



COMPOSITIO MATHEMATICA

On the fourth moment of Hecke–Maass forms and the random wave conjecture

Jack Buttcane and Rizwanur Khan

Compositio Math. **153** (2017), 1479–1511.

[doi:10.1112/S0010437X17007199](https://doi.org/10.1112/S0010437X17007199)



FOUNDATION
COMPOSITIO
MATHEMATICA



LONDON
MATHEMATICAL
SOCIETY
EST. 1865



On the fourth moment of Hecke–Maass forms and the random wave conjecture

Jack Buttcane and Rizwanur Khan

ABSTRACT

Conditionally on the generalized Lindelöf hypothesis, we obtain an asymptotic for the fourth moment of Hecke–Maass cusp forms of large Laplacian eigenvalue for the full modular group. This lends support to the random wave conjecture.

1. Introduction

A central question in arithmetic quantum chaos concerns the distribution of Hecke–Maass forms for the full modular group. The random wave conjecture (RWC) predicts that for large Laplacian eigenvalue, the distribution is close to random. One way to formulate this is to conjecture that on fixed compact sets, the moments of Hecke–Maass cusp forms of large Laplacian eigenvalue asymptotically equal the moments of a random variable with Gaussian distribution. Until now, only numerical work and heuristic arguments (see [HR92, HS01]) have supported this conjecture. On the theoretical side, the fourth moment in particular is a natural and important case to study, as it reduces, via Watson’s formula, to a problem on L -functions. But proving an asymptotic for the fourth moment of Hecke–Maass forms seems to be beyond the reach of current technology. The goal of this paper is to establish such an asymptotic (over $\Gamma \backslash \mathbb{H}$) on the assumption of the generalized Lindelöf hypothesis (GLH). One may wonder what the benefit is of proving one conjecture based on another. One answer is that these conjectures are unrelated. Another answer is of course that the GLH is a much more well accepted conjecture in mathematics and its truth is very firmly believed. This puts the RWC on more solid ground. Our main result is the following.

THEOREM 1.1. *Assume the GLH. Let f be an even or odd Hecke–Maass cusp form for $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ with Laplacian eigenvalue $\lambda = \frac{1}{4} + T^2$, where $T > 0$. Let f be normalized as follows:*

$$\frac{1}{\int_{\Gamma \backslash \mathbb{H}} 1(dx dy/y^2)} \int_{\Gamma \backslash \mathbb{H}} f(z)^2 \frac{dx dy}{y^2} = 1. \quad (1.1)$$

There exists a constant $\delta > 0$ such that

$$\frac{1}{\int_{\Gamma \backslash \mathbb{H}} 1(dx dy/y^2)} \int_{\Gamma \backslash \mathbb{H}} f(z)^4 \frac{dx dy}{y^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^4 e^{-t^2/2} dt + O(T^{-\delta}) \quad (1.2)$$

as $T \rightarrow \infty$.

Received 28 June 2016, accepted in final form 27 February 2017, published online 2 May 2017.

2010 Mathematics Subject Classification 11F12, 11M99 (primary), 81Q50 (secondary).

Keywords: L -functions, automorphic forms, L^4 -norm, quantum chaos, random wave conjecture.

The first author was supported by National Science Foundation grant DMS-1601919.

This journal is © Foundation Compositio Mathematica 2017.

Our result confirms on GLH a prediction of the RWC, with a power saving. Note that f is real-valued because it is assumed to be even or odd (and of weight 0, with trivial nebentypus). The left-hand side of (1.2) will be studied by using Watson’s formula to relate it to a mean value of L -functions. If the GLH were applied at this point, it would immediately yield the upper bound $O(T^\epsilon)$. To go beyond this and obtain an asymptotic with power saving, even on GLH, requires a lot of work and the full power of spectral theory. Some care is also needed to avoid reliance on the Ramanujan conjecture (both at the finite and infinite places). Our proof will show that if not GLH, then at the very least what is required is a subconvexity bound for the L -function associated to the Rankin–Selberg product of an (essentially) fixed Hecke cusp form and the symmetric square lift of f , in the eigenvalue aspect of f . See the discussion following Lemma 5.1 and the last sentence of this paper. This $GL(2) \times GL(3)$ subconvexity problem seems to be very difficult and, interestingly, is also essentially what is required to get a power saving error in the quantum unique ergodicity (QUE) problem. It is safe to say that our problem is more difficult than the quantitative QUE problem.

There have been some unconditional results short of an asymptotic for the fourth moment of automorphic forms on $\Gamma \backslash \mathbb{H}$. In the eigenvalue aspect, an essentially optimal upper bound was proven by Spinu [Spi03] for the fourth moment of truncated Eisenstein series, and by Luo [Luo14] for dihedral Maass forms. Both of these results hinged on the spectral large sieve. For Hecke–Maass forms, Sarnak and Watson announced a sharp upper bound for the fourth moment in [Sar03, Theorem 3], but a proof of this has not appeared. Holomorphic Hecke cusp forms of large weight are expected to be modelled by a complex Gaussian distribution (see [Blo13, Conjecture 1.2]). However, proving an asymptotic for the fourth moment in this case, on GLH or any other reasonable hypothesis, seems to be much harder than the problem for Maass forms. This is because the corresponding mean value of L -functions has a larger ‘log of conductor to log of family size ratio’ (see the discussion below). The best known upper bound for the fourth moment in the weight aspect is far from optimal; see [Blo13].

Let $\{u_j : j \geq 1\}$ denote an orthonormal basis of even and odd Hecke–Maass cusp forms for Γ , ordered by Laplacian eigenvalue $\frac{1}{4} + t_j^2$, where $t_j > 1$. Let $L(s, u_j)$ be the L -function attached to u_j , normalized so that its functional equation relates values at s and $1 - s$. The shape of the mean value of L -functions that we will need to evaluate is essentially

$$\sum_{t_j < 2T} \frac{1}{t_j T^{1/2} (1 + 2T - t_j)^{1/2}} \frac{L(1/2, u_j) L(1/2, u_j \times \text{sym}^2 f)}{L(1, \text{sym}^2 u_j)}. \tag{1.3}$$

The analytic conductors of $L(\frac{1}{2}, u_j)$ and $L(\frac{1}{2}, u_j \times \text{sym}^2 f)$ are t_j^2 and $t_j^2(1 + |4T^2 - t_j^2|)^2$ respectively. Thus, the denominator above is about the same size as the convexity bound for the numerator. In the ‘bulk’ range $T^{1-\epsilon} < t_j < 2T - T^{1-\epsilon}$, which is nearly a dyadic interval, the analytic conductor of the triple product L -function $L(\frac{1}{2}, u_j) L(\frac{1}{2}, u_j \times \text{sym}^2 f)$ is about T^8 , while the sum is over about T^2 forms. Thus the ‘log of conductor to log of family size ratio’ is 4. Our main work will be on treating this bulk range. The remaining ranges will be handled immediately on the GLH.

A similar mean value of triple product L -functions was considered (unconditionally) by Li in [Li09], but there the $GL(3)$ form was fixed, while here it is not (T tends to infinity). Thus, our problem is clearly more complex. A similar mean value was also considered by the authors in [BK17], with $L(\frac{1}{2}, u_j \times \text{sym}^2 f)$ replaced by $L(\frac{1}{2}, u_j \times \chi) L(u_j \times f)$, where χ is a quadratic character. Such a factorization occurs when f is a dihedral form, and this was the motivation for the work in [BK17], although in that paper we were not able to make any direct conclusions about the fourth moment. In this paper, we use the methods of [BK17] together with $GL(3)$

Voronoi summation as a new ingredient to treat the present case where $L(\frac{1}{2}, u_j \times \text{sym}^2 f)$ does not factorize. The present analysis is more delicate, with the ranges of various parameters harder to control (in the same way that many problems in number theory involving the divisor function become more difficult when the divisor function is replaced by Fourier coefficients of cusp forms). For this reason, it is not clear *a priori* that our previous methods would work for this problem. Also, [BK17] used a simplified weight function in place of the one given in (2.3).

2. Reduction to L -functions

Convention. Throughout, ϵ will denote an arbitrarily small positive constant, which may not be the same from one occurrence to another.

To prove Theorem 1.1, the starting point is to express the fourth moment as a mean value of L -functions. Note that since $\int_{\Gamma \backslash \mathbb{H}} 1(dx dy/y^2) = \pi/3$ and $(1/\sqrt{2\pi}) \int_{-\infty}^{\infty} t^4 e^{-t^2/2} dt = 3$, we need to show that with the normalization $\langle f, f \rangle = \langle f^2, 1 \rangle = \pi/3$, where the inner product is the Petersson inner product, we have

$$\langle f^2, f^2 \rangle = \pi + O(T^{-\delta}).$$

By the spectral theorem (see [IK04, Theorem 15.5]) and Parseval’s theorem, we have

$$\langle f^2, f^2 \rangle = \left| \left\langle f^2, \left(\frac{3}{\pi}\right)^{1/2} \right\rangle \right|^2 + \frac{1}{4\pi} \int_{-\infty}^{\infty} \left| \left\langle f^2, E\left(\cdot, \frac{1}{2} + it\right) \right\rangle \right|^2 dt + \sum_{j \geq 1} |\langle f^2, u_j \rangle|^2,$$

where $E(z, s)$ is the standard Eisenstein series. By normalization of f , we have

$$\left| \left\langle f^2, \left(\frac{3}{\pi}\right)^{1/2} \right\rangle \right|^2 = \frac{\pi}{3},$$

and we will show the following result.

LEMMA 2.1. *Let f be as in Theorem 1.1. On the GLH, we have*

$$\int_{-\infty}^{\infty} \left| \left\langle f^2, E\left(\cdot, \frac{1}{2} + it\right) \right\rangle \right|^2 dt \ll T^{-1+\epsilon}. \tag{2.1}$$

Thus it remains to prove the following proposition.

PROPOSITION 2.2. *Let f be as in Theorem 1.1. On the GLH, there exists some $\delta > 0$ such that*

$$\sum_{j \geq 1} |\langle f^2, u_j \rangle|^2 = \frac{2\pi}{3} + O(T^{-\delta}). \tag{2.2}$$

Since $\sqrt{(3/\pi)}f$ has L^2 -norm equal to 1, applying Watson’s formula [Wat02, Theorem 3] (see also [Blo13, p. 2624]) to the inner product of $(3/\pi)f^2$ and u_j , we get

$$\left| \left\langle \frac{3}{\pi} f^2, u_j \right\rangle \right|^2 = \frac{\pi}{2^3} H(t_j) \frac{L(1/2, f \times f \times \bar{u}_j)}{L(1, \text{sym}^2 f)^2 L(1, \text{sym}^2 u_j)},$$

where the L -functions appearing above are defined in the next section and

$$H(t) = \frac{|\Gamma((1/2 + 2iT + it)/2)|^2 |\Gamma((1/2 + 2iT - it)/2)|^2 |\Gamma((1/2 + it)/2)|^4}{|\Gamma((1 + 2iT)/2)|^4 |\Gamma((1 + 2it)/2)|^2}. \tag{2.3}$$

Equivalently, as u_j is real,

$$|\langle f^2, u_j \rangle|^2 = \frac{\pi^3}{72} H(t_j) \frac{L(1/2, u_j) L(1/2, u_j \times \text{sym}^2 f)}{L(1, \text{sym}^2 f)^2 L(1, \text{sym}^2 u_j)}. \tag{2.4}$$

For the weight function $H(t_j)$, we have by Stirling’s approximation (see (3.1) or [Spi03, § 5.1.1]) that

$$H(t) = \frac{8\pi \exp(-\pi q(t, T))}{(1 + |t|) \prod_{\pm} (1 + |2T \pm t|)^{1/2}} \left\{ 1 + O\left(\frac{1}{1 + |t|} + \frac{1}{1 + |2T + t|} + \frac{1}{1 + |2T - t|} \right) \right\}, \tag{2.5}$$

where

$$q(t, T) = \begin{cases} 0 & \text{for } |t| \leq 2T, \\ |t| - 2T & \text{for } |t| > 2T. \end{cases} \tag{2.6}$$

Thus, the right-hand side of (2.4) looks essentially like (1.3).

3. Preliminaries

3.1 Stirling’s approximation

For $\sigma > 0$ fixed, as a first-order approximation we have

$$\begin{aligned} \Gamma(\sigma + i\gamma) &= \sqrt{2\pi} |\sigma + i\gamma|^{\sigma-1/2+i\gamma} \\ &\quad \times \exp\left(-\frac{\pi}{2} |\gamma| + i \text{sgn}(\gamma) \left(\frac{\pi}{2} \left(\sigma - \frac{1}{2}\right) - \gamma\right)\right) (1 + O((1 + |\gamma|)^{-1})), \\ |\Gamma(\sigma + i\gamma)| &= \sqrt{2\pi} (1 + |\gamma|)^{\sigma-1/2} \exp\left(-\frac{\pi}{2} |\gamma|\right) (1 + O((1 + |\gamma|)^{-1})), \end{aligned} \tag{3.1}$$

where $\text{sgn}(\gamma)$ is 1 if γ is positive and -1 if γ is negative. As $|\gamma| \rightarrow \infty$, this gives

$$\Gamma(\sigma + i\gamma) = \sqrt{2\pi} |\gamma|^{\sigma-1/2+i\gamma} \exp\left(-\frac{\pi}{2} |\gamma| + i \text{sgn}(\gamma) \left(\frac{\pi}{2} \left(\sigma - \frac{1}{2}\right) - \gamma\right)\right) (1 + O(|\gamma|^{-1})).$$

3.2 Approximate functional equations

Let $\lambda_j(n)$ and $\lambda_f(n)$ denote the (real) eigenvalues of the n th Hecke operator corresponding to u_j and f respectively, where we write $\lambda_j(-n) = \lambda_j(n)$ for u_j even and $\lambda_j(-n) = -\lambda_j(n)$ for u_j odd. The L -function associated to u_j is given by

$$L(s, u_j) = \sum_{n \geq 1} \frac{\lambda_j(n)}{n^s}$$

for $\Re(s) > 1$. Let $A_f(n, 1) = A_f(1, n)$ be given by

$$\sum_{n \geq 1} \frac{A_f(n, 1)}{n^s} := \zeta(2s) \sum_{n \geq 1} \frac{\lambda_f(n^2)}{n^s}.$$

The right-hand side above equals $L(s, \text{sym}^2 f)$ for $\Re(s) > 1$. Now define $A_f(n, m) = A_f(m, n)$ by the Hecke relations

$$A_f(n, m) = \sum_{v|(n,m)} \mu(v) A_f\left(\frac{n}{v}, 1\right) A_f\left(\frac{m}{v}, 1\right). \tag{3.2}$$

With this, we can define

$$L(s, u_j \times \text{sym}^2 f) = \sum_{m,r \geq 1} \frac{\lambda_j(m) A_f(m, r)}{(r^2 m)^s}$$

for $\Re(s) > 1$.

Kim and Sarnak [Kim03, Appendix 2] have proven the following bounds towards the Ramanujan conjecture:

$$\begin{aligned} |\lambda_j(n)| &\ll n^{7/64+\epsilon}, \\ |A_f(n, 1)| &\ll n^{7/32+\epsilon}. \end{aligned}$$

On average, the best possible bounds are known by [Iwa92, Lemma 1] and [HL94, Lemma 2.2]:

$$\begin{aligned} \sum_{n \leq x} |\lambda_j(n)|^2 &\ll x(|t_j|x)^\epsilon, \\ \sum_{n \leq x} |A(n, 1)|^2 &\ll x(Tx)^\epsilon. \end{aligned} \tag{3.3}$$

This implies by the Cauchy–Schwarz inequality and the Hecke relations (3.2) that

$$\begin{aligned} \sum_{n \leq x} |\lambda_j(n)| &\ll x(|t_j|x)^\epsilon, \\ \sum_{\substack{n \leq x \\ m \leq y}} |A(n, m)| &\ll xy(Txy)^\epsilon. \end{aligned} \tag{3.4}$$

Let

$$\begin{aligned} \Gamma_{\mathbb{R}}(s) &= \pi^{-s/2} \Gamma\left(\frac{s}{2}\right), \\ G_1(s) &= \prod_{\pm} \Gamma_{\mathbb{R}}(s \pm it_j), \\ G_2(s) &= \prod_{\pm} \Gamma_{\mathbb{R}}(s \pm it_j + i2T) \Gamma_{\mathbb{R}}(s \pm it_j) \Gamma_{\mathbb{R}}(s \pm it_j - i2T). \end{aligned}$$

For u_j even we have the functional equations

$$\begin{aligned} L(s, u_j) G_1(s) &= L(1 - s, u_j) G_1(1 - s), \\ L(s, u_j \times \text{sym}^2 f) G_2(s) &= L(1 - s, u_j \times \text{sym}^2 f) G_2(1 - s). \end{aligned}$$

For u_j odd we have the functional equation

$$L(s, u_j) G_1(1 + s) = -L(1 - s, u_j) G_1(2 - s),$$

which implies that $L(\frac{1}{2}, u_j) = 0$. All of these may be found in [Gol06, ch. 3] and [Li09, p. 1670].

We now set up approximate functional equations for the central values. But first we explain what we will need. Usually one takes an approximate functional equation with the shortest possible Dirichlet series. For u_j even, this means taking

$$L\left(\frac{1}{2}, u_j\right) \sim \sum_{n < t_j^{1+\epsilon}} \frac{\lambda_j(n)}{n^{1/2}} + \sum_{n < t_j^{1+\epsilon}} \frac{\lambda_j(n)}{n^{1/2}},$$

and for u_j odd, we could take

$$L\left(\frac{1}{2}, u_j\right) \sim \sum_{n < t_j^{1+\epsilon}} \frac{\lambda_j(n)}{n^{1/2}} - \sum_{n < t_j^{1+\epsilon}} \frac{\lambda_j(n)}{n^{1/2}},$$

which vanishes. Since $\lambda_j(-n) = \lambda_j(n)$ for u_j even and $\lambda_j(-n) = -\lambda_j(n)$ for u_j odd, in both cases we have

$$L\left(\frac{1}{2}, u_j\right) \sim \sum_{n < t_j^{1+\epsilon}} \frac{\lambda_j(n)}{n^{1/2}} + \sum_{n < t_j^{1+\epsilon}} \frac{\lambda_j(-n)}{n^{1/2}}.$$

As we will see below, to understand a mean value of the form $\sum_{T < t_j < 2T} \lambda_j(\pm n)\lambda_j(m)$ using Kuznetsov’s formula, the same sign Kuznetsov (the + sign) leads to a J -Bessel transform while the opposite sign Kuznetsov (the – sign) leads to a K -Bessel transform. Both transforms can be evaluated asymptotically and it turns out that the main term of the K -Bessel transform has no oscillation. We find this easier to work with, so we reduce the analysis involving the same sign terms by taking an approximate functional equation with two Dirichlet series of unequal length as follows:

$$L\left(\frac{1}{2}, u_j\right) \sim \sum_{n < t_j^{1+\epsilon} T^{-\beta}} \frac{\lambda_j(n)}{n^{1/2}} + \sum_{n < t_j^{1+\epsilon} T^\beta} \frac{\lambda_j(-n)}{n^{1/2}}.$$

In this way, the Dirichlet series leading to the same sign terms is shorter. We now state this precisely.

LEMMA 3.1. *For u_j even, we have*

$$L\left(\frac{1}{2}, u_j\right) = 2 \sum_{n \geq 1} \frac{\lambda_j(n)}{n^{1/2}} V_1(n, t_j), \tag{3.5}$$

$$L\left(\frac{1}{2}, u_j \times \text{sym}^2 f\right) = 2 \sum_{m, r \geq 1} \frac{\lambda_j(m) A_f(r, m)}{r m^{1/2}} V_2(r^2 m, t_j), \tag{3.6}$$

where

$$V_i(x, t) = \frac{1}{2\pi i} \int_{(\sigma)} e^{s^2 x^{-s}} \frac{G_i(1/2 + s)}{G_i(1/2)} \frac{ds}{s} \tag{3.7}$$

for any $\sigma > 0$.

Let $0 < \alpha, \beta < \frac{1}{100}$ be some fixed constants to be determined later. For u_j even or odd, and $T^{1-\alpha} < |t_j| < T^{1+\epsilon}$, we have

$$L\left(\frac{1}{2}, u_j\right) = \sum_{\pm} \sum_{n \geq 1} \frac{\lambda_j(\pm n)}{n^{1/2}} V_1^\pm(n, t_j) + O(T^{-1/2+\beta/2+\alpha+\epsilon}), \tag{3.8}$$

where

$$V_1^\pm(x, t) = \frac{1}{2\pi i} \int_{(\sigma)} e^{s^2 (xT^{\pm\beta})^{-s}} \frac{G_1(1/2 + s)}{G_1(1/2)} \frac{ds}{s}$$

for any $\sigma > 0$.

Proof. This follows from [IK04, Theorem 5.3] and the functional equations given above. For (3.5) and (3.6), take $G(u) = e^{u^2}$ and $X = 1$ in that theorem. For (3.8), take $G(u) = e^{u^2}$ and $X = T^\beta$ to get

$$L\left(\frac{1}{2}, u_j\right) = \sum_{\pm} \sum_{n \geq 1} \frac{\lambda_j(\pm n)}{n^{1/2}} \frac{1}{2\pi i} \int_{(\sigma)} e^{s^2} (xT^{\pm\beta})^{-s} \frac{G_1(1/2 + \kappa_j + s)}{G_1(1/2 + \kappa_k)} \frac{ds}{s}, \tag{3.9}$$

where $\kappa_j = 0$ or 1 as u_j is even or odd. By the rapid decay of e^{s^2} in vertical lines, we may restrict the integral above to $|\Im(s)| < T^\epsilon$. By Stirling’s approximation, for $\Re(s) > 0$ fixed, $|\Im(s)| < T^\epsilon$ and $|t| > T^{1-\alpha}$, we have

$$\frac{G_1(1/2 + 1 + s)}{G_1(1/2 + 1)} = \frac{G_1(1/2 + s)}{G_1(1/2)} + O(T^{-1+\alpha+\epsilon}). \tag{3.10}$$

Thus, up to a small error, the ratio of Gamma functions in (3.9) does not depend on κ_j . Also note that the sum in (3.9) can be restricted to $n < T^{1+\beta+\epsilon}$ up to admissible error by Stirling’s approximation. Thus (3.8) follows, using (3.4). \square

Consider those values of $|t|$ that are roughly of size $2T$ but not too close to $2T$. That is, suppose that for some $0 < \alpha < \frac{1}{100}$ to be fixed later, we have

$$T^{1-\alpha} < |t| < 2T - T^{1-\alpha}. \tag{3.11}$$

In the integrals appearing in Lemma 3.1, write $s = \sigma + i\gamma$. By the rapid decay of e^{s^2} in vertical lines, we may restrict these integrals to $|\gamma| < T^\epsilon$. We have by Stirling’s approximation that

$$V_1^\pm(x, t) = \frac{1}{2\pi i} \int_{(\sigma)} e^{s^2} \left(\frac{2\pi x T^{\pm\beta}}{|t|}\right)^{-s} \left(1 + \sum_{n \leq N} \frac{B_n(\sigma, \gamma)}{|t|^n}\right) \frac{ds}{s} + O(T^{-N/2}) \tag{3.12}$$

and

$$\begin{aligned} V_2(x, t) &= \frac{1}{2\pi i} \int_{(\sigma)} e^{s^2} \left(\frac{8\pi^3 x}{|t(4T^2 - t^2)|}\right)^{-s} \left(1 + \sum_{n \leq N} \frac{B_n(\sigma, \gamma)}{|t|^n}\right) \left(1 + \sum_{n \leq N} \frac{C_n(\sigma, \gamma)}{|2T + t|^n}\right) \\ &\quad \times \left(1 + \sum_{n \leq N} \frac{C_n(\sigma, \gamma)}{|2T - t|^n}\right) \frac{ds}{s} + O(T^{-N/2}) \end{aligned} \tag{3.13}$$

for any $N \geq 1$ and some $B_n(\sigma, \gamma)$ and $C_n(\sigma, \gamma)$ polynomial in σ and γ . By thinking of $|t|$ and $|t(4T^2 - t^2)|$ as being of size about T and T^3 in the range (3.11), and taking σ as large as we like in the expressions (3.12)–(3.13), we see that the sums in (3.8) and (3.6) have length about $T^{1\mp\beta}$ and T^3 respectively. We have

$$\frac{\partial^k}{\partial t^k} V_1^\pm(x, t) \ll |t|^{-k+\epsilon} \ll T^{k(-1+\alpha+\epsilon)} \tag{3.14}$$

in the range (3.11) by taking N large enough and $\sigma = \epsilon$. Similarly,

$$\frac{\partial^k}{\partial t^k} V_2(x, t) \ll |t(4T^2 - t^2)|^{-k+\epsilon} \ll T^{k(-3+2\alpha+\epsilon)} \tag{3.15}$$

for $k \geq 0$.

3.3 Kuznetsov trace formula

Let

$$\lambda(n, t) = \sum_{ab=n} \left(\frac{a}{b}\right)^{it}$$

and

$$\mathcal{F}^+(x, t) = \frac{2i}{\sinh(\pi t)} J_{2it}(4\pi x), \quad \mathcal{F}^-(x, t) = \frac{4}{\pi} K_{2it}(4\pi x) \cosh(\pi t).$$

We have

LEMMA 3.2. *Let $h(z)$ be an even, holomorphic function on $|\Im(z)| < \frac{1}{4} + \theta$ with decay $|h(z)| \ll (1 + |z|)^{-2-\theta}$ on that strip, for some $\theta > 0$. Then for $n, m > 0$, we have*

$$\begin{aligned} & \sum_{j \geq 1} \frac{\lambda_j(\pm n)\lambda_j(m)}{L(1, \text{sym}^2 u_j)} h(t_j) + \int_{-\infty}^{\infty} \frac{\lambda(n, t)\lambda(m, -t)}{|\zeta(1 + 2it)|^2} h(t) \frac{dt}{2\pi} \\ &= \delta_{\pm n, m} \int_{-\infty}^{\infty} h(t) \frac{d^*t}{2\pi^2} + \sum_{c \geq 1} \frac{S(\pm n, m, c)}{c} \int_{-\infty}^{\infty} \mathcal{F}^{\pm}\left(\frac{\sqrt{nm}}{c}, t\right) h(t) \frac{d^*t}{2\pi}, \end{aligned} \tag{3.16}$$

where $\delta_{\pm n, m}$ is 1 if $n = m$ and 0 otherwise (thus $\delta_{-n, m}$ is always 0) and $d^*t = \tanh(\pi t)t dt$.

Proof. See [Mot97, Theorems 2.2 and 2.4]. There, the function $h(z)$ must be a holomorphic function on $|\Im(z)| < \frac{1}{2} + \theta$. The relaxation of this condition to $|\Im(z)| < \frac{1}{4} + \theta$ is due to Yoshida [Yos97]. We need this version because $H(t)$, which was defined in § 2, has a pole at $t = \frac{1}{2}i$. \square

For the Kuznetsov trace formula, we will need to asymptotically evaluate some Bessel transforms. The following lemmas, taken from [BK17, § 3.8] together with the correction noted after Lemma 3.4 below, are analogous to the averages of real Bessel functions given in [ILS00, Corollary 8.2]:

$$\begin{aligned} & \sum_{k \equiv 0 \pmod 2} i^k J_{k-1}(x) h\left(\frac{k}{K}\right) \sim \text{oscillatory function supported on } x \gg K^{2-\epsilon}, \\ & \sum_{k \equiv 0 \pmod 2} J_{k-1}(x) h\left(\frac{k}{K}\right) \sim \text{non-oscillatory function supported on } x \asymp K. \end{aligned}$$

LEMMA 3.3. *Let $0 < \alpha < \frac{1}{100}$. For any $x > 0$ and any smooth, even function h compactly supported on $(T^{-\alpha}, T^{\alpha}) \cup (-T^{\alpha}, -T^{-\alpha})$ with derivatives satisfying $\|h^{(k)}\|_{\infty} \ll (T^{\alpha})^k$, we have*

$$\begin{aligned} & \int_{-\infty}^{\infty} \frac{J_{2it}(2\pi x)}{\cosh(\pi t)} h\left(\frac{t}{T}\right) t dt \\ &= \frac{-i\sqrt{2}}{\pi} \frac{T^2}{\sqrt{x}} \Re\left((1+i)\epsilon(x) \int_0^{\infty} th(t)e\left(\frac{-t^2 T^2}{2\pi^2 x}\right) dt \right) + O\left(\frac{x}{T^{3-12\alpha}}\right) + O(T^{-100}). \end{aligned} \tag{3.17}$$

The main term is $O(T^{-100})$ if $x < T^{2-3\alpha}$.

LEMMA 3.4. *Let $0 < \alpha < \frac{1}{100}$. For any $0 < x \leq T^3$ and any smooth even function h , compactly supported on $(T^{-\alpha}, T^{\alpha}) \cup (-T^{\alpha}, -T^{-\alpha})$ with derivatives satisfying $\|h^{(k)}\|_{\infty} \ll (T^{\alpha})^k$,*

we have

$$\int_{-\infty}^{\infty} \sinh(\pi t) K_{2it}(2\pi x) h\left(\frac{t}{T}\right) t dt = \frac{\pi T}{2} \hbar\left(\frac{\pi x}{T}\right) - \frac{i\pi^3}{12T} \hbar^{(3)}\left(\frac{\pi x}{T}\right) + O\left(\frac{x}{T^{4-14\alpha}}\right) + O\left(\frac{x^2}{T^{5-16\alpha}}\right) + O(T^{-100}), \tag{3.18}$$

where $\hbar(y) = yh(y)$.

Note that in [BK17, Lemma 3.8], the error term $O(x^2/T^{5-16\alpha})$ has been erroneously left out. Such a term should be present, as it is for the average of real Bessel functions (see the remark following [Iwa97, Lemma 5.8]). For our purposes we will have $0 < x \leq T^{2+4\alpha}$, so the total error for Lemma 3.4 will be essentially the same as for Lemma 3.3.

3.4 Kuznetsov’s formula for sums of Kloosterman sums

Let Φ be a smooth function compactly supported on the positive reals.

Let

$$\begin{aligned} \hat{\Phi}(k) &= i^k \int_0^\infty J_{k-1}(w) \Phi(w) \frac{dw}{w}, \\ \hat{\Phi}(t) &= \frac{i}{2 \sinh(\pi t)} \int_0^\infty (J_{2it}(w) - J_{-2it}(w)) \Phi(w) \frac{dw}{w}, \\ \check{\Phi}(t) &= \frac{2}{\pi} \cosh(\pi t) \int_0^\infty K_{2it}(w) \Phi(w) \frac{dw}{w}. \end{aligned}$$

For $q \geq 1$ and $k \geq 2$, let $\mathcal{B}_k(q)$ denote an orthonormal basis of weight k holomorphic cusp forms for $\Gamma_0(q)$. For each element g in this basis, let $n^{k/2} \rho_g(n)$ denote the n th Fourier coefficient of g . Let $\mathcal{B}_0(q)$ denote an orthonormal basis of Maass cusp forms for $\Gamma_0(q)$. For each element g in this basis, let $\rho_g(n)$ denote the n th Fourier coefficient of g , and let $\frac{1}{4} + t_g^2$ denote its Laplacian eigenvalue. By a result of Kim and Sarnak [Kim03, Appendix 2] towards Selberg’s eigenvalue conjecture, we have that

$$t_g \in \mathbb{R} \cup \left(-\frac{7}{64}i, \frac{7}{64}i\right). \tag{3.19}$$

Let $\tau_a(n, t)$ denote the n th Fourier coefficient of the Eisenstein series $E_a(s, \frac{1}{2} + it)$ at the cusp \mathfrak{a} of $\Gamma_0(q)$.

LEMMA 3.5. *Keep the notation above and let $q \geq 1$. For positive integers n and m , we have*

$$\begin{aligned} \sum_{c \geq 1} \frac{S(n, m, qc)}{qc} \Phi\left(\frac{4\pi\sqrt{nm}}{c}\right) &= \sum_{\substack{k \geq 2 \\ g \in \mathcal{B}_k(q)}} \hat{\Phi}(k) \frac{(k-1)! \sqrt{nm}}{\pi(4\pi)^{k-1}} \rho_g(n) \overline{\rho_g(m)} \\ &+ \sum_{g \in \mathcal{B}_0(q)} \hat{\Phi}(t_g) \frac{4\pi\sqrt{nm}}{\cosh(\pi t_g)} \rho_g(n) \overline{\rho_g(-m)} \\ &+ \sum_{\mathfrak{a}} \int_{-\infty}^{\infty} \hat{\Phi}(t) \frac{\sqrt{nm}}{\cosh(\pi t)} \tau_{\mathfrak{a}}(n, t) \overline{\tau_{\mathfrak{a}}(-m, t)} dt, \end{aligned}$$

and

$$\begin{aligned} \sum_{c \geq 1} \frac{S(n, -m, qc)}{qc} \Phi\left(\frac{4\pi\sqrt{nm}}{c}\right) &= \sum_{g \in \mathcal{B}_0(q)} \check{\Phi}(t_g) \frac{4\pi\sqrt{nm}}{\cosh(\pi t_g)} \rho_g(n) \overline{\rho_g(m)} \\ &+ \sum_{\mathfrak{a}} \int_{-\infty}^{\infty} \check{\Phi}(t) \frac{\sqrt{nm}}{\cosh(\pi t)} \tau_{\mathfrak{a}}(n, t) \overline{\tau_{\mathfrak{a}}(-m, t)} dt. \end{aligned}$$

Proof. See [IK04, Theorem 16.5], but note that we normalize differently. □

We now record some properties of the transforms of Φ given above, based on the situation we will be in (see (9.18)). Essentially the same result may be found elsewhere; see, for example, [Blo07, Lemma 1].

LEMMA 3.6. *Suppose that $\Phi(w)$ is supported on*

$$X^{-1}T^{-\epsilon} < w < X^{-1}T^\epsilon \tag{3.20}$$

for some $T^{-\epsilon} < X < T^{10}$ and satisfies

$$\Phi^{(k)}(w) \ll YX^kT^\epsilon \tag{3.21}$$

for some $T^{-10} < Y < T^{10}$.

Case 1. If $t \in \mathbb{R}$ and $k \geq 2$ then

$$|\dot{\Phi}(k)|, |\hat{\Phi}(t)|, |\check{\Phi}(t)| \ll YT^\epsilon. \tag{3.22}$$

Unless

$$k, |t| < T^\epsilon,$$

we have

$$|\dot{\Phi}(k)|, |\hat{\Phi}(t)|, |\check{\Phi}(t)| \ll T^{-B}$$

for any $B > 0$.

Case 2. If $t \in (-\frac{7}{64}i, \frac{7}{64}i)$, then

$$|\hat{\Phi}(t)|, |\check{\Phi}(t)| \ll YX^{7/32}T^\epsilon. \tag{3.23}$$

Proof. We demonstrate the claims for $\hat{\Phi}$ only, the other cases being similar. Suppose first that $t \in \mathbb{R}$. Note the bound

$$\frac{J_{2it}(w)}{\sinh(\pi t)} \ll \min\{1, w^{-1/2}\}, \tag{3.24}$$

which follows for $0 < w < 1$ from the power series [GR00, 8.402]

$$J_\nu(w) = \sum_{n \geq 0} \frac{(-1)^n (w/2)^{2n+\nu}}{n! \Gamma(n + \nu + 1)} \tag{3.25}$$

and for $w \geq 1$ by [BM15, Lemma 6]. By (3.20), we may restrict the integral in the definition of $\hat{\Phi}$ to $T^{-11} < w < T^{\epsilon/3}$ and then apply (3.21) with $k = 0$ and (3.24) to get

$$|\hat{\Phi}(t)| \ll YT^\epsilon \int_{T^{-11}}^{T^{\epsilon/3}} \frac{dw}{w} \ll YT^\epsilon.$$

This proves (3.22). Now suppose that $|t| > T^\epsilon$. By the power series (3.25), we have

$$\hat{\Phi}(t) = \frac{i}{2 \sinh(\pi t)} \int_{T^{-11}}^{T^{\epsilon/3}} \Phi(w) \left(\sum_{n \geq 0} \frac{(-1)^n (w/2)^{2n+2it}}{n! \Gamma(n + 2it + 1)} - \sum_{n \geq 0} \frac{(-1)^n (w/2)^{2n-2it}}{n! \Gamma(n - 2it + 1)} \right) \frac{dw}{w}.$$

Integrating by parts k times, we get

$$\hat{\Phi}(t) = \frac{i}{2 \sinh(\pi t)} \int_{T^{-11}}^{T^{\epsilon/3}} \frac{d^k}{dw^k} \left(\frac{\Phi(w)}{w} \right) \sum_{n \geq 0} \frac{(-1)^n (w/2)^{2n+2it} w^k}{n! \Gamma(n+2it+1) \prod_{j=1}^k (2n+2it+j)} dw - \frac{i}{2 \sinh(\pi t)} \int_{T^{-11}}^{T^{\epsilon/3}} \frac{d^k}{dw^k} \left(\frac{\Phi(w)}{w} \right) \sum_{n \geq 0} \frac{(-1)^n (w/2)^{2n-2it} w^k}{n! \Gamma(n-2it+1) \prod_{j=1}^k (2n-2it+j)} dw. \tag{3.26}$$

By Stirling’s approximation, we have for $|t| > T^\epsilon$ and $0 < w < T^{\epsilon/3}$ that

$$\frac{w^{2n}}{\sinh(\pi t) \Gamma(n+1 \pm 2it)} \ll \frac{w^{2n}}{|n+1 \pm 2it|^{n+1/2}} \ll 1.$$

By (3.20), we have that

$$w^k \frac{d^k}{dw^k} \left(\frac{\Phi(w)}{w} \right) \ll Y X T^\epsilon$$

for any $k \geq 0$. Using these bounds in (3.26) and taking k large, we see that $|\hat{\Phi}(t)| \ll |t|^{-B}$ for any $B > 0$ unless $|t| < T^\epsilon$.

Now suppose $t \in (-\frac{7}{64}i, \frac{7}{64}i)$. By [BM15, Lemma 6] and (3.25), we have that

$$\frac{J_{2it}(w)}{\sinh(\pi t)} \ll \min\{T^\epsilon X^{7/32}, w^{-1/2}\}$$

for w in the interval (3.20). This gives (3.23) by the same argument as above for (3.22). □

3.5 Orthonormal basis of newforms

The right-hand side of Lemma 3.5 involves sums over orthonormal bases of cusp forms. We will need these basis elements to be linear combinations of lifts of newforms. Let $S_k(q)$ denote for $k \geq 2$ the space of holomorphic cusp forms of weight $k \geq 2$ and level q , and for $k = 0$ the space of Maass cusp forms of level q . For $d|q$, let $\mathcal{B}_k^*(d)$ denote a basis for the space of newforms of $S_k(d)$, which is orthonormal with respect to the Petersson inner product on $S_k(q)$. For $h \in S_k(d)$ and $b|(q/d)$, let

$$h|_b(z) = b^{k/2} h(bz). \tag{3.27}$$

LEMMA 3.7. *There exists an orthonormal basis for $S_k(q)$ of the form*

$$\bigcup_{d|q} \bigcup_{h \in \mathcal{B}_k^*(d)} \left\{ h_c : c \mid \frac{q}{d} \right\},$$

where

$$h_c = \sum_{b|c} \kappa_f(c, b) h|_b$$

for some numbers $\kappa_f(c, b) \ll q^\epsilon$.

Proof. See [BM15, Lemma 9 and equation (5.6)]. This builds on [ILS00, § 2] and [Rou11]. □

3.6 Voronoi summation

The $GL(3)$ Voronoi summation formula was proven by Miller and Schmid [MS06]. Later, Goldfeld and Li [GL06] gave another proof, and we follow their presentation.

LEMMA 3.8. *Let ψ be a smooth, compactly supported function on the positive real numbers. Let $(b, c) = 1$, and let \bar{b} denote the multiplicative inverse of b modulo c . We have*

$$\begin{aligned} & \sum_{m \geq 1} A_f(m, r) e\left(\frac{m\bar{b}}{c}\right) \psi\left(\frac{mr^2}{M}\right) \\ &= \sum_{\pm} \frac{c}{2} \sum_{\substack{\mathfrak{k} \geq 1 \\ l|cr}} \frac{A_f(\mathfrak{k}, l)}{\mathfrak{k}l} S\left(rb, \pm\mathfrak{k}, \frac{cr}{l}\right) \frac{1}{2\pi i} \int_{(\sigma)} \left(\frac{\mathfrak{k}Ml^2}{c^3r^3}\right)^{1-s} G^\pm(s) \tilde{\psi}(1-s) ds, \end{aligned} \tag{3.28}$$

where $\sigma > 0$, $\tilde{\psi}$ denotes the Mellin transform of ψ and

$$\begin{aligned} G^\pm(s) &= \frac{\Gamma_{\mathbb{R}}(s + 2iT)\Gamma_{\mathbb{R}}(s)\Gamma_{\mathbb{R}}(s - 2iT)}{\Gamma_{\mathbb{R}}(1 - s - 2iT)\Gamma_{\mathbb{R}}(1 - s)\Gamma_{\mathbb{R}}(1 - s + 2iT)} \\ &\mp i \frac{\Gamma_{\mathbb{R}}(1 + s + 2iT)\Gamma_{\mathbb{R}}(1 + s)\Gamma_{\mathbb{R}}(1 + s - 2iT)}{\Gamma_{\mathbb{R}}(2 - s - 2iT)\Gamma_{\mathbb{R}}(2 - s)\Gamma_{\mathbb{R}}(2 - s + 2iT)}. \end{aligned}$$

Suppose that $\|\psi^{(k)}\|_\infty \ll (T^\epsilon)^k$. By integration by parts, we can see that $\tilde{\psi}(1-s) \ll (T^\epsilon)^N (1+|s|)^{-N}$ for any $N \geq 1$. By Stirling’s approximation, we have

$$|G^\pm(s)| \ll (|s + 2iT||s - 2iT|)^{\sigma-1/2}.$$

Thus, we may restrict the integral in (3.28) to $|\Im(s)| < T^\epsilon$. In this range, taking $\sigma = \epsilon$, we get the bound

$$|G^\pm(s)| \ll T^{-1+\epsilon}. \tag{3.29}$$

Moving the line of integration in (3.28) far to the right, we see that the sum on the right-hand side of (3.28) may be restricted to

$$\mathfrak{k} < \frac{c^3r^3T^{2+\epsilon}}{Ml^2},$$

up to an error of $O(T^{-100})$. We also observe that in the range $|\Im(s)| < T^\epsilon$, writing $s = \sigma + i\gamma$, Stirling’s approximation gives

$$\begin{aligned} & \frac{\Gamma_{\mathbb{R}}(s + 2iT)\Gamma_{\mathbb{R}}(s)\Gamma_{\mathbb{R}}(s - 2iT)}{\Gamma_{\mathbb{R}}(1 - s - 2iT)\Gamma_{\mathbb{R}}(1 - s)\Gamma_{\mathbb{R}}(1 - s + 2iT)} \\ &= \left(\frac{T}{\pi}\right)^{2s-1} \frac{\Gamma_{\mathbb{R}}(s)}{\Gamma_{\mathbb{R}}(1 - s)} \left(1 + \sum_{n \leq N} \frac{B_n(\sigma, \gamma)}{T^n} + O(T^{-N/2})\right), \tag{3.30} \\ & \frac{\Gamma_{\mathbb{R}}(1 + s + 2iT)\Gamma_{\mathbb{R}}(1 + s)\Gamma_{\mathbb{R}}(1 + s - 2iT)}{\Gamma_{\mathbb{R}}(2 - s - 2iT)\Gamma_{\mathbb{R}}(2 - s)\Gamma_{\mathbb{R}}(2 - s + 2iT)} \\ &= \left(\frac{T}{\pi}\right)^{2s-2} \frac{\Gamma_{\mathbb{R}}(1 + s)}{\Gamma_{\mathbb{R}}(2 - s)} \left(1 + \sum_{n \leq N} \frac{C_n(\sigma, \gamma)}{T^n} + O(T^{-N/2})\right) \end{aligned}$$

for any $N \geq 1$ and some $B_n(\sigma, \gamma)$ and $C_n(\sigma, \gamma)$ polynomial in σ and γ .

4. Proof of Lemma 2.1

It is more convenient to renormalize f so that $\|f\|_2 = 1$. This does affect what needs to be proved, for in Lemma 2.1 we need only an upper bound. So let

$$f(x + iy) = \rho_f(1) \sum_{n \neq 0} \lambda_f(n) \sqrt{y} K_{iT}(2\pi ny) e(nx) \tag{4.1}$$

denote the Fourier series expansion of f , where

$$\rho_f(1)^2 = \frac{2 \cosh(\pi T)}{L(1, \text{sym}^2 f)}.$$

By unfolding, we have for $\Re(s) > 1$,

$$\langle f^2, E(\cdot, s) \rangle = 2\rho_f(1)^2 \sum_{n \geq 1} \frac{\lambda_f(n)^2}{(2\pi n)^s} \int_0^\infty y^s (K_{iT}(y))^2 \frac{dy}{y}.$$

We have that

$$\sum_{n \geq 1} \frac{\lambda_f(n)^2}{(2\pi n)^s} = \frac{\zeta(s)L(s, \text{sym}^2 f)}{(2\pi)^s \zeta(2s)}$$

by [Gol06, ch. 7],

$$\int_0^\infty y^s (K_{iT}(y))^2 \frac{dy}{y} = \frac{2^{s-3} \Gamma^2(s/2) \Gamma(s/2 + iT) \Gamma(s/2 - iT)}{\Gamma(s)}$$

by [GR00, (6.576)], and

$$\cosh(\pi T) = \frac{\pi}{|\Gamma(1/2 + iT)|^2}.$$

Thus, taking $s = \frac{1}{2} + it$ by analytic continuation, we have

$$\left| \left\langle f^2, E\left(\cdot, \frac{1}{2} + it\right) \right\rangle \right|^2 = \frac{\pi}{4} H(t) \frac{|\zeta(1/2 + it)L(1/2 + it, \text{sym}^2 f)|^2}{L(1, \text{sym}^2 f)^2 |\zeta(1 + 2it)|^2}. \tag{4.2}$$

This may be compared with (2.4).

By (2.5), we may restrict the integral in Lemma 2.1 to $|t| < T^{1+\epsilon}$. The zeta and L -value in the denominator of (4.2) are on the edge of the region of absolute convergence, so they are bounded below by $T^{-\epsilon}$. Thus,

$$\int_{-\infty}^\infty \left| \left\langle f^2, E\left(\cdot, \frac{1}{2} + it\right) \right\rangle \right|^2 dt \ll T^\epsilon \int_{-T^{1+\epsilon}}^{T^{1+\epsilon}} \frac{|\zeta(1/2 + it)L(1/2 + it, \text{sym}^2 f)|^2}{(1 + |t|) \prod_{\pm}(1 + |2T \pm t|)^{1/2}} dt.$$

On the GLH, this is bounded by

$$T^\epsilon \int_{-T^{1+\epsilon}}^{T^{1+\epsilon}} \frac{1}{(1 + |t|) \prod_{\pm}(1 + |2T \pm t|)^{1/2}} dt \ll T^{-1+\epsilon}.$$

5. Proof of Proposition 2.2: applying the trace formula

We first refine what needs to be proved for Proposition 2.2. We can immediately treat the contribution to (2.2) of t_j close to zero and close to $2T$. Let $0 < \alpha < \frac{1}{100}$ be a fixed constant to be determined later. On the GLH, we have by (2.4) and (2.5) that

$$\sum_{|t_j| < T^{1-\alpha}} + \sum_{|t_j - 2T| < T^{1-\alpha}} |\langle f^2, u_j \rangle|^2 \ll T^{-1+\epsilon}$$

$$\times \sum_{|t_j| < T^{1-\alpha}} \frac{1}{1 + |t|} + T^{-3/2+\epsilon} \sum_{|t_j - 2T| < T^{1-\alpha}} \frac{1}{(1 + |t_j - 2T|)^{1/2}}.$$

By Weyl’s law (see [IK04, p. 391]), we have that this is less than $T^{-\alpha/2+\epsilon}$. Thus, we may restrict the left-hand side of (2.2) to values of t_j that are roughly of size $2T$ but not too close to $2T$. We have to take care when making this restriction because we must use functions that will satisfy the conditions of Kuznetsov’s trace formula.

LEMMA 5.1. *Let $0 < \alpha < \frac{1}{100}$ be a fixed constant to be determined later and, define the even function*

$$W(t) = W_\alpha(t) = \left(1 - \exp\left(-\left(\frac{t}{(2T)^{1-\alpha/2}}\right)^{2\lceil 1000/\alpha \rceil}\right)\right) \left(1 - \exp\left(-\left(\frac{4T^2 - t^2}{4T^{2-\alpha/2}}\right)^{2\lceil 1000/\alpha \rceil}\right)\right),$$

where $\lceil x \rceil$ denotes the least integer greater than or equal to x . We have that $H(t)W(t) \ll T^{-100}$ unless

$$T^{1-\alpha} < |t| < 2T - T^{1-\alpha}, \tag{5.1}$$

in which range

$$\frac{d^k}{dt^k} H(t)W(t) \ll T^{-2}(T^{-1+\alpha})^k \tag{5.2}$$

for $k \geq 0$. In the range $T^{1-\alpha/4} < |t| < 2T - T^{1-\alpha/4}$, we have that $W(t) = 1 + O(T^{-100})$.

Proof. Suppose that $|t| \leq T^{1-\alpha}$. Then

$$\left(\frac{t}{(2T)^{1-\alpha/2}}\right)^{2\lceil 1000/\alpha \rceil} \ll T^{-100} \implies 1 - \exp\left(-\left(\frac{t}{(2T)^{1-\alpha/2}}\right)^{2\lceil 1000/\alpha \rceil}\right) \ll T^{-100}$$

and trivially

$$1 - \exp\left(-\left(\frac{4T^2 - t^2}{4T^{2-\alpha/2}}\right)^{2\lceil 1000/\alpha \rceil}\right) \ll 1.$$

Therefore, $H(t)W(t) \ll W(t) \ll T^{-100}$ for $|t| \leq T^{1-\alpha}$. Now suppose that $|t| \geq 2T - T^{1-\alpha}$. If also $|t| > 2T + T^\epsilon$, then by (2.5) we have $H(t) \ll T^{-100}$, and so $W(t)H(t) \ll T^{-100}$. So suppose that $2T - T^{1-\alpha} \leq |t| \leq 2T + T^\epsilon$. Then

$$\left(\frac{4T^2 - t^2}{4T^{2-\alpha/2}}\right)^{2\lceil 1000/\alpha \rceil} \ll T^{-100} \implies 1 - \exp\left(-\left(\frac{4T^2 - t^2}{4T^{2-\alpha/2}}\right)^{2\lceil 1000/\alpha \rceil}\right) \ll T^{-100}$$

and trivially

$$1 - \exp\left(-\left(\frac{t}{(2T)^{1-\alpha/2}}\right)^{2\lceil 1000/\alpha \rceil}\right) \ll 1.$$

This proves the first claim.

For the second claim, observe that for $x > 0$ and $N > 0$, we have $d^k/dx^k \exp(-x^N) \ll 1$. So

$$W^{(k)}(t) \ll (T^{-1+\alpha})^k. \tag{5.3}$$

To prove the same sort of bound for $H(t)$, we need more terms in the Stirling expansion (2.5). In the range (5.1), we have

$$H(t) = \frac{1}{2T^2} H_0\left(\frac{|t|}{2T}\right), \tag{5.4}$$

where

$$\begin{aligned} H_0(x) &= \frac{8\pi}{x(1-x^2)^{1/2}} \\ &\times \left(1 + \sum_{n \leq N} \frac{B_n}{T^n}\right) \left(1 + \sum_{n \leq N} \frac{C_n}{(Tx)^n}\right) \\ &\times \left(1 + \sum_{n \leq N} \frac{D_n}{(2T(1+x))^n}\right) \left(1 + \sum_{n \leq N} \frac{D_n}{(2T(1-x))^n}\right) + O(T^{-N/2}) \end{aligned}$$

for some constants B_n, C_n, D_n . By taking N large enough we see that if $T^{-\alpha} \ll |x| \ll 1 - T^{-\alpha}$, then

$$H_0^{(k)}(x) \ll (T^\alpha)^k. \tag{5.5}$$

Thus,

$$H^{(k)}(t) \ll T^{-2}(T^{-1+\alpha})^k \tag{5.6}$$

in the range (5.1).

For the third claim, suppose that $T^{1-\alpha/4} < |t| < 2T - T^{1-\alpha/4}$. Then

$$\left(\frac{t}{(2T)^{1-\alpha/2}}\right)^{2\lceil 1000/\alpha \rceil} \gg T^{100} \implies 1 - \exp\left(-\left(\frac{t}{(2T)^{1-\alpha/2}}\right)^{2\lceil 1000/\alpha \rceil}\right) = 1 + O(T^{-100})$$

and

$$\left(\frac{4T^2 - t^2}{4T^{2-\alpha/2}}\right)^{2\lceil 1000/\alpha \rceil} \gg T^{100} \implies 1 - \exp\left(-\left(\frac{4T^2 - t^2}{4T^{2-\alpha/2}}\right)^{2\lceil 1000/\alpha \rceil}\right) = 1 + O(T^{-100}).$$

Proposition 2.2 is thus reduced to proving

$$\sum_{j \geq 1} |\langle f^2, u_j \rangle|^2 W(t_j) = \frac{2\pi}{3} + O(T^{-\delta}) \tag{5.7}$$

for some $\delta > 0$.

We remark that since $L(\frac{1}{2}, u_j)$ and $L(\frac{1}{2}, u_j \times \text{sym}^2 f)$ are non-negative (by the work of [LR03] and [Lap03]), the contribution of the very small eigenvalues is at least

$$\sum_{|t_j| < T^\epsilon} |\langle f^2, u_j \rangle|^2 \gg T^{-\epsilon} \sum_{|t_j| < T^\epsilon} \frac{L(1/2, u_j)L(1/2, u_j \times \text{sym}^2 f)}{T}.$$

Thus, even if we were not assuming GLH, a subconvexity bound for $L(\frac{1}{2}, u_j \times \text{sym}^2 f)$ in the T aspect, which is polynomial in $|t_j| < T^\epsilon$, would be required, but this is an unsolved and very difficult problem.

By (2.4), we have

$$\sum_{j \geq 1} |\langle f^2, u_j \rangle|^2 W(t_j) = \frac{\pi^3}{72L(1, \text{sym}^2 f)^2} \sum_{j \geq 1} H(t_j) W(t_j) \frac{L(1/2, u_j) L(1/2, u_j \times \text{sym}^2 f)}{L(1, \text{sym}^2 u_j)}. \tag{5.8}$$

For u_j even, we may use the approximate functional equation (3.6) to write

$$L\left(\frac{1}{2}, u_j\right) L\left(\frac{1}{2}, u_j \times \text{sym}^2 f\right) = 2L\left(\frac{1}{2}, u_j\right) \sum_{m, r \geq 1} \frac{\lambda_j(m) A_f(r, m)}{r m^{1/2}} V_2(r^2 m, t_j).$$

But this equality holds for u_j odd as well, since in this case both sides vanish. Now we may use the approximate functional equation (3.8) for $L(\frac{1}{2}, u_j)$, which holds for both even and odd forms. This idea is an important feature of our proof, which, as mentioned in the remarks following Lemma 3.1, will make the analysis more pleasant. Thus, we get that the right-hand side of (5.8) equals

$$\begin{aligned} & \sum_{\pm} \frac{\pi^3}{72L(1, \text{sym}^2 f)^2} \sum_{j \geq 1} \sum_{n \geq 1} \frac{\lambda_j(\pm n)}{n^{1/2}} L(u_j \times \text{sym}^2 f) H(t_j) W(t_j) V_1^{\pm}(n, t_j) V_2(r^2 m, t_j) \\ & + O\left(T^{-1/2+\beta/2+\alpha+\epsilon} \sum_{T^{1-\alpha} < t_j < T^{1+\epsilon}} |H(t_j) W(t_j)| \left| \sum_{m, r \geq 1} \frac{\lambda_j(m) A_f(r, m)}{r m^{1/2}} V_2(r^2 m, t_j) \right| \right). \end{aligned} \tag{5.9}$$

By (5.2) and (3.7), the error term is

$$O\left(T^{-5/2+\beta/2+\alpha+\epsilon} \sum_{T^{1-\alpha} < t_j < T^{1+\epsilon}} \left| \int_{(\epsilon)} L\left(\frac{1}{2} + s, u_j \times \text{sym}^2 f\right) e^{s^2 \frac{G_2(1/2 + s)}{G_2(1/2)}} \frac{ds}{s} \right| \right),$$

which is $O(T^{-1/2+\beta/2+\alpha/2+\epsilon})$ on the GLH. By the approximate functional equation (3.6), the main term of (5.9) equals

$$\sum_{\pm} \frac{\pi^3}{36L(1, \text{sym}^2 f)^2} \sum_{n, m, r \geq 1} \frac{A_f(m, r)}{r(nm)^{1/2}} \sum_{j \geq 1} \lambda_j(\pm n) \lambda_j(m) H(t_j) W(t_j) V_1^{\pm}(n, t_j) V_2(r^2 m, t_j).$$

Applying the Kuznetsov trace formula to the inner sum gives

$$\sum_{j \geq 1} |\langle f^2, u_j \rangle|^2 W(t_j) = \mathfrak{D} + \mathfrak{E} + \mathfrak{O}^+ + \mathfrak{O}^- + O(T^{-\delta}) \tag{5.10}$$

for some $\delta > 0$, where

$$\mathfrak{D} = \frac{\pi^3}{36L(1, \text{sym}^2 f)^2} \sum_{n, r \geq 1} \frac{A_f(n, r)}{r n} \int_{-\infty}^{\infty} H(t) W(t) V_1^{\pm}(n, t) V_2(r^2 n, t) \frac{d^* t}{2\pi^2}, \tag{5.11}$$

$$\mathfrak{E} = \frac{-\pi^3}{18L(1, \text{sym}^2 f)^2} \int_{-\infty}^{\infty} \sum_{\pm} \sum_{n, m, r \geq 1} \frac{A_f(m, r) \lambda(n, t) \lambda(m, -t)}{r(nm)^{1/2} |\zeta(1 + 2it)|^2} H(t) W(t) V_1^{\pm}(n, t) V_2(r^2 m, t) \frac{dt}{2\pi}$$

and

$$\begin{aligned} \mathfrak{O}^{\pm} &= \frac{\pi^3}{36L(1, \text{sym}^2 f)^2} \sum_{n, m, r \geq 1} \sum_{c \geq 1} \frac{A_f(m, r) S(\pm n, m, c)}{r(nm)^{1/2} c} \\ &\quad \times \int_{-\infty}^{\infty} \mathfrak{J}^{\pm}\left(\frac{\sqrt{nm}}{c}, t\right) H(t) W(t) V_1^{\pm}(n, t) V_2(r^2 m, t) \frac{d^* t}{2\pi}. \end{aligned}$$

By the decay of V_1^\pm and V_2 , we may restrict the sum above to $n < T^{1 \mp \beta + \epsilon}$ and $mr^2 < T^{3 + \epsilon}$. We may also restrict to $c \leq T^3$, say, by a standard method (see [Blo12, Lemma 5], for example). To see this, consider, for instance, \mathcal{O}^+ , and move the line of the t -integral from $\Im(t) = 0$ to $\Im(t) = -\frac{1}{2} + \epsilon$. On the new line, we have by the power series (3.25) that

$$\mathcal{J}^+ \left(\frac{\sqrt{nm}}{c}, t \right) \ll \left(\frac{\sqrt{nm}}{c} \right)^{1-\epsilon}$$

for $\sqrt{nm}/c < 1$, as is the case when $c > T^3$. The rest of the integrand satisfies

$$H(t)W(t)V_1^\pm(n, t)V_2(r^2m, t) \tanh(\pi t) \ll \exp(-\pi q(t, T)),$$

where $q(t, T)$ is given in (2.6). Thus the contribution of the terms with $c > T^3$ is bounded by

$$\sum_{\substack{n \leq T^{1+\epsilon} \\ mr^2 \leq T^{3+\epsilon}}} \sum_{c > T^3} \frac{|A_f(m, r)|}{r(nm)^{1/2}} \frac{|S(\pm n, m, c)|}{c} \int_{-T^{1+\epsilon}}^{T^{1+\epsilon}} \left(\frac{\sqrt{nm}}{c} \right)^{1-\epsilon} dt \ll T^{1+\epsilon} \sum_{c > T^3} \frac{1}{c^{3/2-\epsilon}}.$$

The last bound uses the average version of the Ramanujan bound given in (3.4) and Weil’s bound for the Kloosterman sum. The result is less than a negative power of T .

We will prove that \mathcal{D} yields the main term, while \mathcal{E} and \mathcal{O}^\pm are bounded by a negative power of T .

6. Proof of Proposition 2.2: the diagonal

The goal of this section (see (5.7)) is to show that

$$\mathcal{D} = \frac{2\pi}{3} + O(T^{-\delta})$$

for some $\delta > 0$. This was sketched in [Blo13, § 4] but here we provide the details. By (3.12)–(3.13) and (5.11), we have

$$\begin{aligned} \mathcal{D} &= \frac{\pi^3}{36L(1, \text{sym}^2 f)^2} \int_{-\infty}^{\infty} H(t)W(t) \frac{1}{(2\pi i)^2} \int_{(\epsilon)} \int_{(\epsilon)} e^{s_1^2 + s_2^2} T^{-\beta s_1} \left(\frac{|t|}{2\pi} \right)^{s_1} \left(\frac{|t(4T^2 - t^2)|}{8\pi^3} \right)^{s_2} \\ &\times \sum_{n, r \geq 1} \frac{A_f(n, r)}{r^{1+2s_2} n^{1+s_1+s_2}} \frac{ds_1}{s_1} \frac{ds_2}{s_2} \frac{d^*t}{2\pi^2} (1 + O(T^{-1+\alpha})), \end{aligned} \tag{6.1}$$

where $0 < \alpha < \frac{1}{100}$ is as in Lemma 5.1. By [Gol06, Proposition 6.6.3], we have

$$\sum_{n, r \geq 1} \frac{A_f(n, r)}{r^{1+2s_2} n^{1+s_1+s_2}} = \frac{L(1 + 2s_2, \text{sym}^2 f)L(1 + s_1 + s_2, \text{sym}^2 f)}{\zeta(2 + s_1 + 3s_2)}.$$

Thus, assuming the GLH, we have that the error term in (6.1) contributes

$$O \left(T^{-1+\alpha+\epsilon} \int_{-\infty}^{\infty} |H(t)W(t)| t \tanh(\pi t) dt \right).$$

From Lemma 5.1, the weight function $H(t)W(t)$ is $O(T^{-100})$ unless $T^{1-\alpha} < |t| < 2T - T^{1-\alpha}$. By the estimate for $H(t)$ given in (2.5), we have

$$\int_{T^{1-\alpha}}^{2T - T^{1-\alpha}} |H(t)W(t)| t \tanh(\pi t) dt \ll \int_0^{2T} \frac{1}{(4T^2 - t^2)^{1/2}} dt \ll 1$$

and so the error term of (6.1) is $O(T^{-1/2})$.

Consider the main term of (6.1). Moving the line of integration to $\Re(s_1) = -\frac{1}{10}$, we pick up a simple pole at $s_1 = 0$, getting

$$\mathcal{D} = \frac{\pi^3}{36L(1, \text{sym}^2 f)^2} \int_{-\infty}^{\infty} H(t)W(t) \frac{1}{(2\pi i)^2} \int_{(\epsilon)} e^{s_2^2 \left(\frac{|t(4T^2 - t^2)|}{8\pi^3} \right)^{s_2}} \times \frac{L(1 + 2s_2, \text{sym}^2 f)L(1 + s_2, \text{sym}^2 f)}{\zeta(2 + 3s_2)} \frac{ds_2}{s_2} \frac{d^*t}{2\pi^2} + O(T^{-(1-\beta)/10+\epsilon}). \tag{6.2}$$

The new error term arises by applying GLH on the shifted line of integration, and it is $O(T^{-1/20})$. Now, moving the line of integration to $\Re(s_2) = -\frac{1}{10}$ and picking up a simple pole at $s_2 = 0$, we get

$$\mathcal{D} = \frac{\pi^3}{36L(1, \text{sym}^2 f)^2} \int_{-\infty}^{\infty} \frac{L(1, \text{sym}^2 f)^2}{\zeta(2)} H(t)W(t) \frac{d^*t}{2\pi^2} + O(T^{-(3-\alpha)/10+\epsilon} + T^{-1/20}).$$

The error term is $O(T^{-1/20})$. The main term equals

$$\frac{1}{6\pi} \int_0^{\infty} H(t)W(t)t \tanh(\pi t) dt.$$

We can now restrict the integrand to the range $T^{1-\alpha} < t < 2T - T^{1-\alpha}$, on which interval $\tanh(\pi t) = 1 + O(T^{-100})$. Further, for $T^{1-\alpha/4} < t < 2T - T^{1-\alpha/4}$ we have $W(t) = 1 + O(T^{-100})$. Thus,

$$\begin{aligned} \mathcal{D} &= \frac{1}{6\pi} \int_{T^{1-\alpha}}^{2T-T^{1-\alpha}} H(t)W(t)t dt + O(T^{-1/20}) \\ &= \frac{1}{6\pi} \int_{T^{1-\alpha/4}}^{2T-T^{1-\alpha/4}} H(t)t dt + O(T^{-\alpha/8}). \end{aligned}$$

By (2.5) we have

$$\begin{aligned} \mathcal{D} &= \frac{1}{6\pi} \int_{T^{1-\alpha/4}}^{2T-T^{1-\alpha/4}} \frac{8\pi}{(4T^2 - t^2)^{1/2}} dt + O(T^{-\alpha/8}) \\ &= \frac{4}{3} \left(\arcsin\left(\frac{2T - T^{1-\alpha/4}}{2T}\right) - \arcsin\left(\frac{T^{1-\alpha/4}}{2T}\right) \right) + O(T^{-\alpha/8}) \\ &= \frac{4}{3} \left(\frac{\pi}{2} - 0 \right) + O(T^{-\alpha/8}). \end{aligned}$$

The error term is some negative power of T . This completes the evaluation of the diagonal.

7. Proof of Proposition 2.2: the Eisenstein series contribution

In this section we show that \mathcal{E} is bounded by a negative power of T . We first rewrite the expression for \mathcal{E} using the following approximate functional equations, which are analogous to those in Lemma 3.1:

$$\begin{aligned} \left| \zeta\left(\frac{1}{2} + it\right) \right|^2 &= \sum_{\pm} \sum_{n \geq 1} \frac{\lambda(n, t)}{n^{1/2}} V_1^{\pm}(n, t), \\ \left| L\left(\frac{1}{2} + it, \text{sym}^2 f\right) \right|^2 &= \sum_{m, r \geq 1} \frac{A_f(m, r)\lambda(m, -t)}{rm^{1/2}} V_2(r^2 m, t). \end{aligned}$$

We get that

$$\mathfrak{E} = \frac{-\pi^3}{9L(1, \text{sym}^2 f)^2} \int_0^\infty \frac{|\zeta(1/2 + it)|^2 |L(1/2 + it, \text{sym}^2 f)|^2}{|\zeta(1 + 2it)|^2} H(t)W(t) \frac{dt}{2\pi}.$$

By the decay of the weight function $H(t)W(t)$, we may restrict the integral to $T^{1-\alpha} < t < 2T - T^{1-\alpha}$ and then apply the GLH and the estimate for $H(t)$ given in (2.5) to see that

$$\begin{aligned} \mathfrak{E} &= \frac{-\pi^3}{9L(1, \text{sym}^2 f)^2} \int_{T^{1-\alpha}}^{2T-T^{1-\alpha}} \frac{|\zeta(1/2 + it)|^2 |L(1/2 + it, \text{sym}^2 f)|^2}{|\zeta(1 + 2it)|^2} H(t)W(t) \frac{dt}{2\pi} \\ &\ll T^\epsilon \int_{T^{1-\alpha}}^{2T-T^{1-\alpha}} \frac{1}{t(4T^2 - t^2)^{1/2}} dt \\ &\ll T^{-1+\alpha+\epsilon}. \end{aligned}$$

8. Proof of Proposition 2.2: the short off-diagonal

The goal of this section is to show that \mathfrak{O}^+ is bounded by a negative power of T . We have seen that for any $0 < \alpha < \frac{1}{100}$ we have $H(t)W(t) \ll T^{-100}$ unless $T^{1-\alpha} < |t| < 2T - T^{1-\alpha}$. Thus, in the expression for \mathfrak{O}^+ we may restrict the integral to this range. We may also replace d^*t by $t dt$ because $\tanh t = 1 + O(T^{-100})$ in the given range of t . Thus,

$$\begin{aligned} \mathfrak{O}^+ &\ll T^\epsilon \sum_{\substack{n < T^{1-\beta+\epsilon} \\ mr^2 < T^{3+\epsilon}}} \sum_{c \leq T^3} \frac{A_f(m, r)}{r(nm)^{1/2}} \frac{S(n, m, c)}{c} \\ &\quad \times \int_{T^{1-\alpha}}^{2T-T^{1-\alpha}} + \int_{-2T+T^{1-\alpha}}^{-T^{1-\alpha}} \mathfrak{F}^+ \left(\frac{\sqrt{nm}}{c}, t \right) H(t)W(t) V_1^+(n, t) V_2(r^2 m, t) t dt. \end{aligned}$$

Let Z be any smooth, even function compactly supported on $(T^{-\alpha}, 2 - T^{-\alpha}) \cup (-2 + T^{-\alpha}, -T^{-\alpha})$ with derivatives satisfying

$$\|Z^{(k)}\|_\infty \ll (T^\alpha)^k.$$

Then $Z(t/T)$ is supported on $T^{1-\alpha} < |t| < 2T - T^{1-\alpha}$, on which $(d^k/dt^k)Z(t/T) \ll T^{(-1+\alpha)k}$, and we may use such a function to approximate the characteristic function of this interval. We may absorb $W(t)$ into the function $Z(t/T)$ by property (5.3). Writing

$$H(t) = \frac{1}{2T^2} H_0\left(\frac{t}{T}\right)$$

as in (5.4), we may also absorb $H_0(t/T)$ into $Z(t/T)$ by property (5.5). Thus it suffices to prove that for any function Z as above, we have

$$\frac{1}{T^2} \sum_{n, m, r \geq 1} \sum_{c \leq T^3} \frac{A_f(m, r)}{r(nm)^{1/2}} \frac{S(n, m, c)}{c} \int_{-\infty}^\infty \mathfrak{F}^+ \left(\frac{\sqrt{nm}}{c}, t \right) Z\left(\frac{t}{T}\right) V_1^+(n, t) V_2(r^2 m, t) t dt \ll T^{-\delta}$$

for some $\delta > 0$. We apply Lemma 3.3 to evaluate the Bessel transform. Note that the function $Z(t/T)V_1^+(n, t)V_2(r^2 m, t)$ satisfies the conditions of the lemma by the remarks above and by (3.14)–(3.15).

The contribution of the main term of Lemma 3.3 is $O(T^{-100})$, unless

$$\frac{2\sqrt{nm}}{c} > T^{2-3\alpha}.$$

Since by the decay of $V_1^+(n, t)V_2(r^2m, t)$ we may take $n < T^{1-\beta+\epsilon}$ and $m < T^{3+\epsilon}$ up to $O(T^{-100})$, this imposes

$$c < T^{3\alpha-\beta/2}.$$

We now fix

$$\beta = 7\alpha$$

so that the above condition on c is impossible and the main term is $O(T^{-100})$. This of course leads to an acceptable bound for \mathfrak{O}^+ .

The error term $O(\sqrt{nm}/cT^{3-12\alpha})$ arising from Lemma 3.3 contributes

$$O\left(\frac{1}{T^2} \sum_{\substack{n < T^{1-7\alpha+\epsilon} \\ mr^2 < T^{3+\epsilon}}} \sum_{c \leq T^3} \frac{|A_f(m, r)|}{r(nm)^{1/2}} \frac{|S(n, m, c)|}{c} \frac{\sqrt{nm}}{cT^{3-12\alpha}}\right).$$

By (3.4) and Weil’s bound for the Kloosterman sum, this is

$$O\left(T^{-5+12\alpha+\epsilon} \sum_{\substack{n < T^{1-7\alpha+\epsilon} \\ mr^2 < T^{3+\epsilon}}} \sum_{c \leq T^3} \frac{1}{rc^{3/2}}\right).$$

The innermost c -sum is $O(1)$, so the line above is $O(T^{-1+5\alpha+\epsilon})$, which is admissible as we assume $\alpha < \frac{1}{100}$.

9. Proof of Proposition 2.2: the long off-diagonal

The goal now is to show that \mathfrak{O}^- is bounded by a negative power of T . This proof is the heart of our paper. As in the previous section, it suffices to prove that for any smooth, even function Z compactly supported on

$$(T^{-\alpha}, 2 - T^{-\alpha}) \cup (-2 + T^{-\alpha}, -T^{-\alpha})$$

with derivatives satisfying $\|Z^{(k)}\|_\infty \ll (T^\alpha)^k$, we have

$$\frac{1}{T^2} \sum_{\substack{n < T^{1+7\alpha+\epsilon} \\ mr^2 < T^{3+\epsilon}}} \sum_{c \leq T^3} \frac{A_f(m, r)}{r(nm)^{1/2}} \frac{S(-n, m, c)}{c} \int_{-\infty}^\infty \mathfrak{J}^-\left(\frac{\sqrt{nm}}{c}, t\right) Z\left(\frac{t}{T}\right) V_1^-(n, t)V_2(r^2m, t)t dt \ll T^{-\delta}$$

for some $\delta > 0$. We may replace $V_1^-(n, t)$ and $V_2(r^2m, t)$ by the main terms in their Stirling expansions (3.12) and (3.13) since the lower-order terms can be treated similarly. Thus, we need to show that

$$\frac{1}{T^2} \sum_{\substack{n < T^{1+7\alpha+\epsilon} \\ mr^2 < T^{3+\epsilon}}} \sum_{c \leq T^3} \frac{A_f(m, r)}{r(nm)^{1/2}} \frac{S(-n, m, c)}{c} \int_{-\infty}^\infty \mathfrak{J}^-\left(\frac{\sqrt{nm}}{c}, t\right) Z\left(\frac{t}{T}\right) V\left(\frac{n}{T}, \frac{mr^2}{T^3}, \frac{t}{T}\right)t dt \quad (9.1)$$

is bounded by a negative power of T where

$$V(x_1, x_2; y) = \int_{(\sigma)} e^{s_1^2(2\pi x_1)^{-s_1}|T^{7\alpha}y|^{s_1}} \frac{ds_1}{s_1} \cdot \int_{(\sigma)} e^{s_2^2(8\pi^3 x_2)^{-s_2}|y(4 - y^2)|^{-s_2}} \frac{ds_2}{s_2}$$

for any $\sigma > 0$. We apply Lemma 3.4 to evaluate the Bessel transform in (9.1). The error term arising from this result contributes less than a negative power of T , just as in the previous

section. There are two similar, non-oscillatory main terms in the asymptotic given by Lemma 3.4. It suffices to treat only the leading main term as the other will contribute a factor of T less. Therefore the goal is to bound by a negative power of T the sum

$$\frac{1}{T^2} \sum_{\substack{n < T^{1+7\alpha+\epsilon} \\ mr^2 < T^{3+\epsilon}}} \sum_{c \leq T^3} \frac{A_f(m, r)}{r} \frac{S(-n, m, c)}{c^2} Z\left(\frac{2\pi\sqrt{nm}}{Tc}\right) V\left(\frac{n}{T}, \frac{mr^2}{T^3}, \frac{2\pi\sqrt{nm}}{Tc}\right).$$

Applying a smooth partition of unity, we consider the sum above in dyadic intervals. Let U be a smooth bump function supported on $(1, 2) \times (1, 2)$ and possessing bounded derivatives. It suffices to prove that

$$\frac{1}{T^2} \sum_{n, m, r \geq 1} \sum_{c \geq 1} \frac{A_f(m, r)}{r} \frac{S(-n, m, c)}{c^2} Z\left(\frac{2\pi\sqrt{nm}}{Tc}\right) V\left(\frac{n}{T}, \frac{mr^2}{T^3}, \frac{2\pi\sqrt{nm}}{Tc}\right) U\left(\frac{n}{N}, \frac{mr^2}{M}\right) \ll T^{-\delta} \tag{9.2}$$

for some $\delta > 0$, where

$$N < T^{1+7\alpha+\epsilon}, \quad M < T^{3+\epsilon}, \quad \frac{\sqrt{NM}}{rT} \ll c \ll \frac{\sqrt{NM}}{rT^{1-\alpha}}.$$

The bounds on c are enforced by the function Z and make the condition $c \leq T^3$ redundant. It will be apparent from the proof that as long as α is smaller than some fixed constant, there exists some absolute $\delta > 0$ independent of the value of α . For this reason it will be very convenient to rename α to ϵ and apply the ϵ -convention. Therefore we have

$$N < T^{1+\epsilon}, \quad M < T^{3+\epsilon}, \quad \frac{\sqrt{NM}}{rT} \ll c \ll \frac{\sqrt{NM}}{rT^{1-\epsilon}}. \tag{9.3}$$

9.1 Poisson and Voronoi summation

By separating the n -sum in (9.2) into residue classes mod c and applying the Poisson summation formula, we get that the left-hand side of (9.2) equals

$$\begin{aligned} & \frac{N}{T^2} \sum_{m, r, c \geq 1} \sum_{a \bmod c} \frac{A_f(m, r)}{r} \frac{S(-a, m, c)}{c^3} \\ & \times \sum_{-\infty < j < \infty} e\left(\frac{ja}{c}\right) \int_{-\infty}^{\infty} U\left(\xi, \frac{mr^2}{M}\right) Z\left(\frac{2\pi\sqrt{\xi Nm}}{Tc}\right) \\ & \times V\left(\frac{\xi N}{T}, \frac{mr^2}{T^3}, \frac{2\pi\sqrt{\xi Nm}}{Tc}\right) e\left(\frac{-jN\xi}{c}\right) d\xi. \end{aligned} \tag{9.4}$$

Now, in this expression, writing $x = mr^2/M$, we may replace m by Mx/r^2 . Writing $y = \sqrt{NM}/crT$, we may replace N by $y^2c^2r^2T^2/M$. Thus (9.4) equals

$$\frac{N}{T^2} \sum_{\substack{c, m, r \geq 1 \\ -\infty < j < \infty}} \sum_{a \bmod c} \frac{A_f(m, r)}{r} \frac{S(-a, m, c)}{c^3} e\left(\frac{ja}{c}\right) \psi\left(\frac{mr^2}{M}, \frac{\sqrt{NM}}{crT}, \frac{jc^2r^2T^2}{M}\right) \tag{9.5}$$

where

$$\psi(x, y; u) = \int_{-\infty}^{\infty} U(\xi, x) Z(2\pi y\sqrt{\xi x}) V\left(\frac{\xi N}{T}, \frac{xM}{T^3}, 2\pi y\sqrt{\xi x}\right) e(-\xi y^2 u) d\xi. \tag{9.6}$$

In the integrand above, x and ξ must lie in the interval $(1, 2)$ by the definition of U , while y must lie in the interval $(T^{-\epsilon}, T^\epsilon)$ by the definition of Z . Thus all three variables x, ξ, y should be thought of as roughly constant and bounded away from zero.

We first observe that we may restrict (9.5) to $j \neq 0$. First, note that the contribution of $j = 0$ is nil unless $c = 1$, because $\sum_{a \bmod c} S(-a, m, c) = 0$ for $c > 1$. This leaves the case $j = 0$ and $c = 1$, whose contribution is

$$\begin{aligned} & \frac{N}{T^2} \sum_{m,r \geq 1} \frac{A_f(m,r)}{r} \psi\left(\frac{mr^2}{M}, \frac{\sqrt{NM}}{rT}; 0\right) \\ &= \sum_{m,r \geq 1} \frac{A_f(m,r)}{mr} \frac{mr^2}{M} \frac{NM}{r^2 T^2} \psi\left(\frac{mr^2}{M}, \frac{\sqrt{NM}}{rT}; 0\right) \\ &= \sum_{m,r,v \geq 1} \frac{\mu(v)}{v^2} \frac{A_f(m,1)A_f(r,1)}{mr} \frac{mr^2 v^3}{M} \frac{NM}{r^2 v^2 T^2} \psi\left(\frac{mr^2 v^3}{M}, \frac{\sqrt{NM}}{rvT}; 0\right), \end{aligned} \tag{9.7}$$

where the last equality follows by the Hecke relations (3.2). The sum is trivially $O(T^\epsilon)$, using (3.4), so we must save any negative power of T . We are done, unless

$$v \ll T^\epsilon \quad \text{and} \quad \frac{mr^2 v^3}{M} \gg T^{-\epsilon},$$

in which case, since

$$\frac{mr^2 v^3}{M} \ll T^\epsilon \quad \text{and} \quad T^{-\epsilon} \ll \frac{NM}{r^2 v^2 T^2} \ll T^\epsilon$$

are enforced by the weight function, we have

$$\frac{T^{2-\epsilon}}{N} \ll m \ll \frac{T^{2+\epsilon}}{N}.$$

This shows that the interval of summation of m is at least as long as $T^{1-\epsilon}$, by (9.3). On the GLH we can show in a standard way (see §9.4) that

$$\sum_{m \leq x} A_f(m, 1) \ll x^{1/2+\epsilon}.$$

Using this and partial summation gives the required saving in (9.7).

The next step is to transform the m -sum using GL(3) Voronoi summation (see Lemma 3.8). Writing

$$S(-a, m, c) = \sum_{b \bmod c}^* e\left(\frac{-ab + m\bar{b}}{c}\right),$$

we get by Voronoi summation that the part of (9.5) with $j \neq 0$ equals

$$\begin{aligned} & \frac{N}{T^2} \sum_{\pm} \sum_{c,r,\ell, |j| \geq 1} \frac{A_f(\ell, l)}{\ell l r c^2} \sum_{b \bmod c}^* S\left(rb, \pm \ell, \frac{cr}{l}\right) \sum_{a \bmod c} e\left(\frac{a(j-b)}{c}\right) \\ & \times \frac{1}{4\pi i} \int_{(\epsilon)} \left(\frac{\ell M l^2}{c^3 r^3}\right)^{1-s} G^\pm(s) \int_0^\infty \psi\left(x, \frac{\sqrt{NM}}{crT}; \frac{jcr^2 T^2}{M}\right) x^{-s} dx ds, \end{aligned} \tag{9.8}$$

where $G^\pm(s)$ is as defined in Lemma 3.8. We have that $\sum_{a \bmod c} e(a(j-b)/c)$ equals c if $b \equiv j \pmod c$ and 0 otherwise. In the case $b \equiv j \pmod c$, we have $br \equiv jr \pmod{cr}$. Thus (9.8) equals

$$\begin{aligned} & \frac{N}{T^2} \sum_{\pm} \sum_{\substack{c,r,\mathfrak{k},|j| \geq 1 \\ (j,c)=1 \\ l|cr}} \frac{A_f(\mathfrak{k}, l)}{\mathfrak{k}lrc} S\left(jr, \pm\mathfrak{k}, \frac{cr}{l}\right) \\ & \times \frac{1}{4\pi i} \int_{(\epsilon)} \left(\frac{\mathfrak{k}Ml^2}{c^3r^3}\right)^{1-s} G^\pm(s) \int_0^\infty \psi\left(x, \frac{\sqrt{NM}}{crT}; \frac{jcr^2T^2}{M}\right) x^{-s} dx ds. \end{aligned} \tag{9.9}$$

Recall by the remarks following Lemma 3.8 that, up to negligible error, we can restrict the sum to

$$\frac{\mathfrak{k}Ml^2}{c^3r^3} < T^{2+\epsilon}. \tag{9.10}$$

So, if we set $z = \mathfrak{k}Ml^2/c^3r^3T^2$ then $T^{-6} < z < T^\epsilon$ and we can write $M = zc^3r^3T^2/\mathfrak{k}l^2$ in the fraction jcr^2T^2/M appearing in (9.9), and $\mathfrak{k} = zc^3r^3T^2/Ml^2$. Thus (9.9) equals

$$\frac{NM}{T^2} \sum_{\pm} \sum_{\substack{c,r,\mathfrak{k},|j| \geq 1 \\ (j,c)=1 \\ l|cr}} \frac{lA_f(\mathfrak{k}, l)}{c^4r^4} S\left(rj, \pm\mathfrak{k}, \frac{cr}{l}\right) \Psi^\pm\left(\frac{\sqrt{NM}}{crT}, \frac{\mathfrak{k}Ml^2}{c^3r^3T^2}; \frac{j\mathfrak{k}l^2}{c^2r}\right), \tag{9.11}$$

where

$$\Psi^\pm(y, z; u) = \frac{1}{4\pi i} \int_{(\epsilon)} z^{-s} T^{-2s} G^\pm(s) \int_0^\infty \psi(x, y; z^{-1}u) x^{-s} dx ds. \tag{9.12}$$

The goal is to show that (9.11) is bounded by a negative power of T . Using $G^\pm(s) \ll T^{-1+\epsilon}$, a bound given in (3.29), and repeatedly integrating by parts in (9.6), we have

$$\frac{\partial^k}{\partial u^k} \Psi^\pm(y, z; u) \ll T^{-1+\epsilon} z^{-k} |z^{-1}u|^{-A} \tag{9.13}$$

for any $k \geq 0$ and $A \geq 0$. Thus, up to negligible error, we may assume

$$|u| < zT^\epsilon. \tag{9.14}$$

Using this, we have

$$\frac{\partial^k}{\partial y^k} \Psi^\pm(y, z; u) \ll T^{-1+\epsilon}, \quad \frac{\partial^k}{\partial z^k} \Psi^\pm(y, z; u) \ll T^{-1+\epsilon} z^{-k}. \tag{9.15}$$

9.2 Preparation for Kuznetsov’s formula

We first explain the idea of what we are about to do next. Our task is to bound by a negative power of T a sum like (when $r = l = 1$),

$$\frac{NM}{T^2} \sum_{c,\mathfrak{k},|j| \geq 1} \frac{A_f(\mathfrak{k}, 1)}{c^4} S(j, \pm\mathfrak{k}, c) \Psi^\pm\left(\frac{\sqrt{NM}}{cT}, \frac{\mathfrak{k}M}{c^3T^2}; \frac{j\mathfrak{k}}{c^2}\right). \tag{9.16}$$

By (9.14) and $z < T^\epsilon$, we see that, up to negligible error, we may assume $|j|\mathfrak{k} \ll c^2T^\epsilon$. Also recall that $|\Psi^\pm(y, z; u)| \ll T^{-1+\epsilon}$. By these remarks, the trivial bound for (9.16) is $O(T^{1/2+\epsilon})$, and now we must save this much and a little more. We have already exploited the sums over j

and k (they arose through Poisson and Voronoi summation). Now we will exploit the sum over c . Since the range of c in the sum is at least as large as $T^{-\epsilon} \sqrt{j|k}$, a range sometimes referred to as the Linnik range (see [ST09]), we are in a good position to use Kuznetsov’s formula to transform the sum of Kloosterman sums into a sum of Hecke eigenvalues. This final sum will be estimated under the GLH to complete the proof. To get a feel for how this works, consider the generic ranges $c \asymp T, N \asymp T, M \asymp T^3$ that imply $j \ll T^\epsilon, k \asymp T^2$. Using Kuznetsov’s formula, we will get an identity of the shape

$$\sum_{c \asymp \sqrt{|j|k}} c^{-1} S(\pm j, k, c) \asymp \sum_{t_j < T^\epsilon} \lambda_j(j) \lambda_j(k) + \dots$$

Thus (9.16) essentially becomes, ignoring the j sum of length T^ϵ and keeping in mind that $\Psi^\pm(y, z, w)$ has a factor T^{-1} by (3.30),

$$\sum_{t_j < T^\epsilon} T^{-2} \sum_{k \asymp T^2} A_f(k, 1) \lambda_j(k) + \dots$$

On the GLH, we get cancellation in the k -sum and obtain the required bound.

To carry out the above program, we must first take care of the technicalities posed by the presence of the l and r parameters. First we detect the condition $(j, c) = 1$ in (9.11) using the Möbius function by recalling that $\sum_{d|c, d|j} \mu(d)$ equals 1 if $(j, c) = 1$ and 0 otherwise. Thus we need to bound the following sum, for each sign \pm , by a negative power of T :

$$\frac{NM}{T^2} \sum_{\substack{d, c, r, k, |j| \geq 1 \\ l|cdr}} \frac{l\mu(d)A_f(k, l)}{c^4 r^4 d^4} S\left(rjd, \pm k, \frac{cdr}{l}\right) \Psi^\pm\left(\frac{\sqrt{NM}}{cdrT}, \frac{kMl^2}{c^3 d^3 r^3 T^2}; \frac{jk l^2}{c^2 dr}\right).$$

We can reorder this sum by the greatest common divisor of l and c to say that it equals

$$\frac{NM}{T^2} \sum_{b \geq 1} \sum_{\substack{d, c, r, k, |j| \geq 1 \\ l|cdr \\ (l, c) = b}} \frac{l\mu(d)A_f(k, l)}{c^4 r^4 d^4} S\left(rjd, \pm k, \frac{cdr}{l}\right) \Psi^\pm\left(\frac{\sqrt{NM}}{cdrT}, \frac{kMl^2}{c^3 d^3 r^3 T^2}; \frac{jk l^2}{c^2 dr}\right).$$

Replacing l by lb and c by cb , this sum equals

$$\frac{NM}{T^2} \sum_{b \geq 1} \sum_{\substack{d, c, r, k, |j| \geq 1 \\ l|dr \\ (l, c) = 1}} \frac{l\mu(d)A_f(k, bl)}{c^4 r^4 d^4 b^3} S\left(rjd, \pm k, \frac{cdr}{l}\right) \Psi^\pm\left(\frac{\sqrt{NM}}{cbdrT}, \frac{kMl^2}{c^3 d^3 r^3 bT^2}; \frac{jk l^2}{c^2 dr}\right).$$

Now detecting $(l, c) = 1$ using the Möbius function, the sum above equals

$$\frac{NM}{T^2} \sum_{\substack{b, d, c, r, k, |j| \geq 1 \\ a|dr}} \frac{l\mu(d)\mu(a)A_f(k, abl)}{c^4 r^4 d^4 a^3 b^3} S\left(rjd, \pm k, \frac{cdr}{l}\right) \Psi^\pm\left(\frac{\sqrt{NM}}{cabdrT}, \frac{kMl^2}{c^3 d^3 r^3 abT^2}; \frac{jk l^2}{c^2 dr}\right).$$

Using the Hecke relations (3.2) and writing $1/(cabdr)^3 = (\sqrt{NM}/cabdrT)^3(T/\sqrt{NM})^3$, it suffices to show that

$$\begin{aligned} & \frac{T}{\sqrt{NM}} \sum_{\substack{b,d,r \leq T^2 \\ a|dr \\ v|abl}} |A_f(abl, 1)| \left| \sum_{\substack{k, |j|, c \geq 1}} \frac{A_f(k, 1)}{cdr/l} S\left(drj, \pm kv, \frac{cdr}{l}\right) \right. \\ & \times \left. \left(\frac{\sqrt{NM}}{cabdrT}\right)^3 \Psi^\pm\left(\frac{\sqrt{NM}}{cabdrT}, \frac{kMl^2}{c^3d^3r^3abT^2}; \frac{jkvl^2}{c^2dr}\right) \right| \ll T^{-\delta} \end{aligned} \tag{9.17}$$

for some $\delta > 0$. The given bounds $b, d, r \leq T^2$ are enforced by the weight functions. Also note for later that the desired bound of (9.17) is trivial for the sub-sum restricted to $|jk| \leq T^{1/10}$ using $\Psi^\pm(y, z, u) \ll T^{-1+\epsilon}$ and $T/\sqrt{NM} \ll T^\epsilon/cabdr$ and Kim and Sarnak’s bound

$$A_f(abl, 1) \ll (abl)^{7/32+\epsilon}.$$

We continue to reshape the inner sum in order to apply Lemma 3.5. We write

$$\begin{aligned} & \sum_{c \geq 1} \frac{S(drj, \pm kv, cdr/l)}{cdr/l} \left(\frac{\sqrt{NM}}{cabdrT}\right)^3 \Psi^\pm\left(\frac{\sqrt{NM}}{cabdrT}, \frac{kMl^2}{c^3d^3r^3abT^2}; \frac{jkvl^2}{c^2dr}\right) \\ & = \sum_{c \geq 1} \frac{S(rjd, \pm kv, cdr/l)}{cdr/l} \Phi\left(\frac{4\pi\sqrt{|j|krdv}}{cdr/l}\right), \end{aligned}$$

where, after absorbing the factor $(\sqrt{NM}/cabdrT)^3$ into the function Ψ^\pm , we let

$$\Phi(w) = \Psi^\pm\left(w \frac{\sqrt{NM}}{4\pi ablT\sqrt{|j|krdv}}, w^3 \frac{kM}{ablT^2(4\pi\sqrt{|j|krdv})^3}; w^2 \frac{\text{sgn}(j)}{(4\pi)^2}\right)$$

for $w > 0$. Since $\Psi^\pm(y, z; u)$ is restricted to $T^{-\epsilon} \ll y \ll T^\epsilon$, we have

$$\text{support}(\Phi(w)) \subset \left(T^{-\epsilon} \left(\frac{\sqrt{NM}}{ablT\sqrt{|j|krdv}}\right)^{-1}, \left(\frac{\sqrt{NM}}{ablT\sqrt{|j|krdv}}\right)^{-1} T^\epsilon\right). \tag{9.18}$$

Further, we have seen that $\Psi^\pm(y, z; u)$ may be restricted to $z \ll T^\epsilon$ and $u \ll zT^\epsilon$ up to an error of $O(T^{-100})$. Thus, by (9.13) and (9.15), we have

$$|\Phi^{(k)}(w)| \ll T^{-1+\epsilon}(1+w^{-k})(z^{-1}w^2)^{-A} + O(T^{-100})$$

for any $k \geq 0$ and $A \geq 0$. Equivalently, by (9.18) we have

$$|\Phi^{(k)}(w)| \ll \frac{1}{T^{1-\epsilon}} \left(1 + \left(\frac{\sqrt{NM}}{ablT\sqrt{|j|krdv}}\right)^k\right) \left(\frac{z}{w^2}\right)^A + O(T^{-100}). \tag{9.19}$$

We see that up to an error of $O(T^{-100})$, say, we can assume

$$w^2 \ll zT^\epsilon \Rightarrow \frac{l\sqrt{|j|krdv}}{cdr} \ll T^\epsilon.$$

So, since $\sqrt{NM}/cabdrT$ is restricted to $(T^{-\epsilon}, T^\epsilon)$, we have

$$\frac{\sqrt{NM}}{ablT\sqrt{|j|krdv}} \gg T^{-\epsilon}. \tag{9.20}$$

Thus (9.19) implies

$$|\Phi^{(k)}(w)| \ll \frac{1}{T^{1-\epsilon}} \left(\frac{\sqrt{NM}}{ablT\sqrt{|j|\ell rdv}} \right)^k + O(T^{-100}). \tag{9.21}$$

Finally, we note some bounds on j and ℓ . By (9.14), we may restrict to

$$\frac{j\ell v l^2}{c^2 dr} \ll \frac{\ell M l^2}{c^3 d^3 r^3 abT^{2-\epsilon}},$$

which gives

$$j \ll \frac{M}{cd^2 r^2 abvT^{2-\epsilon}}.$$

We already have the bound (9.10) on ℓ . So since $T^{-\epsilon} < \sqrt{NM}/cabdrT < T^\epsilon$, we may eliminate c and record the bounds

$$\begin{aligned} |j| &\ll \frac{M^{1/2}}{T^{1-\epsilon} N^{1/2} drv}, \\ \ell &\ll \frac{N^{3/2} M^{1/2}}{a^2 b^2 l^2 T^{1-\epsilon}}. \end{aligned}$$

Also note for later use that $\Phi(w)$ depends implicitly on j and ℓ , and by (9.15) we have

$$\frac{\partial^n}{\partial j^n} \Phi(w) \ll T^{-1+\epsilon} j^{-n}, \quad \frac{\partial^n}{\partial \ell^n} \Phi(w) \ll T^{-1+\epsilon} \ell^{-n}. \tag{9.22}$$

9.3 Kuznetsov’s trace formula and the large sieve

Consider the sub-sum of (9.17) consisting of the terms with $j > 0$ and the positive sign case, the rest of the sum being similarly treated. Applying Lemma 3.5 and the remarks following it, we see that it suffices to show that the following expression is bounded by a negative power of T :

$$\begin{aligned} &\frac{T}{\sqrt{NM}} \sum_{\substack{b,d,r \leq T^2 \\ al|dr \\ v|abl}} (abl)^{7/32} \left| \sum_{\substack{K \leq \ell < 2K \\ J \leq j < 2J}} A_f(\ell, 1) \sum_{\substack{g \in \mathcal{B}_0(dr/l) \\ |t_g| < T^\epsilon}} \hat{\Phi}(t_g) \frac{4\pi \sqrt{j\ell drv}}{\cosh(\pi t_g)} \rho_g(\ell v) \overline{\rho_g(jdr)} \right| \\ &+ \frac{T}{\sqrt{NM}} \sum_{\substack{b,d,r \leq T^2 \\ al|dr \\ v|abl}} (abl)^{7/32} \left| \sum_{\substack{K \leq \ell < 2K \\ J \leq j < 2J}} A_f(\ell, 1) \sum_{\substack{g \in \mathcal{B}_k(dr/l) \\ k < T^\epsilon}} \hat{\Phi}(k) \frac{(k-1)! \sqrt{j\ell drv}}{\pi(4\pi)^{k-1}} \rho_g(\ell v) \overline{\rho_g(jdr)} \right| \\ &+ \frac{T}{\sqrt{NM}} \sum_{\substack{b,d,r \leq T^2 \\ al|dr \\ v|abl}} (abl)^{7/32} \left| \sum_{\substack{K \leq \ell < 2K \\ J \leq j < 2J}} A_f(\ell, 1) \sum_{\mathfrak{a}} \int_{|t| < T^\epsilon} \hat{\Phi}(t) \frac{\sqrt{j\ell drv}}{\cosh(\pi t)} \tau_{\mathfrak{a}}(\ell v, t) \overline{\tau_{\mathfrak{a}}(-jdr, t)} dt \right| \end{aligned} \tag{9.23}$$

for any positive integers

$$J < \frac{M^{1/2}}{T^{1-\epsilon} N^{1/2} drv}, \quad K < \frac{N^{3/2} M^{1/2}}{a^2 b^2 l^2 T^{1-\epsilon}} \tag{9.24}$$

with

$$JK > T^{1/10}.$$

For this last assumption, see the remark following (9.17). We only treat the first line of (9.23) as the rest are similar.

Let

$$X = \frac{\sqrt{NM}}{T^{1-\epsilon}abl\sqrt{JKrdv}},$$

so that $\Phi(w)$ is supported on $X^{-1} \ll w \ll X^{-1}$, by (9.18). We first reduce to the case

$$bdr \ll T^\epsilon \quad \text{and} \quad X \ll T^\epsilon.$$

Non-exceptional eigenvalues. Consider the contribution of $t_g \in \mathbb{R}$ to the first line of (9.23). We have the bound

$$|\hat{\Phi}(t_g)| \ll T^{-1+\epsilon},$$

by (9.21) and (3.22). Using the Cauchy–Schwarz inequality, and enlarging the spectral sum to $|t_g| < X$ (recall that $X \gg T^\epsilon$ by (9.20)), we would like to say that the contribution of the non-exceptional eigenvalues to the first line of (9.23) is bounded by

$$\frac{T^\epsilon}{\sqrt{NM}} \sum_{\substack{b,d,r \leq T^2 \\ \frac{al}{dr} \\ v|abl}} (abl)^{7/32} \frac{l}{dr} \left(\sum_{\substack{g \in \mathfrak{B}_0(dr/l) \\ |t_g| < X}} \left| \sum_j \alpha_j \rho_g(j) \right|^2 \right)^{1/2} \left(\sum_{\substack{g \in \mathfrak{B}_0(dr/l) \\ |t_g| < X}} \left| \sum_{\mathfrak{k}} \beta_{\mathfrak{k}} \rho_g(\mathfrak{k}) \right|^2 \right)^{1/2}, \tag{9.25}$$

where

$$\alpha_j = \begin{cases} \left(\frac{dr}{l} \frac{4\pi j}{\cosh(\pi t_g)} \right)^{1/2} & \text{for } Jdr \leq j < 2Jdr, j \equiv 0 \pmod{dr}, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\beta_{\mathfrak{k}} = \begin{cases} A\left(\frac{\mathfrak{k}}{v}, 1\right) \left(\frac{dr}{l} \frac{4\pi \mathfrak{k}}{\cosh(\pi t_g)} \right)^{1/2} & \text{for } Kv \leq \mathfrak{k} < 2Kv, \mathfrak{k} \equiv 0 \pmod{v}, \\ 0 & \text{otherwise.} \end{cases}$$

However, although (9.25) is essentially the right upper bound, to make the argument rigorous we must remember that $\hat{\Phi}(t_g)$ depends implicitly on j and \mathfrak{k} . We can separate variables as follows. Write

$$\hat{\Phi}(t_g) = \hat{\Phi}(t_g, j, \mathfrak{k}).$$

Insert the factors $U_1(j/J)$ and $U_2(\mathfrak{k}/K)$, where $U_1(x_1)$ and $U_2(x_2)$ are smooth bump functions that are compactly supported on $(\frac{1}{2}, \frac{3}{2})$ and equal 1 on $[1, 2]$. Using the Mellin transform, we have

$$\begin{aligned} & U_1\left(\frac{j}{J}\right) U_2\left(\frac{\mathfrak{k}}{K}\right) \hat{\Phi}(t_g, j, \mathfrak{k}) \\ &= \frac{1}{(2\pi i)^2} \int_{(\epsilon)} \int_{(\epsilon)} \left(\frac{j}{J}\right)^{-s} \left(\frac{\mathfrak{k}}{K}\right)^{-s} \\ & \quad \times \int_{1/2}^{3/2} \int_{1/2}^{3/2} x_1^{s_1} x_2^{s_2} U_1(x_1) U_2(x_2) \hat{\Phi}(t_g, x_1 J, x_2 K) \frac{dx_1}{x_1} \frac{dx_2}{x_2} ds_1 ds_2. \end{aligned}$$

We can restrict the integrals to $|\Im(s_1)|, |\Im(s_2)| < T^\epsilon$ by integrating by parts (using (9.22)).

By the spectral large sieve [IK04, Theorem 7.24, equation (7.40)] and (3.3), we have that (9.25) is bounded by

$$\frac{T^\epsilon}{\sqrt{NM}} \sum_{\substack{b,d,r \leq T^2 \\ al|dr \\ v|abl}} (abl)^{7/32} \frac{l}{dr} \left(\left(\frac{dr}{l} X^2 + Jdr \right) J \right)^{1/2} \left(\left(\frac{dr}{l} X^2 + Kv \right) K \right)^{1/2}. \tag{9.26}$$

We expand this out and look at the cross terms one by one. We have

$$\frac{T^\epsilon}{\sqrt{NM}} \sum_{\substack{b,d,r \leq T^2 \\ al|dr \\ v|abl}} (abl)^{7/32} X^2 (JK)^{1/2} \ll \sum_{\substack{b,d,r \leq T^2 \\ al|dr \\ v|abl}} (abl)^{7/32} \frac{(NM)^{1/2}}{T^{2-\epsilon} a^2 b^2 l^2 r d v (JK)^{1/2}}. \tag{9.27}$$

Now, because $N \ll T^{1+\epsilon}$ and $M \ll T^{3+\epsilon}$ and we assume $JK > T^{1/10}$, this is bounded by a negative power of T . The next cross term is

$$\frac{T^\epsilon}{\sqrt{NM}} \sum_{\substack{b,d,r \leq T^2 \\ al|dr \\ v|abl}} (abl)^{7/32} \left(\frac{l}{dr} \right)^{1/2} X J^{1/2} K v^{1/2} \ll \sum_{\substack{b,d,r \leq T^2 \\ al|dr \\ v|abl}} (abl)^{7/32} \frac{N^{3/4} M^{1/4}}{r d a^2 b^2 l^{3/2} T^{3/2-\epsilon}}. \tag{9.28}$$

Thus, the left-hand side is bounded by a negative power of T unless $N \gg T^{1-\epsilon}$, $M \gg T^{3-\epsilon}$, $abl \ll T^\epsilon$, and

$$K \gg T^{-\epsilon} \frac{N^{3/2} M^{1/2}}{a^2 b^2 l^2 T} \gg T^{2-\epsilon}.$$

The next cross term is

$$\frac{T^\epsilon}{\sqrt{NM}} \sum_{\substack{b,d,r \leq T^2 \\ al|dr \\ v|abl}} (abl)^{7/32} \left(\frac{l}{dr} \right)^{1/2} X K^{1/2} J (dr)^{1/2} \ll \sum_{\substack{b,d,r \leq T^2 \\ al|dr \\ v|abl}} (abl)^{7/32} \frac{M^{1/4}}{T^{3/2-\epsilon} N^{1/4} abl^{1/2} dr v}. \tag{9.29}$$

This is less than a negative power of T since $M \ll T^{3+\epsilon}$. The final cross term is

$$\frac{T^\epsilon}{\sqrt{NM}} \sum_{\substack{b,d,r \leq T^2 \\ al|dr \\ v|abl}} (abl)^{7/32} \frac{lv^{1/2}}{(dr)^{1/2}} JK \ll \sum_{\substack{b,d,r \leq T^2 \\ al|dr \\ v|abl}} (abl)^{7/32} \frac{(NM)^{1/2}}{T^{2-\epsilon} a^2 b^2 l (dr)^{3/2} v^{1/2}}. \tag{9.30}$$

So the left-hand side is bounded by a negative power of T unless $N \gg T^{1-\epsilon}$, $M \gg T^{3-\epsilon}$, $abdr \ll T^\epsilon$ (which implies $l \ll T^\epsilon$) and

$$K \gg T^{-\epsilon} \frac{N^{3/2} M^{1/2}}{a^2 b^2 l^2 T} \gg T^{2-\epsilon}.$$

We conclude that (9.26) is bounded by a negative power of T unless

$$T^{1-\epsilon} \ll N \ll T^{1+\epsilon}, \quad T^{3-\epsilon} \ll M \ll T^{3+\epsilon}, \quad b \ll T^\epsilon, \quad T^{2-\epsilon} \ll K \ll T^{2+\epsilon},$$

which is the case we consider now. In these ranges, we deduce from (9.24) that we must have

$$J \ll T^\epsilon, \quad dr \ll T^\epsilon.$$

The last bound holds because were the contrary true, we would have $J \ll T^{-\epsilon}$ and that is impossible for a positive integer. It follows that we also have $a, l, v \ll T^\epsilon$, since these are divisors of small quantities. In summary, we can assume

$$T^{1-\epsilon} \ll N \ll T^{1+\epsilon}, \quad T^{3-\epsilon} \ll M \ll T^{3+\epsilon}, \quad J \ll T^\epsilon, \quad T^{2-\epsilon} \ll K \ll T^{2+\epsilon}, \quad ablvdr \ll T^\epsilon. \tag{9.31}$$

From this we get

$$X \ll T^\epsilon. \tag{9.32}$$

Exceptional eigenvalues. We now consider the contribution of $t_g \in (-\frac{7}{64}i, \frac{7}{64}i)$ to the first line of (9.23). This time we have the bound

$$|\hat{\Phi}(t_g)| \ll T^{-1+\epsilon} X^{7/32}$$

by (9.21) and (3.23). By the Cauchy–Schwarz inequality, the contribution of the possible exceptional eigenvalues is (essentially) bounded by

$$\frac{T^\epsilon}{\sqrt{NM}} \sum_{\substack{b,d,r \leq T^2 \\ a|dr \\ v|abl}} (ablX)^{7/32} \frac{l}{dr} \left(\sum_{\substack{g \in \mathfrak{B}_0(dr/l) \\ |t_g| < 1}} \left| \sum_j \alpha_j \rho_g(j) \right|^2 \right)^{1/2} \left(\sum_{\substack{g \in \mathfrak{B}_0(dr/l) \\ |t_g| < 1}} \left| \sum_{\mathfrak{k}} \beta_{\mathfrak{k}} \rho_g(\mathfrak{k}) \right|^2 \right)^{1/2},$$

where α_j and $\beta_{\mathfrak{k}}$ are as above. By the spectral large sieve again, this is bounded by

$$\frac{T^\epsilon}{\sqrt{NM}} \sum_{\substack{b,d,r \leq T^2 \\ a|dr \\ v|abl}} (ablX)^{7/32} \frac{l}{dr} \left(\left(\frac{dr}{l} + Jdr \right) J \right)^{1/2} \left(\left(\frac{dr}{l} + Kv \right) K \right)^{1/2}. \tag{9.33}$$

We know from (9.20) that $X > T^\epsilon$. Expand (9.33) and consider each cross term. The cross terms corresponding to (9.27)–(9.29) are clearly smaller in size, because each of (9.27)–(9.29) has a factor of X or X^2 , while the cross terms of (9.33) have a factor of $X^{7/32}$ only. Thus it remains only to consider the cross term

$$\frac{T^\epsilon}{\sqrt{NM}} \sum_{\substack{b,d,r \leq T^2 \\ a|dr \\ v|abl}} (ablX)^{7/32} \frac{lv^{1/2}}{(dr)^{1/2}} JK.$$

Since $X > 1$, this is less than

$$\frac{T^\epsilon}{\sqrt{NM}} \sum_{\substack{b,d,r \leq T^2 \\ a|dr \\ v|abl}} (abl)^{7/32} \frac{lv^{1/2}}{(dr)^{1/2}} X^2 JK \ll \sum_{\substack{b,d,r \leq T^2 \\ a|dr \\ v|abl}} (abl)^{7/32} \frac{(NM)^{1/2}}{T^{2-\epsilon} a^2 b^2 l (dr)^{3/2} v^{1/2}}.$$

This is the same as (9.30) and so we arrive at the same conclusions (9.31)–(9.32).

Summary. We have reduced everything to proving that the following sum is bounded by a negative power of T :

$$\frac{T}{\sqrt{NM}} \sum_{K < k < 2K} \hat{\Phi}(t_g) A_f(k, 1) \frac{\sqrt{j k d r v}}{\cosh(\pi t_g)} \rho_g(k v) \overline{\rho_g(j d r)}$$

for any Maass cusp form g of level $q \ll T^\epsilon$ and spectral parameter $t_g \ll T^\epsilon$, and any positive integers $T^{2-\epsilon} < K < T^{2+\epsilon}$, $j, v, d, r \ll T^\epsilon$. In the new ranges we have (remember that $\hat{\Phi}(t_g)$ depends implicitly on k) that

$$\frac{d}{dk} \hat{\Phi}(t_g) \ll T^{-1+\epsilon} k^{-1},$$

by (9.22). Thus by partial summation it suffices to show that

$$T^{-2+\epsilon} \sum_{K < k < 2K} A_f(k, 1) \frac{\sqrt{j k d r v}}{\cosh(\pi t_g)} \rho_g(k v) \overline{\rho_g(j d r)} \ll T^{-\delta} \tag{9.34}$$

for some $\delta > 0$.

9.4 Generalized Lindelöf hypothesis

In order to use L -functions to obtain the required cancellation in (9.34), we must work with primitive Hecke cusp forms. To this end, we can take g to be an element of the special basis described in Lemma 3.7. That is, $g = h_c$ for some newform h of level dividing q . Writing out h_c has a linear combination of $h|_b$ as in the lemma, and using the fact that the coefficients in this linear sum are small, it suffices to prove that

$$T^{-2+\epsilon} \sum_{K < k < 2K} A_f(k, 1) \frac{\sqrt{j k d r v}}{\cosh(\pi t_h)} \rho_{h|_b}(k v) \overline{\rho_{h|_b}(j d r)} \ll T^{-\delta} \tag{9.35}$$

for some newform h of level $q'|q$ (so that $q' < T^\epsilon$) and spectral parameter $t_h = t_g < T^\epsilon$, and some positive integer $b < T^\epsilon$. If $\rho_h(n)$ are the Fourier coefficients of h , then the Fourier coefficients of $h|_b$ (see definition (3.27)) are $\rho_h(n/b)$ for $b|n$ and 0 otherwise. Thus, it suffices to prove that

$$T^{-2+\epsilon} \sum_{K < k < 2K} A_f(bk, 1) \frac{\sqrt{j k d r v}}{\cosh(\pi t_h)} \rho_h(k v) \overline{\rho_h(j d r)} \ll T^{-\delta}$$

for some $\delta > 0$. We have

$$\frac{1}{\cosh(\pi t_g)} \sqrt{k v} \rho_h(k v) \sqrt{j d r} \overline{\rho_h(j d r)} = \frac{1}{\cosh(\pi t_h)} |\rho_h(1)|^2 \lambda_h(k v) \overline{\lambda_h(j d r)},$$

where $\lambda_h(n)$ are the Hecke eigenvalues of h . By the standard (Rankin–Selberg) bound

$$\frac{|\rho_h(1)|^2}{\cosh(\pi t_h)} \ll T^\epsilon,$$

we find that it suffices to prove that

$$T^{-2+\epsilon} \sum_{K < k < 2K} A_f(bk, 1) \lambda_h(k v) \ll T^{-\delta}$$

for some $\delta > 0$. Thus the trivial bound is $O(T^\epsilon)$, and to obtain further cancellation it suffices to partition the interval $[1, 2]$ using smooth bump functions ψ and bound the following by a negative power of T :

$$\begin{aligned} & T^{-2+\epsilon} \sum_{k \geq 1} \psi\left(\frac{k}{K}\right) A_f(bk, 1) \lambda_h(kv) \\ &= T^{-2+\epsilon} \frac{1}{2\pi i} \int_{(2)} \tilde{\psi}(s) K^s D(s) \frac{ds}{s}, \end{aligned} \tag{9.36}$$

where $\tilde{\psi}$ denotes the Mellin transform of ψ and

$$D(s) = \sum_{k \geq 1} \frac{A_f(bk, 1) \lambda_h(kv)}{k^s}.$$

For any integers n, m , let $n|m^\infty$ mean that $p|n \Rightarrow p|m$. By Hecke multiplicativity, we have

$$D(s) = \sum_{\substack{k \geq 1 \\ (k, bv)=1}} \frac{A_f(k, 1) \lambda_h(k)}{k^s} \sum_{\substack{k \geq 1 \\ k|(bv)^\infty}} \frac{A_f(bk, 1) \lambda_h(kv)}{k^s}.$$

Comparing Euler products on both sides, we may write

$$\sum_{\substack{k \geq 1 \\ (k, bv)=1}} \frac{A_f(k, 1) \lambda_h(k)}{k^s} = L(s, \text{sym}^2 f \times h) G_1(s),$$

where $G_1(s)$ is a Dirichlet series which absolutely converges for $\Re s \geq 1 - \delta$ for some $\delta > 0$ and satisfies $G_1(s) \ll 1$ in this half-plane. For any integer n and prime p , let n_p denote the p -part of n . That is, $n_p|n$ and $n_p p \nmid n$. We write

$$G_2(s) = \prod_{p|bv} \left(\frac{A(b_p, 1) \lambda_h(v_p)}{1} + \frac{A(b_p p, 1) \lambda_h(v_p p)}{p^s} + \frac{A(b_p p^2, 1) \lambda_h(v_p p^2)}{p^{2s}} + \dots \right)$$

and note that for $\Re(s) \geq \frac{1}{2}$, we have

$$G_2(s) \ll b^{7/32+\epsilon} v^{7/64+\epsilon} \exp\left(\sum_{p < \log bv} p^{-1/2+7/32+7/64+\epsilon} \right) \ll T^\epsilon.$$

Thus

$$D(s) = L(s, \text{sym}^2 f \times h) G(s),$$

where $G(s) = G_1(s) G_2(s)$ is a Dirichlet series which absolutely converges for $\Re s \geq 1 - \delta$ for some $\delta > 0$ and satisfies $G(s) \ll T^\epsilon$ in this half-plane. Thus, moving the line of integration in (9.36) to $\Re(s) = 1 - \delta$ and applying the GLH bound $L(s, \text{sym}^2 f \times h) \ll (1 + |s|)^\epsilon T^\epsilon$ there completes the proof (we save an absolute power of T , and all ϵ values can be adjusted). Note that what is actually required here is any subconvexity bound for $L(s, \text{sym}^2 f \times h)$.

ACKNOWLEDGEMENT

We thank Gergely Harcos for his encouragement to work on this problem.

REFERENCES

- Blo12 V. Blomer, *Subconvexity for twisted L -functions on $GL(3)$* , Amer. J. Math. **134** (2012), 1385–1421.
- Blo07 V. Blomer, G. Harcos and P. Michel, *Bounds for modular L -functions in the level aspect*, Ann. Sci. Éc. Norm. Supér. (4) **40** (2007), 697–740.
- Blo13 V. Blomer, R. Khan and M. Young, *Distribution of mass of holomorphic cusp forms*, Duke Math. J. **162** (2013), 2609–2644.
- BM15 V. Blomer and D. Milićević, *The second moment of twisted modular L -functions*, Geom. Funct. Anal. **25** (2015), 453–516.
- BK17 J. Buttcane and R. Khan, *A mean value of a triple product of L -functions*, Math. Z. **285** (2017), 565–591.
- Gol06 D. Goldfeld, *Automorphic forms and L -functions for the group $GL(n, \mathbf{R})$* , Cambridge Studies in Advanced Mathematics, vol. 99 (Cambridge University Press, Cambridge, 2006); with an appendix by Kevin A. Broughan.
- GL06 D. Goldfeld and X. Li, *Voronoi formulas on $GL(n)$* , Int. Math. Res. Not. IMRN **2006** (2006), Art. ID 86295, 25.
- GR00 I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, sixth edition (Academic Press Inc., San Diego, CA, 2000); translated from the Russian, edited and with a preface by Alan Jeffrey and Daniel Zwillinger.
- HR92 D. A. Hejhal and B. N. Rackner, *On the topography of Maass waveforms for $PSL(2, \mathbf{Z})$* , Exp. Math. **1** (1992), 275–305.
- HS01 D. A. Hejhal and A. Strömbergsson, *On quantum chaos and Maass waveforms of CM-type*, Found. Phys. **31** (2001), 519–533; invited papers dedicated to Martin C. Gutzwiller, Part IV.
- HL94 J. Hoffstein and P. Lockhart, *Coefficients of Maass forms and the Siegel zero*, Ann. of Math. (2) **140** (1994), 161–181; with an appendix by Dorian Goldfeld, Hoffstein and Daniel Lieman.
- Iwa92 H. Iwaniec, *The spectral growth of automorphic L -functions*, J. Reine Angew. Math. **428** (1992), 139–159.
- Iwa97 H. Iwaniec, *Topics in classical automorphic forms*, Graduate Studies in Mathematics, vol. 17 (American Mathematical Society, Providence, RI, 1997).
- IK04 H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53 (American Mathematical Society, Providence, RI, 2004).
- ILS00 H. Iwaniec, W. Luo and P. Sarnak, *Low lying zeros of families of L -functions*, Publ. Math. Inst. Hautes Études Sci. **91** (2000), 55–131.
- Kim03 H. H. Kim, *Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2* , J. Amer. Math. Soc. **16** (2003), 139–183; with Appendix 1 by Dinakar Ramakrishnan and Appendix 2 by Kim and Peter Sarnak.
- LR03 E. Lapid and S. Rallis, *On the nonnegativity of $L(\frac{1}{2}, \pi)$ for SO_{2n+1}* , Ann. of Math. (2) **157** (2003), 891–917.
- Lap03 E. M. Lapid, *On the nonnegativity of Rankin–Selberg L -functions at the center of symmetry*, Int. Math. Res. Not. IMRN **2003** (2003), 65–75.
- Li09 X. Li, *The central value of the Rankin–Selberg L -functions*, Geom. Funct. Anal. **18** (2009), 1660–1695.
- Luo14 W. Luo, *L^4 -norms of the dihedral Maass forms*, Int. Math. Res. Not. IMRN **2014** (2014), 2294–2304.
- MS06 S. D. Miller and W. Schmid, *Automorphic distributions, L -functions, and Voronoi summation for $GL(3)$* , Ann. of Math. (2) **164** (2006), 423–488.
- Mot97 Y. Motohashi, *Spectral theory of the Riemann zeta-function*, Cambridge Tracts in Mathematics, vol. 127 (Cambridge University Press, Cambridge, 1997).

- Rou11 D. Rouymi, *Formules de trace et non-annulation de fonctions L automorphes au niveau \mathfrak{p}^ν* , *Acta Arith.* **147** (2011), 1–32.
- Sar03 P. Sarnak, *Spectra of hyperbolic surfaces*, *Bull. Amer. Math. Soc. (N.S.)* **40** (2003), 441–478.
- ST09 P. Sarnak and J. Tsimerman, *On Linnik and Selberg’s conjecture about sums of Kloosterman sums*, in *Algebra, arithmetic, and geometry: in honor of Yu. I. Manin, Vol. II*, *Progress in Mathematics*, vol. 270 (Birkhäuser, Boston, MA, 2009), 619–635.
- Spi03 F. Spinu, *The L^4 norm of the Eisenstein series*, PhD thesis, Princeton University (ProQuest LLC, Ann Arbor, MI, 2003).
- Wat02 T. C. Watson, *Rankin triple products and quantum chaos*, PhD thesis, Princeton University (ProQuest LLC, Ann Arbor, MI, 2002).
- Yos97 E. Yoshida, *Remark on the Kuznetsov trace formula*, in *Analytic number theory (Kyoto, 1996)*, *London Mathematical Society Lecture Note Series*, vol. 247 (Cambridge University Press, Cambridge, 1997), 377–382.

Jack Buttcane buttcane@buffalo.edu
 Mathematics Department, 244 Mathematics Building,
 University at Buffalo, Buffalo, NY 14260, USA

Rizwanur Khan rizwanur.khan@qatar.tamu.edu
 Science Program, Texas A&M University at Qatar,
 PO Box 23874, Doha, Qatar