

RANKIN-SELBERG L -FUNCTIONS AND THE REDUCTION OF CM ELLIPTIC CURVES

SHENG-CHI LIU, RIAD MASRI, AND MATTHEW P. YOUNG

ABSTRACT. Let q be a prime and $K = \mathbb{Q}(\sqrt{-D})$ be an imaginary quadratic field such that q is inert in K . If \mathfrak{q} is a prime above q in the Hilbert class field of K , there is a reduction map

$$r_{\mathfrak{q}} : \mathcal{E}ll(\mathcal{O}_K) \longrightarrow \mathcal{E}ll^{ss}(\mathbb{F}_{q^2})$$

from the set of elliptic curves over $\overline{\mathbb{Q}}$ with complex multiplication by the ring of integers \mathcal{O}_K to the set of supersingular elliptic curves over \mathbb{F}_{q^2} . We prove a uniform asymptotic formula for the number of CM elliptic curves which reduce to a given supersingular elliptic curve and use this result to deduce that the reduction map is surjective for $D \gg_{\varepsilon} q^{18+\varepsilon}$. This can be viewed as an analog of Linnik's theorem on the least prime in an arithmetic progression.

We also use related ideas to prove a uniform asymptotic formula for the average

$$\sum_{\chi} L(f \times \Theta_{\chi}, 1/2)$$

of central values of the Rankin-Selberg L -functions $L(f \times \Theta_{\chi}, s)$ where f is a fixed weight 2, level q arithmetically normalized Hecke cusp form and Θ_{χ} varies over the weight 1, level D theta series associated to an ideal class group character χ of K . We apply this result to study the arithmetic of abelian varieties, subconvexity, and L^4 norms of automorphic forms.

1. INTRODUCTION AND STATEMENT OF RESULTS

A problem of great importance in number theory concerns the distribution of primes in primitive residue classes. Given a modulus q and a primitive residue class $a \pmod{q}$, let

$$\pi(x; a, q) = \#\{p \leq x : p \equiv a \pmod{q}\}$$

be the number of primes $p \leq x$ in this residue class. Dirichlet's theorem on primes in arithmetic progressions shows that $\pi(x; a, q) > 0$ for x sufficiently large (in terms of q). It is then natural to ask how large x must be to ensure the existence of such a prime $p \leq x$. A remarkable theorem of Linnik asserts that there exists an absolute constant $L \geq 2$ such that $\pi(x; a, q) > 0$ for $x \gg q^L$. Here L is the famous Linnik "constant." A great amount of effort has been devoted to producing successively smaller numerical values of L . For an extensive discussion of Linnik's theorem, see Chapter 18 of [IK]. Analogs of Linnik's theorem for other arithmetic structures have captured the attention of many mathematicians.

2010 *Mathematics Subject Classification.* 11M41.

Key words and phrases. supersingular elliptic curves, equidistribution, Gross points, Heegner points, mean values of L -functions, L^4 norm.

R. M. and M. Y. were supported by the National Science Foundation under agreement Nos. DMS-1162535 (R. M.) and DMS-1101261 (M. Y.). Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

In this paper we will study an analog of Linnik's theorem for the reduction of CM elliptic curves. Let $\mathcal{E}ll(\mathcal{O}_K)$ be the set of isomorphism classes of elliptic curves over $\overline{\mathbb{Q}}$ with complex multiplication by the ring of integers \mathcal{O}_K of an imaginary quadratic field $K = \mathbb{Q}(\sqrt{-D})$. By the theory of complex multiplication, these curves are defined over the Hilbert class field H_K of K and the Galois group $G_K = \text{Gal}(H_K/K)$ acts simply transitively on $\mathcal{E}ll(\mathcal{O}_K)$. Therefore given a curve $E \in \mathcal{E}ll(\mathcal{O}_K)$, we have $\mathcal{E}ll(\mathcal{O}_K) = \{E^\sigma\}_{\sigma \in G_K}$ where E^σ denotes the Galois action on E , and there are $|G_K| = h(-D)$ such curves where $h(-D)$ is the class number of K . Let q be a prime number and \mathfrak{q} be a prime above q in H_K . If q is inert or ramified in K , the curve E has supersingular reduction modulo \mathfrak{q} . Let $\mathcal{E}ll^{ss}(\mathbb{F}_{q^2}) = \{E_1, \dots, E_n\}$ be the set of isomorphism classes of supersingular elliptic curves defined over \mathbb{F}_{q^2} . Then one has a reduction map

$$r_{\mathfrak{q}} : \mathcal{E}ll(\mathcal{O}_K) \longrightarrow \mathcal{E}ll^{ss}(\mathbb{F}_{q^2}).$$

There is a probability measure on $\mathcal{E}ll^{ss}(\mathbb{F}_{q^2})$ defined by

$$(1.1) \quad \mu_q(E_i) = \frac{w_i^{-1}}{\sum_{j=1}^n w_j^{-1}}$$

where $w_i \in \{1, 2, 3\}$ is the number of units modulo $\{\pm 1\}$ of the endomorphism ring of E_i . Gross [G, Table 1.3] notes that $\prod_{i=1}^n w_i$ divides 6 for q odd, so all but at most two w_i equal 1. Eichler's mass formula states that

$$(1.2) \quad \sum_{j=1}^n w_j^{-1} = \frac{q-1}{12},$$

and thus $\mu_q(E_i) \asymp q^{-1}$.

Results of Gross [G], Iwaniec [I1], and Duke [D1] imply that for fixed q , the points $r_{\mathfrak{q}}(\mathcal{E}ll(\mathcal{O}_K))$ become equidistributed among the isomorphism classes $\mathcal{E}ll^{ss}(\mathbb{F}_{q^2})$ with respect to the measure μ_q as $D \rightarrow \infty$. In fact, this equidistribution holds in the following strong quantitative form: if q is inert in K , then given a curve $E_i \in \mathcal{E}ll^{ss}(\mathbb{F}_{q^2})$ one has

$$(1.3) \quad \frac{\#\{\sigma \in G_K : r_{\mathfrak{q}}(E^\sigma) = E_i\}}{h(-D)} = \mu_q(E_i) + O_q(D^{-\delta})$$

for some absolute $\delta > 0$, where the implied constant is uniform in E_i and ineffective. See also the work of Elkies, Ono and Yang [EOY], which is discussed in more detail below. Michel [Mi, Theorem 3] proved a "sparse" equidistribution version of (1.3), where G_K can be replaced by any subgroup $G < G_K$ of index $\leq D^{1/2115}$. Related equidistribution problems were studied in [Co], [JeK], [V].

The equidistribution result (1.3) implies that for D sufficiently large (in terms of q), the reduction map $r_{\mathfrak{q}}$ is surjective. In analogy with Linnik's theorem, it is then natural to ask how large D must be to ensure the surjectivity. We will answer this question by proving the following (stronger) result concerning the asymptotic distribution of the integers

$$(1.4) \quad N_{D,q,E_i} := \#\{\sigma \in G_K : r_{\mathfrak{q}}(E^\sigma) = E_i\}$$

as $q, D \rightarrow \infty$.

Theorem 1.1. *Let q be an odd prime and $-D < 0$ be a fundamental discriminant such that q is inert in K . Then*

$$(1.5) \quad N_{D,q,E_i} = h(-D)\mu_q(E_i) + O(q^{1/8+\varepsilon}D^{7/16+\varepsilon}),$$

where the implied constant is uniform in E_i and ineffective. Furthermore, we have

$$(1.6) \quad N_{D,q,E_i} \ll D^{1/4+\varepsilon} + \frac{D^{1/2+\varepsilon}}{q^{1/2}}.$$

Assuming the generalized Lindelöf hypothesis for quadratic twists of modular L -functions, we have

$$(1.7) \quad N_{D,q,E_i} = h(-D)\mu_q(E_i) + O(q^\varepsilon D^{1/4+\varepsilon}).$$

The implied constants in (1.6) and (1.7) are effective.

Remark 1.2. Combining (1.5) with Siegel's (ineffective) bound $h(-D) \gg_\varepsilon D^{1/2-\varepsilon}$ immediately yields that

$$(1.8) \quad N_{D,q,E_i} \sim \mu_q(E_i)h(-D)$$

as $q, D \rightarrow \infty$ with the restriction $D \gg_\varepsilon q^{18+\varepsilon}$. For D smaller, with $q^{10} \ll D \ll q^{18+\varepsilon}$, then (1.5) becomes

$$(1.9) \quad N_{D,q,E_i} \ll q^{1/8} D^{7/16+\varepsilon}.$$

In the range $D \ll q^{10}$, the bound (1.6) is strongest.

Theorem 1.1 implies the following

Corollary 1.3. *The reduction map*

$$r_q : \mathcal{E}ll(\mathcal{O}_K) \longrightarrow \mathcal{E}ll^{ss}(\mathbb{F}_{q^2})$$

is surjective for $D \gg_\varepsilon q^{18+\varepsilon}$. Assuming the generalized Lindelöf hypothesis for quadratic twists of modular L -functions, the reduction map is surjective for $D \gg_\varepsilon q^{4+\varepsilon}$.

Elkies, Ono and Yang [EOY] proved that the reduction map is surjective (for q inert or ramified) for all sufficiently large D by relating the integers N_{D,q,E_i} to Fourier coefficients of a weight $3/2$ theta series of level $4q$ studied by Gross [G] and employing results of Iwaniec [I1] and Duke [D1] to give a lower bound for these coefficients. Under the assumption of the generalized Riemann hypothesis, Kane [K1] [K2] effectivized the results of [EOY] and proved that the reduction map is surjective for $D \gg_\varepsilon q^{14+\varepsilon}$. A careful reader may wonder why, assuming Lindelöf, our exponent improves on that of Kane, even though both approaches use essentially optimal bounds on L -functions. The difference lies in the first steps of the proof. Kane uses the theta function framework mentioned above. He decomposes the theta function into a Hecke basis and uses the Kohnen-Zagier [KZ] formula to bound the Fourier coefficients of the weight $3/2$ Hecke eigenforms in the basis. The basis coefficients in this decomposition are then difficult to bound optimally. In contrast, we use a period formula of Gross to directly relate the integers N_{D,q,E_i} to central values of L -functions summed over a weight 2, level q Hecke eigenbasis (see (4.4), (3.1) and (4.7) below). Here the L -values are of degree 4, factoring as the product of two degree 2 L -functions, in contrast to the degree 2 L -function appearing in the Kohnen-Zagier formula. It is straightforward to bound the basis coefficients in this decomposition (see the sentence following (4.7)).

A somewhat complementary question concerning “minimal” CM lifts was studied by Yang in [Y].

The problems studied in Theorem 1.1 are closely related to certain Rankin-Selberg L -functions of arithmetic significance. Let f be an arithmetically normalized Hecke cusp form

of weight 2 and level q . Let Θ_χ be the weight 1 theta series of level D associated to a character χ of the ideal class group $Cl(K)$. We will prove the following uniform asymptotic formula for averages of central values of the Rankin-Selberg L -functions $L(f \times \Theta_\chi, s)$ as χ varies over the ideal class group characters.

Theorem 1.4. *Let $q \neq 2$ be a prime and $-D < 0$ be a fundamental discriminant such that q is inert in K . Then*

$$(1.10) \quad \sum_{\chi \in \widehat{Cl(K)}} L(f \times \Theta_\chi, \tfrac{1}{2}) = \frac{3}{2\pi^3 u^2} \frac{h(-D)^2}{\sqrt{D}} \frac{q}{q-1} L(\text{Sym}^2 f, 1) \\ + O_\varepsilon((qD)^\varepsilon \min(q^{7/8} D^{7/16}, q^{3/4} D^{1/4} + q^{1/4} D^{1/2})).$$

Theorem 1.4 implies the following quantitative nonvanishing result.

Corollary 1.5. *For each $\varepsilon > 0$ there is an ineffective constant $c = c(\varepsilon) > 0$ such that whenever $D \geq cq^{14+\varepsilon}$, we have $L(f \times \Theta_\chi, 1/2) \neq 0$ for some $\chi \in \widehat{Cl(K)}$.*

Remark 1.6. Corollary 1.5 can be used to study the arithmetic of abelian varieties. For example, if A_f denotes the abelian variety associated to f by the Eichler-Shimura construction, a result of Bertolini and Darmon [BD, Theorem B] implies that if $L(f \times \Theta_\chi, 1/2) \neq 0$, the χ -isotypical component $A_f(H_K)^\chi$ of the Mordell-Weil group $A_f(H_K)$ is finite.

Since the conductor of $L(f \times \Theta_\chi, s)$ at $s = 1/2$ is $Q = (qD)^2$, the convexity bound is $L(f \times \Theta_\chi, \frac{1}{2}) \ll_\varepsilon Q^{1/4+\varepsilon}$. These central values are nonnegative, so dropping all but one term in Theorem 1.4 implies the following hybrid subconvexity bound.

Corollary 1.7. *We have*

$$(1.11) \quad L(f \times \Theta_\chi, \tfrac{1}{2}) \ll_\varepsilon (qD)^\varepsilon (q^{3/4} D^{1/4} + q^{1/4} D^{1/2}).$$

A short computation shows this bound is subconvex when $q \asymp D^\eta$ for $0 < \eta < 1$. Michel and Ramakrishnan [MR] were the first to study hybrid subconvexity of the L -functions $L(f \times \Theta_\chi, s)$. By establishing an *exact* formula for the average of $L(f \times \Theta_\chi, 1/2)$ over holomorphic forms f (the opposite average from Theorem 1.4), they also deduced subconvexity for $0 < \eta < 1$. Hybrid subconvexity bounds for more general families of Rankin-Selberg L -functions have been obtained by various methods (see e.g. [FW1], [HoMu], [N], [HoTe]).

For $D \gg q^{14+\varepsilon}$, Theorem 1.4 gives

$$(1.12) \quad L(f \times \Theta_\chi, \tfrac{1}{2}) \ll q^{7/8} D^{7/16+\varepsilon},$$

which is $o(D^{1/2})$ in this range of uniformity. For applications to equidistribution and nonvanishing, it is crucial to have a bound that is subconvex in D alone with q fixed. On the other hand, the estimate (1.9) (which is clearly of an arithmetical nature) uses much of the same technology used to *prove* the hybrid subconvexity bound, so these problems are closely related. A powerful input into the estimate (1.12) is a uniform subconvexity bound of Blomer and Harcos [BH] for $L(f \times \chi_{-D}, 1/2)$, an L -function on $GL_2 \times GL_1$.

We next give a brief outline of the proof of Theorem 1.4 (see Section 2 for more details). In [G, Section 3], Gross used the arithmetic of definite quaternion algebras to define a certain curve X_q which is the disjoint union of n curves Y_i of genus 0 over \mathbb{Q} . Then the Picard group of the curve X_q is given by

$$\text{Pic}(X_q) = \mathbb{Z}e_1 \oplus \cdots \oplus \mathbb{Z}e_n$$

where e_i denotes the class of degree 1 in $\text{Pic}(X_q)$ corresponding to a point on Y_i . Gross [G, Proposition 11.3] established a formula relating the central value $L(f \times \Theta_\chi, 1/2)$ to a certain height pairing of two divisors in $\text{Pic}(X_q) \otimes_{\mathbb{Z}} \mathbb{C}$. We will use Gross's formula to establish an identity of the form

$$\sum_{\chi \in \widehat{\text{Cl}}(K)} L(f \times \Theta_\chi, \tfrac{1}{2}) = c \sum_{\sigma \in G_K} w_\sigma \widetilde{f}(\xi^\sigma)^2,$$

where c is an explicit constant (depending on D , q and f), the w_σ are explicit positive integers (see the sentence following (2.2) for the definition), f is a certain real-valued function on $\text{Pic}(X_q) \otimes_{\mathbb{Z}} \mathbb{R}$ in Jacquet-Langlands correspondence with f , and $\{\xi^\sigma\}_{\sigma \in G_K}$ is the G_K -orbit of a fixed Gross point ξ of discriminant $-D$. After decomposing the sum on the right hand side into a Hecke basis, we are led to estimating an expression of the form

$$\sum_{g \in \mathcal{F}_2(q)} (\widetilde{f}, \widetilde{f}, \widetilde{g}) W_{D, \widetilde{g}},$$

where $\mathcal{F}_2(q)$ is an orthogonal basis of arithmetically normalized Hecke cusp forms of weight 2 and level q , (\cdot, \cdot, \cdot) is a certain trilinear form, and $W_{D, \widetilde{g}}$ is the Weyl sum

$$(1.13) \quad W_{D, \widetilde{g}} := \sum_{\sigma \in G_K} w_\sigma \widetilde{g}(\xi^\sigma).$$

A formula of Gross and Kudla [GK] formula relates $(\widetilde{f}, \widetilde{f}, \widetilde{g})^2$ to the triple product L -function $L(f \times f \times g, 1/2) = L(\text{Sym}^2 f \times g, 1/2)L(g, 1/2)$, while Gross's formula relates $|W_{D, \widetilde{g}}|^2$ to $L(g, 1/2)L(g \times \chi_{-D}, 1/2)$. After an application of Hölder's inequality, we are led to estimating the averages

$$\sum_{g \in \mathcal{F}_2(q)} \frac{L(g \times \chi_{-D}, \tfrac{1}{2})}{L(\text{Sym}^2 g, 1)}, \quad \sum_{g \in \mathcal{F}_2(q)} \frac{L(g, \tfrac{1}{2})^4}{L(\text{Sym}^2 g, 1)},$$

$$\sum_{g \in \mathcal{F}_2(q)} \frac{L(\text{Sym}^2 f \times g, \tfrac{1}{2})^2}{L(\text{Sym}^2 g, 1)}.$$

We estimate the latter two averages using the large sieve inequality for holomorphic cusp forms, along with some deep results from the automorphy of Rankin-Selberg convolutions. To estimate the average with $L(g \times \chi_{-D}, 1/2)$, if q is very small compared to D we apply a hybrid subconvexity bound of Blomer and Harcos [BH], while if q is somewhat large we use the following result which is of independent interest.

Theorem 1.8. *Let q be an odd prime and D be a fundamental discriminant with $(D, q) = 1$. Then*

$$(1.14) \quad \sum_{f \in \mathcal{F}_2(q)} \omega_f L(f \times \chi_D, \tfrac{1}{2}) = 1 + O_\varepsilon((q|D|)^\varepsilon q^{-1}|D|^{1/2}),$$

where $\omega_f = q^{-1+o(1)}$ are the weights occurring in the Petersson trace formula.

Theorem 1.8 is the analog of [LMY, Theorem 1.5] (which considered Maass forms instead of holomorphic forms). Duke [D2] was the first to study first and second moments of this type, though he did not consider the dependence on D . The case of even weight k , $k \geq 4$, could probably be derived with some more refined estimates from work of Jackson and Knightly

[JK] or Kohnen and Sengupta [KoSe], but there are convergence problems in both of their approaches for $k = 2$.

Corollary 1.9. *If $q \gg |D|^{1/2+\delta}$, there exists $f \in \mathcal{F}_2(q)$ with $L(f \times \chi_D, \frac{1}{2}) \neq 0$.*

The opposite problem of first choosing f and then finding D such that $L(1/2, f \times \chi_D) \neq 0$ was studied by Hoffstein and Kontorovich [HK], who showed that this holds with some $|D| \ll q^{1+\varepsilon}$. Such a result can also be derived from Waldspurger’s formula and Riemann-Roch, though this latter method does not give a lower bound on the central value while the moment method of [HK] does (as does Corollary 1.9). In relation to Corollaries 1.5 and 1.9, Michel and Ramakrishnan [MR, Theorem 2] showed that if $q \gg D^{1/2+\delta}$, then there exists an f such that $L(f \times \Theta_\chi, 1/2) \neq 0$. When χ is chosen to be the trivial class group character, then $L(f \times \Theta_\chi, s) = L(f, s)L(f \times \chi_{-D}, s)$ so Michel and Ramakrishnan’s result implies simultaneous nonvanishing.

Corollary 1.10. *We have*

$$(1.15) \quad L(f \times \chi_D, \tfrac{1}{2}) \ll (q + |D|^{1/2})(q|D|)^\varepsilon.$$

Remark 1.11. The bound (1.15) is subconvex for $q^{3/2+\delta} \leq |D| \leq q^A$ with fixed $\delta, A > 0$.

The following result is a bound on the L^4 norm of the real-valued function \tilde{f} on $\text{Pic}(X_q) \otimes_{\mathbb{Z}} \mathbb{R}$ in Jacquet-Langlands correspondence with f described above (see Section 2 below for a more thorough explanation). This is an analog of [LMY, Proposition 1.7] which is a bound on the L^4 norm of a Maass form in the level aspect.

Proposition 1.12. *Suppose \tilde{f} is in Jacquet-Langlands correspondence with $f \in \mathcal{F}_2(q)$, normalized so $\langle \tilde{f}, \tilde{f} \rangle = 1$ (equivalently, $\sum_{i=1}^n w_i |\tilde{f}(e_i)|^2 = 1$). Then we have*

$$(1.16) \quad \|\tilde{f}\|_4^4 := \sum_{i=1}^n w_i^3 |\tilde{f}(e_i)|^4 \ll q^{-1/2+\varepsilon}.$$

The normalization is such that the Lindelöf Hypothesis for triple product L -functions in the level aspect would give $O(q^{-1+\varepsilon})$ as the bound in (1.16). Blomer and Michel [BM] have shown $\|\tilde{f}\|_\infty \ll q^{-\frac{1}{60}+\varepsilon}$, which is the $\lambda = 0$ case in their Theorem 1; see their Remark 1.1.

We conclude the introduction by discussing how our results relate to some existing work. Our analysis is influenced by the beautiful paper of Michel and Venkatesh [MV], where they emphasize the “period formula” approach to asymptotics for families of Rankin-Selberg L -functions. In particular, for fixed q , they give an asymptotic formula for the average in Theorem 1.4 as $D \rightarrow \infty$ using the equidistribution of Gross points. They also discuss the possibility of a more refined analysis which would yield some range of uniformity in q (see [MV, Remark 3.1]).

We proved the analog of Theorem 1.4 for Hecke-Maass newforms in [LMY, Theorem 1.1] and studied level-aspect versions of equidistribution of Heegner points. Although the basic idea of using period formulas to pass to averages of families of L -functions is common to both papers, the methods used here differ in significant ways. For example, we mentioned that Theorem 1.8 is the holomorphic analog of [LMY, Theorem 1.5]. We emphasize that the holomorphic case (with weight 2) has a new analytic difficulty because of slow convergence of the sum of Kloosterman sums in the Petersson formula.

Lastly, we take this opportunity to make a correction and improvement to [LMY]. In [LMY, Corollary 1.3] the word “effective” should be replaced with “ineffective.” Furthermore, in the same corollary, the range of q can be extended to $q \leq cD^{1/14-\varepsilon}$, matching the exponent in Corollary 1.5 here. This improved exponent arises from applying [BH, Theorem 2] in place of [BH, (1.3)] in [LMY, (6.9)] (and subsequent bounds relying on (6.9)). We thank Jack Buttcane for this observation.

2. A FORMULA OF GROSS

In this section we review a period formula of Gross [G], following closely the discussion in [G, Section 3] and [Mi, Section 6]. Let $-D < 0$ be a fundamental discriminant and q be a prime which is inert in $K = \mathbb{Q}(\sqrt{-D})$. Let $Cl(K)$ be the ideal class group, $h(-D)$ be the class number, $\widehat{Cl(K)}$ be the group of ideal class group characters, and \mathcal{O}_K be the ring of integers of K , respectively. Let B be the quaternion algebra over \mathbb{Q} which is ramified at q and ∞ . Fix a maximal order R in B , and let $\{I_1, \dots, I_n\}$ be a set of representatives for the equivalence classes of left R -ideals in B . To each I_i , one associates the maximal right order

$$R_i = \{x \in B : I_i x \subset I_i\}.$$

An optimal embedding of \mathcal{O}_K into R_i is an embedding $\xi : K \hookrightarrow B$ for which $\xi(K) \cap R_i = \xi(\mathcal{O}_K)$. Two optimal embeddings ξ_1 and ξ_2 are conjugate modulo R_i^\times if there is a unit $u \in R_i^\times$ such that $\xi_1(x) = u\xi_2(x)u^{-1}$ for all $x \in \mathcal{O}_K$. A *Gross point* of discriminant $-D$ is an optimal embedding ξ of \mathcal{O}_K into some R_i , modulo conjugation by R_i^\times . Let $h(\mathcal{O}_K, R_i)$ denote the number of R_i^\times -conjugacy classes of optimal embeddings of \mathcal{O}_K into R_i . Then a result of Eichler states that (see [G, eq. (1.12)])

$$\sum_{i=1}^n h(\mathcal{O}_K, R_i) = 2h(-D).$$

In particular, if $\Lambda_{D,q}$ denotes the set of Gross points of discriminant $-D$, we have $\#\Lambda_{D,q} = 2h(-D)$.

The set of left R -ideals $\{I_1, \dots, I_n\}$ corresponds to the set of connected components $\{Y_1, \dots, Y_n\}$ of a curve X_q which is the disjoint union of n curves Y_i of genus 0 over \mathbb{Q} . The Gross points $\Lambda_{D,q}$ can be described geometrically as certain K -valued points on X_q (see [G, pp. 131-132]). Let $\text{Pic}(X_q)$ denote the Picard group of X_q , and let $\text{Pic}^0(X_q)$ denote the subgroup of degree 0 divisors. If e_i denotes the class of degree 1 in $\text{Pic}(X_q)$ corresponding to a point on Y_i , we have

$$\text{Pic}(X_q) = \mathbb{Z}e_1 \oplus \dots \oplus \mathbb{Z}e_n.$$

In this way, a Gross point ξ determines a class e_{i_ξ} in $\text{Pic}(X_q)$. By abuse of notation, we also denote this class by ξ .

There is a height pairing

$$\langle \cdot, \cdot \rangle : \text{Pic}(X_q) \times \text{Pic}(X_q) \rightarrow \mathbb{Z}$$

defined on generators by $\langle e_i, e_j \rangle = w_i \delta_{ij}$, and extended bi-additively to $\text{Pic}(X_q)$. Here $w_i = |R_i^\times|/2$, which agrees with the definition of w_i given in the introduction since the set of left R -ideals $\{I_1, \dots, I_n\}$ corresponds to the set of supersingular elliptic curves $\{E_1, \dots, E_n\}$ in

such a way that $\text{End}(E_i) = R_i$ (see e.g. [Koh, Section 5.3]). We define a probability measure on the set of divisor classes $\{e_1, \dots, e_n\}$ by

$$\mu_q(e_i) = \frac{w_i^{-1}}{\sum_{j=1}^n w_j^{-1}}.$$

Let $S_2(q)$ be the space of cusp forms of weight 2 and level q , and let $\mathcal{F}_2(q)$ be an orthogonal basis of arithmetically normalized Hecke cusp forms for $S_2(q)$. A special case of the Jacquet-Langlands correspondence states that for each form $f \in \mathcal{F}_2(q)$, there is a unique $e_f \in \text{Pic}^0(X_q) \otimes_{\mathbb{Z}} \mathbb{R}$ such that $\langle e_f, e_f \rangle = 1$ and $t_n e_f = \lambda_f(n) e_f$, where t_n denotes the operator on $\text{Pic}(X_q)$ induced by the n -th Hecke correspondence on X_q and $\lambda_f(n)$ is the n -th Hecke eigenvalue of f . We write

$$e_f = \sum_{i=1}^n \nu_i(f) e_i$$

with $\nu_i(f) \in \mathbb{R}$ for $i = 1, \dots, n$. Define

$$e^* = \sum_{i=1}^n \frac{1}{w_i} e_i.$$

Then an orthonormal basis for $\text{Pic}(X_q) \otimes_{\mathbb{Z}} \mathbb{R}$ is given by

$$\left\{ \frac{e^*}{\sqrt{\langle e^*, e^* \rangle}} \right\} \cup \{e_f : f \in \mathcal{F}_2(q)\}.$$

Note that by Eichler's mass formula, $\langle e^*, e^* \rangle = (q-1)/12$.

The geometric description of the Gross points allows one to define a free action of $Cl(K) \cong G_K = \text{Gal}(H_K/K)$ on $\Lambda_{D,q}$, where H_K is the Hilbert class field of K (see [G, p. 133]). Given an ideal class group character $\chi \in \widehat{Cl(K)}$ and a Gross point $\xi \in \Lambda_{D,q}$, let

$$c_\chi = \sum_{\sigma \in G_K} \overline{\chi(\sigma)} \xi^\sigma \in \text{Pic}(X_q) \otimes_{\mathbb{Z}} \mathbb{C}.$$

Given a form $f \in \mathcal{F}_2(q)$, let $L(f \times \Theta_\chi, s)$ be the Rankin-Selberg L -function of f and the weight 1 theta series Θ_χ of level D associated to χ . Then Gross's formula states that¹ (see [G, Proposition 11.2])

$$(2.1) \quad L(f \times \Theta_\chi, \tfrac{1}{2}) = \frac{\langle f, f \rangle_q}{u^2 \sqrt{D}} |\langle c_\chi, e_f \rangle|^2,$$

where u is the number of units in K and

$$\langle f, g \rangle_q := \int_{\Gamma_0(q) \backslash \mathbb{H}} y^2 f(z) \overline{g(z)} \frac{dx dy}{y^2}$$

is the Petersson inner product on $S_2(q)$.

¹Gross's formula is actually stated as

$$L(f \times \Theta_\chi, \tfrac{1}{2}) = \frac{\langle f, f \rangle_q}{u^2 \sqrt{D}} \langle c_{f,\chi}, c_{f,\chi} \rangle,$$

where $c_{\chi,f} := \langle c_\chi, e_f \rangle e_f$ is the projection of c_χ onto the f -isotypical component in $\text{Pic}(X_q) \otimes_{\mathbb{Z}} \mathbb{C}$. With our normalization of e_f this is easily seen to be equivalent to (2.1).

We now give an alternative description of Gross's formula which will be useful for calculations. Let $M_B^{\mathbb{C}}(q)$ be the vector space of \mathbb{C} -valued functions on $\text{Pic}(X_q) \otimes_{\mathbb{Z}} \mathbb{C}$ with the inner product

$$\langle \phi, \psi \rangle := \sum_{i=1}^n w_i \phi(e_i) \overline{\psi(e_i)}.$$

Then the map which sends a generator e_i to its characteristic function $\mathbb{1}_{e_i}$ induces an isomorphism $\text{Pic}(X_q) \otimes_{\mathbb{Z}} \mathbb{C} \cong M_B^{\mathbb{C}}(q)$ defined by

$$e = \sum_{i=1}^n c_i e_i \longmapsto \tilde{e} := \sum_{i=1}^n c_i \mathbb{1}_{e_i}.$$

Moreover, this map is an isometry of inner-product spaces, i.e. $\langle \tilde{e}, \tilde{e}' \rangle = \langle e, e' \rangle$ for any $e, e' \in \text{Pic}(X_q) \otimes_{\mathbb{Z}} \mathbb{C}$. Let $\tilde{f} = \tilde{e}_f$ denote the image of e_f under this isomorphism. Then an orthonormal basis for $M_B^{\mathbb{R}}(q)$ is given by

$$\left\{ \frac{\tilde{e}^*}{\sqrt{\langle \tilde{e}^*, \tilde{e}^* \rangle}} \right\} \cup \{ \tilde{f} : f \in \mathcal{F}_2(q) \}.$$

We can now write Gross's formula as

$$(2.2) \quad L(f \times \Theta_{\chi}, \tfrac{1}{2}) = \frac{\langle f, f \rangle_q}{u^2 \sqrt{D}} \left| \sum_{\sigma \in G_K} \overline{\chi(\sigma)} w_{\sigma} \tilde{f}(\xi^{\sigma}) \right|^2,$$

where by abuse of notation we write w_{σ} for $w_i = |R_i^{\times}|/2$ where ξ^{σ} is an optimal embedding of \mathcal{O}_K into R_i .

By [ILS, Lemma 2.5 and (3.14)], we have

$$(2.3) \quad \langle f, f \rangle_q = \frac{1}{8\pi^3} q L(\text{Sym}^2 f, 1),$$

and so we may write (2.2) as

$$(2.4) \quad L(f \times \Theta_{\chi}, \tfrac{1}{2}) = \frac{q L(\text{Sym}^2 f, 1)}{8\pi^3 u^2 \sqrt{D}} \left| \sum_{\sigma \in G_K} \overline{\chi(\sigma)} w_{\sigma} \tilde{f}(\xi^{\sigma}) \right|^2.$$

3. PERIOD INTEGRAL FORMULAS AND BOUNDS ON L -FUNCTIONS

In this section we evaluate the magnitude of the Weyl sums $W_{D, \tilde{g}}$ (defined in the introduction by (1.13)) and the trilinear forms $(\tilde{f}, \tilde{f}, \tilde{g})$ in terms of L -functions. Applying (2.4) with $\chi = \chi_0$ the trivial ideal class group character, we have

$$(3.1) \quad |W_{D, \tilde{g}}|^2 = \frac{8\pi^3 u^2 \sqrt{D} L(g \times \chi_{-D}, \tfrac{1}{2}) L(g, \tfrac{1}{2})}{q L(\text{Sym}^2 g, 1)}.$$

Using the nonnegativity of $L(g \times \chi_{-D}, 1/2)$, $L(g, 1/2)$ and $L(\text{Sym}^2 g, 1)$, we deduce the following

Lemma 3.1. *For $g \in \mathcal{F}_2(q)$ we have*

$$(3.2) \quad W_{D, \tilde{g}} = \pm \frac{\sqrt{8\pi^3} u D^{1/4} L(g \times \chi_{-D}, \tfrac{1}{2})^{1/2} L(g, \tfrac{1}{2})^{1/2}}{q^{1/2} L(\text{Sym}^2 g, 1)^{1/2}}.$$

The key fact here is that subconvexity for the twisted L -function gives a nontrivial bound on $W_{D,\tilde{g}}$. The current best subconvexity bound which is uniform in q and D is the following result of Blomer and Harcos [BH]:

$$(3.3) \quad L(g \times \chi_{-D}, 1/2) \ll (q^{1/4}D^{3/8} + q^{1/2}D^{1/4})(qD)^\varepsilon.$$

The large sieve inequality for holomorphic cusp forms allows one to deduce Lindelöf on average in the following sense.

Lemma 3.2. [IK, Theorem 7.35] *We have*

$$(3.4) \quad \sum_{g \in \mathcal{F}_2(q)} \frac{L(g, \frac{1}{2})^4}{L(\text{Sym}^2 g, 1)} \ll q^{1+\varepsilon}$$

We now review a period formula of Gross and Kudla [GK] for triple product L -functions. Write the Fourier expansion for $f \in \mathcal{F}_2(q)$ in the form

$$f(z) = \sum_{n=1}^{\infty} \lambda_f(n) n^{1/2} e(nz),$$

where $\lambda_f(n)$ is the n -th Hecke eigenvalue of f . For a prime $p \neq q$, write

$$\lambda_f(p) = \alpha_{f,1}(p) + \alpha_{f,2}(p)$$

where Deligne's bound amounts to $|\alpha_{f,i}(p)| \leq 1$, and the Hecke relation means

$$\alpha_{f,1}(p)\alpha_{f,2}(p) = 1.$$

For $p = q$, $\lambda_f(q) = \pm q^{-1/2}$.

Define the triple product L -function

$$L(f \times g \times h, s) = \prod_p L_p(f \times g \times h, s),$$

where for $p \neq q$,

$$L_p(f \times g \times h, s) = \prod_{i,j,k \in \{1,2\}} (1 - \alpha_{f,i}(p)\alpha_{g,j}(p)\alpha_{h,k}(p)p^{-s})^{-1},$$

and for $p = q$,

$$L_p(f \times g \times h, s) = (1 - \lambda_f(q)\lambda_g(q)\lambda_h(q)q^{-s})^{-1}(1 - \lambda_f(q)\lambda_g(q)\lambda_h(q)q^{1-s})^{-2}.$$

Then the completed triple product L -function is defined by

$$\Lambda(f \times g \times h, s) = q^{5s/2} L_\infty(f \times g \times h, s) L(f \times g \times h, s),$$

where

$$L_\infty(f \times g \times h, s) = (2\pi)^{-(3+4s)} \Gamma(s + \frac{3}{2}) \Gamma(s + \frac{1}{2})^3.$$

This L -function satisfies the functional equation

$$\Lambda(f \times g \times h, s) = \varepsilon_{f,g,h} \Lambda(f \times g \times h, 1-s),$$

where

$$\varepsilon_{f,g,h} = \text{sgn}(\lambda_f(q)\lambda_g(q)\lambda_h(q)) = \pm 1.$$

The Gross-Kudla formula for the central value of the triple product L -function of f, g, h is (see [FW2, Theorem 3.1] and [GK, Corollary 11.3])

$$(3.5) \quad L(f \times g \times h, \tfrac{1}{2}) = 128\pi^5 \frac{\langle f, f \rangle_q \langle g, g \rangle_q \langle h, h \rangle_q}{q} (\tilde{f}, \tilde{g}, \tilde{h})^2,$$

where the trilinear form is defined by

$$(3.6) \quad (\tilde{f}, \tilde{g}, \tilde{h}) = \sum_{i=1}^n w_i^2 \nu_i(f) \nu_i(g) \nu_i(h).$$

Note that our Petersson inner product is $(8\pi^2)^{-1}$ times the Petersson inner product in [FW2, Theorem 3.1]. Using (2.3) we may write (3.5) as

$$(3.7) \quad (\tilde{f}, \tilde{g}, \tilde{h})^2 = \frac{4\pi^4}{q^2} \frac{L(f \times g \times h, \tfrac{1}{2})}{L(\text{Sym}^2 f, 1) L(\text{Sym}^2 g, 1) L(\text{Sym}^2 h, 1)}.$$

Using the factorization $L(f \times f \times g, s) = L(\text{Sym}^2 f \times g, s) L(g, s)$, we deduce

Lemma 3.3. *We have*

$$(3.8) \quad (\tilde{f}, \tilde{f}, \tilde{g}) = \pm \frac{2\pi^2 L(\text{Sym}^2 f \times g, \tfrac{1}{2})^{1/2} L(g, \tfrac{1}{2})^{1/2}}{q L(\text{Sym}^2 f, 1) L(\text{Sym}^2 g, 1)^{1/2}}.$$

4. PROOF OF THEOREM 1.1

To prove Theorem 1.1 we combine the following result with the argument in the first paragraph of [Mi, p. 226].

Theorem 4.1. *Let q be an odd prime and $-D < 0$ be a fundamental discriminant such that q is inert in $K = \mathbb{Q}(\sqrt{-D})$. Given a Gross point $\xi \in \Lambda_{D,q}$ and a class $e_i \in \text{Pic}(X_q)$, define*

$$N_{q,D,e_i} = \#\{\sigma \in G_K : \xi^\sigma = e_i\}.$$

Then

$$(4.1) \quad N_{q,D,e_i} = h(-D)\mu_q(e_i) + O(q^{1/8+\varepsilon} D^{7/16+\varepsilon}),$$

where the implied constant is uniform in e_i . We also have

$$(4.2) \quad N_{q,D,e_i} \ll (Dq)^\varepsilon (D^{1/4} + \frac{D^{1/2}}{q^{1/2}}).$$

Assuming the generalized Lindelöf hypothesis for quadratic twists of modular L -functions, we have

$$(4.3) \quad N_{D,q,E_i} = h(-D)\mu_q(E_i) + O(q^\varepsilon D^{1/4+\varepsilon}).$$

Proof. We begin by showing

$$(4.4) \quad N_{q,D,e_i} = h(-D)\mu_q(e_i) + \frac{1}{w_i} \sum_{g \in \mathcal{F}_2(q)} \langle \tilde{e}_i, \tilde{g} \rangle W_{D,\tilde{g}},$$

which is equivalent to [Mi, (6.3)]. We have

$$(4.5) \quad w_i N_{q,D,e_i} = w_i \#\{\sigma \in G_K : \xi^\sigma = e_i\} = \sum_{\sigma \in G_K} w_\sigma \tilde{e}_i(\xi^\sigma).$$

By decomposing the function \tilde{e}_i into a Hecke basis in $M_B^{\mathbb{R}}(q)$, we have

$$\tilde{e}_i(z) = \frac{\langle \tilde{e}_i, \tilde{e}^* \rangle}{\langle \tilde{e}^*, \tilde{e}^* \rangle} \tilde{e}^*(z) + \sum_{g \in \mathcal{F}_2(q)} \langle \tilde{e}_i, \tilde{g} \rangle \tilde{g}(z).$$

Therefore,

$$w_i N_{q,D,e_i} = \frac{\langle \tilde{e}_i, \tilde{e}^* \rangle}{\langle \tilde{e}^*, \tilde{e}^* \rangle} \sum_{\sigma \in G_K} w_\sigma \tilde{e}^*(\xi^\sigma) + \sum_{g \in \mathcal{F}_2(q)} \langle \tilde{e}_i, \tilde{g} \rangle W_{D,\tilde{g}}.$$

We calculate

$$\sum_{\sigma \in G_K} w_\sigma \tilde{e}^*(\xi^\sigma) = \sum_{\sigma \in G_K} w_\sigma \sum_{i=1}^n \frac{1}{w_i} \tilde{e}_i(\xi^\sigma) = h(-D).$$

Then using $\langle \tilde{e}_i, \tilde{e}^* \rangle = 1$ for all i and the Eichler mass formula $\langle \tilde{e}^*, \tilde{e}^* \rangle = (q-1)/12$, identity (4.4) follows.

Now we turn to the proof of Theorem 4.1. We may assume $q \ll D^{1/2}$ as otherwise the conclusions are trivial. By Cauchy's inequality we have

$$(4.6) \quad |N_{q,D,e_i} - h(-D)\mu_q(e_i)| \leq \frac{1}{w_i} M_1^{1/2} M_2^{1/2},$$

where

$$(4.7) \quad M_1 = \sum_{g \in \mathcal{F}_2(q)} \langle \tilde{e}_i, \tilde{g} \rangle^2, \quad M_2 = \sum_{g \in \mathcal{F}_2(q)} |W_{D,\tilde{g}}|^2.$$

We recall here that $w_i = 1, 2$ or 3 . By Bessel's inequality, $M_1 \leq \langle \tilde{e}_i, \tilde{e}_i \rangle = w_i$. To finish the proof of Theorem 4.1, we now show

$$M_2^{1/2} \ll q^{1/8} D^{7/16} (qD)^\varepsilon.$$

This follows from (3.1), the bound of Blomer-Harcos [BH, Theorem 2] $L(g \times \chi_{-D}, 1/2) \ll q^{1/4} D^{3/8} (qD)^\varepsilon$ (which uses $q \ll D^{1/2}$, as otherwise a different term is dominant, and also that $(q, D) = 1$) and the bound

$$\sum_{g \in \mathcal{F}_2(q)} \frac{L(g, \frac{1}{2})}{L(\text{Sym}^2 g, 1)} \ll q^{1+\varepsilon},$$

which is implied by Lemma 3.2. This shows (4.1).

To show (4.2), we simply quote a result of Michel and Ramakrishnan [MR, Corollary 2], which in our notation states

$$(4.8) \quad M_2 = \sum_{g \in \mathcal{F}_2(q)} |W_{D,\tilde{g}}|^2 \ll \frac{\sqrt{D}}{q} \sum_{g \in \mathcal{F}_2(q)} \frac{L(g \times \chi_{-D}, \frac{1}{2}) L(g, \frac{1}{2})}{L(\text{Sym}^2 g, 1)} \ll (qD)^\varepsilon (\sqrt{D} + \frac{D}{q}).$$

Finally, the bound (4.3) comes from using $M_2 \ll D^{1/2} (qD)^\varepsilon$ which follows from Lindelöf. \square

5. PROOF OF THEOREM 1.4

Using the orthogonality relations for the characters $\widehat{Cl(K)}$, we obtain from (2.4) the identity

$$(5.1) \quad M_f(D) := \sum_{\chi \in \widehat{Cl(K)}} L(f \times \Theta_{\chi, \frac{1}{2}}) = \frac{1}{8\pi^3 u^2} \frac{h(-D)}{\sqrt{D}} q L(\text{Sym}^2 f, 1) \sum_{\sigma \in G_K} w_{\sigma}^2 \tilde{f}(\xi^{\sigma})^2.$$

We claim that

$$(5.2) \quad \sum_{\sigma \in G_K} w_{\sigma}^2 \tilde{f}(\xi^{\sigma})^2 = \frac{h(-D)}{\frac{q-1}{12}} + \sum_{g \in \mathcal{F}_2(q)} (\tilde{f}, \tilde{f}, \tilde{g}) W_{D, \tilde{g}}.$$

Since $\tilde{f} = \sum_{i=1}^n \nu_i(f) \tilde{e}_i$, then $\tilde{f}^2 = \sum_{i=1}^n \nu_i(f)^2 \tilde{e}_i$, and so

$$\sum_{\sigma \in G_K} w_{\sigma}^2 \tilde{f}(\xi^{\sigma})^2 = \sum_{i=1}^n \nu_i(f)^2 w_i^2 N_{q, D, e_i}.$$

Therefore, from (4.4)

$$\sum_{\sigma \in G_K} w_{\sigma}^2 \tilde{f}(\xi^{\sigma})^2 = \sum_{i=1}^n \nu_i(f)^2 w_i \left(\frac{h(-D)}{\frac{q-1}{12}} + \sum_{g \in \mathcal{F}_2(q)} \langle \tilde{e}_i, \tilde{g} \rangle W_{D, \tilde{g}} \right)$$

where we used $\langle \tilde{e}^*, \tilde{e}^* \rangle = (q-1)/12$. Now, $\sum_{i=1}^n w_i \nu_i(f)^2 = \langle \tilde{f}, \tilde{f} \rangle = 1$, and $\langle \tilde{e}_i, \tilde{g} \rangle = w_i \nu_i(g)$, so

$$\sum_{\sigma \in G_K} w_{\sigma}^2 \tilde{f}(\xi^{\sigma})^2 = \frac{h(-D)}{\frac{q-1}{12}} + \sum_{g \in \mathcal{F}_2(q)} \left(\sum_{i=1}^n w_i^2 \nu_i(f)^2 \nu_i(g) \right) W_{D, \tilde{g}},$$

which is precisely (5.2).

The first term in (5.2) equals the stated main term in Theorem 1.4. For the sum over g , we have

Lemma 5.1. *We have*

$$(5.3) \quad \sum_{g \in \mathcal{F}_2(q)} (\tilde{f}, \tilde{f}, \tilde{g}) W_{D, \tilde{g}} \ll \frac{(qD)^{\varepsilon}}{L(\text{Sym}^2 f, 1)} \min \left(\frac{D^{7/16}}{q^{1/8}}, \frac{D^{1/4}}{q^{1/4}} + \frac{D^{1/2}}{q^{3/4}} \right).$$

The error term in Theorem 1.4 is q times larger than the right hand side of (5.3), as desired.

Proof. Combining (3.8) and (3.2) we obtain

$$(5.4) \quad \sum_{g \in \mathcal{F}_2(q)} (\tilde{f}, \tilde{f}, \tilde{g}) W_{D, \tilde{g}} = \frac{D^{1/4}}{q^{3/2} L(\text{Sym}^2 f, 1)} \times \sum_{g \in \mathcal{F}_2(q)} \frac{\theta_{g, f, D}}{L(\text{Sym}^2 g, 1)} L(g \times \chi_{-D}, \frac{1}{2})^{\frac{1}{2}} L(g, \frac{1}{2}) L(\text{Sym}^2 f \times g, \frac{1}{2})^{\frac{1}{2}},$$

where $\theta_{g, f, D} \ll 1$ uniformly in g, f and D . We apply Hölder's inequality with exponents 2, 4, 4, respectively, obtaining

$$(5.5) \quad \sum_{g \in \mathcal{F}_2(q)} (\tilde{f}, \tilde{f}, \tilde{g}) W_{D, \tilde{g}} \ll \frac{D^{1/4}}{q^{3/2} L(\text{Sym}^2 f, 1)} M_3^{1/2} M_4^{1/4} M_5^{1/4},$$

where we have used nonnegativity of central values, and where

$$(5.6) \quad \begin{aligned} M_3 &= \sum_{g \in \mathcal{F}_2(q)} \frac{L(g \times \chi_{-D}, \frac{1}{2})}{L(\text{Sym}^2 g, 1)}, & M_4 &= \sum_{g \in \mathcal{F}_2(q)} \frac{L(g, \frac{1}{2})^4}{L(\text{Sym}^2 g, 1)}, \\ M_5 &= \sum_{g \in \mathcal{F}_2(q)} \frac{L(\text{Sym}^2 f \times g, \frac{1}{2})^2}{L(\text{Sym}^2 g, 1)}. \end{aligned}$$

Lemma 3.2, a consequence of the large sieve inequality for holomorphic cusp forms, states $M_4 \ll q^{1+\varepsilon}$. With similar technology combined with some deep inputs on the automorphy of Rankin-Selberg convolutions, we will show in Section 6 the following

Proposition 5.2. *We have*

$$(5.7) \quad M_5 \ll q^{2+\varepsilon}.$$

For M_3 , we have two different approaches. For q small compared to D we simply apply the best known progress towards Lindelöf for $L(g \times \chi_{-D}, 1/2)$, which is (3.3) due to Blomer and Harcos, and multiply by the number of forms which is $\asymp q$. For q larger we appeal to Theorem 1.8. In all, we obtain

$$(5.8) \quad M_3 \ll (qD)^\varepsilon \min(q^{5/4} D^{3/8} + q^{3/2} D^{1/4}, q + D^{1/2}).$$

Notice that if $q \gg D^{1/2}$, then Theorem 1.8 is an asymptotic formula so the latter term in the min is the optimal choice, while if $q \ll D^{1/2}$ then the first term inside the min in (5.8) may be simplified as $O(q^{5/4} D^{3/8})$. Taking these estimates for granted, we then obtain (5.3) after a short calculation. \square

Now we discuss an alternate arrangement of Hölder's inequality which may be of interest. Applying Hölder's inequality in (5.4) with exponents 4, 4, 2, respectively, we obtain

$$(5.9) \quad \sum_{g \in \mathcal{F}_2(q)} (\tilde{f}, \tilde{f}, \tilde{g}) W_{D, \tilde{g}} \ll \frac{D^{1/4}}{q^{3/2} L(\text{Sym}^2 f, 1)} M_3'^{1/4} M_4'^{1/4} M_5'^{1/2},$$

where M_4 is given already by (5.6), and in addition

$$M_3' = \sum_{g \in \mathcal{F}_2(q)} \frac{L(g \times \chi_{-D}, \frac{1}{2})^2}{L(\text{Sym}^2 g, 1)}, \quad M_5' = \sum_{g \in \mathcal{F}_2(q)} \frac{L(\text{Sym}^2 f \times g, \frac{1}{2})}{L(\text{Sym}^2 g, 1)}.$$

The large sieve inequality for holomorphic cusp forms (see [IK, Theorem 7.24] for example) easily shows $M_3' \ll (q + q^{1/2} D)(qD)^\varepsilon$ and it seems likely that improvements are possible here using current technology. One may hope to show $M_5' \ll q^{1+\varepsilon}$ as this is a family with $\ll q$ elements with conductors of size approximately q^4 ; Buttcane and Khan [BK, Theorem 1.2] proved an estimate of this form for $\mathcal{F}_k(q)$ with k sufficiently large in terms of ε (smaller weights cause some technical difficulties so it is not straightforward to remove this condition). Conditional on this bound on M_5' , one would obtain

$$M_f(D) \ll (q^{1/2} D^{1/4} + q^{3/8} D^{1/2})(qD)^\varepsilon,$$

which would imply a subconvexity bound for $q = D^\eta$ for any fixed $0 < \eta$ (essentially as long as neither q nor D is fixed).

6. PROOF OF PROPOSITION 5.2

In this section we prove Proposition 5.2. The basic idea is to apply the large sieve inequality for holomorphic cusp forms. We begin by collecting some standard facts.

Proposition 6.1. *Let $\lambda_f(n)$ be the n -th Hecke eigenvalue of $f \in \mathcal{F}_2(q)$. Then for any complex numbers a_n , we have*

$$(6.1) \quad \sum_{f \in \mathcal{F}_2(q)} \frac{1}{L(\text{Sym}^2 f, 1)} \left| \sum_{n \leq N} a_n \lambda_f(n) \right|^2 \ll q^\varepsilon (q + N) \sum_{n \leq N} |a_n|^2.$$

By Gelbart and Jacquet [GJ], the symmetric square lift $\text{Sym}^2 f$ is a self-dual automorphic form on GL_3 with Fourier coefficients $A(m, k)$ satisfying

$$A(m, 1) = \sum_{ab^2=m} \lambda_f(a^2),$$

when $q \nmid m$, and

$$A(m, k) = \sum_{d|(m,k)} \mu(d) A\left(\frac{m}{d}, 1\right) A\left(1, \frac{k}{d}\right),$$

when $q \nmid mk$. Xiannan Li [L] showed the following uniform bound

$$(6.2) \quad \sum_{mk^2 \leq N} \frac{|A(m, k)|^2}{mk} \ll (qN)^\varepsilon$$

as a consequence of his uniform convexity bound. Technically, convexity would show (6.2) with mk^2 in the denominator, not mk , but one can use the multiplicativity relations of the Fourier coefficients to derive (6.2). We have

$$(6.3) \quad \begin{aligned} L(\text{Sym}^2 f \times g, s) &= (1 - \lambda_g(q)q^{-s})^{-1} (1 - \lambda_g(q)q^{-(s+1)})^{-1} L^{(q)}(\text{Sym}^2 f \times g, s) \\ &=: \sum_{n=1}^{\infty} \frac{\lambda_{\text{Sym}^2 f \times g}(n)}{n^s}, \end{aligned}$$

where

$$L^{(q)}(\text{Sym}^2 f \times g, s) = \sum_{(mk, q)=1} \frac{A(m, k) \lambda_g(m)}{(mk^2)^s}.$$

The conductor of $L(\text{Sym}^2 f \times g, 1/2)$ is q^4 , as we now briefly explain. Gross and Kudla [GK] showed that the conductor of $L(f \times f \times g, 1/2)$ is q^5 , while the conductor of $L(g, 1/2)$ is q . Therefore, using the factorization $L(f \times f \times g, s) = L(\text{Sym}^2 f \times g, s)L(g, s)$ we make this deduction. By an approximate functional equation [IK, Theorem 5.3], we have

$$(6.4) \quad M_5 \ll \sum_{g \in \mathcal{F}_2(q)} \frac{1}{L(\text{Sym}^2 g, 1)} \left| \sum_{n=1}^{\infty} \frac{\lambda_{\text{Sym}^2 f \times g}(n)}{n^{1/2}} V(n/q^2) \right|^2,$$

where $V(x)$ is some fixed smooth function with rapid decay for $x \rightarrow \infty$. Actually, we prefer to sum over n coprime to q in order to work with $L^{(q)}$ which has a simpler Dirichlet series. Towards this end we write $V(x) = \frac{1}{2\pi i} \int_{(c)} F(s) x^{-s} ds$ where $F(s)$ has rapid decay for

$|\operatorname{Im}(s)| \rightarrow \infty$, and is analytic for $\operatorname{Re}(s) > 0$. By inserting this Mellin formula into (6.4) and factorizing the L -function via (6.3), we have for $\sigma > 1/2$

$$(6.5) \quad M_5 \ll \sum_{g \in \mathcal{F}_2(q)} \frac{1}{L(\operatorname{Sym}^2 g, 1)} \left| \frac{1}{2\pi i} \int_{(\sigma)} \frac{q^{2s} F(s)}{\left(1 - \frac{\lambda_g(q)}{q^s}\right) \left(1 - \frac{\lambda_g(q)}{q^{s+1}}\right)} \sum_{(n,q)=1} \frac{\lambda_{\operatorname{Sym}^2 f \times g}(n)}{n^{1/2+s}} ds \right|^2.$$

Now the sum over n may be truncated at $\ll q^{2+\varepsilon}$ with a small error term (say $O(q^{-100})$) by shifting the contour far to the right if necessary. Having imposed this truncation on n , we may then shift the contour to $\sigma = \varepsilon > 0$ and apply Cauchy-Schwarz to give

$$(6.6) \quad M_5 \ll \int_{(\varepsilon)} |F(s)| \sum_{g \in \mathcal{F}_2(q)} \frac{q^{2\varepsilon}}{L(\operatorname{Sym}^2 g, 1)} \left| \sum_{\substack{(n,q)=1 \\ n \ll q^{2+\varepsilon}}} \frac{\lambda_{\operatorname{Sym}^2 f \times g}(n)}{n^{1/2+s}} \right|^2 |ds| + q^{-100}.$$

By unraveling the definition of Dirichlet series coefficients, and using Cauchy's inequality, we obtain for $\sigma > 1/2$

$$(6.7) \quad \left| \sum_{\substack{(n,q)=1 \\ n \ll q^{2+\varepsilon}}} \frac{\lambda_{\operatorname{Sym}^2 f \times g}(n)}{n^{\sigma+it}} \right|^2 = \left| \sum_{\substack{(mk,q)=1 \\ mk^2 \ll q^{2+\varepsilon}}} \frac{A(m,k)\lambda_g(m)}{m^{\sigma+it} k^{2\sigma+2it}} \right|^2 \\ \leq \zeta(2\sigma) \sum_{k \ll q^{1+\varepsilon}} k^{-2\sigma} \left| \sum_{\substack{(m,q)=1 \\ m \ll q^{2+\varepsilon}/k^2}} \frac{A(m,k)\lambda_g(m)}{m^{\sigma+it}} \right|^2.$$

Inserting (6.7) into (6.6), and using Proposition 6.1, we obtain

$$(6.8) \quad M_5 \ll q^\varepsilon \sum_{k \ll q^{1+\varepsilon}} k^{-1} \left(q + \frac{q^2}{k^2} \right) \sum_{m \leq q^{2+\varepsilon}/k^2} \frac{|A(m,k)|^2}{m}.$$

Then using (6.2) completes the proof of Proposition 5.2.

7. PROOF OF PROPOSITION 1.12

Next we give the proof of Proposition 1.12. Define F via

$$(7.1) \quad F(z) = \sum_{i=1}^n w_i^{1/2} \nu_i(f) \tilde{e}_i(z),$$

so that $\langle F^2, F^2 \rangle = \sum_{i=1}^n w_i^3 \nu_i(f)^4 = \|\tilde{f}\|_4^4$, and

$$(7.2) \quad \langle F^2, \tilde{g} \rangle = \sum_{i=1}^n w_i^2 \nu_i(f)^2 \nu_i(g) = (\tilde{f}, \tilde{f}, \tilde{g}).$$

By Parseval's formula, we have

$$\langle F^2, F^2 \rangle = \frac{\langle F^2, \tilde{e}^* \rangle^2}{\langle \tilde{e}^*, \tilde{e}^* \rangle} + \sum_{g \in \mathcal{F}_2(q)} (\tilde{f}, \tilde{f}, \tilde{g})^2.$$

Note $\langle F^2, \tilde{e}^* \rangle = \sum_{i=1}^n w_i \nu_i(f)^2 = \langle \tilde{f}, \tilde{f} \rangle = 1$. Also recall that $\langle \tilde{e}^*, \tilde{e}^* \rangle = (q-1)/12$, so

$$\frac{\langle F^2, \tilde{e}^* \rangle^2}{\langle \tilde{e}^*, \tilde{e}^* \rangle} \ll q^{-1}.$$

By Lemma 3.3 and Cauchy's inequality, we have

$$(7.3) \quad \sum_{g \in \mathcal{F}_2(q)} (\tilde{f}, \tilde{f}, \tilde{g})^2 \ll q^{-2} M_5^{1/2} M_6^{1/2}, \quad \text{where} \quad M_6 = \sum_{g \in \mathcal{F}_2(q)} \frac{L(g, \frac{1}{2})^2}{L(\text{Sym}^2 g, 1)},$$

and with M_5 as in (5.6). Then by Proposition 5.2 and the bound $M_6 \ll q^{1+\varepsilon}$ (implied by Lemma 3.2), we obtain the bound $O(q^{-1/2+\varepsilon})$. Assuming Lindelöf for triple products, we would have $M_5 \ll q^{1+\varepsilon}$, explaining the remark following Proposition 5.2. This completes the proof.

8. PROOF OF THEOREM 1.8

Let q be an odd prime and D be a fundamental discriminant with $(D, q) = 1$ and associated quadratic character χ_D . Let ω_f be the Petersson weights which occur in the Petersson trace formula, which for weight 2 reads

$$(8.1) \quad \sum_{f \in \mathcal{F}_2(q)} \omega_f \lambda_f(m) \lambda_f(n) = \delta_{m=n} - 2\pi \sum_{c \equiv 0 \pmod{q}} \frac{S(m, n; c)}{c} J_1\left(\frac{4\pi\sqrt{mn}}{c}\right).$$

The Petersson weights satisfy $\omega_f = q^{-1+o(1)}$.

By say Propositions 14.19 and 14.20 of [IK], $f \times \chi_D$ is a Hecke newform of level qD^2 , so the conductor of $L(1/2, f \times \chi_D)$ is also qD^2 . We use a long one-piece approximate functional equation for $L(1/2, f \times \chi_D)$ (as did Duke [D2]), namely

$$(8.2) \quad L(1/2, f \times \chi_D) = \sum_{n=1}^{\infty} \frac{\lambda_f(n) \chi_D(n)}{\sqrt{n}} V(n/X) + O_A((X/qD^2)^{-A}),$$

where $V(x)$ is a fixed (independent of q and D) smooth function with rapid decay, that is, $V(x) \ll_A (1+x)^{-A}$ for any $A > 0$. We take $X = (qD^2)^2$; surprisingly, the method is not particularly sensitive to the length of X . Then

$$(8.3) \quad \sum_{f \in \mathcal{F}_2(q)} \omega_f L(1/2, f \times \chi_D) = V(1/X) - 2\pi S + O((qD)^{-100}),$$

where

$$(8.4) \quad S = \sum_{c \equiv 0 \pmod{q}} c^{-1} S_c, \quad S_c = \sum_{n=1}^{\infty} n^{-1/2} S(n, 1; c) \chi_D(n) V(n/X) J_1\left(\frac{4\pi\sqrt{n}}{c}\right).$$

We will show

$$(8.5) \quad S \ll |D|^{1/2} q^{-1} (|D|q)^\varepsilon.$$

This will suffice to prove Theorem 1.8, since we can choose V to satisfy $V(x) = 1 + O(x)$ for $x \rightarrow 0^+$.

Since $J_1(x) \ll x$, we can use Weil's bound to estimate the terms with $c \geq C \geq q$, giving

$$(8.6) \quad \sum_{\substack{c \equiv 0 \pmod{q} \\ c \geq C}} c^{-1} S_c \ll \sum_{\substack{c \equiv 0 \pmod{q} \\ c \geq C}} X c^{-3/2+\varepsilon} \ll X q^{-1} C^{-1/2+\varepsilon}.$$

If $C \geq X^2/|D|$, then this trivial bound is satisfactory. Although we do not need a strong bound on the tail, it is vital that we can truncate the sum at some C which is polynomial in q and $|D|$.

Next we estimate S_c for c not too large by applying Poisson summation to the sum over n . Note that $f(x) = x^{-1/2}V(x/X)J_1\left(\frac{4\pi\sqrt{x}}{c}\right)$ is continuous on $x \geq 0$ and has rapid decay for x large, but $f(0) \neq 0$ so to simplify the analysis we first apply a dyadic partition of unity to the sum over n . We then write $S_c = \sum_{N \text{ dyadic}} S_c(N)$, where

$$(8.7) \quad S_c(N) = \sum_{n=1}^{\infty} n^{-1/2} S(n, 1; c) \chi_D(n) w_N(n) J_1\left(\frac{4\pi\sqrt{n}}{c}\right),$$

and $w_N(x)$ is a smooth function supported on $[N, 2N]$, say, satisfying $\frac{d^j}{dx^j} w_N(x) \ll N^{-j}$. Now that we have a smooth function, we may easily apply Poisson summation to n modulo $c|D|$, giving

$$(8.8) \quad S_c(N) = \sum_{m \in \mathbb{Z}} \frac{1}{c|D|} a(m; c, D) r(m; c, D),$$

where

$$(8.9) \quad a(m; c, D) = \sum_{x \pmod{c|D|}} \chi_D(x) S(x, 1; c) e\left(\frac{mx}{c|D|}\right),$$

and

$$(8.10) \quad r(m; c, D) = \int_0^{\infty} x^{-1/2} w_N(x) J_1\left(\frac{4\pi\sqrt{x}}{c}\right) e\left(-\frac{mx}{c|D|}\right) dx.$$

By Lemma 10.5 of [LMY], $a(0; c, D) = 0$ for $q|c$, a condition which always holds here. By a tiny strengthening of Lemma 10.5 of [LMY], we have

$$(8.11) \quad |a(m; c, D)| \leq 2^5 c|D|^{1/2} (m, c, D)^{1/2}.$$

Previously we had the bound $4^{\nu+2}c$, where $c = 2^{\nu}c'$, $(c', 2) = 1$, in place of 2^5c , so this requires a little explanation. We need to more carefully treat the case $p = 2$ in the proof of Lemma 10.5 of [LMY]. By a factorization argument, it suffices to consider the case where c and $\pm D$ are both powers of 2. We have, for any p ,

$$(8.12) \quad a(m, p^c, \pm p^D) = \sum_{x \pmod{p^{c+D}}} \sum_{y \pmod{p^c}}^* \chi_{\pm p^D}(x) e\left(\frac{xy + \bar{y}}{p^c}\right) e\left(\frac{mx}{p^{c+D}}\right).$$

When $p = 2$ we have $D = 0, 2$, or 3 . If $D = 0$ then as in (10.34) of [LMY] (which does not require p odd), we have

$$(8.13) \quad a(m; p^c, 1) = p^c e\left(\frac{-\bar{m}}{p^c}\right).$$

If $D = 2$ or 3 and $c < D$ then we quote a trivial bound for simplicity (leading to the factor 2^5). Now assume $c \geq D$. Changing variables $x \rightarrow x + p^c$ multiplies the entire sum by $e\left(\frac{m}{p^D}\right)$, so the sum vanishes unless $p^D | m$. Accordingly, write $m = p^D m_1$. The sum over x is then periodic modulo p^c , so it is the same sum repeated p^D times, so

$$(8.14) \quad a(m, p^c, \pm p^D) = p^D \sum_{x \pmod{p^c}} \sum_{y \pmod{p^c}}^* \chi_{\pm p^D}(x) e\left(\frac{x(y + m_1) + \bar{y}}{p^c}\right).$$

Next write $x = x_1 + p^D x_2$ where x_1 runs modulo p^D and x_2 runs modulo p^{c-D} . The sum over x_2 vanishes unless $y \equiv -m_1 \pmod{p^{c-D}}$, in which case it equals p^{c-D} . At this point, a trivial bound shows $|a(m, p^c, \pm p^D)| \leq p^{2D+c}$. Thus (8.11) holds.

We need to understand the analytic properties of $r(m; c, D)$ where this calculation is different from that of [LMY]. We claim that for $c \gg \sqrt{N}$, we have

$$(8.15) \quad r(m; c, D) \ll c^{-1} N \left(1 + \frac{|m|N}{c|D|}\right)^{-100},$$

while for $c \ll \sqrt{N}$ we have

$$(8.16) \quad r(m; c, D) \ll c \left(1 + \frac{|m|\sqrt{N}}{|D|}\right)^{-100}.$$

The J_1 -Bessel function has two distinct types of behavior. For $c \geq \sqrt{x} \asymp \sqrt{N}$, the J_1 factor is not oscillatory, as $J_1(y) \sim \frac{1}{2}y$ for $y \rightarrow 0$. Hence, in this region, we have that $r(m; c, D)$ is c^{-1} times the Fourier transform of a function satisfying the same derivative bounds as $w_N(x)$, so that (8.15) follows.

In the complementary range $c \ll \sqrt{N}$, we have $J_1(4\pi y) = y^{-1/2} \sum_{\pm} e(\pm 2iy) g_{\pm}(4\pi y)$, where $\frac{d^j}{dy^j} g_{\pm}(y) \ll_j 1$. Thus, in this range we have

$$(8.17) \quad r(m; c, D) = \sum_{\pm} c^{1/2} \int_0^{\infty} x^{-3/4} h_{c,N,\pm}(x) e\left(\frac{\pm 2\sqrt{x}}{c} - \frac{mx}{c|D|}\right) dx,$$

where

$$(8.18) \quad h_{c,N,\pm}(x) = w_N(x) g_{\pm}\left(\frac{4\pi\sqrt{x}}{c}\right).$$

It is easily checked that

$$(8.19) \quad \frac{d^j}{dx^j} h_{c,N,\pm}(x) \ll_j N^{-j}.$$

We treat the estimation of $r(m; c, D)$ using the classical theory of exponential integrals. The basic observation is that unless there is some cancellation in the two phases $\pm 2c^{-1}\sqrt{x}$ and $(c|D|)^{-1}mx$ for some $x \asymp N$ (which would imply $|m| \asymp |D|/\sqrt{N}$), then repeated integration by parts (as in say Lemma 8.1 of [BKY]) shows that

$$(8.20) \quad r(m; c, D) \ll c^{1/2} N^{1/4} \left(\frac{\sqrt{N}}{c} + \frac{|m|N}{c|D|}\right)^{-A},$$

which is a stronger estimate than (8.16). On the other hand, if $|m| \asymp |D|/\sqrt{N}$ (consistent with a stationary point, but we do not need to assume that such a point actually lies inside the support of w_N), then the van der Corput bound shows $r(m; c, D) \ll c$, which agrees with (8.16).

Now we finish the proof of (8.5). We have

$$(8.21) \quad \sum_{\substack{c \equiv 0 \pmod{q} \\ \sqrt{N} \ll c \leq C}} c^{-1} S_c(N) \ll N|D|^{-1/2} \sum_{\substack{c \equiv 0 \pmod{q} \\ \sqrt{N} \ll c \leq C}} c^{-2} \sum_{m \neq 0} (m, c, D)^{1/2} \left(1 + \frac{|m|N}{c|D|}\right)^{-100},$$

and a short calculation shows

$$(8.22) \quad \sum_{\substack{c \equiv 0 \pmod{q} \\ \sqrt{N} \ll c \leq C}} c^{-1} S_c(N) \ll |D|^{1/2+\varepsilon} q^{-1+\varepsilon} C^{\varepsilon}.$$

Pleasantly, this bound is independent of N , and we only need to sum over $O(\log q|D|)$ such values of N , so the same bound holds on $\sum_c c^{-1}S_c$, with c running over the same range. Similarly, we have

$$(8.23) \quad \sum_{\substack{c \equiv 0 \pmod{q} \\ c \ll \sqrt{N}}} c^{-1}S_c(N) \ll |D|^{-1/2} \sum_{\substack{c \equiv 0 \pmod{q} \\ c \ll \sqrt{N}}} \sum_{m \neq 0} (m, c, D)^{1/2} \left(1 + \frac{|m|\sqrt{N}}{|D|}\right)^{-100},$$

and an easy calculation gives a bound of the same form as (8.22). \square

REFERENCES

- [BD] M. Bertolini and H. Darmon, *A rigid analytic Gross-Zagier formula and arithmetic applications*. Ann. of Math. **146** (1997), 111–147.
- [BH] V. Blomer and G. Harcos, *Hybrid bounds for twisted L -functions*. J. Reine Angew. Math. **621** (2008), 53–79.
- [BKY] V. Blomer, R. Khan, and M. Young, *Distribution of mass of Hecke eigenforms*. Duke Math. J. **162** (2013), 2609–2644.
- [BM] V. Blomer and P. Michel, *Sup-norms of eigenfunctions on arithmetic ellipsoids*. Int. Math. Res. Not. IMRN 2011, no. 21, 4934–4966.
- [BK] J. Buttcane and R. Khan, *L^4 norms of Hecke newforms of large level*, preprint 2013, <http://arxiv.org/abs/1305.1850>.
- [Co] C. Cornut, *Mazur’s conjecture on higher Heegner points*. Invent. Math. **148** (2002), 495–523.
- [D1] W. Duke, *Hyperbolic distribution problems and half-integral weight Maass forms*. Invent. Math. **92** (1988), 73–90.
- [D2] W. Duke, *The critical order of vanishing of automorphic L -functions with large level*. Invent. Math. **119** (1995), 165–174.
- [EOY] N. Elkies, K. Ono and T. H. Yang, *Reduction of CM elliptic curves and modular function congruences*. Int. Math. Res. Not. 2005, no. 44, 2695–2707.
- [FW1] B. Feigon and D. Whitehouse, *Averages of central L -values of Hilbert modular forms with an application to subconvexity*. Duke Math. J. **149** (2009), 347–410.
- [FW2] B. Feigon and D. Whitehouse, *Exact averages of central values of triple product L -functions*. Int. J. Number Theory **6** (2010), 1609–1624.
- [GJ] S. Gelbart and H. Jacquet, *A relation between automorphic representations of $GL(2)$ and $GL(3)$* . Ann. Sci. École Norm. Sup. **11** (1978), 471–542.
- [G] B. Gross, *Heights and the special values of L -series*. Number theory (Montreal, Que., 1985), 115–187, CMS Conf. Proc., 7, Amer. Math. Soc., Providence, RI, 1987.
- [GK] B. Gross and S. Kudla, *Heights and the central critical values of triple product L -functions*. Compositio Math. **81** (1992), 143–209.
- [HK] J. Hoffstein and A. Kontorovich, *The first non-vanishing quadratic twist of an automorphic L -series*, preprint 2010, <http://arxiv.org/abs/1008.0839>.
- [HoMu] R. Holowinsky and R. Munshi, *Level Aspect Subconvexity For Rankin-Selberg L -functions*, to appear in Automorphic Representations and L -functions, Tata Institute of Fundamental Research, Mumbai, India, 2012. preprint available at <http://arxiv.org/abs/1203.1300>.
- [HoTe] R. Holowinsky and N. Templier, *First moment of Rankin-Selberg central L -values and subconvexity in the level aspect*. Ramanujan J. **33** (2014), 131–155.
- [I1] H. Iwaniec, *Fourier coefficients of modular forms of half-integral weight*. Invent. Math. **87** (1987), 385–401.
- [IK] H. Iwaniec and E. Kowalski, *Analytic number theory*. American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI, 2004. xii+615 pp.
- [ILS] H. Iwaniec, W. Luo and P. Sarnak, *Low lying zeros of families of L -functions*. Inst. Hautes Études Sci. Publ. Math. No. 91 (2000), 55–131 (2001).
- [JK] J. Jackson and A. Knightly, *Averages of twisted L -functions*, preprint, 24 pages.

- [JeK] D. Jetchev and B. Kane, *Equidistribution of Heegner points and ternary quadratic forms*. Math. Ann. **350** (2011), 501–532.
- [K1] B. Kane, *Representations of integers by ternary quadratic forms*. Int. J. Number Theory **6** (2010), 127–159.
- [K2] B. Kane, *CM liftings of supersingular elliptic curves*. J. Théor. Nombres Bordeaux **21** (2009), 635–663.
- [Koh] D. Kohel, *Endomorphism rings of elliptic curves over finite fields*. Thesis, University of California, Berkeley, December, 1996.
- [KoSe] W. Kohnen and J. Sengupta, *On quadratic character twists of Hecke L -functions attached to cusp forms of varying weights at the central point*. Acta Arith. **99** (2001), 61–66.
- [KZ] W. Kohnen and D. Zagier, *Values of L -series of modular forms at the center of critical strip*. Invent. Math. **64** (1981), 175–198.
- [L] Xiannan Li, *Upper bounds on L -functions at the edge of the critical strip*. IMRN (2010), 727–755.
- [LMY] S.-C. Liu, R. Masri and Matthew P. Young, *Subconvexity and equidistribution of Heegner points in the level aspect*. Compositio Math. **149** (2013), 1150–1174.
- [Mi] P. Michel, *The subconvexity problem for Rankin-Selberg L -functions and equidistribution of Heegner points*. Ann. of Math. **160** (2004), 185–236.
- [MR] P. Michel and D. Ramakrishnan, *Consequences of the Gross-Zagier formulae: stability of average L -values, subconvexity, and non-vanishing mod p* . Number theory, analysis and geometry, 437–459, Springer, New York, 2012.
- [MV] P. Michel and A. Venkatesh, *Heegner points and non-vanishing of Rankin/Selberg L -functions*. Analytic number theory, 169–183, Clay Math. Proc., 7, Amer. Math. Soc., Providence, RI, 2007.
- [N] P. Nelson, *Stable averages of central values of Rankin-Selberg L -functions: some new variants*, preprint 2012, <http://arxiv.org/abs/1202.6313>.
- [V] V. Vatsal, *Uniform distribution of Heegner points*. Invent. Math. **148** (2002), 1–46.
- [Y] T. H. Yang, *Minimal CM liftings of supersingular elliptic curves*. Pure Appl. Math. Q. **4** (2008), 1317–1326.

DEPARTMENT OF MATHEMATICS, WASHINGTON STATE UNIVERSITY, PULLMAN, WA 99164-3113
E-mail address: `scliu@math.wsu.edu`

DEPARTMENT OF MATHEMATICS, MAILSTOP 3368, TEXAS A&M UNIVERSITY, COLLEGE STATION, TX 77843-3368

E-mail address: `masri@math.tamu.edu`

E-mail address: `myoung@math.tamu.edu`