The circle method and bounds for L-functions - IV: Subconvexity for twists of $\mathrm{GL}(3)$ L-functions

By Ritabrata Munshi

To Soumyanetra and Sroutatwisha

Abstract

Let π be an $\mathrm{SL}(3, \mathbb{Z})$ Hecke-Maass cusp form satisfying the Ramanujan conjecture and the Selberg-Ramanujan conjecture, and let χ be a primitive Dirichlet character modulo M, which we assume to be prime for simplicity. We will prove that there is a computable absolute constant $\delta > 0$ such that

$$L\left(\tfrac{1}{2}, \pi \otimes \chi\right) \ll_\pi M^{\frac{3}{4}-\delta}.$$

1. Introduction

Let π be a Hecke-Maass cusp form for $\mathrm{SL}(3, \mathbb{Z})$ of type (ν_1, ν_2) (see [2] and [5]). Let $\lambda(m, n)$ be the normalized (i.e., $\lambda(1, 1) = 1$) Fourier coefficients of π. The Langlands parameters $(\alpha_1, \alpha_2, \alpha_3)$ for π are given by $\alpha_1 = -\nu_1 - 2\nu_2 + 1$, $\alpha_2 = -\nu_1 + \nu_2$ and $\alpha_3 = 2\nu_1 + \nu_2 - 1$. Let χ be a primitive Dirichlet character modulo M. The L-function associated with the twisted form $\pi \otimes \chi$ is given by the Dirichlet series

$$(1) \qquad L(s, \pi \otimes \chi) = \sum_{n=1}^{\infty} \lambda(1, n)\chi(n)n^{-s}$$

in the domain $\sigma = \mathrm{Re}(s) > 1$. The L-function extends to an entire function and satisfies a functional equation with arithmetic conductor M^3. Hence the convexity bound is given by

$$L\left(\tfrac{1}{2}, \pi \otimes \chi\right) \ll_{\pi,\varepsilon} M^{3/4+\varepsilon}.$$

The subconvexity problem for this L-function has been solved in several special cases in [1], [13], [12], [14] and, more recently in [17]. In [1] Blomer established a subconvex bound with exponent $5/8+\varepsilon$ in the case where π is self dual and χ is a quadratic character. In [13], [14] we considered twists of symmetric square

---

lifts by characters χ with prime power modulus $M = q^\ell$ with $\ell > 1$. In [13] it was shown that for every $\ell > 1$, there is a $\delta_\ell > 0$ such that

$$L\left(\tfrac{1}{2}, \pi \otimes \chi\right) \ll_{\pi,\ell,\varepsilon} M^{3/4-\delta_\ell+\varepsilon},$$

and in [14] we proved that

$$L\left(\tfrac{1}{2}, \pi \otimes \chi\right) \ll_{\pi,q,\varepsilon} M^{3/4-1/12+\varepsilon}.$$

The later result, where the base prime q remains fixed and the power ℓ grows, is a non-archimedean analogue of the t-aspect subconvexity (see [9] and [15]). The main result of [17] generalizes that of the unpublished note [12]. In [17] we consider twists of π (which need not be self dual) by characters χ that factorize as $\chi = \chi_1 \chi_2$ with χ_i primitive modulo M_i and $(M_1, M_2) = 1$. Suppose there is a $\delta > 0$ such that $M_2^{1/2+4\delta} < M_1 < M_2^{1-3\delta}$. Then we show that

$$L\left(\tfrac{1}{2}, \pi \otimes \chi\right) \ll_{\pi,q,\varepsilon} M^{3/4-\delta+\varepsilon}.$$

In [12] a similar result was proved for twists of the symmetric square lifts.

   In this paper we will prove a very general statement. However we are going to assume that the form π satisfies the following conditions:

   (R) The Ramanujan conjecture $\lambda(m,n) \ll (mn)^\varepsilon$;
   (RS) The Ramanujan-Selberg conjecture $\mathrm{Re}(\alpha_i) = 0$.

   THEOREM 1. *Let π be a Hecke-Maass cusp form for $\mathrm{SL}(3,\mathbb{Z})$ satisfying conjectures* (R) *and* (RS). *Let χ be a primitive Dirichlet character modulo M. Suppose M is a prime number. Then there is a computable absolute constant $\delta > 0$ such that*

$$L\left(\tfrac{1}{2}, \pi \otimes \chi\right) \ll_\pi M^{\frac{3}{4}-\delta}.$$

   The primality assumption on M is more a technical convenience than an essential requirement. A more general statement without this assumption can be proved using the technique introduced in this paper. Also the exponent can be explicitly computed. In fact one can take $\delta = 1/1612$. Our primary goal here is to present the ideas as clearly as possible without trying to prove the most general statement or the best possible exponent. The conditions (R) and (RS) are quite serious, and their removal is a technical challenge. Indeed unlike the previous papers in this series [16], [17], we do not need to use Deligne type bounds for exponential sums. Instead of estimating exponential sums, we will be required to solve a counting problem, which we tackle in an elementary manner (without recourse to exponential sums).

   The subconvexity problem for L-functions twisted by a Dirichlet character has been studied extensively in the literature. The first instance of such a result is of course the pioneering work of Burgess [3], whose well-known bound

$$L\left(\tfrac{1}{2}, \chi\right) \ll M^{\frac{3}{16}+\varepsilon}$$

still remains unsurpassed. In the case of degree two L-functions the problem was first tackled by Duke, Friedlander and Iwaniec [4] using the amplification technique. Their result has been extended (e.g., to the case of general GL(2) automorphic forms) and improved by several authors in the last two decades. Our theorem provides a GL(3) analogue of the main result of [4].

The present work substantially differs from the previous papers in the series, and one may rightly argue that the way we detect the (diagonal) equation $n = r$ here can hardly be termed a circle method. We use the Petersson trace formula, which gives an expansion of the Kronecker delta symbol in terms of the Fourier coefficients of holomorphic forms and the Kloosterman sums. However the basic set up for the proof of Theorem 1 coincides with that in [16], [17] and [15]. In particular, we use an expansion of the Kronecker delta symbol to separate the oscillation of the GL(3) Fourier coefficients from that of the character. The idea of using the Petersson formula as a substitute of the circle method is also exploited in [18] where we deal with the Rankin-Selberg L-functions.

Remark 1. The approach in this paper gives an unconditional subconvexity result for twists of the symmetric square lifts of $\mathrm{SL}(2, \mathbb{Z})$ holomorphic forms. Indeed in this case (RS) is known and (R) follows from the work of Deligne.

Remark 2. The theorem in fact holds under the weaker assumptions that $\lambda(m, n) \ll (mn)^{\theta_o + \varepsilon}$ and $|\mathrm{Re}(\alpha_i)| \le \eta_o$, with θ_o and η_o sufficiently small. Since we need these parameters to be very small, far from what one can hope to achieve in near future, we refrain from writing it down explicitly. A case of special interest corresponds to symmetric square lifts of $\mathrm{SL}(2, \mathbb{Z})$ Maass forms. In this case, though not sufficient for our purpose, strong bounds are known from the work of Kim and Sarnak ($\theta_o = 7/32$).

Acknowledgements. The author wishes to thank Valentin Blomer, Philippe Michel, Peter Sarnak and Matthew Young for their interest in this work. The author also thanks the anonymous referees for many helpful suggestions that substantially improved the quality of the paper.

2. **The set up**

2.1. *Preliminaries.* Throughout the paper we will adopt the usual ε-convention of analytic number theory. The presence of an ε in the statement of a proposition or lemma will mean that the estimate is valid for any $\varepsilon > 0$ and the implicit implied constant is allowed to depend on that ε. Moreover the value of the ε may differ from one occurrence to another.

Though for the purpose of subconvexity one only needs to consider very special weight functions, the bounds that we establish hold for a larger class

of functions. In fact it is often convenient to prove the results in this more general setting. As such we introduce the following class of functions. Given a sequence of positive numbers $\mathfrak{A} = \{A_1, A_2, \dots\}$, a positive integer h and a vector of positive numbers $\mathbf{H} = (H_1, \dots, H_h)$, we define a class of functions $\mathcal{W}(\mathbf{H}, \mathfrak{A})$. This consists of smooth functions $W : \mathbb{R}^h \to \mathbb{C}$, which are supported in the box $[1, 2]^h$ and which satisfy

$$|W^{(\mathbf{j})}(\mathbf{x})| \le A_{j_1 + \dots + j_k} \mathbf{H}^{\mathbf{j}}$$

for any $\mathbf{j} \in \mathbb{Z}_{\ge 0}^h$. (Here $\mathbf{H}^{\mathbf{j}} = H_1^{j_1} \dots H_h^{j_h}$.) The sequence $\mathfrak{A}$ will not be of any importance in our analysis, and we will drop it from the notation. In fact there is an universal sequence $\mathfrak{A}(\varepsilon)$ depending only on the smallest ε that appears in our analysis, such that all the weight functions considered below fall in the class $\mathcal{W}(\mathbf{H}) = \mathcal{W}(\mathbf{H}, \mathfrak{A}(\varepsilon))$ for some h and $\mathbf{H}$. If $\mathbf{H} = (H, \dots, H)$, then we will, by abuse of notation, simply write $\mathcal{W}(H)$ (or $\mathcal{W}_h(H)$). The subclass of $\mathcal{W}(\mathbf{H})$ consisting of functions with image in $\mathbb{R}_{\ge 0}$ will be denoted by $\mathcal{W}^+(\mathbf{H})$. For notational convenience we will adopt the following convention regarding the analytic weights. At any situation where more than one weights of the same class are involved, the same notation may be used to denote different weight functions (e.g., see Lemma 14).

Our first lemma is regarding the existence of a smooth partition of unity.

LEMMA 1. *There exists a sequence $\mathcal{U} = \{(U, R)\}$ consisting of pairs (U, R) with $U \in \mathcal{W}_1^+(1)$, $R \in \mathbb{R}_{>0}$ such that*

$$\sum_{(U,R)} U\left(\frac{r}{R}\right) = 1 \quad for \quad r \in (0, \infty).$$

Also the collection is such that the sum is locally finite in the sense that for any given $\ell \in \mathbb{Z}$, there are only finitely many (independent of ℓ) pairs with $R \in [2^\ell, 2^{\ell+1}]$.

Once and for all we choose and fix such a smooth partition of unity. Then by a smooth dyadic subdivision of a sum

$$\sum_{r=1}^{\infty} A(r)$$

we will mean

$$\sum_{(U,R)} \sum_{r=1}^{\infty} A(r) U\left(\frac{r}{R}\right).$$

(It is also possible to choose a single weight function U and vary only the range R to obtain such a partition of unity. This is of course not necessary for our purpose.)

We conclude this section by noting some basic properties of the J-Bessel function.

LEMMA 2. *Let $\kappa \geq 2$ and $x > 0$. Then the J-Bessel function splits as*

$$J_\kappa(x) = W_\kappa(x)e^{ix} + \bar{W}_\kappa(x)e^{-ix},$$

where $W_\kappa(x)$ is a smooth function defined on $(0, \infty)$, and it satisfies the following bound:

$$x^j W_\kappa^{(j)}(x) \ll_\kappa \min\{x^{\kappa-1}, x^{-1/2}\}.$$

2.2. *Petersson formula to detect the equation $n = r$.* Now we will explain the expansion of the Kronecker symbol that we will use. Let p be a prime number, and let $k \equiv 3 \bmod 4$ be a positive integer. Let ψ be a character of $\mathbb{F}_p^\times$ satisfying $\psi(-1) = -1 = (-1)^k$. So, in particular, ψ is primitive modulo p. The collection of Hecke cusp forms of level p, weight k and nebentypus ψ is denoted by $H_k(p, \psi)$, and they form an orthogonal basis of the space of cusp forms $S_k(p, \psi)$. Let

$$\omega_f^{-1} = \frac{\Gamma(k-1)}{(4\pi)^{k-1}\|f\|^2}$$

be the spectral weights. The Petersson formula gives

$$\sum_{f \in H_k(p,\psi)} \omega_f^{-1} \lambda_f(n)\overline{\lambda_f(r)} = \delta(n,r) + 2\pi i \sum_{c=1}^\infty \frac{S_\psi(r,n;cp)}{cp} J_{k-1}\left(\frac{4\pi\sqrt{nr}}{cp}\right).$$

This gives an expansion of the Kronecker delta $\delta(n,r)$ (which is the indicator function of the diagonal $n = r$) in terms of the Kloosterman sums

$$S_\psi(a,b;c) = \sum_{\alpha \bmod c}^\star \psi(\alpha)e\left(\frac{\alpha a + \bar\alpha b}{c}\right)$$

and the (Hecke normalized) Fourier coefficients $\lambda_f(n)$ of holomorphic forms f if pk is taken to be sufficiently large (so that the space $S_k(p, \psi)$ is nontrivial).

Let P be a parameter that shall be chosen optimally later as a power of the modulus M. Let

$$P^\star = \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \sum_{\psi \bmod p} (1 - \psi(-1)) = \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \phi(p) \asymp \frac{P^2}{\log P}.$$

LEMMA 3. *For $Pk \gg 1$ (sufficiently large), we have*

$$(2) \quad \delta(n,r) = \frac{1}{P^\star} \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \sum_{\psi \bmod p} (1 - \psi(-1)) \sum_{f \in H_k(p,\psi)} \omega_f^{-1} \lambda_f(n)\overline{\lambda_f(r)}$$

$$- \frac{2\pi i}{P^\star} \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \sum_{c=1}^\infty \frac{1}{cp} \sum_{\psi \bmod p} (1 - \psi(-1)) \, S_\psi(r,n;cp) J_{k-1}\left(\frac{4\pi\sqrt{nr}}{cp}\right).$$

Proof. The lemma follows by taking an average of the Petersson formula over all odd ψ modulo p and all primes in the range $P < p < 2P$. $\qquad\square$

2.3. *Bounds for central values in terms of short smooth sums.* Let $W \in \mathcal{W}_1(1)$, and define the sum

$$S(N, W) = S(N) = \sum_{m,n=1}^{\infty} \lambda(m,n)\chi(n)W\left(\frac{nm^2}{N}\right).$$

For notational simplicity, we will drop W and denote the sum simply by $S(N)$, with the understanding that when we have a sum of several $S(N)$, then the weight function involved may not be same in each occurrence. The Dirichlet series associated with $S(N)$ is given by

$$\sum_{m,n=1}^{\infty} \lambda(m,n)\chi(n)(nm^2)^{-s},$$

and our first task is to relate this series with the twisted L-function $L(s, \pi \otimes \chi)$.

LEMMA 4. *We have*

(3) $$L(3s,\chi)\sum_{m,n=1}^{\infty} \lambda(m,n)\chi(n)(nm^2)^{-s} = L(s, \pi \otimes \chi)L(2s, \tilde{\pi})$$

for $\sigma > 1$, where $\tilde{\pi}$ denotes the dual form.

Proof. The Dirichlet series, which appears on the left-hand side of (3), is given by the Euler product

$$\prod_{p \text{ prime}} \sum_{u,v=0}^{\infty} \lambda(p^u, p^v)\chi(p)^v p^{-(2u+v)s}.$$

For $u, v \geq 1$, we have (the Hecke relations)

$$\lambda(p^u, p^v) = \lambda(p^u, 1)\lambda(1, p^v) - \lambda(p^{u-1}, 1)\lambda(1, p^{v-1}).$$

Consequently we get

$$\sum_{u,v=0}^{\infty} \lambda(p^u, p^v)\chi(p)^v p^{-(2u+v)s}$$

$$= \sum_{u=0}^{\infty} \lambda(p^u, 1)p^{-2us} \sum_{v=0}^{\infty} \lambda(1, p^v)\chi(p)^v p^{-vs} \left\{1 - \chi(p)p^{-3s}\right\}.$$

The lemma follows. $\qquad\square$

Now we relate the sums $S(N)$ with the twisted central values $L(1/2, \pi \otimes \chi)$.

LEMMA 5. *For any $\theta > 0$, we have*

$$L(\tfrac{1}{2}, \pi \otimes \chi) \ll M^{\varepsilon} \sup_N \frac{|S(N)|}{\sqrt{N}} + M^{3/4 - \theta/2 + \varepsilon},$$

where the supremum is taken over N in the range $M^{3/2-\theta} < N < M^{3/2+\theta}$ and the weight functions W (appearing in the sum $S(N)$) belong to $\mathcal{W}_1(1)$.

Before proving the lemma, we will make a couple of remarks regarding the statement of the lemma. The weight functions W involved in $S(N)$ is allowed to change with N. In fact, as we will see below, the weight functions are obtained by taking a smooth dyadic subdivision of a given weight function. Also we will be using (RS) to prove this statement. Using any nontrivial bound towards (RS) one can prove a weaker statement that will be still sufficient for the purpose of the paper.

Proof. Consider the integral

$$I = \frac{1}{2\pi i} \int_{(2)} M^{-3/4} \Lambda(\tfrac{1}{2} + s, \pi \otimes \chi) \Lambda(1 + 2s, \tilde{\pi}) X^s \frac{\mathrm{d}s}{s}.$$

The product of the completed L-functions appearing above is given by

$$M^{-3/4} \Lambda(s, \pi \otimes \chi) \Lambda(2s, \tilde{\pi}) = M^{3s/2 - 3/4} \gamma(s) L(s, \pi \otimes \chi) L(2s, \tilde{\pi}),$$

where

$$\gamma(s) = \pi^{-9s/2} \prod_{j=1}^{3} \Gamma\left(\frac{s - \alpha_j + \delta}{2}\right) \Gamma\left(\frac{2s + \alpha_j}{2}\right),$$

with $\delta = 0$ if $\chi(-1) = 1$ and $\delta = 1$ if $\chi(-1) = -1$. (Note that the Langlands parameters of the dual form $\tilde{\pi}$ are $(-\alpha_3, -\alpha_2, -\alpha_1)$.) We only need the fact that (under (RS)) there are no poles of $\gamma(s)$ in the region $\sigma > 0$. We move the contour, in the definition of I, to $\sigma = -1/2 + \varepsilon$. The residue at $s = 0$ is given by

$$\gamma(1/2) L(\tfrac{1}{2}, \pi \otimes \chi) L(1, \tilde{\pi}).$$

For the integral at $\sigma = -1/2 + \varepsilon$, which is near the edge of the critical strip, we use trivial bounds to get

$$\frac{1}{2\pi i} \int_{(-1/2 + \varepsilon)} M^{-3/4} \Lambda(\tfrac{1}{2} + s, \pi \otimes \chi) \Lambda(1 + 2s, \tilde{\pi}) X^s \frac{\mathrm{d}s}{s} = O\left(M^{3/4} X^{-1/2} (MX)^\varepsilon\right).$$

On the other hand, from Lemma 4 it follows that the initial integral I is given by

$$\sum_{m,n=1}^{\infty} \frac{\lambda(m,n) \chi(n)}{\sqrt{m^2 n}} \frac{1}{2\pi i} \int_{(2)} \gamma(1/2 + s) L(\tfrac{3}{2} + 3s, \chi) \left(\frac{M^{3/2} X}{nm^2}\right)^s \frac{\mathrm{d}s}{s}.$$

We set

$$\mathcal{V}(y) = \frac{1}{2\pi i} \int_{(2)} \gamma(1/2 + s) L(\tfrac{3}{2} + 3s, \chi) y^{-s} \frac{\mathrm{d}s}{s}.$$

For $y \geq M^\varepsilon$, we see that $\mathcal{V}(y) \ll M^{-2013}$ by shifting the contour to the right. For $0 < y < M^\varepsilon$, we shift the contour to $\sigma = \varepsilon$. Differentiating within the integral sign we get

$$y^j \mathcal{V}^{(j)}(y) \ll_j 1.$$

It follows that

$$\gamma(1/2)L(\tfrac{1}{2}, \pi \otimes \chi)L(1, \tilde{\pi}) = \sum_{m,n=1}^{\infty} \frac{\lambda(m,n)\chi(n)}{\sqrt{m^2 n}} \mathcal{V}\left(\frac{nm^2}{XM^{3/2}}\right)$$
$$+ O\left(M^{3/4}X^{-1/2}(MX)^{\varepsilon}\right).$$

Since $L(1, \tilde{\pi}) \gg 1$, taking a smooth dyadic subdivision (as in Section 2.1) and picking $X = M^{\theta}$, we conclude the lemma. Note that for U as in Lemma 1 and $\mathcal{V}$ as above, the function

$$W(x) := U(x)\mathcal{V}\left(\frac{Rx}{XM^{3/2}}\right)$$

belongs to the class $\mathcal{W}_1(1)$. $\square$

One will notice a certain oddity in Lemma 5. The Dirichlet series expansion of the L-function $L(s, \pi \otimes \chi)$ is given by (1). In the usual approximate functional equation, as given in Chapter 5 of [7], one gets an expression for $L(1/2, \pi \otimes \chi)$ in terms of sums of the type

$$\sum_{n=1}^{\infty} \lambda(1,n)\chi(n)W(n/N).$$

In this usual form one only needs to take $N \ll M^{3/2+\varepsilon}$, as the tail makes a negligible contribution. In the above lemma we have given a bound for the twisted central value in terms of a slightly different smooth sums, namely, $S(N)$. This is done at the cost of a larger error of size $O(M^{3/4-\theta+\varepsilon})$. Moreover we are required to consider longer sums $N \ll M^{3/2+\theta+\varepsilon}$. One may wonder what the advantage is of such an expression. The point is that after applying the Petersson formula we will be led to consider the Rankin-Selberg L-function $L(s, \pi \otimes f)$ that is given by the Dirichlet series

$$L(s, \pi \otimes f) = \sum_{m,n=1}^{\infty} \lambda(m,n)\lambda_f(n)(m^2 n)^{-s}.$$

Hence, in hindsight, the introduction of the extra sum over m will turn out to be beneficial.

In the sum $S(N)$ we are going to separate the oscillation of the Fourier coefficients $\lambda(m,n)$ from that of the character $\chi(n)$. Consequently we would like to have a separate smooth weight function for n that will not depend on m. To this end, let V be a smooth function supported in $[M^{-4\theta}, 4]$, with $V(x) = 1$ for $x \in [2M^{-4\theta}, 2]$ and satisfying $y^j V^{(j)}(y) \ll_j 1$, and set

$$(4) \qquad S^{\star}(N) = \sum_{m,n=1}^{\infty} \lambda(m,n)\chi(n)W\left(\frac{nm^2}{N}\right)V\left(\frac{n}{N}\right).$$

LEMMA 6. *For $S^\star(N)$ as above, we have*

(5) $$S^\star(N) = S(N) + O(NM^{-\theta+\varepsilon}).$$

Moreover

$$L(\tfrac{1}{2}, \pi \otimes \chi) \ll M^\varepsilon \sup_N \frac{|S^\star(N)|}{\sqrt{N}} + M^{3/4-\theta/2+\varepsilon},$$

where the supremum is taken over N in the range $M^{3/2-\theta} < N < M^{3/2+\theta}$.

Proof. Using the definition of V and the bound

(6) $$\sum\sum_{m^2 n \le x} |\lambda(m,n)|^2 \ll x^{1+\varepsilon},$$

which follows from the Rankin-Selberg theory, we get the first statement. Now substituting (5) in Lemma 5 and using $N < M^{3/2+\theta}$, we get the second statement. □

2.4. *Conclusion.* We will apply the formula from Lemma 3 directly to the sum $S^\star(N)$, which we first rewrite as

(7) $$S^\star(N) = \sum_{r=1}^\infty \sum_{m,n=1}^\infty \lambda(m,n)\chi(r)\delta(n,r)\, W\left(\frac{nm^2}{N}\right) V\left(\frac{r}{N}\right).$$

We take N in the range $M^{3/2-\theta} \le N \ll M^{3/2+\theta}$, with $\theta > 0$. The parameter θ shall be taken sufficiently small at the end. Applying (2) from Lemma 3 to (7) we get two terms, namely,

$$S^\star(N) = \mathcal{F} - 2\pi i\, \mathcal{O},$$

where

(8) $$\mathcal{F} = \frac{1}{P^\star} \sum_{\substack{P<p<2P \\ p \text{ prime}}} \sum_{\psi \bmod p} (1-\psi(-1)) \sum_{f \in H_k(p,\psi)} \omega_f^{-1}$$

$$\times \sum_{m,n=1}^\infty \lambda(m,n)\lambda_f(n) W\left(\frac{nm^2}{N}\right) \sum_{r=1}^\infty \overline{\lambda_f(r)}\chi(r)\, V\left(\frac{r}{N}\right)$$

and

(9) $$\mathcal{O} = \frac{1}{P^\star} \sum_{\substack{P<p<2P \\ p \text{ prime}}} \sum_{\psi \bmod p} (1-\psi(-1)) \sum_{m,n=1}^\infty \lambda(m,n) W\left(\frac{nm^2}{N}\right)$$

$$\times \sum_{r=1}^\infty \chi(r) V\left(\frac{r}{N}\right) \sum_{c=1}^\infty \frac{S_\psi(r,n;cp)}{cp} J_{k-1}\left(\frac{4\pi\sqrt{nr}}{cp}\right).$$

We pick the weight k to be large, say of size ε^{-1}. The second sum, which we call the off-diagonal, can be nicely bounded if P is taken sufficiently large. On the other hand, to the first sum we will apply the functional equations followed by the Petersson formula. The resulting diagonal term vanishes, and

the off-diagonal term (which we will call dual off-diagonal) can be bounded nicely if P is taken in a suitable range. We will show that there is a choice of P for which both the terms can be bounded satisfactorily.

In the rest of the paper we will prove the following two propositions.

PROPOSITION 1. *Let $\mathcal{O}$ be as defined in (9). Suppose $P \geq N^{1/2+\varepsilon}$ with $\varepsilon > 0$ and $\theta < 1/4$. Then we have*

$$\mathcal{O} \ll \sqrt{N}\frac{M^{3/2+\theta+\varepsilon}}{P}.$$

PROPOSITION 2. *Let $\mathcal{F}$ be as defined in (8). Suppose θ is sufficiently small. Then we have $\theta_0, \delta > 0$, such that for*

$$P = M^{1-\theta_0},$$

we have

$$\mathcal{F} \ll \sqrt{N}M^{3/4-\delta}.$$

Assuming the propositions, let us complete the proof of Theorem 1. From the propositions we conclude that for $\theta > 0$ sufficiently small, we have a $\delta > 0$ such that

$$S^\star(N) \ll \sqrt{N}M^{3/4-\delta}.$$

Substituting this into Lemma 6, we obtain the theorem.

3. Sketch of the proof

In this section we shall present a very rough sketch of the ideas involved in the proof. We need strong bounds for $\mathcal{O}$ and $\mathcal{F}$. Choose P so that $N^{1/2+\varepsilon} < P < M^{1-\varepsilon}$.

3.1. *Analysis of $\mathcal{O}$.* First consider $\mathcal{O}$, which is defined in (9). Roughly speaking it is given by

$$\frac{1}{NP^2} \sum_{\substack{P<p<2P \\ p \text{ prime}}} \sum_{\psi \bmod p} (1 - \psi(-1)) \sum_{n \sim N} \lambda(1,n) \sum_{r \sim N} \chi(r) \sum_{c \sim N/P} S_\psi(r,n;cp).$$

Using the trivial bound (not the Weil bound) for the Kloosterman sum we get that this sum is dominated by $O(N^3/P)$. So after opening the Kloosterman sum we need to make a saving of size N^2/P. Since by our choice $P \gg \sqrt{N}$, we have $(c,p) = 1$, and consequently the Kloosterman sum splits as

$$S_\psi(r,n;cp) = S_\psi(\bar{c}r,\bar{c}n;p)S(\bar{p}r,\bar{p}n;c).$$

Summing over ψ we get

$$\sum_{\psi \bmod p} (1 - \psi(-1)) S_\psi(\bar{c}r,\bar{c}n;p) = \phi(p)\left(e\left(\frac{\bar{c}(r+n)}{p}\right) - e\left(-\frac{\bar{c}(r+n)}{p}\right)\right).$$

This gives us a saving of size P. (In other words, we are saving $\sqrt{P}$ in the ψ sum in addition to the Weil bound for the Kloosterman sums.) Now we are required to save N^2/P^2 in the sum

$$\sum_{\substack{P<p<2P \\ p \text{ prime}}} \sum_{n\sim N} \lambda(1,n) \sum_{r\sim N} \chi(r) \sum_{c\sim N/P} S(\bar{p}r, \bar{p}n; c)\, e\left(\frac{\bar{c}(r+n)}{p}\right).$$

We wish to apply the Poisson summation formula on the sum over r. But before that we shall reduce the 'conductor' of the sum by applying the reciprocity relation

$$e\left(\frac{\bar{c}(r+n)}{p}\right) = e\left(-\frac{\bar{p}(r+n)}{c}\right) e\left(\frac{r+n}{cp}\right).$$

The last factor can be absorbed in the weight function as $cp \sim (r+n)$. So we consider the sum

$$\sum_{\substack{P<p<2P \\ p \text{ prime}}} \sum_{n\sim N} \lambda(1,n) \sum_{c\sim N/P} \sum_{r\sim N} \chi(r) S(\bar{p}r, \bar{p}n; c)\, e\left(-\frac{\bar{p}(r+n)}{c}\right).$$

We break the inner sum into congruence classes modulo cM and apply the Poisson summation formula. The resulting character sums are given by

$$\sum_{a \bmod cM} \chi(a) S(\bar{p}a, \bar{p}n; c)\, e\left(-\frac{\bar{p}(a+n)}{c} + \frac{ar}{cM}\right).$$

Since $c\sim N/P \ll \sqrt{N} \ll M$ and M is assumed to be prime, we have $(c,M)=1$. Hence the character sum splits into a product of two character sums. The one with modulus M is just a Gauss sum. Consequently we find that the contribution of the zero frequency vanishes. The character sum with modulus c can be evaluated explicitly after opening the Kloosterman sum, and we have

$$\sum_{a \bmod c} S(\bar{p}a, \bar{p}n; c)\, e\left(-\frac{\bar{p}(a+n)}{c} + \frac{\bar{M}ar}{c}\right) = c\, e\left(\frac{(\overline{M-pr})rn}{c}\right).$$

So it follows that using the Poisson summation formula we have a saving of size $N/\sqrt{M}$ — in other words, saving $N/\sqrt{cM}$ beyond the Weil bound for the Kloosterman sum modulo c. It remains to save $N\sqrt{M}/P^2 = (M/P)^2$. This is just at the threshold, as to get a satisfactory bound for $\mathcal{F}$ (as we will see below), we need to take P to be just smaller than M. We wish to save $(M/P)^2$ (and a little more) in the sum

$$\sum_{\substack{P<p<2P \\ p \text{ prime}}} \sum_{c\sim N/P} \sum_{n\sim N} \lambda(1,n) \sum_{|r|\ll M/P} \chi(c\bar{r})\, e\left(\frac{(\overline{M-pr})rn}{c}\right).$$

This can be achieved in several ways. One may apply the Voronoi summation formula on the sum over n or directly apply Miller's bound from [10] to the

n sum. However one may refrain from utilizing the fact that the coefficients $\lambda(1,n)$ come from a $\mathrm{SL}(3,\mathbb{Z})$ Hecke-Maass form, and one can obtain a bound that holds for any coefficients $\alpha(n)$ in place of $\lambda(1,n)$ satisfying the bound $\alpha(n) \ll n^{\varepsilon}$ either point-wise or in the L^2 sense. To this end we apply the Cauchy inequality and consider the sum

$$\sum_{\substack{P<p<2P \\ p \text{ prime}}} \sum_{n\sim N} \left| \sum_{c\sim N/P} \sum_{|r|\ll M/P} \chi(c\bar{r})\, e\left(\frac{\overline{(M-pr)}rn}{c} \right) \right|^2.$$

Now we open the absolute square and apply the Poisson summation on the sum over n with modulus cc'. Only the zero frequency survives as $N > cc' \sim N^2/P^2$. We end up with the counting problem

$$N \sum_{\substack{P<p<2P \\ p \text{ prime}}} \sum_{c,c'\sim N/P} \sum_{\substack{|r|,|r'|\ll M/P \\ (M-pr')rc'-(M-pr)r'c\equiv 0 \bmod cc'}} 1.$$

In Section 4.3 we deal with this counting problem and obtain a sufficient bound. One saves the length of the diagonal, i.e., MN/P^2, which is larger than $(M/P)^4$ if $P > M^{3/4+\varepsilon}$. This is the content of Proposition 1.

3.2. *Analysis of $\mathcal{F}$.* Next we turn to $\mathcal{F}$, which is roughly of the form

$$\frac{1}{P^2} \sum_{\substack{P<p<2P \\ p \text{ prime}}} \sum_{\psi \bmod p} (1-\psi(-1)) \sum_{f\in H_k(p,\psi)} \omega_f^{-1} \sum_{n\sim N} \lambda(1,n)\lambda_f(n) \sum_{r\sim N} \overline{\lambda_f(r)}\chi(r).$$

The trivial bound for this sum is N^2, and we need to save N. We use the functional equations to the sums over n and r. The resulting size of the n sum is P^3/N, and that of the r sum is M^2P/N. So we are able to save N^2/MP^2, and it remains to save $MP^2/N = P^2/\sqrt{M}$. However major complications arise due to the root numbers, which involve the Fourier coefficient $\lambda_f(p)$ and the Gauss sums g_ψ (where g_η denotes the Gauss sums associated with the character η). Roughly speaking one is now left with the problem of saving $P^2/\sqrt{M}$ in a sum of the type

$$\sum_{\substack{P<p<2P \\ p \text{ prime}}} \chi(p) \sum_{\psi \bmod p} (1-\psi(-1))\, \bar{\psi}(-M) g_\psi^2 \sum_{f\in H_k(p,\psi)} \omega_f^{-1} \overline{\lambda_f(p)}^2$$

$$\times \sum_{n\sim P^3/N} \lambda(n,1)\overline{\lambda_f(n)} \sum_{r\sim P\sqrt{M}} \bar{\chi}(r)\lambda_f(r).$$

Applying Petersson's formula one gets

$$\sum_{\substack{P < p < 2P \\ p \text{ prime}}} \chi(p) \sum_{\psi \bmod p} (1 - \psi(-1)) \, \bar{\psi}(-M) g_\psi^2 \sum_{n \sim P^3/N} \lambda(n, 1)$$

$$\times \sum_{r \sim P\sqrt{M}} \bar{\chi}(r) \sum_{c \sim P^2/\sqrt{M}} S_\psi(np^2, r; cp).$$

Note that the diagonal term is nonexistent. In this process we make a loss of size $P^2/\sqrt{M}$ (if one opens the Kloosterman sum). So now we need to save P^4/M (and a little more). There are two cases to consider. In the case where $p|c$, we get the required saving quite easily as it turns out that the Kloosterman sum vanishes unless $p|r$. So we already make a saving of size P^2. Summing over ψ we save P more. This is enough for our purpose as $P^3 > P^4/M$ (or $M > P$). In the other case where $p \nmid c$, the Kloosterman sum splits as

$$S_\psi(np^2, r; cp) = S_\psi(0, \bar{c}r; p) S(n, r; c).$$

Observe the curious separation of variables. The first term on the right-hand side does not involve n, and the second term is free of p. This plays a crucial role in our analysis. The first term is just a Gauss sum, and it is given by $\psi(\bar{c}r) g_{\bar{\psi}}$. We are now required to save P^4/M (taking the trivial bound for the Gauss sum, i.e., after opening the Gauss sum) in the sum

$$\sum_{\substack{P < p < 2P \\ p \text{ prime}}} \chi(p) \sum_{n \sim P^3/N} \lambda(n, 1) \sum_{\substack{r \sim P\sqrt{M} \\ p \nmid r}} \bar{\chi}(r)$$

$$\times \sum_{\substack{c \sim P^2/\sqrt{M} \\ p \nmid c}} S(n, r; c) \sum_{\psi \bmod p} (1 - \psi(-1)) \, \bar{\psi}(Mc\bar{r}) g_\psi.$$

The sum over ψ now yields

$$\sum_{\psi \bmod p} (1 - \psi(-1)) \, \bar{\psi}(Mc\bar{r}) g_\psi = \phi(p) \left\{ e\left(\frac{Mc\bar{r}}{p}\right) - e\left(-\frac{Mc\bar{r}}{p}\right) \right\},$$

which gives us a saving of size P. Now we wish to save P^3/M in the sum

$$\sum_{\substack{P < p < 2P \\ p \text{ prime}}} \chi(p) \sum_{n \sim P^3/N} \lambda(n, 1) \sum_{\substack{r \sim P\sqrt{M} \\ p \nmid r}} \bar{\chi}(r) \sum_{\substack{c \sim P^2/\sqrt{M} \\ p \nmid c}} S(n, r; c) e\left(\frac{Mc\bar{r}}{p}\right).$$

Now we apply the Voronoi summation formula on the sum over n. Roughly speaking this yields

$$\sum_{n \sim P^3/N} \lambda(n, 1) S(n, r; c) = \frac{P^3}{c^2 N} \sum_{1 \le n \ll c^3 N/P^3} \lambda(1, n) \sum_{a \bmod c}^{\star} e\left(\frac{ar}{c}\right) S(a, n; c).$$

Let us pretend that the last sum over a is a complete sum without the coprimality condition. Then the right-hand side reduces to

$$\frac{P^3}{cN} \sum_{1 \leq n \ll c^3 N/P^3} \lambda(1,n) e\left(-\frac{\bar{r}n}{c}\right),$$

which is approximately

$$\frac{P}{M} \sum_{1 \leq n \ll P^3} \lambda(1,n) e\left(-\frac{\bar{r}n}{c}\right) = \frac{P}{M} \sum_{\substack{L \text{ dyadic} \\ L < P^3}} \sum_{n \sim L} \lambda(1,n) e\left(-\frac{\bar{r}n}{c}\right).$$

We are taking a smooth partition of the sum into dyadic blocks, as we intend to apply the Voronoi summation again after an application of reciprocity. We consider the L-th block, which is given by

$$\sum_{\substack{P < p < 2P \\ p \text{ prime}}} \chi(p) \sum_{\substack{r \sim P\sqrt{M} \\ p \nmid r}} \bar{\chi}(r) \sum_{\substack{c \sim P^2/\sqrt{M} \\ p \nmid c}} \sum_{n \sim L} \lambda(1,n) e\left(-\frac{\bar{r}n}{c}\right) e\left(\frac{Mc\bar{r}}{p}\right).$$

We need to save L/P. For L large, i.e., $L \sim P^3$, we need to save P^2, which is in fact more than what we were required to save before the Voronoi summation. But we have gained immense structural advantage in the process. In particular, the Kloosterman sum has vanished and we are now able to reduce the conductor by applying reciprocity. Note that the length of the n-sum is proportional to the amount that we need to save. Here we are using the Ramanujan-Selberg conjecture (RS). Applying reciprocity we reduce the above sum to

$$\sum_{\substack{P < p < 2P \\ p \text{ prime}}} \chi(p) \sum_{\substack{r \sim P\sqrt{M} \\ p \nmid r}} \bar{\chi}(r) \sum_{\substack{c \sim P^2/\sqrt{M} \\ p \nmid c}} \sum_{n \sim L} \lambda(1,n) e\left(\frac{\bar{c}n}{r}\right) e\left(-\frac{Mc\bar{p}}{r}\right).$$

Note that $cr \sim P^3 \gg L$ and $pr \sim P^2\sqrt{M} \sim Mc$. Next we apply the Poisson summation on the sum over c and the Voronoi summation on the sum over n. First the Poisson summation gives us a saving of the size $P^{3/2}/M^{3/4}$. Then the Voronoi summation saves $L/P^{3/2}M^{3/4}$. So in total we save $L/M^{3/2}$. One is now required to save $M^{3/2}/P$ in the sum

$$\sum_{\substack{P < p < 2P \\ p \text{ prime}}} \chi(p) \sum_{\substack{r \sim P\sqrt{M} \\ p \nmid r}} \bar{\chi}(r) \sum_{c \sim M/P} \sum_{n < P^3 M^{3/2}/L} \lambda(n,1)$$

$$\times \sum_{a \bmod r} S(a,n;r) e\left(-\frac{Ma\bar{p}}{r} + \frac{ac}{r}\right).$$

The last character sum is given by

$$r e\left(\frac{\overline{M - cppn}}{r}\right).$$

Next we apply the Cauchy inequality to get rid of the Fourier coefficients and reduce the problem to that of saving $M^3/P^2 = (N/P)^2$ in the sum

$$\sum_{c \sim M/P} \sum_{n < P^3 M^{3/2}/L} \left| \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \sum_{\substack{r \sim P\sqrt{M} \\ p \nmid r}} \chi(p\bar{r}) \, e\left(\frac{\overline{M - cppn}}{r} \right) \right|^2.$$

Opening the absolute square we apply the Poisson summation formula on the sum over n. This leads us to a counting problem. 'Ideally' one should be able to save $P^3 M^{3/2}/L > M^{3/2} = N$. But there are few degenerate cases that make large contribution in the count. However we get a satisfactory bound as long as P is taken in a suitable range in $[N^{1/2+\varepsilon}, M^{1-\varepsilon}]$. One will also observe that the present counting problem is same (with different variable sizes) as the one we encountered while dealing with the off-diagonal $\mathcal{O}$. However in the present case we need to analyze the contribution of the nonzero frequencies as well.

In the above sketch we concentrated only on the transition range for the c sum. Due to the rapid decay of the Bessel function, and our choice of large weight k, the tail sum, where c is larger than the transition range, makes a negligible contribution. However for smaller c, we need a detailed analysis to get a satisfactory bound. In this range we gain a little from the size of the Bessel function, but there is an additional oscillation. We analyze these terms in Section 6. The content of this section is summarized in Lemma 18, where we show that by taking P between $[N^{1/2+\varepsilon}, M^{1-\varepsilon}]$ we can obtain a satisfactory bound. Similarly in our analysis of the contribution of the 'wild terms' in the transition range, i.e., those with variables having large common factors with each other, in Section 7 we again need to take $P < M$. The main output of this section is Lemma 24. Also we stress that in our analysis of these sums we make a very small saving compared to our savings in Proposition 1 or in our treatment for c in the transition range. With more work one expects to get better bounds in Sections 6 and 7. Indeed one expects that the optimal choice for P should be near $\sqrt{N}$, rather than near M, as in the standard circle method.

4. The off-diagonal

In this section we will analyze the off-diagonal contribution $\mathcal{O}$ as given in (9). We will prove Proposition 1. Suppose we take $P \geq N^{1/2+\varepsilon}$. Since we are picking k very large, of the order ε^{-1}, and $J_{k-1}(x) \ll x^{k-1}$ (see Lemma 2), the contribution from the tail $c > N^{1/2-\varepsilon}$ is negligibly small. In particular, the contributing c are necessarily coprime with p. We make a dyadic subdivision of the the c-sum (see Section 2.1) and extract the oscillation from the Bessel

function as in Lemma 2. This leads us to the study of the sum

$$\mathcal{O}(m) = \sum_{\substack{P<p<2P \\ p\text{ prime}}} \sum_{\psi \bmod p} (1 - \psi(-1)) \sum_{n=1}^{\infty} \lambda(m,n) \sum_{r=1}^{\infty} \chi(r)$$

$$\times \sum_{\substack{c=1 \\ (c,p)=1}}^{\infty} S_{\psi}(r,n;cp) e\left(\frac{2\sqrt{nr}}{cp}\right) W_0\left(\frac{n}{N_0}, \frac{r}{R}, \frac{c}{C}\right)$$

for any fixed $m \le \sqrt{N}$, where $W_0 \in \mathcal{W}_3(1)$ and

$$N_0 = N/m^2, \quad NM^{-4\theta} \ll R \ll N \quad \text{and} \quad C \ll \sqrt{N_0 R}M^{\varepsilon}/P.$$

From a bound for $\mathcal{O}(m)$ we can conclude a bound for $\mathcal{O}$ via the inequality

$$(10) \qquad\qquad \mathcal{O} \ll \frac{M^{\varepsilon}}{P^2} \sum_{1 \le m \ll \sqrt{N}} \sup \frac{\mathcal{O}(m)}{\sqrt{CP}(N_0 R)^{1/4}},$$

where the supremum is taken over all C and R in the above ranges. (Technically speaking one should also take supremum over a class of weight functions, but that does not affect the bound. This feature will be present throughout this paper.)

4.1. *Sum over ψ and reciprocity.* Our next step is a conductor lowering mechanism. This is one of the most vital steps. Similar tricks were also used in the series of papers [13], [12] and [14]. There a part of the Kloosterman sum could be evaluated as the modulus was powerful. Here the extra average over ψ helps us to evaluate precisely the twisted average value of the Kloosterman sum. This also makes way for the application of reciprocity, which lowers the conductor. We set

$$(11) \qquad W_1^{\pm}(x,y,z) = e\left(\pm\frac{Ry + N_0 x}{Cpz}\right) e\left(\frac{2\sqrt{RN_0 xy}}{Cpz}\right) W_0(x,y,z).$$

Note that the new factors are only mildly oscillating at the transition range for c. Indeed we have $W_1^{\pm}(x,y,z) \in \mathcal{W}(H_1, H_2, H_1 + H_2)$ (see Section 2.1), where

$$H_1 = \left(\frac{N_0}{CP} + \frac{\sqrt{RN_0}}{CP} + 1\right) \quad \text{and} \quad H_2 = \left(\frac{R}{CP} + \frac{\sqrt{RN_0}}{CP} + 1\right).$$

LEMMA 7. *We have*

$$\mathcal{O}(m) = \sum_{\pm}(\pm 1) \sum_{\substack{P<p<2P \\ p\text{ prime}}} \sum_{\substack{c=1 \\ (c,p)=1}}^{\infty} \phi(p) \sum_{n=1}^{\infty} \lambda(m,n)$$

$$\times \sum_{r=1}^{\infty} \chi(r) S(\bar{p}r, \bar{p}n; c) e\left(\mp\frac{\bar{p}(r+n)}{c}\right) W_1^{\pm}\left(\frac{n}{N_0}, \frac{r}{R}, \frac{c}{C}\right),$$

where $W_1^{\pm}$ are as given in (11).

Proof. Using the coprimality $(c, p) = 1$, we get

$$\sum_{\psi \bmod p} (1 - \psi(-1)) \, S_\psi(r, n; cp)$$

$$= S(\bar{p}r, \bar{p}n; c) \sum_{\psi \bmod p} (1 - \psi(-1)) \, S_\psi(\bar{c}r, \bar{c}n; p)$$

$$= \phi(p) S(\bar{p}r, \bar{p}n; c) \, \left(e\left(\frac{\bar{c}(r+n)}{p} \right) - e\left(-\frac{\bar{c}(r+n)}{p} \right) \right).$$

Hence

$$\mathcal{O}(m) = \sum_{\pm}(\pm 1) \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \phi(p) \sum_{n=1}^{\infty} \lambda(m, n) \sum_{r=1}^{\infty} \chi(r)$$

$$\times \sum_{\substack{c=1 \\ (c,p)=1}}^{\infty} S(\bar{p}r, \bar{p}n; c) e\left(\pm\frac{\bar{c}(r+n)}{p} \right) e\left(\frac{2\sqrt{nr}}{cp} \right) W_0\left(\frac{n}{N_0}, \frac{r}{R}, \frac{c}{C} \right).$$

Next we use the reciprocity relation

$$e\left(\pm\frac{\bar{c}(r+n)}{p} \right) = e\left(\mp\frac{\bar{p}(r+n)}{c} \right) e\left(\pm\frac{r+n}{cp} \right).$$

We push the last oscillatory factor into the weight function. The lemma follows.
$\square$

Observe that before the application of the reciprocity relation the modulus for the sum over r was cpM. Using the reciprocity relation we have brought it down to cM.

4.2. *First application of the Poisson summation.* For notational simplicity we will only focus on the contribution of the '+' term in the expression given in Lemma 7. We denote this by $\mathcal{O}_1(m)$. We also set $W_1 = W_1^+$. We start by applying the Poisson summation formula on the sum over r. Let

$$H = \left(1 + \frac{\sqrt{N_0}}{\sqrt{R}} \right) \frac{M^{1+\varepsilon}}{P}.$$

LEMMA 8. *We have* $\mathcal{O}_1(m) \ll |\mathcal{O}_1^\star(m)| + M^{-2013}$, *where*

$$\mathcal{O}_1^\star(m) = \frac{R}{\sqrt{M}} \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \sum_{\substack{c=1 \\ (c,p)=1}}^{\infty} \phi(p) \sum_{n=1}^{\infty} \lambda(m, n)$$

$$\times \sum_{\substack{|r| < H \\ (pr-M,c)=1}} \chi(c\bar{r}) \, e\left(\frac{(\overline{M - pr})rn}{c} \right) W_1^\star\left(\frac{n}{N_0}, \frac{rR}{cM}, \frac{c}{C} \right),$$

with

$$W_1^\star(x, y, z) = \int_{\mathbb{R}} W_1(x, u, z) \, e(-uy) \, \mathrm{d}u.$$

Proof. Consider the sum over r in the expression in Lemma 7 (only the '+' term). Splitting into congruence classes modulo cM we obtain

$$\sum_{a \bmod cM} \chi(a)S(\bar{p}a, \bar{p}n; c)e\left(-\frac{\bar{p}(a+n)}{c}\right)\sum_{r\in\mathbb{Z}}W_1\left(\frac{n}{N_0}, \frac{(a+rcM)}{R}, \frac{c}{C}\right).$$

By the Poisson summation formula we now get

$$\sum_{a \bmod cM} \chi(a)S(\bar{p}a, \bar{p}n; c)e\left(-\frac{\bar{p}(a+n)}{c}\right)$$
$$\times \sum_{r\in\mathbb{Z}}\int_{\mathbb{R}}W_1\left(\frac{n}{N_0}, \frac{(a+xcM)}{R}, \frac{c}{C}\right)e(-rx)\mathrm{d}x.$$

The change of variables $(a+xcM)/R \mapsto y$ reduces the above sum to

$$\frac{R}{cM}\sum_{r\in\mathbb{Z}}\left\{\sum_{a\bmod cM}\chi(a)S(\bar{p}a, \bar{p}n; c)e\left(-\frac{\bar{p}(a+n)}{c}+\frac{ar}{cM}\right)\right\}$$
$$\times \int_{\mathbb{R}}W_1\left(\frac{n}{N_0}, y, \frac{c}{C}\right)e\left(-\frac{rRy}{cM}\right)\mathrm{d}y.$$

Since $C \ll N^{1/2}M^\varepsilon < M$, we have $(c, M) = 1$ (as we are assuming M to be prime and θ to be small, say $\theta < 1/4$). So the character sum splits into a product of two character sums

$$\sum_{\alpha \bmod M}\chi(\alpha)e\left(\frac{\alpha \bar{c}r}{M}\right)\sum_{a\bmod c}S(\bar{p}a, \bar{p}n; c)e\left(-\frac{\bar{p}(a+n)}{c}+\frac{a\bar{M}r}{c}\right).$$

Writing the first sum in terms of the Gauss sum and opening the Kloosterman sum, we get

$$\varepsilon_\chi\chi(c\bar{r})\sqrt{M}\sum_{b\bmod c}^{\star}e\left(\frac{\bar{p}n(\bar{b}-1)}{c}\right)\sum_{a\bmod c}e\left(\frac{\bar{p}a(b-1)}{c}+\frac{a\bar{M}r}{c}\right),$$

where ε_χ is the sign of the Gauss sum for χ. Next we execute the sum over a to arrive at

$$(12)\quad \varepsilon_\chi\chi(c\bar{r})\,c\sqrt{M}\,e\left(\frac{(1-\bar{M}pr-1)\bar{p}n}{c}\right) = \varepsilon_\chi\chi(c\bar{r})\,c\sqrt{M}\,e\left(\frac{(\bar{M}-pr)rn}{c}\right).$$

In particular, this means that the character sum vanishes unless $(pr-M, c) = 1$.

Next we consider the integral. By repeated integration by parts we have

$$\int_{\mathbb{R}}W_1\left(\frac{n}{N_0}, y, \frac{c}{C}\right)e\left(-\frac{rRy}{cM}\right)\mathrm{d}y \ll_j \left[\left(\frac{R}{CP}+\frac{\sqrt{RN_0}}{CP}+1\right)\frac{CM}{rR}\right]^j.$$

Hence the integral is negligibly small if

$$|r| \gg M^\varepsilon\left(\frac{M}{P}+\frac{M\sqrt{N_0}}{P\sqrt{R}}+\frac{CM}{R}\right).$$

The second term 'essentially' dominates the last term as $C \ll \sqrt{N_0 R} M^\varepsilon / P$. Hence the tail $|r| \gg H$ makes a negligible contribution, say of size $O(M^{-2013})$. The lemma follows. $\qquad\square$

We will conclude this subsection by noting a nontrivial bound for the Fourier transform $W_1^\star$ for smaller values of r. Using the explicit form of W_1 as given in (11) and the second derivative test for exponential integrals, we get

$$(13) \qquad \int_{\mathbb{R}} W_1 \left(\frac{n}{N_0}, y, \frac{c}{C} \right) e \left(-\frac{rRy}{cM} \right) dx \ll_j \frac{\sqrt{CP}}{(N_0 R)^{1/4}}.$$

A more elaborate analysis can be carried out using the stationary phase method. It turns out that the contribution of the stationary point nullifies the oscillation coming from the additive character in (12), via the reciprocity relation. This can be used if one wants a better exponent in the main result.

4.3. *Cauchy inequality and second application of Poisson summation.* Using Cauchy's inequality we get

$$(14) \qquad \mathcal{O}_1^\star(m) \ll \frac{RP^2 M^\varepsilon}{\sqrt{M}} \sqrt{\Lambda_m} \sup_{\substack{P<p<2P \\ p \text{ prime}}} \sqrt{\Psi_p},$$

where

$$\Lambda_m = \sum_{n \leq 10N/m^2} |\lambda(m,n)|^2$$

and

$$\Psi_p = \sum_{n=1}^\infty \left| \sum_{\substack{1 \leq c < \infty \\ |r| < H \\ (c, p(pr-M))=1}} \chi(c\bar{r}) \, e \left(\frac{(\overline{M-pr})rn}{c} \right) W_1^\star \left(\frac{n}{N_0}, \frac{rR}{cM}, \frac{c}{C} \right) \right|^2.$$

LEMMA 9. *We have*

$$\mathcal{O}_1^\star(m) \ll \sqrt{\Lambda_m}\, N_0^{1/4} R^{3/4} P^{5/2} \sqrt{CH(C+H)} M^{-1/2+\varepsilon}.$$

Proof. The lemma will follow once we obtain a satisfactory bound for Ψ_p. Consider the expression for Ψ_p as given above. We open the absolute square and apply the Poisson summation formula on the sum over n after splitting the sum into congruence classes modulo cc'. This gives

$$\Psi_p = N_0 \sum_{\substack{1 \leq c, c' < \infty \\ |r|, |r'| < H \\ (c, p(pr-M))=1 \\ (c', p(pr'-M))=1 \\ \overline{(M-pr)c'r-(M-pr')cr'} \equiv n \bmod cc'}} \sum_{n \in \mathbb{Z}} \chi(cr'\overline{c'r}) \, U(n, r, r', c, c'),$$

where

$$U(n, r, r', c, c') = \int_{\mathbb{R}} W_1^\star\left(x, \frac{rR}{cM}, \frac{c}{C}\right) \bar{W}_1^\star\left(x, \frac{r'R}{c'M}, \frac{c'}{C}\right) e\left(\frac{nN_0 x}{cc'}\right) dx.$$

By repeated integration by parts we have

$$U(n, r, r', c, c') \ll_j \left[\left(\frac{N_0}{CP} + \frac{\sqrt{RN_0}}{CP} + 1\right) \frac{C^2}{nN_0}\right]^j.$$

Hence the integral is negligibly small if

$$|n| \gg M^\varepsilon \left(\frac{C}{P} + \frac{C\sqrt{R}}{P\sqrt{N_0}} + \frac{C^2}{N_0}\right).$$

Since $C \ll \sqrt{N_0 R} M^\varepsilon / P$ and $N_0, R \ll N$, we see that the right-hand side is dominated by $O(NM^\varepsilon/P^2)$. So if $P \geq N^{1/2+\varepsilon}$, then the contribution of the nonzero frequencies $n \neq 0$ is negligibly small. Hence

$$\Psi_p \ll N_0 \frac{CP}{\sqrt{N_0 R}} \underset{\substack{1 \leq c, c' \ll C \\ 1 \leq |r|, |r'| < H \\ (c, p(pr-M))=1 \\ (c', p(pr'-M))=1 \\ \overline{(M-pr)}c'r - \overline{(M-pr')}cr' \equiv 0 \bmod cc'}}{\sum\sum\sum\sum} 1 + M^{-2013}.$$

The factor $CP/\sqrt{N_0 R}$ comes from the size of the weight function (see (13)).

We have reduced the problem to counting the number of solutions of the above congruence. This we can estimate quite easily. Let $d = (c, c')$. We write $c = de$ and $c' = de'$, with $(e, e') = 1$. The congruence condition now reduces to

$$\overline{(M - pr)}e'r - \overline{(M - pr')}er' \equiv 0 \bmod dee'.$$

The coprimality $(e, e') = 1$ now forces $e|r$ and $e'|r'$. Accordingly we write $r = es$ and $r' = e's'$. We are now left with the congruence condition

$$(M - pe's')s - (M - pes)s' \equiv 0 \bmod d.$$

We will first study the case where the equality

$$(M - pe's')s = (M - pes)s'$$

holds. This reduces to $M(s - s') = pss'(e' - e)$. Hence $p|s - s'$. By size consideration it now follows that $s - s' = 0$. Consequently we also have $e' = e$. So the equality forces $c = c'$ and $r = r'$. Hence the contribution from this equality to the count is given by

$$M^\varepsilon N_0 \frac{CP}{\sqrt{N_0 R}} CH.$$

Next we study the case where $(M - pe's')s \neq (M - pes)s'$. Here for any given vector (e, e', s, s'), we have $O(M^\varepsilon)$ many d. So it turns out that the

contribution of this part to the count is

$$M^\varepsilon N_0 \frac{CP}{\sqrt{N_0 R}} H^2.$$

Hence

$$\Psi_p \ll M^\varepsilon N_0 \frac{CP}{\sqrt{N_0 R}} H(C + H).$$

The lemma follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

4.4. *Conclusion.* Substituting the bound from Lemma 9 into Lemma 8, we see that the bound of Lemma 9 holds for $\mathcal{O}_1(m)$. The same bound in fact holds for $\mathcal{O}(m) - \mathcal{O}_1(m)$ as well. Substituting this bound in (10) we conclude that

$$\mathcal{O} \ll M^{-1/2+\varepsilon} \sum_{1 \leq m \ll \sqrt{N}} \sup \sqrt{\Lambda_m} \sqrt{RH(C + H)}$$

$$\ll M^{-1/2+\varepsilon} \sum_{1 \leq m \ll \sqrt{N}} \sqrt{\Lambda_m} \left(\frac{N^2 M}{mP^2} + \frac{NM^2}{P^2} \right)^{1/2} \ll \frac{N^{5/4} M^{1/2+\varepsilon}}{P}.$$

The last inequality follows by employing the Cauchy inequality and using (6). Note that to prove this bound, which is sufficient for the purpose of subconvexity, we required neither of the conditions (R) or (RS). However under (R) we have $\Lambda_m \ll M^\varepsilon N/m^2$, and consequently it follows that

$$\mathcal{O} \ll M^{-1/2+\varepsilon} \sum_{1 \leq m \ll \sqrt{N}} \sqrt{\Lambda_m} \left(\frac{N^2 M}{mP^2} + \frac{NM^2}{P^2} \right)^{1/2} \ll \frac{N^{3/2} M^\varepsilon}{P}.$$

Thus we have proved Proposition 1.

There are several ways in which the above estimate can be improved. For example we also have

$$\mathcal{O}_1^\star(m) \ll \frac{RM^\varepsilon}{\sqrt{M}} \sqrt{\Lambda_m} \sqrt{\Psi}$$

where

$$\Psi = \sum_{n=1}^\infty \left| \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \sum_{\substack{1 \leq c < \infty \\ |r| < H \\ (c, p(pr-M))=1}} \phi(p)\chi(c\bar{r})\, e\left(\frac{\overline{(M - pr)}rn}{c} \right) W_1^\star\left(\frac{n}{N_0}, \frac{rR}{cM}, \frac{c}{C} \right) \right|^2.$$

Now we can obtain a bound for Ψ exactly in the same fashion as Lemma 9. As the diagonal is now longer, but the modulus is same as before, we save more. The counting problem turns out to be slightly more involved. This yields the improved bound

$$\mathcal{O} \ll \sqrt{N} \frac{M^{5/4+\theta/2+\varepsilon}}{P}.$$

As our analysis of the dual term $\mathcal{F}$ is much weaker, this extra saving does not lead us to an improved exponent at the end.

5. **Functional equations**

In the rest of the paper we will analyze $\mathcal{F}$, which is given by (8). We will first take a smooth dyadic partition of unity to replace the weight function $V(r/N)$ by a bump function. To this end we apply Lemma 1 to get

$$\mathcal{F} = \frac{1}{P^\star} \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \sum_{\psi \bmod p} (1 - \psi(-1)) \sum_{f \in H_k(p, \psi)} \omega_f^{-1}$$

$$\times \sum_{m,n=1}^{\infty} \lambda(m,n) \lambda_f(n) W\left(\frac{nm^2}{N}\right) \sum_{(U,R)} \sum_{r=1}^{\infty} \overline{\lambda_f(r)} \chi(r) \, V\left(\frac{r}{N}\right) U\left(\frac{r}{R}\right).$$

The function

$$x \mapsto U(x)V(xR/N)$$

belongs to the class $\mathcal{W}_1(1)$. By abuse of notation we will again denote this function by $V(x)$. Moreover we only need to take R in the range $NM^{-4\theta} \ll R \ll N$. Next we apply summation formulas to the sums over (m,n) and r. The summation formulas will be derived from the respective functional equations. (For the sum over r, one may also use the GL(2) Voronoi summation formula directly.)

5.1. *Functional equation for $L(s, f \otimes \chi)$ and related summation formula.* Let

$$\Lambda(s, \bar{f} \otimes \chi) = \left(\frac{M\sqrt{p}}{2\pi}\right)^s \Gamma\left(s + \frac{k-1}{2}\right) L(s, \bar{f} \otimes \chi)$$

be the completed L-function associated with the twisted form $\bar{f} \otimes \chi$. Recall that $(M, p) = 1$. We have the following functional equation ([7, Chap. 14]).

LEMMA 10. *We have*

$$(15) \qquad \Lambda(s, \bar{f} \otimes \chi) = i^k \bar{\psi}(M) \chi(p) \, \frac{g_\chi^2 \bar{g}_\psi}{M\sqrt{p}} \, \lambda_f(p) \, \Lambda(1 - s, f \otimes \bar{\chi}),$$

where g_χ and g_ψ are the Gauss sums associated with χ and ψ respectively.

We will use this functional equation to derive a summation formula for the sum

$$S = \sum_{r=1}^{\infty} \lambda_{\bar{f}}(r) \chi(r) V\left(\frac{r}{R}\right),$$

where $V \in \mathcal{W}_1(1)$. Let $\mathcal{U} = \{(U, \tilde{R})\}$ be a smooth partition of unity as in Lemma 1.

LEMMA 11. *We have*

$$\sum_{r=1}^{\infty} \lambda_{\bar{f}}(r)\chi(r)V\left(\frac{r}{R}\right) = i^k\bar{\psi}(M)\chi(p)g_\chi^2\bar{g}_\psi\,\lambda_f(p)\,\frac{1}{2\pi}\sum_{\mathcal{U}^\star}\sum_{r=1}^{\infty}\frac{\lambda_f(r)\bar{\chi}(r)}{r}U\left(\frac{r}{\tilde{R}}\right)$$

$$\times \frac{1}{2\pi i}\int_{(0)}\tilde{V}(s)\left(\frac{4\pi^2 rR}{M^2 p}\right)^s \frac{\Gamma(1-s+\frac{k-1}{2})}{\Gamma(s+\frac{k-1}{2})}\mathrm{d}s + O(M^{-2013}),$$

where $\mathcal{U}^\star$ is the subset of $\mathcal{U}$ consisting of those pairs $(U, \tilde{R})$ that have $\tilde{R}$ in the range

$$\frac{M^{2-\varepsilon}P}{R} \ll \tilde{R} \ll \frac{M^{2+\varepsilon}P}{R}.$$

Proof. By Mellin inversion we get

$$S = \frac{1}{2\pi i}\int_{(2)}\tilde{V}(s)R^s L(s, \bar{f}\otimes\chi)\mathrm{d}s.$$

Using (15) we get

$$(16) \qquad S = i^k\bar{\psi}(M)\chi(p)\frac{g_\chi^2\bar{g}_\psi}{M\sqrt{p}}\lambda_f(p)\frac{M\sqrt{p}}{2\pi}$$

$$\times \frac{1}{2\pi i}\int_{(2)}\tilde{V}(s)\left(\frac{4\pi^2 R}{M^2 p}\right)^s \frac{\Gamma(1-s+\frac{k-1}{2})}{\Gamma(s+\frac{k-1}{2})}L(1-s, f\otimes\bar{\chi})\mathrm{d}s.$$

We move the contour to $-\varepsilon$, expand the L-function into a series and then use a smooth dyadic partition of unity $\mathcal{U}$, as above, to get

$$S = i^k\bar{\psi}(M)\chi(p)g_\chi^2\bar{g}_\psi\,\lambda_f(p)\,\frac{1}{2\pi}\sum_{\mathcal{U}}\sum_{r=1}^{\infty}\frac{\lambda_f(r)\bar{\chi}(r)}{r}U\left(\frac{r}{\tilde{R}}\right)$$

$$\times \frac{1}{2\pi i}\int_{(-\varepsilon)}\tilde{V}(s)\left(\frac{4\pi^2 rR}{M^2 p}\right)^s \frac{\Gamma(1-s+\frac{k-1}{2})}{\Gamma(s+\frac{k-1}{2})}\mathrm{d}s.$$

The poles of the integrand are located at

$$s = \frac{k+1}{2}+\ell, \quad \text{where } \ell = 0, 1, 2, \dots.$$

For $\tilde{R} \gg M^{2+\varepsilon}P/R$, we shift the contour to the left, and for $\tilde{R} \ll M^{2-\varepsilon}P/R$, we shift the contour to $k/2$. Since k is large (of the size ε^{-1}), we see that the contribution from the above ranges is negligibly small. $\square$

5.2. *Functional equation for $L(s, \pi\otimes f)$ and related summation formula.* Now we consider the Rankin-Selberg convolution $L(s, \pi\otimes f)$, which is given by the Dirichlet series

$$\sum_{m,n=1}^{\infty}\lambda(m,n)\lambda_f(n)(m^2 n)^{-s}$$

in the region of absolute convergences $\operatorname{Re}(s) > 1$. The L-function extends to an entire function. The completed L-function is given by

$$\Lambda(s, \pi \otimes f) = p^{3s/2} \gamma(s) \, L(s, \pi \otimes f),$$

where $\gamma(s)$ is a product of six gamma factors of the type $\Gamma((s + \kappa_j)/2)$. Also each κ_j satisfies $\operatorname{Re}(\kappa_j) > k/2 - 2$ (see [6]). We have the following functional equation.

LEMMA 12. *We have*

(17) $$\Lambda(s, \pi \otimes f) = \iota \left(\frac{\bar{g}_{\bar{\psi}}}{\sqrt{p}} \overline{\lambda_f(p)} \right)^3 \Lambda(1 - s, \tilde{\pi} \otimes \overline{f}),$$

where ι is a root of unity that depends only on the weight k and the Langlands parameters of π.

The (global) ε-factor in the above functional equation is given by the product of the local epsilon factors. For any finite prime q, let ψ_q be an unramified additive character of $\mathbb{Q}_q$. The local component π_q (of π) is an unramified principal series representation of $\operatorname{GL}(3, \mathbb{Q}_q)$ with trivial central character. In other words, $\pi_q = \operatorname{Ind}(\phi_1, \phi_2, \phi_3)$ with ϕ_i unramified character of $\mathbb{Q}_q^\times$ and $\phi_1 \phi_2 \phi_3 = 1$. So (see [8])

$$\begin{aligned}
\varepsilon_q(1/2, \pi \otimes f, \psi_q) &= \prod_{i=1}^{3} \varepsilon_q(1/2, \phi_i \times f, \psi_q) = \prod_{i=1}^{3} \{\phi_i(\star) \varepsilon_q(1/2, f, \psi_q)\} \\
&= \varepsilon_q(1/2, f, \psi_q)^3.
\end{aligned}$$

Here $\star = 1$ if $q \neq p$ and $\star = p$ otherwise. Hence the ε-factor for $L(s, \pi \otimes f)$, up to the archimedean component (which depends only on the weight of f and the Langlands parameters of π) turns out to be the cube of the ε-factor for $L(s, f)$. It is well known (see [7]) that the ε-factor for $L(s, f)$ is given by $\bar{g}_{\bar{\psi}} \bar{\lambda}_f(p)/\sqrt{p}$.

Consider the sum (which we will again temporarily denote by S)

$$S = \sum_{m,n=1}^{\infty} \lambda(m, n) \lambda_f(n) W \left(\frac{m^2 n}{N} \right),$$

with $W \in \mathcal{W}_1(1)$. We will prove the following summation formula.

LEMMA 13. *We have*

$$\sum_{m,n=1}^{\infty}\lambda(m,n)\lambda_f(n)W\left(\frac{m^2n}{N}\right)$$

$$= \iota\psi(-1)g_\psi^3\overline{\lambda_f(p)}^3\sum_{\mathcal{U}^\dagger}\sum_{m,n=1}^{\infty}\frac{\lambda(n,m)\overline{\lambda_f(n)}}{m^2n}U\left(\frac{m^2n}{\tilde{N}}\right)$$

$$\times\frac{1}{2\pi i}\int_{(0)}\tilde{W}(s)\left(\frac{m^2nN}{p^3}\right)^s\frac{\gamma(1-s)}{\gamma(s)}\mathrm{d}s + O(M^{-2013}),$$

where $\mathcal{U}^\dagger$ *is the subset of* $\mathcal{U}$ *consisting of those pairs* $(U,\tilde{N})$ *that have* $\tilde{N}$ *in the range*

$$\frac{P^3M^{-\varepsilon}}{N}\ll\tilde{N}\ll\frac{P^3M^\varepsilon}{N}.$$

Proof. By Mellin inversion we get

$$S = \frac{1}{2\pi i}\int_{(2)}\tilde{W}(s)N^sL(s,\pi\otimes f)\mathrm{d}s.$$

Using functional equation (17) we see that S is given by

$$\iota\psi(-1)\left(\frac{g_\psi}{\sqrt{p}}\right)^3\overline{\lambda_f(p)}^3 p^{3/2}\frac{1}{2\pi i}\int_{(2)}\tilde{W}(s)\left(\frac{N}{p^{3/2}}\right)^s\frac{\gamma(1-s)}{\gamma(s)}L(1-s,\tilde{\pi}\otimes\bar{f})\mathrm{d}s.$$

We move the contour to $-\varepsilon$, expand the *L*-function into a series and then use a partition of unity $\mathcal{U}$, as above, to get

$$\iota\psi(-1)g_\psi^3\overline{\lambda_f(p)}^3\sum_{\mathcal{U}}\sum_{m,n=1}^{\infty}\frac{\lambda(n,m)\overline{\lambda_f(n)}}{m^2n}U\left(\frac{m^2n}{\tilde{N}}\right)$$

$$\times\frac{1}{2\pi i}\int_{(-\varepsilon)}\tilde{W}(s)\left(\frac{m^2nN}{p^3}\right)^s\frac{\gamma(1-s)}{\gamma(s)}\mathrm{d}s.$$

As before, by moving contours we can show that for $\tilde{N}$ outside the range given in the statement of the lemma, the total contribution is negligible. □

5.3. *Application of Petersson formula.* We will conclude this section by proving the following lemma. (Recall our conventions regarding weights from Section 2.1.)

LEMMA 14. *Suppose* $0<\theta<1/24$ *and* $P>N^{1/2+\varepsilon}$. *Then we have*

(18) $$\mathcal{F}\ll M^\varepsilon\sup|\mathcal{O}_{\mathrm{red\,dual}}| + M^{3/2+\theta+\varepsilon}\sqrt{\frac{P}{N}},$$

where

(19)

$$\mathcal{O}_{\text{red dual}} = \frac{RN}{MP^5} \sum_{\substack{P<p<2P \\ p \text{ prime}}} \chi(p) \sum_{\psi \bmod p} (1 - \psi(-1))\, \bar{\psi}(-M) g_\psi^2$$

$$\times \sum_{m,n=1}^{\infty}\sum_{r=1}^{\infty} \bar{\chi}(r)\lambda(n,m) \sum_{\substack{c=1 \\ (c,p)=1}}^{\infty} \frac{S_\psi(np^2, r; cp)}{cp} J_{k-1}\left(\frac{4\pi\sqrt{nr}}{c}\right) W\left(\frac{nm^2}{\tilde{N}}\right) W\left(\frac{r}{\tilde{R}}\right),$$

and the supremum is taken over all R, $\tilde{R}$, $\tilde{N}$ *in the range*

(20) $\quad \dfrac{N}{M^{4\theta}} \ll R \ll N, \quad \dfrac{P^3}{NM^\varepsilon} \ll \tilde{N} \ll \dfrac{P^3 M^\varepsilon}{N} \quad and \quad \dfrac{M^2 P}{RM^\varepsilon} \ll \tilde{R} \ll \dfrac{M^{2+\varepsilon}P}{R}.$

Proof. We apply Lemmas 11 and 13 to (8). This reduces the analyzes of the sum in (8) to that of sums of the type

$$\frac{RN}{MP^5} \sum_{\substack{P<p<2P \\ p \text{ prime}}} \chi(p) \sum_{\psi \bmod p} (1 - \psi(-1))\, \bar{\psi}(-M) g_\psi^2 \sum_{f \in H_k(p,\psi)} \omega_f^{-1} \overline{\lambda_f(p)}^2$$

$$\times \sum_{m,n=1}^{\infty} \lambda(n,m)\overline{\lambda_f(n)} W\left(\frac{nm^2}{\tilde{N}}\right) \sum_{r=1}^{\infty} \bar{\chi}(r)\lambda_f(r) W\left(\frac{r}{\tilde{R}}\right),$$

where R, $\tilde{R}$ and $\tilde{N}$ are in the range (20). The leading factor accounts for the sizes of the denominators appearing on the right-hand side of the summation formulas in Lemmas 11 and 13 and also the sizes of the Gauss sums associated with χ and ψ.

We apply the Petersson formula. The diagonal term vanishes as the equality $r = np^2$ never holds in the above range, as

$$r \ll \tilde{R} \ll M^\varepsilon \frac{M^{2+4\theta}P}{N} \ll M^{1/2+5\theta+\varepsilon}P \ll P^2 M^{-\varepsilon} \ll p^2 M^{-\varepsilon}.$$

The fourth inequality follows from the condition on θ and P. The off-diagonal is given by

(21)

$$\mathcal{O}_{\text{dual}} = \frac{RN}{MP^5} \sum_{\substack{P<p<2P \\ p \text{ prime}}} \chi(p) \sum_{\psi \bmod p} (1 - \psi(-1))\, \bar{\psi}(-M) g_\psi^2$$

$$\times \sum_{m,n=1}^{\infty}\sum_{r=1}^{\infty} \bar{\chi}(r)\lambda(n,m) \sum_{c=1}^{\infty} \frac{S_\psi(np^2, r; cp)}{cp} J_{k-1}\left(\frac{4\pi\sqrt{nr}}{c}\right) W\left(\frac{nm^2}{\tilde{N}}\right) W\left(\frac{r}{\tilde{R}}\right).$$

Since the weight k is large, the contribution of the tail $c > \sqrt{\tilde{R}\tilde{N}_0} M^\varepsilon$ is negligible. Here we are setting $\tilde{N}_0 = \tilde{N}/m^2$. It follows that the terms where $p^2 | c$ make a negligible contribution.

Now let us consider the case where $p\|c$. We write $c = pc'$. In this case the Kloosterman sum splits as

$$S_\psi(np^2, r; cp) = S_\psi(0, \overline{c'}r; p^2) S(n, \bar{p}^2 r; c').$$

The first term on the right-hand side vanishes unless $p|r$, and accordingly we write $r = pr'$. It follows that

$$\sum_{\psi \bmod p} (1 - \psi(-1))\, \bar{\psi}(-M) g_\psi^2 S_\psi(0, \overline{c'}r; p^2)$$

$$= p^2 \sum_{\psi \bmod p} (1 - \psi(-1))\, \psi(\overline{Mc'}r') g_\psi$$

$$= p^2 \phi(p) \left\{ e\left(\frac{Mc'\overline{r'}}{p}\right) - e\left(-\frac{Mc'\overline{r'}}{p}\right) \right\}.$$

So the contribution of those c for which $p\|c$ is dominated by

$$\frac{RNM^\varepsilon}{MP^4} \sum_{\substack{P<p<2P \\ p \text{ prime}}} \sum \sum_{nm^2 \ll \tilde{N}} \sum_{r \ll \tilde{R}/P} |\lambda(n,m)| \sum_{c \ll \sqrt{\tilde{R}\tilde{N}_0} M^\varepsilon/P} \frac{1}{\sqrt{c}}.$$

Trivially estimating the remaining sums (using (R)), we get that the above sum is dominated by

$$O\left(\frac{RNM^\varepsilon}{MP^4} \frac{\sqrt{MP}}{(RN)^{1/4}} \tilde{R}\tilde{N}\right) = O\left(\frac{\sqrt{MP}}{(RN)^{1/4}} M^{1+\varepsilon}\right).$$

(One may avoid (R) by employing the Cauchy inequality and applying (6).) We conclude that

$$\mathcal{O}_{\text{dual}} = \mathcal{O}_{\text{red dual}} + O\left(M^{3/2+\theta+\varepsilon}\sqrt{\frac{P}{N}}\right),$$

where the reduced dual off-diagonal $\mathcal{O}_{\text{red dual}}$ is given by an expression similar to (21) but with the extra coprimality restriction $(c, p) = 1$.   □

Observe that we have used the Weil bound for the Kloosterman sum modulo c. One may avoid the application of the Weil bound by employing the Voronoi summation formula on the n-sum and then evaluating the remaining sums trivially.

6. Dual off-diagonal away from transition

It remains to study (19). We will take a smooth dyadic subdivision of the c-sum in $\mathcal{O}_{\text{red dual}}$. In this section we will show that the contribution of any such subdivision that is away from the transition range, which is marked by $C \sim \sqrt{\tilde{N}_0 \tilde{R}}$, is satisfactory. For larger values of C, the trivial estimation suffices as the size of the Bessel function is small due to the large weight k.

We will see that for smaller size of C, one can get away with a relatively easy estimate.

6.1. *Sum over ψ.* We fix $C, m \geq 1$ and consider sums of the type

(22)
$$
\mathcal{O}_{\pm}(C,m) = \frac{RN}{MP^5} \sum_{\substack{P<p<2P \\ p \text{ prime}}} \phi(p)\chi(p) \sum_{n=1}^{\infty} \sum_{\substack{r=1 \\ (r,p)=1}}^{\infty} \bar{\chi}(r)\lambda(n,m)
$$
$$
\times \sum_{\substack{c=1 \\ (c,p)=1}}^{\infty} \frac{S(n,r;c)}{c} e\left(\pm \frac{Mc\bar{r}}{p}\right) J_{k-1}\left(\frac{4\pi\sqrt{nr}}{c}\right) W\left(\frac{nm^2}{\tilde{N}}\right) W\left(\frac{r}{\tilde{R}}\right) W\left(\frac{c}{C}\right).
$$

For notational simplicity we are using the same notation W for the new smooth weight function on the c sum.

LEMMA 15. *We have*

$$
\mathcal{O}_{\text{red dual}} \ll M^{\varepsilon} \sum_{m=1}^{10\sqrt{\tilde{N}}} \sup |\mathcal{O}_{\pm}(C,m)| + M^{-2013},
$$

where the supremum is taken over all $C \ll M^{\varepsilon}\sqrt{\tilde{N}_0 \tilde{R}}$.

Proof. The Kloosterman sum in (19) factorizes as

$$
S_{\psi}(np^2, r; cp) = S_{\psi}(0, \bar{c}r; p)S(\bar{p}np^2, \bar{p}r; c) = S_{\psi}(0, \bar{c}r; p)S(n, r; c).
$$

Moreover we have

$$
\sum_{\psi \bmod p} (1 - \psi(-1))\, \bar{\psi}(-M)g_{\psi}^2 S_{\psi}(0, \bar{c}r; p) = p \sum_{\psi \bmod p} (1 - \psi(-1))\, \psi(\overline{Mc}r)g_{\psi}
$$
$$
= p\phi(p)\left\{ e\left(\frac{Mc\bar{r}}{p}\right) - e\left(-\frac{Mc\bar{r}}{p}\right)\right\}.
$$

(The sum vanishes unless $(r,p) = 1$.) The lemma follows by taking a smooth dyadic subdivision of the c sum.   □

Observe the curious separation of the variables n and p, which is a consequence of the fact that we are studying a $\mathrm{GL}(d_1) \times \mathrm{GL}(d_2)$ Rankin-Selberg convolution with $d_1 - d_2 = 2$. (In our case $d_1 = 3$ and $d_2 = 1$.) This inbuilt separation of variables will play an important structural role in our analysis of $\mathcal{O}_{\text{dual}}$.

We set $\mathcal{O}(C,m) = \mathcal{O}_{+}(C,m)$, and in the rest of the paper we will only deal with this sum. The other sum with $-$ sign behaves exactly in the same

fashion. Taking absolute values we get

$$|\mathcal{O}(C,m)| \ll \frac{RN}{CMP^5} \sum_{r \in \mathbb{Z}} \sum_{C < c \le 2C} W\left(\frac{r}{\tilde{R}}\right) \Bigg| \sum_{\substack{P < p < 2P \\ p \text{ prime} \\ (cr,p)=1}} \phi(p)\chi(p) e\left(\frac{Mc\bar{r}}{p}\right) \Bigg|$$

$$\times \Bigg| \sum_{n=1}^{\infty} \lambda(n,m) S(n,r;c) J_{k-1}\left(\frac{4\pi\sqrt{nr}}{c}\right) W\left(\frac{nm^2}{\tilde{N}}\right) \Bigg|.$$

This is the point where we use the separation of the variables noted above. Now applying the Cauchy inequality (and exploiting positivity) we get

$$(23) \qquad \mathcal{O}(C,m) \ll \frac{RN}{CMP^5} \sqrt{\Theta_1} \sqrt{\Theta_2},$$

where

$$(24) \qquad \Theta_1 = \sum_{c,r \in \mathbb{Z}} U\left(\frac{c}{C}, \frac{r}{\tilde{R}}\right) \Bigg| \sum_{\substack{P < p < 2P \\ p \text{ prime} \\ (cr,p)=1}} \phi(p)\chi(p) e\left(\frac{Mc\bar{r}}{p}\right) \Bigg|^2$$

and

$$(25) \qquad \Theta_2 = \sum_{r \in \mathbb{Z}} \sum_{C < c \le 2C} W\left(\frac{r}{\tilde{R}}\right) \Bigg| \sum_{1 \le n < 2\tilde{N}} \alpha(n) S(n,r;c) J_{k-1}\left(\frac{4\pi\sqrt{nr}}{c}\right) \Bigg|^2.$$

Here

$$\alpha(n) = \lambda(n,m) W\left(\frac{nm^2}{\tilde{N}}\right),$$

and U is a suitable compactly supported weight function on $(0,\infty)^2$.

6.2. *Bound for* Θ_1. We will consider a slightly general sum

$$\Theta_1^\star = \sum_{c,r \in \mathbb{Z}} U\left(\frac{c}{C}, \frac{r}{\tilde{R}}\right) \Bigg| \sum_{\substack{P < p < 2P \\ p \text{ prime} \\ (cr,p)=1}} \beta(p) e\left(\frac{Mc\bar{r}}{p}\right) \Bigg|^2,$$

with $|\beta(p)| \le p$. Taking $\beta(p) = \phi(p)\chi(p)$, the above sum $\Theta_1^\star$ reduces to Θ_1.

LEMMA 16. *Suppose* $P > \sqrt{N}$. *Then we have*

$$(26) \qquad \Theta_1^\star \ll P^3 \tilde{R}(C + P) M^\varepsilon.$$

Proof. Opening the absolute square we arrive at

$$\sum_{\substack{P < p,p' < 2P \\ p,p' \text{ prime}}} \beta(p)\overline{\beta(p')} \sum_{\substack{c,r \in \mathbb{Z} \\ (cr,pp')=1}} e\left(\frac{Mc\bar{r}}{p} - \frac{Mc\bar{r}}{p'}\right) U\left(\frac{c}{C}, \frac{r}{\tilde{R}}\right).$$

The diagonal $p = p'$ contribution is dominated by $P^3 C\tilde{R}$. Also the coprimality condition $(c, pp') = 1$ can be removed at a cost of an error term of size $P^3 C\tilde{R}$,

which is dominated by the diagonal contribution. We will now apply the Poisson summation formula on the off-diagonal. Breaking into congruence classes modulo pp', we arrive at

$$\sum_{\substack{P<p,p'<2P \\ p\neq p' \text{ prime}}} \beta(p)\overline{\beta(p')} \sum_{\substack{\gamma,\rho \bmod pp' \\ (\rho,pp')=1}} e\left(\frac{M\gamma\overline{\rho}}{p} - \frac{M\gamma\overline{\rho}}{p'}\right)$$
$$\times \sum_{c,r\in\mathbb{Z}} U\left(\frac{\gamma+cpp'}{C}, \frac{\rho+rpp'}{\tilde{R}}\right).$$

Then by the Poisson summation (and standard rescaling) we get

$$C\tilde{R} \sum_{\substack{P<p,p'<2P \\ p\neq p' \text{ prime}}} \frac{\beta(p)\overline{\beta(p')}}{(pp')^2} \sum_{c,r\in\mathbb{Z}} \sum_{\substack{\gamma,\rho \bmod pp' \\ (\rho,pp')=1}} e\left(\frac{M\gamma\overline{\rho}}{p} - \frac{M\gamma\overline{\rho}}{p'} + \frac{c\gamma+r\rho}{pp'}\right)$$
$$\times \int_{\mathbb{R}^2} U(x,y)\, e\left(-\frac{Cc}{pp'}x - \frac{\tilde{R}r}{pp'}y\right) \mathrm{d}x\mathrm{d}y.$$

The complete character sum over γ now yields the relation

$$M\overline{\rho}(p'-p) + c \equiv 0 \bmod pp'.$$

Hence the above sum reduces to

(27) $$C\tilde{R} \sum_{\substack{P<p,p'<2P \\ p\neq p' \text{ prime}}} \frac{\beta(p)\overline{\beta(p')}}{pp'} \sum_{\substack{c,r\in\mathbb{Z} \\ (c,pp')=1}} e\left(-\frac{\overline{c}rM(p'-p)}{pp'}\right)$$
$$\times \int_{\mathbb{R}^2} U(x,y)\, e\left(-\frac{Cc}{pp'}x - \frac{\tilde{R}r}{pp'}y\right) \mathrm{d}x\mathrm{d}y.$$

The integral is negligibly small if $|r| \gg P^2 M^\varepsilon/\tilde{R}$ or if $|c| \gg P^2 M^\varepsilon/C$.

Let $V(x)$ be a smooth bump function with support contained in $[-10, 10]$ and such that $V^{(j)} \ll_j 1$. Set $R^\star = P^2 M^\varepsilon/\tilde{R}$, take $1 \leq c \ll P^2 M^\varepsilon/C$ with $(c,pp')=1$ and consider the sum

$$\sum_{r\in\mathbb{Z}} e\left(-\frac{\overline{c}rM(p'-p)}{pp'}\right) e\left(-\frac{\tilde{R}r}{pp'}y\right) V\left(\frac{r}{R^\star}\right).$$

Here y is a fixed positive number. (The negative values of c are treated in the same fashion.) Applying reciprocity we reduce the above sum to

$$\sum_{r\in\mathbb{Z}} e\left(\frac{\overline{p}\overline{p}'rM(p'-p)}{c}\right) e\left(-\frac{rM(p'-p)}{cpp'} - \frac{\tilde{R}r}{pp'}y\right) V\left(\frac{r}{R^\star}\right).$$

We break the sum into congruence classes modulo c and then apply the Poisson summation formula. This gives (after standard rescaling)

$$\frac{R^\star}{c} \sum_{r \in \mathbb{Z}} \sum_{\rho \bmod c} e\left(\frac{\bar{p}\bar{p}' M(p' - p)\rho + r\rho}{c}\right)$$

$$\times \int_{\mathbb{R}} V(z) e\left(-\frac{R^\star M(p' - p)}{cpp'} z - \frac{\tilde{R}R^\star y}{pp'} z\right) e\left(-\frac{R^\star r}{c} z\right) \mathrm{d}z,$$

which reduces to

$$R^\star \sum_{\substack{r \in \mathbb{Z} \\ c|rpp'+M(p'-p)}} \int_{\mathbb{R}} V(z) e\left(-\frac{R^\star M(p' - p)}{cpp'} z - \frac{\tilde{R}R^\star y}{pp'} z\right) e\left(-\frac{R^\star r}{c} z\right) \mathrm{d}z.$$

By repeated integration by parts we see that the integral is bounded by

$$\ll_j \left(\left(1 + \frac{MP}{c\tilde{R}}\right) \frac{cM^\varepsilon}{R^\star |r|}\right)^j$$

$$\ll_j \left(\left(\frac{c}{R^\star} + \frac{MP}{\tilde{R}R^\star}\right) \frac{M^\varepsilon}{|r|}\right)^j \ll_j \left(\left(\frac{\tilde{R}}{C} + \frac{M}{P}\right) \frac{M^\varepsilon}{|r|}\right)^j.$$

Hence the integral is negligibly small if

$$|r| \gg M^\varepsilon \left(\frac{\tilde{R}}{C} + \frac{M}{P}\right).$$

It follows that (27) is dominated by

$$C\tilde{R}R^\star \sum_{\substack{P<p,p'<2P \\ p \neq p' \text{ prime}}} \sum_{|c| \ll P^2 M^\varepsilon/C} \sum_{\substack{|r| \ll M^\varepsilon(\frac{\tilde{R}}{C}+\frac{M}{P}) \\ c|rpp'+M(p'-p)}} 1 + M^{-2013}.$$

Since $rpp' + M(p' - p)$ never vanishes, the sum is seen to be bounded by

$$M^{1+\varepsilon}CP\tilde{R}R^\star + M^\varepsilon P^2 \tilde{R}^2 R^\star \ll M^{1+\varepsilon}CP^3 + M^\varepsilon P^4 \tilde{R}.$$

Since $P > \sqrt{N}$ and θ is sufficiently small, we find that $\tilde{R} > M$. So the first term is dominated by the diagonal contribution. The lemma follows. $\qquad \square$

6.3. *Bound for* Θ_2. In this subsection we prove the following bound for Θ_2. To this end we will use the Ramanujan conjecture (R), which implies that $|\alpha(n)| \ll M^\varepsilon$.

LEMMA 17. *For $P < M^{1-\varepsilon}$ and θ sufficiently small, we have*

$$\Theta_2 \ll C^2 \tilde{R} \tilde{N}_0 M^\varepsilon.$$

Proof. Opening the absolute square in the sum (25) we arrive at

$$
(28) \qquad \Theta_2 = \sum_{C < c \le 2C} \sum_{1 \le n, n' < 2\tilde{N}_0} \sum \alpha(n)\bar{\alpha}(n') \sum_{r \in \mathbb{Z}} S(n, r; c)S(n', r; c)
$$

$$
\times J_{k-1}\left(\frac{4\pi\sqrt{nr}}{c}\right) J_{k-1}\left(\frac{4\pi\sqrt{n'r}}{c}\right) W\left(\frac{r}{\tilde{R}}\right).
$$

We only need to consider the case where $C \ll M^\varepsilon \sqrt{\tilde{R}\tilde{N}}/m$, as the Bessel function is negligibly small otherwise due to the large weight. For C in this range, we apply the Poisson summation formula on r with modulus c. Now the Fourier transform

$$
\int_{\mathbb{R}} J_{k-1}\left(\frac{4\pi\sqrt{n\tilde{R}}}{c}x\right) J_{k-1}\left(\frac{4\pi\sqrt{n'\tilde{R}}}{c}x\right) W(x) e\left(-\frac{\tilde{R}r}{c}x\right) dx
$$

is bounded by

$$
\ll_j \left[\left(1 + \frac{\sqrt{\tilde{R}\tilde{N}}}{mC}\right)\frac{C}{\tilde{R}r}\right]^j
$$

by repeated integration by parts j times. Since $C \ll M^\varepsilon \sqrt{\tilde{R}\tilde{N}}/m$, it follows that the integral is negligibly small if

$$
|r| \gg \sqrt{\frac{\tilde{N}}{\tilde{R}}}\,\frac{M^\varepsilon}{m}.
$$

Since we are going to choose $P < M^{1-\varepsilon}$, we have $\tilde{R} \gg \tilde{N}M^\varepsilon$, and hence the nonzero frequencies $r \ne 0$ make a negligible contribution. The main contribution comes from the zero frequency that is given by

$$
\tilde{R} \sum_{C < c \le 2C} \frac{1}{c} \sum_{1 \le n, n' < 2\tilde{N}_0} \sum \alpha(n)\bar{\alpha}(n') \sum_{a \bmod c} S(n, a; c)S(n', a; c)
$$

$$
\times \int_{\mathbb{R}} J_{k-1}\left(\frac{4\pi\sqrt{n\tilde{R}x}}{c}\right) J_{k-1}\left(\frac{4\pi\sqrt{n'\tilde{R}x}}{c}\right) W(x)\, dx.
$$

The integral is bounded by

$$
\int_{\mathbb{R}} J_{k-1}\left(\frac{4\pi\sqrt{n\tilde{R}x}}{c}\right) J_{k-1}\left(\frac{4\pi\sqrt{n'\tilde{R}x}}{c}\right) W(x)\, dx \ll \frac{C}{\sqrt{\tilde{R}}(nn')^{1/4}}.
$$

The character sum is given by

$$
\sum_{a \bmod c} S(n, a; c)S(n', a; c) = c\,\mathfrak{c}_c(n - n'),
$$

where $\mathfrak{c}_u(v)$ is the Ramanujan sum modulo u. We obtain the bound

$$
\Theta_2 \ll C\sqrt{\tilde{R}} \sum_{C < c \le 2C} \sum_{1 \le n, n' < 2\tilde{N}_0} \sum \frac{|\alpha(n)||\alpha(n')|}{(nn')^{1/4}}|\mathfrak{c}_c(n - n')|.
$$

The Ramanujan sum can be bounded by the greatest common divisor, i.e., $\mathfrak{c}_c(n - n') \ll (c, n - n')$. Consequently

$$\sum_{c \sim C} |\mathfrak{c}_c(n - n')| \ll \begin{cases} CM^\varepsilon & \text{if } n \neq n', \\ C^2 M^\varepsilon & \text{otherwise.} \end{cases}$$

So it follows (using (R)) that

$$\Theta_2 \ll C\sqrt{\tilde{R}}M^\varepsilon \left\{ C^2 \sum_{n \sim \tilde{N}/m^2} \frac{1}{n^{1/2}} + C \left(\sum_{n \sim \tilde{N}/m^2} \frac{1}{n^{1/4}} \right)^2 \right\}$$

$$\ll C^2 \sqrt{\tilde{R}\tilde{N}_0} \left\{ C + \tilde{N}_0 \right\} M^\varepsilon.$$

Since by our choice $P < M$, it follows that $C + \tilde{N}_0 \ll \sqrt{\tilde{R}\tilde{N}_0}M^\varepsilon$, and we conclude the lemma. $\qquad\square$

6.4. *Estimate for $\mathcal{O}(C, m)$ for C away from transition range.* Recall that we have already noted that the Bessel function in (22) is negligibly small, because of the large weight k, if $C \gg \sqrt{\tilde{R}\tilde{N}_0}M^\varepsilon$. So we need to analyze, for any given m, the contribution of C in the range $C \ll \sqrt{\tilde{R}\tilde{N}_0}M^\varepsilon$.

LEMMA 18. *For $N^{1/2+\varepsilon} < P < M^{1-\varepsilon}$ and θ sufficiently small, we have*

$$\sum_{m > M^{4\theta}} \sup |\mathcal{O}(C, m)| \ll \sqrt{N}M^{3/4 - \theta/2 + \varepsilon},$$

where the supremum is taken over all $C \ll M^\varepsilon \sqrt{\tilde{R}\tilde{N}_0}$. Also we have

$$\sum_{m=1}^{10\sqrt{\tilde{N}}} \sup_{C < P^2/M^{1/2+\theta}} |\mathcal{O}(C, m)| \ll \sqrt{N}M^{3/4 - \theta/2 + \varepsilon}.$$

Proof. Plugging the bounds for Θ_i from Lemmas 16 and 17 into (23), we conclude
(29)
$$\mathcal{O}(C, m) \ll \frac{RNM^\varepsilon}{CMP^5} \sqrt{P^3(C + P)\tilde{R}} \sqrt{C^2 \tilde{R}\tilde{N}_0} \ll M^\varepsilon \left(\frac{M\sqrt{NC}}{mP} + \frac{M\sqrt{N}}{m\sqrt{P}} \right).$$

In the range $C \ll M^\varepsilon \sqrt{\tilde{R}\tilde{N}_0}$, we get

$$\frac{M\sqrt{NC}}{mP} \ll M^\varepsilon \frac{M\sqrt{N}}{mP} (\tilde{R}\tilde{N}_0)^{1/4} \ll \sqrt{N} \frac{M^{3/4 + 3\theta/2 + \varepsilon}}{m^{3/2}}.$$

Summing over m we now conclude the first statement. The second statement also follows from (29). The lemma follows. (To manage the last term in (29) one only needs $\theta < 1/6$.) $\qquad\square$

Substituting the bound from Lemma 18 into Lemma 15, we derive the following corollary.

COROLLARY 1. *For $N^{1/2+\varepsilon} < P < M^{1-\varepsilon}$ and θ sufficiently small, we have*

$$\mathcal{O}_{\text{red dual}} \ll M^{\varepsilon} \sum_{m \leq M^{4\theta}} \sup |\mathcal{O}(C, m)| + \sqrt{N} M^{3/4 - \theta/2 + \varepsilon},$$

where the supremum is taken over all C in the range

(30) $$\frac{P^2}{M^{1/2}} \frac{1}{M^{\theta}} < C < \frac{P^2 M^{1+\varepsilon}}{m\sqrt{NR}} \ll \frac{P^2}{mM^{1/2}} M^{3\theta + \varepsilon}.$$

Later we will be applying Poisson summation on the sum over c. To this end we wish to get rid of the coprimality condition $(c, p) = 1$ in (22). Consider the sum in (22) but with the condition $p | c$ in place of $(c, p) = 1$, i.e.,

(31) $$\mathcal{O}^{\dagger}(C, m) = \frac{RN}{MP^5} \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \phi(p) \chi(p) \sum_{n=1}^{\infty} \sum_{\substack{r=1 \\ (p,r)=1}}^{\infty} \bar{\chi}(r) \lambda(n, m)$$

$$\times \sum_{c=1}^{\infty} \frac{S(n, r; cp)}{cp} J_{k-1}\left(\frac{4\pi\sqrt{nr}}{cp}\right) W\left(\frac{nm^2}{\tilde{N}}\right) W\left(\frac{r}{\tilde{R}}\right) W\left(\frac{cp}{C}\right).$$

We set

$$\mathcal{O}^{\star}(C, m) = \mathcal{O}(C, m) + \mathcal{O}^{\dagger}(C, m),$$

which is exactly the sum in (22) (for a $+$ sign) without the coprimality $(c, p) = 1$ condition.

COROLLARY 2. *For $N^{1/2+\varepsilon} < P < M^{1-\varepsilon}$ and θ sufficiently small, we have*

$$\mathcal{O}_{\text{red dual}} \ll M^{\varepsilon} \sum_{m \leq M^{4\theta}} \sup |\mathcal{O}^{\star}(C, m)| + \sqrt{N} M^{3/4 - \theta/2 + \varepsilon},$$

where the supremum is taken over all C in the range (30).

Proof. We need to show that the term $\mathcal{O}^{\dagger}$ can be absorbed in the error term in Corollary 1. Taking absolute value we get

$$\mathcal{O}^{\dagger}(C, m) \leq \frac{RN}{CMP^4} \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \sum_{r \sim \tilde{R}} \sum_{c \sim C/p}$$

$$\times \left| \sum_{n=1}^{\infty} \lambda(n, m) S(n, r; cp) J_{k-1}\left(\frac{4\pi\sqrt{nr}}{cp}\right) W\left(\frac{nm^2}{\tilde{N}}\right) \right|.$$

Using positivity we glue c and p to arrive at

$$\mathcal{O}^\dagger(C,m) \ll M^\varepsilon \frac{RN}{CMP^4} \sum_{r\sim\tilde{R}} \sum_{C<c\le 4C}$$
$$\times \left| \sum_{n=1}^\infty \lambda(n,m)S(n,r;c)J_{k-1}\left(\frac{4\pi\sqrt{nr}}{c}\right)W\left(\frac{nm^2}{\tilde{N}}\right)\right|.$$

Applying Cauchy inequality we get

$$\mathcal{O}^\dagger(C,m) \ll M^\varepsilon \frac{RN}{CMP^4}\sqrt{C\tilde{R}}\sqrt{\Theta_2} \ll M^\varepsilon \frac{RN}{CMP^5}\sqrt{P^2C\tilde{R}}\sqrt{\Theta_2}.$$

This can be absorbed in the bound given in Lemma 18 or in Corollary 1. $\quad\square$

7. Wild dual off-diagonal in transition

In the rest of the paper we will analyze the contribution of those C that lie in the range (30) for any given $m \le M^{4\theta}$. The Voronoi summation formula now comes into play, and we will be using it more than once. As such the notation starts to become a little messy. At this point we abandon keeping track of the exponent and use the notation

$$M^\star \quad \text{to mean a power of} \quad M^\theta$$

(e.g., $m \ll M^\star$), which will vary from one occurrence to other.

We consider the sum

$$(32) \qquad \mathcal{O}^\star(C,m) = \frac{RN}{CMP^5} \sum_{\substack{P<p<2P \\ p \text{ prime}}} \phi(p)\chi(p) \sum_{\substack{c,r=1 \\ (p,r)=1}}^\infty \bar{\chi}(r)$$
$$\times \sum_{n=1}^\infty \lambda(n,m)S(n,r;c)e\left(\frac{Mc\bar{r}}{p}\right)W\left(\frac{c}{C},\frac{n}{\tilde{N}_0},\frac{r}{\tilde{R}}\right),$$

where

$$W(x,y,z) = J_{k-1}\left(\frac{4\pi\sqrt{\tilde{N}_0\tilde{R}yz}}{Cx}\right)x^{-1}W(x)W(y)W(z).$$

The single variable function W on the right-hand side is as given in (22). In particular, $W(x,y,z) \in \mathcal{W}_3(M^{\hat{\theta}})$ where $\hat{\theta} = 4\theta$. Moreover the weight $W(x,y,z)$ is independent of p.

7.1. *Voronoi summation formula.* The next step involves an application of the Voronoi summation formula (see [11, Th. 1.18] or [9, Prop. 2.1]) on the sum over n. Let

$$\tilde{W}(x,s,z) = \int_0^\infty W(x,y,z)\,y^{s-1}\mathrm{d}y,$$

and for $\ell = 0, 1$, define

$$(33) \qquad \gamma_\ell(s) = \frac{1}{2\pi^{3(s+\frac{1}{2})}} \prod_{i=1}^{3} \frac{\Gamma\left(\frac{1+s+\alpha_i+\ell}{2}\right)}{\Gamma\left(\frac{-s-\alpha_i+\ell}{2}\right)}$$

and set $\gamma_\pm(s) = \gamma_0(s) \mp i\gamma_1(s)$. We define the integral transforms

$$(34) \qquad W_\pm^\star(x, y, z) = \frac{1}{2\pi i} \int_{(\sigma)} y^{-s} \gamma_\pm(s) \tilde{W}(x, -s, z) \mathrm{d}s,$$

where $\sigma > -1 + \max\{-\mathrm{Re}(\alpha_1), -\mathrm{Re}(\alpha_2), -\mathrm{Re}(\alpha_3)\}$. The following lemma gives the Voronoi summation formula.

LEMMA 19. *For W and $W_\pm^\star$ as above, we have*

$$\sum_{n=1}^{\infty} \lambda(m, n) \, e\left(\frac{\alpha n}{c}\right) W\left(\frac{c}{C}, \frac{n}{\tilde{N}_0}, \frac{r}{\tilde{R}}\right)$$

$$= c \sum_{\pm} \sum_{m'|cm} \sum_{n=1}^{\infty} \frac{\lambda(n, m')}{m'n} S(m\bar{\alpha}, \pm n; mc/m') W_\pm^\star\left(\frac{c}{C}, \frac{m'^2 n \tilde{N}_0}{c^3 m}, \frac{r}{\tilde{R}}\right).$$

Since the function $W(x, y, z)$ is smooth and supported in $[1, 2]^3$, the Mellin transform $\tilde{W}(x, s, z)$ is entire in the s variable. Using the bounds for the derivatives of W and using integration by parts, we get

$$\tilde{W}(x, s, z) \ll_j \frac{M^{\hat{\theta}j}}{|s(s+1)\dots(s+j-1)|}.$$

We can now obtain a bound for the integral transform in (34) by shifting the contour to the right and using the Stirling approximation. It follows that $W_\pm^\star(x, y, z)$ is negligibly small if $y \gg M^{3\hat{\theta}+\varepsilon}$. For $0 < y \ll M^{3\hat{\theta}+\varepsilon}$ we shift the contour to the left up to $\sigma = -1 + \varepsilon$. Since we are assuming (RS), there are no poles of the gamma factor in this domain. Differentiating under the integral sign we get

$$(35) \qquad y^j \frac{\partial^j}{\partial y^j} W_\pm^\star(x, y, z) \ll M^{\hat{\theta}(j-1/2)+\varepsilon} \, y$$

for $j \geq 1$. Also $W_\pm^\star(x, y, z) \ll M^\varepsilon y$. It is crucial to assume the (RS) to get tight bounds for $W_\pm^\star(x, y, z)$ for small values of y. This plays an important role in our derivation of Lemma 23 below. (The $-1/2$ in (35) comes from applying the Stirling's approximation to the gamma functions in (33) at the line $\sigma = -1 + \varepsilon$. The size of the gamma functions partly compensates the loss of powers of s when we differentiate (34) under the integral sign.)

Next we will apply the Voronoi summation to establish the following.

LEMMA 20. *We have*

$$\mathcal{O}^\star(C, m) \ll M^\varepsilon \sup |\mathcal{O}(C, m; L, \mathfrak{m}', \mathfrak{d})|,$$

where

(36)

$$\mathcal{O}(C, m; L, \mathfrak{m}', \mathfrak{d}) = \frac{RN}{CMP^5} \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \phi(p)\chi(p) \sum_{\substack{r=1 \\ (p,r)=1}}^{\infty} \bar{\chi}(r) \sum_{c=1}^{\infty} e\left(\frac{Mc\bar{r}}{p}\right)$$

$$\times \frac{\tilde{N}_0}{c} \sum_{\substack{m'|cm \\ m' \sim \mathfrak{m}'}} \frac{m'}{m} \sum_{\substack{d|c \\ d \sim \mathfrak{d}}} \frac{\mu(d)}{d} \sideset{}{^\star}\sum_{\substack{\beta \bmod mc/m' \\ r+\beta m' \equiv 0 \bmod c/d}} \sum_{n=1}^{\infty} \lambda(m',n) e\left(\frac{\bar{\beta}n}{mc/m'}\right) V\left(\frac{c}{C}, \frac{n}{L}, \frac{r}{\tilde{R}}\right).$$

Here the smooth weight function V belongs to the class $\mathcal{W}_3(M^{\hat{\theta}})$. The supremum is taken over all triplets $(L, \mathfrak{m}', \mathfrak{d})$ satisfying

(37)
$$1 \le L \ll \frac{C^3 M^\star}{\mathfrak{m}'^2 \tilde{N}},$$

$\mathfrak{m}' \ll Cm$ *and* $\mathfrak{d} \ll C$.

Proof. We apply Lemma 19 to the sum over n in (32) after opening the Kloosterman sum. (More precisely, one applies the Voronoi summation formula for the dual form $\tilde{\pi}$. The (m,n)-th Fourier coefficient of $\tilde{\pi}$ is given by $\lambda_{\tilde{\pi}}(m,n) = \lambda(n,m)$. The Langlands parameters are given by $(-\alpha_3, -\alpha_2, -\alpha_1)$.) As we observed above, the tail $m'^2 n \gg C^3 m^3 M^{3\hat{\theta}+\varepsilon}/\tilde{N}$ makes a negligible contribution as the integral transform is negligibly small. For smaller values of $m'^2 n$, we take a smooth dyadic subdivision of the n-sum, and a dyadic subdivision of the sum over m' to arrive at (consider only the term with a + sign)

(38)
$$\frac{\tilde{N}_0}{c^2} \sum_{\substack{m'|cm \\ m' \sim \mathfrak{m}'}} \frac{m'}{m} \sum_{n=1}^{\infty} \lambda(m',n) S(m\bar{\alpha}, n; mc/m') V\left(\frac{c}{C}, \frac{n}{L}, \frac{r}{\tilde{R}}\right).$$

Here L needs to be taken in the range

$$1 \le L \ll L^\star = \frac{C^3 m^3 M^{3\hat{\theta}+\varepsilon}}{\mathfrak{m}'^2 \tilde{N}},$$

which in our short-hand notation boils down to (37). Also $V(x,y,z)$ is of the shape

$$(yL/L^\star)^{-1} W_+^\star(x, yL/L^\star, z) U(y),$$

where $U \in \mathcal{W}_1(1)$ comes from the partition of unity. From (35) it follows that $V(x,y,z) \in \mathcal{W}_3(M^{\hat{\theta}})$.

We have applied the Voronoi summation after opening the Kloosterman sum in the initial expression (32). So we eventually get the Fourier transform

of the Kloosterman sum in (38), which is given by

$$\sideset{}{^\star}\sum_{\alpha \bmod c} e\left(\frac{\bar{\alpha}r}{c}\right) S(m\bar{\alpha}, n; mc/m')$$

$$= \sideset{}{^\star}\sum_{\beta \bmod mc/m'} e\left(\frac{\bar{\beta}n}{mc/m'}\right) \sideset{}{^\star}\sum_{\alpha \bmod c} e\left(\frac{\bar{\alpha}(r+\beta m')}{c}\right).$$

The last sum is a Ramanujan sum. Substituting explicit formula for this sum we obtain

$$c\sum_{d|c} \frac{\mu(d)}{d} \sideset{}{^\star}\sum_{\substack{\beta \bmod mc/m' \\ r+\beta m' \equiv 0 \bmod c/d}} e\left(\frac{\bar{\beta}n}{mc/m'}\right).$$

The lemma follows. $\qquad\qquad\square$

Note that the function V, which appears in Lemma 20, involves the (latent) variables m and m' but does not depend on p.

7.2. *Repeating Voronoi summation.* In the rest of this section we will obtain a bound for (36), which will be satisfactory for our purpose when either $\mathfrak{m}'$ or $\mathfrak{d}$ is suitably large. We call these terms 'wild.' Consider the expression in the second line of (36). Suppose we again apply the Voronoi summation formula on the sum over n. (This is the standard reversal process to get rid of the 'wild' terms.) Then we arrive at

$$(39) \qquad \tilde{N}_0 \sum_{\substack{d|c \\ d\sim\mathfrak{d}}} \frac{\mu(d)}{d} \sum_{\pm} \sum_{\substack{m',m''|cm \\ m'\sim\mathfrak{m}'}} \sum_{n=1}^{\infty} \frac{\lambda(n, m'')}{m''n}$$

$$\times \sideset{}{^\star}\sum_{\substack{\beta \bmod mc/m' \\ r+\beta m' \equiv 0 \bmod c/d}} S(m'\beta, \pm n; mc/m'') \, V_{\pm}^{\star}\left(\frac{c}{C}, \frac{m'^2 m''^2 nL}{m^3 c^3}, \frac{r}{\tilde{R}}\right).$$

The integral transform is negligibly small if

$$(40) \qquad\qquad n > \mathfrak{N} = \frac{C^3 M^{\star}}{(m'm'')^2 L}.$$

Let us consider only the $+$ term. We set

$$(41) \qquad \Theta_3(\mathfrak{m}', \mathfrak{d}) = \sup_{u} \sum_{C<c<2C} \sum_{\substack{d|c \\ d\sim\mathfrak{d}}} \frac{1}{d} \sum_{\substack{m',m''|cm \\ m'\sim\mathfrak{m}'}} \sum_{r\in\mathbb{Z}} W\left(\frac{r}{\tilde{R}}\right)$$

$$\times \left| \sum_{n\leq\mathfrak{N}} \frac{\lambda(n, m'')}{m''n} \sideset{}{^\star}\sum_{\substack{\beta \bmod mc/m' \\ r+\beta m' \equiv 0 \bmod c/d}} S(m'\beta, n; mc/m'') \tilde{V}_{+}^{\star}\left(\frac{c}{C}, \frac{m'^2 m''^2 nL}{m^3 c^3}, iu\right) \right|^2.$$

Here the supremum is taken over the range $|u| \ll M^{\hat{\theta}+\varepsilon}$ and W is a nonnegative compactly supported smooth function on $(0, \infty)$ with $W(x) = 1$ for $x \in [1, 2]$.

Also $\tilde{V}^\star_+(x, y, s)$ is the Mellin transform of $V^\star_+(x, y, z)$ in the z variable. The contribution of the $+$ term to (36) will be denoted by $\mathcal{O}^+(C, m; L, \mathfrak{m}', \mathfrak{d})$.

LEMMA 21. *We have*

$$(42) \qquad \mathcal{O}^+(C, m; L, \mathfrak{m}', \mathfrak{d}) \ll M^{\hat{\theta}+\varepsilon} \frac{RN\tilde{N}}{m^2 CMP^5} \, \mathfrak{d}^{-1/2} \sqrt{P^3 C\tilde{R}} \, \sqrt{\Theta_3(\mathfrak{m}', \mathfrak{d})}.$$

Proof. Substituting the expression in (39) (only the term with the $+$ sign) in the second line of (36) we get

$$\frac{RN\tilde{N}}{m^2 CMP^5} \sum_{\substack{P<p<2P \\ p \text{ prime}}} \phi(p)\chi(p) \sum_{c=1}^\infty \sum_{\substack{r=1 \\ (p,r)=1}}^\infty \bar{\chi}(r) e\left(\frac{Mc\bar{r}}{p}\right) \sum_{\substack{d|c \\ d\sim\mathfrak{d}}} \frac{\mu(d)}{d} \sum_{\substack{m',m''|cm \\ m'\sim\mathfrak{m}'}}$$

$$\sum_{n\leq\mathfrak{N}} \frac{\lambda(n, m'')}{m''n} \sideset{}{^\star}\sum_{\substack{\beta \bmod mc/m' \\ r+\beta m'\equiv 0 \bmod c/d}} S(m'\beta, n; mc/m'') \, V^\star_+\left(\frac{c}{C}, \frac{m'^2 m''^2 nL}{m^3 c^3}, \frac{r}{\tilde{R}}\right).$$

Here the sum over n is truncated at $\mathfrak{N}$ at a cost of a negligible error term. Taking inverse Mellin transform to free the variable r from the weight function and then taking absolute values, we bound the above sum by

$$\frac{RN\tilde{N}}{m^2 CMP^5} \int_{-M^{\hat{\theta}+\varepsilon}}^{M^{\hat{\theta}+\varepsilon}} \sum_{C<c<2C} \sum_{\substack{d|c \\ d\sim\mathfrak{d}}} \frac{1}{d} \sum_{\substack{m',m''|cm \\ m'\sim\mathfrak{m}'}} \sum_{\tilde{R}<r<2\tilde{R}} \left| \sum_{\substack{P<p<2P \\ p \text{ prime} \\ (p,r)=1}} \phi(p)\chi(p) e\left(\frac{Mc\bar{r}}{p}\right) \right|$$

$$\times \left| \sum_{n\leq\mathfrak{N}} \frac{\lambda(n, m'')}{m''n} \sideset{}{^\star}\sum_{\substack{\beta \bmod mc/m' \\ r+\beta m'\equiv 0 \bmod c/d}} S(m'\beta, n; mc/m'')\tilde{V}^\star_+\left(\frac{c}{C}, \frac{m'^2 m''^2 nL}{m^3 c^3}, iu\right) \right| du.$$

Recall that the weight function $V^\star_+$ does not depend on p.

Applying the Cauchy inequality we get that the above sum is dominated by

$$M^{\hat{\theta}+\varepsilon} \frac{RN\tilde{N}}{m^2 CMP^5} \mathfrak{d}^{-1/2} \sqrt{\Theta_1} \sqrt{\Theta_3(\mathfrak{m}', \mathfrak{d})},$$

where Θ_1 is as given in (24). Using (26) the lemma follows. $\qquad\square$

We have considered only the $+$ term from (39). The contribution of the $-$ term can be analyzed in a similar fashion. The bound that we obtain is not sensitive to this sign.

7.3. *Bound for Θ_3.* Now we consider $\Theta_3 = \Theta_3(\mathfrak{m}', \mathfrak{d})$. We will see that the estimation of this sum is related with that of the character sum

$$\mathfrak{C}_0 = \sideset{}{^\star}\sum_{\substack{\beta,\beta' \bmod mc/m' \\ \beta m'\equiv\beta' m' \bmod c/d}} \sideset{}{^\star}\sum S(m'\beta, n; mc/m'')S(m'\beta', n'; mc/m'').$$

We first prove the following estimate.

Lemma 22. *We have*

(43) $\mathfrak{C}_0 \ll \dfrac{cmm''}{m'}\Big((cm/m'',n)+(cm/m'',n')\Big)(cm/m'',n-n')$.

Proof. Let p be a prime with $v_p(c)=\gamma$, $v_p(d)=\delta$, $v_p(m)=\mu$, $v_p(m')=\mu'$ and $v_p(m'')=\mu''$. We consider the character sum

$$\sideset{}{^\star}\sum_{\substack{\beta,\beta' \bmod p^{\gamma+\mu-\mu'} \\ \beta p^{\mu'}\equiv\beta' p^{\mu'} \bmod p^{\gamma-\delta}}} \sideset{}{^\star}\sum S(p^{\mu'}\beta a, nb; p^{\gamma+\mu-\mu''})S(p^{\mu'}\beta' a, n'b; p^{\gamma+\mu-\mu''}),$$

where $p \nmid ab$. If $\mu' \geq \gamma - \delta$, then the sum splits into a product of two sums

$$\sideset{}{^\star}\sum_{\beta \bmod p^{\gamma+\mu-\mu'}} S(p^{\mu'}\beta a, nb; p^{\gamma+\mu-\mu''}) \sideset{}{^\star}\sum_{\beta' \bmod p^{\gamma+\mu-\mu'}} S(p^{\mu'}\beta' a, n'b; p^{\gamma+\mu-\mu''}),$$

which can be written as a product of Ramanujan sums

$$\mathfrak{c}_{p^{\gamma+\mu-\mu''}}(n)\mathfrak{c}_{p^{\gamma+\mu-\mu''}}(n')\mathfrak{c}^2_{p^{\gamma+\mu-\mu'}}(p^{\mu''})$$
$$\leq (p^{\gamma+\mu-\mu''},n)(p^{\gamma+\mu-\mu''},n')(p^{\gamma+\mu-\mu'},p^{\mu''})^2.$$

The last term can be bounded by

$$p^{\gamma+\mu-\mu'+\mu''}\Big((p^{\gamma+\mu-\mu''},n)+(p^{\gamma+\mu-\mu''},n')\Big)(p^{\gamma+\mu-\mu''},n-n').$$

On the other hand, if $\mu' < \gamma - \delta$, then we have congruence restriction $\beta \equiv \beta' \bmod p^{\gamma-\delta-\mu'}$, and the above sum boils down to

$$\sideset{}{^\star}\sum_{\beta \bmod p^{\gamma-\delta-\mu'}} \sum_{\beta_1 \bmod p^{\mu+\delta}} S\Big(p^{\mu'}(\beta+\beta_1 p^{\gamma-\delta-\mu'})a, nb; p^{\gamma+\mu-\mu''}\Big)$$
$$\times \sum_{\beta'_1 \bmod p^{\mu+\delta}} S\Big(p^{\mu'}(\beta+\beta'_1 p^{\gamma-\delta-\mu'})a, n'b; p^{\gamma+\mu-\mu''}\Big).$$

Opening the Kloosterman sums we observe that the sums over β_1 and β'_1 vanishes unless $\mu'' = \mu + \delta$. In this case we also need $p^{\mu'}|n$ and n', otherwise the average of the Kloosterman sum vanishes. Set $n = p^{\mu'}\tilde{n}$ and $n' = p^{\mu'}\tilde{n}'$. The character sum now reduces to

$$p^{2(\mu+\delta)} \sideset{}{^\star}\sum_{\beta \bmod p^{\gamma-\delta-\mu'}} S\Big(\beta a, \tilde{n}b; p^{\gamma-\delta-\mu'}\Big) S\Big(\beta a, \tilde{n}'b; p^{\gamma-\delta-\mu'}\Big).$$

As $p \nmid ab$, we can change variables to arrive at

$$p^{2(\mu+\delta)} \sideset{}{^\star}\sum_{\beta \bmod p^{\gamma-\delta-\mu'}} S\Big(\beta, \tilde{n}; p^{\gamma-\delta-\mu'}\Big) S\Big(\beta, \tilde{n}'; p^{\gamma-\delta-\mu'}\Big),$$

which is given by

$$
p^{\gamma+2\mu+\delta-\mu'}\mathfrak{c}_{p^{\gamma-\delta-\mu'}}(\tilde{n}-\tilde{n}') - p^{\gamma+2\mu+\delta-\mu'-1}\mathfrak{c}_{p^{\gamma-\delta-\mu'-1}}\left(\frac{\tilde{n}}{p}-\frac{\tilde{n}'}{p}\right).
$$

This is bounded by

$$
p^{\gamma+\mu-\mu'+\mu''}(p^{\gamma+\mu-\mu''}, n-n').
$$

With this we conclude the lemma. □

Next we will use the above bound for the character sum $\mathfrak{C}_0$ to obtain a bound for Θ_3.

LEMMA 23. *We have*

$$
\Theta_3 \ll \frac{M^\star \tilde{R}}{\mathfrak{m}'}\left(\frac{C^2}{\tilde{N}}+C\right).
$$

Proof. Consider the expression (41). Opening the absolute square we perform Poisson summation on the r sum with modulus c/d. We get

$$
(44) \qquad \Theta_3 = \sup_u \tilde{R} \sum_{\substack{C<c<2C}} \frac{1}{c} \sum_{\substack{d|c \\ d\sim\mathfrak{d}}} \sum_{\substack{m',m''|cm \\ m'\sim\mathfrak{m}'}}
$$

$$
\times \sum_{n,n'\leq\mathfrak{N}} \frac{\lambda(n,m'')}{m''n}\frac{\overline{\lambda(n',m'')}}{m''n'}\tilde{V}_+^\star(\cdots)\bar{\tilde{V}}_+^\star(\cdots)\sum_{r\in\mathbb{Z}}\mathfrak{I}\,\mathfrak{C},
$$

where the character sum is given by

$$
\mathfrak{C} = \sideset{}{^\star}\sum_{\substack{\beta,\beta' \bmod mc/m' \\ \beta m'\equiv\beta'm' \bmod c/d}}\sideset{}{^\star}\sum S(m'\beta,n;mc/m'')S(m'\beta',n';mc/m'')e\left(-\frac{r\beta m'}{c/d}\right)
$$

and the integral is given by

$$
\mathfrak{I} = \int_\mathbb{R} W(z)e\left(-\frac{z\tilde{R}r}{c/d}\right)\mathrm{d}z.
$$

By repeated integration by parts it follows that the integral is negligibly small if

$$
|r| \gg \frac{CM^\varepsilon}{\mathfrak{d}\tilde{R}}.
$$

Given the restriction on the sizes of C and $\tilde{R}$, we see that the nonzero frequencies $r\neq 0$ make a negligible contribution. For $r=0$, we use the trivial bound for the integral $\mathfrak{I} \ll 1$.

From (43) it follows that

$$\Theta_3 \ll \sup_u \tilde{R} \sum_{C<c<2C} \frac{1}{c} \sum_{\substack{d|c \\ d\sim\mathfrak{d}}}$$

$$\times \sum_{\substack{m',m''|cm \\ m'\sim\mathfrak{m}'}} \sum_{n,n'\leq\mathfrak{N}} \frac{|\lambda(n,m'')|}{m''n} \frac{|\lambda(n',m'')|}{m''n'} |\tilde{V}_+^\star(\cdots)||\tilde{V}_+^\star(\cdots)|$$

$$\times \frac{cmm''}{m'}(cm/m'',n)(cm/m'',n-n') + M^{-2013}.$$

Here it is not clear whether one can estimate this sum without taking a point-wise bound for the Fourier coefficients. Direct application of the Cauchy inequality is not helpful as the function $(u,v) \to \gcd(u,v)$ has a large dispersion. Using (R) and (RS) we get

$$(45) \quad \frac{\lambda(n,m'')}{m''n} \frac{\overline{\lambda(n',m'')}}{m''n'} \tilde{V}_+^\star(\cdots)\bar{\tilde{V}}_+^\star(\cdots) \ll \frac{1}{nn'm''^2} \frac{m'^2m''^2nL}{m^3c^3} \frac{m'^2m''^2n'L}{m^3c^3} M^\varepsilon.$$

Substituting this in the above expression, we get

$$\Theta_3 \ll M^\varepsilon \tilde{R} \sum_{C<c<2C} \sum_{d|c} \sum_{\substack{m',m''|cm \\ m'\sim\mathfrak{m}'}} \frac{m'^4m''^3L^2}{c^6m^6} \frac{m}{m'} \sum_{1\leq n,n'\leq\mathfrak{N}} (cm/m'',n)(c/m'',n-n').$$

Next we sum over n and n'. The contribution from the diagonal $n = n'$ is dominated by $c\mathfrak{N}/m''$, and the off-diagonal is dominated by $\mathfrak{N}^2$. Hence

$$\Theta_3 \ll \frac{M^\star \tilde{R}L^2}{\mathfrak{m}'} \sum_{C<c<2C} \sum_{d|c} \sum_{\substack{m',m''|cm \\ m'\sim\mathfrak{m}'}} \frac{m'^4m''^3}{(cm)^6} \left(\frac{c}{m''}\mathfrak{N} + \mathfrak{N}^2\right).$$

Substituting the size of $\mathfrak{N}$ from (40), we get

$$\Theta_3 \ll \frac{M^\star \tilde{R}L^2}{\mathfrak{m}'} \sum_{C<c<2C} \sum_{d|c} \sum_{\substack{m',m''|cm \\ m'\sim\mathfrak{m}'}} \frac{m'^4m''^3}{c^6} \left(\frac{C^4}{m'^2m''^3L} + \frac{C^6}{(m'm'')^4L^2}\right).$$

Now applying the upper bound for L from (37) we arrive at

$$\Theta_3 \ll \frac{M^\star \tilde{R}}{\mathfrak{m}'} \sum_{C<c<2C} \sum_{\substack{m',m''|cm \\ m'\sim\mathfrak{m}'}} \frac{m'^4m''^3}{c^6} \left(\frac{C^4}{m'^2m''^3} \frac{C^3}{\mathfrak{m}'^2\tilde{N}} + \frac{C^6}{(m'm'')^4}\right).$$

Trivially estimating the remaining sums, the lemma follows.   □

7.4. *Conclusion.* The following lemma summarizes the main content of this section.

LEMMA 24. *Suppose* $N^{1/2+\varepsilon} < P < M^{1-\varepsilon}$ *and* $\mathfrak{m}'\mathfrak{d} \geq M^\star$ *(which is a suitable large power of* M^θ*). Then we have*

$$\mathcal{O}(C,m;L,\mathfrak{m}',\mathfrak{d}) \ll \sqrt{N}M^{3/4-10\theta+\varepsilon}.$$

Remark 3. By keeping track of the powers of M^θ in the above calculations one can show that the above statement holds for $\mathfrak{m}'\mathfrak{d} \geq M^{40\theta}$.

Proof. Substituting the bound for Θ_3 (Lemma 23) in (42) we get

$$\mathcal{O}^+(C, m; L, \mathfrak{m}', \mathfrak{d}) \ll M^\varepsilon \frac{RN\tilde{N}}{CMP^5} \, \mathfrak{d}^{-1/2} \sqrt{P^3 C \tilde{R}} \sqrt{\frac{M^\star \tilde{R}}{\mathfrak{m}'} \left(\frac{C^2}{\tilde{N}} + C \right)}$$

$$\ll \frac{M^\star}{\sqrt{\mathfrak{m}'\mathfrak{d}}} \frac{RN\sqrt{\tilde{R}\tilde{N}}}{MP^5} \sqrt{P^3 \tilde{R}(C + \tilde{N})} \ll \frac{M^\star}{\sqrt{\mathfrak{m}'\mathfrak{d}}} M^{3/2} \left(\frac{N}{R} \right)^{1/4}.$$

The lemma follows as the same bound holds for the '$-$' term as well. □

Combining the above lemma with Lemma 20 we draw the following conclusion.

COROLLARY 3. *We have*

$$\mathcal{O}^\star(C, m) \ll M^\varepsilon \sup |\mathcal{O}(C, m; L, \mathfrak{m}', \mathfrak{d})| + \sqrt{N} M^{3/4 - 10\theta + \varepsilon},$$

where the supremum is taken over all L in the range (37), $\mathfrak{m}'\mathfrak{d} \ll M^\star$.

Combining with Corollary 2 and Lemma 14, we conclude the following.

COROLLARY 4. *Suppose $N^{1/2+\varepsilon} < P < M^{1-\varepsilon}$ and $\theta > 0$ sufficiently small. Then we have*

$$\mathcal{F} \ll M^\varepsilon \sum_{m \leq M^{4\theta}} \sup |\mathcal{O}(C, m; L, \mathfrak{m}', \mathfrak{d})| + \sqrt{N} M^{3/4 - \theta/2 + \varepsilon},$$

where the supremum is taken over all L in the range (37), $\mathfrak{m}'\mathfrak{d} \ll M^\star$, *and C in the range* (30).

8. Tamed dual off-diagonal in transition

We now return to (36) the expression we obtained after the first application of the Voronoi summation and dyadic segmentation. In the light of Corollary 4, to complete the proof of Proposition 2 we just need to consider $\mathcal{O}(C, m; L, \mathfrak{m}', \mathfrak{d})$ for small values of the parameters m, $\mathfrak{m}'$ and $\mathfrak{d}$ as given in Corollary 4. Note that for $m = \mathfrak{m}' = \mathfrak{d} = 1$, we have already given a sketch of the proof in Section 3.2. In the rest of the paper we will show that the argument holds even if the parameters are allowed to range over short intervals.

We take C in the transition range (30) and m in the range $1 \leq m \leq M^{\hat{\theta}}$. We write cd in place of c and change the order of summations. It follows that

$$(46) \qquad \sup_{\mathfrak{m}'\mathfrak{d} \leq M^\star} \mathcal{O}(C, m; L, \mathfrak{m}', \mathfrak{d}) \ll M^\varepsilon \frac{RN\tilde{N}_0}{CMP^5} \sup_{d\mathfrak{m}' \leq M^\star} \left| \mathfrak{O}(\cdots) \right|,$$

where

$$(47) \quad \mathfrak{O}(\cdots) = \sum_{\substack{P<p<2P \\ p \text{ prime}}} \phi(p)\chi(p) \sum_{\substack{r=1 \\ (p,r)=1}}^{\infty} \bar{\chi}(r) \sum_{c=1}^{\infty} e\left(\frac{Mcd\bar{r}}{p}\right)$$

$$\times \sum_{\substack{m'|cdm \\ m'\sim\mathfrak{m}'}} \frac{m'}{cdm} \sum_{\substack{\beta \bmod mcd/m' \\ r+\beta m'\equiv 0 \bmod c}}^{\star} \sum_{n=1}^{\infty} \lambda(m',n)e\left(\frac{\bar{\beta}n}{mcd/m'}\right) V\left(\frac{cd}{C},\frac{n}{L},\frac{r}{\tilde{R}}\right).$$

Here the weight function V is as given in (38).

8.1. *Evaluation of character sum and reciprocity.* Consider the character sum (which we again temporarily denote by $\mathfrak{C}$)

$$\mathfrak{C} = \sum_{\substack{\beta \bmod mcd/m' \\ r+\beta m'\equiv 0 \bmod c}}^{\star} e\left(\frac{\bar{\beta}n}{mcd/m'}\right),$$

which appears in (47). If $m = m' = d = 1$, then the character sum can be explicitly evaluated, and it is given by $e(-\bar{r}n/c)$. However in general it is not easy to evaluate the character sum due to the presence of factors m, m' and d. But we have now obtained a good control on the sizes of these factors, and consequently we can evaluate explicitly a large 'portion' of the character sum.

To this end, let $h = (m',c)$. We observe that $\mathfrak{C} = 0$ unless $h|r$. Accordingly we write $m' = hm'_1$, $c = hc_1$ and $r = hr_1$. Let $h_1 = (m'_1,r_1)$, and let us write $r_1 = h_1r_2$ and $m'_1 = h_1m_2$. Hence $(r_2,m_2) = 1$. We get

$$\mathfrak{C} = \sum_{\substack{\beta \bmod mc_1d/m'_1 \\ \beta\equiv -r_2\overline{m}_2 \bmod c_1}}^{\star} e\left(\frac{\bar{\beta}n}{mc_1d/m'_1}\right).$$

It follows that

$$\mathfrak{O}(\cdots) = \sum_{\substack{h_1,m_2 \\ h_1m_2=m'_1|dm}} \frac{m'_1}{dm} \sum_{h\sim\mathfrak{m}'/m'_1} \bar{\chi}(hh_1) \sum_{\substack{r_2=1 \\ (pm_2,r_2)=1}}^{\infty} \bar{\chi}(r_2) \sum_{\substack{P<p<2P \\ p \text{ prime}}} \phi(p)\chi(p)$$

$$\times \sum_{\substack{c_1=1 \\ (r_2m'_1,c_1)=1}}^{\infty} e\left(\frac{Mc_1d\overline{h_1r_2}}{p}\right) \frac{1}{c_1} \sum_{n=1}^{\infty} \lambda(m',n)$$

$$\times \sum_{\substack{\beta \bmod mc_1d/m'_1 \\ \beta\equiv -r_2\overline{m}_2 \bmod c_1}}^{\star} e\left(\frac{\bar{\beta}n}{mc_1d/m'_1}\right) V\left(\frac{c_1hd}{C},\frac{n}{L},\frac{r_2hh_1}{\tilde{R}}\right),$$

where $m' = hm'_1$.

Let $g = (c_1,dm)$. We write $c_1 = gc_2$ and $dm = g_0g'$, where $g_0|g^{\infty}$ and $(g',g) = 1$. Let $f = g'/m'_1$, which is an integer as $m'_1|g_0g'$ but $(c_1,m'_1) = 1$.

Then $\bar{\beta} = -\bar{r}_2 m_2 + \beta_1 c_1$ with $\beta_1 \bmod f g_0$. We have

$$\mathfrak{C} = e\left(-\frac{\bar{r}_2 m_2 n}{c_1 f g_0}\right) \sideset{}{^\dagger}\sum_{\beta_1 \bmod f g_0} e\left(\frac{\beta_1 n}{f g_0}\right),$$

where $\dagger$ implies that $(\beta_1, f) = 1$. In particular, $\mathfrak{C} = 0$ if $(r_2, c_1 f g_0) > 1$. Applying the reciprocity relation to the outer exponential and pulling out the greatest common divisor of β_1 and g_0, we get

$$\mathfrak{C} = e\left(\frac{\overline{c_1 f g_0} m_2 n}{r_2}\right) e\left(-\frac{m_2 n}{c_1 f g_0 r_2}\right) \sum_{g_1 g_2 = g_0} g_1 \sum_{\substack{1 \le \beta_1 < f g_2 \\ (\beta_1, f g_2) = 1}} e\left(\frac{\beta_1 n}{f g_2}\right).$$

We introduce the convention that for $a, b, c \in \mathbb{Z}$ and $c \ne 0$, we have $e(a\bar{b}/c) = 0$ for $(b, c) \ne 1$.

LEMMA 25. *There exists* $h_i \in \mathbb{Z}$, $i = 1, \ldots, 6$, *and* m' *with* $h_i, m' = O(M^\star)$, $(h_5, h_6) = 1$, *such that*

$$\sup_{\mathfrak{m}' \mathfrak{d} \le M^\star} |\mathcal{O}(C, m; L, \mathfrak{m}', \mathfrak{d})| \ll M^\star \frac{R N \tilde{N}}{C M P^5} |\Omega|,$$

where

$$
(48) \qquad \Omega = \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \phi(p) \chi(p) \sum_{\substack{c, r = 1 \\ (r, c h_3 h_5) = 1}}^{\infty} \bar{\chi}(r) e\left(\frac{M c \bar{r} h_1 \bar{h}_2}{p}\right) \frac{1}{c}
$$
$$
\times \sum_{n=1}^{\infty} \lambda(m', n) e\left(\frac{\bar{c} n h_3 \bar{h}_4}{r} + \frac{h_5 n}{h_6}\right) W\left(\frac{c}{C'}, \frac{n}{L}, \frac{r}{\tilde{R}'}\right),
$$

with $W \in \mathcal{W}_3(M^\star)$ *and* $C/M^\star \le C' \le C$, $\tilde{R}/M^\star \le \tilde{R}' \le \tilde{R}$.

Proof. The lemma follows by plugging in the expression for the character sum $\mathfrak{C}$ into $\mathfrak{O}(\cdots)$ and rearranging the sums. Some of the coprimality conditions are then removed using Möbius inversion, which only involves small factors. Here the new weight function is given by

$$W(x, y, z) = e\left(-\frac{m_1' m_2 h^2 h_1}{m} \frac{L}{C\tilde{R}} \frac{y}{xz}\right) V(x, y, z),$$

which is clearly in the class $\mathcal{W}_3(M^\star)$. $\qquad\square$

8.2. *The last application of Voronoi summation.* In the rest of the paper we will obtain a sufficient bound for Ω as defined in (48). We will apply the Voronoi summation formula. The modulus of the additive character is now $r h_6$. Notice that the application of the reciprocity relation has changed the modulus and so the Voronoi summation here is not a reversal process.

LEMMA 26. *There exists $h_i \in \mathbb{Z}$, $i = 1, \ldots, 6$, and m' (not necessarily same as in Lemma 25) with $h_i, m' = O(M^\star)$, and $t \in [-M^\star, M^\star]$, such that*

$$\Omega \ll M^\star L\,|\Delta| + M^{-2013},$$

where

(49)
$$\Delta = \sum_{\substack{P<p<2P \\ p\,\text{prime}}} \phi(p)\chi(p) \sum_{\substack{c,r=1 \\ (r,ch_3h_5)=1}}^{\infty} \bar{\chi}(r) e\left(\frac{Mc\bar{r}h_1\bar{h}_2}{p}\right) \frac{1}{c}$$

$$\times \sum_{m|r} \frac{m}{r^2} \sum_{1 \le n < \mathcal{N}} \frac{\lambda(m'm,n)}{n^{-\varepsilon+it}}\, S(m'\bar{\xi},n;h_3r/m)\, U_t\left(\frac{c}{C'},\frac{r}{\tilde{R}'}\right),$$

with $\xi = \bar{c}\bar{h}_4h_5 + h_6r$, $C/M^\star \le C' \le C$, $\tilde{R}/M^\star \le \tilde{R}' \le \tilde{R}$ and

(50)
$$\mathcal{N} = \frac{M^\star \tilde{R}^3}{m^2 L}.$$

Here the weight function U_t belongs to the class $\mathcal{W}_2(M^\star)$.

Proof. Applying the Voronoi summation, i.e., Lemma 19, on the sum over n in Ω, we get

(51)
$$\Omega = \sum_{\pm} \sum_{\substack{P<p<2P \\ p\,\text{prime}}} \phi(p)\chi(p) \sum_{\substack{c,r=1 \\ (r,ch_3h_5)=1}}^{\infty} \bar{\chi}(r) e\left(\frac{Mc\bar{r}h_1\bar{h}_2}{p}\right) \frac{h_6r}{c}$$

$$\times \sum_{m''|h_6rm'} \sum_{n=1}^{\infty} \frac{\lambda(m'',n)}{m''n}S(m'\bar{\xi},\pm n;m'h_6r/m'')W_{\pm}^{\star}\left(\frac{c}{C'},\frac{m''^2nL}{(h_6r)^3m'},\frac{r}{\tilde{R}'}\right),$$

with ξ as in the statement of the lemma. Recall that (see Lemma 19)

$$W_{\pm}^{\star}\left(\frac{c}{C'},\frac{m''^2nL}{(h_6r)^3m'},\frac{r}{\tilde{R}'}\right) = \frac{1}{2\pi i}\int_{(\sigma)}\left(\frac{m''^2nL}{(h_6r)^3m'}\right)^{-s}\gamma_{\pm}(s)\tilde{W}\left(\frac{c}{C},-s,\frac{r}{\tilde{R}'}\right)\mathrm{d}s.$$

This is negligibly small if

$$n \ge \frac{\tilde{R}^3M^\star}{m''^2L}.$$

For smaller values of n, we shift the contour in the definition of the integral transform to $\sigma = -1 + \varepsilon$, using (RS). The integrand decays rapidly for $t = \mathrm{Im}(s) \gg M^\star$ (as the Mellin transform decays beyond this range), and this part makes a negligible contribution. We now interchange the order of summations and the integral over t. Taking absolute value inside the integral, the lemma follows. Note that

$$U_t\left(\frac{c}{C'},\frac{r}{\tilde{R}'}\right) = \left(\frac{r}{\tilde{R}'}\right)^{3(-\varepsilon+it)}\tilde{W}\left(\frac{c}{C'},1-\varepsilon-it,\frac{r}{\tilde{R}'}\right). \qquad \square$$

In the rest of the paper we will obtain sufficient bounds for the expression in (49), which is uniform with respect to t in the desired range.

8.3. *Reciprocity and Poisson summation.* Next we wish to apply the Poisson summation formula on the sum over c. Recall that c is essentially the modulus of the 'circle method' (Petersson formula) that we applied at the initial stage. After a sequence of applications of summation formulas and reciprocity relations we are finally at the stage where we are able to sum over the modulus again. The variable c appears in the Kloosterman sum in (51). This Kloosterman sum has modulus $h_3 r/m$. Also c appears in the additive character that has modulus p. So apparently the total modulus is too large compared to the length of the sum. However we can now apply the reciprocity again to bring down the modulus.

LEMMA 27. *There exists* $h_i \in \mathbb{Z}$, $i = 1, \ldots, 6$, *and* m' *with* $h_i, m' = O(M^\star)$, *such that*
$$\Delta \ll |\Xi| + M^{-2013},$$
where
$$\Xi = \sum_{\substack{P<p<2P \\ p \text{ prime}}} \phi(p)\chi(p) \sum_{\substack{r=1 \\ (h_3 h_5, r)=1}}^{\infty} \bar{\chi}(r) \sum_{m|r} \frac{m}{r^3} \sum_{1\leq n<\mathcal{N}} \frac{\lambda(m'm,n)}{n^{-\varepsilon+it}} \sum_{|c|\ll\mathcal{C}} \mathfrak{C}\,\mathfrak{I},$$
and
$$\mathcal{C} = M^\star \frac{\tilde{R}}{C}.$$
The character sum is given by
$$\mathfrak{C} = \sum_{\substack{\gamma \bmod h_2 h_3 r \\ (r,\gamma)=1}} e\left(-\frac{Mh_1\bar{p}\gamma}{h_2 r} + \frac{c\gamma}{h_2 h_3 r}\right) S(m'\bar{\xi}, n; h_3 r/m),$$
with $\xi = \bar{\gamma}\bar{h}_4 h_5 + h_6 r$, *and the integral transform is given by*
$$\mathfrak{I} = \int_{\mathbb{R}} V_t\left(x, \frac{r}{\tilde{R}'}\right) e\left(-\frac{C'cx}{h_2 h_3 r}\right) \frac{\mathrm{d}x}{x},$$
with $V_t \in \mathcal{W}_2(M^\star)$.

Proof. In (49) we first use the reciprocity relation
$$e\left(\frac{Mc\bar{r}h_1\bar{h}_2}{p}\right) = e\left(-\frac{Mch_1\bar{p}}{h_2 r}\right) e\left(\frac{Mch_1}{h_2 pr}\right).$$
The last term can be absorbed in the weight function. Accordingly we let
$$V_t(x,y) = e\left(\frac{h_1 C'M}{h_2 p\tilde{R}'}\frac{x}{y}\right) U_t(x,y).$$
Observe that we (still) have $V_t(x,y) \in \mathcal{W}_2(M^\star)$.

We now study the sum over c in (49), which is given by

$$\sum_{\substack{c=1 \\ (r,c)=1}}^{\infty} e\left(-\frac{Mch_1\bar{p}}{h_2r}\right) \frac{1}{c} S(m'\bar{\xi}, n; h_3r/m) V_t\left(\frac{c}{C'}, \frac{r}{\tilde{R}'}\right).$$

We break the sum into congruence classes modulo h_2h_3r and apply the Poisson summation formula. We get

$$\frac{1}{h_2h_3r} \sum_{c\in\mathbb{Z}} \sum_{\substack{\gamma \bmod h_2h_3r \\ (r,\gamma)=1}} e\left(-\frac{Mh_1\bar{p}\gamma}{h_2r} + \frac{c\gamma}{h_2h_3r}\right) S(m'\bar{\xi}, n; h_3r/m)$$

$$\times \int_{\mathbb{R}} V_t\left(x, \frac{r}{\tilde{R}'}\right) e\left(-\frac{C'cx}{h_2h_3r}\right) \frac{\mathrm{d}x}{x},$$

where $\xi = \bar{\gamma}\bar{h}_4 h_5 + h_6 r$. From repeated integration by parts it follows that the integral is negligibly small if $|c| \gg \mathcal{C}$, and the lemma follows.   □

8.4. *Evaluation of character sums.* Now we write $r = r_1 r_2$, with $(r_1, h_2h_3) = 1$ and $r_2|(h_2h_3)^{\infty}$. Accordingly we split $m = m_1 m_2$, with $m_i|r_i$. We set $\zeta = h_3 r_2/m_2$. The character sum $\mathfrak{C}$ splits as a product of two character sums $\mathfrak{C} = \mathfrak{C}_1 \mathfrak{C}_2$. The one with modulus $h_2h_3r_2$ is given by

$$(52) \quad \mathfrak{C}_1 = \sum_{\substack{\gamma \bmod h_2h_3r_2 \\ (r_2,\gamma)=1}} S(m'\overline{\xi r_1/m_1}, n\overline{r_1/m_1}; \zeta) e\left(-\frac{Mh_1\gamma\overline{pr_1}}{h_2r_2} + \frac{\gamma\overline{cr_1}}{h_2h_3r_2}\right),$$

where ξ is as given in Lemma 27. The other sum with modulus r_1 is given by

$$\mathfrak{C}_2 = \sideset{}{^\star}\sum_{\gamma \bmod r_1} S(m'\bar{\eta}, n; r_1/m_1) e\left(-\frac{Mh_1\gamma\overline{ph_2r_2}}{r_1} + \frac{\gamma c\overline{h_2h_3r_2}}{r_1}\right),$$

where $\eta = \xi\zeta^2 \equiv \bar{\gamma}\bar{h}_4 h_5 \zeta^2 \bmod r_1$. For the former sum, we will establish the following bound.

LEMMA 28. *We have*

$$\mathfrak{C}_1 \ll M^{\star} \frac{r_2}{m_2}.$$

Proof. Suppose $p^{\ell} \| r_2$ with $\ell \geq 1$, and suppose $p^k \| h_2h_3$, $p^j \| h_3r_2/m_2$ (so $j \leq \ell + k$). Then we take $A, B, C \in \mathbb{Z}$ with $p \nmid A$ and study the sum

$$\sideset{}{^\star}\sum_{\gamma \bmod p^{\ell+k}} S(m'A\bar{\xi}, B; p^j) e\left(\frac{C\gamma}{p^{\ell+k}}\right).$$

The sum vanishes unless $p^{\ell+k-j}|C$, in which case it reduces to

$$p^{\ell+k-j} \sideset{}{^\star}\sum_{\gamma \bmod p^j} S(m'A\bar{\xi}, B; p^j) e\left(\frac{C'\gamma}{p^j}\right).$$

First consider the case where $\ell > k$, so that $2\ell > j$. Then $\bar{\xi} = \gamma h_4 \bar{h}_5 - (\gamma h_4 \bar{h}_5)^2 h_6 r$. Now if $\ell \geq j$, the character sum reduces to

$$p^{\ell+k-j} \sum_{\gamma \bmod p^j}^{\star} S(m'A\gamma h_4 \bar{h}_5, B; p^j) e\left(\frac{C\gamma}{p^j}\right).$$

Opening the Kloosterman sum we execute the sum over γ, which yields a Ramanujan sum. Using standard bounds for the Ramanujan sum we now get the bound $O(p^{\ell+k}(m', p^j))$ for the character sum. On the other hand, if $j > \ell \geq k$, then we write $\gamma = \gamma_1 + \gamma_2 p^{j-\ell}$ with γ_1 modulo $p^{j-\ell}$, $(\gamma_1, p) = 1$, and γ_2 modulo p^ℓ. Then the character sum reduces to

$$p^{\ell+k-j} \sum_{\substack{\gamma_1 \bmod p^{j-\ell} \\ \gamma_2 \bmod p^\ell}}^{\star} \sum S(m'A((\gamma_1 + \gamma_2 p^{j-\ell})h_4\bar{h}_5 - (\gamma_1 h_4 \bar{h}_5)^2 h_6 r), B; p^j)$$

$$\times e\left(\frac{C\gamma_1}{p^j} + \frac{C\gamma_2}{p^\ell}\right).$$

Opening the Kloosterman sum, executing the sum over γ_2, and trivially estimating the remaining sums we get the bound $O(p^{j+k}(m', p^\ell))$. In the case $\ell < k$ (including when $\ell = 0$) we trivially bound the sum by $O(p^{\ell+k+j}) = O(p^{2k+j})$. Putting the above bounds together we get the lemma. $\qquad \square$

LEMMA 29. *There exists* $h_i \in \mathbb{Z}$, $i = 1, \ldots, 5$, *with* $h_i = O(M^\star)$, *such that*

$$\Xi \ll \frac{M^\star P}{\tilde{R}^2} \sum_{u,v|h_1^\infty} \sum \sum_{m_1=1}^{\infty} \sum_{\substack{\delta_1\delta_2 \ll M^{3/4} \\ (\delta_1 h, \delta_2)=1}} \sum \frac{m_1}{u\delta_2} \Psi + \frac{M^\star \tilde{R}^3}{LM^{1/4}},$$

where $h = h_1 h_2 h_4 h_5$ *and*

(53) $$\Psi = \sum_{|c| \ll \mathcal{C}} \sum_{1 \leq n < \mathcal{N}^\dagger} \left| \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \sum_{|r| \asymp R^\dagger}^{\flat} \upsilon(\cdots)\, \mathfrak{C}_1(n)\, \psi_n(\cdots) \right|.$$

The $\flat$ *indicates the coprimality condition* $(\delta_1 h, r) = 1$, *the factors* $\upsilon(\cdots) \ll 1$, *and do not depend on* n, *and* ψ_n *vanishes unless* $m_1 | (Mh_2 - cp)$, *and in this case we have*

$$\psi_n(\cdots) = \sum_{\substack{\alpha \bmod \delta_2 r \\ \xi p\alpha \equiv (Mh_2-cp)/m_1 \bmod r}}^{\star} e\left(\frac{\bar{\alpha}n}{\delta_2 r}\right),$$

where $\xi = \delta_1 v h_3 \overline{uh_4}$. *The factor* $\mathfrak{C}_1(n)$ *is of the form* (52), *with modulus* $\zeta = h_5 u$. *Also*

$$R^\dagger = \frac{M^\star \tilde{R}}{\delta_1 \delta_2 m_1 uv} \quad and \quad \mathcal{N}^\dagger = \frac{M^\star \tilde{R}^3}{(\delta_1 m_1 v)^2 L}.$$

Proof. Consider the expression for Ξ as given in Lemma 27. We need to explicitly evaluate the character sum $\mathfrak{C}_2$. Opening the Kloosterman sum we get

$$\mathfrak{C}_2 = \sideset{}{^\star}\sum_{\alpha \bmod r_1/m_1} e\left(\frac{\bar{\alpha}n}{r_1/m_1}\right)$$
$$\times \sideset{}{^\star}\sum_{\gamma \bmod r_1} e\left(\frac{m_1 m'\alpha\gamma h_4 \overline{h_5 \zeta^2}}{r_1} - \frac{Mh_1\gamma\overline{ph_2 r_2}}{r_1} + \frac{\gamma c\overline{h_2 h_3 r_2}}{r_1}\right).$$

The sum over γ is a Ramanujan sum. So we get

$$r_1 \sum_{\delta|r_1} \frac{\mu(\delta)}{\delta} \sideset{}{^\star}\sum_{\substack{\alpha \bmod r_1/m_1 \\ m_1 m'\alpha h_4 \overline{h_5 \zeta^2} \equiv (Mh_1 h_3 - cp)\overline{h_2 h_3 pr_2} \bmod r_1/\delta}} e\left(\frac{\bar{\alpha}n}{r_1/m_1}\right).$$

Now we write rm_1 in place of r_1, and we split $\delta = \delta_1\delta_2$ with $\delta_1|m_1$, $(\delta_1,r)=1$ and $\delta_2|r$. The above sum now becomes

$$rm_1 \sum_{\substack{\delta_1|m_1 \\ \delta_2|r \\ (\delta_1,r)=1}} \frac{\mu(\delta_1\delta_2)}{\delta_1\delta_2} \sideset{}{^\star}\sum_{\substack{\alpha \bmod r \\ m_1 m'\alpha h_4 \overline{h_5 \zeta^2} \equiv (Mh_1 h_3 - cp)\overline{h_2 h_3 pr_2} \bmod rm_1/\delta_1\delta_2}} e\left(\frac{\bar{\alpha}n}{r}\right).$$

We substitute this in the expression for Ξ as given in Lemma 27. Then interchange the order of summations and rename variables (e.g., we write $\delta_1 m_1$ in place of m_1 and $\delta_2 r$ in place of r). Finally taking absolute value and using (R), we arrive at the inequality

$$\Xi \ll \frac{M^\star P}{\tilde{R}^2} \sum_{r_2|(h_2 h_3)^\infty} \sum_{m_2|r_2} \sum_{\substack{m_1=1 \\ (h_2 h_3, m_1)=1}}^\infty \sum_{\substack{\delta_1,\delta_2=1 \\ (\delta_1 h_2 h_3,\delta_2)=1}}^\infty \frac{m_1 m_2}{\delta_2 r_2}$$
$$\times \sum_{|c|\ll\mathcal{C}} \sum_{1\leq n<\mathcal{N}^\dagger} \left| \sum_{\substack{P<p<2P \\ p\text{ prime}}} \sum_{\substack{r=1 \\ (\delta_1 h_2 h_3, r)=1}}^\infty \upsilon(\cdots)\,\mathfrak{C}_1\,\psi_n(\cdots) \right|,$$

where $\upsilon(\cdots) = \phi(p)\chi(p\bar{r})\mathfrak{I}/P$ and

$$\psi_n(\cdots) = \sideset{}{^\star}\sum_{\substack{\alpha \bmod \delta_2 r \\ m_1 m'\alpha h_4 \overline{h_5 \zeta^2} \equiv (Mh_1 h_3 - cp)\overline{h_2 h_3 pr_2} \bmod rm_1}} e\left(\frac{\bar{\alpha}n}{\delta_2 r}\right).$$

Note that $\upsilon(\cdots)$ is free of n and is bounded by $O(1)$. Moreover this factor vanishes outside the given range for r.

Finally we show that we only need to consider small values of δ_i. By trivial estimation we get $\psi_n(\cdots) \ll M^\star(\delta_2, n)$. Consequently

$$\frac{M^\star P}{\tilde{R}^2} \sum_{r_2|(h_2h_3)^\infty} \sum_{m_1=1}^\infty \sum_{m_2|r_2} \sum_{\substack{\delta_1 \sim D_1 \\ \delta_2 \sim D_2}} \frac{m_1 m_2}{r_2 \delta_2}$$

$$\times \sum_{|c| \ll \mathcal{C}} \sum_{1 \leq n < \mathcal{N}^\dagger} \left| \sum_{\substack{P < p < 2P \\ p \text{ prime}}} \sum_{|r| \asymp R^\dagger}^\flat \upsilon(\cdots) \, \mathfrak{C}_1 \, \psi_n(\cdots) \right|$$

is dominated by

$$\frac{M^\star P^2}{\tilde{R}^2} \sum_{r_2|(h_2h_3)^\infty} \sum_{m_1=1}^\infty \sum_{m_2|r_2} \sum_{\substack{\delta_1 \sim D_1 \\ \delta_2 \sim D_2}} \frac{m_1 \mathcal{C} R^\dagger}{\delta_2} \sum_{1 \leq n < \mathcal{N}^\dagger} (n, \delta_2) \ll \frac{M^\star \tilde{R}^2 \mathcal{C} P^2}{L D_1^2 D_2}.$$

The lemma follows after renaming the 'small variables.' Lengths of the r and the n sum are derived from those given in Lemma 26. For example, note that m in Lemma 26 has been factorized here as $m = \delta_1 m_1 m_2$, and we have changed m_2 to v in the statement of the lemma. $\qquad\square$

8.5. *Application of Cauchy's inequality and Poisson summation.* Let h_i, $i = 1, \ldots, 5$ be as in the statement of Lemma 29. Let $\delta_1, \delta_2 \in \mathbb{N}$. Set

$$\rho = (Mh_2 - cp)/m_1, \quad \rho' = (Mh_2 - cp')/m_1.$$

Let $\zeta = h_5 u$, $\eta = p\delta_1 h_3 v$, $\eta' = p'\delta_1 h_3 v$, $\mu = h_4 u\rho$ and $\mu' = h_4 u\rho'$. Note that $(\eta, r) = (h_3, r) = O(M^\star)$. Set

$$(54) \qquad \hat{\mathfrak{C}} = \sideset{}{^\star}\sum_{\substack{\alpha \bmod \delta_2 r \\ \eta\alpha \equiv \mu \bmod r \\ \eta'\alpha' \equiv \mu' \bmod r' \\ \bar{\alpha}r' - \bar{\alpha}'r \equiv n \bmod \delta_2 rr'}} \sideset{}{^\star}\sum_{\alpha' \bmod \delta_2 r'} 1,$$

and define

$$(55) \qquad \mathfrak{W} = \sideset{}{^\star}\sum_{\substack{P < p,p' < 2P \\ p,p' \text{ prime}}} \sideset{}{^\star}\sum \sum_{|r|,|r'| \asymp R^\dagger}^\flat \sum^\flat \sum_{|n| \ll N^\star} |\hat{\mathfrak{C}}|,$$

where

$$R^\dagger = \frac{M^\star \tilde{R}}{\delta_1 \delta_2 m_1 uv} \quad \text{and} \quad N^\star = \frac{M^\star}{\delta_2 u} \frac{L}{\tilde{R}}.$$

Here the $\star$ on the p sum indicates the restriction $m_1|(Mh_2 - cp)$.

LEMMA 30. *We have*

$$\Psi \ll \frac{uM^\star \tilde{R}^3}{(\delta_1 m_1 v)^2 L} \sum_{|c| \ll \mathcal{C}} \mathfrak{W}^{1/2} + M^{-2013}.$$

Proof. Applying the Cauchy inequality to the expression on the right-hand side of (53), we get

$$(56) \qquad \Psi = \sum_{|c| \ll \mathcal{C}} \sqrt{\mathcal{N}^\dagger} \sqrt{\Theta_4(\cdots)},$$

where

$$\Theta_4(\cdots) = \sum_{1 \le n < \mathcal{N}^\dagger} \left| \sideset{}{^\star}\sum_{\substack{P < p < 2P \\ p \text{ prime}}} \sideset{}{^\flat}\sum_{r \asymp R^\dagger} \upsilon(\cdots)\, \mathfrak{C}_1(n)\, \psi_n(\cdots) \right|^2.$$

Using positivity we now smooth out the n-sum and then apply the Poisson summation formula after opening the absolute square. The modulus is $\zeta \delta_2 r r'$. (Recall that $\mathfrak{C}_1(n)$ is periodic in n with modulus ζ.) We get

$$\Theta_4(\cdots) \le \frac{\mathcal{N}^\dagger}{\zeta \delta_2} \sideset{}{^\star}\sum_{\substack{P < p, p' < 2P \\ p, p' \text{ prime}}} \sideset{}{^\star}\sum \sideset{}{^\flat}\sum_{|r|, |r'| \asymp R^\dagger} \sideset{}{^\flat}\sum \frac{1}{rr'} \upsilon(\cdots)\bar{\upsilon}(\cdots) \sum_{n \in \mathbb{Z}} \mathfrak{C}\mathfrak{I},$$

where the new character sum is given by

$$\sum_{\beta \bmod \zeta \delta_2 r r'} \mathfrak{C}_1(\beta)\overline{\mathfrak{C}'_1(\beta)}\psi_\beta(\cdots)\overline{\psi'_\beta(\cdots)}\, e\left(\frac{n\beta}{\zeta \delta_2 r r'} \right).$$

Now using the coprimality $(\zeta, \delta_2 r r') = 1$, we split the character sum into a product of two character sums. The one with modulus ζ is estimated trivially using Lemma 28. The other character sum modulo $\delta_2 r r'$ is given by

$$\sideset{}{^\star}\sum_{\substack{\alpha \bmod \delta_2 r \\ \xi p \alpha \equiv (Mh_2 - cp)/m_1 \bmod r}} \sideset{}{^\star}\sum_{\substack{\alpha' \bmod \delta_2 r' \\ \xi p \alpha' \equiv (Mh_2 - cp')/m_1 \bmod r'}} \sum_{\beta \bmod \delta_2 r r'} e\left(\frac{\bar{\alpha}\beta}{\delta_2 r} - \frac{\bar{\alpha}'\beta}{\delta_2 r'} \right) e\left(\frac{\bar{\delta} n \beta}{\delta_2 r r'} \right).$$

This reduces to $\delta_2 r r' \hat{\mathfrak{C}}$. The integral $\mathfrak{I}$ is the Fourier transform of a smooth bump function with compact support. By repeated integration by parts we see that the integral is negligibly small if

$$|n| \gg M^\star \frac{u \delta_2 R^{\dagger 2}}{\mathcal{N}^\dagger} \asymp N^\star.$$

We conclude that

$$\Theta_4(\cdots) \ll u^2 M^\star \mathcal{N}^\dagger \, \mathfrak{W} + M^{-20130},$$

where $\mathfrak{W}$ is as given in (55). Substituting the above bound into (56) and substituting the size of $\mathcal{N}^\dagger$, we conclude the lemma. $\qquad \square$

8.6. *A counting problem.* It remains to estimate $\mathfrak{W}$ for $\delta_1\delta_2 \ll M^{3/4}$.

LEMMA 31. *For $N^{1/2+\varepsilon} < P < M^{1-\varepsilon}$, $\delta_1\delta_2 \ll M^{3/4}$, and θ sufficiently small, we have*

$$\mathfrak{W} \ll M^\star v\, M^{7/4}P^2.$$

Proof. First consider the contribution of the zero frequency $n = 0$. In this case the last congruence in (54) implies that $r = r'$ and $\alpha = \alpha'$. The other two congruences now imply that $\eta'\mu \equiv \eta\mu' \bmod r$. For any (p, p', r) satisfying the above congruence, we have $\hat{\mathfrak{C}} \ll M^\star\delta_2$, as the second congruence in (54) restricts the number of α modulo r by $(\eta, r) \leq M^\star$. If $p = p'$, then there are $O(R^\dagger)$ many choice for r. On the other hand, if $p \neq p'$ (which implies $\eta'\mu \neq \eta\mu'$), there are $O(M^\varepsilon)$ many choice for r. It follows that the contribution of this case, $n = 0$, to (55) is dominated by

$$(57) \qquad\qquad O\left(M^\star\delta_2\left(PR^\dagger + P^2\right)\right).$$

Let $d = (r, r')$, and set $r = ds$ and $r' = ds'$ (so $ss' \neq 0$). Then $d|n$ and we write $n = dk$ with $k \neq 0$. The first two congruences in (54) imply that there are $O(M^\star)$ many α (resp. α') satisfying the congruence modulo r (resp. r'). Also the third congruence condition in (54) implies that $\bar{\alpha}s' - \bar{\alpha}'s \equiv k \bmod \delta_2$. Since $(s, s') = 1$, we see that modulo δ_2 there are $O(\delta_2)$ many pairs (α, α'). Consequently $\hat{\mathfrak{C}} \ll M^\star\delta_2$.

From (54) we conclude that $\eta's \equiv -k\mu' \bmod s'$ and $\eta s' \equiv k\mu \bmod s$. We write the last congruence as an equation:

$$(58) \qquad\qquad \eta s' = k\mu + es.$$

We will now consider the generic case where

$$e \neq 0, \quad \eta'\mu \neq e\mu' \quad \text{and} \quad \eta\mu's' - \eta'\mu s \neq \mu\mu'k.$$

Comparing the sizes of the terms we get $|e| \ll \delta_1 v M^\star P$. Multiplying the equation with η' and using the first congruence, we arrive at $k(\eta'\mu - e\mu') \equiv 0 \bmod s'$. Since $(s', k) = 1$, it follows that $s'|(\eta'\mu - e\mu')$. Suppose (p, p', e) is given with $\eta'\mu \neq e\mu'$. Then there are $O(M^\varepsilon)$ many possible s'. Once s' is obtained, then to find the number of k one looks at the congruence

$$\eta s' \equiv k\mu \bmod |e|$$

if $e \neq 0$. The number of k is given by

$$(e, \mu)\left(1 + \frac{N^\star}{d|e|}\right).$$

If $\eta\mu's' - \eta'\mu s \neq \mu\mu'k$, then there are $O(M^\varepsilon)$ many d. The total contribution of the generic case is given by

$$M^\varepsilon \delta_2 \sum_p \sum_{p'} \sum_{e \neq 0} (e, \mu) \left(1 + \frac{N^\star}{|e|}\right) \ll M^\star \delta_2 \left(\delta_1 v P^3 + N^\star P^2\right).$$

Now we are left with three degenerate cases. First suppose $e = 0$, but $\eta\mu's' - \eta'\mu s \neq \mu\mu'k$ (so that the number of d still remains $O(M^\varepsilon)$). Then $\eta s' = k\mu$. Consequently $p|k$, and once a p and such a k are given, we can solve for s'. Then once p' is given we determine s from the congruence $\eta's \equiv -k\mu' \bmod s'$. The number of such s is $O(M^\star)$ if $p' \nmid k$ and $O(pM^\star)$ if $p'|k$. From size consideration we get that $p|k$ and $p'|k$ imply $p = p'$. So the total contribution of this case to (55) is dominated by

$$M^\star \delta_2 \sum_p \sum_{p'} \frac{N^\star}{p} + M^\star \delta_2 \sum_p p\frac{N^\star}{p} \ll M^\star \delta_2 \, N^\star P.$$

This is absorbed by the generic count.

Next suppose $\eta'\mu = e\mu'$, but $\eta\mu's' - \eta'\mu s \neq \mu\mu'k$. Then $p'|e$, and once p' and such an e are given, we can then solve for p. Then from (58) we get the congruence $k\mu \equiv -es \bmod \eta$. Given s, there are $O(1 + N^\star/P)$ many choices for k. Then s' is solved from equation (58). So the total contribution of this case is dominated by

$$M^\star \delta_2 \sum_s \sum_{p'} \left(1 + \frac{N^\star}{P}\right) \frac{\delta_1 v P}{p'} \ll M^\star \delta_1 \delta_2 v \left(P + N^\star\right)R^\dagger.$$

Now it remains to count the number of solutions of the equation $\eta\mu's' - \eta'\mu s = \mu\mu'k$. Suppose we are given (s, p). Then there are $O(M^\varepsilon)$ many choice for p' as $\mu'|\eta'\mu s$ (and $(\mu', p') = 1$). Suppose we are given (s, p, p'). Let (s', k) and $(\bar{s}', \bar{k})$ be two possible pairs. Then from the equation we get $\mu|\eta(s' - \bar{s}')$. Hence the number of possible pairs (s', k) is bounded by

$$O\left(M^\star \left(1 + \frac{R^\dagger}{d|\mu|}\right)\right).$$

The contribution of this case is dominated by

$$M^\star \delta_2 \sum_d \sum_s \sum_p \left(1 + \frac{R^\dagger}{d|\mu|}\right) \ll M^\star \delta_2 \left(PR^\dagger + R^{\dagger 2}\right).$$

Comparing the bounds obtained above, and using $\delta_1 \delta_2 \ll M^{3/4}$, we are now able to conclude the lemma. $\square$

9. **Conclusion**

The counting function $\mathfrak{W}$ vanishes for uvm_1 large, say for $uvm_1 > M^{2013}$. Substituting the bound from Lemma 31 into Lemma 30 we get

$$\Psi \ll \frac{uM^\star \tilde{R}^3 \mathcal{C} M^{7/8} P}{(\delta_1 m_1)^2 v^{3/2} L}.$$

Substituting into Lemma 29 it follows that

$$\Xi \ll \frac{M^\star P^2 \tilde{R} \mathcal{C} M^{7/8}}{L} + \frac{M^\star \tilde{R}^3}{LM^{1/4}}.$$

From Lemma 27 we see that the same bound holds for Δ. Substituting in Lemma 26 we conclude that

$$\Omega \ll M^\star P^2 \tilde{R} \mathcal{C} M^{7/8} + \frac{M^\star \tilde{R}^3}{M^{1/4}} \ll M^\star M^{19/8} P^2.$$

Then from Lemma 25 it follows that for $N^{1/2+\varepsilon} < P < M^{1-\varepsilon}$ and θ sufficiently small, we have

$$\sup_{\mathfrak{m}'\mathfrak{d}\leq M^\star} |\mathcal{O}(C, m; L, \mathfrak{m}', \mathfrak{d})| \ll \sqrt{N}\,\frac{M^\star M^{21/8}}{P^2}$$

for m in the range of Corollary 4 (at the end of Section 7). From Corollary 4 it now follows that

$$\mathcal{F} \ll M^\varepsilon\,\sqrt{N}\,\left(\frac{M^\star M^{21/8}}{P^2} + M^{3/4-\theta/2}\right).$$

We can now conclude Proposition 2 by picking $P = M^{1-\theta_0}$ with $\theta_0 > 0$ sufficiently small.

References

[1] V. BLOMER, Subconvexity for twisted L-functions on GL(3), *Amer. J. Math.* **134** (2012), 1385–1421. MR 2975240. Zbl 1297.11046. http://dx.doi.org/10.1353/ajm.2012.0032.

[2] D. BUMP, *Automorphic Forms on* GL(3, **R**), *Lecture Notes in Math.* **1083**, Springer-Verlag, New York, 1984. MR 0765698. Zbl 0543.22005.

[3] D. A. BURGESS, On character sums and primitive roots, *Proc. London Math. Soc.* **12** (1962), 179–192. MR 0132732. Zbl 0106.04003. http://dx.doi.org/10.1112/plms/s3-12.1.179.

[4] W. DUKE, J. FRIEDLANDER, and H. IWANIEC, Bounds for automorphic L-functions, *Invent. Math.* **112** (1993), 1–8. MR 1207474. Zbl 0765.11038. http://dx.doi.org/10.1007/BF01232422.

[5] D. GOLDFELD, *Automorphic Forms and L-Functions for the Group* GL(n, **R**), *Cambridge Stud. Adv. Math.* **99**, Cambridge Univ. Press, Cambridge, 2006, with an appendix by Kevin A. Broughan. MR 2254662. Zbl 1108.11039. http://dx.doi.org/10.1017/CBO9780511542923.

[6] R. Holowinsky, R. Munshi, and Z. Qi, Hybrid subconvexity bounds for
     $L(\frac{1}{2}, \mathrm{Sym}^2 f \otimes g)$, 2014, preprint. arXiv `1401.6695`.

[7] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, *Amer. Math. Soc.
     Colloq. Publ.* **53**, Amer. Math. Soc., Providence, RI, 2004. MR `2061214`.
     Zbl `1059.11001`.

[8] H. Jacquet, I. I. Piatetskii-Shapiro, and J. A. Shalika, Rankin-Selberg
     convolutions, *Amer. J. Math.* **105** (1983), 367–464. MR `0701565`. Zbl `0525.`
     `22018`. http://dx.doi.org/10.2307/2374264.

[9] X. Li, Bounds for GL(3) × GL(2) *L*-functions and GL(3) *L*-functions, *Ann. of
     Math.* **173** (2011), 301–336. MR `2753605`. Zbl `05960661`. http://dx.doi.org/10.
     4007/annals.2011.173.1.8.

[10] S. D. Miller, Cancellation in additively twisted sums on GL(n), *Amer. J. Math.*
     128 (2006), 699–729. MR `2230922`. Zbl `1142.11033`. http://dx.doi.org/10.1353/
     ajm.2006.0027.

[11] S. D. Miller and W. Schmid, Automorphic distributions, *L*-functions, and
     Voronoi summation for GL(3), *Ann. of Math.* **164** (2006), 423–488. MR `2247965`.
     Zbl `1162.11341`. http://dx.doi.org/10.4007/annals.2006.164.423.

[12] R. Munshi, Bounds for twisted symmetric square *L*-functions - II, unpublished.

[13] R. Munshi, Bounds for twisted symmetric square *L*-functions, *J. Reine Angew.
     Math.* **682** (2013), 65–88. MR `3181499`. Zbl `6221004`.

[14] R. Munshi, Bounds for twisted symmetric square *L*-functions - III, *Adv. Math.*
     235 (2013), 74–91. MR `3010051`. Zbl `1271.11055`. http://dx.doi.org/10.1016/j.
     aim.2012.11.010.

[15] R. Munshi, The circle method and bounds for *L*-functions—III. *t*-aspect sub-
     convexity for GL(3) *L*-functions, 2013. arXiv `1301.1007`.

[16] R. Munshi, The circle method and bounds for *L*-functions - I, *Math. Ann.*
     358 (2014), 389–401. MR `3158002`. Zbl `06269821`. http://dx.doi.org/10.1007/
     s00208-013-0968-4.

[17] R. Munshi, The circle method and bounds for *L*-functions, II: Subconvexity
     and twists of GL(3) *L*-functions, *Amer. J. Math.* **137** (2015), 791–812. http:
     //dx.doi.org/10.1353/ajm.2015.0018.

[18] R. Munshi, Hybrid subconvexity for Rankin-Selberg *L*-functions, preprint.

Tata Institute of Fundamental Research, Colaba, Mumbai, India
E-mail: rmunshi@math.tifr.res.in