Annals of Mathematics, **160** (2004), 185–236

The subconvexity problem for Rankin-Selberg L-functions and equidistribution of Heegner points

By P. Michel*

À Delphine, Juliette, Anna and Samuel

Abstract

In this paper we solve the subconvexity problem for Rankin-Selberg L-functions $L(f \otimes g, s)$ where f and g are two cuspidal automorphic forms over $\mathbf{Q}$, g being fixed and f having large level and nontrivial nebentypus. We use this subconvexity bound to prove an equidistribution property for incomplete orbits of Heegner points over definite Shimura curves.

Contents

1. Introduction

1.1. *Statement of the results.* Given an automorphic L-function, $L(f, s)$, the subconvexity problem consists in providing good upper bounds for the order of magnitude of $L(f, s)$ on the critical line and in fact, bounds which are stronger than ones obtained by application of the Phragmen-Lindelöf (convexity) principle. During the past century, this problem has received considerable

*This research was supported by NSF Grant DMS-97-29992 and the Ellentuck Fund (by grants to the Institute for Advanced Study), by the Institut Universitaire de France and by the ACI "Arithmétique des fonctions L".

attention and was solved in many cases. More recently it was recognized as a key step for the full solution of deep problems in various fields such as arithmetic geometry or arithmetic quantum chaos (for instance see the end of the introduction of [DFI1] and more recently [CPSS], [Sa2]). For further background on this topic and other examples of applications, we refer to the surveys [Fr], [IS] or [M2].

In this paper we seek bounds which are sharp with respect to the conductor of the automorphic form f. For rank one L-function (i.e. for Dirichlet characters L-functions) this problem was settled by Burgess [Bu] (see also [CI] for a sharp improvement of Burgess bound in the case of real characters). In rank two (i.e. for Hecke L-functions of cuspidal modular forms), the problem was extensively studied and satisfactorily solved during the last ten years by Duke, Friedlander and Iwaniec in a series of papers [DFI1], [DFI2], [DFI3], [DFI4], [DFI5], [DFI6], [DFI7] culminating in [DFI8] with

Theorem 1. *Let f be a primitive cusp form of level q with primitive nebentypus. For every integer $j \geqslant 0$, and every complex number s such that $\Re e s = 1/2$, we have*

$$L^{(j)}(f, s) \ll q^{\frac{1}{4}-\frac{1}{23400}};$$

where the implied constant depends on s, j and on the parameter at infinity of f (i.e. the weight or the eigenvalue of the Laplacian).

Some years ago, motivated by the Birch-Swinnerton-Dyer conjecture and its arithmetic applications, the author, E. Kowalski and J. Vanderkam investigated (amongst other questions) this problem for certain L-functions of rank 4, namely the Rankin-Selberg L-function of two cusp form, one of them being fixed [KMV2].

To set up notation, we consider f and g two (primitive) cusp forms of levels q and D respectively. These are eigenforms of (suitably normalized) Hecke operators $\{T_n\}_{n\geqslant 1}$ with eigenvalues $\lambda_f(n), \lambda_g(n)$ respectively. For all primes p, these eigenvalue can be written as

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

where we denote by χ_f the nebentypus of f, and similarly for g. The Rankin-Selberg L-function is a well defined Euler product of degree 4 , which equals up to finitely many local factors

$$\prod_p \prod_{i,j=1,2} \big(1 - \frac{\alpha_{f,i}(p)\alpha_{g,j}(p)}{p^s}\big)^{-1} = L(\chi_f\chi_g, 2s) \sum_{n\geqslant 1} \frac{\lambda_f(n)\lambda_g(n)}{n^s},$$

with equality if $(q, D) = 1$.

Remark 1.1. According to the Langlands philosophy $L(f \otimes g, s)$ should be associated to a GL_4 automorphic form. Although its standard analytic properties (analytic continuation, functional equation) have been known for a while (from the work of Rankin, Selberg and others, see [J], [JS], [JPPS]), it is only recently that Ramakrishnan established its automorphy in full generality [Ram].

Note that the conductor of this L-function, $Q(f \otimes g)$, satisfies

$$q^2/D^2 \leqslant Q(f \otimes g) \leqslant (qD)^2$$

and $Q(f \otimes g) = (qD)^2$ for $(q, D) = 1$; from these estimates one can obtain the convexity bound

$$L(f \otimes g, s) \ll q^{1/2+\varepsilon} \tag{1.1}$$

for $\Re e s = 1/2$and any $\varepsilon > 0$, the implied constant depending on ε, s, g and the parameters at infinity of f. The subconvexity problem in the q-aspect is to replace the exponent 1/2 above by a strictly smaller one. In [KMV2, Th. 1.1], we could solve this problem under the following additional hypotheses:

- the level of g is square-free and coprime with q (these minor assumptions can be removed; see [M1]),
- f is holomorphic of weight > 1,
- the conductor q^* (say) of the nebentypus of f is not too large; it satisfies i.e. $q^* \leqslant q^\beta$ for some fixed constant $\beta < 1/2$.

In this paper we drop (most of) the two remaining assumptions and, in particular, solve the subconvexity problem when f has weight 0 or 1 and has a primitive nebentypus. We prove here the following:

THEOREM 2. *Let f, g be primitive cusp forms of level q, D and nebentypus χ_f, χ_g respectively. Assume that $\chi_f\chi_g$ is not trivial and also that g is holomorphic of weight $\geqslant 1$. Then, for every integer $j \geqslant 0$, and every complex number s on the critical line $\Re e s = 1/2$,*

$$L^{(j)}(f \otimes g, s) \ll_j q^{\frac{1}{2}-\frac{1}{1057}};$$

moreover the implied constant depends on j, s, the parameters at infinity of f and g (i.e. the weight or the eigenvalue of the Laplacian) and on the level of g.

Remark 1.2. One can check from the proof given below, that the dependence in the parameters s, the parameters at infinity of f, and the level of g, D, is at most polynomial (which may be crucial for certain applications). More precisely the exponent for D is given by an explicit absolute constant, and the exponent for the other parameters is a polynomial (with absolute constants as coefficients) in k_g (the weight of g) of degree at most one (we have made no effort to evaluate the dependence in k_g nor to replace the linear polynomials by absolute constants).

One can note a strong analogy between Theorem 1 and Theorem 2: Indeed the square $L(f,s)^2$ can be seen as the Rankin-Selberg L-function of f against the nonholomorphic Eisenstein series

$$E'(z) := \frac{\partial}{\partial s}E(z,s)_{|s=1/2} = y^{1/2}\log y + 4y^{1/2}\sum_{n\geqslant 1}\tau(n)\cos(2\pi nx)K_0(2\pi ny)$$

or Eisenstein series of weight one. In spite of this analogy, and the fact that our proof borrows some material and ideas from [DFI8], we wish to insist that the bulk of our approach requires completely different arguments (see the outline of the proof below). In fact, our method can certainly be adapted to handle $L(f,s)^2$ as well, thus giving another proof of Theorem 1 by assuming only that χ_f is nontrivial, but we will not carry out the proof here (however, see the discussion at the end of the introduction).

1.2. *Equidistribution of Heegner points.* In many situations, critical values of automorphic L-functions are expected to carry deep arithmetic information. This is specially the case of Rankin-Selberg L-functions, when f is a holomorphic cusp form of weight two and $g = g_\rho$ is the holomorphic weight one cusp form (resp. the weight zero Maass form with eigenvalue $1/4$) corresponding to an odd (resp. an even) Artin representation ρ of dimension two. An appropriate generalization of the Birch-Swinnerton-Dyer conjecture predicts that the central value $L(f\otimes g_\rho, 1/2)$ (eventually the first nonvanishing higher derivative) measures the "size" of some arithmetic cycle lying in the (ρ, f)-isotypic component of a certain Galois-Hecke module associated with a modular curve. For example our results may provide nontrivial upper bounds for the size of the Tate-Shafarevitch group of the associated Galois representations in terms of the conductor of ρ (see for example the paper [GL]).

In particular, for ρ an odd dihedral representation, the Gross-Zagier type formulae which have now been established in many cases [GZ], [G], [Z1], [Z2], [Z3] interpret $L(f\otimes g_\rho, 1/2)$ or its first derivative in terms of the height of *Heegner* divisors. In particular Theorem 2 provides nontrivial upper bounds for these heights, which may give, as we shall see, fairly nontrivial arithmetic information concerning these Heegner divisors, such as equidistribution properties.

For this introduction, we present our application in the most elementary form and refer to Section 6 for a more general statement. Given q a prime, we denote $\mathrm{Ell}^{ss}(\mathbf{F}_{q^2}) = \{e_i\}_{i=1\dots n}$ the finite set of supersingular elliptic curves over $\mathbf{F}_{q^2}$. We have $|\mathrm{Ell}^{ss}(\mathbf{F}_{q^2})| = n = \frac{q-1}{12} + O(1)$. This space is equipped with a "natural" probability measure μ_q given by

$$\mu_q(e_i) = \frac{1/w_i}{\sum_{j=1\dots n} 1/w_j}$$

where w_i is the number of units modulo $\{\pm 1\}$ of the (quaternionic) endomorphism ring of e_i. Note that this measure is not exactly uniform but almost (at least when q is large) since the product $w_1 \dots w_n$ divides 12. Let K be an imaginary quadratic field with discriminant $-D$, for which q is inert; let $\mathrm{Ell}(O_K)$ be the set of elliptic curves over $\overline{\mathbf{Q}}$ with complex multiplication by the maximal order of K. These curves are defined over the Hilbert class field of K, H_K, and the Galois group $G_K = \mathrm{Gal}(H_K/K) = \mathrm{Pic}(O_K)$ acts simply transitively on $\mathrm{Ell}(O_K)$; hence for any curve $E \subset \mathrm{Ell}(O_K)$, we have $\mathrm{Ell}(O_K) = \{E^\sigma\}_{\sigma \in G_K}$. When $\mathfrak{q}|q$ is any prime above q in H_K (recall that q splits completely in H_K), each $E \in \mathrm{Ell}(O_K)$ has good supersingular reduction modulo $\mathfrak{q}$. Hence a reduction map

$$\Psi_{\mathfrak{q}}: \ \mathrm{Ell}(O_K) \to \mathrm{Ell}^{ss}(\mathbf{F}_{q^2}).$$

One can then ask whether the reductions $\{\Psi_{\mathfrak{q}}(E^\sigma)\}_{\sigma \in G_K}$ are evenly distributed on $\mathrm{Ell}^{ss}(\mathbf{F}_{q^2})$ with respect to the measure μ_q as $D \to +\infty$. This is indeed the case, in fact in a stronger form:

THEOREM 3. *Let $G \subset G_K$ any subgroup of index $\leqslant D^{\frac{1}{2115}}$. For each $e_i \in \mathrm{Ell}^{ss}(\mathbf{F}_{q^2})$ and each $E \in \mathrm{Ell}(O_K)$, we have*

$$\frac{|\{\sigma \in G, \ \Psi_{\mathfrak{q}}(E^\sigma) = e_i\}|}{|G|} = \mu(e_i) + O_q(D^{-\eta}) \tag{1.2}$$

for some absolute positive η, the implied constant depending on q only.

To obtain this result, we express (by easy Fourier analysis) the characteristic function of G as a linear combination of characters ψ of G_K. Then the Weyl sums corresponding to this equidistribution problem can be expressed in terms of "twisted" Weyl sums. By a formula of Gross, later generalized by Daghigh and Zhang [G], [Da], [Z3], the twisted Weyl sums are expressed in terms of the central values $L(f \otimes g_\psi, 1/2)$ where f ranges over the fixed set of primitive holomorphic weight two cusp forms of level q, and g_ψ denotes the theta function associated to the character ψ (this is a weight one holomorphic form of level D with primitive nebentypus , $(\frac{-D}{*})$, the Kronecker symbol of K). Now, the subconvexity estimate of Theorem 2 (applied for f fixed and D varying) shows precisely that the Weyl sums are $o(1)$ as $D \to +\infty$ and the equidistribution follows.

Remark 1.3. Note that for the full orbit ($G = G_K$), only the principal character ψ_0 occurs in the above analysis and we have the factorization

$$L(f \otimes g_{\psi_0}, s) = L(f, s) L\left(f \otimes \left(\frac{-D}{*}\right), s\right);$$

in this case, the subconvexity estimate in the D aspect for the central value $L(f\otimes(\frac{-D}{*}),1/2)$ was first proved by Iwaniec [I1].

The result above is a particular instance of the equidistribution problem for Heegner divisors on Shimura curves associated to a *definite* quaternion algebra, namely the quaternion algebra over $\mathbf{Q}$ ramified at q and ∞. For other definite Shimura curves similar results hold *mutatis mutandis*; see Theorem 10 (the reader may consult [BD1] for general background on Heegner points in this context). These results may then be coupled with the methods of Ribet, and Bertolini-Darmon ([Ri], [BD2], [BD3]) to prove equidistribution of (the image of) small orbits of Heegner points in the group of connected components of the Jacobian of a Shimura curve associated to an *indefinite* quaternion algebra at a place of bad reduction or in the set of supersingular points at a place of good reduction. We will not pursue these interpretations here.

In this setting, other equidistribution problems for Heegner divisors have been considered by Vatsal and Cornut [Va], [Co] to study elliptic curves over the anticyclotomic $\mathbf{Z}_p$-extension of K. However the Heegner points considered in these papers were in the same isogeny class (i.e. associated to orders sitting in a *fixed* imaginary quadratic field). The subconvexity bound of the present paper allows for equidistribution statements even when the quadratic field varies.

1.3. *Outline of the proof of Theorem* 2. The beginning of the proof follows [KMV2]. First, we decompose $L(f\otimes g,s)$ into partial sums of the form

$$\mathcal{L}(f\otimes g):=\sum_{n\geqslant 1}\lambda_f(n)\lambda_g(n)W(n)$$

where the $W(n)$ are compactly supported smooth functions, the crucial range being when $n\sim q$. Next we use the amplification method and seek a bound for the second amplified moment

$$\sum_{f'\in\mathcal{F}}\omega_{f'}|\mathcal{L}(f'\otimes g)|^2|\sum_{\ell\leqslant L}\lambda_{f'}(\ell)x_\ell|^2 \tag{1.3}$$

where f' ranges over an appropriate (spectrally complete) family $\mathcal{F}$ of Hecke eigenforms of nebentypus χ_f, containing our preferred form f, $\omega_{f'}$ is an appropriate normalizing factor and the x_ℓ are arbitrary coefficients to be chosen later to amplify the contribution of the preferred form. The choice of the appropriate family $\mathcal{F}$ may be subtle. Specifically, the space of weight one holomorphic forms of given level is too small to make possible an efficient spectral analysis. This structural difficulty was resolved in [DFI8] by embedding the subspace of weight one holomorphic forms into the full spectrum of Maass forms of weight one. At this point, we open (1.3) and convert the resulting sum into sums of

Kloosterman sums using a spectral summation formula (i.e. Petersson's formula or an appropriate extension of Kuznetsov's formula which we borrow from [DFI8]). At this point one needs bounds for expressions of the form

$$\sum_{c\equiv 0(q)}\frac{1}{c}\sum_{m,n\geqslant 1}\overline{\lambda_g}(m)\lambda_g(n)S_\chi(m,\ell n;c)W(m)\overline{W}(\ell n)\mathcal{J}\left(\frac{4\pi\sqrt{\ell mn}}{c}\right)$$

where S_χ denotes the Kloosterman sum twisted by the character $\chi:=\chi_f$ and $\mathcal{J}$ is a kind of linear combination of Bessel type functions. For completeness we add that ℓ can be as large as a small positive power of q and the critical range for the variable c is around q. As in [KMV2] we open the Kloosterman sum and apply a Voronoi type summation formula to the $\lambda_g(m)$ sum, with the effect of replacing the Kloosterman sums by Gauss sums. This yields to an expression of the form

$$\sum_{c\equiv 0(q)}\frac{1}{c^2}\sum_h G_{\chi\chi_g}(h;c)\sum_{\ell m-n=h}\overline{\lambda_g}(m)\lambda_g(n)\mathcal{W}_g(m,n,c), \tag{1.4}$$

where $\mathcal{W}_g$ is a kind of Bessel transform depending on the type at infinity of g. The sum over h above splits naturally into two parts.

The first part corresponds to $h=\ell m-n=0$, its contribution is called the *singular* term. But, since we assume that $\chi\chi_g$ is not trivial, this term vanishes.

Remark 1.4. When $\chi\chi_g$ is trivial the contribution of the singular term is not always small; in fact it may be larger than the expected bound. However one expects as in [DFI8] that, in this case, the contribution is cancelled (up to admissible error term) by the contribution coming from the Eisenstein series. We do not carry this out here since we are mostly interested in cases where the conductor of χ_f is large.

The second part corresponding to $h\neq 0$,

$$\sum_{h\neq 0} G_{\chi\chi_g}(h;c)\sum_{\ell m-n=h}\overline{\lambda_g}(m)\lambda_g(n)\mathcal{W}_g(m,n,c) \tag{1.5}$$

is called the *off-diagonal* term and is the most difficult to evaluate. In order to deal with the shifted convolution sums

$$S_g(\ell,h):=\sum_{\ell m-n=h}\overline{\lambda_g}(m)\lambda_g(n)\mathcal{W}_g(m,n,c), \tag{1.6}$$

one could proceed as in [DFI3], [KMV2], with the δ-symbol method together with Weil's bound for Kloosterman sums. This method and a trivial bound for the Gauss sums $G_{\chi\chi_g}(h;c)$, is sufficient to solve the subconvexity problem as long as the conductor of χ is smaller than q^β for some $\beta<1/2$.

Instead, we handle the sums $S_g(\ell,h)$ by an alternative technique due to Sarnak [Sa2]. His method, which is built on ideas of Selberg [Se], uses the full force of the theory of automorphic forms on $\mathrm{GL}_{2,\mathbf{Q}}$. Sarnak's method consists in expressing (1.6) in terms of the inner product

$$I(s)=\int_{X_0(D\ell)}V_\ell(z)U_h(s,z)d\mu(z) \tag{1.7}$$

where $V_\ell(z)$ is the $\Gamma_0(D\ell)$-invariant function $(\Im m\ell z)^{k/2}\overline{g}(\ell z)(\Im m z)^{k/2}g(z)$ and $U_h(s,z)$ is a nonholomorphic Poincaré series of level $D\ell$. Taking the spectral expansion of $U_h(s,z)$, we transform this sum into

$$\sum_j\langle U_h(.,s),u_j\rangle\langle u_j,\overline{V}_\ell\rangle+\text{“Eisenstein”},$$

where $\{u_j\}_{j\geqslant 1}$ is a Hecke eigenbasis of Maass forms on $X_0(D\ell)$ and “Eisenstein” accounts for the contribution of the continuous spectrum. The scalar product $\langle u_j,\overline{V}_\ell\rangle$ has been bounded efficiently in [Sa1], and the other factor $\langle U_h(.,s),u_j\rangle$ is proportional to the h-th Fourier coefficient $\overline{\rho}_j(h)$ of $u_j(z)$. At this point one uses the following quantitative statement going in the direction of the Ramanujan-Petersson-Selberg conjecture to bound the resulting sums.

Hypothesis H_θ. *For any cuspidal automorphic form π on*

$$\mathrm{GL}_2(\mathbf{Q})\backslash\mathrm{GL}_2(\mathbf{A}_\mathbf{Q})$$

with local Hecke parameters $\alpha_\pi^{(1)}(p)$, $\alpha_\pi^{(2)}(p)$ for $p<\infty$ and $\mu_\pi^{(1)}(\infty)$, $\mu_\pi^{(j)}(\infty)$ there exist the bounds

$$|\alpha_\pi^{(j)}(p)|\leqslant p^\theta,\qquad j=1,2,$$
$$|\Re e\mu_\pi^{(j)}(\infty)|\leqslant\theta,\qquad j=1,2,$$

provided π_p, π_∞ are unramified, respectively.

Note that Hypothesis H_θ is known for $\theta=\frac{7}{64}$ thanks to the works of Kim, Shahidi and Sarnak [KiSh], [KiSa]. When the conductor q^* is small, this value of θ suffices for breaking the convexity bound; in fact it improves greatly the bound of [KMV2, Th. 1.1] (which may be obtained using $H_{1/4}$). Unfortunately, this argument alone is not quite sufficient when q^* is large: even Hypothesis H_0 (which is Ramanujan-Petersson-Selberg's conjecture) allows us only to solve our subconvexity problem as long as q^* is smaller than q^β for some fixed $\beta<1$.

From the discussion above, it is clear that we must also capture the oscillations of the Gauss sums in (1.5); this is reasonable since $G_{\chi\chi_g}(h;c)$ oscillate roughly like $\overline{\chi\chi_g}(h)$ and the length of the h-sum is relatively large (around q). This point is the key observation of the present paper; while this idea seems hard to combine with the δ-symbol technique, it works beautifully with the alternative method of Sarnak. Indeed, an inversion of the summations, reduces the problem to a nontrivial estimate, for each $j \geqslant 1$, of smooth sums of the shape

$$\sum_h \overline{\chi\chi_g}(h)\overline{\rho}_j(h)\tilde{W}(h),$$

where h is roughly of size q: this question reduces to the subconvexity problem for the twisted L-function

$$L(u_j \otimes \chi\chi_g, s), \text{ for } \Re e s = 1/2$$

in the q-aspect! This kind of subconvexity problem was solved by Duke-Friedlander-Iwaniec [DFI1] (when the fixed form is holomorphic) more than ten years ago as one of the first applications of the amplification method. In the appendix to this paper we provide the necessary subconvexity estimate in the case of Maass forms;[1] this estimate together with the Burgess bound (to handle the contribution from the continuous spectrum) is sufficient to finish the proof of Theorem 2.

Remark 1.5. We find rather striking that the solution of the subconvexity problem for our preferred rank four L-functions ultimately reduces to a collection of subconvexity estimates for rank-two and rank-one L-functions. This kind of phenomenon already appeared — implicitly — in [DFI8] where the Burgess estimate was used; in view of the inductive structure of the automorphic spectrum of GL_n (see [MW]), this should certainly be expected when dealing with the subconvexity problem for automorphic forms of higher rank.

Remark 1.6. The proof given here is fairly robust: any subconvex estimate for the $L(u_j \otimes \chi, s)$ in the q aspect (with a polynomial control on the remaining parameters) together with any nontrivial bound toward Ramanujan-Petersson's conjecture (that is H_θ for any fixed $\theta < 1/2$) would be sufficient to solve the given subconvexity problem, although with a weaker exponent.

1.3.1. *Comparison with* [DFI8]. As noted before, Theorem 2 and its proof share many similarities with the main result of [DFI8], but the hearts of the proofs are fairly different. To explain quickly the main differences, consider the subconvexity problem for the Hecke L-function $L(f, s)$. We have the identity

$$(|L(f,s)|^2)^2 = |L(f,s)|^4 = |L(f,s)^2|^2 (= |L(f \otimes E', s)|^2). \tag{1.8}$$

Our method would use the right-hand side of (1.8) and would evaluate the amplified mean square of partial sums of the form

$$\sum_n \lambda_f(n)\tau(n)W(n),$$

[1] See also [H] for a slightly weaker bound, and [CPSS] for another proof, in the holomorphic case, which uses Sarnak's method described above.

while the method of [DFI8] uses the left-hand side of (1.8) and evaluates the amplified mean square of (variants of) the partial sums

$$\sum_n \lambda_f(n)\tau_{\chi_f}(n)W(n),$$

where $\tau_\chi(n) = (1 * \chi_f)(n)$. In this case, the Gauss sums $G_{\chi_f}(h;c)$ of (1.4) are replaced by Ramanujan sums $r(h;c)$, so that for $h = 0$ a singular term appears (see Remark 1.4). This term turns out to be larger than the expected bound, but fortunately, a delicate computation shows that it is compensated by the contribution of the Eisenstein series (see [DFI8, §13]). The main problem then, is to bound the *off-diagonal* term; it is solved by the deep results of [DFI2], [DFI3] on the general determinant equation.

There are some advantages to handling Theorem 1 by the method of the present paper. A first one is technical; as long as χ_f is nontrivial, there is no singular term, hence no matching needs to be verified. However, a critical difference with the present paper is that for $g = E'$ an Eisenstein series, the integral $I(s)$ given in (1.7) has a pole at $s = 1$, which produces a new *off-off-diagonal* term; but as this term is independent of χ_f the resulting contribution is small as long as χ_f is nontrivial (otherwise one expects some matching with the contribution from the continuous spectrum). Another advantage of this method is that once the (many) remaining difficulties have been overcome, it is likely that the saving on the convexity exponent will be at least comparable with the exponent of Theorem 2.

The paper is organized as follows: In the next section, we introduce notation and give some background on automorphic forms, Hecke operators and spectral summation formulas. We recall also some useful lemmas and estimates which are borrowed from [DFI8]. In Section 3 we recall several facts on Rankin-Selberg L-functions and reduce the estimation of $L(f \otimes g, s)$ to that of partial sums. The bound for the second amplified moment of these partial sums starts in Section 4; it follows basically the techniques of [KMV2] and [DFI8]. In Section 5, we handle the shifted convolutions sums (1.5). The proof of Theorem 3 in a more general form is given in Section 6. In the appendix we provide a proof of a subconvexity bound for the L-function of a Maass form g twisted by a primitive character of large level. The result is not new; our main point there is to make explicit the (polynomial) dependence of the bound in the other parameters of g (the level or the eigenvalue), a question for which there is no available reference. Indeed, the polynomial control in the other parameters is crucial for the solution of our subconvexity problem.

Acknowledgments. During the course of this project, I visited the Institute for Advanced Study (during the academic year 1999–2000 and the first semester of 2000–2001), the Mathematics Department of Caltech (in April 2001), and the American Institute of Mathematics (in May 2001). I gratefully acknowledge these institutions for their hospitality and support. I wish to thank E. Ullmo, S. W. Zhang and D. Ramakrishnan for several discussions related to the equidistribution problem for Heegner points and my colleagues and friends E. Kowalski and J. Vanderkam with whom I began a fairly extensive study of Rankin-Selberg L-functions. I also thank the referee for his thorough review of the manuscript and his suggestions about many slips in earlier versions of the text. During the two years of this project, J. Friedlander, H. Iwaniec and P. Sarnak generously shared with me their experience, ideas and even the manuscripts (from the roughest to the most polished versions) of their respective ongoing projects; I thank them heartily for this, for their encouragement and their friendship.

2. A review of automorphic forms

In this section we collect various facts about automorphic Maass forms. Our main reference is [DFI8] which contains a very clear exposition of the whole theory.

The group $\mathrm{SL}_2(\mathbf{R})$ acts on the upper half-plane by linear-fractional transformations

$$\gamma z = \frac{az+b}{cz+d}, \text{ if } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

For $\gamma \in \mathrm{SL}_2(\mathbf{R})$ we define

$$j_\gamma(z) = \frac{cz+d}{|cz+d|} = \exp(i \arg(cz+d)),$$

and for any integer $k \geqslant 0$ an action of weight k on the functions $f : \mathbf{H} \to \mathbf{C}$ by

$$f_{|_k \gamma}(z) = j_\gamma(z)^{-k} f(\gamma z).$$

For $q \geqslant 1$, we consider Γ the congruence subgroup $\Gamma_0(q)$, and a Dirichlet character $\chi(\mathrm{mod}\ q)$; such a χ defines a character of Γ by

$$\chi\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \chi(d) = \overline{\chi}(a), \text{ for } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma.$$

2.1. *Maass forms.* A function $f : \mathbf{H} \to \mathbf{C}$ is said to be Γ-automorphic of weight k and nebentypus χ if and only if it satisfies

$$f_{|_k \gamma}(z) = \chi(\gamma) f(z) \tag{2.1}$$

for all $\gamma \in \Gamma$. We denote $\mathcal{L}_k(q,\chi)$ the L^2-space of such automorphic functions with respect to the Petersson inner product

$$\langle f, g \rangle = \int_{\Gamma\backslash\mathbf{H}} f(z)\overline{g}(z)\frac{dxdy}{y^2}.$$

By the theory of Maass and Selberg $\mathcal{L}_k(q,\chi)$ admits a spectral decomposition into the eigenspace of the Laplacian of weight k

$$\Delta_k = y^2\left(\frac{\partial^2}{\partial^2 x}+\frac{\partial^2}{\partial^2 y}\right) - iky\frac{\partial}{\partial x}.$$

The spectrum of Δ_k has two components: a discrete part spanned by the square integrable smooth eigenfunctions of Δ_k (the Maass cusp forms), and a continuous spectrum spanned by the Eisenstein series. The Eisenstein series are indexed by the *singular cusps* $\{\mathfrak{a}\}$ and are given by:

$$E_{\mathfrak{a}}(z,s) = \sum_{\gamma\in\Gamma_{\mathfrak{a}}\backslash\Gamma} \overline{\chi}(\gamma) j_{\sigma_{\mathfrak{a}}^{-1}\gamma}(z)^{-k}(\Im m(\sigma_{\mathfrak{a}}^{-1}\gamma z))^s$$

where $\sigma_{\mathfrak{a}}$ is a scaling matrix for the cusp $\mathfrak{a}$. Recall that the scaling matrix of a cusp $\mathfrak{a}$ is the unique matrix (up to right translations) such that

$$\sigma_{\mathfrak{a}}\infty = \mathfrak{a},\quad \sigma_{\mathfrak{a}}^{-1}\Gamma_{\mathfrak{a}}\sigma_{\mathfrak{a}} = \Gamma_\infty = \left\{\pm\begin{pmatrix} 1 & b \\ & 1\end{pmatrix},\ b\in\mathbf{Z}\right\},$$

and that a cusp $\mathfrak{a}$ is *singular* whenever

$$\chi\left(\sigma_{\mathfrak{a}}\begin{pmatrix} 1 & 1 \\ & 1\end{pmatrix}\sigma_{\mathfrak{a}}^{-1}\right) = 1,\ \text{or}\ (-1)^k.$$

The Eisenstein series $E_{\mathfrak{a}}(z,s)$ admit analytic continuation to the whole complex plane without pole for $\Re es \geqslant 1/2$ and are eigenfunctions of Δ_k with eigenvalue $\lambda(s)=s(1-s)$. The Maass cusp forms generate the cuspidal part of $\mathcal{L}_k(q,\chi)$ which we denote $\mathcal{C}_k(q,\chi)$. A Maass cusp form f has exponential decay and a Fourier expansion at every cusp. We only need Fourier expansion at infinity, this takes the form

$$f(z) = \sum_{\substack{n=-\infty\\ n\neq 0}}^{+\infty} \rho_f(n) W_{\frac{n}{|n|}\frac{k}{2},it}(4\pi|n|y)e(nx) \tag{2.2}$$

where $W_{\alpha,\beta}(y)$ is the Whittaker function, and $(1/2+it)(1/2-it)$ is the eigenvalue of f. The Eisenstein series have a similar Fourier expansion

$$\begin{aligned} E_{\mathfrak{a}}(z,1/2+it) = {} & \delta_{\mathfrak{a}} y^{1/2+it} + \phi_{\mathfrak{a}}(1/2+it)y^{1/2-it} \\ & + \sum_{\substack{n=-\infty\\ n\neq 0}}^{+\infty} \rho_{\mathfrak{a}}(n,t) W_{\frac{n}{|n|}\frac{k}{2},it}(4\pi|n|y)e(nx), \end{aligned}\tag{2.3}$$

where $\delta_{\mathfrak{a}}=0$, unless $\mathfrak{a}=\infty$, in which case $\delta_\infty=1$ and $\varphi_{\mathfrak{a}}(1/2+it)$ is the entry $(\infty,\mathfrak{a})$ of the scattering matrix.

2.2. *Holomorphic forms.* Let $\mathcal{S}_k(q,\chi)$ denote the space of holomorphic cusp forms of weight k, level q and nebentypus χ, i.e. the space of holomorphic functions $F:\mathbf{H}\rightarrow\mathbf{C}$ which satisfy

$$F(\gamma z)=\chi(\gamma)(cz+d)^kF(z) \tag{2.4}$$

for every $\gamma=\begin{pmatrix} a & b \\ c & d\end{pmatrix}\in\Gamma$ and which vanish at every cusp. This space is equipped with the Petersson inner product:

$$\langle F,G\rangle_k=\int_{\Gamma\backslash\mathbf{H}}F(z)\overline{G}(z)y^k\frac{dxdy}{y^2}.$$

Such a form has a Fourier expansion at ∞,

$$F(z)=\sum_{n\geqslant 1}\rho_F(n)n^{\frac{k}{2}}e(nz). \tag{2.5}$$

From the automorphy relations (2.4) one can deduce the following Voronoi-type summation formula (see [KMV2] and Section 7 for a more general formulas of the same type).

LEMMA 2.1. *Let $W:\mathbf{R}^+\rightarrow\mathbf{C}$ be a smooth function with compact support. Let $c\equiv 0(q)$ and a be an integer coprime with c. For $g\in\mathcal{S}_k(q,\chi)$,*

$$c\sum_{n\geqslant 1}\sqrt{n}\rho_g(n)e\left(n\frac{a}{c}\right)W(n)$$
$$=2\pi i^k\overline{\chi}(a)\sum_{n\geqslant 1}\sqrt{n}\rho_g(n)e\left(-n\frac{\overline{a}}{c}\right)\int_0^\infty W(x)J_{k-1}\left(\frac{4\pi\sqrt{nx}}{c}\right)dx.$$

It will be useful to quote the following properties of the Bessel function $J_k(x)$ for $k\geqslant 0$ (see [GR], [Wa]). We have

$$J_k(x)=e^{ix}V_k(x)+e^{-ix}\overline{V}_k(x) \tag{2.6}$$

where V_k satisfies

$$x^jV_k^{(j)}(x)\ll_j k^{2+j}\frac{1}{(1+x)^{1/2}} \tag{2.7}$$

for $j,k,x\geqslant 0$, the implied constant depending only on j. In fact, holomorphic forms can be embedded isometrically into the space of Maass forms of weight k:

LEMMA 2.2. *For $F(z)\in\mathcal{S}_k(q,\chi)$ the function $y^{k/2}F(z)$ belongs to $\mathcal{C}_k(q,\chi)$. More precisely the map $F(z)\rightarrow f(z):=y^{k/2}F(z)$ is a surjective isometry (relatively to the Petersson inner products) onto the eigenspace of Maass cusp forms of weight k with eigenvalue $\frac{k}{2}(1-\frac{k}{2})$; moreover the Fourier coefficients agree for all $n\in\mathbf{Z}$,*

$$\rho_F(n)=\rho_f(n).$$

From this lemma, it follows that $L(F\otimes g,s)=L(f\otimes g,s)$; so for the purpose of proving Theorem 2 we may and will assume that the varying form f is a Maass form of some weight $k\geqslant 0$.

2.3. *Spectral summation formulas.* Given $B_k(q,\chi)=\{u_j\}_{j\geqslant 1}$ an orthonormal basis of $\mathcal{C}_k(q,\chi)$ formed of Maass cusp forms with eigenvalues $\lambda_j=1/4+t_j^2$ and Fourier coefficients $\rho_j(n)$; the following spectral summation formula (borrowed from [DFI8, Prop. 5.2]) is an important tool for harmonic analysis on $\mathcal{L}_k(q,\chi)$. For any real number r, and any integer k we set

$$h(t)=h(t,r)=\frac{4\pi^3}{|\Gamma(1-\frac{k}{2}-ir)|^2}.\frac{1}{\mathrm{ch}\pi(r-t)\mathrm{ch}\pi(r+t)}. \tag{2.8}$$

Proposition 2.1. *For any positive integers m,n and any real r,*

$$\sqrt{mn}\sum_{j\geqslant 1}h(t_j)\overline{\rho}_j(m)\rho_j(n)+\sqrt{mn}\sum_{\mathfrak{a}}\frac{1}{4\pi}\int_{\mathbf{R}}h(t)\overline{\rho}_{\mathfrak{a}}(m,t)\rho_{\mathfrak{a}}(n,t)dt$$
$$=\delta_{m,n}+\sum_{c\equiv 0(q)}\frac{S_\chi(m,n;c)}{c}I\left(\frac{4\pi\sqrt{mn}}{c}\right)$$

where $S_\chi(m,n;c)$ is the Kloosterman sum

$$S_\chi(m,n;c)=\sum_{x(c),(x,c)=1}\overline{\chi}(x)e\left(\frac{m\overline{x}+nx}{c}\right),$$

and $I(x)$ is the Kloosterman integral

$$I(x)=I(x,r)=-2x\int_{-i}^{i}(-i\zeta)^{k-1}K_{2ir}(\zeta x)d\zeta.$$

In fact this formula is not quite sufficient for our purpose. In order to gain convergence over the c variable, an extra averaging over r is needed, and to achieve this, we follow the choice of [DFI8, §14]. Given A a fixed large real number we set

$$q(r)=\frac{r\mathrm{sh}2\pi r}{(r^2+A^2)^8}\left(\mathrm{ch}\frac{\pi r}{2A}\right)^{-4A}. \tag{2.9}$$

Integrating $q(r)h(t,r)$ over r we form

$$\mathcal{H}(t)=\int_{\mathbf{R}}h(t,r)q(r)dr \tag{2.10}$$

and correspondingly

$$\mathcal{I}(x)=\int_{\mathbf{R}}I(x,r)q(r)dr. \tag{2.11}$$

Hence, we deduce from Proposition 2.1 the following refined formula:

PROPOSITION 2.2. *For any positive integers m, n,*

$$\sqrt{mn}\sum_{j\geqslant 1}\mathcal{H}(t_j)\overline{\rho}_j(m)\rho_j(n)+\sqrt{mn}\sum_{\mathfrak{a}}\frac{1}{4\pi}\int_{\mathbf{R}}\mathcal{H}(t)\overline{\rho}_{\mathfrak{a}}(m,t)\rho_{\mathfrak{a}}(n,t)dt$$
$$=c_A\delta_{m,n}+\sum_{c\equiv 0(q)}\frac{S_\chi(m,n;c)}{c}\mathcal{I}\left(\frac{4\pi\sqrt{mn}}{c}\right)$$

where $\mathcal{H}$ and $\mathcal{I}$ are defined above and $c_A=\hat{q}(0)$ is the integral of q over $\mathbf{R}$.

We collect below the following estimates for $\mathcal{I}$ and $\mathcal{H}$ (see [DFI8, §§14 and 17]).

For t real or purely imaginary,

$$\mathcal{H}(t)>0,\ \mathcal{H}(t)\gg(1+|t|)^{k-16}e^{-\pi t}. \tag{2.12}$$

For all $j\geqslant 0$, we have

$$x^j\mathcal{I}^{(j)}(x)\ll_j\left(\frac{x}{1+x}\right)^{A+1}(1+x)^{1+j}. \tag{2.13}$$

One can also use more general forms of the above spectral summation formula to provide upper bounds for the Fourier coefficients of Maass forms; for instance, the following bound follows immediately from [DI, §§5.3 (5.6) (5.7) and (1.25)]:

LEMMA 2.3. *For $k=0$ and for any positive integer n, any $\varepsilon, T\geqslant 1$,*

$$\sum_{\substack{u_j\in B_0(q,\chi)\\|t_j|\leqslant T}}\frac{n|\rho_j(n)|^2}{\mathrm{ch}(\pi t_j)}\ll_\varepsilon T^2+(nqT)^\varepsilon\frac{(n,q)^{1/2}n^{1/2}}{q} \tag{2.14}$$

where the implied constant depends on ε only.

2.4. *Hecke operators.* The Hecke operators $\{T_n\}_n\geqslant 1$ are defined by

$$T_nf(z)=\frac{1}{\sqrt{n}}\sum_{ad=n}\chi(a)\sum_{b(d)}f\left(\frac{az+b}{d}\right).$$

They act on the L^2-space of Maass forms of weight k and in fact act on both $\mathcal{C}_k(q,\chi)$ and $\mathcal{E}_k(q,\chi)$. They satisfy the Hecke multiplicative relations:

$$T_mT_n=\sum_{d|(m,n)}\chi(d)T_{mnd^{-2}}, \tag{2.15}$$

and, in particular, commute with each other. They also commute with Δ_k and for $(n,q)=1$, T_n is a normal, because $T_n^*=\chi(n)T_n$; that is for all $f,g\in\mathcal{L}_k(q,\chi)$,

$$\langle T_nf,g\rangle=\overline{\chi}(n)\langle f,T_ng\rangle. \tag{2.16}$$

A Maass cusp form which is also an eigenfunction of the T_n for all $(n,q)=1$ will be called a *Hecke-Maass* cusp form and an orthonormal basis of $\mathcal{C}_k(q,\chi)$ made of Hecke-Maass cusp forms will be called a *Hecke eigenbasis*. The problem of the dimension of the Hecke eigenspace is well understood by Atkin-Lehner theory [AL], [ALi], [Li1]. By a primitive form we mean a Hecke-Maass cusp form which is orthogonal to the space of old forms and (unless otherwise specified) which has L^2-norm 1. By the Strong Multiplicity One Theorem, a primitive form is automatically an eigenform of all the Hecke operators.

For f an Hecke-Maass cusp form, with Hecke eigenvalues given by

$$T_n f = \lambda_f(n) f,$$

we have from (2.15),

$$\lambda_f(m)\lambda_f(n) = \sum_{d|(m,n)} \chi(d)\lambda_f(mnd^{-2}), \tag{2.17}$$

$$\lambda_f(mn) = \sum_{d|(m,n)} \mu(d)\chi(d)\lambda_f(m/d)\lambda_f(n/d), \tag{2.18}$$

for all $(mn,q)=1$ and these relations hold for all m,n if f is primitive. From (2.16) we also have

$$\lambda_f(n) = \chi(n)\overline{\lambda_f}(n), \tag{2.19}$$

for all $(n,q)=1$. Finally the action of Hecke operators on the Fourier expansion can be computed explicitly and for a Hecke-Maass cusp form we have:

$$\sqrt{m}\rho_f(m)\lambda_f(n) = \sum_{d|(m,n)} \chi(d)\rho_f\left(\frac{m}{d}\frac{n}{d}\right)\sqrt{\frac{mn}{d^2}}, \tag{2.20}$$

and

$$\sqrt{mn}\rho_f(mn) = \sum_{d|(m,n)} \mu(d)\chi(d)\rho_f\left(\frac{m}{d}\right)\sqrt{\frac{m}{d}}\lambda_f\left(\frac{n}{d}\right), \tag{2.21}$$

for all $m,n \geqslant 1$ with $(n,q)=1$. In particular, for all $(n,q)=1$,

$$\rho_f(n)\sqrt{n} = \rho_f(1)\lambda_f(n), \tag{2.22}$$

and for f primitive the relations (2.20), (2.21) and (2.22) are valid for all $n \geqslant 1$.

Remark 2.1. For the classical weight k holomorphic modular forms the Hecke operators T_n have a slightly different definition, and not too surprisingly this action commutes with the isometry $F(z) \to f(z) = y^{k/2}F(z)$ and in particular for F a primitive cusp form, $y^{k/2}F$ is also primitive and we have, for all n,

$$\lambda_F(n) = \lambda_f(n).$$

Remark 2.2. The Hecke operators also act on the space of Eisenstein series, but unless χ is *primitive* (for this case see [DFI8]) the Eisenstein series $E_{\mathfrak{a}}(z,s)$ are NOT eigenvectors of the T_n, $(n,q)=1$. The problem of diagonalizing the Hecke operators in the space of Eisenstein series was studied by Rankin in a series of papers [Ra1], [Ra2], [Ra3]; however we will not need any of these results.

2.5. *Bounds for Fourier coefficients of cusp forms.* In this section, we recall trivial and nontrivial bounds for Hecke eigenvalues and Fourier coefficients of automorphic forms. Given g a primitive cusp form of level D, weight k and eigenvalue $1/4+t_g^2$ (by convention g is L^2-normalized) from [DFI8] and [HL], we have

(2.23)
$$\frac{D^{-\varepsilon}(1+|t_g|)^{k/2-\varepsilon}}{\sqrt{D}}\mathrm{ch}\left(\frac{\pi t_g}{2}\right)\ll_\varepsilon \rho_g(1)\ll_\varepsilon \frac{D^{\varepsilon}(1+|t_g|)^{k/2+\varepsilon}}{\sqrt{D}}\mathrm{ch}\left(\frac{\pi t_g}{2}\right).$$

For Hecke eigenvalues, Hypothesis H_θ gives the individual bound[2]

$$|\lambda_g(n)|\leqslant \tau(n)n^{\theta}; \tag{2.24}$$

hence for all $n\neq 0$ we have by (2.22)

$$\rho_g(n)\ll_\varepsilon \frac{(Dn)^{\varepsilon}(1+|t_g|)^{k/2+\varepsilon}}{\sqrt{D}}n^{\theta-1/2}\mathrm{ch}\left(\frac{\pi t_g}{2}\right). \tag{2.25}$$

If g is holomorphic of weight $k\geqslant 1$, it follows from the work of Eichler-Shimura-Igusa, Deligne, Deligne-Serre that the Ramanujan-Petersson bound holds true:

$$|\lambda_g(n)|\leqslant \tau(n). \tag{2.26}$$

In general it turns out that the Ramanujan-Petersson bound is true on average by the theory of Rankin-Selberg and some auxiliary arguments (see [DFI8, §19]); we have for all $N\geqslant 1$ and all $\varepsilon>0$

$$\sum_{n\leqslant N}|\lambda_g(n)|^2\ll_\varepsilon (D(|t_g|+1)N)^{\varepsilon}N. \tag{2.27}$$

It will be also useful to introduce the following function

$$\sigma_g(n):=\sum_{d|n}|\lambda_g(d)|.$$

Note first that this function is almost multiplicative; by (2.17) and (2.18) we have

$$(mn)^{-\varepsilon}\sigma_g(mn)\ll\sigma_g(m)\sigma_g(n)\ll(mn)^{\varepsilon}\sigma_g(mn) \tag{2.28}$$

[2]Note that this bound remains true (trivially) for n a ramified prime.

for all $\varepsilon > 0$, and from (2.27) we have

$$\sum_{n\leqslant N}\sigma_g(n)^2 \ll_\varepsilon (q(1+|t_g|N))^\varepsilon N, \tag{2.29}$$

for all $N, \varepsilon > 0$. In the above estimates the implied constants depend only on ε.

For technical purposes it will also be useful to have a substitute of (2.25) when g is an L^2-normalized Hecke-Maass form of L^2 but not necessarily primitive. More precisely we have the following improvement over (2.14):

Proposition 2.3. *Let $B_0(q,\chi) = \{u_j\}_{j\geqslant 0}$ be a (orthonormal) Hecke-eigenbasis. Assume that Hypothesis H_θ holds; for any $T \geqslant 1$, $n \geqslant 1$ and any $\varepsilon > 0$,*

$$\sum_{\substack{u_j\in B_0(q,\chi)\\ |t_j|\leqslant T}} \frac{n|\rho_j(n)|^2}{\mathrm{ch}(\pi t_j)} \ll_\varepsilon (nqT)^\varepsilon T^2 n^{2\theta} \tag{2.30}$$

where the implied constant depends on ε only.

Proof. By the Atkin-Lehner theory, each Hecke-eigenspace is indexed by the primitive forms $g(z) \in \mathcal{C}_0(q^*q', \tilde{\chi})$ where q' ranges over the divisors of q/q^* (q^* the conductor of χ and $\tilde{\chi}$ is the character induced by χ^*); for each eigenspace, any element of any orthonormal basis $\{g_{(d)}(z),\ d|q/(q^*q')\}$ is a linear combinations of the $g(dz)$ where d ranges over the divisors of $q/(q^*q')$

$$g_{(d)}(z) = \sum_{d'|q/(q^*q')} \alpha_g(d,d')g(dz).$$

For uniformity we extend the above notation to all the divisors of q; namely we set $\alpha_g(d,d') = 0$ for each pair (d,d') of divisors of q which are not divisors of q/q_g and consequently we set $g_{(d)} = 0$ if d is not a divisor of q/q_g. With this convention, we have by (2.22)

$$\begin{aligned} n^{1/2}\rho_{(d)}(n) &= \sum_{d'|(q,n)} (d')^{1/2}\alpha_g(d,d')(n/d')^{1/2}\rho_g(n/d') \\ &= \rho_g(1)\sum_{d'|(q,n)} (d')^{1/2}\alpha_g(d,d')\lambda_g(n/d') := \rho_g(1)\beta_g(d,n), \end{aligned}$$

say. By Möebius inversion, we have for $d'|q$

$$(d')^{1/2}\alpha_g(d,d') = \sum_{d''|d'} \beta_g(d,d'')\lambda_g^{(-1)}(d'/d'')$$

where $\lambda_g^{(-1)}$ denotes the Möbius inverse of $\lambda_g(n)$: this is a multiplicative function given that for each prime p, by

$$\lambda_g^{(-1)}(p) = -\lambda_g(p),\ \lambda_g^{(-1)}(p^2) = \tilde{\chi}(p),\ \textit{and}\ \lambda_g^{(-1)}(p^k) = 0\ \textit{if}\ k \geqslant 3.$$

In particular we have from H_θ that $|\lambda_g^{(-1)}(n)| \leqslant \tau(n)n^\theta$. From the above discussion, it follows that

$$\sum_{\substack{u_j\in B_0(q,\chi)\\ |t_j|\leqslant T}} \frac{n|\rho_j(n)|^2}{\mathrm{ch}(\pi t_j)} = \sum_{q'|q/q^*} \sum_{\substack{g\\ |t_g|\leqslant T}} \frac{|\rho_g(1)|^2}{\mathrm{ch}(\pi t_g)} \sum_{d|q} |\beta_g(d,n)|^2, \tag{2.31}$$

and in particular when $n = d'|q$ we obtain from (2.14) the bound

$$\sum_{q'|q/q^*} \sum_{\substack{g\\ |t_g|\leqslant T}} \frac{|\rho_g(1)|^2}{\mathrm{ch}(\pi t_j)} \sum_{d|q} |\beta_g(d,d')|^2 \ll_\varepsilon (qT)^\varepsilon (T^2 + \frac{d'}{q}) \leqslant (qT)^\varepsilon T^2. \tag{2.32}$$

More generally we have

$$\begin{aligned}\sum_{d|q} |\beta_g(d,n)|^2 &= \sum_{d|q} |\sum_{d'|(q,n)} (d')^{1/2}\alpha_g(d,d')\lambda_g(n/d')|^2\\ &= \sum_{d|q} |\sum_{d''|(q,n)} \beta_g(d,d'') \sum_{d'|(n,q)/d''} \lambda_g\left(\frac{n}{d'd''}\right) \lambda_g^{(-1)}(d')|^2\\ &\ll_\varepsilon n^\varepsilon \sum_{d''|(q,n)} (\frac{n}{d''})^{2\theta} \sum_{d|q} |\beta_g(d,d'')|^2\end{aligned}$$

by Cauchy-Schwarz and H_θ. From (2.31), the last inequality and (2.32) we conclude the proof of Proposition 2.3. □

3. Rankin-Selberg L-functions

Our basic reference for Rankin-Selberg L-functions is the book of Jacquet [J]. Given f and g two primitive forms of level q and D respectively, the Rankin-Selberg L-function is a degree four Euler product

(3.1)
$$L(f\otimes g,s) = \sum_{n\geqslant 1} \frac{\lambda_{f\otimes g}(n)}{n^s} = \prod_p L_p(f\otimes g,s) = \prod_p \prod_{i=1}^4 \left(1-\beta_{f\otimes g}^{(i)}(p)p^{-s}\right)^{-1}$$

which is absolutely convergent for $\Re e s > 1$. In view of Lemma 2.2 and Remark 2.1 we may assume that f is a Maass form of some weight $k \geqslant 0$, with eigenvalue $1/4 + t_f^2$.

Remark 3.1. Although we will not use this fact, it is useful to know that by [Ram], $L(f\otimes g,s)$ is the L-function of a GL_4 automorphic form, which we denote by $f\otimes g$.

By direct inspection of the possible cases one can check that

$$|\beta_{f\otimes g}^{(i)}(p)| \leqslant p^{2\theta},$$

and for all $p \not| (q, D)$,

$$L_p(f \otimes g, s) = \prod_{i,j=1,2} \left(1 - \frac{\alpha_{f,i}(p)\alpha_{g,j}(p)}{p^s}\right)^{-1}.$$

In particular we have the following factorization for $\Re es > 1$,

$$L(f \otimes g, s) = \left(\sum_{d|D^\infty} \frac{\gamma(d)}{d^s}\right) L(\chi_f\chi_g, 2s) \sum_{n\geqslant 1} \frac{\lambda_f(n)\lambda_g(n)}{n^s} \tag{3.2}$$

with

$$\gamma_{f\otimes g}(d) \ll_\varepsilon d^{2\theta+\varepsilon}. \tag{3.3}$$

From now on we assume that $f \neq g$; then $L(f \otimes g, s)$ admits analytic continuation over $\mathbf{C}$ with no poles and it has a functional equation of the form

$$\Lambda(f \otimes g, s) = \varepsilon(f \otimes g)\Lambda(\overline{f} \otimes \overline{g}, 1 - s) \tag{3.4}$$

where $\varepsilon(f \otimes g)$ is some complex number of modulus one and

$$\Lambda(f \otimes g, s) = (Q(f \otimes g))^{s/2} L_\infty(f \otimes g, s) L(f \otimes g, s).$$

Here $L_\infty(f \otimes g, s)$ is the local factor at infinity

$$\begin{aligned} L_\infty(f \otimes g, s) &= L_\infty(\overline{f} \otimes \overline{g}, s) \\ &= \prod_{i=1,\dots,4} \Gamma_{\mathbf{R}}(s + \mu_{f\otimes g,i}(\infty)), \quad \text{with } \Gamma_{\mathbf{R}}(s) = \pi^{-s/2}\Gamma(s/2), \end{aligned}$$

and the integer $Q = Q(f \otimes g)$ is called the conductor of $f \otimes g$ and satisfies

$$Q(f \otimes g) \leqslant q^2 D^2. \tag{3.5}$$

From hypothesis H_θ and by inspection of the possible cases we verify that

$$\Re es \mu_{f\otimes g,i}(\infty) \geqslant -2\theta, \ i = 1, \dots, 4;$$

in particular $L_\infty(f \otimes g, s)$ is holomorphic for $\Re es > 2\theta$.

3.1. *Approximating $L(f \otimes g, s)$ by partial sums.* We proceed as in [DFI8, §9]. For $A_0 \geqslant 1$ large (to be defined later), set

$$G(u) = (\cos \frac{\pi u}{4A_0})^{-5A_0}. \tag{3.6}$$

By a contour shift we infer from the functional equation (3.4) that for $\Re es = 1/2$,

$$L(f\otimes g, s) = \sum_{n\geqslant 1} \frac{\lambda_f(n)\lambda_g(n)}{n^s} W_s\left(\frac{n}{\sqrt{Q}}\right) + \omega_{f\otimes g}(s) \sum_{n\geqslant 1} \frac{\overline{\lambda_f}(n)\overline{\lambda_g}(n)}{n^{1-s}} \tilde{W}_{1-s}\left(\frac{n}{\sqrt{Q}}\right)$$

where

$$\begin{aligned}\omega_{f\otimes g}(s) &= \varepsilon(f\otimes g)Q^{\frac{1-2s}{2}}\frac{L_\infty(f\otimes g,1-s)}{L_\infty(f\otimes g,s)},\\ W_s(y) &= \sum_{d|D^\infty}\frac{\gamma_{f\otimes g}(d)}{d^s}V_s(dy),\\ V_s(y) &= \frac{1}{2\pi i}\int_{(1)}\frac{L_\infty(f\otimes g,s+u)}{L_\infty(f\otimes g,s)}L(\chi_f\chi_g,2s+2u)\frac{G(u)}{u}y^{-u}du,\end{aligned}$$

and $\tilde{W}_s$ is defined like W_s except that $\gamma_{f\otimes g}(d)$ and $\chi_f\chi_g$ are replaced by $\overline{\gamma_{f\otimes g}}(d)$ and $\overline{\chi_f\chi_g}$.

Remark 3.2. For $\Re es = 1/2$, $|\omega_{f\otimes g}(s)| = 1$. Define

$$P = \prod_{i=1\ldots 4}(|s| + |\mu_{f\otimes g,i}(\infty)|)^{1/2}. \tag{3.7}$$

We have (compare with [DFI8, Lemma 9.2]) the following:

LEMMA 3.1. *Assume (for simplicity) that $\chi_f\chi_g$ is not trivial . For $\Re es = 1/2$ and for any $j \geqslant 0$,*

$$y^jW_s^{(j)}(y) \ll_{j,A} \log(1+qD|s|)^2P^j(1+\frac{y}{P})^{-A_0}.$$

Remark 3.3. If $\chi_f\chi_g$ is the trivial character, the bound above is valid with an extra factor $\log(1+y^{-1})$.

Proof. From (3.3) the series

$$\sum_{d|D^\infty}\frac{|\gamma_{f\otimes g}(d)|}{d^{1/2}}$$

converges and, so it suffices to prove the lemma for the function V_s. We shift the u contour to $\Re es = B$ with $B = -1/(\log(1+qD|s|))$ or $B = A_0$ and differentiate j times in y to get

$$\begin{aligned}y^jV_s^{(j)}(y) \ll_j y^{-B}&\int_{(B)}\left|\frac{L_\infty(f\otimes g,s+u)}{L_\infty(f\otimes g,s)}L(\chi_f\chi_g,2s+2u)\frac{u^jG(u)}{u}du\right|\\ &+\delta_{j=0,B<0}|L(\chi_f\chi_g,2s)|.\end{aligned}$$

Setting $s_i = s+\mu_{f\otimes g,i,j}(\infty)$ and $\sigma_i = \Re es_i$, we have by Stirling's formula,

$$\begin{aligned}\frac{\Gamma_{\mathbf{R}}(s+\mu_{f\otimes g,i}(\infty)+u)}{\Gamma_{\mathbf{R}}(s+\mu_{f\otimes g,i}(\infty))} &\ll_B \frac{|s_i+u|^{\frac{\sigma_i+B-1}{2}}}{|s_i|^{\frac{\sigma_i-1}{2}}}\exp\left(\frac{\pi}{4}(|s_i|-|s_i+u|)\right)\\ &\ll_{B,j} |u|^{-j}|s_i|^{\frac{j+B}{2}}\exp\left(\frac{\pi}{4}|u|\right).\end{aligned}$$

Hence,

$$y^j V_s^{(j)}(y) \ll_j y^{-B} P^{j+B} \int_{(B)} \exp(\pi|u|) \left| L(\chi_f\chi_g, 2s+2u) \frac{G(u)}{u} du \right| + \delta_{j=0,B<0} |L(\chi_f\chi_g, 2s)|.$$

By definition of $G(u)$, the integral is absolutely convergent and bounded by $\ll_{A_0} 1$ if $B = A_0$ and by $\ll_{A_0} \log^2(1 + qD|s|)$ for $B = -1/(\log(1 + qD|s|))$. The lemma follows by choosing $B = A_0$ if $y \geqslant P$ and $B = -1/(\log(1+qD|s|))$ otherwise. □

Applying a smooth partition of unity we derive that

$$L(f \otimes g, s) \ll \log(1 + qD|s|)^2 \sum_N \frac{|L_{f\otimes g}(N)|}{\sqrt{N}} \Big(1 + \frac{N}{PDq}\Big)^{-A_0} \tag{3.8}$$

where $L_{f\otimes g}(N)$ are sums of type

$$L_{f\otimes g}(N) = \sum_n \lambda_f(n)\lambda_g(n)W(n)$$

with $W(x)$ a smooth function supported on $[N/2, 5N/2]$ for $N = 2^\nu$, $\nu \geqslant -1$, such that for all $j \geqslant 0$

$$x^j W^{(j)}(x) \ll_{j,A_0} P^j. \tag{3.9}$$

By taking A_0 large enough, we see that Theorem 2 follows from Theorem 4 below, which gives a bound for the partial sums $L_{f\otimes g}(N)$.

Theorem 4. *Let g be a primitive holomorphic form of weight $k \geqslant 1$. For any $N \geqslant 1/2$ and any smooth function W supported on $[N/2, 5N/2]$ bounded by* 1 *and satisfying* (3.1),

$$\begin{aligned} L_{f\otimes g}(N) &\ll (qN)^\varepsilon[(qN)^{1/2} + N \left(\frac{N}{q}\right)^{E+B/2+3}] q^{-1/4(22(2C+2E+B+9)+11)} \\ &\ll (qN)^\varepsilon[(qN)^{1/2} + N \left(\frac{N}{q}\right)^{4}] q^{-1/1056} \end{aligned} \tag{3.10}$$

where the exponents B, C, E are as specified in (5.19) *and the implied constant depends on ε, k, P, D.*

Now, we obtain from this theorem and (3.8) the bound given in Theorem 2 for the zero-th derivative. By convexity we deduce the same bound for s in a $1/\log q$ neighborhood of the critical line and by Cauchy's formula we deduce the bound for $\Re es = 1/2$ for all the derivatives.

4. The amplified second moment

In this section we make the first reductions toward the proof of Theorem 4. In particular we perform amplification of the partial sum $L_{f\otimes g}(N)$ by averaging its amplified mean square over a well chosen family. Before doing so we need to transform slightly these sums. The reason for these apparently unmotivated transformations is to avoid the fact that Eisenstein series $E_{\mathfrak{a}}(z,s)$ are not Hecke eigenfunctions.

We denote by χ the character χ_f of our original form f. We consider the following linear form

$$L_{f\otimes g}(\vec{x},N)=\rho_f(1)\Big(\sum_{\ell\leqslant L}x_\ell\lambda_f(\ell)\Big)L_{f\otimes g}(N)$$

for any vector $\vec{x}=(x_1,\dots,x_\ell,\dots,x_L)\in\mathbf{C}^L$ with L some small power of q, the coefficients x_ℓ satisfying

$$(\ell,qD)\neq 1\Longrightarrow x_\ell=0. \tag{4.1}$$

From (2.17) for f followed by (2.18) for g we have

$$\begin{aligned}L_{f\otimes g}(\vec{x},N)&=\rho_f(1)\sum_\ell x_\ell\sum_n W(n)\lambda_g(n)\lambda_f(\ell)\lambda_f(n)\\ &=\rho_f(1)\sum_\ell x_\ell\sum_{de=\ell}\chi(d)\sum_{ab=d}\mu(a)\chi_g(a)\lambda_g(b)\sum_n W(adn)\lambda_g(n)\lambda_f(aen)\end{aligned}$$

and from (2.22) we obtain

$$\begin{aligned}&L_{f\otimes g}(\vec{x},N)\\ &\quad=\sum_\ell x_\ell\sum_{de=\ell}\chi(d)\sum_{ab=d}\mu(a)\chi_g(a)\lambda_g(b)\sum_n W(adn)\lambda_g(n)\sqrt{aen}\rho_f(aen).\end{aligned} \tag{4.2}$$

Note that the last expression makes perfectly good sense even if f is not a Hecke-eigenform. Hence we define for f any cusp form $L_{f\otimes g}(\vec{x},N)$ by the equality (4.2). We may also extend this definition for the Eisenstein series $E_{\mathfrak{a}}(z,1/2+t^2)$ and we denote $L_{\mathfrak{a},t,g}(\vec{x},N)$ the corresponding linear form (obtained by replacing $\rho_f(aen)$ by $\rho_{\mathfrak{a}}(aen,t)$ above).

Next we choose an orthonormal basis $B_k([q,D],\chi)$ of automorphic cusp forms of level $[q,D]$ – the least common multiple of q and D – and nebentypus the character (mod $[q,D]$) induced by χ. We average the quadratic form $|L_{f\otimes g}(\vec{x},N)|^2$ over it together with the Eisenstein series to form the "spectrally complete" quadratic form

$$Q_k(\vec{x},N):=\sum_j\mathcal{H}(t_j)|L_{u_j\otimes g}(\vec{x},N)|^2+\sum_{\mathfrak{a}}\frac{1}{4\pi}\int_{\mathbf{R}}\mathcal{H}(t)|L_{\mathfrak{a},t,g}(\vec{x},N)|^2dt$$

where $\mathcal{H}(t)$ is as defined in (2.10). Our goal is the following estimate for the complete quadratic form

THEOREM 5. *Assume g is primitive and holomorphic of level D. With the above notation, for all $\varepsilon > 0$,*

$$(LNq)^{-\varepsilon}Q_k(\vec{x},N) \ll N||\vec{x}||_2^2 + ||\vec{x}||_1^2 L^{2C+2E+B+9}N^2 q^{\theta-3/2}(q^*)^{\left(\frac{1}{2}-\frac{1}{22}-\theta\right)}\left(\frac{N}{q}\right)^{2E+B+6} + ||\vec{x}||_1^2 L^{2C'+2E'+B'+9}N^2 q^{\theta-3/2}(q^*)^{\left(\frac{1}{2}-\frac{1}{8}-\theta\right)}\left(\frac{N}{q}\right)^{2E'+B'+6}.$$

In the above expression,

$$||\vec{x}||_1 = \sum_{\ell\leqslant L}|x_\ell|, \text{ and } ||\vec{x}||_2^2 = \sum_{\ell\leqslant L}|x_\ell|^2;$$

the exponent θ equals $\frac{7}{64}$ and the exponents B, C, E, B', C', E' are as specified in (5.19) *and* (5.20); *moreover the implied constant depends on ε, k, P and D only.*

Remark 4.1. Considering a family slightly bigger than the obvious one enables us to simplify considerably the forthcoming computations (see §4.1.2).

Proof of Theorem 4 (*derivation from Theorem* 5). We choose an orthonormal basis $B_k([q,D],\chi)$ containing our preferred (now old) form $\frac{f}{\sqrt{[\Gamma_0([q,D]):\Gamma_0(q)]}}$. By positivity (in particular that of $\mathcal{H}(t)$, see (2.12)) we deduce that

$$\frac{|\rho_f(1)|^2}{[\Gamma_0([q,D]):\Gamma_0(q)]}\mathcal{H}(t_f)|\sum_{\ell\leqslant L}x_\ell\lambda_f(\ell))|^2|L_{f\otimes g}(N)|^2 \leqslant Q_k(\vec{x},N)$$

and from (2.12) and (2.23) we have

$$\frac{|\rho_f(1)|^2}{[\Gamma_0([q,D]):\Gamma_0(q)]}\mathcal{H}(t_f) \gg_\varepsilon \frac{(qD+|t_f|)^{-\varepsilon}}{[q,D](1+|t_f|)^{16}}.$$

Hence

$$|\sum_{\ell\leqslant L}x_\ell\lambda_f(\ell)|^2|L_{f\otimes g}(N)|^2 \ll_{D,P,k,\varepsilon} (LNq)^\varepsilon\Big[qN\sum_\ell|x_\ell|^2 + L^{2C+2E+B+9}(\sum_\ell|x_\ell|)^2\frac{N^2}{q^{\frac{1}{22}}}\left(\frac{N}{q}\right)^{2E+B+6} + L^{2C'+2E'+B'+9}(\sum_\ell|x_\ell|)^2\frac{N^2}{q^{\frac{1}{8}}}\left(\frac{N}{q}\right)^{2E'+B'+6}\Big].$$

To conclude we choose the standard amplifier

$$x_\ell = \begin{cases} \lambda_f(p)\overline{\chi}(p) & \text{if} \quad \ell = p,\ (p,qD)=1,\ \sqrt{L}/2 < p \leqslant \sqrt{L} \\ -\overline{\chi}(p) & \text{if} \quad \ell = p^2,\ (p,qD)=1,\ \sqrt{L}/2 < p \leqslant \sqrt{L} \\ 0 & \text{otherwise.} \end{cases}$$

From the relation $\lambda_f(p)^2 - \lambda_f(p^2) = \chi(p)$,

$$|\sum_{\ell \leqslant L} x_\ell \lambda_f(\ell))| \gg L^{1/2}/\log^2 L$$

for $L \geqslant (\log qD)^2$, and from (2.27),

$$\sum_\ell |x_\ell| + \sum_\ell |x_\ell|^2 \ll (q(1+|t_j|)L)^\varepsilon L^{1/2}.$$

To finish the proof of Theorem 4, we note first that N can be taken smaller than $q^{1+\frac{1}{4224}}$; otherwise the trivial bound $L_{f\otimes g}(N) \ll_{\varepsilon,P,D,k} (qN)^\varepsilon N$ is stronger than (3.10); then we conclude with the three inequalities above, by choosing

$$L = q^{\frac{1}{22(2C+2E+B+9)+11}} = q^{1/264}. \qquad \square$$

It remains to prove Theorem 5 for which we spend the rest of this section.

4.1. *Analysis of the quadratic form $Q(\vec{x}, N)$.* By Proposition 2.2 we have

(4.3)

$$\begin{aligned} Q(\vec{x},N) = &\sum_{\ell_1,\ell_2} \overline{x}_{\ell_1} x_{\ell_2} \sum_{\substack{a_1b_1e_1=\ell_1\\ a_2b_2e_2=\ell_2}} \mu(a_1)\mu(a_2)\chi\chi_g(\overline{a_1}a_2)\chi(\overline{b_1}b_2)\overline{\lambda_g}(b_1)\lambda_g(b_2) \\ &\times \big(S^D\big(\big(\begin{matrix} a_1 & b_1 & e_1 \\ a_2 & b_2 & e_2 \end{matrix}\big), N\big) + \sum_{c\equiv 0([q,D])} \frac{1}{c^2} S^{ND}\big(\big(\begin{matrix} a_1 & b_1 & e_1 \\ a_2 & b_2 & e_2 \end{matrix}\big), N; c\big)\big) \\ = &\, c_A Q^D(\vec{x},N) + Q^{ND}(\vec{x},N), \end{aligned}$$

say, where c_A is the constant defined in Proposition 2.2,

$$S^D\left(\left(\begin{matrix} a_1 & b_1 & e_1 \\ a_2 & b_2 & e_2 \end{matrix}\right), N\right) = \sum_{a_1e_1m=a_2e_2n} \overline{\lambda_g}(m)\lambda_g(n)\overline{W}(a_1d_1m)W(a_2d_2n)$$

and

(4.4)

$$\begin{aligned} S^{ND}\left(\left(\begin{matrix} a_1 & b_1 & e_1 \\ a_2 & b_2 & e_2 \end{matrix}\right), N; c\right) = c\sum_{m,n} \overline{\lambda_g}(m)\lambda_g(n) S_\chi(a_1e_1m, a_2e_2n; c) \\ \times \mathcal{I}\left(\frac{4\pi\sqrt{a_1a_2e_1e_2mn}}{c}\right)\overline{W}(a_1d_1m)W(a_2d_2n) \end{aligned}$$

where $d_1 = a_1b_1$ and $d_2 = a_2b_2$.

4.1.1. *The diagonal term.* Applying (2.21) in the reverse direction we find that

$$Q^D(\vec{x},N) = \sum_{\substack{d_1,e_1\\ d_2,e_2}} \overline{\chi}(d_1)\overline{x}_{d_1e_1}\chi(d_2)x_{d_2e_2} \sum_{e_1m=e_2n} \overline{\lambda_g}(d_1m)\lambda_g(d_2n)\overline{W}(d_1m)W(d_2n).$$

From (2.27), (2.28) and (2.29),

(4.5)
$$
\begin{aligned}
Q^D(\vec{x},N) &\ll \sum_{\substack{d,d_1,e_1\\ d_2,e_2}} |x_{dd_1e_1}||x_{dd_2e_2}| \sum_{n\geqslant 1} |\lambda_g(d_1e_2n)\lambda_g(d_2e_1n)||W(d_1e_2n)W(d_2e_1n)| \\
&\ll_\varepsilon (qP)^\varepsilon \sum_{\substack{d,d_1,e_1\\ d_2,e_2}} |x_{dd_1e_1}||x_{dd_2e_2}|\sigma_g(d_2)\sigma_g(e_1)\sigma_g(d_1)\sigma_g(e_2)\frac{N^{1+\varepsilon}}{\mathrm{Max}(d_2e_1,d_1e_2)} \\
&\ll_\varepsilon (qNP)^{2\varepsilon} N \sum_{d,\ell_1,\ell_2} |x_{d\ell_1}||x_{d\ell_2}|\frac{\sigma_g(\ell_1)\sigma_g(\ell_2)}{\sqrt{\ell_1\ell_2}} \ll_\varepsilon (qNP)^{3\varepsilon} N \sum_{\ell} |x_\ell|^2.
\end{aligned}
$$

4.1.2. *The nondiagonal term.* We transform (4.4) further by applying the Voronoi summation formula to the n variable. For this, we set $e=(a_2e_2,c)$, $c^*=c/e$, $e^*=a_2e_2/e$ so that $(c^*,e^*)=1$. Opening the Kloosterman sum, we have from (4.4),

$$
\begin{aligned}
S^{ND}\left(\begin{pmatrix} a_1 & b_1 & e_1\\ a_2 & b_2 & e_2\end{pmatrix}, N; c\right) &= ec^* \sum_{\substack{x(c)\\ (x,c)=1}} \chi(\overline{x}) \sum_m \overline{\lambda_g}(m) e\left(\frac{a_1e_1m\overline{x}}{c}\right) \\
&\times \sum_n \lambda_g(n) e\left(\frac{e^*xn}{c^*}\right)\mathcal{I}\left(\frac{4\pi\sqrt{a_1a_2e_1e_2mn}}{c}\right)\overline{W}(a_1d_1m)W(a_2d_2n).
\end{aligned}
$$

By (4.1), we have $(e,qD)=1$, hence $D|[q,D]|c^*$, so we apply Lemma 2.1 with the effect of replacing the additive character $e(\frac{e^*xn}{c^*})$ above by

$$
\overline{\chi_g(e^*x)}e\left(-\frac{\overline{e^*x}n}{c^*}\right) = \chi_g(\overline{e^*x})e\left(-\frac{\overline{e^*x}en}{c}\right).
$$

Hence

$$
\begin{aligned}
&S^{ND}\left(\begin{pmatrix} a_1 & b_1 & e_1\\ a_2 & b_2 & e_2\end{pmatrix}, N; c\right) \\
&\qquad = e\overline{\chi_g}(e^*)\sum_{m,n}\overline{\lambda_g}(m)\lambda_g(n)G_{\chi\chi_g}(a_1e_1m - e\overline{e^*}n;c)\mathcal{J}(m,n)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{J}(x,y) = 2\pi i^{k_g}\overline{W}(a_1d_1x)&\int_0^\infty W(a_2d_2u) \\
&\times\mathcal{I}\left(\frac{4\pi\sqrt{a_1a_2e_1e_2xu}}{c}\right)J_{k_g-1}\left(\frac{4\pi e\sqrt{yu}}{c}\right)du.
\end{aligned}
\tag{4.6}
$$

To proceed further we factor c as follows:

$$
c = c^\sharp c^\flat, \text{ where } c^\flat := \prod_{\substack{p|c\\ v_p(c)<v_p(a_2e_2)}} p^{v_p(c)};
\tag{4.7}
$$

in particular one has the following:

$$(c^\sharp, c^\flat) = 1, \ c^\flat | e, \ (c^\sharp, e^*) = 1,$$

and the Gauss sum factors accordingly (remember that $(e, qD) = 1$):

$$\begin{aligned} G_{\chi\chi_g}(a_1e_1m - e\overline{e^*}n; c) &= \chi\chi_g(c^\flat e^*) G_{\chi\chi_g}(a_1e_1m - en; c^\sharp) r(a_1e_1m - e\overline{e^*}n; c^\flat) \\ &= \chi\chi_g(c^\flat e^*) G_{\chi\chi_g}(a_1e_1m - en; c^\sharp) r(a_1e_1m; c^\flat), \end{aligned}$$

where

$$r(a_1e_1m; c^\flat) = \sum_{\substack{x(c^\flat) \\ (x,c^\flat)=1}} e\left(\frac{a_1e_1m}{c^\flat}\right) = \sum_{f|(a_1e_1m, c^\flat)} f\mu\left(\frac{c^\flat}{f}\right)$$

denotes the Ramanujan sum. Hence

$$\begin{aligned} &S^{ND}\left(\begin{pmatrix} a_1 & b_1 & e_1 \\ a_2 & b_2 & e_2 \end{pmatrix}, N; c\right) \\ &= \chi(e^*)\chi\chi_g(c^\flat) e \sum_{f|c^\flat} f\mu\left(\frac{c^\flat}{f}\right) \sum_h G_{\chi\chi_g}(h; c^\sharp) \sum_{\substack{e^*a_1e_1m - en = h \\ a_1e_1m \equiv 0(f)}} \overline{\lambda_g(m)} \lambda_g(n) \mathcal{J}(m,n); \end{aligned}$$

the congruence $a_1e_1m \equiv 0(f)$ is equivalent to $m \equiv 0(f^*)$ where $f^* := f/(a_1e_1, f)$. Using (2.18), we infer that

$$\begin{aligned} (4.8) \quad &S^{ND}\left(\begin{pmatrix} a_1 & b_1 & e_1 \\ a_2 & b_2 & e_2 \end{pmatrix}, N; c\right) \\ &= \chi(e^*)\chi\chi_g(c^\flat) e \sum_{f|c^\flat} f\mu\left(\frac{c^\flat}{f}\right) \sum_{f'|f^*} \mu(f') \overline{\chi_g(f')\lambda_g(f^*/f')} \Sigma(a_1e_1e^*f'f^*, e), \end{aligned}$$

where

$$\Sigma(a_1e_1e^*f'f^*, e) = \sum_h G_\chi(h; c^\sharp) S_h(a_1e_1e^*f'f^*, e) \tag{4.9}$$

and

$$S_h(a_1e_1e^*f'f^*, e) = \sum_{a_1e_1e^*f'f^*m - en = h} \overline{\lambda_g}(m) \lambda_g(n) \mathcal{J}(f'f^*m, n).$$

Since $\chi\chi_g$ is not the trivial character, $G_{\chi\chi_g}(0; c)S_0 = 0$, and we are left with evaluating (4.9) over the frequencies $h \neq 0$. This will be done in Theorem 6.

First we analyze the properties of $\mathcal{J}$; to simplify the notation we set

$$a = a_1d_1, \ b = a_2d_2, \ d = a_1a_2e_1e_2,$$

Lemma 4.1. *Let*

$$\Theta = \left(\frac{d}{ab}\right)^{1/2}\frac{N}{c}, \quad X_0 = \frac{N}{a},$$
$$Y_0 = P^2\frac{b}{e^2}\frac{c^2}{N}(1+\Theta^2) = P^2\frac{d}{e^2}\left(\frac{\Theta^2+1}{\Theta^2}\right)X_0, \quad Z = P+\Theta.$$

For any $\alpha, \beta, \nu \geqslant 0$,

$$x^\alpha y^\beta \frac{\partial^\alpha}{\partial^\alpha x}\frac{\partial^\beta}{\partial^\beta y}\mathcal{J}(x,y) \ll P^\beta Z^\alpha \frac{N}{b}\left(\frac{\Theta}{1+\Theta}\right)^{A+1}(1+\Theta)(1+\frac{y}{Y_0})^{-\nu-1/4},$$

the implied constant depending on α, β, ν *and (polynomially) on* k_g *and* A *is the constant fixed in* (2.9) *which also appears in* (2.13)*. Recall also that as a function of* x, $\mathcal{J}(x,y)$ *is supported on* $[X_0/2, 5X_0/2]$.

Proof. By a trivial estimation of the integral (4.6) using (2.7), (2.13) , (3.9) and that $x \sim N/a$, we see that

$$\mathcal{J}(x,y) \ll \frac{N}{b}\left(\frac{\Theta}{1+\Theta}\right)^{A+1}(1+\Theta)\left(1+\frac{ye^2N}{bc^2}\right)^{-1/4}.$$

Using the decomposition (2.6), we integrate by parts 2ν times the exponential. Using again (2.7), (2.13) and (3.9) we obtain

$$\mathcal{J}(x,y) \ll \frac{N}{b}\left(\frac{\Theta}{1+\Theta}\right)^{A+1}(1+\Theta)\left(1+\frac{y}{Y_0}\right)^{-\nu-1/4}. \tag{4.10}$$

Differentiating in x and y we obtain the desired conclusion. □

We now bound $\Sigma(a_1e_1e^*f'f^*, e)$ by applying Theorem 6 (to be proved in the forthcoming section), with the following choice of parameters (to avoid confusion the parameters of Theorem 6 are noted in boldface):

$$\Theta = \left(\frac{d}{ab}\right)^{1/2}\frac{N}{c}, \;\; \mathbf{l_1} = a_1e_1e^*f'f^* = f'f^*\frac{d}{e}, \;\; \mathbf{l_2} = e,$$

$$\mathbf{Z} = P+\Theta, \;\; \mathbf{X} = \frac{\mathbf{l_1}}{f'f^*}X_0 = \frac{d}{e}X_0, \;\; \mathbf{Y} = \mathbf{l_2}Y_0 = P^2\left(\frac{1+\Theta^2}{\Theta^2}\right)\mathbf{X}$$

$$\mathbf{q} = \mathrm{Cond}(\chi\chi_g)\;, \mathbf{c} = c^\sharp, \;\; \mathbf{F}(x,y) = \mathcal{J}(f'f^*x/\mathbf{l_1}, y/\mathbf{l_2});$$

and we obtain that (4.9) is bounded by (remember that $f'f^*d$ is coprime with qD and that $ff' \leqslant (c^\flat)^2$)

$$\begin{aligned}
&\ll (Lc)^\varepsilon \frac{d^{C+\frac{3}{2}}(f^*f')^{C+\frac{1}{2}}(df^*f')_1^{\frac{2}{11}}}{abe} \frac{(1+\Theta)^{k_g+3(2E+B+5)+1-A}}{\Theta^{k_g+2(2E+B+5)-A}} N^2\Big(\frac{c}{c^\flat}\Big)^{\theta+\frac{1}{2}}(q^*)^{\frac{1}{2}-\frac{1}{22}-\theta}\\
&\quad +(Lc)^\varepsilon \frac{d^{C'+\frac{3}{2}}(f^*f')^{C'+\frac{1}{2}}(df^*f')_1^{\frac{1}{2}}}{abe} \frac{(1+\Theta)^{k_g+3(2E'+B'+5)+1-A'}}{\Theta^{k_g+2(2E'+B'+5)-A'}} N^2\Big(\frac{c}{c^\flat}\Big)^{\theta+\frac{1}{2}}(q^*)^{\frac{1}{2}-\frac{1}{8}-\theta}\\
&\ll (Lc)^\varepsilon \frac{d^{C+E+B/2+\frac{9}{2}}(c^\flat)^{2C+\frac{15}{11}}(d)_1^{\frac{2}{11}}}{(ab)^{E+B/2+4}e} N^2\left(\frac{c}{c^\flat}\right)^{\theta+\frac{1}{2}}\left(\frac{N}{c}\right)^{2E+B+6}(q^*)^{\frac{1}{2}-\frac{1}{22}-\theta}\\
&\quad +(Lc)^\varepsilon \frac{d^{C'+E'+B'/2+\frac{9}{2}}(c^\flat)^{2C'+2}(d)_1^{\frac{1}{2}}}{(ab)^{E'+B'/2+4}e} N^2\left(\frac{c}{c^\flat}\right)^{\theta+1/2}\left(\frac{N}{c}\right)^{2E'+B'+6}(q^*)^{\frac{1}{2}-\frac{1}{8}-\theta},
\end{aligned}$$

the implied constants depending on D, k_g, P, ε only. Here we have used (4.10) with $A = k_g + 3(2E + B + 5) + 1$ and $A' = k_g + 3(2E' + B' + 5) + 1$, and we have bounded ff' by $(c^\flat)^2$; recall that $(d)_1$ denotes the factor of d defined as in (5.18).

Hence we deduce from this bound and (4.8) the upper bound

$$\begin{aligned}
&\sum_{c\equiv 0([q,D])} \frac{1}{c^2} S^{ND}\Big(\Big(\begin{matrix} a_1 & b_1 & e_1\\ a_2 & b_2 & e_2\end{matrix}\Big), N; c\Big)\\
&\qquad \ll_{D,k_g,P,\varepsilon} (Lq)^\varepsilon \frac{d^{C+E+B/2+\frac{9}{2}}(d)_1^{\frac{2}{11}}}{(ab)^{E+B/2+4}e} N^2 q^{\theta-\frac{3}{2}}\left(\frac{N}{q}\right)^{2E+B+6}(q^*)^{\frac{1}{2}-\frac{1}{22}-\theta}\\
&\qquad\quad +(Lq)^\varepsilon \frac{d^{C'+E'+B'/2+\frac{9}{2}}(d)_1^{\frac{1}{2}}}{(ab)^{E'+B'/2+4}e} N^2 q^{\theta-\frac{3}{2}}\left(\frac{N}{q}\right)^{2E'+B'+6}(q^*)^{\frac{1}{2}-\frac{1}{8}-\theta}.
\end{aligned}$$

Collecting all the terms (see (4.3)) we deduce

$$\begin{aligned}
Q^{ND}(\vec{x}, N) \ll_{D,k_g,P,\varepsilon} &(qLN)^\varepsilon ||\vec{x}||_1^2 L^{2C+2E+B+9} N^2 q^{\theta-\frac{3}{2}}(q^*)^{\left(\frac{1}{2}-\frac{1}{22}-\theta\right)}\left(\frac{N}{q}\right)^{2E+B+6}\\
+ \quad &(qLN)^\varepsilon ||\vec{x}||_1^2 L^{2C'+2E'+B'+9} N^2 q^{\theta-\frac{3}{2}}(q^*)^{\left(\frac{1}{2}-\frac{1}{8}-\theta\right)}\left(\frac{N}{q}\right)^{2E'+B'+6}.
\end{aligned}$$

This estimate together with (4.5) concludes the proof of Theorem 5.

5. A shifted convolution problem

In this section, which is the bulk of this paper, we consider the following shifted convolution problem: Let χ be a primitive character of modulus q, $1 < c \equiv 0(q)$. Let $\ell_1, \ell_2 \geqslant 1$ be two integers, and g be a primitive holomorphic cusp form of weight k and level D with some nebentypus, which is arithmetically normalized. That is, g has the Fourier expansion

$$g(z)=\sum_{n\geqslant 1}\lambda_g(n)n^{\frac{k-1}{2}}e(nz)$$

where $\lambda_g(n)$ denotes the n-th Hecke-eigenvalue. Let $F(x,y)$ be a smooth function supported on $[X/2,5X/2[\times[1/2,+\infty[$ which satisfies

$$x^\alpha y^\beta\frac{\partial^\alpha}{\partial^\alpha x}\frac{\partial^\beta}{\partial^\beta y}F(x,y)\ll Z^{\alpha+\beta}\left(1+\frac{y}{Y}\right)^{-\nu} \tag{5.1}$$

for some $Z,X,Y\geqslant 1$ and for all $\nu,\alpha,\beta\geqslant 0$, the implied constant depending on α,β,ν only.

We consider the sums

$$\Sigma(\ell_1,\ell_2):=\sum_{h\neq 0}G_\chi(h;c)S_h(\ell_1,\ell_2)$$

where $G_\chi(h;c)$ is the Gauss sum of the (induced) character $\chi \mod (c)$ and

$$S_h(\ell_1,\ell_2)=\sum_{\ell_1m-\ell_2n=h}\overline{\lambda_g}(m)\lambda_g(n)F(\ell_1m,\ell_2n). \tag{5.2}$$

Our goal in this section is to prove

THEOREM 6. *With the above notation, let $X'=\min(X,Y)$, $Y'=\mathrm{Max}(X,Y)$. The following upper bound holds*

$$\begin{aligned}\Sigma(\ell_1,\ell_2)\ll{}&(c\ell_1\ell_2)^\varepsilon Z^{2E+B+5}(\ell_1\ell_2)^{C+\frac12}(\ell_1\ell_2)_1^{\frac{2}{11}}(\ell_1\ell_2,q)^{\frac{1}{11}}\\&\times\left(\frac{Y'}{X'}\right)^{\frac{k-1}{2}+2E+B+5}\left(\frac{c}{q}\right)^\theta q^{\frac12-\frac{1}{22}}Y'c^{\frac12}\\&+(c\ell_1\ell_2)^\varepsilon Z^{2E'+B'+5}(\ell_1\ell_2)^{C'+\frac12}(\ell_1\ell_2)_1^{\frac12}(\ell_1\ell_2,q)^{\frac14}\\&\times\left(\frac{Y'}{X'}\right)^{\frac{k-1}{2}+2E'+B'+5}\left(\frac{c}{q}\right)^\theta q^{\frac12-\frac18}Y'c^{\frac12}\end{aligned}$$

for all $\varepsilon>0$, the constant implied depending on ε and g only. Here θ is any number such that Hypothesis H_θ is satisfied; the integer $(\ell_1\ell_2)_1$ (a factor of $\ell_1\ell_2$) is defined by the formula (5.18) *and the exponents B,C,E,B',C',E' are the ones given in* (5.19), (5.20).

Proof. First by a smooth dyadic partition of unity on the y variable we reduce the proof to the case where $F(x,y)$ is compactly supported on $[X/2,5X/2]\times[Y/2,5Y/2]$ and by symmetry we assume that $X'=X\leqslant Y=Y'$. We consider the following unique factorization

$$c=qq'c',\quad (c',q)=1,\ q'|q^\infty.$$

Now,

$$G_\chi(h;c) = \chi(c')G_\chi(h;qq')r(h;c')$$

where $r(h;c') = \sum_{d|(c',h)} d\mu(c'/d)$ denotes the Ramanujan sum. Moreover $G_\chi(h;qq') = 0$ unless $q'|h$ in which case

$$G_\chi(h;qq') = \overline{\chi}(h/q')q'G_\chi(1;q);$$

hence

$$\Sigma(\ell_1,\ell_2) = \chi(c')q'G_\chi(1;q)\sum_{d|c'} d\mu(c'/d)\overline{\chi}(d)\sum_{h\neq 0}\overline{\chi}(h)S_{hq'd}. \tag{5.3}$$

Our treatment of $\Sigma(\ell_1,\ell_2)$ begins with the method of Sarnak [Sa2] which we summarize below. This method is based on the analytic properties of the series

$$D(g,s;\ell_1,\ell_2,h) = \sum_{\ell_1 m - \ell_2 n = h} \overline{\lambda_g}(m)\lambda_g(n)\left(\frac{\sqrt{\ell_1\ell_2 mn}}{\ell_1 m + \ell_2 n}\right)^{k-1}(\ell_1 m + \ell_2 n)^{-s}.$$

Indeed,

$$S_h = \frac{1}{2\pi i}\int_{(2)} D(g,s;\ell_1,\ell_2,h)\widehat{F}(h,s)ds \tag{5.4}$$

with

$$\begin{aligned}\widehat{F}(h,s) &= \int_0^{+\infty} F\left(\frac{u+h}{2},\frac{u-h}{2}\right)\left(\frac{4u^2}{u^2-h^2}\right)^{\frac{k-1}{2}} u^s\frac{du}{u}\\ &= \int_{X-h}^{5X-h} F\left(\frac{u-h}{2},\frac{u+h}{2}\right)\left(4+\frac{2h}{u-h}-\frac{2h}{u+h}\right)^{\frac{k-1}{2}} u^s\frac{du}{u}.\end{aligned} \tag{5.5}$$

from the support property of F; in particular we have $\widehat{F}(h,s) = 0$ if $|h| \gg Y$.

Following [Sa2, Appendix], we set $N = D\ell_1\ell_2$ and express $D(g,s;h)$ in terms of the integral of the $\Gamma_0(N)$ invariant function

$$V(z) = (\ell_1 y)^{k/2}\overline{g}(\ell_1 z)(\ell_2 y)^{k/2}g(\ell_2 z)$$

against an appropriate Poincaré series

$$U_h(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma}(\Im m\gamma z)^s e(h\Re e\gamma z);$$

precisely,

(5.6)

$$I = \langle U_h, \overline{V}\rangle = \int_{\Gamma_0(N)\backslash\mathbf{H}} U_h(z,s)V(z)\frac{dxdy}{y^2} = \frac{\Gamma(s+k-1)}{(2\pi)^{s+k-1}(\ell_1\ell_2)^{-\frac12}}D(g,s;h).$$

On the other hand U_h can be decomposed spectrally (at least formally): we pick $B_0(N) = \{u_j\}_{j\geqslant 0}$ an orthonormal Hecke-eigenbasis of $\mathcal{C}_0(N,\chi_0)$ (where

u_0 is the constant function) and assume also that the u_j are eigenforms of the reflection operator $R.u_j(z) := u_j(-\overline{z}) = \varepsilon_j u_j(z)$ where $\varepsilon_j = \pm 1$. By Parceval, we have

(5.7)
$$\begin{aligned} I &= \sum_{j\geqslant 0} \langle U_h(.,s), u_j\rangle\langle u_j, \overline{V}\rangle + \text{Eisenstein Contr.} \\ &= \sum_{j\geqslant 1} \frac{2^{s-1}\overline{\rho_j}(h)}{|h|^{s-1}} \Gamma\left(\frac{s-\frac{1}{2}+it_j}{2}\right)\Gamma\left(\frac{s-\frac{1}{2}-it_j}{2}\right)\langle u_j, \overline{V}\rangle + \text{Eisenstein Contr.} \end{aligned}$$

and Eisenstein Contr. (the contribution from the continuous spectrum) is given by

(5.8)
$$\frac{1}{4\pi}\sum_{\mathfrak{a}}\int_0^\infty \frac{2^{s-1}\overline{\rho_{\mathfrak{a},t}}(h)}{|h|^{s-1}}\Gamma\left(\frac{s-\frac{1}{2}+it}{2}\right)\Gamma\left(\frac{s-\frac{1}{2}-it}{2}\right)\langle E_{\mathfrak{a}}(.,\tfrac{1}{2}+it), \overline{V}\rangle dt.$$

(the reader should note that the quantity "$\rho_j(h)$" in [Sa2] equals $2\rho_j(h)|h|^{1/2}$ in the present paper.) From [Sa2], (18), we have

$$\langle u_j, \overline{V}\rangle \ll_g \sqrt{N}(1+|t_j|)^{k+1}e^{-\frac{\pi}{2}|t_j|}, \tag{5.9}$$

and the same bound holds for Eisenstein series

$$\langle E_{\mathfrak{a}}(.,\tfrac{1}{2}+it), \overline{V}\rangle \ll_g \sqrt{N}(1+|t|)^{k+1}e^{-\frac{\pi}{2}|t|}.$$

Now calculations similar to those of [ILS, pp. 71–75] show[3] that one can choose the Hecke-eigenbasis $B_0(N) = \{u_j\}_{j\geqslant 0}$ such that the bound (2.25) holds:

$$\rho_{u_j}(h) \ll_\varepsilon \frac{(|h|N(1+|t_j|))^\varepsilon}{\sqrt{N}}\,\mathrm{ch}\left(\frac{\pi t_j}{2}\right)|h|^{\theta-1/2} \tag{5.10}$$

for all u_j; eventually by Weyl's law and the above estimates we obtain Theorem A.1 of [Sa2].

Theorem 7. *For any $\theta_1 > \theta$, $D(g,s;\ell_1,\ell_2,h)$ extends holomorphically to the half-plane $\{s\in\mathbf{C},\ \sigma := \Re e s \geqslant 1/2+\theta_1\}$ and satisfies in this region the upper bound*

$$D(g,s;\ell_1,\ell_2,h) \ll_{\varepsilon,g} (h\ell_1\ell_2)^\varepsilon(\ell_1\ell_2)^{1/2}|h|^{1/2+\theta_1-\sigma}(1+|t|)^3,$$

where $s=\sigma+it$ and the implied constant depends on ε and g only.

[3]For simplicity we shall not reproduce these (tedious) computations here but use instead the averaged version (2.30) of (5.10).

From the above result we can shift the contour in (5.4) to $\Re es = 1/2+\theta+\varepsilon$, and after integrating by parts $\widehat{F}(h,s)$, five times in u, we obtain that

$$\widehat{F}(h,s) \ll_\nu \frac{Z^5}{|s|^5}\left(\frac{Y}{X}\right)^{\frac{k-1}{2}+5} Y^{1/2+\theta+\varepsilon};$$

hence

$$S_h \ll_\varepsilon Z^5(D\ell_1\ell_2YZ)^\varepsilon D(\ell_1\ell_2)^{1/2}\left(\frac{Y}{X}\right)^{\frac{k-1}{2}+5} Y^{1/2+\theta}.$$

From (5.3) we obtain (since $dq'|h| \ll Y$),

$$\Sigma(\ell_1,\ell_2) \ll (cD\ell_1\ell_2Y)^\varepsilon DZ^5(\ell_1\ell_2q)^{1/2}\left(\frac{Y}{X}\right)^{\frac{k-1}{2}+5} Y^{3/2+\theta}, \tag{5.11}$$

the implied constant depending only on ε and g.

Remark 5.1. One can see easily that (5.11) is much stronger than the bound of Theorem 6 when q is small and in particular yields much better subconvexity exponents than the one given by Theorem 2 for small conductors. In fact, for the purpose of breaking convexity for Rankin-Selberg L-functions , any bound for $\Sigma(\ell_1,\ell_2)$ with $Y^{3/2+\theta}$ replaced by $Y^{2-\delta}$ for any fixed $\delta>0$ would suffice. One can see that the bound (5.11) is sufficient as long as $q \leqslant Y^{1-2\theta-\delta}$ for any fixed $\delta>0$. Taking back the notation from the introduction we see that the Ramanujan-Petersson conjecture (i.e. $\theta=0$ is admissible) would solve the convexity problem for Rankin-Selberg L-functions as long as $q^* \leqslant q^{1-\delta}$ for any fixed $\delta>0$. In the rest on this section we will solve the problem for all $\hat{q}$ unconditionally by exploiting the averaging over h and the oscillations of $\overline{\chi}(h)$.

From the above analysis we see that

$$\begin{aligned}\sum_{h\neq 0}\overline{\chi}(h)S_{dq'h} &= (\ell_1\ell_1)^{-\frac{1}{2}}\\ &\times \frac{1}{2\pi i}\int_{(2)} \frac{2^{s-1}(2\pi)^{s+k-1}}{\Gamma(s+k-1)}\Big(\text{Discrete Contr.} + \text{Eisenstein Contr.}\Big)ds\end{aligned} \tag{5.12}$$

where

$$\begin{aligned}\text{Discrete Contr.} = &\sum_{j\geqslant 1}\Gamma\left(\frac{s-\frac{1}{2}+it_j}{2}\right)\Gamma\left(\frac{s-\frac{1}{2}-it_j}{2}\right)\langle u_j,\overline{V}\rangle\\ &\times\sum_{h\neq 0}\frac{\overline{\chi}(h)\overline{\rho_j}(dq'h)}{|dq'h|^{s-1}}\widehat{F}(dq'h,s)\end{aligned} \tag{5.13}$$

is the contribution from the discrete part of the spectrum and Eisenstein Contr. is the similar expression coming from the continuous spectrum. In the next Subsections 5.1 and 5.2 we evaluate both contributions.

5.1. *The discrete spectrum contribution.* We handle here the discrete part (5.13) and more precisely the inner sum

$$\Sigma_j(\chi,s) := \sum_{h\neq 0} \frac{\overline{\chi}(h)\overline{\rho_j}(dq'h)}{|dq'h|^{s-1}} \widehat{F}(dq'h,s)$$

which has analytic continuation to $\Re e s \geqslant 1/2+\theta_1$. We handle here the contribution corresponding to $h<0$, the other one being similar. We abuse the notation slightly by using the same notation $\Sigma_j(\chi,s)$ for the sum running over $h<0$. Now,

$$\Sigma_j(\chi,s) = \varepsilon_j\chi(-1)\frac{1}{2\pi i}\int_{(2)} L_j^{dq'}(\chi,z+s-1)(dq')^{-(z+s-1)}\widetilde{F}(z,s)dz$$

where

(5.14)
$$\widetilde{F}(z,s) = \int_0^\infty\int_0^\infty F\left(\frac{u+h}{2},\frac{u-h}{2}\right)(4+\frac{2h}{u-h}-\frac{2h}{u+h})^{\frac{k-1}{2}}h^z u^s\frac{dhdu}{hu},$$

and

$$L_j^{dq'}(\chi,s) = \sum_{h\geqslant 1}\frac{\overline{\chi}(h)\overline{\rho_j}(dq'h)}{h^s}.$$

$L_j^{dq'}(\chi,s)$ is (up to a shift by 1/2) essentially the L-function of u_j twisted by χ. We will see in the next subsection that $L_j^{dq'}(\chi,s)$ has analytic continuation to the half-plane $\Re e s\geqslant 0$, and satisfies for $\Re e s=0$ the following bound:

(5.15)
$$\begin{aligned} &L_j^{dq'}(\chi,s)\\ &\ll_\varepsilon R_j(dq',N;0)(cN|s|(1+|t_j|))^\varepsilon|1+s|^E(1+|t_j|)^B N^C N_1^{2/11}(N,q)^{1/11}q^{1/2-1/22}\\ &\quad+R_j(dq',N;0)(cN|s|(1+|t_j|))^\varepsilon|1+s|^{E'}(1+|t_j|)^{B'}N^{C'}N_1^{1/2}(N,q)^{1/4}q^{1/2-1/8}\end{aligned}$$

for any $\varepsilon>0$, where N_1 and the exponents B, C, E, B', C', E' are as given in (5.18), (5.19) and (5.20) and

$$R_j(dq',N;\sigma) = \sum_{dq'|h|(dq'N)^\infty}\frac{|\rho_j(h)|}{h^\sigma}$$

(by (2.30), $R_j(dq',N;\sigma)$ is converging for $\sigma>-1/4$). We evaluate $\Sigma_j(\chi,s)$ on the line $\Re e s=1/2+\theta_1$ with $\theta_1=\theta+\varepsilon$. First we shift the z contour in (5.14) to $\Re e z=1/2-\theta_1$. Then we integrate by parts α times in u and β times in h in (5.14) and apply (5.1) with $\nu=\frac{k-1}{2}+\alpha+\beta+3/2+\varepsilon$ to gain convergence in the h variable. Now, we obtain

$$\widetilde{F}(z,s)\ll_k\left(\min_{0\leqslant\alpha,\beta\leqslant 100}\left(\frac{ZY}{X|s|}\right)^\alpha\left(\frac{ZY}{X|z|}\right)^\beta\right)\left(\frac{Y}{X}\right)^{\frac{k-1}{2}}Y.$$

In the above bound, α and β don't need to be integers. We take $\beta = E+1+\varepsilon$ to ensure convergence of the z integral and apply (5.15) to get

$$\Sigma_j(\chi,s) \ll_\varepsilon R_j(dq',N;0)(cN|s|(1+|t_j|))^\varepsilon |s|^{E-\alpha}(1+|t_j|)^B N^C (N,q)^{1/11} N_1^{2/11} \times Z^{\alpha+E+1}\left(\frac{Y}{X}\right)^{\frac{k-1}{2}+\alpha+E+1} q^{\frac{1}{2}-\frac{1}{22}} Y + \dots$$

where $\dots$ contains the similar term involving the exponents B', C', E'. We plug this bound, (5.9) and (5.10) into (5.13) and use the following estimate

$$\begin{aligned}
&\sum_{j\geqslant 1} \frac{|\Gamma\left(\frac{s-\frac{1}{2}+it_j}{2}\right)\Gamma\left(\frac{s-\frac{1}{2}-it_j}{2}\right)|}{|\Gamma(s+k-1)|} \frac{(1+|t_j|)^{B+k+1} R_j(dq',N;0)}{e^{\pi|t_j|/2}} \\
&\leqslant \Bigg(\sum_{j\geqslant 1} \frac{(1+|t_j|)^{2(B+k+1)}|\Gamma\left(\frac{s-\frac{1}{2}+it_j}{2}\right)\Gamma\left(\frac{s-\frac{1}{2}-it_j}{2}\right)|}{|\Gamma(s+k-1)|}\Bigg)^{1/2} \\
&\quad\times \sum_{dq'|h|(dq'N)^\infty} \Bigg(\sum_{j\geqslant 1} \frac{|\Gamma\left(\frac{s-\frac{1}{2}+it_j}{2}\right)\Gamma\left(\frac{s-\frac{1}{2}-it_j}{2}\right)|}{|\Gamma(s+k-1)|} \frac{|\rho_j(h)|^2}{e^{\pi|t_j|}}\Bigg)^{1/2} \\
&\ll_{\varepsilon,k,D} (dq'N|s|)^\varepsilon |2s|^{B+3} N^{1/2} (dq')^{\theta-1/2} = (dq'N|s|)^\varepsilon N^{1/2}(dq')^{\theta-1/2}|2s|^{B+3}
\end{aligned}$$

by Weyl's law for the spectrum and (2.30). We choose $\alpha = E+B+4+\varepsilon$ to have convergence in the s integral and we infer from (5.9) and the last estimate that the discrete spectrum contribution to $\sum_{h\neq 0}\overline{\chi}(h)S_{dq'h}$ in (5.12) is bounded by

$$\begin{aligned}
&\ll_{\varepsilon,k,D} \frac{(c\ell_1\ell_2)^\varepsilon}{(dq')^{1/2-\theta}} Z^{2E+B+5}(\ell_1\ell_2)^{C+\frac{1}{2}}(\ell_1\ell_2)_1^{2/11}(\ell_1\ell_2,q)^{\frac{1}{11}}\left(\frac{Y}{X}\right)^{\frac{k-1}{2}+2E+B+5} q^{\frac{1}{2}-\frac{1}{22}}Y \\
&+\frac{(c\ell_1\ell_2)^\varepsilon}{(dq')^{1/2-\theta}} Z^{2E'+B'+5}(\ell_1\ell_2)^{C'+\frac{1}{2}}(\ell_1\ell_2)_1^{1/2}(\ell_1\ell_2,q)^{\frac{1}{4}}\left(\frac{Y}{X}\right)^{\frac{k-1}{2}+2E'+B'+5} q^{\frac{1}{2}-\frac{1}{8}}Y.
\end{aligned} \tag{5.16}$$

It remains to prove (5.15) which we do in the next subsection.

5.1.1. *Bounds for twisted L-functions.* Recall that u_j is a Hecke-eigenform and denote $\tilde{u}_j$ the primitive form (of some level N' dividing N) underlying u_j. For any $n \geqslant 1$, we denote $\lambda_j(n)$ the n-th Hecke-eigenvalue of $\tilde{u}_j$; in particular, for $(n,N)=1$, it coincides with the n-th Hecke eigenvalue of u_j. We have the further factorization

$$(5.17) L_j^{dq'}(\chi, s) = \left(\sum_{h|(dq'N)^\infty} \frac{\overline{\chi}(h)\overline{\rho_j}(dq'h)}{h^s}\right)\left(\sum_{(n,dq'N)=1} \frac{\overline{\chi}(n)\overline{\lambda_j}(n)}{n^{s+1/2}}\right),$$

$$= \left(\sum_{h|(dq'N)^\infty} \frac{\overline{\chi}(h)\overline{\rho_j}(dq'h)}{h^s}\right)$$

$$\times \left(\prod_{p|dq'N}\left(1 - \frac{\overline{\chi}(p)\overline{\lambda_j}(p)}{p^{s+1/2}} + \frac{\chi_0(p)}{p^{2s+1}}\right)\right) L(\overline{\tilde{u}_j.\chi}, s+1/2)$$

where χ_0 denotes the trivial character modulo N' and

$$L(\overline{\tilde{u}_j.\chi}, s) = \sum_n \frac{\overline{\chi}(n)\overline{\lambda_j}(n)}{n^s}$$

is the twisted L-function of $\overline{\tilde{u}_j}$ by $\overline{\chi}$. By Hypothesis H_θ and (2.30) the product of two first factors of (5.17) has analytic continuation to the half-plane $\Re es \geqslant -1/2 + \theta + \delta$ for any fixed $\delta > 0$, and is bounded in this domain by

$$\ll_{\varepsilon,\delta} (dq'N(1+|t_j|))^\varepsilon R_j(dq', N; \Re es).$$

On the other hand $L(\overline{\tilde{u}_j.\chi}, s+1/2)$ has analytic continuation to $\mathbf{C}$ and what we need is an upper bound for it when s is on the shifted critical line $\Re es = 0$. It turns out that the convexity bound is just insufficient for us. The subconvexity problem for twisted L-functions $L(g \otimes \chi, s)$ in the conductor aspect was solved for the first time in [DFI1] for g holomorphic and of level one with the subconvexity exponent $1/2 - 1/22$. Recently, Cogdell, Piatetski-Shapiro and Sarnak solved the problem by another method (based on Theorem 7) for g still holomorphic, but of any level and with the better subconvexity exponent $1/2 - 7/130$ [CPSS]. The case (of main interest for us) where g is a weight zero Maass form of any level was recently settled by G. Harcos [H] by a variant of the δ-symbol method.

THEOREM 8. *Let g be a fixed weight zero primitive Maass form, and χ be a primitive character of modulus q. For $\Re es = 1/2$,*

$$L(g.\chi, s) \ll q^{1/2 - 1/54 + \varepsilon}$$

for any $\varepsilon > 0$, the implied constant depending on ε, s and g.

Unfortunately this bound does not display explicitly the dependence in $|s|$ or in the parameters of g. For our purpose an explicit polynomial dependence is crucial; lacking a reference, we provide in the appendix a refinement of the method of [DFI1] yielding:

THEOREM 9. *Let g be a weight zero primitive Maass form of level N and eigenvalue $1/4+t^2$, and let χ be a primitive character of modulus q. Set*

$$N_1 := \prod_{\substack{p^\alpha || N \\ \alpha\geqslant 2}} p^{\alpha-2}. \tag{5.18}$$

For $\Re e s = 1/2$,

$$\begin{aligned}L(g.\chi,s) \ll_\varepsilon &(|s|(1+|t|)Nq)^\varepsilon |s|^E (1+|t|)^B N^C N_1^{2/11}(N,q)^{1/11} q^{1/2-1/22+\varepsilon}\\ &+(|s|(1+|t|)Nq)^\varepsilon |s|^{E'}(1+|t|)^{B'} N^{C'} N_1^{1/2}(N,q)^{1/4}q^{1/2-1/8+\varepsilon}\end{aligned}$$

for any $\varepsilon>0$, the implied constant depending on ε only and with the following values for the exponents

$$B = 14/11,\quad C = 1/4,\quad E = 4/11, \tag{5.19}$$

$$B' = 7/2,\quad C' = 11/16,\quad E' = 1. \tag{5.20}$$

From this result, we deduce (5.15) for $\Re e s = 0$.

5.2. *The continuous spectrum contribution.* The arguments for the contribution from the continuous spectrum follows the same lines. The only point we need to check is that a bound similar to (5.15) holds in the case of Eisenstein series $E_{\mathfrak{a}}(z,t)$ for the corresponding L-function

$$L^{dq'}_{\mathfrak{a},t}(\chi,s) := \sum_{h\geqslant 1}\frac{\overline{\chi}(h)\overline{\rho_{\mathfrak{a}}}(dq'h,t)}{h^s}.$$

Recall that the cusps of $\Gamma_0(N)$ are uniquely represented by the rationals

$$\frac{u}{w},\ w|N,\ (u,w)=1,\ 1\leqslant u\leqslant (w,N/w).$$

In the half-space $\Im m t<0$, we have ([DI] (1.17) and p.247)

$$\begin{aligned}\rho_{\mathfrak{a}}(h,t) = &\frac{\pi^s|h|^{it-1/2}}{\Gamma(1/2+it)}\left(\frac{(w,N/w)}{wN}\right)^{1/2+it}\sum_{(\gamma,N/w)=1}\gamma^{-1-2it}\\ &\times \sum_{\substack{\delta(\gamma w),(\delta,\gamma w)=1\\ \delta\gamma\equiv u \bmod (w,N/w)}} e(-h\frac{\delta}{\gamma w})\end{aligned}$$

and (either by the general theory of Eisenstein series or in this case by the standard zero-free region for Dirichlet L-functions) the $\rho_{\mathfrak{a}}(h,t)$ have analytic continuation to $\Im m t = 0$ with at most one simple pole at $t=-i/2$ (for $n=0$ only). This can be seen by a direct computation which may be cumbersome for a general cusp $\mathfrak{a}$. In particular note that the Fourier coefficients $\rho_{\mathfrak{a}}(h,t)$ are not proportional to a multiplicative function, the reason being that $E_{\mathfrak{a}}(z,1/2+it)$ is not an eigenfunction of the Hecke operators (even of these T_n for n coprime with the level). The problem of diagonalizing Eisenstein series is studied thoroughly by Rankin in [Ra1], [Ra2], [Ra3], [Ra4], but we will not use his results. We restore multiplicativity by decomposing the γ sum according to the characters modulo $(w,N/w)$:

$$\sum_{(\gamma,N/w)=1}\frac{1}{\gamma^{1+2it}}\sum_{\substack{\delta(\gamma w),(\delta,\gamma w)=1\\ \delta\gamma\equiv u \bmod (w,N/w)}} e\left(-h\frac{\delta}{\gamma w}\right)$$
$$=\frac{1}{\varphi((w,N/w))}\sum_{\psi \bmod (w,N/w)}\overline{\psi}(-u)\sum_{(\gamma,N/w)=1}\frac{\psi(\gamma)}{\gamma^{1+2it}}G_\psi(h;\gamma w).$$

For each character $\psi \bmod (w,N/w)$ we denote w^* its conductor, and decompose $w=w^*w'w''$ with $w'|w^{*\infty},\ (w'',w^*)=1$. Accordingly the Gauss sum factors as follows:

$$\begin{aligned}G_\psi(h;\gamma w)&=\psi(\gamma w'')G_\psi(h;w^*w')r(h;\gamma w'')\\&=\delta_{w'|h}w'\psi(\gamma w'')G_\psi(h/w';w^*)r(h/w';\gamma w'').\end{aligned}$$

Hence

$$\sum_{(\gamma,N/w)=1}\frac{\psi(\gamma)}{\gamma^{1+2it}}\ldots=\delta_{w'|h}w'\psi(w'')G_\psi(1;w^*)\Bigg(\sum_{\substack{\gamma|N^\infty\\(\gamma,N/w)=1}}\frac{\psi^2(\gamma)}{\gamma^{1+2it}}r(h;\gamma w'')\Bigg)$$
$$\times\frac{\overline{\psi}(h/w')}{L^{(N)}(\psi^2,1+2it)}\sum_{d|h,(d,N)=1}\psi^2(d)d^{-2it},$$

where the superscript (N) indicates that the local factors at the primes dividing N have been removed. From this computation we deduce the bound ($t\in\mathbf{R}$):

$$\rho_{\mathfrak{a}}(h,t)\ll_\varepsilon(|h|N(1+|t|))^\varepsilon\mathrm{ch}\left(\frac{\pi|t|}{2}\right)\frac{(h,w)^{1/2}(w,N/w)}{\sqrt{wN|h|}}.$$

We can analyze $L^{dq'}_{\mathfrak{a},t}(\chi,s)$ as before, this time with $L(\widetilde{u}_j.\chi,s)$ replaced by

$$\sum_n\frac{\overline{\chi}(n)}{n^s}\sum_{ad=n}\overline{\psi}(a)\psi(d)(a/d)^{it}=L(\overline{\chi\psi},s-it)L(\overline{\chi}\psi,s+it)$$

and Theorem 9 replaced by the Burgess bound

$$|L(\overline{\chi\psi},s-it)L(\overline{\chi}\psi,s+it)|\ll_\varepsilon(|s|+|t|)q^{1/2-1/8+\varepsilon}\tag{5.21}$$

for $\Re es=1/2$ and $t\in\mathbf{R}$. Gathering these estimates we deduce that the contribution from the continuous spectrum to $\sum_{h\neq0}\overline{\chi}(h)S_{dq'h}$ in (5.12) is also bounded by (5.16); and by (5.3) we conclude the proof of Theorem 6. □

6. Equidistribution of Heegner points

In this section we apply our subconvexity estimates to prove equidistribution results for Heegner points on Shimura curves associated to definite quaternion algebras over $\mathbf{Q}$. For more details, we refer to the papers of Gross [G] and of Bertolini and Darmon [BD1].

6.1. *Definite Shimura curves.* We consider $q = q_1 \dots q_r$ a fixed square-free number and a fixed factorization $q = q^- q^+$ with q^- having an odd number of prime factors. Let B_{q^-} be the quaternion algebra ramified at the primes dividing q^- and at ∞. We fix $R_1 = R_{q^+,q^-}$ an Eichler order of B_{q^-}, $I_1 = R_1, I_2, \dots, I_n \subset R_1$ a set of representatives of (left) ideals classes and we denote R_i the right order of I_i.

This set corresponds to the set of connected components of a certain conic curve denoted X_{q^+,q^-} in [BD1]. We denote $\mathrm{Pic}(X_{q^+,q^-}) = \mathbf{Z}e_1 \oplus \dots \mathbf{Z}e_n$ the group of divisor classes where e_i corresponds to the class of a single point supported on the i-th component and $\mathrm{Pic}^0(X_{q^+,q^-})$ the kernel of the degree map. $\mathrm{Pic}(X_{q^+,q^-})$ is equipped with the inner pairing

$$\langle\ ,\ \rangle : \mathrm{Pic}(X_{q^+,q^-}) \times \mathrm{Pic}(X_{q^+,q^-}) \to \mathbf{Z}$$

given by $\langle e_i, e_j \rangle = \delta_{i,j} w_i$ with $w_i = |R_i^\times/\{\pm 1\}|$. The curve X_{q^+,q^-} is endowed with an action of a Hecke algebra $\mathbf{T}_{q^+,q^-}$ ([BD1, 1.5]) by correspondences, and the Hecke operators T_n, $(n,q) = 1$ are self-adjoint for the induced action on $\mathrm{Pic}(X_{q^+,q^-})$. Moreover (as a consequence of Eichler's trace formula) it is known that the image of this Hecke algebra into $\mathrm{End}(\mathrm{Pic}^0(X_{q^+,q^-}))$ is isomorphic to the Hecke algebra of $S_2^{q^-\mathrm{new}}(q)$ (the space of weight 2 holomorphic cusp forms of level q which are new at q^-). In particular (as a special case of the Jacquet-Langlands correspondence), for each primitive form $f \in S_2^p(q)$ there is a unique $e_f \in \mathrm{Pic}^0(X_{q^+,q^-}) \otimes_{\mathbf{Z}} \mathbf{R}$ such that $\langle e_f, e_f \rangle = 1$ and such that $T_n e_f = \lambda_f(n) e_f$ for every $(n,q) = 1$. For $q'|q^+$ we denote $\pi_{q'} : X_{q^+,q^-} \Rightarrow X_{q',q^-}$ the degeneracy map induced by the inclusion $R_{q',q^-} \subset R_{q^+,q^-}$ and $\pi_{q'}^*$ the map $\pi_{q'}^* : \mathrm{Pic}^0(X_{q^+,q^-}) \Rightarrow \mathrm{Pic}^0(X_{q',q^-})$ induced by contravariant functoriality. A basis of $\mathrm{Pic}^0(X_{q^+,q^-})$ is given by

$$\mathcal{B}_{q^+,q^-} = \{W_d^+ \pi_{q'}^*(e_f),\ dq'|q^+,\ f \in S_2^p(q'q^-)\}$$

where W_d^+ is the Atkin-Lehner operator ([BD1, 1.5]) and f ranges over the primitive forms of level $q'q^-$ for $q'd|q^+$.

6.2. *Gross-Heegner points.* Let K be an imaginary quadratic field of discriminant $-D$; denoting O_K the ring of integers, $\mathrm{Pic}(O_{K,c})$ the ideal class group and H_K the Hilbert class field, we have

$$D^{1/2-\varepsilon} \ll_\varepsilon |\mathrm{Pic}(O_K)| \ll D^{1/2} \log D, \tag{6.1}$$

the lower and upper bounds following from the Class Number Formula and Siegel's theorem.

A Gross-Heegner point (associated to the maximal order O_K)[4] is an optimal embedding $\xi : O_K \Rightarrow R_i$ of O_K into some R_i modulo conjugation by $R_i^\times$. By a well know recipe, a given Gross point determines a point in X_{q^+,q^-} and we still denote (with an abuse of notation) by ξ its natural image in $\mathrm{Pic}(X_{q^+,q^-})$, which is some e_{i_ξ}. The set $H_{q^+,q^-}(1)$ of Gross points is nonempty if and only if every prime p dividing q^- is inert in K and every p dividing q^+ is split (a condition which we assume for the rest of this section). In this case $H_{q^+,q^-}(1)$ is endowed with a free and transitive action of $\{\pm 1\}^r \times \mathrm{Pic}(O_K)$. For ξ a Gross point, and χ a character of $\mathrm{Pic}(O_K)$ we denote

$$\xi_\psi := \sum_{\sigma \in \mathrm{Pic}(O_K)} \overline{\psi}(\sigma)\xi^\sigma \in \mathrm{Pic}(X_{q^+,q^-}) \otimes_{\mathbf{Z}} \mathbf{C}$$

the ψ-eigen-component of ξ. The following formula due to Gross when q and D are primes ([G], p. 164) and subsequently generalized by Daghigh [Da] and Zhang [Z3] relates the central value of Rankin-Selberg L-functions to the position of ξ_ψ in $\mathrm{Pic}(X_{q^+,q^-}) \otimes_{\mathbf{Z}} \mathbf{C}$; more precisely for f a primitive form of level q,

$$|\langle \xi_\psi, e_f\rangle|^2 = \sqrt{D}\eta_f \frac{L(f\otimes g_\psi, 1/2)}{\langle f, f\rangle}; \tag{6.2}$$

here g_ψ is the theta series (of weight one, level D and nebentypus $\left(\frac{-D}{*}\right)$) associated to the character ψ and η_f is a certain positive factor depending on f only.

Theorem 3 is a particular case of the following:

THEOREM 10. *Let K be a quadratic field such that every prime p dividing q^- is inert in K and every p dividing q^+ is split. Consider a Gross point $\xi \in H_{q^+,q^-}(1)$ and a subgroup $G \subset \mathrm{Pic}(O_K)$ of index $\leqslant D^{\frac{1}{2115}}$; then as $D \to +\infty$, the orbit $G.\xi$ becomes equidistributed in the set $\{e_1, \ldots, e_n\}$ relative to the measure given by*

$$\mu_q(\{e_i\}) = w_i^{-1}/(\sum_{i=1}^n w_i^{-1}).$$

More precisely there exists an absolute constant $\eta > 0$ such that

$$\frac{|\{\sigma\in G,\ \xi^\sigma = e_i\}|}{|G|} = \mu_q(\{e_i\}) + O_q(D^{-\eta})$$

for any $i \in \{1, \ldots, n\}$. Here, the implied constant depends on q only.

[4] For simplicity we consider this case only; the general case of Gross-Heegner points with CM by a nonmaximal order is similar.

Proof. We consider the basis of $\mathrm{Pic}(X_{q^+,q^-})\otimes_{\mathbf{Z}}\mathbf{R}$ given by $\{e_0\}\cup\mathcal{B}_{q^+,q^-}$ where

$$e_0=(\sum_{i=1}^n w_i^{-1})^{-1/2}\sum_{i=1}^n w_i^{-1}e_i.$$

From the decomposition

$$e_i=\langle e_0,e_i\rangle e_0+\sum_{dq'|q^+}\sum_{f\in S_2^p(q'q-)}x_{f,d}W_d^+\pi_{q'}^*e_f$$

we deduce that

(6.3)

$$\begin{aligned} w_i\frac{|\{\sigma\in G,\ \xi^\sigma=e_i\}|}{|G|}&=\langle e_i,\frac{1}{|G|}\sum_{\sigma\in G}\xi^\sigma\rangle\\ &=\frac{1}{\sum_{i=1}^n w_i^{-1}}+\sum_{dq'|q^+}\sum_f x_{f,d}\langle W_d^+\pi_{q'}^*e_f,\frac{1}{|G|}\sum_{\sigma\in G}\xi^\sigma\rangle. \end{aligned}$$

Now,

$$\langle W_d^+\pi_{q'}^*e_f,\frac{1}{|G|}\sum_{\sigma\in G}\xi^\sigma\rangle=\langle\pi_{q'}^*e_f,\frac{1}{|G|}\sum_{\sigma\in G}W_d^+\xi^\sigma\rangle=\langle e_f,\frac{1}{|G|}\sum_{\sigma\in G}\xi'^\sigma\rangle$$

where $\xi'=\pi_{q'}(W_d^+\xi)$ defines a Gross point on X_{q',q^-}. By Fourier inversion,

$$\frac{1}{|G|}\sum_{\sigma\in G}\xi'^\sigma=\frac{1}{|G_K|}\sum_{\psi\in\widehat{G}_K}\Big(\frac{1}{|G|}\sum_{\sigma'\in G}\psi(\sigma')\Big)\xi'_\psi.$$

Since $\frac{1}{|G|}\sum_{\sigma'\in G}\psi(\sigma')$ is the characteristic function of characters which are trivial on G we deduce from (6.2) that

$$|\langle\frac{1}{|G|}\sum_{\sigma\in G}\xi^\sigma,W_d^+\pi_{q'}^*e_f\rangle|\ll\frac{1}{|G|}\operatorname*{Max}_\psi|D|^{1/4}\frac{\sqrt{c_f}L(f\otimes g_\psi,1/2)^{1/2}}{\langle f,f\rangle^{1/2}}.$$

When ψ is a real character g_ψ is an Eisenstein series and

$$L(f\otimes g_\psi,1/2)=L(f\otimes\chi_1,1/2)L(f\otimes\chi_2,1/2)$$

for some real Dirichlet characters χ_1,χ_2 such that $\chi_1,\chi_2=\big(\frac{-D}{*}\big)$; in this case, we use the bound of Theorem 3 of [DFI3] for each twist. When ψ is a complex character, g_ψ is cuspidal and we use the bound provided by Theorem 2 instead. In all cases

$$|\langle\frac{1}{|G|}\sum_{\sigma\in G}\xi^\sigma,W_d^+\pi_{q'}^*e_f\rangle|\ll_q\frac{|D|^{1/2-\frac{1}{2114}}}{|G|}\ll_q D^{-\eta}$$

for some $\eta>0$ if the index of G is $\leqslant|D|^{\frac{1}{2115}}$. The proof follows from this estimate and (6.3). $\square$

We now deduce Theorem 3 of the introduction from Theorem 10 applied for $q = q^-$ a prime number. The ideal $I_1, \dots, I_n$ corresponds to the n isomorphism classes of supersingular elliptic curves $e_1, \dots, e_n$ over $\overline{\mathbf{F}_q}$ and in this identification $End(e_i) = R_i$. Fix $\mathfrak{q}$ a prime in H_K above q. For $E \in \mathrm{Ell}(O_K)$ the reduction mod $\mathfrak{q}$, $\Psi_{\mathfrak{q}}$, defines an optimal embedding $\xi_{\mathfrak{q},E} : O_K \to R_{i(E)}$ by reduction of the endomorphism. Moreover (see [BD1, p. 120]) the action of $\mathrm{Pic}(O_K) = G_K$ commutes with the reduction map; for any $\sigma \in \mathrm{Pic}(O_K)$,

$$\xi_{\mathfrak{q},E}^{\sigma} = \xi_{\mathfrak{q},E^{\sigma}} = \Psi_{\mathfrak{q}}(E^{\sigma}) \tag{6.4}$$

and Theorem 3 follows.

7. Appendix

In this appendix we provide a proof for Theorem 9 which yields a subconvexity estimate in the q aspect for the Hecke L-function of a weight zero, primitive Maass form g, (which we normalize here by setting $\rho_g(1) = 1$), twisted by a primitive Dirichlet character χ of conductor q. Besides the subconvexity estimate, the main feature is an explicit polynomial dependence in the other parameters of g and of the complex variable s. We denote by D, and $1/4 + t^2$, respectively, the level of g and the eigenvalue of the Laplacian and we assume for simplicity that the nebentypus of g is trivial. Our proof follows closely [DFI1], [DFI2].

We prove here:

THEOREM 11. *Let g be a weight zero primitive Maass form of level D, trivial nebentypus, and eigenvalue $1/4 + t^2$, and let χ be a primitive character of modulus q. Denote by $g \otimes \chi$ the twist of g by χ. For $\Re e s = 1/2$,*

$$\begin{aligned} L(g \otimes \chi, s) \ll_{\varepsilon} &(|s|(1+|t|)Dq)^{\varepsilon} |s|^{E} (1+|t|)^{B} D^{C} D_1^{2/11} (D,q)^{1/11} q^{1/2-1/22+\varepsilon} \\ &+ (|s|(1+|t|)Dq)^{\varepsilon} |s|^{E'} (1+|t|)^{B'} D^{C'} D_1^{1/2} (D,q)^{1/4} q^{1/2-1/8+\varepsilon} \end{aligned}$$

for any $\varepsilon > 0$, the implied constant depending on ε only. Here D_1 is the integer defined in (7.5) below and the values of the exponents are given by

$$B = 14/11, \quad C = 1/4 \quad, E = 4/11, \tag{7.1}$$

$$B' = 7/2, \quad C' = 1/2 + 3/16 \quad, E' = 1. \tag{7.2}$$

Recall that $g \otimes \chi$ is a weight zero primitive Maass form with eigenvalue $1/4 + t^2$ and level Q dividing Dq^2, with nebentypus the Dirichlet character of modulus Q induced by (the underlying primitive character of) χ^2. The basic property of $g \otimes \chi$ is that for n coprime with (q, D), the n-th Hecke-eigenvalue satisfy

$$\lambda_{g \otimes \chi}(n) = \chi(n) \lambda_g(n).$$

Moreover we have the factorization

$$L(g\otimes\chi,s)=\sum_{n\geqslant 1}\frac{\lambda_{g\otimes\chi}(n)}{n^s}=\Big(\sum_{n|(qD)^\infty}\frac{\gamma_{g\otimes\chi}(n)}{n^s}\Big)\sum_{n\geqslant 1}\frac{\lambda_g(n)\chi(n)}{n^s}$$

where (by Hypothesis H_θ) the coefficients $\gamma_{g\otimes\chi}(n)$ satisfy

$$\gamma_{g\otimes\chi}(n)\ll_\varepsilon n^{\theta+\varepsilon}$$

for every $\varepsilon>0$, the implied constant depending only on ε. Its L-function $L(g\otimes\chi,s)$ satisfies a functional equation of the form

$$Q^{s/2}L_\infty(g,s)L(g\otimes\chi,s)=w(g\otimes\chi,)Q^{(1-s)/2}L_\infty(g,1-s)L(\overline{g\otimes\chi},1-s)$$

where $|w(g\otimes\chi)|=1$,

$$L_\infty(g,s)=\Gamma_{\mathbf{R}}(s+it+\frac{1-\varepsilon_g}{2})\Gamma_{\mathbf{R}}(s-it+\frac{1-\varepsilon_g}{2})$$

and ε_g is the eigenvalue of g under the involution $g(z)\to g(-\overline{z})$). Proceeding as in Section 3 we approximate $L(g\otimes\chi,s)$ for s on the critical line by partial sums of length $\simeq Q^{1/2}$ and obtain the following estimate

$$L(g\otimes\chi,s)\ll_{A_0}\log(qD|s|)\sum_N\frac{|L_{g.\chi}(N)|}{\sqrt{N}}\Big(1+\frac{N}{P\sqrt{Q}}\Big)^{-A_0}$$

where A_0 is a constant that can be taken arbitrarily large, $P=|s|+|t|$, $N=2^\nu$, $\nu\geqslant -1$, and

$$L_{g.\chi}(N)=\sum_n\chi(n)\rho_g(n)n^{1/2}V(n).$$

Here $V(x)=V_N(x)$ is some smooth function supported on $[N/2,5N/2]$, such that for all $j\geqslant 0$,

$$x^jV^{(j)}(x)\ll_{j,A_0}P^j.$$

In particular the convexity bound gives

$$L(g\otimes\chi,s)\ll_\varepsilon(qD(1+|t|)|s|)^\varepsilon q^{1/2}D^{1/4}(|s|+|t|)^{1/2}.$$

We now bound $L_{g.\chi}(N)$, by using the amplification and the δ-symbol methods of [DFI1], [DFI2] with the appropriate generalization given in [KMV2], [M1]. To this end we consider the quadratic form

$$Q(\vec{x},N)=\sideset{}{^*}\sum_{\chi(q)}|L_{g.\chi}(N)|^2|\sum_{\ell\leqslant L}x_\ell\chi(\ell)|^2$$

where the χ range over the primitive characters of modulus q, and the x_ℓ are complex numbers of modulus less than 1 such that

$$x_\ell=0,\ \text{unless}\ \ (\ell,qD)=1\text{ and }L/2<\ell\leqslant L. \tag{7.3}$$

We prove below

(7.4)
$$Q(\vec{x},N)\ll_\varepsilon (qPN)^\varepsilon(qNL+(1+|t|)^2(|t|^{17/4}+|s|^{5/4})DD_1(D_0,q)^{1/2}L^{2+7/4}N^{7/4},$$

where

$$D_0=\prod_{p|D^w}p^{v_p(D^w)-1}=\prod_{p|D}p^{v_p(D)-1},\quad D_1=\prod_{\substack{p|D\\ v_p(D)\geqslant 2}}p^{v_p(D)-2}. \tag{7.5}$$

From the trivial bound

$$|L_{g.\chi}(N)|^2|\sum_{\ell\leqslant L}x_\ell\chi(\ell)|^2\leqslant Q(\vec{x},N)$$

we obtain (on choosing the classical amplifier given by $x_\ell=\overline{\chi}(\ell)$ for $\ell\in\,]L/2,L]$, such that $(\ell,D)=1$ and $\ell=0$ otherwise), the bound

$$L_{g.\chi}(N)\ll_\varepsilon (qPN)^\varepsilon(\Big(\frac{qN}{L}\Big)^{1/2}$$
$$+(1+|t|)(|t|^{17/8}+|s|^{5/8})(DD_1)^{1/2}(D_0,q)^{1/4}L^{7/8}N^{7/8}.$$

We then set

$$M=(1+|t|)(|t|^{17/8}+|s|^{5/8})(DD_1)^{1/2}(D_0,q)^{1/4}$$

and choose $L=1+q^{4/11}M^{-8/11}N^{-3/11}$ so that

$$\frac{L_{g.\chi}(N)}{\sqrt{N}}\ll_\varepsilon (qPN)^\varepsilon\Big(q^{1/2-4/22}M^{4/11}N^{3/22}+MN^{3/8}\Big),$$

which is sufficient.

7.1. *Treatment of the quadratic form.* We have (see for example [DFI1, p. 4])

$$Q(\vec{x},N)\leqslant\phi(q)\sum_{h\equiv 0(q)}\sum_{\ell_1,\ell_2\leqslant L}x_{\ell_1}\overline{x}_{\ell_2}S_h(\ell_1,\ell_2)$$

with

$$S_h(\ell_1,\ell_2)=\sum_{\ell_1m-\ell_2n=h}\lambda_g(m)\lambda_g(n)V(m)V(n).$$

By (2.27) the contribution from the $h=0$ term is bounded by

$$\ll_\varepsilon (DLN(1+|t|))^\varepsilon qN\sum_\ell|x_\ell|^2\ll_\varepsilon (DLN(1+|t|))^\varepsilon qNL. \tag{7.6}$$

For $h\neq 0$, we proceed to bound $S_h(\ell_1,\ell_2)$; we can rewrite this sum as

$$S_h(\ell_1,\ell_2)=\sum_{\ell_1m-\ell_2n=h}\lambda_g(m)\lambda_g(n)F(\ell_1m,\ell_2n),$$

with

$$F(x,y) = V(x/\ell_1)V(y/\ell_2)\varphi(x-y-h)$$

where $\varphi(u)$ is a smooth function supported on $|u| < U = LNP^{-1}$ such that $\varphi(0) = 1$ and $\varphi^{(i)}(u) \ll_i U^{-i}$ for all $i \geqslant 0$. In particular,

$$F^{(i,j)}(x,y) \ll_{i,j} U^{-(i+j)}$$

for all $i, j \geqslant 0$. From the δ-symbol method we get

$$S_h(\ell_1,\ell_2) = \sum_{1\leqslant c<C} \sum_{\substack{d(c)\\(d,c)=1}} e\left(\frac{-dh}{c}\right) \sum_{m,n} \lambda_g(m)\lambda_g(n) e\left(\frac{d\ell_1 m - d\ell_2 n}{c}\right) E(m,n)$$

with $C = 2U^{1/2} \leqslant 2\sqrt{LN/P}$ and $E(x,y) = F(\ell_1 x, \ell_2 y)\Delta_c(\ell_1 x - \ell_2 y - h)$ where $\Delta_c(u)$ is the function defined in (11) of [DFI2].

7.1.1. *A summation formula.* We will transform the above sum by means of a summation formula, for this we need the following refinement of Theorem A.4 of [KMV2]. We define the "wild" part of D to be

$$D^w := \prod_{\substack{p|D\\v_p(D)>1}} p^{v_p(D)}.$$

For g primitive of level D we have, by [Li1, Th. 3, p. 295],

$$\rho_g(n) = 0 \text{ whenever } (n, D^w) \neq 1. \tag{7.7}$$

Proposition 7.1. *Let D be a positive integer, and g a primitive weight zero Maass form of level D and trivial nebentypus. For $(a,c)=1$, set*

$$c^\flat := \prod_{\substack{p|(c,D^w)\\v_p(c)<v_p(D)}} p^{v_p(c)},\ c = c^\sharp c^\flat,\ D^\sharp = (c^\sharp, D),\ D^\flat = D/D^\sharp, \tag{7.8}$$

$\tilde{D} = \mathrm{l.c.m.}[D, (c^\flat)^2]$, $\tilde{D}^\flat = \tilde{D}/D^\sharp$ *(note that* $(c^\sharp, c^\flat) = (D^\sharp, D^\flat) = (D^\sharp, \tilde{D}^\flat) = 1$*).*

For $F \in C^\infty(\mathbf{R}^{*+})$ a Schwartz class function vanishing in a neighborhood of zero we have the identity

$$c\sum_{n\geqslant 1} \rho_g(n) n^{1/2} e(n\frac{a}{c}) F(n) \tag{7.9}$$

$$= \frac{1}{\sqrt{\tilde{D}^\flat}} \sum_{\pm} (\varepsilon_g)^{\pm 1} \sum_{n\geqslant 1} \kappa_g(n,a,c) e\left(\frac{\pm n c^\flat \overline{a\tilde{D}^\flat}}{c^\sharp}\right) \int_0^\infty F(x) J^{\pm}_{2it}\left(\frac{4\pi}{c^\sharp}\sqrt{\frac{nx}{\tilde{D}^\flat}}\right) dx.$$

In the above expression,

- ε_g *denotes the eigenvalue of g under the reflection operator,*

- $$J^-_{2it}(x) = \frac{-\pi}{\mathrm{ch}(\pi t)}(Y_{2it}(x) + Y_{-2it}(x)),\ J^+_{2it}(x) = 4\mathrm{ch}(\pi t)K_{2it}(x).$$

- $$\kappa_g(n,a,c) = \frac{c^\flat}{\varphi(c^\flat)}\sum_{\psi(c^\flat)}\psi(a)\psi(c^\sharp)G_{\overline{\psi}}(c^\flat)\rho_{g.\psi_{|W_{\tilde{D}^\flat}}}(n)n^{1/2},$$

where ψ runs over the Dirichlet characters of modulus $c^\flat$, $G_{\overline{\psi}}(c^\flat)$ is the Gauss sum, $g.\psi$ denotes the twist of g by the character ψ (which is a form of level $\tilde{D}$) and $W_{\tilde{D}^\flat}$ denotes the Atkin-Lehner operator acting on forms of level $\tilde{D}$.

Proof. Since $(c^\sharp, c^\flat) = 1$,

$$S(g,\frac{a}{c}) = \sum_n \rho_g(n)n^{1/2}e\left(\frac{na}{c^\sharp c^\flat}\right)F(n) = \sum_n \rho_g(n)n^{1/2}e\left(\frac{na\overline{c^\sharp}}{c^\flat} + \frac{na\overline{c^\flat}}{c^\sharp}\right).$$

By our assumption (7.7) the n-sum runs over integers coprime with $c^\flat$, which allows us to transform easily the additive character $e\left(\frac{na\overline{c^\sharp}}{c^\flat}\right)$ into multiplicative ones

$$S(g,\frac{a}{c}) = \frac{1}{\varphi(c^\flat)}\sum_{\psi(mod\ c^\flat)}\psi(a\overline{c^\sharp})G_{\overline{\psi}}(c^\flat)\sum_n \rho_g(n)n^{1/2}\psi(n)e\left(\frac{na\overline{c^\flat}}{c^\sharp}\right)F(n).$$

By [ALi, Prop. 3.1], the twisted form $g.\psi$ has level $\tilde{D}$ and nebentypus ψ^2. Since $(c^\sharp, \tilde{D}) = D^\sharp$ is coprime with $\tilde{D}/D^\sharp = \tilde{D}^\flat$, we may apply Theorem A.4 of [KMV2], and the proof follows. □

7.1.2. *Transformation of the double sum.* In view of the above summation formula we set

$$c_1 = c/(c,\ell_1),\ c_2 = c/(c,\ell_2),\ l_1 = \ell_1/(c,\ell_1),\ l_2 = \ell_2/(c,\ell_2)$$

and apply Proposition 7.1 to both variables to get a sum of four terms of the form

$$\sum_{1\leqslant c<C}\frac{1}{c_1c_2\tilde{D}^\flat}\sum_{n_1,n_2\geqslant 1}\Sigma^{\pm,\pm}(n_1,n_2)I^{\pm,\pm}(n_1,n_2)$$

with

$$\begin{aligned}(7.10)\quad \Sigma^{\pm,\pm}(n_1,n_2) = &\sum_{d(c)}e\left(\frac{-dh}{c}\right)\sum_{n_1,n_2\geqslant 1}\kappa_g(n_1,dl_1,c_1)\kappa_g(n_2,dl_2,c_2)\\ &\times e\left(\frac{\pm n_1c^\flat\overline{dl_1\tilde{D}^\flat}}{c_1^\sharp} + \frac{\pm n_2c^\flat\overline{dl_2\tilde{D}^\flat}}{c_2^\sharp}\right)\end{aligned}$$

(7.11)
$$I^{\pm,\pm}(n_1,n_2)=\int_0^\infty\int_0^\infty E(x,y)J_{2it}^{\pm}\left(\frac{4\pi}{c_1^\sharp}\sqrt{\frac{n_1x}{\tilde{D}^\flat}}\right)J_{2it}^{\pm}\left(\frac{4\pi}{c_2^\sharp}\sqrt{\frac{n_2y}{\tilde{D}^\flat}}\right)dxdy,$$

and where $c_1=c_1^\sharp c_1^\flat$, $c_2=c_2^\sharp c_2^\flat$ as in (7.8). Note also that since $(\ell_1\ell_2,D)=1$, we have $c_1^\flat=c_2^\flat:=c^\flat$; hence the corresponding decompositions for D and $\tilde{D}$ coincide; $D=D_1^\sharp D_1^\flat=D_2^\sharp D_2^\flat$ with $D_1^\flat=D_2^\flat=D^\flat$, and $\tilde{D}_2^\flat=\tilde{D}_2^\flat=\tilde{D}^\flat$.

For each $m,n\geqslant 1$,

$$\Sigma^{\pm,\pm}(n_1,n_2)=\left(\frac{c^\flat}{\varphi(c^\flat)}\right)^2$$
$$\times\sum_{\psi,\psi'(c^\flat)}\psi\psi'(c^\sharp)G_{\overline{\psi}}(c^\flat)G_{\overline{\psi}'}(c^\flat)\rho_{g.\psi_{|W_{\tilde{D}^\flat}}}(n_1)n_1^{1/2}\rho_{g.\psi_{|W_{\tilde{D}^\flat}}}(n_2)n_2^{1/2}$$
$$\times\sum_{d(c)}\psi\psi'(d)e\left(\frac{-dh}{c}\pm\frac{n_1c^\flat\overline{dl_1\tilde{D}^\flat}}{c_1^\sharp}\pm\frac{n_2c^\flat\overline{dl_2\tilde{D}^\flat}}{c_2^\sharp}\right)$$

the innermost d sum is a Kloosterman type sum and is bounded by

$$\ll_\varepsilon c^{1/2+\varepsilon}(c,h)^{1/2};$$

hence

$$\Sigma^{\pm,\pm}(n_1,n_2)\ll_\varepsilon c^{1/2+\varepsilon}(c,h)^{1/2}(c^\flat)^2\sum_{\psi(c^\flat)}|\rho_{g.\psi_{|W_{\tilde{D}^\flat}}}(n_1)n_1^{1/2}|^2 +|\rho_{g.\psi_{|W_{\tilde{D}^\flat}}}(n_2)n_2^{1/2}|^2. \tag{7.12}$$

We now bound the analytic term $I^{\pm,\pm}(n_1,n_2)$ using the following estimates for the Bessel type functions (cf. [I2, p. 227]); for $\sigma,r\in\mathbf{R}$,

$$Y_{\sigma+2it}(y)\ll y^{-1/2}e^{\frac{\pi}{2}|t|},\ K_{\sigma+2it}(y)\ll y^{-1/2}e^{-y} \tag{7.13}$$

for $y>1+|\sigma|^2+4|t|^2$, the implied constant being absolute. In particular we get

$$J_{2it}^{\pm}(y)\ll y^{-1/2} \tag{7.14}$$

for $y>1+4|t|^2$, the implied constant being absolute. For the remaining range we will also use the following general bound:

$$J_{2it}^{\pm}(y)\ll_\varepsilon\left(\frac{1+|t|}{y}\right)^{\pm\varepsilon} \tag{7.15}$$

for every $r\in\mathbf{R}$ and $\varepsilon>0$, the implied constant depending only on ε. These last bounds follow from the integral representations

$$J_{2it}^{+}(y)=2\mathrm{ch}(\pi t)\frac{1}{2\pi i}\int_{(\varepsilon)}\Gamma\left(\frac{s}{2}+it\right)\Gamma\left(\frac{s}{2}-it\right)\left(\frac{y}{2}\right)^{-s}ds$$

and

$$J^-_{2it}(y) = \frac{1}{2\pi i}\int_{(\varepsilon)} \Gamma\left(\frac{s}{2}+it\right)\Gamma\left(\frac{s}{2}-it\right)\cos(\pi\frac{s}{2})\left(\frac{y}{2}\right)^{-s} ds$$

(for J^- we obtain the bound by shifting the contour to $\Re es = -\varepsilon$: we meet two poles at $s = \pm 2it$ whose residues are bounded by $(y/(1+|t|))^{\varepsilon}$). By several integrations by parts, using the recurrence relations

$$(z^\nu K_\nu(z))' = -z^\nu K_{\nu-1}(z),\ (z^\nu Y_\nu(z))' = -z^\nu Y_{\nu-1}(z),$$

the estimate $E^{(i,j)} \ll_{i,j} (cC)^{-(1+i+j)}\ell_1^i\ell_2^j$ and (7.13), we see that $I^{\pm,\pm}(n_1,n_2)$ is very small unless

$$n_i \ll_\varepsilon (qP)^\varepsilon \frac{(c_i^\sharp)^2\tilde{D}^\flat}{N}\left(1+|t|^4+\frac{C^2}{c^2}P^2\right),\ i=1,2. \tag{7.16}$$

When the variables $n_i,\ i=1,2$ satisfy (7.16), we use the bounds (7.14) and (7.15) without integrating by parts to get

$$\begin{aligned}
&I^{\pm,\pm}(n_1,n_2)\\
&\quad\ll_\varepsilon (Pq)^\varepsilon ||E||_1\left(1+\frac{\sqrt{Nn_1}}{(1+|t|^2)c_1^\sharp\sqrt{\tilde{D}^\flat}}\right)^{-1/2}\left(1+\frac{\sqrt{Nn_2}}{(1+|t|^2)c_2^\sharp\sqrt{\tilde{D}^\flat}}\right)^{-1/2}\\
&\quad\ll_\varepsilon (Pq)^{2\varepsilon}\frac{N}{L}\left(1+\frac{\sqrt{Nn_1}}{(1+|t|^2)c_1^\sharp\sqrt{\tilde{D}^\flat}}\right)^{-1/2}\left(1+\frac{\sqrt{Nn_2}}{(1+|t|^2)c_2^\sharp\sqrt{\tilde{D}^\flat}}\right)^{-1/2},
\end{aligned}$$

by (30) of [DFI2]. Using this bound, (7.14) and

$$\sum_{n\leqslant X}|\rho_{g.\psi|W_{\tilde{D}^\flat}}(n)n^{1/2}|^2 \ll_\varepsilon (qDX(1+|t|))^\varepsilon X$$

we obtain

$$\begin{aligned}
S_h(\ell_1,\ell_2) &\ll_\varepsilon (qDP)^\varepsilon \sum_{c\leqslant C}\frac{(c,h)^{1/2}c^{1/2}c^{\flat 3}}{c_1^\sharp c_2^\sharp\tilde{D}^\flat}\frac{N}{L}(1+|t|)^2\\
&\qquad\times\frac{(c_1^\sharp c_2^\sharp\tilde{D}^\flat)^{1/2}}{N^{1/2}}\frac{(c_1^\sharp c_2^\sharp\tilde{D}^\flat)^{3/2}}{N^{3/2}}\left(|t|^6+\frac{P^3C^3}{c^3}\right)\\
&\ll_\varepsilon (qDP)^\varepsilon\frac{D(1+|t|)^2}{LN}\sum_{c\leqslant C}(c,h)^{1/2}c^{2+1/2}c^{\flat 3}\left(|t|^6+\frac{P^3C^3}{c^3}\right)\\
&\ll_\varepsilon (qDP)^\varepsilon DD_1(D_0,h)^{1/2}(1+|t|)^2(|t|^6+P^3)\frac{(LN)^{3/4}}{P^{7/4}}
\end{aligned}$$

where D_0, D_1 are as defined in (7.5) (we have used here that $c_i^\sharp c^\flat = c_i \leqslant c$ and $\tilde{D}^\flat \leqslant \tilde{D} \leqslant DD_1$). Summing over $h \equiv 0(q)$, $h \neq 0$ and ℓ_1,ℓ_2 we obtain finally (7.4).

In the (improbable) case where $2it \in \mathbf{R}$ (*i.e.* g is an exceptional eigenform and so $2it \in [-2\theta, 2\theta]$) we proceed as above to obtain the same bound. In particular we use in the range

$$n_i \ll_\varepsilon (qP)^\varepsilon \frac{(c_i^\sharp)^2 \tilde{D}^\flat}{N} \frac{C^2}{c^2} P^2,\ i = 1, 2,$$

the bound

$$J^{\pm}_{2it}(y) \ll \min(y^{-2\theta}, y^{-1/2}) \leqslant y^{-1/2}$$

(since one can take $\theta \leqslant 1/4$) the implied constant being absolute.

Université Montpellier II, Montpellier, France
E-mail address: michel@math.univ-montp2.fr

References

[AL] A. Atkin and J. Lehner, Hecke operators on $\Gamma_0(N)$, *Math. Ann.* **185** (1970), 134–160.

[ALi] A. Atkin and W. Li, Twists of newforms and pseudo-eigenvalues of W-operators, *Invent. Math.* **48** (1978), 221–243.

[BD1] M. Bertolini and H. Darmon, Heegner points of Mumford-Tate curves, *Invent. Math.* **126** (1996), 413–456.

[BD2] ———, A rigid analytic Gross-Zagier formula and arithmetic applications (with an appendix by Bas Edixhoven), *Ann. of Math.* **146** (1997), 111–147.

[BD3] ———, Iwasawa's main conjecture for elliptic curves over anticyclotomic $\mathbf{Z}_p$-extensions, *Ann. of Math.*, to appear.

[Bu] D. A. Burgess, On character sums and L-series, *Proc. London Math. Soc.* **12** (1962), 193–206.

[Co] C. Cornut, Mazur's conjecture on higher Heegner points, *Invent. Math.* **148** (2002), 495–523.

[CPSS] J. Cogdell, I. I. Piatetskii-Shapiro, and P. Sarnak, Estimates for Hilbert modular L-functions and applications, in preparation.

[CI] B. Conrey and H. Iwaniec, The cubic moment of central values of automorphic L-functions, *Ann. of Math.* **151** (2000), 1175–1216.

[Da] H. Daghigh, Modular forms, quaternion algebras, and special values of L-functions, Ph.D. thesis, Mac Gill University, 1998.

[DI] J.-M. Deshouillers and H. Iwaniec, Kloosterman sums and Fourier coefficients of cusp forms, *Invent. Math.* **70** (1982/83), 219–288.

[Du] W. Duke, Hyperbolic distribution problems and half-integral weight Maass forms, *Invent. Math.* **92** (1988), 73–90.

[DFI1] W. Duke, J. Friedlander, and H. Iwaniec, Bounds for automorphic L-functions, *Invent. Math.* **112** (1993), 1–8.

[DFI2] ———, A quadratic divisor problem, *Invent. Math.* **115** (1994), 209–217.

[DFI3] ———, Bounds for automorphic L-functions, II, *Invent. Math.* **115** (1994), 219–239.

[DFI4] W. Duke, J. Friedlander, and H. Iwaniec, Class group L-functions, *Duke Math. J.* **79** (1995), 1–56.

[DFI5] ———, Bilinear forms with Kloosterman fractions, *Invent. Math.* **128** (1997), 23–43.

[DFI6] ———, Representations by the determinant and mean values of L-functions, in *Sieve Methods, Exponential Sums, and their Applications in Number Theory* (Cardiff, 1995), 109–115, *London Math. Soc. Lecture Note Ser.* **227**, Cambridge Univ. Press, Cambridge, 1997.

[DFI7] W. Duke, J. Friedlander, and H. Iwaniec, Bounds for automorphic L-functions, III, *Invent. Math.* **143** (2001), 221–248.

[DFI8] ———, The subconvexity problem for Artin L-functions, *Invent. Math.* **149** (2002), 489–577.

[Fr] J. Friedlander, Bounds for L-functions, *Proc. of the Internat. Congress of Mathematicians* **1**, **2** (Zürich, 1994), 363–373, Birkhäuser, Basel, 1995.

[GL] D. Goldfeld and D. Lieman, Effective bounds on the size of the Tate-Shafarevich group, *Math. Res. Lett.* **3** (1996), 309–318.

[G] B. Gross, Heights and the special values of L-series, in *Number Theory* (Montreal, Que., 1985), 115–187, *CMS Conf. Proc.* **7** A. M. S., Providence, RI, 1987.

[GR] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press Inc., San Diego, CA (1980).

[GZ] B. Gross and D. Zagier, Heegner points and derivatives of L-series, *Invent. Math.* **84** (1986), 225–320.

[H] G. Harcos, An additive problem in the Fourier coefficients of Maass forms, *Math. Ann.* **326** (2003), 347–365; arXiv:math.NT/0101096.

[HL] J. Hoffstein and P. Lockhart, Coefficients of Maass forms and the Siegel zero (with an appendix by D. Goldfield, J. Hoffstein, and D. Lieman), *Ann. of Math.* **140** (1994), 161–181.

[I1] H. Iwaniec, Fourier coefficients of modular forms of half-integral weight, *Invent. Math.* **87** (1987), 385–401.

[I2] ———, *Introduction to the Spectral Theory of Automorphic Forms*, *Biblioteca de la Revista Matemática Iberoamericana*, Revista Matemática Iberoamericana, Madrid, 1995.

[ILS] H. Iwaniec, W. Luo, and P. Sarnak, Low lying zeros of families of L-functions, *Publ. Math. IHES* **91** (2000), 55–131, (2001).

[IS] H. Iwaniec and P. Sarnak, Perspectives in the analytic theory of L-functions, Special Volume, Part II *GAFA 2000* (Tel Aviv, 1999), *Geom. Funct. Anal.* **2000**, 705–741.

[J] H. Jacquet, *Automorphic Forms of* GL(2), *Part II*, *Lecture Notes in Math.* **278**, Springer-Verlag, New York, 1972.

[JS] H. Jacquet and J. A. Shalika, On Euler products and the classification of automorphic representations. I, *Amer. J. Math.* **103** 1981), 499–558.

[JPPS] H. Jacquet, I. I. Piatetskii-Shapiro, and J. A. Shalika, Rankin-Selberg convolutions, *Amer. J. Math.* **105** (1983), 367–464.

[Ki] H. Kim, Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2, *J. Amer. Math. Soc.* **16** (2003), 139–183 (electronic).

[KiSa] H. Kim and P. Sarnak, Appendix: Refined estimates towards the Ramanujan and Selberg conjectures, *J. Amer. Math. Soc.*, to appear.

[KiSh] H. Kim and F. Shahidi, Cuspidality of symmetric powers with applications, *Duke Math. J*. **112** (2002), 177–197.

[KMV1] E. Kowalski, P. Michel, and J. VanderKam, Mollification of the fourth moment of automorphic L-functions and arithmetic applications, *Invent. Math*. **142** (2000), 95–151.

[KMV2] ———, Rankin-Selberg L-functions in the level aspect, *Duke Math. J*. **114** (2002), 123–191.

[Li1] W. Li, Newforms and functional equations, *Math. Ann*. **212** (1975), 285–315.

[Li2] ———, L-series of Rankin type and their functional equations, *Math. Ann*. **244** (1979), 135–166.

[Lin] Y. V. Linnik, *Ergodic Properties of Algebraic Fields*, *Ergebnisse Math*. **45**, Springer-Verlag, New York (1968).

[M1] P. Michel, Complements to "On the Rankin-Selberg L-functions in the level aspect", Unpublished Notes; Available at http://www.math.univ-montp2.fr/∼michel/publi.html.

[M2] ———, Familles de fonctions L de formes et applications, *Proc. Journées Arithmétiques de Lille*, *J. Th. Nombres Bordeaux* **15** (2003), 275–307.

[MW] C. Mœglin and J.-L. Waldspurger, *Spectral Decomposition and Eisenstein Series*: *Une paraphrase de l'Ecriture*, *Cambridge Tracts in Math*. **113**, Cambridge Univ. Press, Cambridge (1995).

[Ram] D. Ramakrishnan, Modularity of the Rankin-Selberg L-series, and multiplicity one for SL(2), *Ann. of Math*. **152** (2000), 45–111.

[Ra1] R. A Rankin, Diagonalizing Eisenstein series. I, in *Analytic Number Theory* (Allerton Park, IL, 1989), 429–450, *Progr. Math*. **85** (1990), Birkhäuser Boston, Boston, MA.

[Ra2] ———, Diagonalizing Eisenstein series. II. A tribute to Emil Grosswald, in *Number Theory and Related Analysis*, 525–537, *Contemp. Math*. **143** 143 (1993), A. M. S., Providence, RI.

[Ra3] ———, Diagonalizing Eisenstein series. III, in *Discrete Groups and Geometry* (Birmingham, 1991), 196–208, *London Math. Soc. Lecture Note Ser*. **173** (1992), Cambridge Univ. Press, Cambridge.

[Ra4] ———, Diagonalizing Eisenstein series. IV, in *The Rademacher Legacy to Mathematics* (University Park, PA, 1992), 107–118, *Contemp. Math*. **166** (1994), A. M. S., Providence, RI.

[Ri] K. A. Ribet, On modular representations of $\mathrm{Gal}(\overline{Q}/Q)$ arising from modular forms, *Invent. Math*. **100** (1990), 431–476.

[Sa1] P. Sarnak, Integrals of products of eigenfunctions, *IMRN* **6** (1994), 251–260.

[Sa2] ———, Estimates for Rankin-Selberg L-functions and Quantum Unique Ergodicity, *J. Funct. Anal*. **184** (2001), 419–453.

[Se] A. Selberg, On the estimation of Fourier coefficients of modular forms (1965), *Proc. Sympos. Pure Math*. **VIII**, 1–15, A. M. S., Providence, RI.

[Va] V. Vatsal, Uniform distribution of Heegner points, *Invent. Math*. **148** (2002), 1–46.

[Wa] J.-L. Waldspurger, Sur les coefficients de Fourier des formes modulaires de poids demi-entier, *J. Math. Pures et Appliquées* **60** (1981), 374–484.

[Z1] S. Zhang, Heights of Heegner cycles and derivatives of L-series, *Invent. Math*. **130** (1997), 99–152.

[Z2] S. Zhang, Heights of Heegner points on Shimura curves, *Ann. of Math.* **153** (2001), 27–147.

[Z3] ———, Gross-Zagier formula for GL(2), *Asian J. Math.* **5** (2001), 183–290.

(Received November 22, 2001)
(Revised December 5, 2002)