Compositio Mathematica **135:** 331–356, 2003.

Rank 0 Quadratic Twists of a Family of Elliptic Curves

GANG YU
Department of Mathematics, The University of Michigan, Ann Arbor, MI 48105 U.S.A
e-mail: gangyu@math.lsa.umich.edu

(Received: 25 July 2001; accepted in final form: 15 November 2001)

Abstract. In this paper, we consider a family of elliptic curves over $\mathbb{Q}$ with 2-torsion part $\mathbb{Z}_2$. We prove that, for every such elliptic curve, a positive proportion of quadratic twists have Mordell–Weil rank 0.

Mathematics Subject Classifications (2000). 11G05, 11L40, 14H52.

Key words. elliptic curves, 2-descent procedure, character sums.

1. Introduction

While it still is not known if there are elliptic curves over $\mathbb{Q}$ of arbitrarily large rank, it is generally believed that curves with large ranks comprise a small 'proportion' of all elliptic curves. In particular, Goldfeld [3] conjectured that the average rank of the quadratic twists of any given elliptic curve over $\mathbb{Q}$ is 1/2. A quick consequence of this is that, for any elliptic curve over $\mathbb{Q}$, asymptotically, there are at least half of the quadratic twists of this curve which have rank 0. Thus a comparatively weaker conjecture states that, for any elliptic curve over $\mathbb{Q}$, the rank 0 quadratic twists comprise a positive proportion of all quadratic twists of the given curve. In the general case, this conjecture, though much weaker than the other famous ones related to elliptic curves, is still open.

Suppose E is an elliptic curve over $\mathbb{Q}$ defined by the equation $y^2 = f(x)$ where $f(x) \in \mathbb{Z}[x]$ is a cubic polynomial. As usual, we denote by $E(\mathbb{Q})$ the Mordell–Weil group of E over $\mathbb{Q}$ and by $r(E(\mathbb{Q}))$ (simply as $r(E)$) the rank of $E(\mathbb{Q})$. For a non-zero integer D, by E_D we denote the Dth quadratic twist given by the equation $Dy^2 = f(x)$. For integer $r \geqslant 0$, and a positive real number X, we define

$$M_E^r(X) := \#\{D : |D| \leqslant X, r(E_D) = r\}.$$

With this notation, the problem is then saying that, for any fixed elliptic curve E over $\mathbb{Q}$, we should have $M_E^0(X) \gg X$ for sufficiently large X.

There have been numerous papers treating this problem.* Because of the work of Kolyvagin [10], most of them are focusing on the nonvanishing of the L-functions (see [1] for a good survey). In light of the work of Shimura [19] and Waldspurger [23], people have been able to get some partial results. With the knowledge about the Fourier coefficients of some new forms, James [7], [8] proved that the quadratic twists for some given curve over $\mathbb{Q}$ have rank 0 for a positive proportion of square-free numbers. James' method was later extended by other authors to some other family of elliptic curves (see [9], [21], etc.). For E a general elliptic curve over $\mathbb{Q}$, the current best unconditional result is due to Ono [15], he proved that $M_E^0(X) \gg X(\log X)^{c-1}$ for some $c > 0$. In [24], Wong proved that there is an infinite family of non-isomorphic elliptic curves such that for each curve a positive proportion of the quadratic twists have rank 0.

There is also another approach via the first descent. In a series of two papers, [4] and [5], Heath-Brown considered the average size of the 2-Selmer groups of congruent number curves $E_D : y^2 = x^3 - D^2x$. As a consequence of the main result of [4], a positive proportion of the curves E_D have rank 0.

For the more general curve E over $\mathbb{Q}$ with 2-torsion $\mathbb{Z}_2 \times \mathbb{Z}_2$, in [25], we generalized Heath-Brown's method and showed that, for such a curve, the average size of the 2-Selmer groups of the quadratic twists E_D with D running over some arithmetic progressions is 12. Along with Monsky's result [12] on the parity of the 2-Selmer rank, this implies that a positive proportion of quadratic twists E_D have rank 0.

While it works well for the curves with 2-torsion $\mathbb{Z}_2 \times \mathbb{Z}_2$, the idea of bounding the average size of 2-Selmer groups may not be sufficient to prove that $M_E^0(X) \gg X$ for a general elliptic curve over $\mathbb{Q}$. This is because the average size of the 2-Selmer groups of the quadratic twists may be too large. The method of directly attacking the resulting homogeneous spaces, however, still works for some families of elliptic curves with 2-torsion other than $\mathbb{Z}_2 \times \mathbb{Z}_2$. In particular, for some elliptic curves over $\mathbb{Q}$ with 2-torsion $\mathbb{Z}_2$, we can achieve this by bounding the average size of the Selmer groups of the quadratic twists corresponding to 2-isogenies.

In this paper, we consider a special family of elliptic curves. Suppose $b \geqslant 2$ is an integer, not a perfect square and admitting a solution $(u, v) \in \mathbb{Z}^2$ for the equation

$$u^2 - bv^2 = -4. \tag{1.1}$$

(One should note that, if such a solution exists, then there are infinitely many pairs of (u, v) satisfying (1.1).) For every fixed b and v satisfying (1.1), we consider a curve E given by the equation

$$E : y^2 = x(x^2 + ax + b), \quad a = bv. \tag{1.2}$$

*We remark that the positive proportional problem for rank 1 quadratic twists has also been studied by some authors. For example, assuming the Riemann Hypothesis, Iwaniec and Sarnak [6] proved that, for any elliptic curve E over $\mathbb{Q}$ and $r = 0, 1$, $M_E^r(X) \gg X$ for sufficiently large X; Unconditionally, Vatsal [22] proved this for $E = X_0(19)$.

We note that b and $a^2 - 4b$ have the same squarefree kernel. In other words, b and $a^2 - 4b$ differ by a perfect square factor. In this paper, we shall prove the following result.

THEOREM 1.1. ⋆*For an elliptic curve E satisfying the above conditions with $(a, b) \neq (4, 2)$, there exists a constant $X_0 > 0$, such that for every $X > X_0$, we have*

$$M_E^0(X) \gg X. \tag{1.3}$$

The idea to prove Theorem 1.1 is that, for a positive proportion of quadratic twists of E, we bound the admissible homogeneous spaces resulting from 2-isogenies by considering the local solvability. In general, it is not very hard (but sometimes a little complicated!) to get asymptotic formulas for the average size of the Selmer groups of the quadratic twists corresponding to the 2-isogenies. Just to prove the Theorem sufficiently and avoid the unnecessary complication, however, it is not necessary to pursue an asymptotic formula. Usually, in considering the local solvability of a homogeneous space, we shall not consider $\mathbb{Q}_2$. We remark that, in many cases, the upper bound we get this way actually gives the main term of the asymptotic formula. The great save is the heavy labor spent on discussing the solvability in $\mathbb{Q}_2$.

2. Rational 2-Isogenies and the Corresponding Selmer Groups

For an elliptic curve E over $\mathbb{Q}$ given by the equation $E: y^2 = x(x^2 + ax + b)$ with $a, b \in \mathbb{Z}$, we know there is a rational 2-isogeny $\phi: E \to \widehat{E}$ onto

$$\widehat{E}: y^2 = x(x + \hat{a}x + \hat{b}),$$

where $\hat{a} = -2a$ and $\hat{b} = a^2 - 4b$. There is also a dual isogeny $\psi: \widehat{E} \to E$ such that $\psi \circ \phi = [2]_E$ and $\phi \circ \psi = [2]_{\widehat{E}}$.

The first descent by these 2-isogenies yields the following short exact sequences:

$$\begin{aligned} &0 \to \psi(\hat{E}(\mathbb{Q})) \to E(\mathbb{Q}) \xrightarrow{\alpha} \mathbb{Q}^\times/\mathbb{Q}^{\times 2}, \\ &0 \to \phi(E(\mathbb{Q})) \to \widehat{E}(\mathbb{Q}) \xrightarrow{\beta} \mathbb{Q}^\times/\mathbb{Q}^{\times 2}, \end{aligned} \tag{2.1}$$

where the map $\alpha : E(\mathbb{Q}) \to \mathbb{Q}^\times/\mathbb{Q}^{\times 2}$ is defined as: $\alpha(O) = 1 \cdot \mathbb{Q}^{\times 2}$, $\alpha((0, 0)) = b \cdot \mathbb{Q}^{\times 2}$ and $\alpha(P) = x \cdot \mathbb{Q}^{\times 2}$ for all points $P = (x, y) \in E(\mathbb{Q})$ different from (0,0) and the identity $\mathcal{O}$. For simplicity, we denote the image of α (which is isomorphic to $E(\mathbb{Q}/\psi(\widehat{E}(\mathbb{Q}))))$ by $W(\widehat{E}/\mathbb{Q})$. The definition of $\beta: \widehat{E}(\mathbb{Q}) \to \mathbb{Q}^\times/\mathbb{Q}^{\times 2}$ is similar and, also, $\mathrm{im}\beta := W(E/\mathbb{Q})$.

Tate showed that the Mordell-Weil rank of E (thus $\widehat{E}$ also) over $\mathbb{Q}$ is given by

$$2^{r(E(\mathbb{Q}))+2} = \#W(\widehat{E}/\mathbb{Q}) \cdot \#W(E/\mathbb{Q}). \tag{2.2}$$

⋆The same result can be proved for the curve E with $(a, b) = (4, 2)$ by a method similar to what we use to prove Theorem 1.1. It is excluded here for the sake of a uniform proof of Theorem 1.1. One notes that, for every odd D, $r(E_D)$ is predicted to be odd!

Thus, to bound $r(E(\mathbb{Q}))$, it suffices to give an upper bound for $\#W(\widehat{E}/\mathbb{Q}) \cdot \#W(E/\mathbb{Q})$. With an elementary argument (cf. [20] or [11], for instance), one can see that $W(\widehat{E}/\mathbb{Q})$ consists of the classes $b_1 \cdot \mathbb{Q}^{\times^2}$, where b_1 is a squarefree integer (could be negative!) with $b_1 b_2 = b$, such that the homogeneous space

$$T^{(\psi)}(b_1)\colon N^2 = b_1 M^4 + aM^2 e^2 + b_2 e^4 \tag{2.3}$$

has a non-trivial primitive solution in integers $N, M, e \in \mathbb{N}$. (Here the primitive means that $(N, e) = (M, e) = 1$.) Similarly, $W(E/\mathbb{Q})$ consists of the classes $\hat{b}_1 \cdot \mathbb{Q}^{\times^2}$, where $\hat{b}_1$ is a squarefree integer with $\hat{b}_1 \hat{b}_2 = a^2 - 4b$, such that the homogeneous space

$$T^{(\phi)}(\hat{b}_1)\colon X^2 = \hat{b}_1 Y^4 - 2aY^2 Z^2 + \hat{b}_2 Z^4 \tag{2.4}$$

has a non-trivial primitive solution in integers $X, Y, Z \in \mathbb{N}$.

The classes $b_1 \cdot \mathbb{Q}^{\times^2}$ such that (2.3) is everywhere (including $\mathbb{Q}_\infty$) locally solvable consist the Selmer group of $\widehat{E}$ corresponding to the 2-isogeny ψ (or say the ψ-part of the Selmer group of $\widehat{E}$ over $\mathbb{Q}$), usually denoted by $S^{(\psi)}(\widehat{E}/\mathbb{Q})$. Similarly, we have the ϕ-part of the Selmer groups of E over $\mathbb{Q}$, $S^{(\phi)}(E/\mathbb{Q})$.

We remark here that, from the classic Kummer exact sequence

$$0 \to W(E/\mathbb{Q}) \to S^{(\phi)}(E/\mathbb{Q}) \to \text{Ш}(E/\mathbb{Q})[\phi] \to 0, \tag{2.5}$$

$W(E/\mathbb{Q})$ is a subgroup of $S^{(\phi)}(E/\mathbb{Q})$. The factor group $\text{Ш}(E/\mathbb{Q})[\phi]$ is the ϕ-part of the Tate-Shafarevich group of E over $\mathbb{Q}$. Similarly, $W(\widehat{E}/\mathbb{Q})$ is a subgroup of $S^{(\psi)}(\widehat{E}/\mathbb{Q})$.

We now state a result about the parity of the sum of the 2-ranks of $S^{(\phi)}(E/\mathbb{Q})$ and $S^{(\psi)}(\widehat{E}/\mathbb{Q})$. This appeared as a conjecture in [11], but turned out to be a corollary of a result in [12].

LEMMA 2.1. *Suppose $E/\mathbb{Q}$ permits a rational 2-isogeny. Then*

$$\operatorname{rank}(S^{(\phi)}(E/\mathbb{Q})) + \operatorname{rank}(\mathbf{S}^{(\psi)}(\hat{\boldsymbol{E}}/\mathbb{Q})) \equiv v \pmod 2, \tag{2.6}$$

where $w = (-1)^v$ is the root number in the functional equation of the Hasse-Weil L-function of E.

For an elliptic curve $E/\mathbb{Q}$ with conductor N_E, we know that the functional equation for the Hasse–Weil L-function of E is given by

$$\left(\frac{\sqrt{N_E}}{2\pi}\right)^s \Gamma(s) L(E, s) = w_E \left(\frac{\sqrt{N_E}}{2\pi}\right)^{2-s} \Gamma(2-s) L(E, 2-s), \tag{2.7}$$

where w_E is the root number. For an integer D prime to N_E, the functional equation of the L-function of the quadratic twist E_D is given by

$$\left(\frac{D\sqrt{N_E}}{2\pi}\right)^s \Gamma(s) L(E_D, s) = w_{E_D} \left(\frac{D\sqrt{N_E}}{2\pi}\right)^{2-s} \Gamma(2-s) L(E_D, 2-s). \tag{2.8}$$

Here the root number w_{E_D} is related to w_E by

$$w_{E_D} = w_E \cdot \left(\frac{-D}{-N_E}\right), \tag{2.9}$$

where $(-D/\cdot)$ is the Kronecker character. Thus, for the quadratic twists E_D with D running over a fixed class in $(\mathbb{Z}/4N_E\mathbb{Z})^\times$, the parity of $\operatorname{rank}(S^{(\phi)}(E_D/\mathbb{Q})) + \operatorname{rank}(S^{(\psi)}(\widehat{E}_D/\mathbb{Q}))$ is fixed. Henceforth, C_E is defined as a fixed positive integer, divisible by $4N_E$ for a given elliptic curve $E/\mathbb{Q}$, h is an integer prime to C_E. For a positive real number X, we define

$$S(X; h) := \{0 < D \leqslant X : D \equiv h \,(\mathrm{mod} C_E),\ D \text{ squarefree}\}. \tag{2.10}$$

For Theorem 1.1, we shall not compute how large the proportion could be. It doesn't harm if we choose a special arithmetic progression $h \pmod{C_E}$ so that we can simplify the proof a little bit. For this sake, we assume that h satisfies that

$$\left(\frac{-2vh}{p}\right) = -1 \quad \text{for odd prime } p|u. \tag{2.11}$$

We note that there is no any conflict for this restriction since every odd prime divisor of u divides N_E but doesn't divide v. Also we note that, with such an assumption, we still have room for the restriction on h so that w_{E_D} for $D \in S(X; h)$ takes a fixed value, either 1 or -1.

In the next sections, we shall devote to proving the following Theorem 2.2 which, along with Lemma 2.1, implies Theorem 1.1.

THEOREM 2.2. *Suppose E/Q is an elliptic curve given by the Weierstrass equation*

$$E: y^2 = x(x^2 + ax + b),$$

where $a, b \in \mathbb{Z}$ and satisfy the conditions (1.1) *and* (1.2)*, and b is squarefree. Then for a fixed h prime to C_E, satisfying* (2.11)*, and sufficiently large X, we have*

$$\sum_{D \in S(X;h)} \#S^{(\phi)}(E_D/\mathbb{Q}) \cdot \#S^{(\psi)}(\widehat{E}_D/\mathbb{Q}) \leqslant (13 + o(1))\#S(X; h). \tag{2.12}$$

Proof of Theorem 1.1. For the curve E in Theorem 1.1, suppose $b = \tilde{b}\delta^2$, where $\tilde{b}$ is squarefree, then E is actually in the form

$$E: y^2 = x(x^2 + \tilde{b}\delta^2 vx + \tilde{b}\delta^2), \tag{2.13}$$

where v is an integer such that, with some u, (u, v) is a solution of (1.1). We note that the δth quadratic twist of E is given by the equation

$$E_\delta: y^2 = x(x^2 + \tilde{b}\delta vx + \tilde{b}). \tag{2.14}$$

For E_δ, the conditions (1.1) and (1.2) are satisfied. We apply Theorem 2.2 to E_δ: let $C_{E_\delta} = 4N_{E_\delta}$, h is any integer such that (2.11) (for the corresponding new u) is satisfied and

$$w_{E_\delta} \cdot \left(\frac{-h}{-N_{E_\delta}}\right) = 1. \tag{2.15}$$

This ensures that, for every $D \in S(X; h)$, $\operatorname{rank}(S^{(\phi)}(E_{\delta D}/\mathbb{Q})) + \operatorname{rank}(S^{(\psi)}(\widehat{E}_{\delta D}/\mathbb{Q}))$ be even. In other words, $\#(S^{(\phi)}(E_{\delta D}/\mathbb{Q})) \cdot \#(S^{(\psi)}(\widehat{E}_{\delta D}/\mathbb{Q}))$ can only take values 2^{2+2k}, for $k = 0, 1, 2, \ldots$. Hence, Theorem 2.2 implies that, for a positive proportion of $D \in S(X; h)$,

$$\#(S^{(\phi)}(E_{\delta D}/\mathbb{Q})) \cdot \#(S^{(\psi)}(\widehat{E}_{\delta D}/\mathbb{Q})) = 4. \tag{2.16}$$

Together with (2.2), (2.16) yields that, for a positive portion of $D \in S(X; h)$, we have $r(E_{\delta D}(\mathbb{Q})) = 0$. Note that $S(X; h)$ is of positive asymptotic density among integers, this gives Theorem 1.1 for curve E_δ. It is clear that this also implies that $M_E^0(X) \gg X$ for sufficiently large X. □

Remark 2.3. For every given curve E satisfying (1.1) and (1.2), while we don't give an asymptotic formula in Theorem 2.2 as the upper bound (2.12) is sufficient for Theorem 1.1, it is not hard to get asymptotic formulas respectively for the average orders of $S^{(\phi)}(E_D/\mathbb{Q})$ and $S^{(\psi)}(\widehat{E}_D/\mathbb{Q})$ as D runs over $S(X; h)$. It turns out that the average orders of $S^{(\phi)}(E_D/\mathbb{Q})$ and $S^{(\psi)}(\widehat{E}_D/\mathbb{Q})$ are always some constants depending on E. For instance, without proof, we can state the following results.

EXAMPLE 2.4. Suppose E is the elliptic curve given by the Weierstrass equation

$$E: y^2 = x(x^2 + 4x + 2).$$

Then for sufficiently large X, we have

$$\sum_{D \in S_1(X)} \#S^{(\phi)}(E_D/\mathbb{Q}) = (6 + o(1))\#S_1(X) \tag{2.17}$$

and

$$\sum_{D \in S_1(X)} \#S^{(\psi)}(\widehat{E}_D/\mathbb{Q}) = (3 + o(1))\#S_1(X) \tag{2.18}$$

where $S_1(X)$ is the set of the odd positive squarefree integers not exceeding X. Moreover, the average orders of $S^{(\phi)}(E_D/\mathbb{Q})$ and $S^{(\psi)}(\widehat{E}_D/\mathbb{Q})$ with D running over even squarefree integers are both 4.

We remark that, while the proof for Example 2.4 is much easier than Theorem 2.2, the results of (2.17) and (2.18), combining with Cauchy's inequality, don't imply Theorem 1.1. This is almost always the case for every curve satisfying conditions (1.1) and (1.2).

3. An Upper Bound of $\#S^{(\varphi)}(E_D/\mathbb{Q})\#S^{(\psi)}(\widehat{E}_D/\mathbb{Q})$

The curve E in this section satisfies the conditions in Theorem 2.2. First of all, we note from (2.3) that, corresponding to $W(\widehat{E}_D(\mathbb{Q}))$, we have homogeneous spaces $T^{(\psi)}(b_1)^\star$ given by

$$T^{(\psi)}(b_1): N^2 = b_1M^4 + aDM^2e^2 + \frac{bD^2}{b_1}e^4, \tag{3.1}$$

where $|b_1|$ is a divisor of bD^2. For $D \in S(X; h)$, we have $(D, b) = 1$. Thus, we suppose that $b_1 = \tilde{B}_0\tilde{D}_0$, where $|\tilde{B}_0|$ is a squarefree divisor of b and $\tilde{D}_0 > 0$ is a divisor of D. Write $D = \tilde{D}_0\tilde{D}_1$, $b = \tilde{B}_0\tilde{B}_1$, then (3.1) becomes

$$T^{(\psi)}(\tilde{B}_0\tilde{D}_0): N^2 = \tilde{B}_0\tilde{D}_0M^4 + \tilde{B}_0\tilde{B}_1v\tilde{D}_0\tilde{D}_1M^2e^2 + \tilde{B}_1\tilde{D}_0{\tilde{D}_1}^2e^4. \tag{3.2}$$

Note that $\tilde{D}_0$ is squarefree, thus (3.2) is essentially

$$T^{(\psi)}(\tilde{B}_0\tilde{D}_0): \tilde{D}_0N^2 = \tilde{B}_0M^4 + \tilde{B}_0\tilde{B}_1v\tilde{D}_1M^2e^2 + \tilde{B}_1{\tilde{D}_1}^2e^4. \tag{3.3}$$

Thus, we have

$$S^{(\psi)}(\widehat{E}_D/\mathbb{Q}) = \{\tilde{B}_0\tilde{D}_0\mathbb{Q}^{\times 2} : (3.3) \text{ is everywhere locally solvable}\}. \tag{3.4}$$

To give a description for $S^{(\phi)}(E_D/\mathbb{Q})$, we suppose first that $a^2 - 4b = bu^2$, where $a = bv$ as before, u is a positive integer. From (1.1), we know $(b, u^2) = 1$ or 2. With this notation, we similarly have homogeneous spaces $T^{(\phi)}(\bar{B}_0u_0\bar{D}_0)$ contained in $W(E/\mathbb{Q})$:

$$T^{(\phi)}(\bar{B}_0u_0\bar{D}_0): \bar{D}_0X^2 = \bar{B}_0u_0Y^4 - 2\bar{B}_0\bar{B}_1v\bar{D}_1Y^2Z^2 + \bar{B}_1u_0{u_1}^2{\bar{D}_1}^2Z^4, \tag{3.5}$$

where $D = \bar{D}_0\bar{D}_1$, $b = \bar{B}_0\bar{B}_1$ and $u = u_0u_1$. In case $(b, u) = 2$, to avoid repetition, we may set either $\bar{B}_0$ or u_0 always be odd. (Note here u_0 is squarefree and $\bar{B}_0$ could be negative.) With this convention, we note that, because of (2.11), $T^{(\phi)}(\bar{B}_0u_0\bar{D}_0)$ doesn't have a nontrivial solution in $\mathbb{Q}_p$ for every odd $p \mid u_0$. Hence, for a everywhere locally solvable homogeneous space given by (3.5), the odd part of u_0 must be 1. In other words, in considering $S^{(\phi)}(E_D/\mathbb{Q})$, we only need consider the everywhere locally solvable homogeneous spaces

$$T^{(\phi)}(\bar{B}_0\bar{D}_0): \bar{D}_0X^2 = 2^k\bar{B}_0Y^4 - 2\bar{B}_0\bar{B}_1v\bar{D}_1Y^2Z^2 + 2^{-k}\bar{B}_1u^2{\bar{D}_1}^2Z^4, \tag{3.6}$$

where k takes possible values 0 and 1 only if b is odd and u is even, otherwise k is always 0. We thus conclude that

$$S^{(\phi)}(E_D/\mathbb{Q}) = \{2^k\bar{B}_0\bar{D}_0\mathbb{Q}^{\times 2} : \ (3.6) \text{ is everywhere locally solvable}\}. \tag{3.7}$$

To measure $\#S^{(\phi)}(E_D/\mathbb{Q})\#S^{(\psi)}(\widehat{E}_D/\mathbb{Q})$, we shall count the numbers of homogeneous spaces given by (3.3) and (3.5) respectively that have a nontrivial solution in every $\mathbb{Q}_p$, including $\mathbb{Q}_\infty = \mathbb{R}$. We suppose in (3.3) and (3.5) that $(\tilde{D}_0, \bar{D}_0) = D_0$ and $(\tilde{B}_0, \bar{B}_0) = B_0$.

*Without any confusion, we abuse the 2-isogeny notation. For instance, the map ψ here is the 2-isogeny from $\widehat{E}_D$ to E_D.

We write $\tilde{D}_0 = D_0D_1,\ \bar{D}_0 = D_0D_2,\ D = D_0D_1D_2D_3,\ \tilde{B}_0 = B_0B_1,\ \bar{B}_0 = B_0B_2,\ b = B_0B_1B_2B_3$. Here all D_j's are positive integers. B_0 is positive, but B_1, B_2 and B_3 could be negative. With this reformulation, (3.3) and (3.6) then respectively become

$$D_0D_1N^2 = B_0B_1M^4 + B_0B_1B_2B_3vD_2D_3M^2e^2 + B_2B_3(D_2D_3)^2e^4 \tag{3.8}$$

and

$$D_0D_2X^2 = 2^kB_0B_2Y^4 - 2B_0B_1B_2B_3vD_1D_3Y^2Z^2 + 2^{-k}B_1B_3u^2(D_1D_3)^2Z^4. \tag{3.9}$$

Hence we have

$$\begin{aligned}&\#S^{(\phi)}(E_D/\mathbb{Q})\#S^{(\psi)}(\widehat{E}_D/\mathbb{Q})\\ &\quad = \#\left\{\begin{array}{l} b = B_0B_1B_2B_3,\ D = D_0D_1D_2D_3,\ k = 0, 1: \\ (3.8)\text{ and }(3.9)\text{ are everywhere locally solvable}\end{array}\right\}.\end{aligned} \tag{3.10}$$

Note $\{1\cdot\mathbb{Q}^{\times 2}, b\cdot\mathbb{Q}^{\times 2}\}$ is a subgroup of both $S^{(\phi)}(E_D/\mathbb{Q})$ and $S^{(\psi)}(\widehat{E}_D/\mathbb{Q})$, thus, in case b is even, half of the elements $b_1\cdot\mathbb{Q}^{\times 2}$ of $S^{(\phi)}(E_D/\mathbb{Q})$ (also of $S^{(\psi)}(\widehat{E}_D/\mathbb{Q})$) are with b_1 even. Therefore, by only considering these $b_1\cdot\mathbb{Q}^{\times 2}$ in both $S^{(\phi)}(E_D/\mathbb{Q})$ and $S^{(\psi)}(\widehat{E}_D/\mathbb{Q})$, we have, if b is even,

$$\begin{aligned}&\#S^{(\phi)}(E_D/\mathbb{Q})\#S^{(\psi)}(\widehat{E}_D/\mathbb{Q})\\ &\quad = 4\cdot\#\left\{\begin{array}{l} b = B_0B_1B_2B_3,\ D = D_0D_1D_2D_3,\ 2\mid B_0: \\ (3.8)\text{ and }(3.9)\text{ with } k = 0\text{ are everywhere} \\ \text{locally solvable}\end{array}\right\}.\end{aligned} \tag{3.10$'$}$$

We now discuss the necessary conditions for (3.8) and (3.9) to have solutions in every local field. Without loss of generality, we can assume that v is positive. (In case v is negative, the discussion is similar and comes up with a same result.) In the following, p always stands for an odd prime. For an integer, by n' we denote the odd part of n, i.e., the largest (positive) odd integer that divides n.

I. $p \mid B_0$. From (3.8) and (3.9), we note that p has to satisfy

$$\left(\frac{B_2B_3D_0D_1}{p}\right) = \left(\frac{2^kB_1B_3D_0D_2}{p}\right) = 1. \tag{3.11}$$

Therefore, for (3.8) and (3.9) to both be solvable in every $\mathbb{Q}_p$ for $p \mid B_0$, we need

$$\prod_{p\mid B_0}\frac{1}{4}\left(1+\left(\frac{B_2B_3D_0D_1}{p}\right)\right)\left(1+\left(\frac{2^kB_1B_3D_0D_2}{p}\right)\right) = 1, \tag{3.12}$$

which is equivalent to

$$4^{-\omega(B_0')}\sum_{B_0'=B_{0,1}B_{0,2}B_{0,3}B_{0,4}}\left(\frac{B_2B_3D_0D_1}{B_{0,2}}\right)\left(\frac{2^kB_1B_3D_0D_2}{B_{0,3}}\right)\left(\frac{2^kB_1B_2D_1D_2}{B_{0,4}}\right) = 1, \tag{3.13}$$

where $\omega(n)$, as usual, denotes the number of distince prime divisors of n.

We remark that the above sum takes value 0 or 1, thus serves as a character function for those B_0 satisfying condition (3.11).

II. $p \mid B_1$. Similarly, from (3.8) and (3.9), we deduce that

$$\left(\frac{B_2B_3D_0D_1}{p}\right) = \left(\frac{2^kB_0B_2D_0D_2}{p}\right) = 1, \tag{3.14}$$

thus we need

$$\prod_{p|B_1'} \frac{1}{4}\left(1+\left(\frac{B_2B_3D_0D_1}{p}\right)\right)\left(1+\left(\frac{2^kB_0B_2D_0D_2}{p}\right)\right) = 1, \tag{3.15}$$

or, equivalently,

$$4^{-\omega(B_1')} \sum_{B_1'=B_{1,1}B_{1,2}B_{1,3}B_{1,4}} \left(\frac{B_2B_3D_0D_1}{B_{1,2}}\right)\left(\frac{2^kB_0B_2D_0D_2}{B_{1,3}}\right)\left(\frac{2^kB_0B_3D_1D_2}{B_{1,4}}\right) = 1. \tag{3.16}$$

III. $p \mid B_2$. (3.8) and (3.9) respectively imply that

$$\left(\frac{B_0B_1D_0D_1}{p}\right) = 1 \quad \text{and} \quad \left(\frac{2^kB_1B_3D_0D_2}{p}\right) = 1. \tag{3.17}$$

Thus, we require the condition attached to B_2

$$\prod_{p|B_2} \frac{1}{4}\left(1+\left(\frac{B_0B_1D_0D_1}{p}\right)\right)\left(1+\left(\frac{2^kB_1B_3D_0D_2}{p}\right)\right) = 1, \tag{3.18}$$

or, equivalently,

$$4^{-\omega(B_2')} \sum_{B_2'=B_{2,1}B_{2,2}B_{2,3}B_{2,4}} \left(\frac{B_0B_1D_0D_1}{B_{2,2}}\right)\left(\frac{2^kB_1B_3D_0D_2}{B_{2,3}}\right)\left(\frac{2^kB_0B_3D_1D_2}{B_{2,4}}\right) = 1. \tag{3.19}$$

IV. $p \mid B_3$. Similarly, by checking the solvablity in $\mathbb{Q}_p$ of (3.8) and (3.9) for every $p \mid B_3$, we require the condition attached to B_3

$$4^{-\omega(B_3')} \sum_{B_3'=B_{3,1}B_{3,2}B_{3,3}B_{3,4}} \left(\frac{B_0B_1D_0D_1}{B_{3,2}}\right)\left(\frac{2^kB_0B_2D_0D_2}{B_{3,3}}\right)\left(\frac{2^kB_1B_2D_1D_2}{B_{3,4}}\right) = 1. \tag{3.20}$$

V. $p \mid D_0$. First we note that (3.8) needs to have a solution modulo p. In other words,

$$\left(B_0B_1M^2+\frac{aD_2D_3e^2}{2}\right)^2-\frac{(aD_2D_3)^2e^4}{4}+b(D_2D_3)^2e^4\equiv 0 \pmod p \tag{3.21}$$

is solvable. Thus, a^2-4b is a quadratic residue modulo p, which is equivalent to saying that

$$\left(\frac{b}{p}\right)=1 \tag{3.22}$$

since a^2-4b and b differ by a perfect square factor. Suppose $\sqrt{b}$ stands for any square root of b modulo p, then the left-hand side of (3.21) becomes

$$\left(B_0B_1M^2+\frac{(a+\sqrt{b}u)D_2D_3e^2}{2}\right)\left(B_0B_1M^2+\frac{(a-\sqrt{b}u)D_2D_3e^2}{2}\right). \tag{3.23}$$

We note that, under the condition (3.22),

$$\left(\frac{2B_0B_1D_2D_3(-a+\sqrt{b}u)}{p}\right)\left(\frac{2B_0B_1D_2D_3(-a-\sqrt{b}u)}{p}\right)=1, \tag{3.24}$$

thus, no matter how we choose the square root $\sqrt{b}$ of b modulo p, in addition to (3.22), we also need

$$\left(\frac{-2B_0B_1D_2D_3(a+\sqrt{b}u)}{p}\right)=1. \tag{3.25}$$

Now, by considering (3.9) modulo p, we see that

$$(2^kB_0B_2Y^2-aD_1D_3Z^2)^2-4b(D_1D_3)^2Z^4\equiv 0 \pmod p \tag{3.26}$$

should be solvable in $(Y,Z)\in(\mathbb{Z}/p\mathbb{Z})^{\times 2}$. With p satisfying (3.22), the left-hand side of (3.26) splits into the product of two factors $2^kB_0B_2Y^2-(a\pm 2\sqrt{b})D_1D_3Z^2$. Thus, either one of the factors is reducible modulo p. Still, note that

$$\begin{aligned}&\left(\frac{2^kB_0B_2D_1D_3(a+2\sqrt{b})}{p}\right)\left(\frac{2^kB_0B_2D_1D_3(a-2\sqrt{b})}{p}\right)\\&\quad=\left(\frac{a^2-4b}{p}\right)=\left(\frac{bu^2}{p}\right)=1,\end{aligned} \tag{3.27}$$

thus, no matter how we choose the square root of b modulo p, for (3.9) to have a nontrivial solution in $\mathbb{Z}_p$, we need

$$\left(\frac{2^k(a+2\sqrt{b})B_0B_2D_1D_3}{p}\right)=1. \tag{3.28}$$

In conclusion, for every prime divisor p of D_0, p should satisfy (3.22), (3.25) and (3.28). We simply translate this as

D_0 completely splits in $\mathbb{Q}(\sqrt{b})$ and for $\Psi:=a+\sqrt{b}u$, $\Omega:=a+2\sqrt{b}$, the sum

$$4^{-\omega(D_0)} \sum_{D_0=D_{0,1}D_{0,2}D_{0,3}D_{0,4}} \left(\frac{-2\Psi B_0B_1D_2D_3}{D_{0,2}}\right)\left(\frac{2^k\Omega B_0B_2D_1D_3}{D_{0,3}}\right)\times$$
$$\times \left(\frac{-2^{1+k}\Psi\Omega B_1B_2D_1D_2}{D_{0,4}}\right) \text{ is equal to 1.} \tag{3.29}$$

VI. $p \mid D_1$. Similar to the discussion for $p \mid D_0$, from the solvability of (3.8) modulo p, we need

$$\left(\frac{b}{p}\right) = \left(\frac{-2(a+\sqrt{b}u)B_0B_1D_2D_3}{p}\right) = 1, \tag{3.30}$$

where the choice of $\sqrt{b}$ is still not important. For (3.9) to have a solution modulo p, it is simple to see that we need

$$\left(\frac{2^kB_0B_2D_0D_2}{p}\right) = 1 \quad \text{or} \quad \left(\frac{2^kB_1B_3D_0D_2}{p}\right) = 1. \tag{3.31}$$

Note under the condition $(b/p) = 1$, the two Jacobi symbols in (3.31) always take the same value, thus we summerize the restriction on D_1 as

D_1 completely splits in $\mathbb{Q}(\sqrt{b})$ and for $\Psi := a + \sqrt{b}u$, the sum

$$4^{-\omega(D_1)} \sum_{D_1=D_{1,1}D_{1,2}D_{1,3}D_{1,4}} \left(\frac{-2\Psi B_0B_1D_2D_3}{D_{1,2}}\right)\left(\frac{2^kB_0B_2D_0D_2}{D_{1,3}}\right)\times$$
$$\times \left(\frac{-2^{1+k}\Psi B_1B_2D_0D_3}{D_{1,4}}\right) \text{ is equal to 1.} \tag{3.32}$$

VII. $p \mid D_2$. Discussing the solvability of (3.9) modulo p, we have

$$\left(\frac{b}{p}\right) = \left(\frac{2^k(a+2\sqrt{b})B_0B_2D_1D_3}{p}\right) = 1, \tag{3.33}$$

where the choice of $\sqrt{b}$ is not important. Under the condition that $(b/p) = 1$, (3.8) is nontrivially solvable modulo p if and only if

$$\left(\frac{B_0B_1D_0D_1}{p}\right) = 1. \tag{3.34}$$

Similar to that for D_0 and D_1, the restrictions on D_2 would be that D_1 completely splits in $\mathbb{Q}(\sqrt{b})$ and for $\Omega := a + 2\sqrt{b}$, the sum

$$4^{-\omega(D_2)} \sum_{D_2=D_{2,1}D_{2,2}D_{2,3}D_{2,4}} \left(\frac{B_0B_1D_0D_1}{D_{2,2}}\right)\left(\frac{2^k\Omega B_0B_2D_1D_3}{D_{2,3}}\right)\times$$
$$\times \left(\frac{2^k\Omega B_1B_2D_0D_3}{D_{2,4}}\right) \text{ is equal to 1.} \tag{3.35}$$

VIII. $p \mid D_3$. For a prime divisor p of D_3, for (3.8) to have a solution modulo p, one needs

$$\left(\frac{B_0B_1D_0D_1}{p}\right)=1 \quad \text{or} \quad \left(\frac{B_2B_3D_0D_1}{p}\right)=1. \tag{3.36}$$

Similarly, for (3.9) to have a solution modulo p, one needs

$$\left(\frac{2^kB_0B_2D_0D_2}{p}\right)=1 \quad \text{or} \quad \left(\frac{2^kB_1B_3D_0D_2}{p}\right)=1. \tag{3.37}$$

We note that the product of the two Jacobi symbols in (3.36) is equal to (b/p). Thus, if $(b/p)=-1$, then (3.36) is automatically satisfied since the two Jacobi symbols take opposite signs; if $(b/p)=1$, then we simplify (3.36) as $(B_0B_1D_0D_1/p)=1$. We also have exactly the same situation in (3.37). Thus, the restrictions on every prime divisor p of D_3 are

$$\left(\frac{b}{p}\right)=\left(\frac{B_0B_1D_0D_1}{p}\right)=\left(\frac{2^kB_0B_2D_0D_2}{p}\right)=1 \quad \text{or} \quad \left(\frac{b}{p}\right)=-1. \tag{3.38}$$

In other words, this is equivalent to

$$\frac{1}{2}\left(1-\left(\frac{b}{p}\right)\right)+\frac{1}{8}\left(1+\left(\frac{b}{p}\right)\right)\left(1+\left(\frac{B_0B_1D_0D_1}{p}\right)\right)\times \\ \times\left(1+\left(\frac{2^kB_0B_2D_0D_2}{p}\right)\right)=1. \tag{3.39}$$

Thus the restriction attached to D_3 is that the product of the left-hand side of (3.39) over all prime divisors of D_3 be 1. If multiplied out, the product turns out to be the sum

$$8^{-\omega(D_3)}\sum_{D_3=D_{3,1}D_{3,2}\cdots D_{3,8}}5^{\omega(D_{3,1})}(-3)^{\omega(D_{3,2})}\left(\frac{b}{D_{3,2}}\right)\left(\frac{B_0B_1D_0D_1}{D_{3,3}}\right)\left(\frac{2^kB_0B_2D_0D_2}{D_{3,4}}\right)\times \\ \times\left(\frac{2^kB_1B_2D_1D_2}{D_{3,5}}\right)\left(\frac{B_2B_3D_0D_1}{D_{3,6}}\right)\left(\frac{2^kB_1B_3D_0D_2}{D_{3,7}}\right)\left(\frac{2^kB_0B_3D_1D_2}{D_{3,8}}\right). \tag{3.40}$$

Corresponding to any integer that completely splits over $\mathbb{Q}(\sqrt{b})$, we formally define $\Psi:=a+\sqrt{b}u$ and $\Omega:=a+2\sqrt{b}$. With the above discussions on the local solvability, we have

LEMMA 3.1. *Suppose all the conditions given in Theorem* 2.2*, then we have*

$$\sum_{D\in S(X;h)}\#S^{(\phi)}(E_D/\mathbb{Q})\cdot\#S^{(\psi)}(\widehat{E}_D/\mathbb{Q})\leqslant S_h(X). \tag{3.41}$$

Here $S_h(X)$ *is defined as*

$$S_h(X):=\sum_{k,\vec{b}}f(k,\vec{b})\sum_{\substack{D\in S(X;h)\\ D=\vec{D}}}g(k,\vec{b},\vec{D})h(\vec{D}), \tag{3.42}$$

where, in the sum, k takes value 0 *or* 1 *if b is odd and u is even and is always equal to* 0 *otherwise, where the notation $\vec{b}$ stands for the factorizations*

$$b = B_0B_1B_2B_3,\ \ B_0 > 0 \quad \textit{and} \quad B_i' = \prod_{j=1}^{4} B_{i,j} \quad \textit{for } i = 0, 1, 2, 3,$$

and $\vec{D}$ represents the factorizations $D = D_0D_1D_2D_3$ where $D_0D_1D_2$ completely splits over $\mathbb{Q}(\sqrt{b})$ and

$$D_i = \prod_{j=1}^{4} D_{i,j} \quad \textit{for } i = 0, 1, 2 \quad \textit{and} \quad D_3 = \prod_{j=1}^{8} D_{3,j},$$

and where

$$f(k, \vec{b}) = \left(\frac{1}{4}\right)^{\omega(b')} \prod_{i=0}^{3} \left(\frac{2^k}{B_{i,3}B_{i,4}}\right) \left\{ \prod_{\substack{j=1 \\ \{s,t\}=\{1,2,3\}\setminus\{j\}}}^{3} \left(\frac{B_sB_t}{B_{0,j+1}B_{j,j+1}}\right)\left(\frac{B_0B_j}{B_{s,j+1}B_{t,j+1}}\right) \right\}, \tag{3.43}$$

where

$$\begin{aligned} g(k, \vec{b}, \vec{D}) = {} & \left\{ \prod_{i=0}^{3} \left(\frac{D_0}{B_{i,2}B_{i,3}}\right)\left(\frac{D_1}{B_{i,2}B_{i,4}}\right)\left(\frac{D_2}{B_{i,3}B_{i,4}}\right) \right\} \left(\frac{-2\Psi}{D_{0,2}D_{0,4}D_{1,2}D_{1,4}}\right) \times \\ & \times \left\{ \prod_{i=0}^{2} \left(\frac{2^k}{D_{i,3}D_{i,4}}\right) \right\} \left(\frac{2^k}{D_{3,4}D_{3,5}D_{3,7}D_{3,8}}\right)\left(\frac{b}{D_{3,2}}\right) \times \\ & \times \left(\frac{\Omega}{D_{0,3}D_{0,4}D_{2,3}D_{2,4}}\right)\left(\frac{B_0B_1}{D_{0,2}D_{1,2}D_{2,2}D_{3,3}}\right)\left(\frac{B_0B_2}{D_{0,3}D_{1,3}D_{2,3}D_{3,4}}\right) \times \\ & \times \left(\frac{B_1B_2}{D_{0,4}D_{1,4}D_{2,4}D_{3,5}}\right)\left(\frac{B_2B_3}{D_{3,6}}\right)\left(\frac{B_1B_3}{D_{3,7}}\right)\left(\frac{B_0B_3}{D_{3,8}}\right), \end{aligned} \tag{3.44}$$

and where

$$\begin{aligned} h(\vec{D}) = {} & \left(\frac{1}{4}\right)^{\omega(D_0D_1D_2)} \left(\frac{1}{8}\right)^{\omega(D_3)} 5^{\omega(D_{3,1})}(-3)^{\omega(D_{3,2})} \left(\frac{D_0D_1}{D_{2,2}D_{3,3}D_{3,6}}\right)\left(\frac{D_2D_3}{D_{0,2}D_{1,2}}\right) \times \\ & \times \left(\frac{D_0D_2}{D_{1,3}D_{3,4}D_{3,7}}\right)\left(\frac{D_1D_3}{D_{0,3}D_{2,3}}\right)\left(\frac{D_1D_2}{D_{0,4}D_{3,5}D_{3,8}}\right)\left(\frac{D_0D_3}{D_{1,4}D_{2,4}}\right). \end{aligned} \tag{3.45}$$

Moreover, if b is even, (3.10′) *combining with* (3.41)–(3.45) *implies*

$$\sum_{D \in S(X;h)} \#S^{(\phi)}(E_D/\mathbb{Q}) \cdot \#S^{(\psi)}(\widehat{E}_D/\mathbb{Q}) \leqslant \widetilde{S}_h(X). \tag{3.46}$$

Here $\widetilde{S}_h(X)$ is defined as

$$\widetilde{S}_h(X) := 4\sum_{\vec{b}} f(k,\vec{b}) \sum_{\substack{D\in S(X;h)\\ D=\vec{D}}} g(k,\vec{b},\vec{D})h(\vec{D}), \tag{3.42}$$

where the functions f, g and h are given by (3.43), (3.44) *and* (3.45) *with* $k=0$ *and, in the sum, where the notation* $\vec{b}$ *stands for the factorizations*

$$b = B_0B_1B_2B_3,\ \ B_0>0,\ \ 2\mid B_0 \quad \textit{and} \quad B_i' = \prod_{j=1}^{4} B_{i,j} \quad \textit{for } i=0,1,2,3.$$

4. Two Lemmas on Character Sum Estimates

In this section we prove two lemmas about character sum estimates that will be frequently referred to in the next sections.

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

$$S := \sum_{m,n} a_m b_n \Big(\frac{m}{n}\Big),$$

where the summation is subject to

$$M < m \leqslant 2M, \quad N < n \leqslant 2N, \quad mn \leqslant X \quad \textit{and} \quad mn \equiv h \,(\mathrm{mod}\ q).$$

Then we have

$$S \ll MN^{\frac{15}{16}+\epsilon} + M^{\frac{15}{16}+\epsilon}N, \tag{4.1}$$

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

Proof. This is essentially Lemma 4 of [4], proved based on the work of Burgess [2]. To sketch a proof, we first write S as

$$S = \sum_{\substack{i,j \,(\mathrm{mod}\ q)\\ ij\equiv h\,(\mathrm{mod}\ q)}} \sum_{m\equiv i\,(\mathrm{mod}\ q)} \sum_{n\equiv j\,(\mathrm{mod}\ q)} a_m b_n \Big(\frac{m}{n}\Big).$$

Then it is clear that we only need to prove (4.1) with ϵ replaced by $2\epsilon/3$ and S replaced by a similar sum with the restriction $mn \equiv h \,(\mathrm{mod}\ q)$ discarded. (One can set $a_m = 0$ if $m \not\equiv i \,(\mathrm{mod}\ q)$, similarly for b_n.) Note m and n are symmetric because of the quadratic reciprocity law, we thus can assume $M \geqslant N$, without loss of generality. By Cauchy's inequality, we have

$$|S|^2 \ll M \sum_{m\leqslant M} \sum_{n_1\leqslant N} \sum_{n_2\leqslant N} b_{n_1}\overline{b_{n_2}} \Big(\frac{m}{n_1n_2}\Big). \tag{4.2}$$

We now recall a special case ($r=2$) of the Theorem 2 of Burgess [2], which states that, if χ is a nonprincipal character modulo k, then for any integer N and positive integer H, then for any $\eta>0$

$$\sum_{n=N+1}^{N+H} \chi(n) \ll H^{1/2}k^{3/16+\eta}, \tag{4.3}$$

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

Note the terms in (4.2) with n_1n_2 being a perfect square contribute $O(M^2N^{1+\epsilon})$. For the other terms, note then $(\cdot/n_1n_2)$ gives a non-principal character modulo n_1n_2, so by replacing η in (4.3) by $2\epsilon/3$, we have an upper bound

$$\ll M\sum_{n_1}\sum_{n_2} M^{\frac{1}{2}}(n_1n_2)^{\frac{3}{16}+\frac{2\epsilon}{3}} \ll M^{\frac{3}{2}}N^{\frac{19}{8}+\frac{4\epsilon}{3}} \ll M^2N^{\frac{15}{8}+\frac{4\epsilon}{3}},$$

which, together with the contribution from the diagonal terms and (4.2), gives the desired result. □

LEMMA 4.2. *Suppose s is a fixed rational number. Let N be sufficiently large. Then for arbitrary positive integers q, r and any nonprincipal character $\chi(\mathrm{mod}\ q)$, we have*

$$\sum_{n\leqslant x,(n,r)=1} \mu^2(n)s^{\omega(n)}\chi(n) \ll x\tau(r)\exp(-\eta\sqrt{\log x}) \tag{4.4}$$

with a positive constant $\eta = \eta_{s,N}$, uniformly for $q \leqslant \log^N x$ and where $\tau(\cdot) = \tau_2(\cdot)$ is the usual divisor function, with $\tau_k(n)$ being the number of representations of n as the product of k ordered positive integers.

Proof. Without loss of generality, we suppose $s \neq 0$ and write

$$|s| = \frac{t}{d} \quad \text{with } t, d \in \mathbb{N} \quad \text{and} \quad (t, d) = 1. \tag{4.5}$$

For z on the half plane $\mathrm{Re}(z) > 1$, let

$$f(z) := \sum_{\substack{n=1\\(n,r)=1}}^{\infty} \frac{\mu^2(n)s^{\omega(n)}\chi(n)}{n^z}. \tag{4.6}$$

For $f(z)$, we have the Euler product expansion

$$f(z) = \prod_{p|r}\left(1+\frac{s\chi(p)}{p^z}\right),$$

thus

$$\begin{aligned}(f(z))^d &= \prod_{p|r}\left(1+\frac{s\chi(p)}{p^z}\right)^{-d}\prod_{p}\left(1+\frac{\mathrm{d}s\chi(p)}{p^z}+\frac{d(d-1)s^2\chi^2(p)}{2p^{2z}}+\cdots\right)\\ &= \prod_{p|r}\left(1+\frac{s\chi(p)}{p^z}\right)^{-d}(L(z,\chi))^{\mathrm{d}s}G(z,\chi),\end{aligned} \tag{4.7}$$

where $G(z, \chi)$, depending on s, is analytic, nonvanishing and absolutely convergent on the half-plane $\mathrm{Re}(z) > \frac{3}{4}$. On the right of $\mathrm{Re}(z) = \frac{3}{4}$, in a zero-free region of $L(z, \chi)$, we have a proper analytic branch of $(L(z, \chi))^s(G(z, \chi))^{\frac{1}{d}}$. Hence, in this region, we simply note $f(z)$ as, from (4.7),

$$f(z) = \prod_{p|r} \left(1 + \frac{\chi(p)}{sp^z}\right)^{-1} G_0(z, \chi) L(z, \chi)^s, \tag{4.8}$$

where $G_0(z, \chi)$ is analytic and absolutely convergent on $\mathrm{Re}(z) > \frac{3}{4}$. By Perron's formula, letting $b = 1 + (\log x)^{-1}$ and $T = \exp(\sqrt{\log x})$, we get

$$\sum_{n \leqslant x, (n,r)=1} \mu^2(n) s^{-\omega(n)} \chi(n) = \frac{1}{2i\pi} \int_{b-iT}^{b+iT} f(z) \frac{x^z}{z} \, \mathrm{d}z + \mathrm{O}\left(x \exp(-0.9\sqrt{\log x})\right). \tag{4.9}$$

From Siegel's Theorem, we know that for any $\epsilon > 0$, there is a constant $c(\epsilon) > 0$ such that the possible Siegel zero β of $L(z, \chi)$ satisfies

$$\beta > 1 - c(\epsilon) q^{-\epsilon}. \tag{4.10}$$

On the other hand, there is a constant $c > 0$ such that $L(z, \chi)$ has no imaginary root in the domain

$$\mathrm{Re}(z) > \frac{c}{\log T} \quad \text{and} \quad |\mathrm{Im}(z)| \leqslant T. \tag{4.11}$$

If letting ϵ in (4.10) be $1/3N$, then we see that, for sufficiently large x, there is no root for $L(z, \chi)$ in the region (4.11). We shall consider the rectangular contour with vertices $b \pm iT$, $1 - c(\log T)^{-1} \pm iT$, where c is the coefficient in (4.11). Since $f(z)$ is analytic in the rectangle, we can bound the integral in (4.9) via bounding the integrals on the other three sides.

We note that in the rectangle, the classic estimates for $L(z, \chi)$ give us

$$|L(z, \chi)|^{\pm 1} < c' \log T \tag{4.12}$$

for some constant $c' > 0$ which depends only on N. It is easy to see that, from (4.12), the integrals on $[1 - c(\log T)^{-1} + iT, b + iT]$ and $[1 - c(\log T)^{-1} - iT, b - iT]$ contribute an error admissible for (4.4). Again, from (4.12), the integral on the line $[1 - c(\log T)^{-1} - iT, 1 - c(\log T)^{-1} + iT]$ is bounded by

$$\begin{aligned} & x^{1-c(\log x)^{-\frac{1}{2}}} (\log x)^{\frac{|s|}{2}} \prod_{p|r} \left(1 + \frac{|s|}{p^{1-c(\log x)^{-1/2}}}\right) \int_{-T}^{T} \frac{1}{|t| + c(\log T)^{-1}} \, \mathrm{d}t \\ & \quad \ll x^{1-0.9c(\log x)^{-\frac{1}{2}}} \tau(r), \end{aligned} \tag{4.13}$$

which also satisfies (4.4). □

5. Some Error Terms

With the expression in lemma 3.1, we are ready to get an asymptotic formula for $S_h(X)$ (and also $\tilde{S}_h(X)$). Instead of summing over D, we sum over the 20 new variables $D_{i,j}$, subject to the conditions that each D_{ij} is squarefree, that they are pairwise coprime and that their product D satisfies

$$D \leqslant X, \qquad D \equiv h \pmod{C_E}, \tag{5.1}$$

and

$$\prod_{i=0}^{2}\prod_{j=1}^{4} D_{i,j} \text{ completely splits over } \mathbb{Q}(\sqrt{b}). \tag{5.2}$$

To make it more convenient in treating the error terms, we divide the range of each $D_{i,j}$ into dyadic intervals $(A_{i,j}, 2A_{i,j}]$ with $A_{i,j}$ running over powers of 2 and $1 \ll \prod A_{i,j} \ll X$. This gives us $O(\log^{20} X)$ nonempty subsums, each written as $S_h(\vec{A})$, $\vec{A}$ referring to the 20-tuple of numbers $A_{i,j}$. Further, we shall with a brief notation $S_h(k, \vec{b}, \vec{A})$ define the sum of $g(k, \vec{b}, \vec{D})h(\vec{D})$ with k, $\vec{b}$ fixed and the $D_{i,j}$'s running over the $A_{i,j}$'s.

We borrow the terminology from [4], two variables m and n are called 'linked' if exactly one of the Jacobi symbols (m/n) and (n/m) occurs in $f(k, \vec{b})g(k, \vec{b}, \vec{D})h(\vec{D})$, 'joined' if both of the Jacobi symbols occur in the summand. Furthermore, we call m and n 'independent' if neither of the Jacobi symbols occurs.

As we expect that the main term of the asymptotic formula for $S_h(X)$ (and $\tilde{S}_h(X)$) is of magnitude X, we treat every subsum of $S_h(X)$ (and $\tilde{S}_h(X)$) of order $O(X(\log X)^{-c})$ for any $c > 0$ as an error term. If for some linked pair $D_{i,j}$ and $D_{s,t}$ we have $A_{i,j}$, $A_{s,t}$ both greater than $(\log X)^{610}$, say, then by Lemma 4.1, the corresponding $S_h(\vec{A})$ is trivially bounded by

$$\begin{aligned}
&\sum_{\substack{\vec{D}'=\vec{D}\setminus\{D_{i,j},D_{s,t}\}\\ D' \ll X/(A_{i,j}A_{s,t})}} A_{i,j}A_{s,t}(\log X)^{-38} \\
&\quad \ll A_{i,j}A_{s,t}(\log X)^{-38} \sum_{D' \ll X/(A_{i,j}A_{s,t})} \tau_{18}(D') \\
&\quad \ll X(\log X)^{-21},
\end{aligned} \tag{5.3}$$

which, summed over $\vec{A}$, gives a negligible contribution to $S_h(X)$.

We henceforth set

$$M := (\log X)^{610} \quad \text{and} \quad T := \exp((\log X)^{0.001}). \tag{5.4}$$

Now we suppose that, in some $S_h(k, \vec{b}, \vec{A})$, $A_{i,j} > T$, and $\{D_{s,t}\}_{(s,t)\in\Im(i,j)}$ are all the D-variables linked to $D_{i,j}$. From the estimate (5.3), we know that the subsums $S_h(k, \vec{b}, \vec{A})$ with any $A_{s,t} > M$ give rise to an error term for $S_h(X)$ (and $\tilde{S}_h(X)$). Moreover, from the fact that for comprime odd integers m and n,

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = \tfrac{1}{2}(1 + \chi_4(m) + \chi_4(n) - \chi_4(mn)) \tag{5.5}$$

where $\chi_4(\cdot)$ is the nontrivial character modulo 4, the relationship between a pair of joined variables is reduced to being independent. Thus, if assuming all the variables $D_{s,t}$ linked to $D_{i,j}$ be less than M, fixing all the variables other than $D_{i,j}$ and summing over $D_{i,j}$ first, we see that the summands for $D_{i,j}$ are characters with moduli being $O((\log X)^A)$ for some $A > 0$. If the character is nontrivial, then from Lemma 4.2, the contribution of the corresponding subsum is negligible for $S_h(X)$ (and $\tilde{S}_h(X)$).

We note that the characters $\chi_4(D_{i,j})$, the Jacobi symbols formed by $D_{i,j}$ and $D_{s,t}$, and the Jocobi symbols related to 2^k and the factors of b don't annihilate each other, thus, for those $S_h(k, \vec{b}, \vec{A})$ with $A_{i,j} > T$ which finally give a major contribution to $S_h(X)$, we must have $D_{s,t} = 1$ for all $(s, t) \in \Im(i, j)$, and all the factors of b linked to $D_{i,j}$ must be 1.

We conclude the above discussion as the following lemma.

LEMMA 5.1. *There exists a constant $c > 0$, such that*

$$\sum_{k,\vec{b}} \sum_{\vec{A}} S_h(k, \vec{b}, \vec{A}) \ll X(\log X)^{-c}, \tag{5.6}$$

where the sum over $k, \vec{b}, \vec{A}$ is for all sets in which there are linked variables $D_{i,j}$ and $D_{s,t}$ with $A_{i,j} \geqslant T$, and $D_{s,t} > 1$ or there is another nontrivial integer – either 2^k or a factor of b linked with $D_{i,j}$.

Before excluding other error terms, we note that those $S_h(k, \vec{b}, \vec{A})$ with $A_{3,2} > T$ are negligible. This is because $D_{3,2}$ is linked with b. Thus, in the next, we always suppose that $A_{3,2} \leqslant T$.

We divide the sums $S_h(k, \vec{b}, \vec{A})$ into two categories: those with $A_{3,1} > T$ and those with $A_{3,1} \leqslant T$.

Case 1. $A_{3,1} \leqslant T$. We first note that, in this case, the total contribution of all the sums $S_h(k, \vec{b}, \vec{A})$ with at most 7 numbers $A_{i,j} > T$ is admissible for the error term of $S_h(X)$. To see this, suppose we have precisely m variables $D_{i,j} > T$ with $i = 0, 1, 2$, n variables $D_{3,j} > T$, and $m + n \leqslant 7$, all the other variables being $\leqslant T$. For convenience, a brief notation $t/\mathbb{Q}(\sqrt{b})$ will be used to stand for the condition that integer t completely splits in $\mathbb{Q}(\sqrt{b})$. Then we have

$$\begin{aligned}
&\sum_{k,\vec{b}} \sum_{\vec{A}} S_h(k, \vec{b}, \vec{A}) \\
&\ll_b \sum_{s \leqslant T^{20}} \tau_{20}(s) \sum_{\substack{t_1 \cdots t_m t_{m+1} \cdots t_{m+n} \leqslant X/s \\ t_1 \cdots t_m / \mathbb{Q}(\sqrt{b})}} \left(\frac{1}{4}\right)^{\omega(t_1 \cdots t_m)} \left(\frac{1}{8}\right)^{\omega(t_{m+1} \cdots t_{m+n})} \\
&\ll \sum_{s \leqslant T^{20}} \tau_{20}(s) \sum_{\substack{wz \leqslant X/s \\ w/\mathbb{Q}(\sqrt{b})}} \left(\frac{m}{4}\right)^{\omega(w)} \left(\frac{n}{8}\right)^{\omega(z)} \\
&= \sum_{s \leqslant T^{20}} \tau_{20}(s) \sum_{w_1 w_2 z \leqslant X/s} \left(\frac{m}{8}\right)^{\omega(w_1 w_2)} \left(\frac{b}{w_2}\right) \cdot \left(\frac{n}{8}\right)^{\omega(z)},
\end{aligned} \tag{5.7}$$

where the summation over k, $\vec{b}$ and $\vec{A}$ is subject to the conditions described above. In the sums we also have discarded the congruence restriction but keep the squarefree restriction on the variables. From Lemma 4.2, the subsum of the last formula in (5.7) with $w_2 > T$ is bounded by

$$\sum_{s \leqslant T^{20}} \tau_{20}(s) \sum_{w_1 z \leqslant X/(sT)} \left(\frac{m}{8}\right)^{\omega(w_1)} \left(\frac{n}{8}\right)^{\omega(z)} \frac{X}{s w_1 z} \exp(-\eta\sqrt{\log T})$$
$$\ll X \exp(-\eta\sqrt{\log T}) \sum_{s \leqslant T^{20}} \frac{\tau_{20}(s)}{s} \sum_{g \leqslant X/(sT)} \frac{1}{g} \left(\frac{m+n}{8}\right)^{\omega(g)}$$
$$\ll X \exp(-\eta\sqrt{\log T})(\log T)^{20}. \tag{5.8}$$

Thus from (5.7) and (5.8), we have

$$\sum_{k,\vec{b},\vec{A}} S_h(k, \vec{b}, \vec{A})$$
$$\ll \sum_{s \leqslant T^{21}} \tau_{21}(s) \sum_{w_1 z \leqslant X/s} \left(\frac{m}{8}\right)^{\omega(w_1)} \left(\frac{n}{8}\right)^{\omega(z)} + X \exp(-\eta\sqrt{\log T})(\log T)^{20}$$
$$\ll \sum_{s \leqslant T^{21}} \tau_{21}(s) \sum_{g \leqslant X/s} \left(\frac{m+n}{8}\right)^{\omega(g)} + X \exp(-\eta\sqrt{\log T})(\log T)^{20}$$
$$\ll \sum_{s \leqslant T^{21}} \tau_{21}(s) \cdot \frac{X}{s} (\log X)^{\frac{m+n}{8}-1} + X \exp(-\eta\sqrt{\log T})(\log T)^{20}$$
$$\ll X(\log X)^{-\frac{1}{8}} (\log T)^{21} \ll X(\log X)^{-\frac{1}{10}}, \tag{5.9}$$

which is an error term for $S_h(X)$.

Now, it is a little tedious but technically easy to examine that for any choice of 8 variables $D_{i,j}$, there are at least a pair of them linked to each other. Thus, from Lemma 5.1, the case that 8 variables $D_{i,j}$ are simultaneously greater than T is also excluded. Combining this with (5.9), we conclude the following lemma.

LEMMA 5.2. *For the constant c in Lemma* 5.1*, we have*

$$\sum_{\vec{A}} S_h(\vec{A}) \ll X(\log X)^{-c}, \tag{5.10}$$

where the sum over $\vec{A}$ is for all sets in which $A_{3,1} \leqslant T$.

Case 2. $A_{3,1} > T$. In view of Lemma 5.1, the whole contribution from those $S_h(k, \vec{b}, \vec{A})$ with $D_{0,2}D_{1,2}D_{0,3}D_{2,3}D_{1,4}D_{2,4} \neq 1$ is $O(X(\log X)^{-c})$ since $D_{3,1}$ is linked with $D_{0,2}D_{1,2}D_{0,3}D_{2,3}D_{1,4}D_{2,4}$. Excluding these error terms, we shall suppose in the following that

$$D_{0,2}D_{1,2}D_{0,3}D_{2,3}D_{1,4}D_{2,4} = 1. \tag{5.11}$$

We note that an argument similar to (5.7)–(5.9) shows that the total contribution of all the sums $S_h(\vec{A})$ with at most 3 numbers $A_{i,j} > T$ is bounded by $O(X(\log X)^{-c})$ for some constant $c > 0$, which yields an error term for $S_h(X)$. We can also check that, for any choice of 5 variables $D_{i,j}$, with $D_{3,1}$ included and those variables appearing in (5.11) excluded, there must be two of them linked to each other, or there is at least one variable $D_{i,j}$ linked to a nontrivial divisor of b. With this criterion and by

checking the 'linked relationship' between the variables, we can exclude all the cases except those listed in the following lemma.

LEMMA 5.3. *There exists a constant $c > 0$, such that*

$$\sum_{\substack{\vec{A} \\ A_{3,1} > T}} S_h(\vec{A}) \ll X(\log X)^{-c}, \tag{5.12}$$

where the sum over $\vec{A}$ is for all sets except for those in which there are precisely 4 numbers $A_{i,j} > T$, where the index (i, j) runs over the four pairs given in each of the following seven cases:

(1) (0,1), (1,1), (2,1), (3,1);
(2) (1,1), (1,3), (3,1), (3,4);
(3) (2,1), (2,2), (3,1), (3,3);
(4) (3,1), (3,3), (3,4), (3,5);
(5) (3,1), (3,4), (3,6), (3,8);
(6) (3,1), (3,5), (3,6), (3,7);
(7) (3,1), (3,3), (3,7), (3,8).

6. Proof of Theorem 2.2 for Odd *b*

Since b is odd, we have $b' = b$ throughout the section. From Lemmas 5.2 and 5.3, the main term of $S_h(X)$ comes up from seven subsums, each corresponding to one of the seven cases listed in Lemma 5.3. We write the subsums $S_h^j(X)$, $j = 1, \ldots, 7$, with $S_h^{(j)}(X)$ being the subsum that has precisely four large variables with indices described in case (j) in Lemma 5.3. Note the seven subsums don't overlap, we thus have

$$S_h(X) = \sum_{j=1}^{7} S_h^{(j)}(X) + \mathrm{O}(X(\log X)^{-c}). \tag{6.1}$$

Estimate of $S_h^{(1)}(X)$. In this case we have $D_{i,1} > T$, $i = 0, 1, 2, 3$. By checking the 'linked relationship' between variables, we see that, apart from an error term $\mathrm{O}(X(\log X)^{-c})$ by Lemma 5.1, $S_h^{(1)}(X)$ is equal to the subsum of $S_h(X)$ with $D_{i,1} > T$, $i = 0, 1, 2, 3$, $D_{3,2} \leqslant T$, all the other D-variables being 1, and $B_{i,j} = 1$ for $j = 2, 3, 4$. Thus, from Lemma 3.1, we have

$$S_h^{(1)}(X) = C_1 \sum_{\substack{D_0 D_1 D_2 D_{3,1} D_{3,2} \in S(X;h) \\ D_i > T, i=0,1,2,3 \\ D_{3,2} \leqslant T \\ D_0 D_1 D_2 / \mathbb{Q}(\sqrt{b})}} \left(\frac{1}{4}\right)^{\omega(D_0 D_1 D_2)} \left(\frac{5}{8}\right)^{\omega(D_{3,1})} \left(\frac{-3}{8}\right)^{\omega(D_{3,2})} \times$$
$$\times \left(\frac{b}{D_{3,2}}\right) + \mathrm{O}(X(\log X)^{-c}), \tag{6.2}$$

where

$$C_1 := 4^{-\omega(b')} \sum_{b=B_0B_1B_2B_3}^{*} 1, \tag{6.3}$$

where the asterisk indicates that $B_0 > 0$, and B_j, $j = 1, 2, 3$ take appropriate sign such that (3.8) and (3.9) are also solvable in $\mathbb{R}$.

By eliminating the condition $D_0D_1D_2/\mathbb{Q}(\sqrt{b})$, we see that the sum in (6.2) is equal to

$$\begin{aligned}
&\sum_{\substack{\tilde{D}D_{3,1}D_{3,2}\in S(X;h)\\ D_{3,1},\tilde{D}>T,\\ D_{3,2}\leqslant T\\ \tilde{D}/\mathbb{Q}(\sqrt{b})}} \left(\frac{3}{4}\right)^{\omega(\tilde{D})}\left(\frac{5}{8}\right)^{\omega(D_{3,1})}\left(\frac{-3}{8}\right)^{\omega(D_{3,2})}\left(\frac{b}{D_{3,2}}\right)\\
&\quad= \sum_{\substack{\tilde{D}_1\tilde{D}_2D_{3,1}D_{3,2}\in S(X;h)\\ D_{3,1},\tilde{D}_1\tilde{D}_2>T,\\ D_{3,2}\leqslant T}} \left(\frac{3}{8}\right)^{\omega(\tilde{D}_1\tilde{D}_2)}\left(\frac{b}{\tilde{D}_2}\right)\cdot\left(\frac{5}{8}\right)^{\omega(D_{3,1})}\left(\frac{-3}{8}\right)^{\omega(D_{3,2})}\left(\frac{b}{D_{3,2}}\right).
\end{aligned} \tag{6.4}$$

By Lemma 4.2, we can relax the restriction on the sizes of the variables, the resulting error for $S_h^{(1)}(X)$ being $\mathrm{O}(X\exp(-\eta\sqrt{\log T})$. With this and from (6.2) and (6.4), we have, apart from an error $\mathrm{O}(X(\log X)^{-c})$,

$$\begin{aligned}
S_h^{(1)}(X) &= C_1 \sum_{\tilde{D}_1\tilde{D}_2D_{3,1}D_{3,2}\in S(X;h)} \left(\frac{3}{8}\right)^{\omega(\tilde{D}_1\tilde{D}_2)}\left(\frac{b}{\tilde{D}_2}\right)\cdot\left(\frac{5}{8}\right)^{\omega(D_{3,1})}\left(\frac{-3}{8}\right)^{\omega(D_{3,2})}\left(\frac{b}{D_{3,2}}\right)\\
&= C_1 \sum_{\tilde{D}_1D_{3,1}\bar{D}\in S(X;h)} \left(\frac{3}{8}\right)^{\omega(\tilde{D}_1)}\left(\frac{5}{8}\right)^{\omega(D_{3,1})}\left(\frac{b}{\bar{D}}\right)\sum_{\bar{D}=\tilde{D}_2D_{3,2}}\left(\frac{-3}{8}\right)^{\omega(D_{3,2})}\left(\frac{3}{8}\right)^{\omega(\tilde{D}_2)}.
\end{aligned} \tag{6.5}$$

Since the inner sum in (6.5) about $\bar{D}$ is equal to 1 if $\bar{D} = 1$ and 0 otherwise, we thus have

$$\begin{aligned}
S_h^{(1)}(X) &= C_1 \sum_{\tilde{D}_1D_{3,1}\in S(X;h)} \left(\frac{3}{8}\right)^{\omega(\tilde{D}_1)}\left(\frac{5}{8}\right)^{\omega(D_{3,1})} + \mathrm{O}(X(\log X)^{-c})\\
&= C_1\cdot \#S(X;h) + \mathrm{O}(X(\log X)^{-c}).
\end{aligned} \tag{6.6}$$

Estimate of $S_h^{(2)}(X)$. For this subsum with $D_{1,1}$, $D_{1,3}$, $D_{3,1}$, $D_{3,4} > T$, by using Lemma 5.1 to exclude the negligible subsums with a total contribution $\mathrm{O}(X(\log X)^{-c})$, we conclude that the major term of $S_h^{(2)}(X)$ comes up with all the D-variables except $D_{1,1}$, $D_{1,3}$, $D_{3,1}$, $D_{3,2}$, $D_{3,4}$ and $D_{3,7}$ being trivial, and

$$2^k = B_0 = B_2 = B_{1,2} = B_{1,4} = B_{3,2} = B_{3,4} = 1. \tag{6.7}$$

We remark that, under the condition (6.7), $2^k B_1 B_3 = b$. Furthermore, by relaxing the restriction on the sizes of the D-variables, we simply get that

$$S_h^{(2)}(X) = C_2 \sum_{\substack{D_{1,1}D_{1,3}D_{3,1}D_{3,2}D_{3,4}D_{3,7}\in S(X;h)\\ D_{1,1}D_{1,3}/\mathbb{Q}(\sqrt{b})}} \left(\frac{1}{4}\right)^{\omega(D_{1,1}D_{1,3})} \left(\frac{5}{8}\right)^{\omega(D_{3,1})} \times$$

$$\times \left(\frac{1}{8}\right)^{\omega(D_{3,4}D_{3,7})} \left(\frac{-3}{8}\right)^{\omega(D_{3,2})} \left(\frac{b}{D_{3,2}D_{3,7}}\right) + \mathrm{O}(X(\log X)^{-c}), \tag{6.8}$$

where

$$C_2 := 4^{-\omega(b')} \sum_{\substack{b=B_1B_3, B_1, B_3\in\mathbb{Z}\\ B_1'=B_{1,1}B_{1,3}, B_3'=B_{3,1}B_{3,3}}}^{*} 1 \tag{6.9}$$

with the asterisk in the summation indicating that B_1 takes appropriate sign such that (3.8) is solvable in $\mathbb{R}$. With exactly the same method we used in (6.4)–(6.6), we have

$$S_h^{(2)}(X) = C_2 \cdot \#S(X; h) + \mathrm{O}(X(\log X)^{-c}). \tag{6.10}$$

Estimate of $S_h^{(3)}(X)$. We use Lemma 5.1 to exclude error terms, and the major term of $S_h^{(3)}(X)$ arises from the subsum with all the D-variables except $D_{2,1}$, $D_{2,2}$, $D_{3,1}$, $D_{3,2}$, $D_{3,3}$ and $D_{3,6}$ being 1, and

$$B_0 = B_1 = B_{2,3} = B_{2,4} = B_{3,3} = B_{3,4} = 1. \tag{6.11}$$

Note then $b = B_2B_3$, $B_2, B_3 \in \mathbb{Z}$ and $B_2' = B_{2,1}B_{2,2}$, $B_3' = B_{3,1}B_{3,2}$. With a discussion similar to that applied to $S_h^{(1)}(X)$, we have

$$S_h^{(3)}(X) = C_3 \cdot \#S(X; h) + \mathrm{O}(X(\log X)^{-c}), \tag{6.12}$$

where

$$C_3 := \sum_k 4^{-\omega(b')} \sum_{\substack{b=B_2B_3\\ B_2,B_3\in\mathbb{Z}\\ B_2'=B_{2,1}B_{2,2}\\ B_3'=B_{3,1}B_{3,2}}}^{*} 1. \tag{6.13}$$

Again, the asterisk in (6.13) indicates that the summation is also subject to that (3.9) be solvable in $\mathbb{R}$.

Estimate of $S_h^{(j)}(X)$, $j = 4, 5, 6, 7$. The four subsums $S_h^{(j)}(X)$, $j = 4, 5, 6, 7$ are in the same shape. By a transform of variables, they are exactly equal to each other. Because of this, we can only estimate $S_h^{(4)}(X)$. Again, by using Lemma 5.1 to exclude an error of $\mathrm{O}(X(\log X)^{-c})$, we see that the main term of $S_h^{(4)}(X)$ arises from the subsum with all the D-variables except $D_{3,j}$, $j = 1, \ldots, 8$ being 1 and

$$2^k = B_0 = B_1 = B_2 = 1. \tag{6.14}$$

Under the condition (6.14), we see that $b = B_3$, thus we have

$$S_h^{(4)}(X) = 4^{-\omega(b')}\left\{\sum_{b'=\prod_{j=1}^4 B_{3,j}} 1\right\} \cdot \sum_{\prod_{j=1}^8 D_{3,j}\in S(X;h)} \left(\frac{5}{8}\right)^{\omega(D_{3,1})}\left(\frac{-3}{8}\right)^{\omega(D_{3,2})} \times$$
$$\times \left(\frac{1}{8}\right)^{\omega(D_{3,2}\cdots D_{3,8})}\left(\frac{b}{D_{3,2}D_{3,6}D_{3,7}D_{3,8}}\right) + \mathrm{O}(X(\log X)^{-c})$$
$$= \#S(X;h) + \mathrm{O}(X(\log X)^{-c}). \tag{6.15}$$

Hence, for $j = 4, 5, 6, 7$, we have

$$S_h^{(j)}(X) = \#S(X;h) + \mathrm{O}(X(\log X)^{-c}). \tag{6.16}$$

Collecting the asymptotic formulas (6.6), (6.10), (6.12) and (6.16) together and noticing (6.1), we thus have

$$S_h(X) = (C_1 + C_2 + C_3 + 4)\cdot \#S(X;h) + \mathrm{O}(X(\log X)^{-c}). \tag{6.17}$$

From (6.3), we see that

$$C_1 = \begin{cases} 2, & \text{if } v \text{ is odd,} \\ 4, & \text{if } v \text{ is even.} \end{cases} \tag{6.18}$$

And (6.9) implies that

$$C_2 = \begin{cases} 2, & \text{if } v \text{ is positive,} \\ 1, & \text{if } v \text{ is negative.} \end{cases} \tag{6.19}$$

Also from (6.13) we have

$$C_3 = \begin{cases} 1, & \text{if } v \text{ is odd and positive,} \\ 2, & \text{if } v \text{ is odd and negative,} \\ 2, & \text{if } v \text{ is even and positive,} \\ 4, & \text{if } v \text{ is even and negative.} \end{cases} \tag{6.20}$$

From (6.17)–(6.20), we conclude that, in case b is odd,

$$S_h(X) \leqslant (13 + o(1))\#S(X;h), \tag{6.21}$$

which proves Theorem 2.2 for odd b.

7. Proof of Theorem 2.2 for Even *b*

We shall simply sketch a proof for the case that b is even. We start from (3.46). First we note that all the estimates about the error terms in Section 5 are also valid for $\widetilde{S}_h(X)$. Thus the leading terms of $\widetilde{S}_h(X)$ come up from the cases listed in Lemma 5.3. Moreover, in the summation for $\widetilde{S}_h(X)$, B_0 is always nontrivial (since it is divisible by 2). Thus, for a major contribution, a subsum must have all the D-variables linked with B_0 running over small intervals. For this reason, we see that, among all the seven cases listed in Lemma 5.3, all but cases (1) and (6) actually make a

contribution at most $O(X(\log X)^{-c})$ for $\widetilde{S}_h(X)$. Therefore, with a discussion similar to that for $S_h^{(1)}(X)$ and $S_h^{(6)}(X)$ in section 6, we have

$$\widetilde{S}_h(X) = 4 \cdot (\widetilde{C}_1 + 1) \cdot \#S(X; h) + O(X(\log X)^{-c}), \tag{7.1}$$

where

$$\widetilde{C}_1 := 4^{-\omega(b')} \sum_{\substack{b=B_0B_1B_2B_3 \\ B_0>0,\ 2|B_0}}^{*} 1, \tag{7.2}$$

with the asterisk indicating that the factorization also guarantees the solvability of (3.8) and (3.9) in $\mathbb{R}$. It is easy to see that $\widetilde{C}_1 = 2$, thus we have

$$\widetilde{S}_h(X) = 12 \cdot \#S(X; h) + O(X(\log X)^{-c}), \tag{7.3}$$

which proves Theorem 2.2 for even b.

8. Some Further Remarks

We note that the conditions (1.1) and (1.2) have restricted the curves in consideration to a very small family. One may expect that the method works for a larger family of elliptic curves with a rational 2-torsion point. Based on our discussion (3.30)–(3.40) concerning the solvability in $\mathbb{Q}_p$ for $p \mid D$, however, we can see that such a restriction is crucial and, for a curve E (with a rational 2-torsion point) other than those given in Theorem 2.2, the average size of $\#S^{(\phi)}(E_D/\mathbb{Q}) \cdot \#S^{(\psi)}(\widehat{E}_D/\mathbb{Q})$ would be too large – it is unbounded. Actually, we can prove that, for such a curve E, the average size of $\#S^{(\phi)}(E_D/\mathbb{Q}) \cdot \#S^{(\psi)}(\widehat{E}_D/\mathbb{Q})$ with $|D| \leqslant X$ would have order of magnitude $(\log X)^{\frac{1}{16}}$. Without an actual proof, one can see that with the following heuristic: with the condition $b \equiv a^2 - 4b (\mathrm{mod}\ \mathbb{Q}^{\times})^2$ (that is only used in (3.27)) being removed, the probability for (3.8) and (3.9) to have a nontrivial solution in $\mathbb{Q}_p$ for $p \mid D_0$ changes from $\frac{1}{8}$ to $\frac{3}{16}$, which results in the change of the factor that brings up the major term of sum (2.12) from $1^{\omega(D)}$ to $(\frac{17}{16})^{\omega(D)}$. Therefore, to prove (1.3) for such a curve, one has to take the (negative) contribution of the Tate–Shafarevich groups into consideration, and this would be another work for further study.

We also remark that Theorem 2.2 yields the following conditional result.

THEOREM 8.1. *Suppose E is an elliptic curve satisfying the conditions* (1.1) *and* (1.2). *Assuming the parity conjecture for the Mordell–Weil ranks, we have $M_E^1(X) \gg X$ for sufficiently large X.*

Acknowledgements

The author is grateful to Professors Carl Pomerance and Trevor Wooley for their encouragement, to Professor Frank Lemmermeyer for clarifying the conjecture involved in his paper [11], and to the referee for many helpful comments.

References

1. Bruinier, J. H., James, K., Kohnen, W., Ono, K., Skinner, C. and Vatsal, V.: Congruence properties of values of L-functions and applications, *Topics in Number Theory* (*University Park, PA, 1997*), Math. Appl. 467, Kluwer Acad. Publ., Dordrecht, 1999, pp. 115–125.
2. Burgess, D. A.: On character sums and L-series, II, *Proc. London Math. Soc.* (*3*) **13** (1963), 524–536.
3. Goldfeld, D.: Conjectures on elliptic curves over quadratic fields, In: *Number Theory,* (*Carbondale 1979*), Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979, Lecture Notes in Math. 751, Springer, Berlin, 1979, pp. 108–118.
4. Heath-Brown, D. R.: The size of Selmer groups for the congruent number problem I, *Invent. Math.* **111**(1) (1993), 171–195.
5. Heath-Brown, D. R.: The size of Selmer groups for the congruent number problem II, *Invent. Math.* **118**(2) (1994), 331–370.
6. Iwaniec, H. and Sarnak, P.: The non-vanishing of central values of automorphic L-functions and Landau–Siegel zeros, *Israel J. Math.* **120** (2000), 155–177.
7. James, K.: An example of an elliptic curve with a positive density of prime quadratic twists which have rank zero, In: *Topics in Number Theory* (*University Park, PA, 1997*), Math. Appl. 467, Kluwer Acad. Publ., Dordrecht, 1999, pp. 223–227.
8. James, K.: L-series with non-zero central critical value, *J. Amer. Math. Soc.* **11**(3) (1998), 635–641.
9. Kohnen, W.: On the proportion of quadratic character twists of L-functions attached to cusp forms not vanishing at the central point, *J. Reine Angew. Math.* **508** (1999), 179–187.
10. Kolyvagin, V. A.: Finiteness of $E(\mathbb{I})$ and the Tate–Shafarevich group of $E(\mathbb{Q})$ for a subclass of Weil curves, *Izv. Akad. Nauk SSSR Ser. Mat* **52** (1988), 522–540, 670–671.
11. Lemmermeyer, F.: On Tate-Shafarevich groups of some elliptic curves, In: *Algebraic Number Theory and Diophantine Analysis* (*Graz 1998*), pp. 277–291.
12. Monsky, P.: Generalizing the Birch-Stephens theorem. I. Modular curves, *Math. Z.* **221**(3) (1996), 415–420.
13. Ono, K. Rank zero quadratic twists of modular elliptic curves, *Compositio Math.* **104**(3) (1996), 293–304.
14. Ono, K.: Twists of elliptic curves, *Compositio Math.* **106**(3) (1997), 349–360.
15. Ono, K.: Nonvanishing of quadratic twists of modular L-functions and applications to elliptic curves, *J. Reine Angew. Math.* **533** (2001), 81–97.
16. Ono, K. and Skinner, C.: Fourier coefficients of half-integral weight modular forms modulo ℓ, *Ann. of Math. (2)* **147**(2) (1998), 453–470.
17. Ono, K. and Skinner, C.: Non-vanishing of quadratic twists of modular L-functions, *Invent. Math.* **134**(3) (1998), 651–660.
18. Schmitt, S.: Computation of the Selmer groups of certain parametrized elliptic curves, *Acta Arith.* **78** (1997), 241–254.
19. Shimura, G.: On modular forms of half integral weight, *Ann. of Math. (2)* **97** (1973), 440–481.

20. Silverman, J.: *The Arithmetic of Elliptic Curves*, Grad. Text Math. 106, Springer, New York, 1986.
21. Vatsal, V.: Canonical periods and congruence formulae, *Duke Math. J.* **98**(2) (1999), 397–419.
22. Vatsal, V.: Rank one twists of a certain elliptic curve, *Math Ann.* **311** (1998), 791–794.
23. Waldspurger, J. L. Sur les coefficients de Fourier des formes modulaires de poids demi-entier, *J. Math. Pures. Appl.* **60** (1981), 375–484.
24. Wong, S.: Elliptic curves and class number divisibility, *Internat. Math. Res. Notices* No. 12 (1999), 661–672.
25. Yu, G.: On the quadratic twists of a family of elliptic curves, Preprint.