SHIFTED CONVOLUTION SUMS FOR $GL(3) \times GL(2)$

RITABRATA MUNSHI

Abstract. For the shifted convolution sum

$$D_h(X) = \sum_{m=1}^{\infty} \lambda_1(1,m)\lambda_2(m+h)V\left(\frac{m}{X}\right),$$

where $\lambda_1(1,m)$ are the Fourier coefficients of a $SL(3, \mathbb{Z})$ Maass form $\pi_1$, and $\lambda_2(m)$ are those of a $SL(2, \mathbb{Z})$ Maass or holomorphic form $\pi_2$, and $1 \leq |h| \ll X^{1+\varepsilon}$, we establish the bound

$$D_h(X) \ll \pi_1, \pi_2, \varepsilon X^{1-\frac{1}{20}+\varepsilon}.$$

The bound is uniform with respect to the shift $h$.

1. Introduction

The shifted convolution sum

$$\sum_{m=1}^{\infty} \lambda_1(m)\lambda_2(m+h)V\left(\frac{m}{X}\right),$$

with $GL(2)$ Fourier coefficients $\lambda_i(m)$, has been investigated extensively by several authors since Selberg’s seminal paper [25]. Non-trivial bound of this sum often has deep implications, e.g. subconvexity and equidistribution (QUE) (see [2], [6], [7], [11], [12], [13], [14], [16], [17], [18], [20], [24]).

In this paper we will consider a higher rank analogue -

$$D_h(X) := \sum_{m=1}^{\infty} \lambda_1(1,m)\lambda_2(m+h)V\left(\frac{m}{X}\right)$$

where $\lambda_1(1,m)$ are the Fourier coefficients of a $SL(3, \mathbb{Z})$ Hecke-Maass cusp form $\pi_1$, and $\lambda_2(m)$ are those of a $SL(2, \mathbb{Z})$ Hecke-Maass or Hecke holomorphic cusp form $\pi_2$. We will take $V$ to be smooth and compactly supported in $[1,2]$. Also we will take $0 \leq h \ll X^{1+\varepsilon}$, as for larger shifts the trivial bound most often suffices. (Pitt [23] has considered a similar sum with $\tau_3(m)$ in place of the Fourier coefficients $\lambda_1(1,m)$.) Applying Cauchy and estimates coming from Rankin-Selberg theory we obtain the following (trivial) estimate

$$D_h(X) \ll \pi_1, \pi_2 X^{1+\varepsilon}.$$  

(1)

Our main theorem gives a nontrivial power saving over this estimate.

Theorem 1. For $0 \leq h \ll X^{1+\varepsilon}$, we have

$$D_h(X) \ll \pi_1, \pi_2, \varepsilon X^{1-\frac{1}{20}+\varepsilon}.$$
Remark 6. This seemingly simple idea has other important applications. In [22] we apply this idea to several subconvexity problems.

In the theorem, the dependence of the implied constant on the conductors of \( \pi_1 \) and \( \pi_2 \) can be explicitly given, though we do not try to do this here. In fact, in some of the estimates that we use, e.g. Lemma 2, the implied constants are independent of the conductor of the form. Moreover it is not necessary to assume that \( \pi_2 \) is of full level. The same bound holds for general \( \pi_2 \). It is expected that extra cancellation can be obtained by averaging over \( h \). However we will not take up this issue in this paper.

Acknowledgements. I thank Valentin Blomer, Tim Browning, Gergely Harcos and Matthew Young for their helpful comments.

2. Preliminaries

2.1. Preliminaries on \( SL(3, \mathbb{Z}) \) Maass forms. We shall first recall some basic facts about \( SL(3, \mathbb{Z}) \) automorphic forms. Our need is minimal and, in fact, the Voronoi summation formula (of Miller and Schmid [21], and Goldfeld and Li [10]) is all that we will be using. Suppose \( \pi_1 \) is a Maass form of type \((\nu_1, \nu_2)\) for \( SL(3, \mathbb{Z}) \) which is an eigenfunction of all the Hecke operators with Fourier coefficients \( \lambda_1(m_1, m_2) \), normalized so that \( \lambda_1(1,1) = 1 \) (for details see Goldfeld’s book [9]). We introduce the Langlands parameters \((\alpha_1, \alpha_2, \alpha_3)\), defined by

\[
\alpha_1 = -\nu_1 - 2\nu_2 + 1, \quad \alpha_2 = -\nu_1 + \nu_2 \quad \text{and} \quad \alpha_3 = 2\nu_1 + \nu_2 - 1.
\]

The Ramanujan-Selberg conjecture predicts that \( |\Re(\alpha_i)| = 0 \), and from the work of Jacquet and Shalika we at least know that \( |\Re(\alpha_i)| < \frac{1}{2} \).

Let \( g \) be a compactly supported function on \((0, \infty)\), and let \( \tilde{g}(s) = \int_0^\infty g(x)x^{-s}dx \) be the Mellin transform. For \( \sigma > -1 + \max\{ -\Re(\alpha_1), -\Re(\alpha_2), -\Re(\alpha_3) \} \) and \( \ell = 0, 1 \) define

\[
G_\ell(y) = \frac{1}{2\pi i} \int_{(\sigma)} (\pi^3 y)^{-s} \Gamma \left( \frac{1+s+\alpha_1+\ell}{2} \right) \Gamma \left( \frac{1+s+\alpha_2}{2} \right) \Gamma \left( \frac{1+s+\alpha_3+\ell}{2} \right) \tilde{g}(-s)ds
\]

and set

\[
G_+ (y) = \frac{1}{2\pi^{3/2}} (G_0(y) - iG_1(y)), \quad \text{and} \quad G_- (y) = \frac{1}{2\pi^{3/2}} (G_0(y) + iG_1(y)).
\]

Lemma 1. Let \( g \) be a compactly supported smooth function on \((0, \infty)\), we have

\[
\sum_{m=1}^\infty \lambda_1(1, m) e_q(am) g(m) = q \sum_{m_1 \mid q} \sum_{m_2=1}^\infty \frac{\lambda_1(m_2, m_1)}{m_1 m_2} S(\bar{a}, m_2; q/m_1) G_+ \left( \frac{m_1^2 m_2}{q^3} \right) + q \sum_{m_1 \mid q} \sum_{m_2=1}^\infty \frac{\lambda_1(m_2, m_1)}{m_1 m_2} S(\bar{a}, -m_2; q/m_1) G_- \left( \frac{m_1^2 m_2}{q^3} \right),
\]

where \((a, q) = 1\), and \( \bar{a} \) denotes the multiplicative inverse of \( a \mod q \). Also \( e_q(x) = e^{2\pi i x/q} \).

Remark 1. If \( g \) is supported in \([X, 2X]\), satisfying \( x^j g^{(j)}(x) \ll 1 \), then the sums on the right hand side of (3) are essentially supported on \( m_1^2 m_2 \ll q^3(qX)^{\varepsilon}/X \) (where the implied constant depends on the form \( \pi_1 \) and \( \varepsilon \)). The contribution from the terms with \( m_1^2 m_2 \gg q^3(qX)^{\varepsilon}/X \) is negligibly small (i.e. \( O((qX)^{-N}) \) for any \( N > 0 \)). This follows by estimating the integral \( G_\ell(y) \) by shifting the contour to the right. For smaller values of \( m_1^2 m_2 \) we shift the contour to left upto \( \sigma = -\frac{1}{2} \) (using the result of Jacquet and Shalika) to obtain

\[
G_\pm \left( \frac{m_1^2 m_2}{q^3} \right) \ll \sqrt{X m_1^2 m_2}/q^\varepsilon.
\]

The following lemma is well-known.

Lemma 2. We have

\[
\sum_{n \leq x} |\lambda_1(1, n)|^2 \ll x^{1+\varepsilon},
\]

where the implied constant depends on the form \( \pi_1 \) and \( \varepsilon \).
2.2. Preliminaries on $SL(2, \mathbb{Z})$ Maass forms. Next we shall recall the Voronoi summation formula for $SL(2, \mathbb{Z})$ automorphic forms. For the sake of exposition we only present the case of Maass forms. The case of holomorphic forms is just similar (or even simpler). Furthermore, for technical simplicity, we only restrict to the case of full level. Let $\pi_2$ be a Maass cusp form with Laplace eigenvalue $\frac{1}{4} + \mu^2 \geq 0$, and with Fourier expansion

$$\sqrt{y} \sum_{n \neq 0} \lambda_2(n) K_{i\mu}(2\pi |n|y)e(nx).$$

We will use the following Voronoi type summation formula (see Meurman [19]).

**Lemma 3.** Let $h$ be compactly supported smooth function on $(0, \infty)$. We have

$$\sum_{n=1}^{\infty} \lambda_2(n)e_q(an)h(n) = \frac{1}{q} \sum_{\pm} \lambda_2(\mp n)e_q(\pm \bar{a}n) H^\pm \left( \frac{n}{q^2} \right)$$

where $\bar{a}$ is the multiplicative inverse of $a$ mod $q$, and

$$H^-(y) = -\frac{\pi}{\cosh \pi \mu} \int_0^\infty h(x) \{Y_{2i\mu} + Y_{-2i\mu}\} (4\pi \sqrt{xy}) \, dx$$

$$H^+(y) = 4 \cosh \pi \mu \int_0^\infty h(x) K_{2i\mu} (4\pi \sqrt{xy}) \, dx.$$

**Remark 2.** If $h$ is supported in $[Y, 2Y]$, satisfying $y^j h^{(j)}(y) \ll j$, then the sums on the right hand side of (5) are essentially supported on $n \ll q^2(qY)^{\varepsilon}/Y$ (where the implied constant depends on the form $\pi_2$ and $\varepsilon$). The contribution from the terms with $n \gg q^2(qY)^{\varepsilon}/Y$ is negligibly small. For smaller values of $n$ we will use the trivial bound $H^\pm(n/q^2) \ll Y$.

3. Applying the circle method

3.1. A version of the circle method. We will be using a variant of the circle method, with overlapping intervals, which has been investigated by Jutila ([15], [16]). For any set $S \subset \mathbb{R}$, let $\mathbb{I}_S(x) = 1$ for $x \in S$ and 0 otherwise. For any collection of positive integers $Q \subset [1, Q]$ (which we call the set of moduli), and a positive real number $\delta$ in the range $Q^{-2} \ll \delta \ll Q^{-1}$, we define the function

$$\tilde{I}_{Q, \delta}(x) = \frac{1}{2L} \sum_{q \in Q \mod q} \sum_{\pm} \mathbb{I}_{[\delta^{-1}, \delta + \delta]}(x),$$

where $L = \sum_{q \in Q} \phi(q)$. This is an approximation for $\mathbb{I}_{[0,1]}$ in the following sense:

**Lemma 4.** We have

$$\int_0^1 \left| 1 - \tilde{I}_{Q, \delta}(x) \right|^2 \, dx \ll \frac{Q^{2+\varepsilon}}{\delta L^2}.$$

To prove this, let

$$a_n := \int_0^1 \tilde{I}_{Q, \delta}(x)e(-nx) \, dx = \frac{1}{2L} \sum_{q \in Q} \left[ \sum_{\pm} \mathbb{I}_{[\delta^{-1}, \delta + \delta]}(x) \right] \int_{-\delta}^{\delta} e(-nx) \, dx$$

be the $n$-th Fourier coefficient of $\tilde{I}_{Q, \delta}(x)$. The sum over $a \mod q$ is the Ramanujan sum $c_q(n)$, which can be bounded as $|c_q(n)| \leq \sum_{d \mid (n,q)} d$. Clearly $a_0 = 1$. For $n \neq 0$, by estimating the integral trivially we get that

$$|a_n| \ll \frac{1}{L} \sum_{q \in Q} \sum_{d \mid (n,q)} d \ll \frac{1}{L} \sum_{q \leq Q} \sum_{d \mid (n,q)} d \ll \frac{1}{L} \sum_{d \mid n} d \sum_{q \leq Q \mod d} \frac{1}{d} \ll \frac{Q |n|^{\varepsilon}}{L}.$$

On the other hand we can also bound the integral by $\ll |n|^{-1}$, and obtain

$$|a_n| \ll \frac{Q |n|^{\varepsilon}}{\delta L |n|}.$$
By Parseval we have
\[ \int_0^1 \left| 1 - \tilde{I}_{Q,\delta}(x) \right|^2 \, dx = \int_0^1 \left| \sum_{n \neq 0} a_n e(nx) \right|^2 \, dx = \sum_{n \neq 0} |a_n|^2. \]

Now to conclude the lemma we apply 7 for $|n| \leq \delta^{-1}$, and 8 for $|n| > \delta^{-1}$.

### 3.2. Setting up the circle method.

Let $W$ be a smooth function supported in $[1/2, 3]$ satisfying $W(x) = 1$ for $x \in [1, 2]$, and let $Y = X + h$. (We note that $X \leq Y \ll X^{1+\varepsilon}$.) Then we have
\[ D_h(X) = \sum_{m,n=1}^{\infty} \lambda_1(1,m)\lambda_2(n)V \left( \frac{m}{X} \right) 2\pi \frac{n}{Y} \delta(n, m + h) \]
\[ = \int_0^1 e(xh) \left[ \sum_{m=1}^{\infty} \lambda_1(1,m)e(xm)V \left( \frac{m}{X} \right) \right] \left[ \sum_{n=1}^{\infty} \lambda_2(n)e(-xn)W \left( \frac{n}{Y} \right) \right] \, dx. \]

Let $Q$ be a collection of moduli of size $Q$. Suppose $|Q| \gg Q^{1-\varepsilon}$, so that $L = \sum_{q \in Q} \phi(q) \gg Q^{2-\varepsilon}$. Let
\[ \tilde{D}_h(X, Q) := \int_0^1 \tilde{I}_{Q,\delta}(x)e(xh) \left[ \sum_{m=1}^{\infty} \lambda_1(1,m)e(xm)V \left( \frac{m}{X} \right) \right] \left[ \sum_{n=1}^{\infty} \lambda_2(n)e(-xn)W \left( \frac{n}{Y} \right) \right] \, dx. \]

For convenience we will use the shorthand notation $\tilde{D}_h(X)$ in place of $\tilde{D}_h(X, Q)$. Using the definition of the approximating function $\tilde{I}_{Q,\delta}(x)$, we get
\[ \tilde{D}_h(X) = \frac{1}{L} \sum_{q \in Q} \sum_{a \mod q} \epsilon_q(ah) \sum_{m,n \in \mathbb{Z}} \lambda_1(1,m)\lambda_2(n)\epsilon_q(a(m - n))F(m, n) \]
where
\[ F(x, y) = V \left( \frac{x}{X} \right) W \left( \frac{y}{Y} \right) \frac{1}{2\delta} \int_{-\delta}^{\delta} e(\alpha(x - y)) \, d\alpha. \]

We choose $\delta = Y^{-1}$ so that
\[ \left| \frac{\partial^{i+j}}{\partial x^i \partial y^j} F(x, y) \right| \ll_{i,j} \frac{1}{X^i Y^j}. \]

In circle method we approximate $D_h(X)$ by $\tilde{D}_h(X)$, and then try to estimate the latter sum. Lemma 4 gives a way to estimate the error of replacing $D_h(X)$ by $\tilde{D}_h(X)$. More precisely we have
\[ \left| D_h(X) - \tilde{D}_h(X) \right| \ll \int_0^1 \left[ \sum_{m=1}^{\infty} \lambda_1(1,m)e(xm)V \left( \frac{m}{X} \right) \right] \left[ \sum_{n=1}^{\infty} \lambda_2(n)e(-xn)W \left( \frac{n}{Y} \right) \right] \left| 1 - \tilde{I}(x) \right| \, dx, \]

Using the well-known point-wise uniform bound
\[ \sum_{n=1}^{\infty} \lambda_2(n)e(-xn)W \left( \frac{n}{Y} \right) \ll_{\pi, \varepsilon} Y^{1+\varepsilon} \]
it follows that the right hand side of 11 is bounded by
\[ \ll Y^{1+\varepsilon} \int_0^1 \left[ \sum_{m=1}^{\infty} \lambda_1(1,m)e(xm)V \left( \frac{m}{X} \right) \right] \left| 1 - \tilde{I}(x) \right| \, dx. \]

Now we apply Cauchy and Lemma 4 to conclude
\[ D_h(X) = \tilde{D}_h(X) + O \left( \frac{\sqrt{XYQ(XYQ)^{1+\varepsilon}}}{\sqrt{\delta L}} \right). \]

As the moduli set $Q$ is such that $L \gg Q^{2-\varepsilon}$, and $\delta = Y^{-1}$, it follows that the above error term is smaller than the trivial bound 1, if we choose $Q = Y^{\frac{1}{2}+\delta}$ for any $\delta > 0$. Indeed with this choice we have the following:

**Lemma 5.** We have
\[ D_h(X) = \tilde{D}_h(X) + O \left( X^{1-\delta+\varepsilon} \right). \]
4. Estimation of $\tilde{D}_h(X)$

4.1. Applying the Voronoi summation formulas. Fix $\alpha \in [-\delta, \delta]$ and set

$$\tilde{D}_{h,\alpha}(X) = \frac{1}{L} \sum_{q \in \mathcal{Q}} \sum_{a \equiv m \mod q}^* e_q(ah) \sum_{m, n \in \mathbb{Z}} \lambda_1(1,m)\lambda_2(n) e_q(a(m-n)) V \left( \frac{m}{X} \right) W \left( \frac{n}{Y} \right) e(\alpha(m-n)),$$

so that $\tilde{D}_h(X) = \frac{1}{2\delta} \int_{-\delta}^\delta \tilde{D}_{h,\alpha}(X) d\alpha$. Now we apply Voronoi summations on the sums over $m$ and $n$. This process gives rise to several terms as noted in Section $2$ - Lemma $3$ and Lemma $5$. As far as our analysis is concerned all the terms are of equal complexity, and so we just focus our attention on one such term -

$$\tilde{D}_{h,1}(X) = \frac{1}{L} \sum_{q \in \mathcal{Q}} \sum_{m_1|q, m_2=1}^\infty \frac{\lambda_1(m_2,m_1)}{m_1m_2} \times \sum_{n=1}^\infty \lambda_2(-n)S(m_1,m_2,n,h;q)G_+ \left( \frac{m_1^2m_2}{q^3} \right) H^+ \left( \frac{n}{q^2} \right),$$

where the character sum is given by

$$S(m_1,m_2,n,h;q) := \sum_{a \equiv m \mod q}^* e_q(ah)e_q(-\bar{a}n)S(\bar{a},m_2;q/m_1).$$

Also here we are taking

$$g(x) = V \left( \frac{x}{X} \right) e(\alpha x), \quad \text{and} \quad h(y) = W \left( \frac{y}{Y} \right) e(-\alpha y).$$

The functions $G_+$ and $H^+$ are defined in Lemma $3$ and Lemma $5$ respectively.

**Remark 3.** Suppose we establish square-root cancellation in the character sum $S(m_1,m_2,n,h;q)$ in (13) Then estimating the remaining sums trivially using the decay in the functions $G_+$ and $H^+$, we get that

$$\tilde{D}_{h,1}(X) \ll Y^{1+2\delta}(XY)^\varepsilon \ll X^{1+2\delta+\varepsilon}.$$
of \( f(x, y) = q_1hx - q_1nx^{-1} + x^{-1}y + m_2y^{-1} \in \mathbb{F}_{q_2}[x, y, (xy)^{-1}] \). In the generic case where \( q_2 \not| nm_2 \), the polyhedron \( \Delta(f) \) is given by the 4-gon in \( \mathbb{R}^2 \) with vertices \((1, 0), (-1, 0), (-1, 1) \) and \((0, -1) \). Hence \( \dim \Delta(f) = 2 \). Also it is easily seen that \( f \) is nondegenerate with respect to \( \Delta(f) \). Adolphson and Sperber show that this condition is sufficient to conclude that

\[
\sum_{a, b \in \mathbb{F}_{q_2}^*} e_{q_2}(q_1ah - q_1\bar{a}n + b\bar{a} + m_2\bar{b}) \ll q_2
\]

in the light of the general results of Deligne.

**Lemma 6.** For \( q = q_1q_2 \) a product of two primes, \( m_1| q \), and \( (n, q_1q_2) = 1 \), we have

\[
S(m_1, m_2, n, h; q) \ll \frac{q}{\sqrt{m_1}} \sqrt{\left( \frac{q}{m_1}, m_2 \right)}.
\]

We will now use the above lemma to estimate the following sub-sum of \( D_{h, \mathbf{a}, 1} \),

\[
D_{h, \mathbf{a}, 1}^1(X) = \frac{1}{L} \sum_{q_1 \in \mathbb{Q}_1} \sum_{q_2 \in \mathbb{Q}_2} \sum_{m_2 = 1}^\infty \frac{\lambda_1(m_2, q_1)}{q_1m_2} \lambda_2(-n)S(q_1, m_2, n, h; q_1q_2)G_+(\frac{m_2}{q_1q_2})H^+\left(\frac{n}{q_1q_2}\right).
\]

Using Lemma 6 and the properties of the weight functions (see Section 2), it follows that up to a negligible error term (i.e. \( O(X^{-N}) \) for any \( N > 0 \)) we have

\[
D_{h, \mathbf{a}, 1}^1(X) \ll \frac{X^2}{LQ_1^2} \sum_{q_1 \in \mathbb{Q}_1} \sum_{q_2 \in \mathbb{Q}_2} \sum_{m_2 = 1}^\infty \frac{|\lambda_1(m_2, q_1)|}{\sqrt{m_2}} \sqrt{\frac{q_2}{m_2}} \sum_{|n| \leq X^{2\delta + \varepsilon}} |\lambda_2(-n)|.
\]

(We will take \( Q_1, Q_2 \gg X^{2\delta + \varepsilon} \), see the remark below, so that the coprimality condition \( (q_1, n) = 1 \) of the lemma is satisfied.) Using Cauchy inequality and the Rankin-Selberg theory the sum over \( n \) is bounded by \( X^{2\delta + \varepsilon} \). To do the sum over \( m_2 \), we apply Cauchy to get

\[
\sum_{m_2 = 1}^\infty \frac{|\lambda_1(m_2, q_1)|}{\sqrt{m_2}} \sqrt{\frac{q_2}{m_2}} \leq \left[ \sum_{m_2 = 1}^\infty \frac{|\lambda_1(m_2, q_1)|}{\sqrt{m_2}} \right]^{\frac{1}{2}} \left[ \sum_{m_2 = 1}^\infty \frac{(q_2/m_2)}{m_2} \right]^{\frac{1}{2}}.
\]

The last sum is clearly bounded by \( X^\varepsilon \). To bound the middle sum we plug in the inequality

\[
|\lambda_1(m_2, q_1)| \leq 2|\lambda_1(m_2, 1)|^2|\lambda_1(q_1, 1)|^2 + 2|\lambda_1(m_2, 1)|^2,
\]

which is a consequence of the Hecke relation, and apply Lemma 2. It follows that up to a negligible error term we have

\[
D_{h, \mathbf{a}, 1}^1(X) \ll \frac{X^{2\delta + \varepsilon}}{LQ_1^2} \sum_{q_1 \in \mathbb{Q}_1} \sum_{q_2 \in \mathbb{Q}_2} (|\lambda_1(q_1, 1)| + 1) \ll \frac{X^{2\delta + \varepsilon}}{Q_1^2}.
\]

The last inequality follows from another application of Cauchy inequality and Lemma 2. We summarize the outcome of the above analysis in the following:

**Lemma 7.** We have

\[
D_{h, \mathbf{a}, 1}^{m_1 \not| m_2}(X) := \frac{1}{L} \sum_{q \in \mathbb{Q}} \sum_{m_1| q} \sum_{m_2 = 1}^\infty \frac{\lambda_1(m_2, m_1)}{m_1m_2} \lambda_2(-n)S(m_1, m_2, n, h; q)G_+(\frac{m_2q_1}{q^3})H^+\left(\frac{n}{q^2}\right) \ll X^{2\delta + \varepsilon} \min\{Q_1, Q_2\}^{\varepsilon}.
\]

**Remark 4.** Suppose \( \min\{Q_1, Q_2\} \gg X^{2\delta + \varepsilon} \), then the bound given in the above lemma is at least as good as the bound we have for the error term in (12).
Applying the bound from Lemma 5, we can also get a satisfactory bound for the contribution from the small values of \(m_2\) even when \(m_1 = 1\).

**Lemma 8.** We have

\[
\hat{D}_{h,\alpha,1}^{m_1=1}(X, M) := \frac{1}{L} \sum_{\nu \in Q} \sum_{m_2 \leq M} \frac{\lambda_1(m_2, 1)}{m_2} \times \sum_{n=1}^{\infty} \lambda_2(-n)S(1, m_2, n, h; q)G_+ \left( \frac{m_2}{q^3} \right) H^+ \left( \frac{n}{q^2} \right) \ll X^{\frac{\sqrt{M}X^{\varepsilon}}{X^{\frac{1}{4}+\varepsilon}}}.
\]

**Remark 5.** This is at least as good as the bound we have for the error term in (12) if \(M \leq X^{\frac{1}{2}-3\delta}\).

5. **Estimation of \(\hat{D}_h(X)\) : Final analysis**

5.1. **Applying Cauchy and Poisson.** It remains to analyse the size of the sum

\[(17) \quad \hat{D}_{h,\alpha,1}^2(X, M) = \frac{1}{L} \sum_{\nu \in Q} \sum_{m \sim M} \frac{\lambda_1(m, 1)}{m} \sum_{n=1}^{\infty} \lambda_2(-n)S(1, m, n, h; q)G_+ \left( \frac{m}{q^3} \right) H^+ \left( \frac{n}{q^2} \right),\]

where \(S(1, m, n, h; q)\) is defined in (14). Here \(m \sim M\) means that \(m\) runs over the integers in the dyadic segment \([M, 2M]\). Also by Lemma 5, it is enough to take \(X^{\frac{1}{2}-3\delta} < M < X^{\frac{1}{2}+3\delta+\varepsilon}\). We have already noted that the square-root cancellation in the character sum is not enough for our purpose, and we need to prove cancellation in the average. To this end we will exploit heavily the factorization of the moduli set \(Q\). We have

\[
\hat{D}_{h,\alpha,1}^2(X, M) \ll \frac{1}{LM} \sum_{\nu_1 \in Q_2} \sum_{n \in X^{1/4+\varepsilon}} |\lambda_2(-n)| \times \sum_{m \sim M} \left| \sum_{q_1 \in Q_1} S(1, m, n, h; q)G_+ \left( \frac{m}{q^3} \right) H^+ \left( \frac{n}{q^2} \right) \right|,
\]

where \(q = q_1q_2\). To get rid of the Fourier coefficients, we apply Cauchy twice, Lemma 2 and its well-known \(GL(2)\) version. With this we arrive at the following

\[(18) \quad \hat{D}_{h,\alpha,1}^2(X, M) \ll \frac{\sqrt{Q_2}X^{1/4+\varepsilon}}{L\sqrt{M}} \left\{ \sum_{\nu_1 \in Q_2} \sum_{n \in X^{1/4+\varepsilon}} \hat{D}_{h,\alpha,1}^2(X, M; n, q_2) \right\}^{\frac{1}{2}},\]

where

\[
\hat{D}_{h,\alpha,1}^2(X, M; n, q_2) = \sum_{m \in \mathbb{Z}} \left( \frac{m}{M} \right) \sum_{\nu_1 \in Q_1} S(1, m, n, h; q_1q_2)G_+ \left( \frac{m}{q_1^3q_2^2} \right) H^+ \left( \frac{n}{q_1^2q_2} \right).
\]

Here \(F\) is non-negative smooth function on \((0, \infty)\), supported on \([1/2, 3]\), and such that \(F(x) = 1\) for \(x \in [1, 2]\).

**Remark 6.** It is quite natural to split the set of moduli at this point. Indeed if \(Q_1 = 1\) then we do not have enough points of summation inside the absolute value square to get more cancellation beyond the square root cancellation in the character sum \(S(1, m, n, h; q_1q_2)\). This term shows up as the diagonal contribution. On the other hand if \(Q_2 = 1\), or \(Q_1 = Q\), then when we apply Poisson summation on the sum over \(m\) after opening the absolute square, the size of the modulus is \(Q^2 = X^{1/2+\delta}\), which is too large compared to the range of summation of \(m\) and so the saving from Poisson is not enough. Hence the off-diagonal term is not satisfactory.
Opening the absolute square and interchanging the order of summations we get
\[
\tilde{D}_{\alpha,1}^2(X, M; n, q_2) = \sum_{q_1 \in \mathbb{Q}_1} \sum_{\alpha \mod q_1} H^+ \left( \frac{n}{q_1 q_2} \right) \tilde{H}^+ \left( \frac{n}{q_1 q_2} \right)
\times \sum_{m \in \mathbb{Z}} F \left( \frac{m}{M} \right) S(1, m, n, h; q_1 q_2) \tilde{S}(1, m, n, h; q_1 q_2) G_+ \left( \frac{m}{q_1 q_2} \right) \tilde{G}_+ \left( \frac{m}{q_1 q_2} \right).
\]

Applying Poisson summation on the sum over \(m\) with modulus \(q_1 \tilde{q}_1 q_2\), we get
\[
\frac{M}{q_2} \sum_{q_1 \in \mathbb{Q}_1} \sum_{\tilde{q}_1 \in \mathbb{Q}_1} \frac{1}{q_1 \tilde{q}_1} H^+ \left( \frac{n}{q_1 q_2} \right) \tilde{H}^+ \left( \frac{n}{q_1 q_2} \right) \sum_{m \in \mathbb{Z}} T(n, m, h; q_1, \tilde{q}_1, q_2) \mathcal{I}(m; q_1, \tilde{q}_1, q_2),
\]
where the character sum is given by
\[
T(n, m, h; q_1, \tilde{q}_1, q_2) = \sum_{\alpha \mod q_1 \tilde{q}_1 q_2} S(1, \alpha, n, h; q_1 q_2) \tilde{S}(1, \alpha, n, h; q_1 q_2) e_{q_1 \tilde{q}_1 q_2}(m\alpha),
\]
and the integral is given by
\[
\mathcal{I}(m; q_1, \tilde{q}_1, q_2) = \int_{\mathbb{R}} F(x) G_+ \left( \frac{x M}{q_1 q_2} \right) \tilde{G}_+ \left( \frac{x M}{q_1 q_2} \right) e_{q_1 \tilde{q}_1 q_2}(-mMx) dx.
\]

Integrating by parts repeatedly we get that the integral is negligibly small unless \(|m| \ll Q_1 Q X^\epsilon / M\).
Observe that differentiating under the integral sign in (2), one can show that \(y^j G_{\epsilon}^j(y) \ll j \sqrt{yX}\). So for \(|m| \ll Q_1 Q X^\epsilon / M\) we have the bound
\[
\mathcal{I}(m; q_1, \tilde{q}_1, q_2) \ll \frac{MX}{Q^3}.
\]

The following lemma now follows from (19).

Lemma 9. For any \(N > 0\), we have the bound
\[
\tilde{D}_{\alpha,1}^2(X, M; n, q_2) \ll \frac{M^2 X^3}{Q_1 Q^4} \sum_{q_1 \in \mathbb{Q}_1} \sum_{\tilde{q}_1 \in \mathbb{Q}_1} \sum_{|m| \ll M Q X^\epsilon} |T(n, m, h; q_1, \tilde{q}_1, q_2)| + X^{-N}.
\]

5.2. Estimating the character sums. First consider the case where \(q_1 \neq \tilde{q}_1\). Then the character sum \(T(n, m, h; q_1, \tilde{q}_1, q_2)\) splits into a product of three character sums with moduli \(q_1, q_1, q_2\) respectively. The sum modulo \(q_1\) is given by
\[
T_1 = \sum_{\alpha \mod q_1} \sum_{\beta \mod q_1} e_{q_1}(\tilde{q}_2 h \beta - \tilde{q}_2 n \beta) S(q_2 \tilde{\beta}, q_2 \alpha; q_1) e_{q_1}(\tilde{q}_1 \tilde{q}_2 m \alpha).
\]

Opening the Kloosterman sum and executing the sum over \(\alpha\) we arrive at an explicit expression of this character sum in terms of Kloosterman sums, namely
\[
T_1 = q_1 S(\tilde{q}_2 h, -\tilde{q}_2 n(\tilde{q}_1 \tilde{m})/q_1)
\]
if \((m, q_1) = 1\), and \(T_1 = 0\) otherwise. The sum modulo \(q_1\), which we denote by \(\tilde{T}_1\), can be computed in exactly the same fashion. Next we consider the sum modulo \(q_2\), which is given by
\[
T_2 = \sum_{\alpha \mod q_2} \sum_{\beta \mod q_2} \sum_{\gamma \mod q_2} e_{q_2}(\tilde{q}_1 h \beta - \tilde{q}_1 n \beta - \tilde{q}_1 h \gamma + \tilde{q}_1 n \gamma) S(\tilde{q}_1 \tilde{\beta}, \tilde{q}_1 \alpha; q_2) S(\tilde{q}_1 \tilde{\gamma}, \tilde{q}_1 \alpha; q_2) e_{q_2}(\tilde{q}_1 \tilde{q}_1 m \alpha).
\]

Opening the Kloosterman sums we execute the sum over \(\alpha\) to get
\[
T_2 = q_2 \sum_{\delta \mod q_2} \sum_{\beta \mod q_2} \sum_{\gamma \mod q_2} e_{q_2}(\tilde{q}_1 h \beta - \tilde{q}_1 n \beta - \tilde{q}_1 h \gamma + \tilde{q}_1 n \gamma + \tilde{q}_1 \beta \delta - \tilde{q}_1 q_1 \gamma)(\tilde{q}_1 \delta + m),
\]
where the double asterisk on the sum over \(\delta\) indicates that \(\delta\) and \(\tilde{q}_1 \delta + m\) are invertible modulo \(q_2\). To get square-root cancellation in the remaining character sum we shall appeal to the work of Bombieri and Sperger (4) (in particular see Section IV. of (4)).

Using the notation of (4), we set
\[
f(x, y, z) = a(x) + y + \frac{b(x)}{y} + z + \frac{c(x)}{z}
\]
where
\[ a(x) = 0, \quad b(x) = q_1^2 h \left( \frac{1}{x} - n \right), \quad \text{and} \quad c(x) = \tilde{q}_1^2 h \left( \frac{q_1}{q_1 x + m} - n \right). \]

Let \( V \) be the quasi-projective variety in \( \mathbb{A}_3(F_{q_2}) \) defined by \( x \neq 0, -\tilde{q}_1 m, y \neq 0, \) and \( z \neq 0. \) Then
\[ T_2 = q_2 \sum_{(x,y,z) \in V(q_{q_2})} e_{q_2}(f(x, y, z)). \]

From the main result of [4], it follows that if \( q_2 \nmid q_1 - \tilde{q}_1 \) or \( q_2 \nmid m, \) then \( T_2 \ll q_2^{\frac{5}{2}}. \) Otherwise using the Weil bound for Kloosterman sums we have \( T_2 \ll q_2^3. \) Putting the following bound.

**Lemma 10.** For \( q_1 \neq \tilde{q}_1, \) the character sum \( T(n, m, h; q_1, \tilde{q}_1, q_2) \) vanishes unless \( (m, q_1 \tilde{q}_1) = 1, \) in which case we have
\[ T(n, m, h; q_1, \tilde{q}_1, q_2) \ll q_1^2 q_2^3 (m, q_2)^{\frac{1}{2}}. \]

If \( q_1 = \tilde{q}_1, \) then the character sum \( T(n, m, h; q_1, \tilde{q}_1, q_2) \) splits as a product of two character sums. The one with modulus \( q_2, \) has already been analysed above and it satisfies the bound \( \ll q_2^{\frac{5}{2}} (m, q_2)^{\frac{1}{2}}. \) The other sum with modulus \( q_1 \) is given by
\[ T_{q_1=\tilde{q}_1} = \sum_{\alpha \mod q_1} \sum_{\delta \mod \tilde{q}_1} \sum_{\beta, \gamma \mod q_1} \epsilon_1 (q_1 \tilde{h} \beta - q_2 n \beta - q_2 h \gamma + q_2 n \gamma) S(q_1 \tilde{h} \beta, q_2 \alpha; q_1) S(q_2 \tilde{h} \gamma, q_2 \alpha; q_1) e_{q_1} (q_2 m \alpha). \]

As before we open the Kloosterman sums and execute the sum over \( \alpha. \) It follows that the sum vanishes unless \( q_1 | m. \) So let \( m = q_1 m'. \) Then we arrive at
\[ T_{q_1=\tilde{q}_1} = q_1 \sum_{\delta \mod \tilde{q}_1} \sum_{\beta, \gamma \mod q_1} \epsilon_1 (q_1 \tilde{h} \beta - q_2 n \beta - q_2 h \gamma + q_2 n \gamma + q_2 \tilde{h} \beta - q_2 \tilde{h} \gamma (\delta + m')). \]

The remaining sum is just like \( T_2, \) with different parameters.

**Lemma 11.** The character sum \( T(n, m, h; q_1, q_2, q_2) \) vanishes unless \( q_1 | m, \) in which case we have
\[ T(n, q_1 m', h; q_1, q_2) \ll q_1^2 q_2^{\frac{5}{2}} \sqrt{(m', q_1 q_2)}. \]

### 5.3 Final estimation and conclusion of the theorem.

It follows from Lemma 10 that
\[ \frac{M^2 X^3}{Q_1 Q_1} \sum_{q_1 \in Q_1, \tilde{q}_1 \in Q_1, |m| \ll \frac{q_1}{Q_1} X} |T(n, m, h; q_1, \tilde{q}_1, q_2)| \ll \frac{M^2 X^3 Q_1^3 Q_2^2}{Q_1 Q_1} \sum_{1 \leq |m| \ll \frac{q_1}{Q_1} X} \sqrt{(m, q_2)} \ll \frac{M X^3 Q_1^2 Q_2^2}{Q^1} X^\varepsilon. \]

Applying Lemma 11 it follows that
\[ \frac{M^2 X^3}{Q_1 Q_1} \sum_{q_1 \in Q_1, \tilde{q}_1 \in Q_1, |m| \ll \frac{q_1}{Q_1} X} |T(n, m, h; q_1, \tilde{q}_1, q_2)| = \frac{M^2 X^3}{Q_1 Q_1} \sum_{q_1 \in Q_1} \sum_{1 \leq |m| \ll \frac{q_1}{Q_1} X} |T(n, q_1 m, h; q_1, q_2)| + \frac{M^2 X^3}{Q_1 Q_1} \sum_{q_1 \in Q_1} |T(n, 0, h; q_1, q_2)| \ll \frac{M X^3}{\sqrt{Q}} X^\varepsilon + \frac{M^2 X^3}{Q} X^\varepsilon \ll \frac{M X^3}{Q} X^\varepsilon. \]

The above two bounds (20), (21), together with Lemma 9 imply that
\[ D_{n, \alpha, 1}^2 (X, M; n, q_2) \ll \frac{M X^3 Q_1^2}{\sqrt{Q_2}} X^\varepsilon + \frac{M^2 X^3}{Q_1 Q_2} X^\varepsilon. \]

Plugging this estimate in (18) we get the following:
Lemma 12. For $Q_1Q_2 = Q = X^{\frac{1}{2} + \delta}$ and $M \ll X^{\frac{1}{2} + 3\delta + \varepsilon}$, we have
\[
\tilde{D}^2_{\eta, \alpha, 1}(X, M) \ll \left( \frac{X^{\delta}}{Q_2} + \frac{X^{2\delta}}{Q_1} \right) X^{1+\varepsilon}.
\]

For given $\delta$ the optimal break up of $Q$ is given by $Q_1 = X^{\frac{1}{2} + \delta}$ and $Q_2 = X^\delta$. To obtain the optimal value for $\delta$, recall that we are assuming $\min\{Q_1, Q_2\} \gg X^{2\delta + \varepsilon}$. So we have the restriction $\delta < \frac{1}{10}$. Moreover comparing the above bound with the bound for the error term in (12), we get that the optimal choice is given by
\[
\delta = \frac{1 - 5}{20},
\]
i.e. $\delta = \frac{1}{10}$. Our main theorem now follows from Lemma 5, Lemma 7, Lemma 8 and Lemma 12.

REFERENCES

[1] A. Adolphson; S. Sperber: Exponential sums on $(G_m)^n$, Invent. Math. 101 (1990), 63–79.
[2] V. Blomer: Shifted convolution sums and subconvexity bounds for automorphic $L$-functions, Int. Math. Res. Not. 73 (2004), 3905–3926.
[3] V. Blomer; G. Harcos: The spectral decomposition of shifted convolution sums, Duke Math. J. 144 (2008), 321–339.
[4] E. Bombieri; S. Sperger: On the estimation of certain exponential sums, Acta Arith. 69 (1995), 329–358.
[5] J. Denef; F. Loeser: Weights of exponential sums, intersection cohomology, and Newton polyhedra, Invent. Math. 106 (1991), 275–294.
[6] W. Duke; J.B. Friedlander; H. Iwaniec: Bounds for automorphic $L$-functions, Invent. Math. 112 (1993), 1–8.
[7] W. Duke; J.B. Friedlander; H. Iwaniec: A quadratic divisor problem, Invent. Math. 115 (1994), 209–217.
[8] W. Duke; J.B. Friedlander; H. Iwaniec: Bilinear forms with Kloosterman fractions, Invent. Math. 128 (1997), 23–43.
[9] D. Goldfeld: Automorphic forms and $L$-functions for the group $GL(n, \mathbb{R})$, with an appendix by Kevin A. Broughan. Cambridge Studies in Advanced Mathematics 99. Cambridge University Press, Cambridge, 2006.
[10] D. Goldfeld; X. Li: Voronoi formulas on $GL(n)$, Int. Math. Res. Not. (2006), Art. ID 86295, 25 pp.
[11] G. Harcos: An additive problem in the Fourier coefficients of cusp forms, Duke Math. J. 146 (2009), 401–448.
[12] R. Holowinsky; R. Munshi: Level aspect subconvexity for Rankin-Selberg $L$-functions. (Preprint)
[13] M. Jutila: Transformations of exponential sums, Proceedings of the Amalfi Conference on Analytic Number Theory (Maiori 1989), Univ. Salerno, Salerno, (1992) 263–270.
[14] M. Jutila: The additive divisor problem and its analogs for Fourier coefficients of cusp forms. I., Math. Z. 223 (1996), 435–461; II., ibid 225 (1997), 625–637.
[15] E. Kowalski; P. Michel; J. VanderKam: RankinSelberg $L$-functions in the level aspect, Duke Math. J. 114 (2002), 123–191.
[16] W. Luo; P. Sarnak: Mass equidistribution for Hecke eigenforms, Comm. Pure Appl. Math. 56 (2003), 874–891.
[17] T. Meurman: On exponential sums involving the Fourier coefficients of Maass wave forms, J. Reine Angew. Math. 384 (1988), 192–207.
[18] P. Michel: The subconvexity problem for Rankin-Selberg $L$-functions and equidistribution of Heegner points. II., Invent. math. 163 (2006), 581–655.
[19] R. Holowinsky: A sieve method for shifted convolution sums, Duke Math. J. 146 (2009), 401–448.
[20] R. Holowinsky; R. Munshi: Level aspect subconvexity for Rankin-Selberg $L$-functions. (Preprint)
[21] M. Jutila: Transforms of exponential sums, Proceedings of the Amalfi Conference on Analytic Number Theory (Maiori 1989), Univ. Salerno, Salerno, (1992) 263–270.
[22] M. Jutila: The additive divisor problem and its analogs for Fourier coefficients of cusp forms. I., Math. Z. 223 (1996), 435–461; II., ibid 225 (1997), 625–637.
[23] E. Kowalski; P. Michel; J. VanderKam: RankinSelberg $L$-functions in the level aspect, Duke Math. J. 114 (2002), 123–191.
[24] W. Luo; P. Sarnak: Mass equidistribution for Hecke eigenforms, Comm. Pure Appl. Math. 56 (2003), 874–891.
[25] T. Meurman: On exponential sums involving the Fourier coefficients of Maass wave forms, J. Reine Angew. Math. 384 (1988), 192–207.
[26] P. Michel: The subconvexity problem for Rankin-Selberg $L$-functions and equidistribution of Heegner points, Annals of Math. 160 (2004), 185–236.
[27] S. Miller; W. Schmid: Automorphic distributions, $L$-functions, and Voronoi summation for $GL(3)$, Annals of Math. 164 (2006), 423–488.
[28] R. Munshi: The circle method and bounds for $L$-function. (Preprint)
[29] N. J. E. Pitt: On shifted convolution sums of $c^3(s)$ with automorphic $L$-functions, Duke Math. J. 77 (1995), 383–406.
[30] P. Sarnak: Estimates for Rankin-Selberg $L$-functions and Quantum Unique Ergodicity, Journal of Functional Analysis 184 (2001), 419–453.
[31] A. Selberg: On the estimation of Fourier coefficients of modular forms, Proc. Sympos. Pure Math. 8 (1965), 1–15.

School of Mathematics, Tata Institute of Fundamental Research, 1 Dr. Homi Bhabha Road, Colaba, Mumbai 400005, India.

E-mail address: rmunshi@math.tifr.res.in