_p(b) \text{ is even, then } \left(\tfrac{ah}{p}\right) = -1. \\
&(Cc'): \text{ If } p|(a-b), \text{ then } \left(\tfrac{-bh}{p}\right) = -1.
\end{aligned}$$

We remark that the above conditions for a, b and h are in Yu's paper, except for (Cc'), which is slightly stronger than the original condition (Cc) from his paper.

Theorem 2. *For $a, b \in \mathbb{Z}$ such that $a - b \equiv 1 \pmod 2, a > 0, b > 0, \gcd(a, b) = 1$, and ab is not a square, let D be the conductor of $E : y^2 = x(x + a)(x + b)$. Fix an integer h such that $\gcd(h, D) = 1$ and that a, b, h satisfy the conditions (Ca), (Cb) and (Cc'). For $X > 0$ and $n \in S(X, h, D)$, let E_n be the elliptic curve defined by (1), and denote*

$$\#Ш(E_n/\mathbb{Q})[\phi] = 2^{t(n,\phi)}, \quad \#Ш(E'_n/\mathbb{Q})[\hat{\phi}] = 2^{t(n,\hat{\phi})}.$$

Then $t(n,\phi)=0$ *for almost all* $n\in S(X,h,D)$*, as* $X\to\infty$*. Moreover, for any integer* $k>0$*, one has*

$$\sum_{n\in S(X,h,D)} t(n,\hat{\phi})^k = \#S(X,h,D)\left(\frac{\log\log X}{2}\right)^k + O_k\left(X\left(\log\log X\right)^{k-1}\right).$$

In particular by taking $k=1$, we see that the mean value of the 2-rank of the large Tate-Shafarevich groups is $\frac{1}{2}\log\log X+O(1)$. Since $Ш(E_n'/\mathbb{Q})[\hat{\phi}]\subset Ш(E_n'/\mathbb{Q})[2]$, Theorem 2 shows that the 2-part of the Tate-Shafarevich group $Ш(E_n'/\mathbb{Q})$ can be arbitrarily large.

There are three main ingredients in the proofs of the above results. First, we employ Heath-Brown's method based on character sums to obtain asymptotic formulas on the size of the Selmer groups. Second, we use a graphical method, which plays an essential role in isolating the main contribution and reducing the complexity of the problem. Third, by combining our results with a result obtained by Yu ([32]) we obtain information on the corresponding Tate-Shafarevich groups.

Acknowledgement The authors want to express their gratitude to the referee for the careful reading and many useful suggestions.

2. Preliminaries

2.1. **Selmer groups and Tate-Shafarevich groups.** In this section we recall the formulation of Selmer groups and Tate-Shafarevich groups. Proofs can be found in Silverman's book ([27]). Let $\phi: E\longrightarrow E'$ be an isogeny between two elliptic curves E and E' over $\mathbb{Q}$. For the cases of interest to us, ϕ is defined over $\mathbb{Q}$ and $E[\phi]$, the kernel of ϕ consists of $\mathbb{Q}$-rational points. Via Galois cohomology, the short exact sequence of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-modules

$$0\longrightarrow E[\phi]\longrightarrow E(\overline{\mathbb{Q}})\xrightarrow{\phi} E'(\overline{\mathbb{Q}})\longrightarrow 0$$

yields the commutative diagrams (For details, the reader is referred to chapter X in [27])

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \frac{E'(\mathbb{Q})}{\phi(E(\mathbb{Q}))} & \longrightarrow & H^1(G_{\bar{\mathbb{Q}}/\mathbb{Q}}, E[\phi]) & \longrightarrow & H^1(G_{\bar{\mathbb{Q}}/\mathbb{Q}}, E)[\phi] & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow \pi_1 & & \downarrow \pi_2 & & \\
0 & \longrightarrow & 0 & \longrightarrow & \prod_v H^1(G_v, E) & \longrightarrow & \prod_v H^1(G_v, E) & \longrightarrow & 0,
\end{array}$$

where the homomorphisms π_1, π_2 are defined naturally by local consideration. The kernel of π_1 is the $\phi-$Selmer group $Sel^{(\phi)}(E/\mathbb{Q})$ and the kernel of π_2 (without the restriction $[\phi]$) is the Tate-Shafarevich group $Ш(E/\mathbb{Q})$. Here $Ш(E/\mathbb{Q})[\phi]$ is the ϕ-kernel of $Ш(E/\mathbb{Q})$. By the snake lemma one obtains the short exact sequence

$$0 \longrightarrow \frac{E'(\mathbb{Q})}{\phi(E(\mathbb{Q}))} \longrightarrow Sel^{(\phi)}(E/\mathbb{Q}) \longrightarrow Ш(E/\mathbb{Q})[\phi] \longrightarrow 0.$$

The group $\frac{E'(\mathbb{Q})}{\phi(E(\mathbb{Q}))}$ is directly related with the rank of the elliptic curve over $\mathbb{Q}$, which is difficult to compute in general. The Tate-Shafarevich group $Ш(E/\mathbb{Q})$ is also very mysterious. It appears naturally in the Birch and Swinnerton-Dyer conjecture, and measures the degree of deviation from the Hasse principle. Even the finiteness of the group is not known in general. Various families of elliptic curves with large Tate-Shafarevich groups were identified by a number of authors (see Aoki [2], Atake [3], Bölling [4], Cassels [5], Kloosterman [22], Kramer [23], Lemmermeyer [24],[25]). Moments and heuristic results were considered by Delaunay ([7],[6]). Effective bounds on the size of the Tate-Shafarevitch groups were obtained by Goldfeld and Szpiro ([15]), Goldfeld and Lieman ([14]). By contrast, the Selmer group $Sel^{(\phi)}(E/\mathbb{Q})$ is a local object and is relatively easy to handle in principle. By computing the Selmer group, one can obtain information on the rank of the elliptic curve and the Tate-Shafarevich group.

2.2. **2-descent and Selmer groups.** The 2-descent method is explained in the last chapter of Silverman's book ([27]) in general. For our particular case of E_n in (1), this can be specified as follows (see also [3]).

For a square-free positive integer n, define a finite set S of prime divisors of the rational number field $\mathbb{Q}$ by

$$S = \{\infty\} \bigcup \{p : p|ab(a-b)n\}.$$

Let M be the multiplicative subgroup of $\mathbb{Q}^*/(\mathbb{Q}^*)^2$ generated by -1 and the prime divisors of $(a-b)n$, and let M' be the multiplicative subgroup of $\mathbb{Q}^*/(\mathbb{Q}^*)^2$ generated by -1 and the prime divisors of abn. For each $d \in M$ ($d' \in M'$) we have homogeneous spaces C_d (respectively $C'_{d'}$) defined by

$$C_d : dw^2 = t^4 - 2(a+b)\frac{n}{d}t^2z^2 + (a-b)^2\frac{n^2}{d^2}z^4\,,$$

$$C'_{d'} : dw^2 = t^4 + 4(a+b)\frac{n}{d}t^2z^2 + 16ab\frac{n^2}{d^2}z^4\,.$$

The Selmer group $Sel^{(\phi)}(E_n/Q)$ (respectively $Sel^{(\hat{\phi})}(E'_n/Q)$) measures the possibility of C_d ($C'_{d'}$) having non-trivial solutions in the local field $\mathbb{Q}_v$ for all $v \in S$. Namely,

$$Sel^{(\phi)}(E_n/Q) \cong \{d \in M : C_d(\mathbb{Q}_v) \neq \emptyset \text{ for all } v \in S\}\ ,$$

$$Sel^{(\hat{\phi})}(E'_n/Q) \cong \{d' \in M' : C'_{d'}(\mathbb{Q}_v) \neq \emptyset \text{ for all } v \in S\}\,,$$

where $C_d(\mathbb{Q}_v) \neq \emptyset$ ($C'_{d'}(\mathbb{Q}_v) \neq \emptyset$) means that the homogeneous space C_d ($C'_{d'}$) has non-trivial solutions $(w,t,z) \neq (0,0,0)$ in $\mathbb{Q}_v$.

For the rank of the elliptic curve E_n we obtain the formula (see pp 286, [3])

$$\begin{aligned}\operatorname{rank}(E_n(\mathbb{Q})) &= \dim_{\mathbb{F}_2} Sel^{(\phi)}(E_n/\mathbb{Q}) - \dim_{\mathbb{F}_2} \text{Ш}(E_n/Q)[\phi] \\ &+ \dim_{\mathbb{F}_2} Sel^{(\hat{\phi})}(E'_n/\mathbb{Q}) - \dim_{\mathbb{F}_2} \text{Ш}(E'_n/Q)[\hat{\phi}] - 2.\end{aligned}$$

Thus we can calculate the rank from the dimensions of the Selmer groups and the Tate-Shafarevich groups.

2.3. **A graphical method.** We use standard terminology in graph theory ([18]). Let $G = (V, A)$ be a simple directed graph where $V = V(G) = \{v_1, \cdots, v_m\}$ is the set of vertices of G, and $A = A(G)$ is the set of arcs in G. We denote an arc $(v_i, v_j) \in A$ by $\overrightarrow{v_i v_j}$. The adjacency matrix of G is defined by

$$M(G) = (a_{ij})_{1 \leqslant i,j \leqslant m} \ ,$$

where

$$a_{ij} = \begin{cases} 1, & \text{if } \overrightarrow{v_i v_j} \in A \quad (1 \leqslant i \neq j \leqslant m) \\ 0, & \qquad\qquad \text{otherwise} \ . \end{cases}$$

For the vertex v_i, $1 \leq i \leq m$, let $d_i = \sum_{j=1}^{m} a_{ij}$. The Laplace matrix of the graph G is defined by

$$L(G) = \mathrm{diag}(d_1, \cdots, d_m) - M(G) \, .$$

The term "odd graph" has been used by Feng, Xue and one of the authors in their study of new families of non-congruent numbers ([11],[12],[13]). It is also used by Faulkner and James to compute the size of the Selmer groups ([10]).

Definition 1. *Let $G = (V, A)$ be a directed graph. A partition of vertices $V_1 \bigcup V_2 = V$ is called odd if either there exists a vertex $v_1 \in V_1$ such that $\#\{v_1 \to V_2\}$, the total number of arcs from v_1 to vertices in V_2 is odd, or there exists $v_2 \in V_2$ such that $\#\{v_2 \to V_1\}$ is odd. Otherwise the partition $V_1 \bigcup V_2 = V$ is called even. The graph G is called odd if all non-trivial partitions $\{V_1, V_2\} \neq \{V, \emptyset\}$ of V are odd.*

We need the following counting lemma, which can be derived by the same idea used in the proof of Lemma 2.2 in [11].

Lemma 1. *Let $G = (V, A)$ be a directed graph, $V = \{v_1, \ldots, v_{s+t}\}$ $(s, t \geq 0)$. Then the number of even partition $\{V_1, V_2\}$ of V such that $\{v_{s+1}, \ldots, v_{s+t}\} \subset V_2$ is equal to the number of vectors $(x_1, \ldots, x_s) \in \mathbb{F}_2^s$ such that $L(G) \cdot (x_1, \ldots, x_s, 0, \ldots, 0)^T = \mathbf{0}$.*

2.4. **Generalized Erdös-Kac Theorem.** For a positive integer n, let $\omega(n)$ be the number of distinct prime divisors of n. The remarkable theorem of Erdős and Kac ([9]) is that, for any $\gamma \in \mathbb{R}$,

$$\lim_{X\to\infty} \frac{1}{X}\#\left\{n : 1 \le n \le X, \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \le \gamma\right\} = G(\gamma) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\gamma} e^{-\frac{t^2}{2}}\,\mathrm{d}t\,.$$

There is very rich literature on various aspects of the Erdős-Kac theorem. Interested readers can refer to Granville and Soundararajan's paper [16] for the most recent account and Elliot's monograph [8] for a comprehensive treatment of the subject.

We will use the following generalization of Erdös-Kac Theorem obtained by Liu ([26]). For completeness we reproduce the statement here. Let S be an infinite subset of $\mathbb{N}$. For $X \in \mathbb{R}, X > 1$, define

$$S(X) = \{n \le X : n \in S\}.$$

We assume that S satisfies the cardinality condition

$$\left|S(X^{1/2})\right| = o\left(|S(X)|\right), \tag{3}$$

where $|S(X)|$ is the cardinality of $S(X)$. Let $f : S \longrightarrow \mathbb{N}$ be a map. For each prime l, write

$$\frac{1}{|S(X)|}\#\left\{n \in S(X) : f(n) \text{ is divisible by } l\right\} = \lambda_l(X) + e_l(X),$$

and for any u-tuples of distinct primes $(l_1, l_2, \ldots, l_u)$, write

$$\frac{1}{|S(X)|}\#\left\{n \in S(X) : f(n) \text{ is divisible by } l_1 l_2 \cdots l_u\right\} = \prod_{i=1}^{u} \lambda_{l_i}(X) + e_{l_1 l_2 \cdots l_u}(X).$$

We will use abbreviated notations λ_l, e_l and $e_{l_1 l_2 \cdots l_u}$ below.

Suppose there exist absolute constants β and c with $0 < \beta \le 1$ and $c > 0$, and a function $Y = Y(X) < X^{\beta}$ such that the following conditions hold:

(i) For each $n \in S(X)$, the number of distinct prime divisors l of $f(n)$ with $l > X^{\beta}$ is bounded uniformly.

(ii) $\sum_{Y<l\le X^\beta}\lambda_l = o((\log\log X)^{1/2})$, where the sum is over primes l.

(iii) $\sum_{Y<l\le X^\beta}|e_l| = o((\log\log X)^{1/2})$.

(iv) $\sum_{l\le Y}\lambda_l = c\log\log X + o((\log\log X)^{1/2})$.

(v) $\sum_{l\le Y}\lambda_l^2 = o((\log\log X)^{1/2})$.

(vi) For $r\in\mathbb{N}$, let $u=1,2,\dots,r$. We have

$$\sum{}''|e_{l_1\cdots l_u}| = o((\log\log X)^{-r/2}),$$

where $\sum''$ extends over all u-tuples of distinct primes $(l_1,l_2,\dots,l_u)$ with $l_i\le Y$. (Notice that the condition (4) in Liu's paper [26] is actually $c=1$. However there is no essential difference by introducing the constant $c>0$ here.)

Lemma 2. (Theorem 3, [26]) *Let S be an infinite subset of $\mathbb{N}$ satisfying condition (3) and $f:S\to\mathbb{N}$. Suppose there exist absolute constants β, c with $0<\beta\le 1$, $c>0$ and $Y=y(X)<X^\beta$ such that the conditions (i)–(vi) hold. Then for $\gamma\in\mathbb{R}$, we have*

$$\lim_{X\to\infty}\frac{1}{|S(X)|}\#\left\{n\in S(X):\frac{\omega(f(n))-c\log\log n}{\sqrt{c\log\log n}}\le\gamma\right\}=G(\gamma).$$

2.5. **Additional lemmas.** The following results proved by Heath-Brown ([19]) and generalized by Yu will be used several times in our proofs.

Lemma 3. (Lemma 2.2 in [31], Lemma 4.1 in [29]) *Suppose $\epsilon>0$ is any fixed number, X,M and N are sufficiently large real numbers, and $\{a_m\},\{b_n\}$ are two complex sequences, supported on odd integers, satisfying $|a_m|,|b_n|\le 1$. Fix positive integers h, q satisfying $\gcd(h,q)=1$ and $q\le\{\min(M,N)\}^{\epsilon/3}$. Let*

$$S:=\sum_{m,n}a_mb_n\left(\frac{m}{n}\right),$$

where the summation is subject to

$$M\le m<2M, N\le n<2N, mn\le X \text{ and } mn\equiv h \pmod q.$$

Then we have

$$S \ll MN^{15/16+\epsilon} + M^{15/16+\epsilon}N\,,$$

where the constant involved in the $\ll$ symbol depends on ϵ only.

Lemma 4. (Lemma 4.2, [29]) *Suppose s is a fixed rational number. Let N be sufficiently large. Then for arbitrary positive integers q, r and any nonprincipal character $\chi \pmod q$, we have*

$$\sum_{n \le X, \gcd(n,r)=1} \mu^2(n) s^{\omega(n)} \chi(n) \ll X\tau(r)\exp(-\eta\sqrt{\log X})$$

with a positive constant $\eta = \eta_{s,N}$, uniformly for $q \le \log^N X$. Here τ is the usual divisor function and μ is the Möbius function.

Lemma 5. (Lemma 2.4, [31]) *Let s and C be two positive integers, and $A > 0$ be any fixed number. For $X > 1$, let $T \le \exp(\sqrt{\log X})$ and $M, N \ge T$ be given. There exists some constant $\eta > 0$ such that, for any positive integer r, any integer h prime to C, and any distinct characters $\chi_1, \chi_2 \pmod q$, where $q \ll (\log X)^A$, we have*

$$\sum_{m,n} \mu^2(m)\mu^2(n) s^{-\omega(m)-\omega(n)} \chi_1(m)\chi_2(n) \ll X\tau(r)\exp\left(-\eta\sqrt{\log T}\right)\log X,$$

where the sum is over coprime variables satisfying the conditions

$$M < m \le 2M,\ N < n \le 2N,\ mn \le X,\ mn \equiv h \pmod C,\ \gcd(mn, r) = 1,$$

and the constant involved in the $\ll$-symbol depends on s and C only.

3. Solvability conditions of homogeneous spaces

The problem of finding the size of the Selmer groups $Sel^{(\phi)}(E_n/\mathbb{Q})$ $(Sel^{(\hat{\phi})}(E'_n/\mathbb{Q}))$ is equivalent to the problem of determining how many homogeneous spaces C_d (respectively $C'_{d'}$) have non-trivial solutions over certain local fields. We collect solvability conditions for C_d and $C'_{d'}$ in the following two lemmas.

Lemma 6. *Let $a, b \in \mathbb{Z}$ with $ab(a-b) \neq 0$ and $\gcd(a,b)=1$. Let n be a square-free integer with $\gcd(n, ab(a-b)) = 1$, and $M \subseteq \mathbb{Q}^*/\mathbb{Q}^{*2}$, the multiplicative subgroup generated by -1 and the prime divisors of $(a-b)n$. Let p denote an odd prime number. For any $d \in M$, one has:*

(i) For $p|n, p|d$: $\left(\dfrac{ab}{p}\right) = 1$ and $\left(\dfrac{an/d}{p}\right) = 1 \iff C_d(\mathbb{Q}_p) \neq \emptyset$.

(ii) For $p|n, p \nmid d$: $\left(\dfrac{d}{p}\right) = 1 \iff C_d(\mathbb{Q}_p) \neq \emptyset$.

(iii) For $p|(a-b), p|d$: $\left(\dfrac{-bn}{p}\right) = 1 \iff C_d(\mathbb{Q}_p) \neq \emptyset$.

(iv) Suppose $a+b \geq 0$ or $ab < 0$. If $d < 0$, then $C_d(\mathbb{R}) = \emptyset$.

Proof. (i) Let $p|n$ and $p|d$. Suppose (w,t,z) is a non-trivial solution of C_d : $dw^2 = t^4 - 2(a+b)\frac{n}{d}t^2z^2 + (a-b)^2\frac{n^2}{d^2}z^4$ over Q_p. Then (p^2w, pt, pz) is also a non-trivial solution. We may assume that $0 \leq \min\{v_p(w), v_p(t), v_p(z)\} \leq 1$, and also, if $v_p(w) \geq 2$, then $\min\{v_p(t), v_p(z)\} = 0$, where v_p is the p-adic exponential valuation, normalized by $v_p(p) = 1$. From the above equation one knows that the minimum of the four values

$$1 + 2v_p(w), \quad 4v_p(t), \quad v_p(a+b) + 2v_p(t) + 2v_p(z), \quad 4v_p(z),$$

is attained for at least two of them. Therefore $v_p(t) = v_p(z) = 0 < 1 + 2v_p(w)$. One has

$$t^4 - 2(a+b)\frac{n}{d}t^2z^2 + (a-b)^2\frac{n^2}{d^2}z^4 \equiv 0 \pmod{p}.$$

This implies that

$$\left(u^2 - (a+b)\frac{n}{d}\right)^2 \equiv 4ab\frac{n^2}{d^2} \pmod{p} \tag{4}$$

where $u = t/z \in \mathbb{Z}_p^*$, and one must have

$$\left(\frac{ab}{p}\right) = 1.$$

Let $\sqrt{ab} \in \mathbb{Z}_p^*$ be one of the square roots of $ab \in \mathbb{Z}_p$. The equation (4) implies that

$$u^2 \equiv \left((a+b) \pm 2\sqrt{ab}\right)\frac{n}{d} \pmod p.$$

Hence

$$\left(\frac{\left((a+b)+2\sqrt{ab}\right)\frac{n}{d}}{p}\right) = 1 \text{ or } \left(\frac{\left((a+b)-2\sqrt{ab}\right)\frac{n}{d}}{p}\right) = 1.$$

Since $\left(\frac{ab}{p}\right) = 1$, the above condition is equivalent to

$$\left(\frac{an/d}{p}\right) = 1.$$

On the other hand, if $(\frac{ab}{p}) = 1$ and $(\frac{an/d}{p}) = 1$, from the above argument, the equation

$$t^4 - 2(a+b)\frac{n}{d}t^2 + (a-b)^2\frac{n^2}{d^2} = 0$$

is solvable for $t \in (\mathbb{Z}/p\mathbb{Z})^*$. By Hensel's lemma this leads to a non-trivial solution $(0, t, 1)$ of C_d over $\mathbb{Q}_p$.

(ii) Let $p|n, p \nmid d$. If $\left(\frac{d}{p}\right) = 1$, it is easy to see that one has a solution $(w, 1, 0) \in \mathbb{Z}_p^3$ for C_d. On the other hand, suppose (w, t, z) is a non-trivial solution of C_d over Q_p with $0 \le \min\{v_p(w), v_p(t), v_p(z)\} \le 1$ and $v_p(w) \ge 2 \implies \min\{v_p(t), v_p(z)\} = 0$. The minimum of the four values

$$2v_p(w), \quad 4v_p(t), \quad 1 + v_p(a+b) + 2v_p(t) + 2v_p(z), \quad 2 + 4v_p(z),$$

is attainted for at least two of them. Either $v_p(w) = v_p(t) = 0 \le v_p(z)$, in which case one has

$$dw^2 \equiv t^4 \pmod p,$$

and this implies that

$$\left(\frac{d}{p}\right) = 1,$$

or $v_p(w) = 1, v_p(z) = 0$ and $v_p(t) \ge 1$, in which case one has $d\frac{w^2}{p^2} \equiv (a-b)^2\frac{n^2}{p^2d^2}z^4 \pmod p$, and one still gets the condition $\left(\frac{d}{p}\right) = 1$.

(iii) Let $p|(a-b), p|d$. Suppose (w,t,z) is a non-trivial solution of C_d over Q_p with $0 \leq \min\{v_p(w), v_p(t), v_p(z)\} \leq 1$ and $v_p(w) \geq 2 \Longrightarrow \min\{v_p(t), v_p(z)\} = 0$. The minimum of the four values

$$1+2v_p(w), \quad 4v_p(t), \quad -1+2v_p(t)+2v_p(z), \quad 2v_p(a-b)-2+4v_p(z),$$

is attained for at least two of them. Since $v_p(a-b) \geq 1$, one obtains

$$1+2v_p(w) = -1+2v_p(t)+2v_p(z) = \min .$$

Dividing a suitable power of p on both sides and then taking the equation modulo p, one has

$$\frac{d}{p}w^2 \equiv -2(a+b)\frac{np}{d}z^2 \pmod{p},$$

for some $w, z \in (\mathbb{Z}/p\mathbb{Z})^*$. Therefore

$$\left(\frac{-2(a+b)n}{p}\right) = \left(\frac{-bn}{p}\right) = 1.$$

On the other hand if $(\frac{-bn}{p}) = 1$, one can see that $C_d(\mathbb{Q}_p) \neq \emptyset$ by using Hensel's lemma.

(iv) The proof is clear. This completes the proof of Lemma 6.

Lemma 7. *Let $a, b \in \mathbb{Z}$ with $ab(a-b) \neq 0$ and $\gcd(a,b) = 1$. Let n be a square-free integer with $\gcd(n, ab(a-b)) = 1$, and $N \subseteq \mathbb{Q}^*/\mathbb{Q}^{*2}$ the multiplicative subgroup generated by the prime divisors of n. Let p denote an odd prime number. For any $d \in N$, one has:*

(i) For $p|n, p|d$: $\left(\frac{ab}{p}\right) = 1$ and $\left(\frac{-an/d}{p}\right) = -1 \Longleftrightarrow C'_d(\mathbb{Q}_p) = \emptyset$.

(ii) For $p|n, p \nmid d$: $\left(\frac{ab}{p}\right) = 1$ and $\left(\frac{d}{p}\right) = -1 \Longleftrightarrow C'_d(\mathbb{Q}_p) = \emptyset$.

(iii) For $p|(a-b), p \nmid d$: $\left(\frac{-bn}{p}\right) = 1$ and $\left(\frac{d}{p}\right) = -1 \Longleftrightarrow C'_d(\mathbb{Q}_p) = \emptyset$.

(iv) For $p|ab, p \nmid d$: $C'_d(Q_p) \neq \emptyset$.

(v) $C'_d(\mathbb{R}) \neq \emptyset$. Moreover $d \equiv 1 \pmod{8}$, then $C'_d(\mathbb{Q}_2) \neq \emptyset$.

Proof. The proof of Lemma 7 is similar to that of Lemma 6 and will be left to the reader.

4. Averaging the size of Selmer groups $Sel^{(\phi)}(E_n/\mathbb{Q})$

The main purpose of this section is to prove the following lemma.

Lemma 8. *Let $a, b \in \mathbb{Z}$ with $ab(a-b) \neq 0, \gcd(a,b) = 1$ and ab not a square. Define*

$$C_0 = \prod_{p|ab(a-b)} p,$$

and let h and C be coprime integers such that $C_0|C$. For $X > 0$, let $S(X,h,C)$ be the set defined in (2) and for $n \in S(X,h,C)$, consider the elliptic curve E_n given by (1). Let $N \subseteq \mathbb{Q}^*/\mathbb{Q}^{*2}$ be the multiplicative subgroup generated by the prime divisors of n and denote

$$\#\left(Sel^{(\phi)}(E_n/Q)\bigcap N\right) = 2^{\hat{s}(n,\phi)}.$$

Then

$$\hat{s}(n,\phi) = 0$$

for almost all $n \in S(X,h,C)$ as $X \to \infty$.

Once Lemma 8 is proved, from the definition of the Selmer group $Sel^{(\phi)}(E_n/\mathbb{Q})$ we will have

$$0 \leq s(n,\phi) \leq \hat{s}(n,\phi) + \omega(a-b) + 1,$$

which implies that

$$s(n,\phi) \leq \omega(a-b) + 1 \tag{5}$$

for almost all $n \in S(X,h,C)$, as $X \to \infty$.

If $c = \gcd(a,b) > 1$, one may consider the elliptic curve

$$E_n : y^2 = x(x + a'n')(x + b'n'),$$

where $a = a'c, b = b'c, n' = nc$, and one can see that the inequality (5) still holds true for almost all $n \in S(X, h, C)$, as $X \to \infty$. This completes the proof of the first part of Theorem 1.

A more careful analysis of Lemma 6 shows that, if one assumes further the following three conditions on a and b in Lemma 8:

- $a + b \geq 0$ or $ab < 0$,
- $a - b \equiv 1 \pmod 2$,
- If $p \mid (a - b)$, then $(\frac{-bh}{p}) = -1$,

then $s(n, \phi) = \hat{s}(n, \phi)$, in which case $s(n, \phi) = 0$ for almost all $n \in S(X, h, C)$ as $X \to \infty$.

The proof of Lemma 8 is similar to that in [28], where the key idea which is based on character sums was initiated by Heath-Brown ([19], [20]) to study the size of the 2-Selmer groups of elliptic curves related with the congruent number problem. His method has been generalized by Yu ([29],[30],[31],[32]) to study the size of Selmer groups for other families of elliptic curves. Since we will treat several similar sums later in this paper, for the sake of completeness we present a proof of Lemma 8 below.

Proof of Lemma 8. By Lemma 6, one has

$$2^{\hat{s}(n,\phi)} \leq \sum_{n=dd'} \prod_{p|d} \frac{1}{4} \left(\left(\frac{ab}{p} \right) + 1 \right) \left(\left(\frac{ad'}{p} \right) + 1 \right) \prod_{p|d'} \frac{1}{2} \left(\left(\frac{d}{p} \right) + 1 \right).$$

Expanding the product on the right hand side one has

$$\begin{aligned} 2^{\hat{s}(n,\phi)} &\leq \sum_{n=D_0D_1D_2D_3D_4D_5} 4^{-\omega(D_0D_1D_2D_3)} 2^{-\omega(D_4D_5)} \left(\frac{b}{D_1D_2} \right) \left(\frac{a}{D_2D_3} \right) \\ &\quad \times \left(\frac{D_4}{D_1} \right) \left(\frac{D_1}{D_4} \right) \left(\frac{D_4}{D_3} \right) \left(\frac{D_3}{D_4} \right) \left(\frac{D_5}{D_1} \right) \left(\frac{D_5}{D_3} \right) \left(\frac{D_0}{D_4} \right) \left(\frac{D_2}{D_4} \right) \\ &= \sum_{\mathbf{D}} g(\mathbf{D}), \end{aligned}$$

where $\mathbf{D} = (D_0, D_1, D_2, D_3, D_4, D_5)$ is subject to the condition that $n = D_0D_1D_2D_3D_4D_5$. Since n is square-free, all the D_i's are square-free and pairwise coprime. Our goal is to estimate

$$\sum_{n\in S(X,h)} \sum_{\mathbf{D}} g(\mathbf{D})\,.$$

We sum over the six variables D_i, subject to the conditions that each D_i is square-free, that they are pairwise coprime, and that their product n satisfies

$$n \le X,\ n \equiv h \pmod{C}\,.$$

We divide the range of each variable D_i into dyadic intervals $[A_i, 2A_i)$, where A_i runs over powers of 2. There are $O(\log^6 X)$ many non-empty subsums, which we shall write in the form $S(\mathbf{A})$, where $\mathbf{A} = (A_0, A_1, A_2, A_3, A_4, A_5)$. Here we may assume that

$$1 \le \prod_{i=1}^{5} A_i \ll X.$$

Following Heath-Brown ([19], [20]), we shall describe the variables D_i and D_j as being "linked" if exactly one of the Jacobi symbols

$$\left(\frac{D_i}{D_j}\right), \quad \left(\frac{D_j}{D_i}\right)$$

occurs in the expression for $g(\mathbf{D})$. One sees that $(D_1, D_5), (D_3, D_5), (D_0, D_4)$ and (D_2, D_4) are the pairs of linked variables.

4.1. **Case one.** Consider the linked variables D_1, D_5. Suppose $A_1, A_5 \ge (\log X)^{224}$. We may write $g(\mathbf{D})$ in the form

$$g(\mathbf{D}) = \left(\frac{D_5}{D_1}\right) a(D_5)b(D_1),$$

where the function $a(D_5)$ depends on all other variables D_i except D_1, and $b(D_1)$ depends on all other variables D_i except D_5. Moreover we have

$$|a(D_5)|, |b(D_1)| \le 1.$$

We can now write

$$|S(\mathbf{A})| = \sum_{D_0,D_2,D_3,D_4} \left| \sum_{D_1,D_5} \left(\frac{D_5}{D_1}\right) a(D_5)b(D_1) \right|.$$

As a consequence of Lemma 3 one finds that

$$S(\mathbf{A}) \ll A_0A_2A_3A_4A_1A_5 \left\{\min(A_1, A_5)\right\}^{-1/32} \ll X(\log X)^{-7}.$$

Similar results hold for other linked variables. Therefore

Lemma 9. *We have*

$$S(\mathbf{A}) \ll X(\log X)^{-7}$$

whenever there is a pair of linked variables with $A_i, A_j \geq (\log X)^{224}$.

4.2. **Case two.** We now examine the case when $A_1 \geq (\log X)^{224}$ while $A_5 < (\log X)^{224}$. Using quadratic reciprocity we put $g(\mathbf{D})$ in the form

$$g(\mathbf{D}) = 4^{-\omega(D_1)} \left(\frac{D_1}{D_5}\right) \chi(D_1)c\,,$$

where χ is a character with modulus dividing $8b$. χ may depend on the variables D_i other than D_1, and the factor c is independent of D_1 and satisfies $|c| \leq 1$. It follows that

$$|S(\mathbf{A})| \leq \sum_{D_0,D_2,D_3,D_4,D_5} \left| \sum_{D_1} 4^{-\omega(D_1)} \left(\frac{D_1}{D_5}\right) \chi(D_1) \right|, \tag{6}$$

where the inner sum is restricted by the conditions that D_1 must be square-free and coprime to all the other variables D_0, D_2, D_3, D_4, D_5.

We remove the condition $D_1 \equiv h' \pmod{C}$ from the inner sum on the right side of (6) and insert instead a factor

$$\frac{1}{\phi(C)} \sum_{\psi \pmod{C}} \psi(D_1)\overline{\psi(h')}.$$

Employing Lemma 4 one has

$$\begin{aligned} S(\mathbf{A}) &\ll A_1 \exp(-\eta\sqrt{\log A_1}) \sum_{D_0,D_2,D_3,D_4,D_5} \tau(D_0D_2D_3D_4D_5) \\ &\ll A_1 \exp(-\eta\sqrt{\log A_1}) \prod_{D_i, i\neq 1} \sum_{D_i} \tau(D_i) \\ &\ll X(\log X)^5 \exp(-\eta\sqrt{\log A_1}), \end{aligned}$$

provided that $D_5 \neq 1$ and D_5 times the modulus of χ is $\ll \log^N A_1$ for some $N > 0$. We summarize the above results as follows.

Lemma 10. *For any constant κ with $0 < \kappa < 1$ one has*

$$S(\mathbf{A}) \ll X(\log X)^{-7}$$

whenever there are linked variables D_i, D_j for which

$$A_i \geq \exp\{(\log X)^\kappa\}$$

and $D_j > 1$.

4.3. **Case Three.** For any $0 < \kappa < 1$ denote

$$C = \exp\{(\log X)^\kappa\}\ . \tag{7}$$

Let $\sum'$ indicate the condition that $A_0, A_1, A_2, A_3 \leq C$, $A_4 \leq C$ or $A_5 \leq C$. Then

$$\sum_{\mathbf{A}}{}' |S(\mathbf{A})| \leq 2 \sum_{D_i \leq 2C, 0\leq i\leq 4} 4^{-\omega(D_0)} \cdots 4^{-\omega(D_3)} 2^{-\omega(D_4)} \sum_{D_5 \leq \frac{X}{D_0\cdots D_4}} 2^{-\omega(D_5)}.$$

We now use the bounds ([17])

$$\sum_{n\leq X} \gamma^{\omega(n)} \ll X(\log X)^{\gamma-1},$$

and

$$\sum_{n\leq X} \frac{\gamma^{\omega(n)}}{n} \leq \prod_{p\leq X}\left(1+\frac{\gamma}{p}\right) \ll (\log X)^\gamma,$$

which are valid for any fixed $\gamma > 0$. Since

$$\frac{X}{D_0 \cdots D_4} \gg XC^{-5} \gg X^{1/2},$$

one has $\log(XC^{-5}) \gg \log X$. Therefore

$$\begin{aligned}
\sum_{\mathbf{A}}{}' |S(\mathbf{A})| &\ll \sum_{D_i \le 2C, 0 \le i \le 4} 4^{-\omega(D_0)} \cdots 4^{-\omega(D_3)} 2^{-\omega(D_4)} \frac{X}{D_0 \cdots D_4} (\log X)^{-1/2} \\
&\ll X(\log X)^{-1/2} \left(\sum_{n \le 2C} \frac{4^{-\omega(n)}}{n} \right)^4 \left(\sum_{n \le 2C} \frac{2^{-\omega(n)}}{n} \right) \\
&\ll X(\log X)^{-1/2} (\log 2C)^{\frac{1}{4} \cdot 4} (\log 2C)^{\frac{1}{2}} \ll X(\log X)^{-\frac{1}{2} + \kappa\frac{3}{2}}.
\end{aligned}$$

Let $\sum''$ indicate the condition that $A_4, A_5 \le C$ and at least one of A_0, A_1, A_2, A_3 is less than C. Then

$$\begin{aligned}
\sum_{\mathbf{A}}{}'' |S(\mathbf{A})| &\le \sum_{D_0 D_1 D_2 D_3 D_4 D_5 \le X} 4^{-\omega(D_0)} \cdots 4^{-\omega(D_3)} 2^{-\omega(D_4)} 2^{-\omega(D_5)} \\
&= \sum_{mn \le X} 4^{-\omega(m)} 2^{-\omega(n)} \left(\sum_{D_0 D_1 D_2 D_3 = m} 1 \right) \left(\sum_{D_4 D_5 = n} 1 \right) \\
&\le \sum_{n \le (2C)^2} 1 \sum_{m \le X/n} 4^{-\omega(m)} \sum_{D_0 D_1 D_2 D_3 = m} 1.
\end{aligned}$$

Write

$$m_1 = \prod_{D_i < 2C} D_i, \quad m_2 = \prod_{D_i \ge 2C} D_i,$$

so that $m_1 \le (2C)^4$. One has

$$\begin{aligned}
\sum_{\mathbf{A}}{}'' |S(\mathbf{A})| &\ll \sum_{n \le (2C)^2} 1 \sum_{m_1 \le (2C)^4} \sum_{m_2 \le \frac{X}{m_1 n}} \left(\frac{3}{4} \right)^{\omega(m_2)} \\
&\ll \sum_{n \le (2C)^2} 1 \sum_{m_1 \le (2C)^4} \frac{X}{m_1 n} (\log X)^{-1/4} \\
&\ll X(\log X)^{-1/4} (\log 2C)^2 \ll X(\log X)^{-\frac{1}{4} + 2\kappa}.
\end{aligned}$$

We summarize our results as follows.

Lemma 11. *Choosing $\kappa = \frac{1}{40}$, we have*

$$\sum_{\mathbf{A}} |S(\mathbf{A})| \ll X(\log X)^{-1/5}\,,$$

where the sum over $\mathbf{A}$ *runs over all sets in which either $A_0, A_1, A_2, A_3 \le C$ and at least one of A_4, A_5 is $\le C$, or $A_4, A_5 \le C$ and at least one of A_0, A_1, A_2, A_3 is $\le C$, or there are linked variables D_i and D_j with $D_i \ge C$ and $D_j > 1$.*

4.4. **The remaining cases.** The cases where the sums $S(\mathbf{A})$ are not handled by Lemma 11 are as follows.

(1) $A_4, A_5 \ge C \Longrightarrow D_0 = D_1 = D_2 = D_3 = 1$.
(2) $A_4 \ge C, A_5 < C \Longrightarrow D_0 = D_2 = D_5 = 1$, A_1 or $A_3 \ge C$.
(3) $A_4 < C, A_5 > C \Longrightarrow D_1 = D_3 = D_4 = 1$, A_0 or $A_2 \ge C$.
(4) $A_4, A_5 \le C \Longrightarrow A_0, A_1, A_2, A_3 \ge C$ and $D_4 = D_5 = 1$.

Case (1). With $D_0 = D_1 = D_2 = D_3 = 1$ the function $g(\mathbf{D})$ reduces to $2^{-\omega(D_4)}2^{-\omega(D_5)}$. The sum is

$$\sum_{D_4, D_5} 2^{-\omega(D_4)}2^{-\omega(D_5)},$$

where D_4, D_5 are subject to the conditions

$$D_4, D_5 > C, \quad n = D_4 D_5 \equiv h \pmod{C}, \quad n \text{ square-free}, \quad n \le X.$$

We can remove the condition $D_4, D_5 > C$ with an error

$$\begin{aligned} &\le\ 2\sum_{D_4 \le C} 2^{-\omega(D_4)} \sum_{D_5 \le \frac{X}{D_4}} 2^{-\omega(D_5)} \ll X(\log X)^{-1/2} \sum_{D_4 \le C} \frac{2^{-\omega(D_4)}}{D_4} \\ &\ll\ X(\log X)^{-\frac{1}{2}+\frac{1}{2}\kappa} \ll X(\log X)^{-1/5}. \end{aligned}$$

Since $n = D_4 D_5$ is square-free it factors as $D_4 D_5$ in exactly $2^{\omega(n)}$ different ways. We therefore obtain

$$\sum_{n \in S(X,h,C)} 1 + O\left(X(\log X)^{-1/5}\right) = \#S(X,h,C) + O\left(X(\log X)^{-1/5}\right).$$

Case (2). With $D_0 = D_2 = D_5 = 1$ the function $g(\mathbf{D})$ reduces to

$$f(\mathbf{D}) = 4^{-\omega(D_1D_3)}2^{-\omega(D_4)}\left(\frac{b}{D_1}\right)\left(\frac{a}{D_3}\right)\left(\frac{D_4}{D_1}\right)\left(\frac{D_1}{D_4}\right)\left(\frac{D_4}{D_3}\right)\left(\frac{D_3}{D_4}\right),$$

and the conditions for $\mathbf{A}$ are $A_4 \geq C$ and at least one of $A_1, A_3 \geq C$. If $A_1 \leq C$ or $A_3 \leq C$, using an argument similar to that in Case Three one sees that

$$S(\mathbf{A}) \ \ll \ X(\log X)^{-1/5}.$$

If $A_1, A_3, A_4 \geq C$, one can apply Lemma 5 to obtain that the total contribution of these sums is $O(X(\log X)^{-1})$. Case (3) and (4) can be treated in a similar way and the total contribution is $O(X(\log X)^{-1/5})$. Therefore we conclude that

$$\sum_{n\in S(X,h,C)} 2^{\hat{s}(n,\phi)} \leq \#S(X,h,C) + O\left(X(\log X)^{-1/5}\right)$$

as $X \to \infty$.

4.5. **Analyzing the result.** For any integer $r \geq 0$, let

$$a_r = \#\{n \in S(X,h,C) : \hat{s}(n,\phi) = r\}.$$

The above inequality is

$$\sum_{r\geq 0} 2^r a_r \ \leq \ \#S(X,h,C) + O\left(X(\log X)^{-1/5}\right),$$

hence

$$\sum_{r\geq 1} 2^{r-1} a_r \ \leq \ \sum_{r\geq 1}(2^r - 1)a_r = O\left(X(\log X)^{-1/5}\right).$$

One has

$$a_r = O\left(X(\log X)^{-1/5}2^{-r}\right), r \geq 1,$$

and

$$\sum_{r\geq 1} a_r = O\left(X(\log X)^{-1/5}\right).$$

Therefore

$$\hat{s}(n, \phi) = 0$$

for almost all $n \in S(X, h, C)$ as $X \to \infty$. The completes the proof of Lemma 8.

5. Averaging the size of Selmer groups $Sel^{(\hat{\phi})}(E'_n/\mathbb{Q})$

Let $h, C, S(X, h, C), s(n, \hat{\phi})$ be defined as in Theorem 1. For simplicity we first assume $\gcd(a, b) = 1$. For any $n \in S(X, h, C)$ consider the elliptic curve E_n defined in (1). Let $N \subseteq \mathbb{Q}^*/\mathbb{Q}^{*2}$ be the multiplicative subgroup generated by prime divisors of n, and denote

$$\#\left(Sel^{(\hat{\phi})}(E'_n/Q) \bigcap N\right) = 2^{\hat{s}(n,\hat{\phi})}.$$

From the definition of the Selmer group $Sel^{(\hat{\phi})}(E'_n/Q)$ one has

$$\hat{s}(n, \hat{\phi}) \le s(n, \hat{\phi}) \le \hat{s}(n, \hat{\phi}) + \omega(ab) + 1. \tag{8}$$

It is enough to study the asymptotic behavior of $\hat{s}(n, \hat{\phi})$ for $n \in S(X, h, C)$, as $X \to \infty$. For any $d \in N$, by Lemma 7, $C'_d(\mathbb{Q}_p) \neq \phi$ for any $p|ab$. Denote the cardinality of the set of $d \in N$ such that $C'_d(\mathbb{Q}_p) \neq \phi$ for any $p|n$ by $2^{s_1(n)}$, and the cardinality of the set of $d \in N$ such that $d \equiv 1 \pmod 8$ and $C'_d(\mathbb{Q}_p) \neq \phi$ for any $p|(a-b)n$ by $2^{s_2(n)}$. One can consider $s_1(n)$ as the 2-rank of the set of $d \in N$ with respect to restrictions (i) and (ii) of Lemma 7, and $s_2(n)$ as the 2-rank of the set of $d \in N$ with respect to restrictions (i), (ii), (iii) and (v) of Lemma 7. One sees that

$$s_2(n) \le \hat{s}(n, \hat{\phi}) \le s_1(n). \tag{9}$$

We will treat $s_1(n)$ first.

5.1. **Analysis of $s_1(n)$.** We will take a graphical approach to study the asymptotic behavior of $s_1(n)$. For $n \in S(X, h, C)$, let

$$n = p_1 \cdots p_t q_1 \cdots q_s$$

be its prime factorization, where the prime numbers p_i, q_j satisfy the conditions

$$\left(\frac{ab}{p_i}\right) = 1, \quad \left(\frac{ab}{q_j}\right) = -1.$$

We construct a graph $G_1(n) = (V, A)$ with

$$V = \{p_1, \ldots, p_t, q_1, \ldots, q_s, a\},$$

$$\begin{aligned} A = & \left\{\overrightarrow{p_i p_r} \, : \, \left(\tfrac{p_r}{p_i}\right) = -1, \ 1 \le i \ne r \le t\right\} \\ & \bigcup \left\{\overrightarrow{p_i q_j} \, : \, \left(\tfrac{q_j}{p_i}\right) = -1, \ 1 \le i \le t, \ 1 \le j \le s\right\} \\ & \bigcup \left\{\overrightarrow{p_i a} \, : \, \left(\tfrac{-a}{p_i}\right) = -1, \ 1 \le i \le t\right\}. \end{aligned}$$

One can see from (i) and (ii) of Lemma 7 that, for any $d \in N$, $C'_d(\mathbb{Q}_p) \neq \emptyset$ for any $p|n$ if and only if the partition

$$\{p : p|d\} \bigcup \left(\left\{p : p \,\middle|\, \frac{n}{d}\right\} \bigcup \{a\}\right)$$

is an even partition. Hence the number $2^{s_1(n)}$ is the number of even partitions

$$V = V_1 \bigcup V_2$$

of the graph $G_1(n)$ with the condition that $a \in V_2$. Putting the vertices in order as $p_1, \ldots, p_t, q_1, \ldots, q_s, a$ and letting $M_1(n)$ be the Laplace matrix of the graph $G_1(n)$, by Lemma 1, one obtains that $2^{s_1(n)}$ equals the number of vectors $(x_1, \ldots, x_t, y_1, \ldots, y_s) \in \mathbb{F}_2^{t+s}$ such that

$$M_1(n)(x_1, \ldots, x_t, y_1, \ldots, y_s, 0)^T = \mathbf{0}.$$

We may write the matrix $M_1(n)$ explicitly as

$$M_1(n) = \begin{bmatrix} & & * \\ A & B & \vdots \\ & & * \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \\ 0\cdots 0 & 0\cdots 0 & 0 \end{bmatrix},$$

where A is a $t \times t$ matrix and B is a $t \times s$ matrix. One has

$$s_1(n) = t + s - \mathrm{rank}_{\mathbb{F}_2}[AB].$$

Denoting

$$\hat{s}_1(n) = t - \mathrm{rank}_{\mathbb{F}_2}[AB],$$

we will prove the following.

Lemma 12. $\hat{s}_1(n) = 0$ *for almost all* $n \in S(X, h, C)$*, as* $X \to \infty$.

Proof of Lemma 12. First one sees immediately $\hat{s}_1(n) \geq 0$. Next, when $n \equiv 1 \pmod 4$, one constructs a graph (V,A) with

$$V = \{p_1, \ldots, p_t, q_1, \ldots, q_s, a\},$$

$$\begin{aligned} A = &\left\{\overrightarrow{p_i p_r} : \left(\frac{p_i}{p_r}\right) = -1,\ 1 \leq i \neq r \leq t\right\} \\ &\bigcup \left\{\overrightarrow{p_i q_j} : \left(\frac{p_i}{q_j}\right) = -1,\ 1 \leq i \leq t,\ 1 \leq j \leq s\right\} \\ &\bigcup \left\{\overrightarrow{p_i a} : \left(\frac{a}{p_i}\right) = -1,\ 1 \leq i \leq t\right\}. \end{aligned}$$

When $n \equiv 3 \pmod 4$, one constructs another graph (V,A) with

$$V = \{p_1, \ldots, p_t, q_1, \ldots, q_s, a\},$$

$$\begin{aligned} A = &\left\{\overrightarrow{p_i p_r} \,:\, \left(\tfrac{p_i}{p_r}\right) = -1,\ 1 \le i \neq r \le t\right\} \\ &\bigcup \left\{\overrightarrow{p_i q_j} \,:\, \left(\tfrac{p_i}{q_j}\right) = -1,\ 1 \le i \le t,\ 1 \le j \le s\right\} \\ &\bigcup \left\{\overrightarrow{p_i a} \,:\, \left(\tfrac{-a}{p_i}\right) = -1,\ 1 \le i \le t\right\}. \end{aligned}$$

Let $M_1'(n)$ be the Laplace matrix in either of these two cases. Putting the vertices in order as $p_1, \ldots, p_t, q_1, \ldots, q_s, a$, we may write the Laplace matrix $M_1'(n)$ in the form

$$M_1'(n) = \begin{bmatrix} & & * \\ A^T & C & \vdots \\ & & * \\ & 0 & 0 \\ B^T & \vdots & \vdots \\ & 0 & 0 \\ 0\cdots 0 & 0\cdots 0 & 0 \end{bmatrix}.$$

Denote by $\epsilon(n)$ the set of even partitions $V = V_1 \bigcup V_2$ of the graph such that $\{q_1, \ldots, q_s, a\} \subset V_2$. Then $\#\epsilon(n)$ is equal to the number of vectors $(x_1, \ldots, x_t)$ such that

$$M_1'(n)(x_1, \ldots, x_t, 0, \ldots, 0, 0)^T = \mathbf{0}.$$

Therefore

$$\#\epsilon(n) = 2^{t - \mathrm{rank}_{\mathbb{F}_2}\left(\begin{smallmatrix} A^T \\ B^T \end{smallmatrix}\right)} = 2^{\hat{s}_1(n)}.$$

One can see that the set $\epsilon(n)$ corresponds to the set of $d \in N$ subject to the following conditions: if $n \equiv 1 \pmod 4$, then:

(i) For $p|n, p|d$: $\left(\dfrac{ab}{p}\right) = 1$ and $\left(\dfrac{a}{p}\right)\left(\dfrac{p}{n/d}\right) = 1$.

(ii) For $p|n, p \nmid d$: $\left(\dfrac{p}{d}\right) = 1$.

If $n \equiv 3 \pmod 4$, then

$$\begin{array}{ll} \text{(i)} & \text{For } p|n, p|d: \left(\dfrac{ab}{p}\right) = 1 \text{ and } \left(\dfrac{-a}{p}\right)\left(\dfrac{p}{n/d}\right) = 1\,. \\ \text{(ii)} & \text{For } p|n, p \nmid d: \left(\dfrac{p}{d}\right) = 1\,. \end{array}$$

From the above description one sees that if $n \equiv 1 \pmod 4$, then

$$2^{\hat{s}_1(n)} = \sum_{n=dd'} \prod_{p|d} \frac{1}{4}\left(\left(\frac{ab}{p}\right)+1\right)\left(\left(\frac{a}{p}\right)\left(\frac{p}{d'}\right)+1\right) \prod_{p|d'} \frac{1}{2}\left(\left(\frac{p}{d}\right)+1\right),$$

and if $n \equiv 3 \pmod 4$, then

$$2^{\hat{s}_1(n)} = \sum_{n=dd'} \prod_{p|d} \frac{1}{4}\left(\left(\frac{ab}{p}\right)+1\right)\left(\left(\frac{-a}{p}\right)\left(\frac{p}{d'}\right)+1\right) \prod_{p|d'} \frac{1}{2}\left(\left(\frac{p}{d}\right)+1\right).$$

In both of these two cases one can apply Heath-Brown's method as in Section 4 to obtain the asymptotic formula

$$\sum_{n \in S(X,h,C)} 2^{\hat{s}_1(n)} = \#S(X,h,C) + O\left(X(\log X)^{-1/5}\right),$$

as $X \to \infty$, and this implies that

$$\hat{s}_1(n) = 0$$

for almost all $n \in S(X,h,C)$, as $X \to \infty$. This completes the proof of Lemma 12. Since $s_1(n) = s + \hat{s}_1(n)$, by Lemma 12, $s_1(n) = s$ for almost all $n \in S(X,h,C)$, as $X \to \infty$.

5.2. **Analysis of $s_2(n)$.** For $n \in S(X,h,C)$, let

$$n = p_1 \cdots p_t q_1 \cdots q_s$$

be its prime factorization, where the prime numbers p_i, q_j satisfy the conditions

$$\left(\frac{ab}{p_i}\right) = 1, \quad \left(\frac{ab}{q_j}\right) = -1.$$

We construct a graph $G_2(n) = (V, A)$ with

$$V = \{p_1, \ldots, p_t, q_1, \ldots, q_s, a, e_1, e_2\} \bigcup \{p : p|(a-b)\},$$

$$\begin{aligned} A = &\left\{\overrightarrow{p_i q_j} : \left(\tfrac{q_j}{p_i}\right) = -1,\ 1 \le i \le t,\ 1 \le j \le s\right\} \\ &\bigcup \left\{\overrightarrow{p_i a} : \left(\tfrac{-a}{p_i}\right) = -1,\ 1 \le i \le t\right\} \\ &\bigcup \{\overrightarrow{e_1 p} : p \equiv 3 \pmod 4,\ p|n\} \\ &\bigcup \{\overrightarrow{e_2 p} : p \equiv \pm 3 \pmod 8,\ p|n\} \\ &\bigcup \left\{\overrightarrow{pq} : \left(\tfrac{q}{p}\right) = -1, p|(a-b), q|n\right\}. \end{aligned}$$

One can see from (i), (ii), (iii) and (v) of Lemma 7 that for any $d \in N$, if the partition

$$\{p : p|d\} \bigcup \left(\left\{p : p \left| \frac{n}{d}\right.\right\} \bigcup \{a, e_1, e_2\} \bigcup \{p : p|(a-b)\}\right)$$

is an even partition, then $C'_d(\mathbb{Q}_p) \neq \emptyset$ for any $p|(a-b)n$. Hence $2^{s_2(n)}$ is at least the number of even partitions

$$V = V_1 \bigcup V_2$$

of the graph $G_2(n)$ with the condition that $\{a, e_1, e_2\} \bigcup \{p : p|(a-b)\} \subset V_2$. Putting the vertices in order as $p_1, \ldots, p_t, q_1, \ldots, q_s, a, e_1, e_2$ and p for $p|(a-b)$, and letting $M_2(n)$ be the Laplace matrix of the graph $G_2(n)$, by Lemma 1 one obtains that $2^{s_2(n)}$ is at least the number of vectors $(x_1, \ldots, x_t, y_1, \ldots, y_s) \in \mathbb{F}_2^{t+s}$ such that

$$M_2(n)(x_1, \ldots, x_t, y_1, \ldots, y_s, 0, 0, 0, \mathbf{0})^T = \mathbf{0}.$$

We may write the matrix $M_2(n)$ explicitly as

$$M_2(n) = \begin{bmatrix} & & * & 0 & 0 & 0 \\ A & B & \vdots & \vdots & \vdots & \vdots \\ & & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 \cdots 0 & 0 \cdots 0 & 0 & 0 & 0 & 0 \\ * \cdots * & * \cdots * & 0 & * & 0 & 0 \\ * \cdots * & * \cdots * & 0 & 0 & * & 0 \\ * \cdots * & * \cdots * & 0 & 0 & 0 & * \end{bmatrix},$$

where $M_2(n)$ is a $(t+s+3+\omega(a-b))^2$ matrix, A is a $t \times t$ matrix and B is a $t \times s$ matrix. One has

$$s_2(n) \geq t + s - \operatorname{rank}_{\mathbb{F}_2} \begin{bmatrix} A & B \\ * & * \\ * & * \\ * & * \end{bmatrix} \geq s - 2 - \omega(a-b) + t - \operatorname{rank}_{\mathbb{F}_2}[AB].$$

One sees from Lemma 12 that $\hat{s}_1(n) = t - \operatorname{rank}_{\mathbb{F}_2}[AB] = 0$ for almost all $n \in S(X,h,C)$, as $X \to \infty$; hence $s_2(n) \geq s - 2 - \omega(a-b)$ for almost all $n \in S(X,h,C)$, as $X \to \infty$. From the inequalities (9) and (8), one concludes that

Lemma 13.

$$s(n, \hat{\phi}) = \sum_{\substack{p|n\,, \\ (\frac{ab}{p})=-1}} 1 + O(1)$$

for almost all $n \in S(X,h,C)$, as $X \to \infty$.

It is easy to see that Lemma 13 also holds true if $\gcd(a,b) > 1$.

For $n \in S(X, h, C)$, denote

$$h(n) = \sum_{\substack{p|n\,, \\ (\frac{ab}{p})=-1}} 1,$$

where ab is not a square. Let

$$S = \{n \in \mathbb{N} : n \text{ is square-free and } n \equiv h \pmod{C}\}.$$

Define the map $f : S \to \mathbb{N}$ as

$$f(n) = \prod_{\substack{p|n \\ (\frac{ab}{p})=-1}} p$$

for any $n \in S$. Then

$$h(n) = \omega(f(n)).$$

One can verify that the set S and the function f satisfy all the conditions listed in Lemma 2 with constant $c = \frac{1}{2}$. Therefore for $n \in S(X, h, C)$ and $X \to \infty$, $h(n)$, as well as $s(n, \hat{\phi})$ satisfies the desired Gaussian distribution, with mean and variance $\frac{1}{2} \log \log n$. This proves the second part of Theorem 1. The proof of Theorem 1 is now complete.

6. Proof of Theorem 2

Recall that $\phi : E_n \to E'_n$ is a 2-isogeny and $\hat{\phi} : E'_n \longrightarrow E_n$ is the dual 2-isogeny. Hence $\hat{\phi} \circ \phi = [\,2\,]$, one has the following commutative diagrams (see pp 97, [1]):

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & \frac{E'_n(\mathbb{Q})}{\phi(E_n(\mathbb{Q}))} \arrow[r] \arrow[d] & Sel^{(\phi)}(E_n/\mathbb{Q}) \arrow[r] \arrow[d] & Ш(E_n/\mathbb{Q})[\phi] \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \frac{E_n(\mathbb{Q})}{2E_n(\mathbb{Q}))} \arrow[r] \arrow[d] & Sel^{(2)}(E_n/\mathbb{Q}) \arrow[r] \arrow[d] & Ш(E_n/\mathbb{Q})[2] \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \frac{E_n(\mathbb{Q})}{\hat{\phi}(E'_n(\mathbb{Q}))} \arrow[r] \arrow[d] & Sel^{(\hat{\phi})}(E'_n/\mathbb{Q}) \arrow[r] \arrow[d] & Ш(E'_n/\mathbb{Q})[\hat{\phi}] \arrow[r] \arrow[d] & 0 \\
 & 0 \arrow[r] & \hat{C} \arrow[r] \arrow[d] & \hat{C} \arrow[r] \arrow[d] & 0 \\
 & & 0 & 0 &
\end{tikzcd}$$

For $n \in S(X, h, D)$, denote as in the theorem

$$\#Ш(E_n/\mathbb{Q})[\phi] = 2^{t(n,\phi)},\ \#Ш(E'_n/\mathbb{Q})[\hat{\phi}] = 2^{t(n,\hat{\phi})},\ \#Ш(E_n/\mathbb{Q})[2] = 2^{t(n)},$$

and

$$\#Sel^{(\phi)}(E_n/\mathbb{Q}) = 2^{s(n,\phi)},\ \#Sel^{(\hat{\phi})}(E'_n/\mathbb{Q}) = 2^{s(n,\hat{\phi})},\ \#Sel^{(2)}(E_n/\mathbb{Q}) = 2^{s(n)}.$$

From the above commutative diagrams one has the inequality

$$0 \le t(n, \phi) \le s(n, \phi),$$

which immediately implies that $t(n, \phi) = 0$ for almost all $n \in S(X, h, D)$ as $X \to \infty$ by Theorem 1, under the stronger assumptions of Theorem 2. One also has the relation

$$s(n,\hat{\phi}) - s(n) \le t(n,\hat{\phi}) \le s(n,\hat{\phi}),$$

hence

$$\sum_{n\in S(X,h,D)} \Big(s(n,\hat{\phi}) - s(n)\Big)^k \le \sum_{n\in S(X,h,D)} t(n,\hat{\phi})^k \le \sum_{n\in S(X,h,D)} s(n,\hat{\phi})^k \,. \tag{10}$$

By Lemma 13, $s(n,\hat{\phi}) = h(n) + O(1)$, where

$$h(n) = \sum_{\substack{p|n\,,\\ (\frac{ab}{p})=-1}} 1.$$

We will prove in the next section that for any $k \in \mathbb{N}$

$$\sum_{n\in S(X,h,D)} h(n)^k = \#S(X,h,D)\left(\frac{\log\log X}{2}\right)^k + O_k\Big(X\left(\log\log X\right)^{k-1}\Big)\,. \tag{11}$$

6.1. **k-th moment of the h-function.** To establish the asymptotic formula (11), we first prove the case when $k=1$, which is essentially the following lemma.

Lemma 14. *For $X>0$ and non-zero integers c, h, C such that c is not a square, $\left(\prod_{p|c} p\right)|C$ and $\gcd(h,C)=1$, let the set $S(X,h,C)$ be defined in (2). For any $n \in S(X,h,C)$, define the function*

$$h(n) = \sum_{\substack{p|n\,,\\ (\frac{c}{p})=-1}} 1\,.$$

Then

$$\sum_{n\in S(X,h,C)} h(n) = \#S(X,h,C)\,\left(\frac{\log\log X}{2}\right) + O(X)$$

as $X \to \infty$.

Proof. First we write

$$\sum_{n\in S(X,h,C)} h(n) = \sum_{\substack{n\le X\\ n\equiv h \pmod{C}}} \mu^2(n)h(n).$$

Removing the condition $n \equiv h \pmod{C}$ by inserting the factor

$$\frac{1}{\phi(C)} \sum_{\psi \pmod{C}} \psi(n)\overline{\psi(h)},$$

where ϕ is the Euler-ϕ function, and interchanging the summation one has

$$\sum_{n\in S(X,h,C)} h(n) = \frac{1}{\phi(C)} \sum_{\psi \pmod{C}} \overline{\psi(h)} \sum_{n\le X} \mu^2(n)h(n)\psi(n).$$

For the character $\psi \pmod{C}$, denote

$$S(\psi, X) = \sum_{n\le X} \mu^2(n)h(n)\psi(n).$$

If $\psi \neq 1$, one has

$$S(\psi, X) = \sum_{n\le X} \mu^2(n)\psi(n) \sum_{\substack{p\mid n\\ (\frac{c}{p})=-1}} 1 = \sum_{\substack{p\le X\\ (\frac{c}{p})=-1}} \psi(p) \sum_{\substack{m\le X/p\\ \gcd(m,p)=1}} \mu^2(m)\psi(m).$$

By Lemma 4,

$$\sum_{\substack{m\le X/p\\ \gcd(m,p)=1}} \mu^2(m)\psi(m) \ll \frac{X}{p} \exp\left(-\eta\sqrt{\log(X/p)}\right).$$

Since

$$\begin{aligned}\sum_{p\le\sqrt{X}} \frac{1}{p\exp\left(\eta\sqrt{\log(X/p)}\right)} &\le \exp\left(-\eta\sqrt{(\log X)/2}\right) \sum_{p\le\sqrt{X}} p^{-1}\\ &\ll \exp\left(-\eta\sqrt{\log X}\right) \log\log X \ll 1,\end{aligned}$$

and

$$\sum_{\sqrt{X}<p\le X} \frac{1}{p\exp\left(\eta\sqrt{\log(X/p)}\right)} \le \sum_{\sqrt{X}<p\le X} p^{-1} \ll 1,$$

one has

$$S(\psi, X) \ll X.$$

When $\psi = 1$, one has

$$S(1, X) \;=\; \sum_{\substack{n \le X \\ \gcd(n,C)=1}} \mu^2(n) \sum_{\substack{p|n \\ (\frac{c}{p})=-1}} 1 = \sum_{\substack{p \le X \\ (\frac{c}{p})=-1 \\ \gcd(p,C)=1}} \sum_{\substack{m \le X/p \\ \gcd(m,pC)=1}} \mu^2(m)\,.$$

For any integer $r \ge 1$, denote

$$A(r, X) = \sum_{\substack{n \le X \\ \gcd(n,r)=1}} \mu^2(n).$$

We define the multiplicative function g by convolution $g = \mu^2 * \mu$. One sees that $\mu^2 = 1 * g$ and for any prime p,

$$g(p^m) = \begin{cases} 0 & : \quad m = 1, \\ -1 & : \quad m = 2, \\ 0 & : \quad m \ge 3\,. \end{cases}$$

Then

$$A(r, X) \;=\; \sum_{\substack{n \le X \\ \gcd(n,r)=1}} \sum_{d|n} g(d) = \sum_{\substack{d \le X \\ \gcd(d,r)=1}} g(d) \sum_{\substack{m \le X/d \\ \gcd(m,r)=1}} 1 = \sum_{\substack{n \le \sqrt{X} \\ \gcd(n,r)=1}} \mu(n) \sum_{\substack{m \le X/n^2 \\ \gcd(m,r)=1}} 1\,.$$

Since

$$\sum_{\substack{m \le X \\ \gcd(m,r)=1}} 1 \;=\; \sum_{d|r} \mu(d) \cdot \left[\frac{X}{d}\right] = \sum_{d|r} \mu(d) \cdot \left(\frac{X}{d} + O(1)\right) = \frac{\phi(r)X}{r} + O(\tau(r)),$$

and

$$\sum_{\substack{n \le \sqrt{X} \\ \gcd(n,r)=1}} \frac{\mu(n)}{n^2} = \frac{6}{\pi^2} \prod_{p|r} (1 - p^{-2})^{-1} + O(X^{-1/2}),$$

one obtains that

$$A(r,X) \;=\; \frac{6X}{\pi^2}\prod_{p|r}(1+p^{-1})^{-1} + O\left(\sqrt{X}\tau(r)\right).$$

Using this result we have

$$\begin{aligned}
S(1,X) &= \sum_{\substack{p\le X\\ (\frac{c}{p})=-1\\ \gcd(p,C)=1}} A(pC, X/p)\\
&= \sum_{\substack{p\le X\\ (\frac{c}{p})=-1\\ \gcd(p,C)=1}} \left(\frac{6X}{\pi^2 p}\,(1+p^{-1})^{-1}\prod_{q|C}(1+q^{-1})^{-1} + O\left(\sqrt{\frac{X}{p}}\,\tau(pC)\right)\right)\\
&= \frac{6X}{\pi^2}\prod_{q|C}(1+q^{-1})^{-1} \sum_{\substack{p\le X\\ (\frac{c}{p})=-1\\ \gcd(p,C)=1}} \frac{1}{p+1} + O\left(X^{1/2}\sum_{p\le X}p^{-1/2}\right).
\end{aligned}$$

Since

$$\begin{aligned}
\sum_{\substack{p\le X\\ (\frac{c}{p})=-1\\ \gcd(p,C)=1}} \frac{1}{p+1} &= \sum_{\substack{p\le X\\ (\frac{c}{p})=-1\\ \gcd(p,C)=1}} \frac{1}{p} + O(1) = \frac{1}{2}\sum_{\substack{p\le X\\ \gcd(p,C)=1}} \frac{1-(\frac{c}{p})}{p} + O(1)\\
&= \frac{\log\log X}{2} + O(1),
\end{aligned}$$

by Merten's estimate, and

$$\sum_{p\le X}p^{-1/2} \le \left(\sum_{p\le X}1\right)^{1/2}\cdot\left(\sum_{p\le X}p^{-1}\right)^{1/2} \ll \left(\frac{X}{\log X}\right)^{1/2}(\log\log X)^{1/2},$$

one obtains that

$$S(1,X) = \frac{3}{\pi^2}\prod_{p|C}(1+p^{-1})^{-1}\; X\log\log X + O(X)\,.$$

Finally, one concludes that

$$
\begin{aligned}
\sum_{n \in S(X,h,C)} h(n) &= \frac{1}{\phi(C)} \left(S(1,X) + \sum_{\substack{\psi \pmod{C} \\ \psi \neq 1}} \overline{\psi(h)} S(\psi, X) \right) \\
&= \frac{3X \log\log X}{\pi^2 \phi(C) \prod_{p|C}(1+p^{-1})} + O(X)\,.
\end{aligned}
$$

Since

$$
\begin{aligned}
\#S(X,h,C) &= \sum_{\substack{n \leq X \\ n \equiv h \pmod{C}}} \mu^2(n) = \frac{1}{\phi(C)} \sum_{\psi \pmod{C}} \overline{\psi(h)} \sum_{n \leq X} \mu^2(n)\psi(n) \\
&= \frac{1}{\phi(C)} \sum_{\substack{n \leq X \\ \gcd(n,C)=1}} \mu^2(n) + \frac{1}{\phi(C)} \sum_{\substack{\psi \pmod{C} \\ \psi \neq 1}} \overline{\psi(h)} \sum_{n \leq X} \mu^2(n)\psi(n) \\
&= \frac{1}{\phi(C)} \left(\frac{6X}{\pi^2} \prod_{p|C} (1+p^{-1})^{-1} + O(X^{1/2}) \right) + O\left(X \exp\left(-\eta \sqrt{\log X} \right) \right) \\
&= \frac{6X}{\pi^2 \phi(C) \prod_{p|C}(1+p^{-1})} + O\left(X \exp\left(-\eta \sqrt{\log X} \right) \right)\,,
\end{aligned}
$$

one immediately sees that

$$
\sum_{n \in S(X,h,C)} h(n) = \#S(X,h,C) \left(\frac{\log\log X}{2} \right) + O(X)\,.
$$

This completes the proof of Lemma 14. Now we can prove

Lemma 15. *Assume the conditions of Lemma 14. Then for any integer $k \geq 1$,*

$$
\sum_{n \in S(X,h,C)} h(n)^k = \#S(X,h,C) \left(\frac{\log\log X}{2} \right)^k + O_k\left(X \left(\log\log X \right)^{k-1} \right)
$$

as $X \to \infty$.

Proof. For $k = 1$, this is established in Lemma 14. For $k \geq 2$, we recall the following high-power analogues of the Turán-Kubilius inequalities (see [8] or [21]) for the additive function h,

$$\frac{1}{X}\sum_{n\le X}|h(n)-A(X)|^k \ll B(X)^k+\sum_{p^m\le X}\frac{|h(p^m)|^k}{p^m},$$

where

$$A(X)=B^2(X)=\sum_{p^m\le X}\frac{h(p^m)}{p^m}=\frac{\log\log X}{2}+O(1),$$

by the argument in Lemma 14. For $k\ge 2$ one has

$$\begin{aligned}\sum_{n\le X}\left|h(n)-\frac{\log\log X}{2}\right|^k &\ll_k \sum_{n\le X}|h(n)-A(X)|^k+\sum_{n\le X}\left|A(X)-\frac{\log\log X}{2}\right|^k\\ &\ll_k XB(X)^k+X \ll_k X(\log\log X)^{k/2}.\end{aligned}$$

Therefore

$$\begin{aligned}\sum_{n\in S(X,h,C)} h(n)^k &= \sum_{n\in S(X,h,C)}\left(h(n)-\frac{\log\log X}{2}+\frac{\log\log X}{2}\right)^k\\ &= \left(\frac{\log\log X}{2}\right)^k \#S(X,h,C)+k\left(\frac{\log\log X}{2}\right)^{k-1}\sum_{n\in S(X,h,C)}\left(h(n)-\frac{\log\log X}{2}\right)\\ &+ O_k\left(\max_{0\le r\le k-2}\left\{(\log\log X)^r\sum_{n\in S(X,h,C)}\left|h(n)-\frac{\log\log X}{2}\right|^{k-r}\right\}\right).\end{aligned}$$

The second term is

$$O_k\left(X(\log\log X)^{k-1}\right)$$

by Lemma 14, while for any $0\le r\le k-2$, one has

$$\begin{aligned}(\log\log X)^r\sum_{n\in S(X,h,C)}\left|h(n)-\frac{\log\log X}{2}\right|^{k-r} &\ll_k (\log\log X)^r\, X\,(\log\log X)^{(k-r)/2}\\ &\le X\,(\log\log X)^{k-1}.\end{aligned}$$

Putting these two error terms together we complete the proof of Lemma 15. This completes the proof of the asymptotic formula (11).

Using (11) one obtains

$$\sum_{n\in S(X,h,D)} s(n,\hat{\phi})^k = \#S(X,h,D)\left(\frac{\log\log X}{2}\right)^k + O_k\left(X\left(\log\log X\right)^{k-1}\right).$$

Recal the following result obtained by Yu (Theorem 2, [32])

$$\sum_{n\in S(X,h,D} 2^{s(n)} = (3+o(1))\#S(X,h,D),$$

which implies that

$$\sum_{n\in S(X,h,D)} s(n)^k = O_k(X).$$

The magnitude of the left hand side of (10) is

$$\sum_{n\in S(X,h,D)} s(n,\hat{\phi})^k + O_k\left(\max_{0\le r\le k-1}\left\{\sum_{n\in S(X,h,D)} s(n,\hat{\phi})^r s(n)^{k-r}\right\}\right),$$

and for any r with $0\le r\le k-1$, one obtains that

$$\begin{aligned}\sum_{n\in S(X,h,D)} s(n,\hat{\phi})^r s(n)^{k-r} &\le \left(\sum_{n\in S(X,h,D)} s(n,\hat{\phi})^{2r}\right)^{1/2}\left(\sum_{n\in S(X,h,D)} s(n)^{2(k-r)}\right)^{1/2}\\ &\ll_k \left(\sum_{n\in S(X,h,D)} s(n,\hat{\phi})^{2r}\right)^{1/2} X^{1/2}\\ &\ll_k \left(X(\log\log X)^{2r}\right)^{1/2}(X)^{1/2} \le X(\log\log X)^{k-1}.\end{aligned}$$

Therefore

$$\sum_{n\in S(X,h,D)} t(n,\hat{\phi}_i)^k = \#S(X,h,D)\left(\frac{\log\log X}{2}\right)^k + O_k\left(X\left(\log\log X\right)^{k-1}\right),$$

which completes the proof of Theorem 2.

References

[1] N. Aoki, *On the 2-Selmer groups of elliptic curves arising from the congruent number problems*, Comment. Math. Univ. St. Paul. 48(1999), 77–101.

[2] N. Aoki, *On the Tate-Shafarevich group of semistable elliptic curves with a rational 3-torsion*, Acta Arith. 112(2004), no. 3, 209–227.

[3] D. Atake, *On elliptic curves with large Tate-Shafarevich groups*, J. Number theory. 87(2001), 282–300.

[4] R. Bölling, *Die Ordnung der Schafarewitsch-Tate-Gruppe kann beliebig groß werden*, Math. Nachr. 67(1975), 157–179.

[5] J. W. S. Cassels, *Arithmetic on curves of genus 1. VI. The Tate-Shafarevich group can be arbitrarily large*, J. Reine Angew. Math. 214(1963), 65–70.

[6] C. Delaunay, *Heuristics on Tate-Shafarevitch Groups of Elliptic Curves Defined over* $\mathbb{Q}$, Experiment. Math. 10(2001), no. 2, 191-Ũ196.

[7] C. Delaunay, *Moments of the orders of Tate-Shafarevich groups*, Int. J. Number Theory. 1(2005), no. 2, 243–264.

[8] P.D.T.A. Elliott, *High-power analogues of the Turán-Kubilius inequality, and an application to number theory*, Canad. J. Math. 32(1980), no. 4, 893–907.

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

[10] B. Faulkner, K. James, *A graphical approach to computing Selmer groups of congruent number curves*, Ramanujan J. 14(2007), no. 1, 107–129.

[11] K. Feng, *Non-congruent numbers, odd graph and BSD conjecture on* $y^2 = x^3 - n^2x$, Acta. Arith. 80(1996), 71–83.

[12] K. Feng and M. Xiong, *On elliptic curves* $y^2 = x^3 - n^2x$ *with rank zero*, J. of Number Theory 109(2004), no 1, 1–26.

[13] K. Feng, Y. Xue, *New series of odd non-congruent numbers*, to appear in Science in China (Series A), 2006.

[14] D. Goldfeld, D. Lieman, *Effective bounds on the size of the Tate-Shafarevich group*, Math. Res. Lett. 3(1996), no. 3, 309–318.

[15] D. Goldfeld, L. Szpiro, *Bounds for the order of the Tate-Shafarevich group*, Compositio Math. 97(1995), no. 1-2, 71-87.

[16] A. Granville, K. Soundararajan, *Sieving and the Erdős-Kac theorem*, Equidistribution in number theory, an introduction, 15–27, NATO Sci. Ser. II Math. Phys. Chem., 237, Springer, Dordrecht, 2007.

[17] R.R. Hall, G. Tenenbaum, "Divisors", Cambridge Tracts in Mathematics, 90. Cambridge University Press, Cambridge, 1988.

[18] J.M. Harris, J.L. Hirst, M.J. Mossignhoff, "Combinatorics and graph theory", Springer-Verlag, Berlin, 2000.

[19] D.R. Heath-Brown, *The size of Selmer groups for the congruent number problem, I*, Invent. math. 111(1993), 171–195.

[20] D.R. Heath-Brown, *The size of Selmer groups for the congruent number problem, II*, Invent. math. 118(1994), 331–370.

[21] A. Hildebrand, *Sur Les Moments D'une Fonction Additive*, Ann. Inst. Fourier, Grenoble, 33, 3(1983), 1–22.

[22] R. Kloosterman, *The p-part of the Tate-Shafarevich groups of elliptic curves can be arbitrarily large*, J. Théor. Nombres Bordeaux. 17(2005), no. 3, 787–800.

[23] K. Kramer, *A family of semistable elliptic curves with large Tate-Shafarevitch groups*, Proc. Amer. Math. Soc. 89(1983), 379–386.

[24] F. Lemmermeyer, *On Tate-Shafarevich groups of some elliptic curves*, Proc. Conf., Graz, 1998.

[25] F. Lemmermeyer, R. Mollin, *On Tate-Shafarevich groups of* $y^2 = x(x^2 - k^2)$, Acta Math. Univ. Comenian. (N.S.) 72(2003), no. 1, 73–80.

[26] Y. Liu, *Prime analogues of the Erdös-Kac theorem for elliptic curves*, J. Number Theory 119(2006), no. 2, 155–170.

[27] J.H. Silverman, "The Arithmetic of Elliptic Curves", GTM 106, Springer-Vetlag, 1986.

[28] M. Xiong, A. Zaharescu, *Selmer groups and Tate-Shafarevich groups for the congruent number problem*, Preprint.

[29] G. Yu, *Rank 0 quadratic twists of a family of elliptic curves*, Compositio Math. 135(2003), no. 3, 331–356.

[30] G. Yu, *Average size of 2-Selmer groups of a family of elliptic curves. I*, Trans. Amer. Math. Soc. 358(2006), no. 4, 1563–1584.

[31] G. Yu, *Average size of 2-Selmer groups of elliptic curves. II*, Acta. Arith. 117(2005), no. 1, 1–33.

[32] G. Yu, *On the quadratic twists of a family of elliptic curves*, Mathematika 52(2005), no. 1-2, 139–154 (2006).

Maosheng Xiong: Department of Mathematics, Eberly College of Science, Pennsylvania State University, McAllister Building, University Park, PA 16802 USA

E-mail address: xiong@math.psu.edu

Alexandru Zaharescu: Institute of Mathematics of the Romanian Academy, P.O. Box 1–764, 70700 Bucharest, Romania, and Department of Mathematics, University of Illinois at Urbana-Champaign, 273 Altgeld Hall, MC-382, 1409 W. Green Street, Urbana, IL 61801 USA

E-mail address: zaharesc@math.uiuc.edu