

Mollification of the fourth moment of automorphic L -functions and arithmetic applications

E. Kowalski, P. Michel, J. VanderKam

March 16, 2003

Abstract

In this paper, we compute the twisted fourth power moment of modular L -functions of prime level near the critical line. This allows us to prove some new non-vanishing results on the central values of automorphic L -functions, in particular those obtained by base change from $GL_2(\mathbf{Q})$ to $GL_2(K)$ for K a cyclic field of low degree.

Contents

1	Introduction	1
2	Auxiliary Lemmas	9
3	The twisted fourth moment	11
4	The off-off-diagonal main term	18
5	Mollification at the center of the critical strip	32
6	Removing the harmonic weight	42
7	Simultaneous non-vanishing	45
8	Appendix: Removing the exponential factor	47

1 Introduction

There has been considerable recent research, most notably the pioneering work of Iwaniec and Sarnak, into the asymptotics of the second moments of families of L -functions in their critical strip. A particular family of interest has been forms of fixed weight (usually two) and high level (usually prime). In this case, the central value of these L -functions has a deep relation to the arithmetic of the underlying modular curve. In this paper we increase the degree of difficulty and extend these methods to finding asymptotics for the fourth moment of this family. This allows for some new non-vanishing results, and suggests a framework for providing several others.

Let q be a prime number, and let $S_2(q)^*$ denote the set of primitive Hecke eigenforms of weight 2 relative to the subgroup $\Gamma_0(q)^1$. Any $f(z) \in S_2(q)^*$ admits a Fourier expansion of the form

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

where we normalize so that $\lambda_f(1) = 1$. The coefficients $\lambda_f(n)$ are real and moreover the $\sqrt{n} \lambda_f(n)$ for $n \geq 1$ generate a totally real number field K_f (see [Sh1]). To f is attached an L -function with Euler product

$$L(f, s) := \sum_{n \geq 1} \lambda_f(n) n^{-s} = \prod_p \left(1 - \frac{\lambda_f(p)}{p^s} + \frac{\varepsilon_q(p)}{p^{2s}} \right)^{-1}, \quad (1)$$

where $\varepsilon_q(\cdot)$ is the trivial character mod q . This series is absolutely convergent when $\Re s$ is large, and admits analytic continuation to all of \mathbf{C} . It satisfies the functional equation

$$\Lambda(f, s) = \varepsilon_f \Lambda(f, 1 - s), \quad (2)$$

where

$$\Lambda(f, s) := \widehat{q}^s \Gamma\left(s + \frac{1}{2}\right) L(f, s), \quad \widehat{q} := \frac{q^{1/2}}{2\pi}, \quad \varepsilon_f = -q^{1/2} \lambda_f(q) = \pm 1.$$

Throughout this paper, we use two methods of averaging over $S_2(q)^*$, the ‘‘harmonic’’ and ‘‘natural,’’ defined as

$$\sum_f^h \alpha_f := \sum_{f \in S_2(q)^*} \frac{\alpha_f}{4\pi(f, f)}, \quad \sum_f^n \alpha_f := \sum_{f \in S_2(q)^*} \frac{\alpha_f}{|S_2(q)^*|},$$

with (f, g) the Petersson inner product on the space $\Gamma_0(q) \backslash \mathbf{H}$. We will also frequently use the notation \sum_f^* to denote results which apply to both averages. The focus of this paper is on the twisted harmonic fourth power moment of $L(f, s)$ on the critical line:

$$M(\ell) := \sum_f^h |L(f, s)|^4 \lambda_f(\ell), \quad s = 1/2 + \mu, \quad (3)$$

with μ imaginary (to save notation, we suppress the dependence of $M(\ell)$ on μ). We should point out that the methods of this paper also apply when $1/2 + \mu$ is away from the critical line, but much simpler methods are sufficient unless $\Re \mu$ is close to zero.

The question of bounding $M(\ell)$ was considered by Duke-Friedlander-Iwaniec [DFI II], and for $\ell \ll q^{1/15}$ they obtained (essentially) the best possible bound:

$$M(\ell) \ll_{\varepsilon, \mu} q^\varepsilon \ell^{-1/2} \quad (4)$$

for all $\varepsilon > 0$. The main value in studying the twisted moments is that it enables us to evaluate on average the terms $|L(f, s)|^4 M(f, s)$ with

$$M(f, s) := \sum_{m \leq M} \frac{x_m}{m^s} \lambda_f(m)$$

¹we have chosen to restrict ourselves to this kind of modular forms because in that case $S_2(q)^*$ is (rather than being included in) an orthogonal basis for the Hilbert space of cuspidal forms of weight 2 and level q : this allows to apply directly Petersson’s formula, saving us a lot of technical complications

a finite Dirichlet polynomial. There are two very different applications for this, depending on the choice of the coefficients x_m .

- Amplification: the x_m 's can be chosen so that, for a specific $f_0 \in S_2(q)^*$, $M(f_0, s)$ is large. This was developed by Friedlander and Iwaniec to obtain (among other things) improvements in the convexity bounds for L -functions on the critical line. For this application, a bound of the type (4) is sufficient, and as a corollary one obtains the following deep result of [DFI II]:

Theorem 1.1 *For some $B > 0$ and for all $\mu \in i\mathbf{R}$ the following bound holds²: for all $\varepsilon > 0$*

$$L(f, 1/2 + \mu) \ll_{\varepsilon, \mu} e^{\pi|\mu|/2} (1 + |\mu|)^B q^{1/4 - 1/192 + \varepsilon}.$$

- Mollification: the x_m 's can be chosen so that the $M(f, s)$'s compensate for the deviation of $L(f, s)$ from its average size. This was invented by Landau and Bohr, and used very fruitfully by Selberg, then Levinson and Conrey to study the zeros of Riemann's ζ function. Very recently, this technique has been used (for the second power moment) in the context of automorphic L -functions to obtain non-vanishing results for these L -functions at the center of the critical strip or unconditional bounds for the average order of vanishing of $L(f, s)$ at $s = 1/2$ [IS, KM1, KM2, KMV, V1, V2]. For mollification, a bound like (4) is not sufficient and we need a precise asymptotic expansion for $M(\ell)$.

Our first step in this asymptotic expansion is to isolate the main terms for $M(\ell)$, which we do by following [DFI II] closely:

Theorem 1.2 *There exists some $B > 0$ such that if $l < q$ and $\mu \in i\mathbf{R}$, we have for $s = 1/2 + \mu$ and any $\varepsilon > 0$*

$$M(\ell) = M^D(\ell) + M^{OD}(\ell) + M^{OOD}(\ell) + O_\varepsilon(q^\varepsilon e^{2\pi|\mu|} (1 + |\mu|)^B (\ell^{3/4} q^{-1/12} + \ell^{17/8} q^{-1/4} + \ell q^{-1/6})).$$

Here $M^D(\ell)$ is the “diagonal” main term defined in (18), $M^{OD}(\ell)$ is the “off-diagonal” main term defined in (32) if $\mu \neq 0$, (33) if $\mu = 0$, and $M^{OOD}(\ell)$ is the “off-off-diagonal” main term defined in (35).

From the definitions of $M^D(\ell)$ and $M^{OD}(\ell)$ it is not too difficult to infer that they are $O_\varepsilon(e^{2\pi|\mu|} (1 + |\mu|)^B q^\varepsilon \ell^{-1/2})$. It turns out that $M^{OOD}(\ell)$ satisfies the same bound (up to some remainder terms which are dominated by those of Theorem 1.2), and this is sharp. Another proof is given (in greater generality) in the Erratum [DFI III]: in the original paper [DFI II], this term was incorrectly bounded by $q^\varepsilon \ell^{-1}$, see the remark at the beginning of section 3.4 below).

The bulk of this work is the evaluation of these main terms, the most difficult being $M^{OOD}(\ell)$. An asymptotic expansion for $M^D(\ell) + M^{OD}(\ell)$ is given at the end of section 3, while one for $M^{OOD}(\ell)$ is given in Theorem 4.1. Using the specific values $\ell = 1$, $\mu = 0$ in these asymptotics, one can obtain an asymptotic formula for the fourth moment:

²The dependance in μ was not stated explicitly in [DFI II] but can be checked either by looking at the arguments of [DFI II] or from the present paper.

Corollary 1.3 For all $\varepsilon > 0$

$$M(1) = \sum_f^h L(f, 1/2)^4 = P(\log q) + O_\varepsilon(q^{-1/12+\varepsilon}),$$

with P a polynomial of degree 6 and leading coefficient $\frac{1}{60\pi^2}$.

Remark. As remarked by the referee, there are very few other examples of families of L -functions where the fourth moment at the critical point can be evaluated precisely: firstly the fourth moment of Zeta on the critical line by Ingham [I1] (see also [HB1] where the error term saves an exponent) and secondly the fourth moment of Dirichlet L -function by Heath-Brown [HB2] (where the error term save a power of logarithm). Moreover a distinctive feature of our family is that it should admit an “orthogonal” symmetry (in the terminology of Katz-Sarnak [KS]) in contrast with the “unitary” symmetry enjoyed by the above mentioned families. In particular the value $1/60\pi^2$ may be relevant in the context of the theory of the moments for L -functions with orthogonal symmetry as it provides an “experimental data” to test the theoretical model (see the paper of Conrey and Farmer [CF]).

Remark. the exponent $1/12$ matches the one obtained by Estermann [E] for the additive divisor problem:

$$\sum_{n \leq x} \tau(n)\tau(n+h) = xP_h(\log x)(1 + O_{\varepsilon,h}(x^{-1/12+\varepsilon})) \text{ for } P_h \text{ some quadratic polynomial;}$$

this is not a coincidence since Estermann’s estimate rely on Weil’s bound for Kloostermann’s sums, and the present error term which is obtained directly from [DFI II], ultimately depends on the wide generalization of Estermann’s result [DFI I] thus also relying on Weil’s bound (see also the works of Deshouillers-Iwaniec and Motohashi [DI, Mo] which far beyond this result by using the spectral theory of automorphic forms).

More generally, these asymptotics enable us to input a mollifier $M(f, s)$ whose length is a small positive power of q .

Theorem 1.4 Set $M = \hat{q}^\Delta$, and let $M(f)$ be the mollifier defined in Section 5, equation (64) with the choice $P(x) = x^3$. Let $F(\Delta)$ denote the function

$$F(\Delta) = \frac{1024}{\Delta^{10}} \left(648\left(\frac{\Delta}{2}\right)^6 + \frac{51696}{35}\left(\frac{\Delta}{2}\right)^5 + \frac{6408}{5}\left(\frac{\Delta}{2}\right)^4 + \frac{2772}{5}\left(\frac{\Delta}{2}\right)^3 + 126\left(\frac{\Delta}{2}\right)^2 + \frac{72}{5}\left(\frac{\Delta}{2}\right) + \frac{27}{10} \right) \quad (5)$$

Then for all $0 < \Delta < 1/30$ we have

$$\sum_f^h L(f, 1/2)^4 M(f)^4 = (1 + o(1))F(\Delta) \left(\frac{\zeta(2)}{\log q} \right)^4 \quad (6)$$

For the natural average, the same is true in the smaller range $0 < \Delta < 1/30$:

$$\sum_f^n L(f, 1/2)^4 M(f)^4 = (1 + o(1))F(\Delta) \left(\frac{\zeta(2)}{\log q} \right)^4. \quad (7)$$

Remark. The discrepancy in the range of validity between the harmonic and the natural average can be explained by the fact (first noted by H. Iwaniec) that for the fourth power moment removing the harmonic weight has a cost (which was not the case for the first and the second moments). In particular, breaking the “convexity” bound for the size of $L(f, 1/2)$ is needed, so that Theorem 1.1 is essential to our argument. As pointed out in that paper, if one could show that λ_f is not too small too often, more precisely, if

$$\sum \lambda_f(\ell)^2 \gg Lq^{-\epsilon}$$

(the sum over $\ell < L$ prime to q), then one could replace $1/192$ by $1/120$ in Theorem 1.1, and then the range of validity would be $\Delta < 1/30$ for the natural average also. In section 6 we explain in greater detail how one passes from the harmonic to the natural average.

Remark. Our methods give a similar result for mollified moments at points other than $1/2$. In particular, $F(\Delta)$ is replaced by a function $F_\mu(\Delta)$ which is uniformly bounded for μ near the critical line (for $|\mu|$ large it will have lower degree in $\log q$, but approach (7) as $\mu \rightarrow 0$). Lacking a useful interpretation of these values, we do not reproduce the proof in detail here.

1.1 Applications to non-vanishing

As is usually the case, our evaluations of mollified moments can be used to infer non-vanishing results for large fractions of a family at the central point. The fourth moment does not contribute anything to the non-vanishing of the original family $S_2(q)^*$ (this is analyzed thoroughly in [IS]), but it does allow us to consider non-vanishing of products of twists. Let χ be a fixed primitive character of conductor D , coprime with q , and consider the twisted L function

$$L(f \otimes \chi, s) = \sum_n \frac{\lambda_f(n)\chi(n)}{n^s} = \prod_p \left(1 - \frac{\lambda_f(p)\chi(p)}{p^s} + \frac{\chi^2(p)}{p^{2s}}\right)^{-1}.$$

This L function comes from a modular form of level qD^2 and nebentypus χ^2 (see [I1]) and thus satisfies the functional equation:

$$\Lambda(f \otimes \chi, s) = \chi(-q)\varepsilon_f \Lambda(f \otimes \bar{\chi}, 1 - s), \quad (8)$$

where

$$\Lambda(f \otimes \chi, s) := \widehat{q}_D^s \Gamma(s + \frac{1}{2}) L(f \otimes \chi, s), \quad \widehat{q}_D := \frac{q^{1/2} D}{2\pi}.$$

We use Theorem 1.4 along with the techniques of [IS] and [KMV] to prove non-vanishing of the product $L(f, 1/2)L(f \otimes \chi, 1/2)$ for a positive proportion of the f 's.

Theorem 1.5 *Let χ be a fixed primitive character with $\chi^2 \neq 1$. There exists an absolute constant $c > 0$ (independent of χ) such that as $q \rightarrow +\infty$ through primes,*

$$\sum_{\substack{f \\ L(f, 1/2)L(f \otimes \chi, 1/2) \neq 0}}^* 1 \geq c + o_\chi(1),$$

“*” meaning either the harmonic or the natural average.

The constant c we obtain is extremely small (see section 7 for the exact value). The theorem also applies (with a slightly better constant) when χ is real and $\chi(-q) = 1$, but this result was already obtained by Iwaniec and Sarnak [IS] (with a much larger, although still quite small, constant $c = 1/1812$) using only second moment estimates. Their proof, which is quite clever, depends crucially on some particular features of the sign of the functional equations involved; for complex characters such a favorable situation does not occur, so to handle this question it seems, we have to rely on fourth moment estimates. Note that for χ quadratic and $\chi(-q) = -1$, so that $L(f, 1/2)L(f \otimes \chi, 1/2) = 0$ identically, the results of [IS] and [KM2] may be used, again without fourth-moment analysis, to show that

$$\sum_{f, (L(f)L(f \otimes \chi))'(1/2) \neq 0}^n 1 \geq 3/8 + o_\chi(1);$$

For completeness let's sketch a proof of that fact: since q ranges over the primes such that $\chi(-q) = -1$, we have $(L(f)L(f \otimes \chi))'(1/2) = L(f, 1/2)L'(f \otimes \chi, 1/2)$ if f is even (i.e. $\varepsilon_f = 1$) or $L(f \otimes \chi, 1/2)L'(f, 1/2)$ if f is odd ($\varepsilon_f = -1$). Consider the subset $S_2(q)^{*odd}$ of odd forms, it was proved in [IS] that

$$|\{f \in S_2(q)^{*odd}, L(f \otimes \chi, 1/2) \neq 0\}| \geq (\frac{1}{2} + o(1))|S_2(q)^{*odd}|$$

while it was proved in [KM2] that

$$|\{f \in S_2(q)^{*odd}, L'(f, 1/2) \neq 0\}| \geq (\frac{7}{8} + o(1))|S_2(q)^{*odd}|$$

so that by intersecting the two sets

$$|\{f \in S_2(q)^{*odd}, L(f \otimes \chi, 1/2)L'(f, 1/2) \neq 0\}| \geq (\frac{3}{8} + o(1))|S_2(q)^{*odd}|$$

the same works for even forms by taking the appropriate variants.

One can obtain from Theorem 1.5 some interesting corollaries. It was proven by Shimura [Sh2] that for σ any \mathbf{Q} -linear automorphism of \mathbf{C} ,

$$L(f \otimes \chi, 1/2) \neq 0 \Leftrightarrow L(f^\sigma \otimes \chi^\sigma, 1/2) \neq 0$$

where

$$L(f^\sigma \otimes \chi^\sigma, s) = \sum_n \frac{\sqrt{n}^{-1} \sigma(\lambda_f(n) \sqrt{n}) \sigma(\chi(n))}{n^s}.$$

Let d be the order of χ . If the number field K_f generated by the coefficients of f is linearly disjoint from the d -th cyclotomic field $\mathbf{Q}(\zeta_d)$, then for all χ primitive of order d we have

$$L(f \otimes \chi, 1/2) \neq 0 \Leftrightarrow \prod_{\sigma} L(f \otimes \chi^\sigma, 1/2) \neq 0,$$

where σ ranges over the automorphisms of $\mathbf{Q}(\zeta_d)$.

At a simpler level, since $\overline{L(f \otimes \chi, 1/2)} = L(f \otimes \bar{\chi}, 1/2)$, the twists by χ and $\bar{\chi}$ can only vanish simultaneously.

Hence when $d = 3$, Theorem 1.5 immediately implies the following:

Corollary 1.6 *There exists an absolute constant $c > 0$, such that: for χ a fixed primitive character of order 3, when $q \rightarrow +\infty$ through primes,*

$$\sum_{\substack{f \\ L(f,1/2)L(f \otimes \chi,1/2)L(f \otimes \bar{\chi},1/2) \neq 0}}^* 1 \geq c + o_K(1).$$

“*” meaning either the natural or the harmonic average.

In other words, by the Kronecker-Weber theorem, we have obtained a non-vanishing result for the central value of base change L -functions to cyclic cubic fields:

Corollary 1.7 *Let K/\mathbf{Q} be a cyclic cubic field. There exists an absolute constant $c > 0$ (independent of K) such that as $q \rightarrow +\infty$ through primes,*

$$\sum_{f, L(f_K, 1/2) \neq 0}^* 1 \geq c + o_K(1),$$

where $L(f_K, s)$ is the L -function of the base change of f from $GL_2(\mathbf{Q})$ to $GL_2(K)$.

This is immediate, as $L(f_K, s) = L(f, s)L(f \otimes \chi, s)L(f \otimes \bar{\chi}, s)$ for some primitive character χ of order three.

1.2 Possible extensions

The feasibility of mollifying the fourth powers of $L(f, 1/2)$ indicates that it may be possible to mollify fourth powers of other families of this type, or to mollify the second moment of certain families of degree-four L -functions. In particular, it is not too difficult to replace $L(f, 1/2)^4$ in our methods by $|L(f \otimes \chi_1, 1/2)|^2 |L(f \otimes \chi_2, 1/2)|^2$ for χ_1, χ_2 fixed characters of conductor relatively prime to q (one must replace Proposition 2.3 by a more general expression, as can be found in [I2]) Let us turn now to a brief discussion of some possible uses for this and other extensions of the method.

- In Theorem 1.5, it is much more natural to mollify the square product $|L(f, 1/2)L(f \otimes \chi, 1/2)|^2$, rather than the fourth power $|L(f, 1/2)|^4$, so it is likely that an extension to character twists would significantly improve the value of c .
- A second possibility, using the inequality

$$\left| \sum_f \mathcal{L}_1(f) \mathcal{L}_2(f) \mathcal{L}_3(f) \right| \leq \left(\sum_{\mathcal{L}_1(f) \mathcal{L}_2(f) \mathcal{L}_3(f) \neq 0} 1 \right)^{1/4} \left(\sum_f |\mathcal{L}_1(f)|^4 \cdot \sum_f |\mathcal{L}_2(f)|^4 \cdot \sum_f |\mathcal{L}_3(f)|^4 \right)^{1/4}$$

is to prove non-vanishing of triple products $L(f \otimes \chi_1, 1/2)L(f \otimes \chi_2, 1/2)L(f \otimes \chi_3, 1/2)$. In particular, for χ a primitive character of order 4 or 5, by taking $\mathcal{L}_1(f) = L(f, 1/2)M(f)$, $\mathcal{L}_2(f) = L(f \otimes \chi, 1/2)M(f \otimes \chi)$, and $\mathcal{L}_3(f) = L(f \otimes \chi^2, 1/2)M(f \otimes \chi^2)$ for suitably chosen mollifiers, it is possible to extend Corollary 1.6 to quartic or quintic cyclic number fields (in the former case, under the assumption that the field is not biquadratic) so as to obtain

Let K a galois field of degree ≤ 5 over \mathbf{Q} , which is not biquadratic, then as $q \rightarrow +\infty$, there exists a positive constant c (independent of K) such that as $q \rightarrow +\infty$

$$\sum_{f, L(f_K, 1/2) \neq 0}^* 1 \geq (1 + \chi(-q))c + o_K(1),$$

where $\chi(-q) = 0$ if $[K : \mathbf{Q}]$ is odd, and for $[K : \mathbf{Q}]$ even, χ is the Kronecker character associated to the unique quadratic field contained in K . Moreover if $[K : \mathbf{Q}]$ is even, and q range over the primes such that $\chi(-q) = -1$ (so that $L(f_K, 1/2) = 0$), we have

$$\sum_{f, L'(f_K, 1/2) \neq 0}^* 1 \geq c + o_K(1).$$

To do so also require asymptotics for the first moment of the mollified triple product $L(f \otimes \chi_1, 1/2)M(f \otimes \chi_1)L(f \otimes \chi_2, 1/2)M(f \otimes \chi_2)L(f \otimes \chi_3, 1/2)M(f \otimes \chi_3)$. While these do not appear explicitly in the literature, they are quite similar to the asymptotics of the double product, which we give in section 7 (it should not be a surprise to the reader that a third moment is considerably simpler than a fourth moment). The one feature of triple products which does not arise for double products is that the ε_f factors introduce an additional off-diagonal contribution. This phenomenon is studied in a slightly different setting in [KM2], sections 2.4.2 through 2.4.6.

If we consider a biquadratic extension K/\mathbf{Q} , since there is no conjugation to force simultaneous non-vanishing, one would have be able to mollify sixth moments to give this type of result by similar techniques.

- A third possibility comes from the theory of Heegner cycles on modular curves. For ε_K a quadratic character (corresponding to the field $K = \mathbf{Q}(\sqrt{\varepsilon_K(-1)D})$), the product $L(f, s)L(f \otimes \varepsilon_K, s)$ is again the L function of the base change of f from $GL_2(\mathbf{Q})$ to $GL_2(K)$. For any character χ of the ideal class group of K , one may then consider $L(f_K \otimes \chi, s)$, the L function of the twist of f_K by χ . From the work of Gross, Kohnen, and Zagier, [GZ, G, GKZ], it is known that the central values $L(f_K \otimes \chi, 1/2)$ (if $\chi(-q) = 1$) or $L'(f_K \otimes \chi, 1/2)$ (if $\varepsilon(-q) = -1$) correspond to heights of (f, χ) -eigencomponents of Heegner cycles living in certain Hecke modules. It is possible to generalize our techniques to obtain non-vanishing results of these central values, to bound on average the “analytic rank” of these L -functions, or to find linear independence results along the lines of [V2]. Observe that when χ is not real (in particular non-trivial) it is very unlikely that $L(f_K \otimes \chi, s)$ factors as a product of two L functions of degree 2 as $L(f_K, s)$ does.
- In fact, all these exemples belong to the more general problem of mollifying the square of $L(f \otimes g, s)$, the L functions of the Rankin-Selberg convolutions of our family of forms f against a *fixed* automorphic (not necessarily cuspidal) form g over GL_2 .
- Finally, the feasibility of mollifying the fourth power averaged over the family of automorphic forms of fixed weight suggests that there may be other families where these techniques apply, such as Dirichlet L -functions of primitive characters of a given modulus (the second moment is evaluated in [IS2, MV]) or Dirichlet L -functions of quadratic characters (the

second moment has been evaluated in [So]). Evaluating the fourth moment of the latter family should be more or less equal in difficulty to evaluating the second moment of $L(f \otimes \varepsilon_K, 1/2)$ when ε_K varies over the quadratic characters of conductor less than D and f is a fixed modular form. For this problem, some bounds are known (see [PP]) but asymptotics sufficient for mollification would be of considerable value. We hope to return to this problem, as well as the others mentioned above, in a later work.

1.3 Final comments

As a final remark, let us contemplate what was achieved here concerning the question of mollification. An essential parameter of a mollifier is its (logarithmic) length $\Delta := \log M / \log \hat{q}$. For the family $S_2(q)^*$, mollification of the second moment can be performed effectively for a very general class of mollifiers provided that Δ is strictly less than 1. As noted by Iwaniec and Sarnak, going beyond this limit for general mollifiers would have striking consequences for the study of Siegel's zeros (see [IS] for discussion of this topic). By writing $M'(f, s) = L(f, s)M(f, s)^2$, our fourth moment calculations can be re-interpreted as second moment calculations for $M'(f, s)L(f, s)$, with the mollifier length going beyond the former limit (essentially $\Delta' \sim 1 + 1/15$). Unfortunately, our results cannot be used to say anything about the zeros of $L(\chi, s)$, because our mollifier has a very peculiar shape, with most of its coefficients (those coming from the $L(f, s)$) being “smooth,” that is, avoiding the difficult oscillation of the Möbius function μ . In [IS], the authors were able to go beyond $\Delta \sim 1$ for non-smooth mollifiers either by averaging over the level q or by assuming the generalized Riemann Hypothesis. It was their success in this, along with the remarkable technology of [DFI II], that indicated that it might be possible to handle the present smooth case.

This paper is organized as follows. After some preliminary lemmas in section 2, we turn to the proof of Theorem 1.2 in section 3, where we also compute the asymptotics of the diagonal and off-diagonal terms. Section 4, which is the main technical point of the paper, is devoted to the asymptotics of the off-off-diagonal term. In section 5 we apply the theorems to compute the harmonic average of the mollified fourth power moments at the critical point (Theorem 1.4). In section 6 we show how to remove the harmonic weight from (6) to obtain (7). In section 7 we sketch the second moment calculations needed to prove Proposition 1.5. We end with an appendix, where we discuss how to remove the annoying factor $e^{2\pi|\mu|}$ from Theorem 1.2.

Acknowledgements. Some portions of our results come from formal computations which were performed with the MAPLE system. Copies of the code can be sent upon request to the authors.

It is a pleasure to express our deep gratitude to Henryk Iwaniec and Peter Sarnak for sharing their insight and experience during the preparation of this paper.

Some of these results were announced during a series of lectures given by the second author at the Collège de France (Cours Peccot), he wishes to thank this institution for its hospitality and support. The third author was supported by NSF grant DMS-9804517.

2 Auxiliary Lemmas

We list some results about modular forms and their coefficients which will be cited repeatedly in what is to come. We begin with Hecke's recursion formula for primitive forms (which is equivalent to (1), the Euler product):

Lemma 2.1 For $m, n \geq 1$ and $f \in S_2(q)^*$

$$\lambda_f(m)\lambda_f(n) = \sum_{d|(m,n)} \varepsilon_q(d)\lambda_f\left(\frac{mn}{d^2}\right). \quad (9)$$

The next lemma is a particular case of Petersson's trace formula, for weight two and for q prime (so that $S_2(q)^*$ is an orthogonal basis of $S_2(q)$).

Lemma 2.2 For $m, n \geq 1$ one has

$$\sum_{f \in S_2(q)^*}^h \lambda_f(m)\lambda_f(n) = \delta_{m,n} - 2\pi \sum_{c \geq 1} \frac{S(m, n; cq)}{cq} J_1\left(\frac{4\pi\sqrt{mn}}{cq}\right) \quad (10)$$

where $\delta_{m,n}$ is the Kronecker symbol,

$$S(m, n; c) = \sum_{\substack{a \pmod c \\ (a,c)=1}} e\left(\frac{ma + n\bar{a}}{c}\right)$$

is the classical Kloosterman sum, and $J_1(x)$ is the Bessel function of order 1. Moreover we have the asymptotic formula

$$\sum_{f \in S_2(q)^*}^h \lambda_f(m)\lambda_f(n) = \delta_{m,n} + O((m, n, q)^{1/2}(mn)^{1/2}q^{-3/2}). \quad (11)$$

The inequality follows easily from Weil's bound on Kloosterman sums and the bound $J_1(x) \ll x$. We will also make use of Jutila's extension ([J], Theorem 1.7) of the Voronoi summation formula.

Proposition 2.3 (Jutila). Let $t : \mathbf{R}^+ \rightarrow \mathbf{C}$ be a C^∞ function which vanishes in the neighborhood of 0 and is rapidly decreasing at infinity. Then for $c \geq 1$ and $d\bar{d} \equiv 1 \pmod{c}$,

$$\begin{aligned} c \sum_{m \geq 1} \tau(m) e\left(\frac{dm}{c}\right) t(m) &= 2 \int_0^{+\infty} \left(\log \frac{\sqrt{x}}{c} + \gamma\right) t(x) dx \\ &\quad - 2\pi \sum_{h \geq 1} \tau(h) e\left(-\frac{\bar{d}h}{c}\right) \int_0^{+\infty} Y_0\left(\frac{4\pi\sqrt{hx}}{c}\right) t(x) dx \\ &\quad + 4 \sum_{h \geq 1} \tau(h) e\left(\frac{\bar{d}h}{c}\right) \int_0^{+\infty} K_0\left(\frac{4\pi\sqrt{hx}}{c}\right) t(x) dx. \end{aligned}$$

Through the same methods as [J], one can also show a similar formula for sums of $\eta_\mu(m)$. While this is not needed in the main body of this paper, it will play an important role in removing the exponential dependence on μ in the appendix.

3 The twisted fourth moment

To prove Theorem 1.2 we shall refine the method of [DFI II]. We first need to express the square $L(f, s)^2$ for s on the critical line as an (almost) finite sum. First note the following, which is implied by (9):

$$L(f, s)^2 = \zeta_q(2s) \sum_{n \geq 1} \tau(n) \frac{\lambda_f(n)}{n^s}.$$

Let $\mu \in i\mathbf{R}$. The functional equation (2) and a contour shift imply that

$$G(\mu)\Gamma^2(1 + \mu)\hat{q}^{2\mu}L(f, 1/2 + \mu)^2 = \sum_n \tau(n)\lambda_f(n)n^{-1/2}W(n/\hat{q}^2) \quad (12)$$

with

$$\begin{aligned} W(y) &= W_\mu(y) + W_{-\mu}(y) \\ W_\mu(y) &= \frac{1}{2\pi i} \int_{(3)} G(s)\Gamma^2(1 + s)\zeta_q(1 + 2s)y^{-s} \frac{ds}{s - \mu} \end{aligned}$$

where $G(s)$ is an even polynomial with real coefficients (depending on $|\mu|$) such that

$$\begin{aligned} G(s + \mu)\Gamma^2(1 + s + \mu) &\text{ is holomorphic for } \Re s \geq -A \\ G(\mu) = G(-\mu) &= |\Gamma(1 + \mu)|^{-2} \end{aligned} \quad (13)$$

for some large fixed constant A . To simplify exposition, we also assume that $G(s)$ admits a zero at $s = 1$ of order A , and a zero at $s = 1 \pm \mu$ of order two. A shift of the line of integration to $\Re s = C$, together with Stirling's formula and (13), shows that there exists a large $B > 0$ for which³

$$W(y) \ll_C e^{\pi|\mu|}(1 + |\mu|)^B y^{-C}. \quad (14)$$

In the appendix we discuss how to remove the exponential factor from (14).

We now look to reexpress $M(\ell)$ so that we can use Lemma 2.2. Since we will be taking ℓ less than a small power of q , we may assume that $(\ell, q) = 1$. We start the proof with formula (12), which gives, using (13),

$$M(\ell) = \sum_f^h \sum_m \frac{\tau(m)\lambda_f(m)}{m^{1/2}} W\left(\frac{m}{\hat{q}^2}\right) \sum_n \frac{\tau(n)\lambda_f(\ell)\lambda_f(n)}{n^{1/2}} W\left(\frac{n}{\hat{q}^2}\right).$$

Applying (9), this is

$$M(\ell) = \sum_{de=\ell} \frac{1}{d^{1/2}} \sum_{m,n} \frac{\tau(m)\tau(dn)}{(mn)^{1/2}} W\left(\frac{m}{\hat{q}^2}\right) W\left(\frac{dn}{\hat{q}^2}\right) \sum_f^h \lambda_f(m)\lambda_f(ne).$$

The recursion formula

$$\tau(dn) = \sum_{a|(d,n)} \mu(a)\tau(d/a)\tau(n/a) \quad (15)$$

³In the rest of the paper the value of B may change from line to line.

allows us to write

$$M(\ell) = \sum_{de=\ell} \frac{1}{d^{1/2}} \sum_{ab=d} \frac{\mu(a)}{a^{1/2}} \tau(b) \sum_{m,n} \frac{\tau(m)\tau(n)}{(mn)^{1/2}} W\left(\frac{m}{\hat{q}^2}\right) W\left(\frac{adn}{\hat{q}^2}\right) \sum_f^h \lambda_f(m) \lambda_f(aen).$$

We apply Petersson's formula (10) to obtain $M(\ell) := M^D(\ell) + M^{ND}(\ell)$ with

$$M^D(\ell) = \sum_{de=\ell} \frac{1}{d^{1/2}} \sum_{ab=d} \frac{\mu(a)}{a^{1/2}} \tau(b) \sum_n \frac{\tau(aen)\tau(n)}{n(ae)^{1/2}} W\left(\frac{aen}{\hat{q}^2}\right) W\left(\frac{adn}{\hat{q}^2}\right) \quad (16)$$

and

$$M^{ND}(\ell) = -2\pi \sum_{de=\ell} \frac{1}{d^{1/2}} \sum_{ab=d} \frac{\mu(a)}{\sqrt{a}} \tau(b) \quad (17)$$

$$\sum_{q|c} \frac{1}{c^2} \left[c \sum_{m,n} \frac{\tau(m)\tau(n)}{\sqrt{mn}} W\left(\frac{m}{\hat{q}^2}\right) W\left(\frac{adn}{\hat{q}^2}\right) S(m, aen; c) J_1\left(\frac{4\pi\sqrt{aemn}}{c}\right) \right].$$

This last term will be evaluated using the techniques of [DFI II], from which we borrow some notation. In what follows let $T(c)$ denote the expression contained in brackets, so that

$$M^{ND}(\ell) := -2\pi \sum_{de=\ell} \frac{1}{d^{1/2}} \sum_{ab=d} \frac{\mu(a)}{a^{1/2}} \tau(b) \sum_{q|c} \frac{1}{c^2} T(c).$$

3.1 Evaluation of the diagonal term

By inverting (15), we find that the main diagonal term equals

$$M^D(\ell) = \frac{1}{\ell^{1/2}} \sum_{de=\ell} \sum_n \frac{\tau(dn)\tau(en)}{n} W\left(\frac{dn}{\hat{q}^2}\right) W\left(\frac{en}{\hat{q}^2}\right). \quad (18)$$

By (14) and $\tau(n) \ll n^\epsilon$, this is no larger than $e^{2\pi|\mu|}(1+|\mu|)^B q^\epsilon \ell^{-1/2}$, as expected. We hold off on evaluating the asymptotics of this term until later, as it will turn out that it combines with the first off-diagonal term in a convenient manner.

3.2 Preliminary evaluation of the non-diagonal terms

We adjust $M^{ND}(\ell)$ so that we can use the methods of [DFI II] (most of the methods in the next three sections can be found there, up to notational adjustments). First we make a smooth partition of unity and write

$$\frac{1}{(mn)^{1/2}} W\left(\frac{m}{\hat{q}^2}\right) W\left(\frac{dan}{\hat{q}^2}\right) = \sum_{M,N} F_{M,N}(m, n)$$

the sum being performed over some indices $M, N \geq 1$ and

$$F_{M,N}(m, n) := \frac{1}{(mn)^{1/2}} W\left(\frac{m}{\hat{q}^2}\right) W\left(\frac{dan}{\hat{q}^2}\right) F_M(m) F_N(n)$$

where $F_M(m)$ denotes a smooth function compactly supported in $[M/2, 3M]$ such that for all $j \geq 0$

$$x^j F_M^{(j)}(x) \ll_j 1.$$

Moreover, we assume that $\sum_{M \leq X} 1 \ll \log X$. Accordingly we split

$$T(c) = \sum_{M,N} T_{M,N}(c)$$

with

$$T_{M,N}(c) = c \sum_{m,n} \tau(m)\tau(n)S(m, aen; c)F_{M,N}(m, n)J_1\left(\frac{4\pi\sqrt{aemn}}{c}\right) \quad (19)$$

and define also

$$T_{M,N} = \sum_{q|c} \frac{1}{c^2} T_{M,N}(c).$$

Note that the derivatives of $F_{M,N}$ satisfy the bounds

$$x^i y^j \frac{\partial^i}{\partial^i x} \frac{\partial^j}{\partial^j y} F_{M,N}(x, y) \ll e^{2\pi|\mu|} (1 + |\mu|)^B (MN)^{-1/2} (\log q)^{i+j} \left(\frac{\hat{q}^2}{x}\right)^A \left(\frac{\hat{q}^2}{day}\right)^{A'} \quad (20)$$

for all $i, j, A, A' \geq 0$. By taking $i = j = 0$ and either A or A' large in (20), we have

$$\sum_{M+N \gg q^{1+\varepsilon}} T_{M,N} \ll_{\varepsilon, A} e^{2\pi|\mu|} (1 + |\mu|)^B q^{-A},$$

for any $\varepsilon > 0$ and any $A > 0$. Thus we may assume that

$$M \leq q^{1+\varepsilon}, \quad N \leq q^{1+\varepsilon}. \quad (21)$$

It will also prove convenient to remove large values of c through the large sieve inequality, as given in [DFI II], Proposition 1, and section 12. Using these methods, the contribution of $T_{M,N}(c)$ for $c \geq C$ and $M, N \ll q^{1+\varepsilon}$ is

$$\sum_{\substack{q|c \\ c \geq C}} \frac{1}{c^2} T_{M,N}(c) \ll_{\varepsilon} e^{2\pi|\mu|} (1 + |\mu|)^B q^{\varepsilon} \frac{\ell^{3/4} (MN)^{1/4}}{C^{1/2}}. \quad (22)$$

As in [DFI II], we take $C = \min(q^{2/3} M^{1/2}, q^{7/6})$ (note that the entire sum on c is dropped when $M < q^{2/3}$). The error taken on by doing this is at most

$$e^{2\pi|\mu|} (1 + |\mu|)^B \ell^{3/4} N^{1/4} q^{-1/3+\varepsilon} \ll e^{2\pi|\mu|} (1 + |\mu|)^B \ell^{3/4} q^{-1/12+\varepsilon}. \quad (23)$$

Accordingly, in the next two sections we attach a smooth compact function $\eta_C(c)$ vanishing for $c \geq 2C$ and equal to 1 for $c \leq C$.

3.3 Applying the summation formula

Now we apply Jutila's summation formula (2.3) on the m variable with the effect of splitting

$$T_{M,N}(c) = T_{M,N}^*(c) + T_{M,N}^+(c) + T_{M,N}^-(c) \quad (24)$$

according to [DFI II] p 228.

The "trivial" bound (43) of [DFI II] suffices to bound $T_{M,N}^*(c)$ for our needs (by our different normalization of $F_{M,N}$ we have an extra factor of $(MN)^{-1/2}$ compared to [DFI II]) by

$$\sum_{q|c} c^{-2} T_{M,N}^*(c) \ll_{\varepsilon} e^{2\pi|\mu|} (1 + |\mu|)^B q^{\varepsilon} q^{-2} (MN)^{1/2} \ll_{\varepsilon} e^{2\pi|\mu|} (1 + |\mu|)^B q^{-1+\varepsilon}.$$

The other two terms of (24) require considerably more effort and contribute main terms in two different ways.

3.4 Evaluation of the off-diagonal terms

We continue to apply the arguments of [DFI II]. It turns out that there is an error in Section 9, p. 229: a factor of $s = ae$ coming from a change of variable (namely, replacing x with $h + sy$) has been omitted. Correcting this, we obtain from [DFI II], equations (44), (46), and (47) that

$$\begin{aligned} T_{M,N}^-(c) + T_{M,N}^+(c) &= \phi(c) \sum_n \tau(aen) \tau(n) G^-(aen, n) \\ &\quad + \sum_{1 \leq w} \frac{(ae, w)}{w^2} \sum_{h \neq 0} S(0, h; c) S(0, h; w) \left[V^-(h) + V^+(h) \right] \\ &\quad + O_{\varepsilon}(e^{2\pi|\mu|} (1 + |\mu|)^B q^{\varepsilon} P^{11/4} \ell^{3/4} \frac{N^{1/4}}{M^{1/2}} c^2) \\ &:= T^{OD}(c) + T^{OOD}(c) + Err(c), \end{aligned} \quad (25)$$

(remark that, unlike [DFI II], we have not introduced the summation condition $w < c$) with

$$G^-(z, y) = -2\pi \int_0^{\infty} Y_0\left(\frac{4\pi\sqrt{xz}}{c}\right) J_1\left(\frac{4\pi\sqrt{x(aey)}}{c}\right) F_{M,N}(x, y) dx,$$

$$\begin{aligned} V^{\pm}(h) &= \int_0^{\infty} \int_0^{\infty} \delta_{h \mp aey > 0} [\log(h \mp aey) - \lambda_{1,w}] [\log(aey) - \lambda_{ae,w}] \\ &\quad \times K^{\pm}\left(\frac{4\pi\sqrt{x(h \mp aey)}}{c}\right) J_1\left(\frac{4\pi\sqrt{x(aey)}}{c}\right) F_{M,N}(x, y) dx dy, \end{aligned} \quad (26)$$

with

$$\begin{aligned} K^-(y) &= -2\pi J_0(y), \quad K^+(y) = 4K_0(y), \\ \lambda_{ae,w} &= 2\gamma + \log \frac{ae w^2}{(ae, w)^2}, \quad P = 1 + \frac{(aeMN)^{1/2}}{c}. \end{aligned}$$

The contribution of $Err(c)$ over $c \ll C$ is bounded by

$$\begin{aligned} \sum_{q|c} c^{-2} Err(c) &\ll_{\varepsilon} e^{2\pi|\mu|} (1+|\mu|)^B q^{\varepsilon} \left(\ell^{3/4} \frac{N^{1/4} C}{M^{1/2} q} + \ell^{17/8} \frac{M^{7/8} N^{13/8}}{q^{11/4}} \right) \\ &\ll_{\varepsilon} e^{2\pi|\mu|} (1+|\mu|)^B q^{\varepsilon} \left(\ell^{3/4} q^{-1/12} + \ell^{17/8} q^{-1/4} \right). \end{aligned} \quad (27)$$

It is the $\ell^{3/4} q^{-1/12}$ term which accounts for our choice of C and leads to the most important error term in Theorem 1.2 (and to the value of $1/192$ in Theorem 1.1). Improvements in the admissible size of ℓ will have to come primarily from improvements in this bound.

3.5 Expanding the c sum

We thus wish to evaluate

$$\sum_{q|c} \eta_C(c) c^{-2} \left[T^{OD}(c) + T^{OOD}(c) \right].$$

We can now extend to the wider range $c \ll C_{max} = q^A$ for some large A (while we no longer need the strict $c \ll C$ bound, there are still some advantages to keeping the sum finite). The off-off-diagonal term can be bounded using the arguments in [DFI II], sections 9 through 12 (remembering the extra factor of ae , and that our F differs from theirs by a factor of $(MN)^{-1/2}$):

$$\begin{aligned} \sum_{q|c} (1 - \eta_C(c)) c^{-2} T^{OOD}(c) &\ll_{\varepsilon} e^{2\pi|\mu|} (1+|\mu|)^B q^{\varepsilon} \ell^{1/2} MN \sum_{q|c} (1 - \eta_C(c)) c^{-2} \\ &\ll_{\varepsilon} e^{2\pi|\mu|} (1+|\mu|)^B q^{\varepsilon} \ell^{1/2} \frac{MN}{qC} \\ &\ll_{\varepsilon} e^{2\pi|\mu|} (1+|\mu|)^B q^{\varepsilon} \ell^{1/2} q^{-1/6} \end{aligned}$$

For the off-diagonal term, we start with equation (54) of [DFI II], which gives

$$\sum_{c \equiv 0(q)} (1 - \eta_C(c)) c^{-2} T^{OD}(c) = -2\pi \sum_n \tau(aen) \tau(n) \int_0^{\infty} Y_0(4\pi\sqrt{aent}) J_1(4\pi\sqrt{aent}) S_n(t) dt$$

with the same definition of $S_n(t)$ except for a factor of $1 - \eta_C(c)$ in the c sum (in particular, this means that $S_n(t)$ is supported in $[0, 2M/C^2]$). From this we obtain

$$\begin{aligned} \sum_{c \equiv 0(q)} (1 - \eta_C(c)) c^{-2} T^{OD}(c) &\ll_{\varepsilon} e^{2\pi|\mu|} (1+|\mu|)^B q^{\varepsilon} \frac{N^{1/2}}{M^{1/2}} \int_0^{2M/C^2} |\log t| \frac{M^{1/2}}{t^{1/2+\varepsilon}} (aeNt)^{-1/4} dt \\ &\ll_{\varepsilon} e^{2\pi|\mu|} (1+|\mu|)^B q^{\varepsilon} (ae)^{-1/4} q^{-1/12}. \end{aligned}$$

Thus we may extend the c sums at acceptable cost, so that the non-diagonal terms come from

$$\begin{aligned} \sum_{q|c} \frac{1}{c^2} T(c) &= \sum_{c \equiv 0(q)} \frac{\phi(c)}{c^2} \sum_n \tau(sn) \tau(n) G^-(aen, n) \\ &\quad + \sum_{q|c} \frac{1}{c^2} \sum_{w \geq 1} \frac{(ae, w)}{w^2} \sum_{h \neq 0} S(0, h; c) S(0, h; w) \left[V^-(h) + V^+(h) \right] \\ &\quad + O_{\varepsilon}(q^{\varepsilon} e^{2\pi|\mu|} (1+|\mu|)^B (\ell^{3/4} q^{-1/2} + \ell^{1/2} q^{-1/6})). \end{aligned} \quad (28)$$

Moreover, using (14), we may reintroduce the contribution from $M + N \gg q^{1+\varepsilon}$ at the cost of an error term which is $\ll_{A,\varepsilon} e^{2\pi|\mu|}(1+|\mu|)^B q^{-A}$ for all A . Thus we may combine all of the F_M 's into one function F , replacing the $F_{M,N}$ in the formulas for $G^-(z, y)$ and $V^\pm(h)$ with

$$F(x, y) := \frac{1}{\sqrt{xy}} W\left(\frac{x}{\hat{q}^2}\right) W\left(\frac{ady}{\hat{q}^2}\right) F(x) F(y), \quad (29)$$

where $F(x)$ is a smooth function, supported in $[1/2, \infty)$ with $F(x) = 1$ for $x \geq 1$. For the off-diagonal term (the first line of (28)) we may also replace the product $F(x)F(y)$ by 1 at the cost of an error term which is $O_\varepsilon(e^{2\pi|\mu|}(1+|\mu|)^B q^{-1+\varepsilon} \ell^{1/2})$. We will eventually do likewise for the off-off-diagonal term, but this requires more difficult arguments and will not be dealt with until section 4.1.

3.6 Evaluation of the off-diagonal term

Putting the first line of (28) into (17) and inverting (15), we have

$$M^{OD}(\ell) = -2\pi \sum_{de=\ell} \frac{1}{d^{1/2}} \sum_{ab=d} \frac{\mu(a)}{a^{1/2}} \tau(b) M^{OD}(\ell, ae) = \frac{-2\pi}{\ell^{1/2}} \sum_{de=\ell} \sum_n \frac{\tau(dn)\tau(en)}{n} W\left(\frac{dn}{\hat{q}^2}\right) Y(en) \quad (30)$$

with

$$Y(n) := - \int_0^\infty Y_0(y) J_1(y) \left[\sum_{c \equiv 0(q)} \frac{\phi(c)}{c} W\left(\frac{c^2 y^2}{(4\pi)^2 \hat{q}^2 n}\right) \right] dy.$$

The Mellin transform for W gives

$$Y(n) = - \int_0^\infty Y_0(y) J_1(y) \frac{1}{2\pi i} \int_{(3)} G(s) \Gamma^2(1+s) \zeta_q(1+2s) \left(\frac{(4\pi)^2 \hat{q}^2 n}{y^2} \right)^s \sum_{q|c} \frac{\phi(c)}{c^{1+2s}} \frac{2s ds}{s^2 - \mu^2} dy.$$

We can extend the c sum to include $c > q^A$, at negligible cost for $\Re s = 3$, making it

$$q^{-1-2s} \phi(q) \frac{\zeta(2s)}{\zeta_q(1+2s)}.$$

We can then switch the order of integration and use the identity

$$\int_0^\infty Y_0(y) J_1(y) y^{-2s} dy = \frac{1}{2\sqrt{\pi}} \frac{\Gamma(s)}{\Gamma(1/2-s)} \frac{\Gamma(1-s)^2}{\Gamma(1+s)^2}. \quad (31)$$

The orthogonality property for the Bessel functions used in [DFI II], section 12, is encapsulated in (31), since this expression vanishes at $s = 1/2$. We thus have

$$Y(n) = - \frac{\phi(q)}{2\pi q} \frac{1}{2\pi i} \int_{(3)} G(s) \Gamma^2(1-s) \frac{\Gamma(s) \zeta(2s)}{\pi^{2s-1/2} \Gamma(1/2-s)} \left(\frac{n}{\hat{q}^2}\right)^s \frac{2s ds}{s^2 - \mu^2}$$

The functional equation

$$\zeta(2s) \Gamma(s) = \pi^{2s-1/2} \zeta(1-2s) \Gamma(1/2-s)$$

lets us convert this to

$$Y(n) = -\frac{\phi(q)}{2\pi q} \frac{1}{2\pi i} \int_{(3)} G(s) \Gamma^2(1-s) \zeta(1-2s) \left(\frac{n}{\hat{q}^2}\right)^s \frac{2s ds}{s^2 - \mu^2}.$$

We can now shift the contour to $\Re s = -1/2$, meeting poles at $s = \pm\mu$ (the factor of $2s$ cancels the pole of $\zeta(1-2s)$ at $s = 0$ unless $\mu = 0$) contributing residues of

$$G(\mu) \Gamma^2(1+\mu) \zeta(1+2\mu) \left(\frac{n}{\hat{q}^2}\right)^{-\mu} + G(-\mu) \Gamma^2(1-\mu) \zeta(1-2\mu) \left(\frac{n}{\hat{q}^2}\right)^{\mu}.$$

In the resulting contour integral, changing s into $-s$ gives an expression which equals $-W(\frac{n}{\hat{q}^2})$, up to a factor of $1 + O(q^{-1})$. Thus we have

$$Y(n) = -\frac{G(\mu)}{2\pi} \left(-W\left(\frac{n}{\hat{q}^2}\right) + \Gamma^2(1+\mu) \zeta_q(1+2\mu) \left(\frac{n}{\hat{q}^2}\right)^{-\mu} + \Gamma^2(1-\mu) \zeta_q(1-2\mu) \left(\frac{n}{\hat{q}^2}\right)^{\mu} \right) + O_\varepsilon(e^{2\pi|\mu|} (1+|\mu|)^B q^{-1+\varepsilon}).$$

Putting this back in (30), we get

$$M^{OD}(\ell) = -M^D(\ell) + O_\varepsilon(e^{2\pi|\mu|} (1+|\mu|)^B q^{-1+\varepsilon}) + \frac{G(\mu)}{\ell^{1/2}} \sum_{de=\ell} \sum_n \frac{\tau(dn)\tau(en)}{n} W\left(\frac{dn}{\hat{q}^2}\right) \left[\Gamma^2(1+\mu) \zeta(1+2\mu) \left(\frac{en}{\hat{q}^2}\right)^{-\mu} + \Gamma^2(1-\mu) \zeta(1-2\mu) \left(\frac{en}{\hat{q}^2}\right)^{\mu} \right]. \quad (32)$$

For $\mu = 0$ we take the limit of this expression, getting

$$M^{OD}(\ell) = -M^D(\ell) + \frac{1}{\ell^{1/2}} \sum_{de=\ell} \sum_n \frac{\tau(dn)\tau(en)}{n} W\left(\frac{dn}{\hat{q}^2}\right) \log\left(\frac{\hat{q}^2}{en}\right) + O_\varepsilon(q^{-1+\varepsilon}). \quad (33)$$

For future reference we write this in the following form

$$M^D(\ell) + M^{OD}(\ell) = O_\varepsilon(q^{-1+\varepsilon}) + Res_{t=0} \frac{1}{\ell^{1/2}} \sum_{de=\ell} \sum_n \frac{\tau(dn)\tau(en)}{n} W\left(\frac{dn}{\hat{q}^2}\right) \frac{2G(t)\Gamma^2(1+t)\zeta(1+2t)}{t} \left(\frac{\hat{q}^2}{en}\right)^t \quad (34)$$

From (32) or (33), along with the usual bounds for W and τ , we have

$$M^D(\ell) + M^{OD}(\ell) \ll_\varepsilon q^\varepsilon e^{2\pi|\mu|} (1+|\mu|)^B \ell^{-1/2}.$$

We can use (33) to evaluate the contribution of the diagonal and off-diagonal terms to $M(1)$ at $\mu = 0$. Setting $\ell = 1$, we get a main term of

$$\sum_n \frac{\tau(n)^2}{n} W\left(\frac{n}{\hat{q}^2}\right) \log \frac{\hat{q}^2}{n}.$$

Using the integral formula for W , this is

$$\frac{1}{2\pi i} \int_{(3)} G(s) \Gamma(1+s)^2 \zeta_q(1+2s) \partial_s \left[\hat{q}^{2s} \sum_n \frac{\tau(n)^2}{n^{1+s}} \right] \frac{2ds}{s}$$

$$= \frac{1}{2\pi i} \int_{(3)} G(s)\Gamma(1+s)^2 \zeta_q(1+2s) \hat{q}^{2s} \frac{\zeta(1+s)^4}{\zeta(2+2s)} [\log \hat{q}^2 + 4\frac{\zeta'}{\zeta}(1+s) - 2\frac{\zeta'}{\zeta}(2+2s)] \frac{2ds}{s}.$$

A contour shift to $\Re s = -1/2$ leaves a contour integral whose value is bounded by a constant times bounded by $q^{-1/2}$, plus the contribution of multiple poles at $s = 0$, which clearly gives a polynomial in $\log q$. The lead term comes from replacing all $\zeta(1+s)$ factors by $1/s$, and winds up as

$$\frac{1}{2\zeta(2)} \text{Res}_{s=0} \frac{q^s}{s^6} (\log q - \frac{4}{s}) = \frac{(\log q)^6}{360\zeta(2)} = \frac{(\log q)^6}{60\pi^2}.$$

The off-off-diagonal term will also contribute a main term to the fourth moment, but, as we shall see, $M^{OOD}(1)$ has lead term $(\log q)^2$, and thus does not affect the lead term of this polynomial.

4 The off-off-diagonal main term

We thus return to the analysis of the off-off-diagonal term, which from the previous section is

$$M^{OOD}(\ell) = -2\pi \sum_{de=\ell} \frac{1}{d^{1/2}} \sum_{ab=e} \frac{\mu(a)}{a^{1/2}} \tau(b) \tag{35}$$

$$\sum_{q|c < q^A} \frac{1}{c^2} \sum_w \frac{(ae, w)}{aew^2} \sum_{h \neq 0} S(0, h; c) S(0, h; w) [V^-(h) + V^+(h)],$$

with V^\pm and F defined as in (29). In this section, which is the most delicate part of this paper, we find the asymptotic value of this expression. As the proof is rather lengthy, we first provide an outline. What we will do is:

- convert the Ramanujan sums to divisibility conditions, splitting into two cases;
- bring in the Mellin transforms of J_1 , Y_0 , and K_0 to evaluate V^+ and V^- ;
- include $1 - F(y)$ at negligible cost, and perform the y integral;
- include $1 - F(x)$ at negligible cost, and perform the x integral;
- evaluate the J_1 integral through a contour shift;
- show that the remaining integrand is an even function in two variables, and thus equals its residue at the origin;
- and bound the second case through more elementary methods.

We change the Ramanujan sums into divisibility conditions using

$$\sum_c \sum_{(a,c)=1} f(a, c) = \sum_g \mu(g) \sum_{a,c} f(gc, ga),$$

and

$$\sum_{a=0}^{c-1} e(ah/c) = c\delta_{c|h}.$$

The same applies to $S(0, h; w)$, so

$$M^{OOD}(\ell) = -2\pi \sum_{de=\ell} \frac{1}{d^{1/2}} \sum_{ab=e} \frac{\mu(a)}{a^{1/2}} \tau(b) \sum_{g,v} \frac{\mu(g)}{g^2} \frac{\mu(v)}{v^2} \sum_{\substack{q|cg \\ cg < q^A}} \frac{1}{c} \sum_w \frac{(ae, wv)}{w} \sum_{\substack{[w,c]|h \\ h \neq 0}} [V^-(h) + V^+(h)].$$

There are two cases to consider, coming from $q|g$ and $(q, g) = 1$ (once again, the primality of q simplifies our argument considerably). The former will only contribute a remainder term, which we consider in section 4.4. It is the latter which contributes the main term to the expression. For the next three sections, we will assume that $(q, g) = 1$, so $q|c$ and $q|h$.

4.1 Evaluating V^- and V^+

We now look to evaluate the $V^-(h) + V^+(h)$ portion of the off-off-diagonal term. As a first step, we replace negative values of h (which only contribute to V^-) by their absolute value, and replacing both V 's by their integral formulations. We rewrite for positive h , after the change of variable $ae y/h \rightarrow y$:

$$\begin{aligned} V^-(h) + V^+(h) &= \frac{h^{1/2}}{(ae)^{1/2}} \int_0^\infty \int_0^\infty (xy)^{-1/2} J_1\left(\frac{4\pi\sqrt{hxy}}{cg}\right) W\left(\frac{x}{\hat{q}^2}\right) W\left(\frac{dhy}{e\hat{q}^2}\right) F(x) F\left(\frac{hy}{ae}\right) \\ &\quad \times \left[-2\pi Y_0\left(\frac{4\pi\sqrt{hx(1+y)}}{cg}\right) L^{++} - 2\pi \delta_{y>1} Y_0\left(\frac{4\pi\sqrt{hx(y-1)}}{cg}\right) L^{-+} \right. \\ &\quad \left. + 4\delta_{y<1} K_0\left(\frac{4\pi\sqrt{hx(1-y)}}{cg}\right) L^{+-} \right] dx dy, \end{aligned} \quad (36)$$

with

$$L^{\pm\pm} = (\log(h(\pm 1 \pm y)) - \lambda_{1, wv})(\log(hy) - \lambda_{ae, wv}).$$

To simplify the integration, we bring in the Mellin transforms of J_1 , Y_0 , and K_0 (note in particular the similarity between the last two).

$$\begin{aligned} J_1(x) &= -\frac{1}{2\pi i} \int_{(-0.1)} \frac{\pi}{\Gamma(1+z)\Gamma(2+z)\sin(\pi z)} \left(\frac{x}{2}\right)^{1+2z} dz, \\ Y_0(x) &= -\frac{1}{2\pi} \frac{1}{2\pi i} \int_{(*)} 2^{2\beta} \Gamma(\beta)^2 x^{-2\beta} \cos \pi\beta 2d\beta, \\ K_0(x) &= \frac{1}{4} \frac{1}{2\pi i} \int_{(0.7)} 2^{2\beta} \Gamma(\beta)^2 x^{-2\beta} 2d\beta. \end{aligned} \quad (37)$$

The $(*)$ contour can be taken as a path including almost all of $\Re e\beta = -1$, except when $|\Im m\beta| < 1$, where it curves to hit the real axis at $\beta > 0$. We need this extra care because $\Gamma(\beta)^2 \cos \pi\beta$ does not decay quickly enough for large values of $\Im m\beta$ unless $\Re\beta$ is negative. Fortunately, we will almost always be able to ignore this difficulty.

We insert (37) into (36). With the given values of z and β , all of the integrals are absolutely convergent, so we may move the x and y integrals inside the z and β integrals. We thus have

$$V^-(h) + V^+(h) = -\frac{2\pi}{(2\pi i)^2} \int_{(0.7)} \int_{(-0.1)} \frac{\Gamma(\beta)^2 (4\pi)^{1+2z-2\beta} 2^{-1-2z+2\beta}}{\Gamma(1+z)\Gamma(2+z)\sin \pi z} (cg)^{-1-2z+2\beta} h^{1+z-\beta}$$

$$\begin{aligned} & \times \int_0^\infty x^{z-\beta+1} W\left(\frac{x}{\hat{q}^2}\right) F(x) \frac{dx}{x} \\ & \int_0^\infty y^{z+1} W\left(\frac{dhy}{e\hat{q}^2}\right) F\left(\frac{hy}{ae}\right) \left[\frac{\cos \pi \beta}{(1+y)^\beta} L^{++} + \delta_{y>1} \frac{\cos \pi \beta}{(y-1)^\beta} L^{-+} + \delta_{y<1} \frac{L^{+-}}{(1-y)^\beta} \right] \frac{dy}{y} dz d\beta. \end{aligned} \quad (38)$$

We have shifted the (*) contour to $\Re e\beta = 0.7$, which can be justified by the rapid decay of the x integral in β .

We next look to evaluate the x and y integrals in (38). This would be much easier if we could replace the $F(x)F(y)$ by 1 on $[0, \infty)^2$, so we start by showing that we can include the contributions from small x and y at negligible cost.

4.1.1 Expanding the y integral

We first look to replace $F(hy/ae)$ by 1. Since $\Re ez > -1$, there is no convergence difficulty in taking y near zero in (38), so we need only calculate the contribution to the integral from $y < ae/h \ll \ell/q$. Since this is much less than one, we need not worry about poles from $(1 \pm y)^{-\beta}$, so we can dominate the y integral in this range by $(ae/h)^{1+\Re ez}$. We may shift the β contour to $\Re e\beta = 1 + \delta$ and the z contour to $-\delta$ without crossing any poles. The x , z , and β integrals are thus bounded by $e^{2\pi|\mu|}(1+|\mu|)^B (gc)^{1+4\delta} h^{-1-\delta+\varepsilon}$. Putting this into the v, w, c, g, h sums, the contribution from small y is dominated by

$$\begin{aligned} & e^{2\pi|\mu|}(1+|\mu|)^B \ell^{1/2} \sum_{g,v,w} \frac{1}{v^2} \frac{1}{w} g^{-1+4\delta} \sum_{q|c, gc < q^A} c^{4\delta} \sum_{[c,w]|h} h^{-1-\delta+\varepsilon} \\ & \ll e^{2\pi|\mu|}(1+|\mu|)^B \ell^{1/2} q^{-1+\varepsilon} q^{3\delta A}. \end{aligned}$$

Taking $\delta < 1/6A$ is enough to make this smaller than the error terms already identified, so we may replace $F(hy/ae)$ by 1.

This makes it feasible to evaluate the y integral. It appears that there will be problems with convergence, since we need $\Re e\beta > 1$ for convergence at infinity and $\Re e\beta < 1$ for convergence near $y = 1$. However, the double zero of $\cos \pi\beta + 1$ at $\beta = 1$ allows us to circumvent this difficulty with a contour shift.⁴ To simplify notation, we suppress this technicality and write the y integral as a single integral. Using the integral formula for $W(dhy/e\hat{q}^2)$ (with $\Re t = 0.7$ so that we have convergence at $y = 0$), the y -integral becomes

$$\begin{aligned} & \left(\frac{\hat{q}^2 e}{dh} \right)^t (-\partial_t + \log h - \lambda_{ae, wv}) \\ & \int_0^\infty y^{1+z-t} \left[\cos \pi\beta (-\partial_\beta + \log h - \lambda_{1, wv}) \left(\frac{1}{(1+y)^\beta} + \frac{\delta_{y>1}}{(y-1)^\beta} \right) + \right. \\ & \left. (-\partial_\beta + \log h - \lambda_{1, wv}) \frac{\delta_{y<1}}{(1-y)^\beta} \right] \frac{dy}{y}. \end{aligned} \quad (39)$$

⁴More precisely, we separate a small neighborhood of $y = 1$ from the integral, evaluate it at $\Re e\beta < 1$ and the rest at $\Re e\beta > 1$, then shift contours to connect the two without crossing any poles.

Each of the y integrals is a known Mellin transform, they can be found in [Ob] 1.2.19, 1.2.20, and 1.2.21. Setting $\Lambda = \log h - \lambda_{1,wv}$ to save space, the last two lines of (39) equal

$$\begin{aligned} & \cos \pi\beta(-\partial_\beta + \Lambda) \frac{\Gamma(1+z-t)\Gamma(-1-z+t+\beta)}{\Gamma(\beta)} \left(1 + \frac{\sin \pi(t-z)}{\sin \pi\beta}\right) - \\ & (-\partial_\beta + \Lambda) \frac{\Gamma(1+z-t)\Gamma(-1-z+t+\beta)}{\Gamma(\beta)} \frac{\sin \pi(-z+t+\beta)}{\sin \pi\beta}, \end{aligned}$$

where we have used the identity $\Gamma(x)\Gamma(1-x) = \pi/\sin \pi x$ to make the arguments of the Γ functions more consistent. This can be simplified further, but for the time being it is enough to note that the first and third terms have poles (and after differentiation by t , double poles) at $1+z-t=0$.

4.1.2 Expanding the x integral

We replace $F(x)$ by 1 in (38). To justify this, we must first examine the contribution coming from replacing $F(x)$ by $F_1(x) = 1 - F(x)$. Once again we do this through contour shifts, which start with $\Re\beta = 1 + \delta$, $\Re t = 0.7$, $\Re z = -0.1$. Using the value of the y integral given above, the key parts of the integrand (which we then sum over c, g, h, v , and w) look like

$$\begin{aligned} & g^{-1-2z+2\beta} c^{-2-2z+2\beta} h^{1+z-\beta-t} \hat{q}^{2t} \left(\frac{e}{d}\right)^t \times \\ & [\Gamma(1+z-t)H_1(t, z, \beta) + \Gamma'(1+z-t)H_2(t, z, \beta)] \int_0^\infty x^{z-\beta+1} W\left(\frac{x}{\hat{q}^2}\right) F(x) \frac{dx}{x}, \end{aligned}$$

with H_1, H_2 analytic with rapid decay in the directions of interest. We shift $\Re\beta$ to 0.3, so that there are no convergence issues coming from the integral of $x^{z-\beta}$ near zero. We want $\sum h^{1+z-\beta-t}$ absolutely convergent, so we to shift t to the right, say to $\Re t = 1.7$, to compensate. This will give a new contour, along with single and double poles at $t = 1 + z$.

- First we look at what comes from the $\Re t = 1.7$ portion. Since $x < 1$, the x integral is dominated by a power of $\log q$, and we can factor $q^{-1-z+\beta}$, which is negligible, from the c, h , and q terms. The h sum is absolutely convergent, as is the c sum (thanks to t being so large), and this is enough. Shifting to $\Re t = 1.5$ and $\Re(\beta - z) = 0.5$ lets us bound this term by $eq^{-1/2} \ll \ell q^{-1/2}$, no matter what upper limit we take on the c, g sum.
- As for the poles at $t = 1 + z$, we need another contour shift, this time in z and β . We shift z to the right (passing a pole at $z = 0$ from $1/\sin \pi z$, but $G(t) = G(1+z)$ has a zero to compensate), then shift β to the right, so that $z - \beta > -1$ at all stages of the shift. Stopping at $\Re z = 0.5 + 2\epsilon$ and $\Re\beta = 1 + \epsilon$, say, the x integral is again bounded by a power of $\log q$, and both the sum on $h^{-\beta}$ and $c^{-2-2z+2\beta}$ sums are absolutely convergent. The power of q involved works out to be $q^{-1-z+\beta+\epsilon}$, so this term is dominated by $O_\epsilon(q^\epsilon e^{2\pi|\mu|} (1+|\mu|)^B \ell q^{-1/2})$, again small enough to ignore.

Thus we can extend the x integral to the entire positive axis in both cases. By the inverse Mellin transform, so long as $\Re(z - \beta) > -1$,

$$\int_0^\infty x^{z-\beta} W(x/\hat{q}^2) dx = \frac{G(1+z-\beta)\Gamma(2+z-\beta)^2 \zeta_q(3+2z-2\beta) \hat{q}^{2+2z-2\beta} 2(1+z-\beta)}{(1+z-\beta)^2 - \mu^2}.$$

To simplify notation, we change variables, setting $\beta = 1 + z - s$. Putting the values of the x and y integrals into our formula for $M^{OOD}(\ell)$, we have

$$\begin{aligned}
M^{OOD}(\ell) &= \frac{2\pi}{\ell^{1/2}} \sum_{de=\ell} \sum_{ab=d} \tau(b) \frac{\mu(a)}{a} \sum_g \frac{\mu(g)}{g^2} \sum_{q|c} \frac{1}{c} \sum_v \frac{\mu(v)}{v^2} \sum_w \frac{(ae, wv)}{w} \sum_{[c,w]|h} \\
&\frac{1}{(2\pi i)^3} \int_{(1.7)} \int_{(0.6)} \int_{(-0.1)} \frac{G(t)\Gamma(1+t)^2 \zeta_q(1+2t) G(s)\Gamma(1+s)^2 \zeta_q(1+2s) \hat{q}^{2s+2t}}{(t^2 - \mu^2)(s^2 - \mu^2)} \\
&(2\pi)^{-1+2s} (cg)^{1-2s} h^{s-t} \left(\frac{e}{d}\right)^t \frac{\pi\Gamma(1+z-s)^2}{\Gamma(1+z)\Gamma(2+z)\sin\pi z} (-\partial_t + \log h - \lambda_{ae, wv}) \\
&\left[-\cos(\pi(z-s))(\partial_s + \Lambda) \frac{\Gamma(1+z-t)\Gamma(t-s)}{\Gamma(1+z-s)} \left(1 + \frac{\sin(\pi(t-z))}{\sin(\pi(s-z))}\right) \right. \\
&\left. + (\partial_s + \Lambda) \frac{\Gamma(1+z-t)\Gamma(t-s)}{\Gamma(1+z-s)} \frac{\sin(\pi(t-s))}{\sin(\pi(s-z))} \right] dz 4s ds 2t dt, \tag{40}
\end{aligned}$$

plus the poles arising from $t = 1 + z$, coming from the shift of t to the right⁵, and some negligible error terms. We now examine the z integral, we return to the full expression later.

4.2 Shifting the z contour

For the time being, we ignore everything in (40) which does not depend on z . Along vertical strips, the integrand decays as $|z|^{-1-s-t}$, and similarly along horizontal strips, so as long as the real parts of s and t are positive we can safely shift in either direction. We choose to shift to $\Re z = +\infty$ because the residue calculations are simpler. We pick up poles from three sources: $1/\sin\pi z$, $1/\sin\pi(s-z)$, and $\Gamma(1+z-t)$. This last source gives residues which exactly cancel those from $t = 1+z$ when t was shifted to the right in the last section, since the signs are opposite (this is not a surprise, the shift of t was an artificial step taken to ensure convergence of various sums). Thus we can restrict our attention to the first two, which give poles at $z = k$ and $z = k+s$ for $k \geq 0$. We first write the entire z -integral, without the s and t portions, for ease of reference.

$$\begin{aligned}
&\frac{\pi}{2\pi i} \int_{(-0.1)} \frac{\Gamma(1+z-s)^2 \Gamma(1+z-t)}{\Gamma(2+z)\Gamma(1+z)\sin\pi z} \\
&\left[-\cos(\pi(z-s))(\partial_s + \Lambda) \frac{\Gamma(t-s)}{\Gamma(1+z-s)} \left(1 + \frac{\sin\pi(t-z)}{\sin\pi(s-z)}\right) \right. \\
&\left. + (\partial_s + \Lambda) \frac{\Gamma(t-s)}{\Gamma(1+z-s)} \frac{\sin\pi(t-s)}{\sin\pi(s-z)} \right]. \tag{41}
\end{aligned}$$

⁵We do not compute these residues, since they will soon be cancelled by another shift.

4.2.1 Poles at $z = k$

The poles from $1/\sin(\pi z)$ are all simple, with a residue of $(-1)^k \frac{1}{\pi}$ times the result of plugging in $z = k$ to the integrand of (41). Rotation by πk multiplies sine and cosine of an angle by $(-1)^k$, so (by checking how often z appears in an angle in (41)), all of the factors of $(-1)^k$ cancel, so the residue at $z = k$ is (the $-$ sign arises because we are shifting the contour to the right):

$$-\sum_{k=0}^{\infty} \frac{\Gamma(1+k-s)^2}{\Gamma(2+k)\Gamma(1+k)} \left[-\cos \pi s (\partial_s + \Lambda) \frac{\Gamma(1+k-t)\Gamma(t-s)}{\Gamma(1+k-s)} \left(1 + \frac{\sin \pi t}{\sin \pi s}\right) + (\partial_s + \Lambda) \frac{\Gamma(1+k-t)\Gamma(t-s)}{\Gamma(1+k-s)} \frac{\sin(\pi(t-s))}{\sin(\pi s)} \right]. \quad (42)$$

The “ $+\Lambda$ ” terms give

$$\cos \pi s + \cot \pi s \sin \pi t - (\sin \pi t \cos \pi s - \sin \pi s \cos \pi t) / \sin \pi s = \cos \pi s + \cos \pi t$$

times

$$\Gamma(t-s) \sum_{k=0}^{\infty} \frac{\Gamma(1+k-s)\Gamma(1+k-t)}{\Gamma(2+k)\Gamma(1+k)} = \Gamma(t-s)\Gamma(t+s) \frac{\Gamma(1-s)\Gamma(1-t)}{\Gamma(1+s)\Gamma(1+t)},$$

where the last step is Gauss’s hypergeometric identity. The ∂_s terms are similar. Applying ∂_s to the trigonometric portions yields

$$\pi \sin \pi t \Gamma(t-s)\Gamma(t+s) \frac{\Gamma(1-s)\Gamma(1-t)}{\Gamma(1+s)\Gamma(1+t)},$$

while applying it to the Γ functions gives

$$(\cos \pi s + \cos \pi t) \sum_{k=0}^{\infty} \frac{\Gamma(1+k-s)\Gamma(1+k-t)}{\Gamma(2+k)\Gamma(1+k)} \left(-\frac{\Gamma'}{\Gamma}(t-s) + \frac{\Gamma'}{\Gamma}(1+k-s) \right).$$

The sum can be evaluated again through the hypergeometric identity, differentiated in either t or s , making the total contribution from these poles is

$$\Gamma(t-s)\Gamma(t+s) \frac{\Gamma(1-s)\Gamma(1-t)}{\Gamma(1+s)\Gamma(1+t)} \times \left(\pi \sin \pi t + (\cos \pi s + \cos \pi t) \left(\Lambda + \frac{\Gamma'}{\Gamma}(1+s) + \frac{\Gamma'}{\Gamma}(1-s) - \frac{\Gamma'}{\Gamma}(t+s) - \frac{\Gamma'}{\Gamma}(t-s) \right) \right). \quad (43)$$

Note that, after bringing in the factor of $\Gamma(1+s)^2$ from (40), this is already even in s .

4.2.2 Poles at $z = k + s$

As it turns out, these contribute exactly zero to the expression, which greatly simplifies our task. Replacing $\sin \pi(t-z)$ with $\sin \pi(t-s) \cos \pi(s-z) + \cos \pi(t-s) \sin \pi(s-z)$, we find that the relevant terms are

$$(1 - \cos^2 \pi(z-s)) (\partial_s + \Lambda) \frac{\Gamma(t-s) \sin \pi(t-s)}{\Gamma(1+z-s) \sin \pi(s-z)}$$

$$+\pi \cos \pi(z-s) \frac{\Gamma(t-s) \sin \pi(t-s)}{\Gamma(1+z-s) \sin \pi(s-z)} \sin \pi(s-z).$$

When $z = k + s$, the first line has a double zero from the $1 - \cos^2$ factor, so there is no pole. Obviously, the $\sin \pi(s-z)$ terms cancel in the second line, so there is no pole for it either. This simplification arises completely from taking the Y_0 and K_0 terms together, and including both positive and negative values of h into one sum. Had we treated them separately, there would have been considerably more terms to consider (which, of course, would eventually cancel).

4.3 The rest of the calculation

Using (43), replacing $2\pi\hat{q}$ by $q^{1/2}$, and cancelling a few factors of 2π , we have

$$M^{OOD}(\ell) = \frac{2}{(2\pi i)^2} \int_{(1.7)} \int_{(0.6)} \frac{G(s)G(t)\Gamma(1+s)\Gamma(1-s)\Gamma(1+t)^2\zeta_q(1+2s)\zeta_q(1+2t)q^{s+t}}{(s^2-\mu^2)(t^2-\mu^2)(2\pi)^{2t}} \quad (44)$$

$$\frac{1}{\ell^{1/2}} \sum_{abe=\ell} \tau(b) \frac{\mu(a)}{a} \left(\frac{e}{ab}\right)^t \sum_g \frac{\mu(g)}{g^{1+2s}} \sum_{q|c} \frac{1}{c^{2s}} \sum_v \frac{\mu(v)}{v^2} \sum_w \frac{(ae, wv)}{w} \sum_{c,w|h} \frac{1}{h^{t-s}} (-\partial_t + \log h - \lambda_{ae, wv})$$

$$\left(\pi \sin \pi t + (\cos \pi s + \cos \pi t) \left(\frac{\Gamma'}{\Gamma}(1 \pm s) - \frac{\Gamma'}{\Gamma}(t \pm s) + \log h - \lambda_{1, wv} \right) \right)$$

$$\Gamma(t-s)\Gamma(t+s) \frac{\Gamma(1-t)}{\Gamma(1+t)} 2s ds 2t dt.$$

We now look to simplify the integer sums, primarily by removing as many divisibility conditions as possible. Everything appearing after $(e/ab)^t$ is a function only of ae , not any of a , b , or e individually, so we can set $A = ae$, making the sum on a, b, e

$$\frac{1}{\ell^{1/2+t}} \sum_{Ab=\ell} A^{2t} \tau(b) \prod_{p|A} (1 - p^{-1-2t}).$$

By Mobius inversion,

$$\sum_{c,w} f(c,w) \sum_{c,w|h} g(h) = \sum_{u,x} \mu(u) \sum_{c,w} f(uxc, uxw) \sum_h g(cwu^2xh),$$

so we can simplify the $c, w|h$ condition by introducing two new sums. This means that we replace $q|c$ by $q|cux$, but if $q|ux$, the resulting expression is small enough to ignore (in particular, $q|x$ gives the same terms except with an extra factor of q^{-1} , while the $q|u$ case can be immediately bounded by $q^{-2.1}\ell^{0.6}$). Thus we can replace c by cq , factoring q^{-s-t} from the c sum. We extend the c, g, u , and x sums out to infinity, since these extra terms contribute at most $\ell^{-1/2+t}q^{A(1-s-t)}$, which is negligible. Since $(q, g) = 1$, the g sum is thus $\zeta_q(1+2s)^{-1}$, exactly cancelling the term in the first line. We thus have, up to negligible error,

$$M^{OOD}(\ell) = \frac{2}{(2\pi i)^2} \int_{(1.7)} \int_{(0.6)} \frac{G(s)G(t)\Gamma(1+s)\Gamma(1-s)\Gamma(1+t)^2\zeta_q(1+2t)}{(s^2-\mu^2)(t^2-\mu^2)(2\pi)^{2t}} \quad (45)$$

$$\frac{1}{\ell^{1/2}} \sum_{Ab=\ell} \frac{A^{2t}}{\ell^t} \tau(b) \prod_{p|A} (1 - p^{-1-2t}) \sum_c \frac{1}{c^{t+s}} \sum_v \frac{\mu(v)}{v^2} \sum_x \frac{1}{x^{1+t+s}} \sum_{(u,q)=1} \frac{\mu(u)}{u^{1+2t}}$$

$$\sum_w \frac{(A, uvwx)}{w^{1+t-s}} \sum_h \frac{1}{h^{t-s}} (-\partial_t + \log qcwu^2 xh - \lambda_{A, uvwx})$$

$$\left(\pi \sin \pi t + (\cos \pi s + \cos \pi t) \left(\frac{\Gamma'}{\Gamma}(1 \pm s) - \frac{\Gamma'}{\Gamma}(t \pm s) + \log qcwu^2 xh - \lambda_{1, uvwx} \right) \right)$$

$$\Gamma(t-s)\Gamma(t+s) \frac{\Gamma(1-t)}{\Gamma(1+t)} 2sds2tdt.$$

Note that we have now completely removed powers of q from the expression. Thus the usual method of evaluation—shifting contours to get a negative power of q and taking the residues passed along the way—cannot work here. As it turns out, we will still be able to evaluate everything through a contour shift, since it will be even in both s and t . We discuss this more later, after additional adjustments to the integer sums.

Since u is squarefree, we can factor it into u_1 , which divides A , and u_2 , relatively prime to A . The former contributes $\mu(u_1)/u_1^{2t}$, the latter completes the $\zeta_q(1+2t)^{-1}$ factor. Replacing A by Au_1 then lets us combine the u_1 and b sums. We can also combine wx into w and ch into c , since their only appearances in the logarithms are as a product. The integer sums are then

$$\frac{1}{\ell^{1/2+t}} \sum_{A|\ell} A^{2t} \sum_c \frac{\sigma_{2s}(c)}{c^{t+s}} \sum_v \frac{\mu(v)}{v^2} \sum_w \frac{(A, vw)\sigma_{2s}(w)}{w^{1+t+s}}, \quad (46)$$

times a few logarithms. We can factor v as we did u , replacing A by Av_2 . The identity

$$\sum_{A,w} (A, w) f(A, w) = \sum_{d,g} d\mu(g) \sum_{A,w} f(dgA, dgw) = \sum_D \phi(D) \sum_{A,w} f(DA, Dw)$$

lets us eliminate the resulting (A, w) factor.

$$\sum_{D,w} (Dw)^{-s} \sigma_{2s}(Dw) f(D, w) = \sum_a \mu(a) \sum_{D,w} [D^{-s} \sigma_{2s}(D)] [w^{-s} \sigma_{2s}(w)] f(Da, wa)$$

(compare this with (15), which is the case $s = 0$). Using $\Psi(x)$ to denote $\sum_{p|x} \frac{\log p}{p-1}$, we have

$$M^{OOD}(\ell) = \frac{3}{(2\pi i)^2} \int_{(1.7)} \int_{(0.6)} \frac{G(s)G(t)\Gamma(1+s)\Gamma(1-s)\Gamma(1+t)^2}{(s^2 - \mu^2)(t^2 - \mu^2)(2\pi)^{2t}} \quad (47)$$

$$\frac{1}{\ell^{1/2+t}} \sum_{aAD|\ell} A^{2t} \frac{\phi(A)}{A} D^t \frac{\sigma_{2s}(D)}{D^s} \frac{\phi(aD)}{aD} \frac{\mu(a)}{a} \sum_{(v,aAD)=1} \frac{\mu(v)}{v^2}$$

$$\sum_c \frac{\sigma_{2s}(c)}{c^{t+s}} \sum_w \frac{\sigma_{2s}(w)}{w^{1+t+s}} (-\partial_t - 2\gamma + \log cq - \log w - \log \ell + \log D - 2\log v + 2\Psi(aD))$$

$$\left(\pi \sin \pi t + (\cos \pi s + \cos \pi t) \left[\frac{\Gamma'}{\Gamma}(1+s) + \frac{\Gamma'}{\Gamma}(1-s) \right. \right.$$

$$\left. \left. - \frac{\Gamma'}{\Gamma}(t+s) - \frac{\Gamma'}{\Gamma}(t-s) - 2\gamma + \log cq - \log w - \log D - 2\log a - 2\log v + 2\Psi(A) \right] \right)$$

$$\Gamma(t-s)\Gamma(t+s) \frac{\Gamma(1-t)}{\Gamma(1+t)} 2sds2tdt.$$

Note that this entire expression (except for sds) is even in s . Thus the contour integral at $\Re s = -0.6$ is the negative of the integral at $\Re s = 0.6$, meaning that the latter is exactly half the sum of the residues passed during the shift from one to the other. The only such poles are at $s = \pm\mu$, and each has the same residue, so the s integral equals the residue at $s = \mu$. When $\mu = 0$, the integral is half the residue at $s = 0$.

We next show that the integrand is also even in t . We can think of it as having two parts, the ℓ -dependent and the analytic. The former consists of most of the integer sums (including the logarithms of D , a , and v), and the latter consists of the ζ functions, Γ functions, and the sums on c and w . We will show that each part is even in t separately, which allows us to evaluate it through a contour shift. This evenness should not be a surprise, despite how complicated the integrand is. The t integral came originally from the functional equation for $\Lambda(f, s)^2$, an even function, and all we have done since is average it over f . It is possible that one could write the entire fourth moment, not just this off-off-diagonal term, as an even integrand in t , and this might assist in controlling the remainder terms.

4.3.1 The ℓ -dependent part

We ignore the various Γ factors, as well as everything associated with c and w , collecting them into the constants C_1 and C_2 . We want to show that

$$\frac{1}{\ell^t} \sum_{aAD|\ell} A^{2t} \frac{\phi(A)}{A} D^t \frac{\sigma_{2s}(D)}{D^s} \frac{\phi(aD)}{aD} \frac{\mu(a)}{a} \sum_{(v,aAD)=1} \frac{\mu(v)}{v^2} \quad (48)$$

$$[C_1 - \log D - 2 \log a - 2 \log v + 2\Psi(A)][C_2 + \log D - 2 \log v + 2\Psi(aD)]$$

is even in t . Letting $c = \ell/aAD$, we can rewrite (48) as $(A/ac)^t$ times an expression not depending on t . Thus it is enough to show that switching A with ac does not change anything in that expression. In particular, nothing with D matters except where it is mixed with a (so in particular, we can assume that D is squarefree). Also, $\mu(a) \log a$, $\mu(v) \log v$, $\Psi(A)$, and $\Psi(aD)$ do not depend on how often a prime divides the number, merely whether it divides it at all. Thus we can also assume that A is squarefree, as with ac . Given a prime p , let (e_1, e_2, e_3) denote the number of times it divides A , D , and ac , respectively. We need only worry about $0 \leq e_i \leq 1$, so there are only eight cases. Since $(A/ac)^t = p^{(e_1 - e_3)t}$, we can also ignore those cases where $e_1 = e_3$, so we need only consider four cases.

- The case $(1, 0, 0)$ yields

$$p^t \frac{p-1}{p} (C_1 + \frac{2 \log p}{p-1}) C_2.$$

- The case $(0, 0, 1)$ splits into three cases, based on $p|a$, $p|v$, or neither, giving

$$p^{-t} \left(-\frac{p-1}{p^2} (C_1 - 2 \log p) (C_2 + \frac{2 \log p}{p-1}) - \frac{1}{p^2} (C_1 - 2 \log p) (C_2 - 2 \log p) + C_1 C_2 \right).$$

This equals the previous expression except for the negation of t .

- The case $(1, 1, 0)$ yields

$$\sigma_{2s}(p) p^{-s} p^t \frac{(p-1)^2}{p^2} (C_1 - \log p + \frac{2 \log p}{p-1}) (C_2 + \log p + \frac{2 \log p}{p-1}).$$

- The case $(0, 1, 1)$ splits into two cases, $p|a$ or $(p, av) = 1$, giving

$$\begin{aligned} & \sigma_{2s}(p)p^{-s}p^{-t} \left(-\frac{p-1}{p^2}(C_1 - \log p - 2 \log p)(C_2 + \log p + \frac{2 \log p}{p-1}) \right. \\ & \quad \left. + \frac{p-1}{p}(C_1 - \log p)(C_2 + \log p + \frac{2 \log p}{p-1}) \right). \end{aligned}$$

Again, this matches the previous expression except that p^{-t} becomes p^t .

Thus this portion is even in t . One must also check cross-terms of the form $(\log p_1)(\log p_2)$, but they work in exactly the same way.

4.3.2 The analytic part

We now restrict our attention to the other parts of (47). To simplify notation, we let

$$C_3 = -2\gamma + \log q - \log \ell + \log D - 2 \log v + 2\Psi(aD),$$

$$C_4 = -2\gamma + \log q - \log D - 2 \log a - 2 \log v + 2\Psi(A) + \Gamma'/\Gamma(1 \pm s).$$

We want to analyze the behavior of

$$\begin{aligned} & \frac{\Gamma(1+t)^2}{(2\pi)^{2t}} \sum_c \frac{\sigma_{2s}(c)}{c^{t+s}} \sum_w \frac{\sigma_{2s}(w)}{w^{1+t+s}} (-\partial_t + \log \frac{c}{w} + C_3) \\ & \left(\pi \sin \pi t + (\cos \pi s + \cos \pi t) \left(-\frac{\Gamma'}{\Gamma}(t+s) - \frac{\Gamma'}{\Gamma}(t-s) + \log \frac{c}{w} + C_4 \right) \right) \\ & \quad \Gamma(t-s)\Gamma(t+s) \frac{\Gamma(1-t)}{\Gamma(1+t)}. \end{aligned}$$

We apply the $-\partial_t$ to the expression, and put the resulting $\Gamma(1 \pm t)'$ terms (whose sum is obviously even) into C_3 . This leaves us with $\Gamma(1+t)\Gamma(1-t)$, which is obviously even, times

$$\begin{aligned} & \frac{\Gamma(t+s)\Gamma(t-s)}{(2\pi)^{2t}} \sum_c \frac{\sigma_{2s}(c)}{c^{t+s}} \sum_w \frac{\sigma_{2s}(w)}{w^{1+t+s}} \tag{49} \\ & \left[-\pi^2 \cos \pi t + \pi \sin \pi t (C_3 + C_4 + 2 \log \frac{c}{w} - 2 \frac{(\Gamma(t+s)\Gamma(t-s))'}{\Gamma(t+s)\Gamma(t-s)}) \right. \\ & \quad + (\cos \pi s + \cos \pi t) \left(\frac{(\Gamma(t+s)\Gamma(t-s))'}{\Gamma(t+s)\Gamma(t-s)} \right)' \\ & \quad + (\cos \pi s + \cos \pi t) \left(C_3 + \log \frac{c}{w} - \frac{(\Gamma(t+s)\Gamma(t-s))'}{\Gamma(t+s)\Gamma(t-s)} \right) \\ & \quad \left. \left(C_4 + \log \frac{c}{w} - \frac{(\Gamma(t+s)\Gamma(t-s))'}{\Gamma(t+s)\Gamma(t-s)} \right) \right]. \end{aligned}$$

Here all derivatives are to be taken in the t variable. The w sum gives $\zeta(1+t+s)\zeta(1+t-s)$ and the c sum will give $\zeta(t+s)\zeta(t-s)$, in both cases possibly differentiated. We adjust the c sum through the asymmetric functional equation (see for example [D])

$$\frac{\Gamma(t)\zeta(t)}{(2\pi)^t} = \frac{\zeta(1-t)}{2\cos(\pi t/2)},$$

which implies

$$\frac{\Gamma(t-s)\Gamma(t+s)\zeta(t-s)\zeta(t+s)}{(2\pi)^{2t}} = \frac{\zeta(1-t-s)\zeta(1-t+s)}{4\cos\frac{\pi}{2}(t+s)\cos\frac{\pi}{2}(t-s)} = \frac{\zeta(1-t-s)\zeta(1-t+s)}{2(\cos\pi t + \cos\pi s)}.$$

We group the terms in (49) based on derivatives and factors of C_3 and C_4 .

1. The very first term, combined with the C_3C_4 term, gives

$$(-\pi^2 \cos \pi t + C_3C_4(\cos \pi s + \cos \pi t)) \frac{\zeta(1+t+s)\zeta(1+t-s)\zeta(1-t+s)\zeta(1-t-s)}{2(\cos \pi t + \cos \pi s)}. \quad (50)$$

2. The $(C_3 + C_4)$ terms give

$$(C_3 + C_4) \left(\pi \sin \pi t \frac{\zeta(1+t+s)\zeta(1+t-s)\zeta(1-t+s)\zeta(1-t-s)}{2(\cos \pi t + \cos \pi s)} \right. \quad (51)$$

$$+ \frac{(\cos \pi s + \cos \pi t)}{(2\pi)^{2t}} \left(- \left(\frac{(2\pi)^{2t}\zeta(1-t+s)\zeta(1-t-s)}{2(\cos \pi t + \cos \pi s)} \right)' \zeta(1+t+s)\zeta(1+t-s) \right.$$

$$\left. \left. + \frac{(2\pi)^{2t}\zeta(1-t+s)\zeta(1-t-s)}{2(\cos \pi t + \cos \pi s)} (\zeta(1+t+s)\zeta(1+t-s))' \right) \right).$$

The derivative of $(2\pi)^{2t}\zeta(1-t-s)\zeta(1-t+s)$, added to the derivative of $\zeta(1+t-s)\zeta(1+t+s)$, gives

$$(C_3 + C_4) \frac{\zeta(1+t+s)\zeta(1+t-s)\zeta(1-t+s)\zeta(1-t-s)}{2} \quad (52)$$

$$\left(-2 \log 2\pi + \frac{\zeta'}{\zeta}(1+t+s) + \frac{\zeta'}{\zeta}(1+t-s) + \frac{\zeta'}{\zeta}(1-t+s) + \frac{\zeta'}{\zeta}(1-t-s) \right).$$

The derivative of $1/(2\cos \pi t + 2\cos \pi s)$ cancels exactly with the first line of (51), so (52) is the entire contribution.

3. The $\pi \sin \pi t$ term (other than the $C_3 + C_4$) gives

$$\frac{2\pi \sin \pi t}{(2\pi)^{2t}} \left(- \frac{(2\pi)^{2t}\zeta(1-t+s)\zeta(1-t-s)}{2(\cos \pi t + \cos \pi s)} \right)' \zeta(1+t+s)\zeta(1+t-s) \quad (53)$$

$$+ \frac{(2\pi)^{2t}\zeta(1-t+s)\zeta(1-t-s)}{2(\cos \pi t + \cos \pi s)} (\zeta(1+t+s)\zeta(1+t-s))'.$$

4. The remaining $\cos \pi s + \cos \pi t$ terms give

$$\begin{aligned} & \frac{\cos \pi s + \cos \pi t}{(2\pi)^{2t}} \left(\left(\frac{(2\pi)^{2t} \zeta(1-t+s) \zeta(1-t-s)}{2(\cos \pi t + \cos \pi s)} \right)'' \zeta(1+t+s) \zeta(1+t-s) \right. \\ & - 2 \left(\frac{(2\pi)^{2t} \zeta(1-t+s) \zeta(1-t-s)}{2(\cos \pi t + \cos \pi s)} \right)' (\zeta(1+t+s) \zeta(1+t-s))' \\ & \left. + \frac{(2\pi)^{2t} \zeta(1-t+s) \zeta(1-t-s)}{2(\cos \pi t + \cos \pi s)} (\zeta(1+t+s) \zeta(1+t-s))'' \right). \end{aligned} \quad (54)$$

We group this by number of derivatives of the $\cos \pi t$, $(2\pi)^{2t}$, or ζ terms. Differentiating the \cos term in the denominator of the first two lines of (54) once gives exactly the negative of (53), so these terms drop. Differentiating it twice in the first line gives

$$\zeta(1+t+s) \zeta(1+t-s) \zeta(1-t+s) \zeta(1-t-s) \left(\frac{\pi^2 \cos \pi t}{2(\cos \pi s + \cos \pi t)} - \frac{\pi^2 \sin^2 \pi t}{(\cos \pi s + \cos \pi t)^2} \right). \quad (55)$$

Note that the first term exactly cancels the first term of (50). Differentiating the $(2\pi)^{2t}$ expression in the first two lines of (54) exactly once gives

$$\begin{aligned} & -2 \log 2\pi \zeta(1+t+s) \zeta(1+t-s) \zeta(1-t+s) \zeta(1-t-s) \\ & \left(\frac{\zeta'}{\zeta}(1+t+s) + \frac{\zeta'}{\zeta}(1+t-s) + \frac{\zeta'}{\zeta}(1-t+s) + \frac{\zeta'}{\zeta}(1-t-s) \right). \end{aligned} \quad (56)$$

Differentiating it twice in the first line of (54) gives

$$2(\log 2\pi)^2 \zeta(1+t+s) \zeta(1+t-s) \zeta(1-t+s) \zeta(1-t-s). \quad (57)$$

We are left with differentiating only the ζ terms. These give

$$\frac{\zeta(1+t+s) \zeta(1+t-s) \zeta(1-t+s) \zeta(1-t-s)}{2} \quad (58)$$

$$\left(\sum \frac{\zeta''}{\zeta}(1 \pm t \pm s) + 2 \sum' \frac{\zeta'}{\zeta}(1 \pm t \pm s) \frac{\zeta'}{\zeta}(1 \pm t \pm s) \right),$$

where the \sum' means that we only take the six pairs of distinct \pm patterns.

Combining (50), (52), (55), (56), (57), and (58) we find that the analytic part is

$$\begin{aligned} & \frac{1}{2} \zeta(1+t+s) \zeta(1+t-s) \zeta(1-t+s) \zeta(1-t-s) \\ & \left(C_3 C_4 + (C_3 + C_4) (-2 \log 2\pi + \sum \frac{\zeta'}{\zeta}) - \frac{2\pi^2 \sin^2 \pi t}{(\cos \pi t + \cos \pi s)^2} \right. \\ & \left. - 4 \log 2\pi \sum \frac{\zeta'}{\zeta} + 4(\log 2\pi)^2 + \sum \frac{\zeta''}{\zeta} + 2 \sum' \frac{\zeta'}{\zeta} \frac{\zeta'}{\zeta} \right), \end{aligned} \quad (59)$$

which is obviously even in t . Thus the t integral equals half the sum of the residues between $\Re t = 1.7$ and $\Re t = -1.7$. It appears that this involves poles at $t = \pm 1, \pm 1 \pm \mu$, but these can be eliminated by setting $G(t) = 0$ to sufficient order at these points. Thus all that contributes is the poles at $t = \pm \mu$, which have equal residues, so the main term equals the residue at $t = \mu$. When $\mu = 0$, the main term is half the residue at $t = 0$.

This completes the derivation of the main off-off-diagonal term. Before writing the complete expression explicitly in Theorem 4.1, we need to take care of the case $q|g$, which arose at the start of this section.

4.4 The case $q|g$

We now suppose that $q|g$, so that we may replace g by gq , with $(g, q) = 1$. We return to the original application of Proposition 2.3 with regards to the off-off-diagonal terms, letting F again be smooth with support on $(1/2, \infty)$ with $F(x) = 1$ when $x \in [1, \infty)$. We will not need to use the results of [DFI I] on the n sum, so we only use $F(m)$. After removing irrelevant constants, we need to bound

$$\frac{1}{q^2} \sum_g \frac{\mu(g)}{g^2} \sum_{cg < q^{A-1}} \frac{1}{c} \sum_{m,n} \frac{\tau(m)\tau(n)}{n^{1/2}} W\left(\frac{adn}{q}\right) \int_0^\infty J_1\left(\frac{\sqrt{aenv}}{qcg}\right) [-2\pi\delta_{m \equiv aen(c)} Y_0\left(\frac{\sqrt{mv}}{qcg}\right) + 4\delta_{m \equiv -aen(c)} K_0\left(\frac{\sqrt{mv}}{qcg}\right)] W\left(\frac{v^2}{q}\right) F(v^2) dv.$$

The congruence condition on m, n is now quite mild, and we will be able to get sufficient cancellation through multiplicative character sums.

We assume for the moment that $(m, c) = (ae, c) = 1$. Summing over all multiplicative characters modulo c , we have

$$\sum_{m,n} \delta_{m \equiv \pm aen(c)} F(m, n) = \frac{1}{\phi(c)} \sum_{\chi} \sum_{m,n} \chi(m) \bar{\chi}(\pm aen) F(m, n).$$

We use the Mellin transforms of W , J_1 , K_0 , and Y_0 , changing the order of integration to perform the v integral first (everything is absolutely convergent so this is justified). We thus wish to bound

$$\begin{aligned} & -2\pi \sum_g \sum_c \sum_{\chi} \bar{\chi}(ae) \sum_{m,n} \tau(m)\tau(n) \chi(m) \bar{\chi}(n) \\ & \frac{1}{(2\pi i)^3} \int_{(1.2)} \int_{(*)} \int_{(-0.1)} \frac{G(t)\Gamma(1+t)^2 \zeta_q(1+2t)\Gamma(\beta)^2}{(t^2 - \mu^2)\Gamma(1+z)\Gamma(2+z) \sin \pi z} (ad)^{-t} (ae)^{1/2+z} \\ & q^{-3-2z+2\beta+t} \frac{\mu(g)}{g^{3+2z-2\beta}} \frac{1}{\phi(c)c^{2+2z-2\beta}} (\cos \pi\beta + \chi(-1)) \\ & m^{-\beta} n^{-t+z} \int_0^\infty x^{1+z-\beta} W\left(\frac{x}{q}\right) F_M(x) \frac{dx}{x} dz d\beta 2t dt. \end{aligned}$$

As with the main term, the inner x integral decays quickly in β , so that we may shift the β contour to $\Re \beta = 1.2$. This makes the m sum absolutely convergent, so that we may move the c , g , m , and n sums inside the integrals, giving

$$-\frac{\pi}{(2\pi i)^3} \int_{(1.2)} \int_{(1.2)} \int_{(-0.1)} \frac{G(t)\Gamma(1+t)^2 \zeta_q(1+2t)\Gamma(\beta)^2}{(t^2 - \mu^2)\Gamma(1+z)\Gamma(2+z) \sin \pi z} (ad)^{-t} (ae)^{1/2+z}$$

$$q^{-3-2z+2\beta+t} \sum_g \frac{\mu(g)}{g^{3+2z-2\beta}} \sum_c \sum_\chi \frac{1}{\phi(c)c^{2+2z-2\beta}} (\cos \pi\beta + \chi(-1))$$

$$L(\chi, \beta)^2 L(\bar{\chi}, t-z)^2 \int_0^\infty x^{1+z-\beta} W\left(\frac{x}{q}\right) F_M(x) \frac{dx}{x} dz d\beta 2t dt. \quad (60)$$

We shift the β contour to the left, to $\Re\beta = \delta$. The pole at $\beta = 1$ from $L(\chi, \beta)^2$, when χ is principal, is cancelled by the double zero of $\cos \pi\beta + \chi(-1)$. We shift the z contour to $-\delta$ and the t contour to 1. By the convexity principle, $L(\beta, \chi)^2$ is dominated by $c^{1-\delta}$, while $L(t-z, \chi)^2$ is bounded. Taking absolute values, performing the absolutely convergent integrals (note that we can now replace F by 1 in the v integral with negligible loss), and summing over all χ , we can bound (60) by

$$e^{2\pi|\mu|} (1 + |\mu|)^B \ell^{1/2-\delta} q^{-1+2\delta} \sum_c c^{-1+3\delta}.$$

Since $c \ll q^{A-1}$, this is bounded by

$$e^{2\pi|\mu|} (1 + |\mu|)^B \ell^{1/2} q^{-1-2\delta+3\delta A},$$

which can be made negligible by taking δ small enough.

4.4.1 Common divisors

So far we have assumed that $(m, c) = 1$ and $(ae, c) = 1$. However, the general case is really no different. We wish to bound

$$\sum_c \frac{1}{c^{2+2z-2\beta}} \sum_{m \equiv \pm aen(c)} \frac{\tau(m)\tau(n)}{m^\beta n^{t-z}}. \quad (61)$$

Letting $d = (c, ae)$, we replace c, m by cd, md , so (61) becomes

$$\sum_{d|ae} \frac{1}{d^{2+2z-\beta}} \sum_{(c, ae)=1} \frac{1}{c^{2+2z-2\beta}} \sum_{\substack{m \equiv \pm aen/d \\ (\text{mod } c)}} \frac{\tau(md)\tau(n)}{m^\beta n^{t-z}}.$$

We then let $r = (m, c) = (n, c)$, so that (61) is now

$$\sum_{d|ae} \frac{1}{d^{2+2z-\beta}} \sum_{(r, ae)=1} \frac{1}{r^{2+z-\beta+t}} \sum_{(c, ae)=1} \frac{1}{c^{2+2z-2\beta}} \sum_{\substack{m \equiv \pm aen/d \\ (m, c)=1}} \frac{\tau(mdr)\tau(nr)}{m^\beta n^{t-z}}.$$

We have

$$\sum_m \frac{\tau(mh)\chi(m)}{m^\beta} = L(\chi, \beta)^2 \prod_{p^e || h} (e + 1 - e\chi(p)p^{-\beta}). \quad (62)$$

For the values of z, β , and t where we evaluate these terms, the r sum is absolutely convergent, and the d sum, which is only over divisors of ae , will not have any effect on the bounds given. This completes the analysis of the off-off-diagonal terms, and with it the asymptotic evaluation of the twisted fourth moments of the L -functions inside the critical strip.

4.5 The complete off-off-diagonal main term

Collecting all that has been done in this section, we obtain

Theorem 4.1 *For some absolute constant B , for all $\varepsilon > 0$ and $\ell < q$, the off-off-diagonal term $M^{OOD}(\ell)$ defined in (35) equals*

$$M^{OOD}(\ell) = \frac{1}{\ell^{1/2}} \frac{1}{(2\pi i)^2} \int_{(1.7)} \int_{(0.6)} I(\ell; s, t) \frac{2sds}{s^2 - \mu^2} \frac{2sdt}{t^2 - \mu^2} + O_\varepsilon(q^\varepsilon e^{2\pi|\mu|} (1 + |\mu|)^B \frac{\ell}{q^{1/2}})$$

where

$$\begin{aligned} I(\ell; s, t) := & G(s)G(t)\Gamma(1+s)\Gamma(1-s)\Gamma(1+t)\Gamma(1-t) \prod \zeta(1 \pm s \pm t) \quad (63) \\ & \times \frac{1}{\ell^t} \sum_{aAD|\ell} A^{2t} \frac{\phi(A)}{A} D^t \frac{\sigma_{2s}(D)}{D^s} \frac{\phi(aD)}{aD} \frac{\mu(a)}{a} \sum_{(v, aAD)=1} \frac{\mu(v)}{v^2} \\ & \times \left[\left(\log q - \log \frac{\ell v^2}{D} + 2\Psi(aD) - 2\gamma + \frac{\Gamma'}{\Gamma}(1 \pm s) \right) \right. \\ & \quad \times \left(\log q - \log Dv^2 + 2\Psi(A) - 2\gamma + \frac{\Gamma'}{\Gamma}(1 \pm t) \right) \\ & \quad + \left(2\log q - \log \ell v^4 + 2\Psi(A) + 2\Psi(aD) - 4\gamma + \frac{\Gamma'}{\Gamma}(1 \pm s) \right. \\ & \quad \quad \left. + \frac{\Gamma'}{\Gamma}(1 \pm t) \right) \left(-2\log 2\pi + \sum \frac{\zeta'}{\zeta}(1 \pm s \pm t) \right) \\ & \quad - \frac{2\pi^2 \sin^2 \pi t}{(\cos \pi t + \cos \pi s)^2} + 4(\log 2\pi)^3 - 4\log 2\pi \sum \frac{\zeta'}{\zeta}(1 \pm s \pm t) \\ & \quad \left. + \sum \frac{\zeta''}{\zeta}(1 \pm s \pm t) + 2 \sum' \frac{\zeta'}{\zeta}(1 \pm s \pm t) \frac{\zeta'}{\zeta}(1 \pm s \pm t) \right]; \end{aligned}$$

moreover $I(\ell; s, t)$ is an even function in both s and t .

Shifting the s and t contours to (ε) and (2ε) , respectively, we have

Corollary 4.2 *For $\ell \leq q^{1/3}$ and for all $\varepsilon > 0$,*

$$M^{OOD}(\ell) \ll_\varepsilon q^\varepsilon e^{2\pi|\mu|} (1 + |\mu|)^B \ell^{-1/2}.$$

Theorem 4.1 is sufficient for use in mollification. To find the asymptotics of $M(1)$ at $\mu = 0$ requires more calculation, but no more thought, one simply takes $\ell = 1$ and calculates residues. The lead term goes as $(\log q)^2$, significantly smaller than the diagonal and off-diagonal contributions.

5 Mollification at the center of the critical strip

In this section we use our evaluation of the twisted fourth power moment to evaluate mollified fourth power moments at the center of the critical strip (that is, for the special case $\mu = 0$). The same techniques apply for other values of μ , but since they have no particular arithmetic significance, we simplify the argument by focusing on the central point.

5.1 The mollifier

Following [IS] and [KMV], we take as our mollifier

$$M(f) = \sum_{m < M} \frac{\mu(m)\lambda_f(m)}{\psi(m)m^{1/2}} P\left(\frac{\log M/m}{\log M}\right), \quad (64)$$

where M is a (non-integer) small power of q , $\psi(m)$ is the product of $1 + 1/p$ over all primes dividing m and $P(x) = x^3$.

Some comment is in order about our choice of $P(x)$. For mollification of the second moment, the optimal choice (see [IS]) is $P(x) = x^2$, while in [KMV] it is shown that to mollify the second moment of derivatives optimally, one must take polynomials of higher degree (while still requiring $P(0) = P'(0) = 0$). It turns out here that, in order to have the fourth moment comparable to the square of the second moment, $P(x)$ must vanish to at least third degree at the origin. This is due to the $(\log q)^6$ and $(\log q)^5$ terms in the diagonal and off-diagonal contributions. For the $(\log q)^6$ term, taking $P(x) = x^2$ (or any other polynomial with non-vanishing second derivative at the origin) leads to a $(\log q)^6/(\log M)^8$ term in the mollified fourth moment (see section 5.3 below). Taking $P(x) = x^3$ makes that particular term go as $(\log q)^6/(\log M)^{12}$, while introducing new terms which are at most $(\log q)^6/(\log M)^{10}$. The mollified second moment (using either $P(x) = x^2$ or $P(x) = x^3$) goes as $(\log q)^{-2}$, so we want a fourth moment which goes as $(\log q)^{-4}$, making the x^2 mollifier inappropriate.

This requirement of greater smoothness in the mollifier coefficients is not surprising: had $P(x) = x$ been used in [IS], the first moment would have been asymptotic to $(\log q)^{-1}$, as would the second moment, making the non-vanishing arguments less successful. It should be pointed out that we have not formally optimized the choice of $P(x)$ as we did in [KMV]. However, given that x^2 was optimal for the second moment, it is likely that the simplest polynomial possible is also optimal for the fourth (and it certainly makes exposition easier).

Using the Mellin transform, we have

$$M(f) = \frac{1}{(\log M)^3} \frac{6}{2\pi i} \int_{(2)} \sum_{m < M} \frac{\mu(m)\lambda_f(m)}{\psi(m)} \frac{M^z dz}{m^z z^4}.$$

We first simplify $M(f)^4$ by use of the Hecke recursion. Replacing $\lambda_f(m_1)\lambda_f(m_2)$ by $\sum \lambda_f(m_1 m_2/d^2)$ and changing variables gives

$$\begin{aligned} M(f)^4 &= \frac{1}{(\log M)^{12}} \frac{1296}{(2\pi i)^4} \int_{(2)^4} \sum_{\ell} \frac{\lambda_f(\ell)}{\ell^{1/2}} \sum_g \frac{1}{g} \\ &\quad \sum_{\ell_1 \ell_3 = \ell} \sum_{(d_1, g\ell_1)=1} \frac{\mu(d_1)^2}{\psi(d_1)^2 d_1^{1+z_1+z_2}} \sum_{(d_3, g\ell_3)=1} \frac{\mu(d_3)^2}{\psi(d_3)^2 d_3^{1+z_3+z_4}} \\ &\quad \sum_{m_1 m_2 = g\ell_1} \frac{\mu(m_1)\mu(m_2)}{\psi(m_1)\psi(m_2) m_1^{z_1} m_2^{z_2}} \sum_{m_3 m_4 = g\ell_3} \frac{\mu(m_3)\mu(m_4)}{\psi(m_3)\psi(m_4) m_3^{z_3} m_4^{z_4}} \frac{M^{z_1+z_2+z_3+z_4} d\vec{z}}{z_1^4 z_2^4 z_3^4 z_4^4}. \end{aligned} \quad (65)$$

Evaluating the mollified fourth moment is thus simply a matter of replacing $\lambda_f(\ell)$ by $M(\ell)$ in (65) this formula and performing a series of contour shifts. In order to simplify the technical steps to come, we introduce the following conventions:

- We use $\nu(\vec{z}, m)$ for any arithmetic function of the form

$$\nu(\vec{z}, m) = \prod_{p|m} \left(1 + \frac{1}{p} f(\vec{z}, p)\right)$$

with $|f(\vec{z}, \mu, p)| = O(p^{3/4})$ uniformly for \vec{z} in the domain $\Re z_i \geq -1/4$.

- We use $\eta(\vec{z})$ to denote any Euler product of the form

$$\eta(\vec{z}) = \prod_p \left(1 + \frac{1}{p^2} f(\vec{z}, p)\right)$$

absolutely convergent and bounded (together with its partial derivatives) for \vec{z} in the domain $\Re z_i \geq -1/4$.

The idea is that terms of this nature will not affect the degrees of any poles we might encounter, and thus we can ignore them until we need to compute exact constants. Our notations are “generic” in the sense that the exact value of $\nu(\vec{z}, m)$ or $\eta(\vec{z})$ may change from line to line.

These conventions allow us to ignore the various relative primality restrictions in (65) “up to factors of $\eta(\vec{z})$ and $\nu(l, \vec{z})$,” so we may use the more compact form

$$M(f)^4 = \frac{1296}{(\log M)^{12}} \frac{1}{(2\pi i)^4} \int_{(2)^4} \sum_{\ell} \frac{\lambda_f(\ell)}{\ell^{1/2}} \nu(\vec{z}, \ell) \eta(\vec{z}) \left(\sum_{\substack{\ell=m_1 m_2 \times \\ m_3 m_4}} \frac{\mu(m_1)\mu(m_2)\mu(m_3)\mu(m_4)}{m_1^{z_1} m_2^{z_2} m_3^{z_3} m_4^{z_4}} \right) \prod_{i < j} \zeta(1 + z_i + z_j) \frac{M^{z_1+z_2+z_3+z_4} d\vec{z}}{z_1^4 z_2^4 z_3^4 z_4^4}. \quad (66)$$

The d_1 and d_3 sums give $\zeta(z_1 + z_2)\zeta(z_3 + z_4)$ while the g sum gives $\zeta(z_1 + z_3)\zeta(z_1 + z_4)\zeta(z_2 + z_3)\zeta(z_2 + z_4)$.

5.2 Using the twisted fourth moment

We now reduce $M(\ell)$ by throwing out all terms which contribute fewer than the maximum possible number of powers of $\log q$ (or $\log M$) to the mollified fourth moment. The remainder terms of Theorem 1.2, summed against $\ell^{-1/2+\epsilon}$ for $\ell < M^4 \sim q^{2\Delta}$, contribute at most

$$q^{5\Delta/2-1/12+\epsilon} + q^{21\Delta/4-1/4+\epsilon},$$

which is negligible so long as $\Delta < 1/30$. We may thus restrict our attention to the main terms $M^D(\ell) + M^{OD}(\ell)$ and $M^{OOD}(\ell)$, given in (34) and (63).

5.2.1 The diagonal and off-diagonal terms

We start with the main contribution of $M^D(\ell) + M^{OD}(\ell)$, as given by (34), namely

$$\frac{1}{\ell^{1/2}} \frac{1}{2\pi i} \int_{(3)} \text{Res}_{t=0} J(s, t; \ell) \frac{4ds}{st}$$

with

$$J(s, t; \ell) := G(s)G(t)\Gamma^2(1+t)\Gamma^2(1+s)\zeta_q(1+2s)\zeta(1+2t)\hat{q}^{2(s+t)}K(s, t; \ell),$$

$$K(s, t; \ell) = \sum_{abd=\ell} \frac{\mu(a)}{a^{1+s+2t}} \frac{\tau(b)}{b^t} \frac{1}{d^s} \sum_n \frac{\tau(adn)\tau(n)}{n^{1+s+t}}.$$

Now

$$\sum_n \frac{\tau(dn)\tau(n)}{n^s} = \frac{\zeta^4(s)}{\zeta(2s)} \tau(d) \prod_{p^\alpha \parallel d} \left(1 - \frac{2\alpha}{(\alpha+1)(p^s+1)}\right)$$

so that

$$K(s, t; \ell) = \frac{\zeta^4(1+s+t)}{\zeta(2(1+s+t))} \sum_{abd=\ell} \frac{\mu(a)}{a^{1+s+2t}} \frac{\tau(b)}{b^t} \frac{\tau(ad)}{d^s} \nu(1+s+t, ad). \quad (67)$$

Thus the contribution of these terms to the mollified fourth power is

$$\begin{aligned} & \frac{1296}{(\log M)^{12}} \frac{1}{(2\pi i)^4} \int_{(2)^4} \int_{(2)} \operatorname{Res}_{t=0} \sum_{\ell} \frac{J(s, t; \ell)}{\ell} \nu(\vec{z}, \ell) \eta(\vec{z}) \left(\sum_{\substack{\ell=m_1 m_2 \\ \times m_3 m_4}} \frac{\mu(m_1)\mu(m_2)\mu(m_3)\mu(m_4)}{m_1^{z_1} m_2^{z_2} m_3^{z_3} m_4^{z_4}} \right) \\ & \prod_{i<j} \zeta(1+z_i+z_j) \frac{4ds}{st} \frac{M^{z_1+z_2+z_3+z_4} d\vec{z}}{z_1^4 z_2^4 z_3^4 z_4^4}. \end{aligned} \quad (68)$$

The ℓ sum is then

$$\begin{aligned} & G(s)G(t)\Gamma^2(1+s)\Gamma^2(1+t)\zeta_q(1+2s)\zeta(1+2t)\hat{q}^{2(s+t)} \\ & \times \eta(s, t, \vec{z}) \frac{\zeta^4(1+s+t)}{\zeta(2(1+s+t))} \prod_{i=1}^4 [\zeta(1+z_i+s)\zeta(1+z_i+t)]^{-2}. \end{aligned}$$

To evaluate this, we first shift all of the z contours to $\Re z = 1/2$, then evaluate the t residue, then shift the s contour to $\Re s = -1/2$. The resulting contour integral is bounded by

$$O_\varepsilon(q^\varepsilon M^2 q^{-1/2})$$

which is admissible as long as $\Delta < 1/2$. This allows us to replace in (68) the s -integral by the residues of the integrand at $s = 0$:

$$\begin{aligned} & \frac{1296}{(\log M)^{12}} \frac{1}{(2\pi i)^4} \int_{(1/2)^4} \operatorname{Res}_{s,t=0} \left[\frac{4}{st} \eta_1(s, t, \vec{z}) G(s)G(t) \frac{\Gamma(1+s)^2 \Gamma(1+t)^2 \zeta_q(1+2s)\zeta(1+2t)}{\prod_{i=1}^4 [\zeta(1+z_i+s)\zeta(1+z_i+t)]^2} \right. \\ & \left. \hat{q}^{2(s+t)} \frac{\zeta^4(1+s+t)}{\zeta(2(1+s+t))} \right] \prod_{i<j} \zeta(1+z_i+z_j) \frac{M^{z_1+z_2+z_3+z_4} d\vec{z}}{z_1^4 z_2^4 z_3^4 z_4^4}. \end{aligned} \quad (69)$$

The exact value of the $s = t = 0$ residues is messy enough that there is no value in printing it, but we should point out that the only contributions to the main term will come from differentiations in s or t of either q^{s+t} or functions of the form $\zeta(1+z_i)$. All else, such as the $G(s)\Gamma(1+s)^2$, the η_1 , or the $\zeta(2+2s+2t)$, will contribute fewer powers of $\log q$ or $\log M$, and can thus be ignored.

Let us take a moment, then, to examine the total degree in $\log q$ and $\log M$ which we expect out of this expression after taking the residues at the origin (which is where the z_i integrals will be evaluated as well). The twelve factors of $\zeta(1+*)$ in the numerator contribute twelve poles, the

sixteen factors of $\zeta(1+*)$ in the denominator contribute sixteen zeros, and the factors of s , t , and z_i^4 contribute another eighteen poles. We thus have a total of fourteen net poles, evaluated over the course of six variables. Since each pole can give at most one extra factor of $\log q$ (in the case of s and t poles) or $\log M$ (for the z_i poles), we expect the main term to involve at most eight powers of $\log q$ or $\log M$ from the integrand, multiplied by the $(\log M)^{-12}$ in front. Thus anything which gives less than this can be ignored. Note also that the s and t residues by themselves will contribute at most $(\log q)^6$, so we will need at least six poles from the z integrands to bring the total up to eight powers. This will be important later, when many of the z_i poles are simple and thus do not contribute enough.

5.2.2 The off-off-diagonal term

We consider now the off-off diagonal term: by Theorem 4.1 (using the evenness of $I(\ell; s, t)$), the main term of $M^{OOD}(\ell)$ equals

$$\frac{1}{4} \text{Res}_{s=t=0} \frac{I(\ell; s, t)}{\ell^{1/2}} \frac{2}{s} \frac{2}{t} = \text{Res}_{t=0} \frac{I(\ell; 0, t)}{\ell^{1/2} t},$$

so that the contribution of the off-off-diagonal term is given by

$$\begin{aligned} & \frac{1296}{(\log M)^{12}} \frac{1}{(2\pi i)^4} \int_{(2)^4} \text{Res}_{t=0} \sum_{\ell} \frac{I(0, t; \ell)}{t} \frac{\nu(\vec{z}, \ell)}{\ell} \left(\sum_{\substack{\ell=m_1 m_2 \times \\ m_3 m_4}} \frac{\mu(m_1)\mu(m_2)\mu(m_3)\mu(m_4)}{m_1^{z_1} m_2^{z_2} m_3^{z_3} m_4^{z_4}} \right) \\ & \times \eta(\vec{z}) \prod_{i < j} \zeta(1 + z_i + z_j) \frac{M^{z_1+z_2+z_3+z_4} d\vec{z}}{z_1^4 z_2^4 z_3^4 z_4^4}. \end{aligned} \quad (70)$$

Again, we are only interested in the main term upon sliding the z contours past the origin. This allows us to simplify this expression considerably, since we know that differentiation of any term other than a power of M or a $\zeta(1+*)$ will lead to lower orders of magnitude. Among other things, this lets us drop expressions like $\log v$, γ , or $\Gamma'/\Gamma(1 \pm s)$ from the formula for $I(0, t; \ell)$. Thus we are left with (70) where $\frac{1}{t} I(0, t; \ell)$ is replaced by

$$\begin{aligned} & \text{Res}_{t=0} \frac{1}{t^5} \frac{1}{\ell^t} \sum_{a|AD|\ell} A^{2t} \frac{\phi(A)}{A} D^t \tau(D) \frac{\phi(aD)}{aD} \frac{\mu(a)}{a} \sum_{v, (v, aAD)=1} \frac{\mu(v)}{v^2} \\ & \times \left[\left((\log q)^2 - \log q \log \ell + \left(\log \frac{\ell}{D} \right) \log D \right) + \frac{4}{t^2} \right], \end{aligned} \quad (71)$$

the $\frac{4}{t^2}$ coming from the sums of ζ''/ζ and $(\zeta'/\zeta)(\zeta'/\zeta)$. The series

$$\sum_{\ell} \frac{1}{\ell^{1+t}} \sum_{a|AD|\ell} A^{2t} \frac{\phi(A)}{A} D^t \tau(D) \frac{\phi(aD)}{aD} \frac{\mu(a)}{a} \sum_{(v, aAD)=1} \frac{\mu(v)}{v^2} \left(\sum_{\substack{\ell=m_1 m_2 \times \\ m_3 m_4}} \frac{\mu(m_1)\mu(m_2)\mu(m_3)\mu(m_4)}{m_1^{z_1} m_2^{z_2} m_3^{z_3} m_4^{z_4}} \right)$$

can be rewritten as

$$\eta_2(t, \vec{z}) \left[\prod_i \zeta(1 + z_i + t) \zeta(1 + z_i - t) \zeta(1 + z_i)^2 \right]^{-1}.$$

We also need to check how the $-\log \ell$ and $-\log D$ factors contribute. Up to terms of lower order, the $-\log \ell$ corresponds to differentiating at $u = 0$ the product

$$\left[\prod_i \zeta(1+u+z_i+t)\zeta(1+u+z_i-t)\zeta(1+u+z_i)^2 \right]^{-1},$$

so we may replace $-\log \ell$ in (71) by

$$-\sum_i \frac{\zeta'}{\zeta}(1+z_i+t) + \frac{\zeta'}{\zeta}(1+z_i-t) + 2\frac{\zeta'}{\zeta}(1+z_i).$$

Similarly, $-\log D$ can be replaced by

$$-2\sum_i \frac{\zeta'}{\zeta}(1+z_i)$$

and so the main term is given by

$$\begin{aligned} & \frac{1}{(\log M)^{12}} \frac{1296}{(2\pi i)^4} \int_{(2)^4} \operatorname{Res}_{t=0} \frac{\eta_2(0, \vec{z}) M^{z_1+z_2+z_3+z_4} \prod \zeta(1+z_i+z_j)}{t^5 \prod_i z_i^4 \zeta(1+t+z_i)\zeta(1-t+z_i)\zeta(1+z_i)^2} \\ & \left((\log q)^2 + (\log q) \left(\sum_i 2\frac{\zeta'}{\zeta}(1+z_i) + \frac{\zeta'}{\zeta}(1+z_i+t) + \frac{\zeta'}{\zeta}(1+z_i-t) \right) \right. \\ & \left. + 2\sum_i \frac{\zeta'}{\zeta}(1+z_i) \left(\sum_i \frac{\zeta'}{\zeta}(1+z_i+t) + \frac{\zeta'}{\zeta}(1+z_i-t) \right) + \frac{4}{t^2} \right) d\vec{z}. \end{aligned} \quad (72)$$

Again, we check for total number of poles. There are six poles from the ζ functions in the numerator, sixteen zeros from the ζ functions in the denominator, and another twenty-one poles from the factors of t and z_i^4 in the denominator. The expression in parentheses gives either two extra poles or two extra powers of $\log q$, for a net of thirteen poles over five integrals. Thus, as with the diagonal and off-diagonal terms, the main term goes as $(\log M)^{-4}$ and we can ignore anything which contributes less. We next study the contour shifts of the z integrals more carefully, as they involve a phenomenon which has not occurred in the study of lower moments (namely, more than one sequence of residues contributing to the main term).

5.3 Shifting the z contours

We evaluate (69) and (72) by shifting the z contours, one at a time, past the origin. We show that the main terms come only from various sequences of poles passed in this process, but this requires some care, due to the presence of $\zeta(1+z_i)$ terms in the denominator. We wish to avoid any poles coming from zeros of these factors in our contour shift. Under the Riemann Hypothesis (or even a “quasi-RH” giving a zero-free strip), this is simple: we shift z_1 to $\Re z_1 = -\delta$, shift the other z 's to $\Re z_i = \delta/4$, and the resulting contour integral is dominated by $M^{-\delta/4}$, which is small enough to ignore.

Without the Riemann Hypothesis, the contour shifts are still feasible, just more intricate. Let C_γ denote a contour lying to the left of the line $\Re z = 0$ but to the right of all zeros of $\zeta(1+z)$, and

on which $|(1/\zeta(1+z))^{(j)}| + |\zeta(1+z)^{(j)}| \ll \log(1+|\Im mz|)^B$, for some fixed B , for all derivatives $j \leq 8$. By the prime number theorem one may take C_γ of the form

$$C_\gamma = \left\{ -\frac{C}{\log(|u|+2)} + iu, u \in \mathbf{R} \right\}$$

for some positive constant $C < \frac{1}{4} \log 2$. We shift the z contours to C_γ , it will turn out that these bounds are sufficient for our needs. To avoid, for as long as possible, poles coming from $z_i = z_j$, we start with $z_i \in -C_\gamma/(6-i)$, $i = 1 \dots 4$, then (after taking the s and t residues) we shift the z_1 contour to C_γ . In the process we get residues coming from the poles located at $z_1 = 0, z_1 = -z_2, z_1 = -z_3, z_1 = -z_4$ to which we will return later. The new integral of $z_1 \in C_\gamma$ is small: we shift the 3 other contours to $\Re z_i = 1/(6-i) \log q, i = 2, 3, 4$ and bound the resulting integrals by $(\log q \log(1+|z_1|))^{B'}$ for some fixed B' depending on B . The z_1 -integral along C_γ is thus bounded by

$$\begin{aligned} & \int_{\mathbf{R}} (\log q)^{B'} (\log(2+|u|))^{B'} M^{-C/2 \log(2+|u|)} \frac{du}{(1+|u|)^4} \\ & \ll (\log q)^{B'} e^{B' \log(\log q) - b_1 (\log M)^{1/2}} \ll e^{-\delta \log^{1/2} q} \end{aligned}$$

for some absolute $\delta = \delta(B', \Delta) > 0$. This is certainly sufficient, so we may ignore the resulting z_1 contour and deal only with the residues passed during the shift which are integrals in the variables z_2, z_3, z_4 .

We repeat this process by shifting one of z_2, z_3, z_4 to C_γ ; the rule being that we next shift z_i where i is the smallest index for which M^{z_i} appears in the integrand to a positive power. So usually we shift z_2 next, but for the $z_1 = -z_2$ pole, where the M factor is $M^{z_3+z_4}$, we shift z_3 instead. In the case of the z_2 shift, we pass poles at $z_2 = 0, z_2 = -z_3$, and $z_2 = -z_4$, and get a contour integral at C_γ which can be bounded by $e^{-\delta_2 \log^{1/2} q}$. We can thus ignore the contour and repeat the process. There are five different types of pole sequences which can occur in this process, we now list them.

1. One possibility is that the first two poles eliminate powers of M from the integrand. There are three such sequences, namely $(z_1 = -z_2, z_3 = -z_4)$, $(z_1 = -z_3, z_2 = -z_4)$, and $(z_1 = -z_4, z_2 = -z_3)$. In each case we are left with an integral in two variables with no dependence on M , with a factor of at most $(\log q)^6/(\log M)^{12}$ out front. Since the integrals will be absolutely bounded, their contribution is at most $O((\log q)^{-6})$, too small to be part of the main term. Note that had we taken $P(x) = x^2$ in the mollifier, these poles would have gone as $O((\log q)^{-2})$, larger than the main term.
2. Another possibility is that the first three poles eliminate powers of M from the integrand. There are nine such sequences, up to permutation of indices (and order) they look like either $(z_1 = 0, z_2 = 0, z_3 = -z_4)$ or $(z_1 = -z_2, z_3 = z_2, z_4 = -z_4)$. These lead to the same situation as the previous case, not enough poles have been hit yet to get enough powers of $\log M$, and the remaining poles cannot increase the order of magnitude.
3. A third possibility is that there is only a factor of M^{-z_i} after the third pole is taken. There are eight such sequences, up to permutation of indices and order they all look like $(z_1 = 0, z_2 = -z_3, z_4 = -z_3)$. These are easier, we shift the remaining variable to $-C_\gamma$ and get a negligible contour integral, without passing any more poles.

4. We can also have a factor of M^{-2z_i} after three poles, there are three sequences leading to this, all of which look like $(z_1 = -z_2, z_3 = -z_2, z_4 = -z_2)$. The same argument as case 3 still applies, we shift to $-C_\gamma$ without passing poles.
5. Finally, there are five sequences which contribute a main term (this is rather different from the second moment, where $z_1 = z_2 = 0$ is the only main term, $z_1 = -z_2$ being like case one here). Those five sequences are:

$$\begin{aligned}
&(z_1 = 0, z_2 = 0, z_3 = 0, z_4 = 0), \\
&(z_1 = 0, z_2 = -z_3, z_4 = z_3, z_3 = 0), \\
&(z_1 = -z_2, z_3 = z_2, z_2 = 0, z_4 = 0), \\
&(z_1 = -z_2, z_3 = 0, z_4 = z_2, z_2 = 0), \\
&(z_1 = -z_3, z_2 = 0, z_4 = z_3, z_3 = 0).
\end{aligned}$$

We will discuss the exact value of the residues from these sequences in a moment, first we look to calculate the constants which arise from the various η and ν functions.

5.4 Computation of the constants

In this section, we compute the exact value of the constants $\eta_1(0, 0, \vec{0})/\zeta(2)$ and $\eta_2(0, \vec{0})$ from (69) and (72). We do this by factoring over primes (both of these functions come from Euler products). Since we already know which $\zeta(1 + z_i + \dots)$'s will appear (or, equivalently, which factors of $(1 - p^{-1-z_i-\dots})^{-1}$ will arise) we will simplify by setting $s = t = z_i = 0, i = 1 \dots 4$ and performing formal computations with the Euler products. In particular, this means that we will also be allowing factors of the form $\zeta(1) = \prod_p (1 - 1/p)^{-1}$.

We start with the mollifier: we define the formal dirichlet series

$$\begin{aligned}
\sum_{\ell} \frac{\lambda_f(\ell)}{\ell^{1/2}} \Theta(\ell) &:= \sum_{\ell} \frac{\lambda_f(\ell)}{\ell^{1/2}} \sum_g \frac{1}{g} \sum_{\ell_1 \ell_3 = \ell} \left(\sum_{(d_1, g\ell_1)=1} \frac{\mu(d_1)^2}{\psi(d_1)^2 d_1} \right) \left(\sum_{(d_3, g\ell_3)=1} \frac{\mu(d_3)^2}{\psi(d_3)^2 d_3} \right) \\
&\times \left(\sum_{m_1 m_2 = g\ell_1} \frac{\mu(m_1)\mu(m_2)}{\psi(m_1)\psi(m_2)} \right) \left(\sum_{m_3 m_4 = g\ell_3} \frac{\mu(m_3)\mu(m_4)}{\psi(m_3)\psi(m_4)} \right),
\end{aligned}$$

which is (65) evaluated at $z_1 = z_2 = z_3 = z_4 = 0$. Since m_1, m_2, m_3, m_4 are all square-free, ℓg^2 must be fifth-power free, so we are left with a finite calculation for each prime which might divide ℓ . Performing the d and m sums, we have

$$\begin{aligned}
\Theta(\ell) &= \prod_p \frac{1}{(1 + p^{-1})^4} \sum_{\ell_1 \ell_3 = \ell} \sum_g \frac{1}{g} \prod_{(p, g\ell_1)=1} (1 + 3p^{-1} + p^{-2}) \prod_{p|g\ell_1} (-2 - 2p^{-1}) \prod_{p^2||g\ell_1} 1 \prod_{p^3|g\ell_1} 0 \quad (73) \\
&\prod_{(p, g\ell_3)=1} (1 + 3p^{-1} + p^{-2}) \prod_{p|g\ell_3} (-2 - 2p^{-1}) \prod_{p^2||g\ell_3} 1 \prod_{p^3|g\ell_3} 0.
\end{aligned}$$

Breaking into cases based on $p^e \parallel \ell$, the formal Euler factor for the mollifier is

$$\begin{array}{r}
e \\
0 \\
1 \\
2 \\
3 \\
4 \\
\geq 5
\end{array}
\begin{array}{l}
(1 + p^{-1})^4 \Theta(p^e) \\
(1 + 10p^{-1} + 20p^{-2} + 10p^{-3} + p^{-4}) \\
(-4 - 20p^{-1} - 20p^{-2} - 4p^{-3}) \\
(6 + 15p^{-1} + 6p^{-2}) \\
-4(1 + p^{-1}) \\
1 \\
0.
\end{array}$$

We turn now to the twisted power $M(\ell)$, which we calculate in a similar fashion. To evaluate $M^D(\ell) + M^{OD}(\ell)$, let $\Phi(\ell)$ denote the multiplicative function defined in 67 for $s = t = 0$:

$$\Phi(\ell) = \sum_{ab=\ell} \tau(b) \prod_{\substack{p^\alpha \parallel a \\ \alpha > 0}} (1 - \frac{1}{p})(\alpha + 1 - \frac{2\alpha}{p+1}). \quad (74)$$

Similarly for the off-off-diagonal term we may also consider the arithmetic function defined at the top line of (71) (which, to facilitate comparison with (74), we write with D replaced by b):

$$\begin{aligned}
\Phi'(\ell) &= \zeta(2) \sum_{aAb|\ell} \tau(b) \frac{\phi(A)}{A} \frac{\phi(ab)}{ab} \frac{\mu(a)}{a} \sum_{v, (v, aAb)=1} \frac{\mu(v)}{v^2} \\
&= \sum_{aAcb=\ell} \tau(b) \frac{\phi(A)}{A} \frac{\phi(ab)}{ab} \frac{\mu(a)}{a} \prod_{p|aAb} (1 - \frac{1}{p^2})^{-1}.
\end{aligned}$$

We now show that $\Phi = \Phi'$. Clearly it is enough to consider ℓ a prime power. If $\ell = 1$ both equal 1, and if $1 < b < \ell = p^e$, one can quickly convert Φ' to Φ (the b 's in the two expressions correspond). Finally, one can check directly that the sum of the terms $b = 1$ and $b = \ell = p^e > 1$ is

$$e \frac{2p^2 - p + 1}{p^2 + p} + 2 - \frac{1}{p}$$

for both Φ and Φ' . Thus $\eta_1(0, 0, \vec{0})/\zeta(2) = \eta_2(0, \vec{0})$ and the constant in front of the residues coming from these factors is

$$\sum_{\ell} \frac{1}{\ell} \Theta(\ell) \frac{\Phi(\ell)}{\zeta(2)} = \sum_l \frac{1}{l} \Theta(l) \frac{\Phi'(l)}{\zeta(2)} = \prod_p \sum_{e=0}^4 \frac{1}{p^e} \Theta(p^e) \frac{\Phi(p^e)}{\zeta(2)}.$$

Again breaking into cases based on $p^e = \ell$, for our cases of interest we have

$$\begin{array}{r}
e \\
0 \\
1 \\
2 \\
3 \\
4
\end{array}
\begin{array}{l}
(1 - p^{-2})\Phi(\ell) \\
(1 - \frac{1}{p^2}) \\
4(1 - p^{-1}) \\
(1 - p^{-1})(10 - 5p^{-1} + p^{-2}) \\
(1 - p^{-1})(20 - 16p^{-1} + 4p^{-2}) \\
(1 - p^{-1})(35 - 35p^{-1} + 10p^{-2}).
\end{array}$$

We point out here for future reference that one can actually evaluate $\Phi(p^e)$ in closed form:

$$\Phi(p^e) = (e+1) + \frac{e(e+1)(e+5)}{6}(1-p^{-1}) - \frac{e(e+1)(e+2)}{3} \frac{1-p^{-1}}{p(1+p^{-1})}. \quad (75)$$

Using the two tables, we find that the constant in front of the residue calculations is

$$\sum_l \frac{1}{\ell} \Theta(\ell) \frac{\Phi(\ell)}{\zeta(2)} = \prod_p \frac{(1-p^{-1})^6}{(1+p^{-1})^4} = \frac{\zeta^4(2)}{\zeta(1)^{10}} = \frac{\eta_1(0, 0, \vec{0})}{\zeta(2)\zeta(1)^{10}} = \frac{\eta_2(0, \vec{0})}{\zeta(1)^{10}}. \quad (76)$$

Since we had found six $\zeta(1+z_i+\dots)$ factors in the numerator and sixteen in the denominator, the $\zeta(1)^{-10}$ agrees with expectations, and the constant by which we will multiply the residues is $\zeta(2)^4$. There is a $\zeta(2)$ in the first moment and a $\zeta(2)^2$ in the second moment (see [IS]), so this is exactly as expected.

5.5 Calculating residues

We now have done enough to be able to calculate the residues which contribute to the main term. We take (69) and (72) and replace every $\zeta(1+z)$ by $1/z$, every function not involving a $\zeta(1+*)$ or a power of q or M by the constant $\zeta(2)^4$, and every contour integral by one of the five sequences of residues. Using the MAPLE “residue” command and replacing $\log M$ by $\frac{\Delta}{2} \log q$ (which is all that matters for the main term) we find that the $(z_1 = 0, z_2 = 0, z_3 = 0, z_4 = 0)$ pole sequence contributes

$$\frac{\zeta(2)^4}{(\log q)^4} \left(\frac{2}{\Delta}\right)^9 \left(648\left(\frac{\Delta}{2}\right)^5 + \frac{52128}{35}\left(\frac{\Delta}{2}\right)^4 + \frac{6624}{5}\left(\frac{\Delta}{2}\right)^3 + \frac{3024}{5}\left(\frac{\Delta}{2}\right)^2 + 144\left(\frac{\Delta}{2}\right) + \frac{72}{5}\right). \quad (77)$$

The $(z_1 = -z_2, z_3 = z_2, z_2 = 0, z_4 = 0)$ pole sequence gives

$$\frac{\zeta(2)^4}{(\log q)^4} \left(\frac{2}{\Delta}\right)^{10} \left(\frac{108}{5}\left(\frac{\Delta}{2}\right)^5 + \frac{378}{5}\left(\frac{\Delta}{2}\right)^4 + \frac{477}{5}\left(\frac{\Delta}{2}\right)^3 + \frac{99}{2}\left(\frac{\Delta}{2}\right)^2 + \frac{54}{5}\left(\frac{\Delta}{2}\right) + \frac{27}{10}\right), \quad (78)$$

while the other three (which are symmetric under a change of variable order) each contribute

$$-\frac{\zeta(2)^4}{(\log q)^4} \left(\frac{2}{\Delta}\right)^9 \left(\frac{396}{35}\left(\frac{\Delta}{2}\right)^4 + \frac{198}{5}\left(\frac{\Delta}{2}\right)^3 + \frac{243}{5}\left(\frac{\Delta}{2}\right)^2 + \frac{45}{2}\left(\frac{\Delta}{2}\right) + \frac{18}{5}\right). \quad (79)$$

Adding together (77), (78), and three copies of (79) shows that the mollified fourth moment is

$$\begin{aligned} \sum_f^h L(f, \frac{1}{2})^4 M(f)^4 &= \frac{\zeta(2)^4}{(\log q)^4} (1 + O(\frac{1}{\log q})) \times \\ &\times \left(\frac{2}{\Delta}\right)^{10} \left(648\left(\frac{\Delta}{2}\right)^6 + \frac{51696}{35}\left(\frac{\Delta}{2}\right)^5 + \frac{6408}{5}\left(\frac{\Delta}{2}\right)^4 + \frac{2772}{5}\left(\frac{\Delta}{2}\right)^3 + 126\left(\frac{\Delta}{2}\right)^2 + \frac{72}{5}\left(\frac{\Delta}{2}\right) + \frac{27}{10}\right). \end{aligned} \quad (80)$$

This is the term given in theorem 1.4 (6). For comparison, the first and second moments with this mollifier (using the formula for general mollifiers in [KMV]) are

$$\frac{3\zeta(2)}{\log q} \left(\frac{2}{\Delta}\right) \quad \text{and} \quad \frac{\zeta(2)^2}{(\log q)^2} \left(\frac{2}{\Delta}\right)^3 (12 + 18\left(\frac{\Delta}{2}\right)),$$

respectively. For small Δ , the first moment goes as Δ^{-1} , the second moment as Δ^{-3} , and the fourth moment as Δ^{-10} (this part comes only from (78)), so any $\Delta < 1/30$ gives a fourth moment which is much larger than the square of the second moment. Thus further improvements on the number of non-vanishing $L(f, 1/2)$'s are out of the question (this is completely to be expected, it would be startling for the fourth moment to be a stronger bound than the second). On the other hand, notice what happens for large Δ (think of this as being able to mollify the non-vanishing elements completely). The square of the first moment approaches half the second moment, and the square of the second moment approaches half the fourth moment, which is what one would expect, since it is generally believed (see in particular [Br]) that half of these L -functions do not vanish for large q .

It is, in principle, possible to combine these results with the methods of [KMV] to find the asymptotic fourth moments of the k th derivative of $\Lambda(f, s)$ for any $k \geq 0$. One useful feature is that the quality of mollification (namely, the size of Δ) becomes unimportant for large k . In effect, we can think of this as allowing M to be arbitrarily large. Thus it is likely that, despite the considerable restrictions on Δ , the fourth moments of the k th derivatives will approach one-half the square of the second moments as k gets large. Since there is no particular significance to these asymptotics, we have not pursued this approach further.

6 Removing the harmonic weight

In this section we explain the necessary arguments to remove the harmonic weight $1/4\pi(f, f)$ from (6) in Theorem 1.4 so that (7) also holds. This technique has been performed several times (see [IS, KM1, KM2]) so we will give only a sketch of the proof. The basic idea is that $4\pi(f, f)$ equals (up to admissible terms)

$$\frac{|S_2(q)^*|}{\zeta(2)} L(\text{sym}^2 f, 1)$$

where

$$L(\text{sym}^2 f, s) = \sum_n \frac{\rho_f(n)}{n^s} = \sum_n \frac{1}{n^s} \sum_{m^2 l = n} \varepsilon_q(m) \lambda_f(l^2)$$

Since we only need this value at the edge of the critical strip, we can approximate $L(\text{sym}^2 f, 1)$ by a very short Dirichlet polynomial (for $\kappa < 1$)

$$\omega_f(q^\kappa) = \sum_{n \leq q^\kappa} \frac{\rho_f(n)}{n} = \sum_{m^2 l \leq q^\kappa} \varepsilon_q(m) \frac{\lambda_f(l^2)}{m^2 l} = \sum_{m^2 l \leq q^\kappa} \frac{\lambda_f(l^2)}{m^2 l}. \quad (81)$$

This is possible unconditionally so long as we work on average. We borrow from [KM1] a general proposition which is the axiomatization of the method described in [KM2].

Proposition 6.1 *For q prime, let $(\alpha_f)_{f \in S_2(q)^*}$ be a family of complex numbers satisfying*

$$\sum_f^h |\alpha_f| \ll (\log q)^A \quad \text{for some absolute } A > 0 \quad (82)$$

$$\text{Max}_{f \in S_2(q)^*} \frac{|\alpha_f|}{4\pi(f, f)} \ll q^{-\delta} \quad \text{for some absolute } \delta > 0 \quad (83)$$

Then for all $\kappa > 0$ there exists a $\gamma = \gamma(\kappa, \delta) > 0$ such that

$$\sum_f^n \alpha_f = \frac{1}{\zeta(2)} \sum_f^h \omega_f(q^\kappa) \alpha_f + O(q^{-\gamma}).$$

We want to apply this proposition to $\alpha_f = (L(f, 1/2)M(f))^4$. Condition (82) follows immediately from Theorem 1.4. As for condition (83): it is known from [GHL] that

$$1/4\pi(f, f) \ll \frac{\log q}{q}.$$

Trivially, $|M(f)| \ll q^{\Delta/4} \log^3(q)$. To bound $|L(f, 1/2)|$ we appeal to Theorem 1.1, so that for all $\varepsilon > 0$

$$\text{Max}_{f \in S_2(q)^*} \frac{|M(f)L(f, 1/2)|^4}{4\pi(f, f)} \ll_\varepsilon q^{-1/48 + \Delta + \varepsilon}.$$

Thus (83) is satisfied so long as $\Delta < 1/48$. Note that the convexity bound $|L(f, 1/2)| \ll_\varepsilon q^{1/4 + \varepsilon}$ would have been insufficient for our purpose.

Thus we are left with estimating the sum

$$\frac{1}{\zeta(2)} \sum_f^h L(f, 1/2)^4 M(f)^4 \omega_f(q^\kappa)$$

We take κ to satisfy $4(\Delta/2) + \kappa < 1/15$. This lets us treat $\omega_f(q^\kappa)$ as a perturbation of the mollifier and use the techniques of section 5. Specifically, we write

$$\Theta(\vec{z}, \ell) := \nu(\vec{z}, \ell) \left(\sum_{\ell = m_1 m_2 m_3 m_4} \frac{\mu(m_1)\mu(m_2)}{m_1^{z_1} m_2^{z_2}} \frac{\mu(m_3)\mu(m_4)}{m_3^{z_3} m_4^{z_4}} \right)$$

where ν is as in (66). Using the Hecke recursion, we rewrite the ‘‘perturbed’’ mollifier $M(f)^4 \omega_f(q^\kappa)$ as

$$\begin{aligned} M(f)^4 \omega_f(q^\kappa) &= \frac{1296}{\log(M)^4} \frac{1}{(2\pi i)^4} \int_{(2)^4} \sum_{m^2 (dn')^{1/2} \leq q^\kappa} \frac{\delta_{dn'=\square}}{dm^2} \sum_\ell \frac{\lambda_f(n'\ell)}{(n'\ell)^{1/2}} \Theta(\vec{z}, d\ell) \\ &\quad \times \eta(\vec{z}) \prod_{i < j} \zeta(1 + z_i + z_j) \frac{M^{z_1 + z_2 + z_3 + z_4} d\vec{z}}{z_1^4 z_2^4 z_3^4 z_4^4}. \end{aligned} \quad (84)$$

To compute $\zeta(2)^{-1} \sum_f^h L(f, 1/2)^4 M(f)^4 \omega_f(q^\kappa)$ we use (84), replacing $\lambda_f(n'\ell)$ by $M^D(n'l) + M^{OD}(n'l) + M^{OOD}(n'l)$. We deal here with the sum of the two first terms (see section 5.2.1), the off-off-diagonal term is similar. We write

$$\Phi(s, t; \ell) := \sum_{abd=\ell} \frac{\mu(a)}{a^{1+s+2t}} \frac{\tau(b)}{b^t} \frac{1}{d^s} \prod_{p^\alpha || ad} \left(\alpha + 1 - \frac{2\alpha p^{-(1+s+t)}}{1 + p^{-(1+s+t)}} \right)$$

for the arithmetic function defined in formula (67). Including $\omega_f(q^\kappa)$ simply involves replacing

$$\sum_\ell \frac{J(s, t; \ell)}{\ell} \Theta(\vec{z}, \ell)$$

in (68) by

$$\begin{aligned} & \sum_{m^2(dn')^{1/2} \leq q^\kappa} \frac{\delta_{dn'=\square}}{dn'm^2} \sum_{\ell} \frac{J(s, t, \ell n')}{\ell} \Theta(\vec{z}, d\ell) \\ &= \left[\sum_{m^2 n \leq q^\kappa} \frac{1}{n^2 m^2} \sum_{dn'=n^2} \Theta(\vec{z}, d) \Phi(s, t; n') \nu(s, t, \vec{z}; d, n') \right] \sum_{\ell} \frac{J(s, t; \ell)}{\ell} \Theta(\vec{z}, \ell) \end{aligned} \quad (85)$$

with

$$\nu(s, t, \vec{z}; d, n') = \prod_{p|dn'} (1 + O(p^{-1/4}))$$

uniformly for $\Re s, \Re t, \Re z_i \geq -1/4$. Since the sum is against n^2 , the first factor and all its partial derivatives of low order are bounded by an absolute constant in the domain of the contour shifts (so this factor brings no new poles). We can remove the constraint $m^2 n \leq q^\kappa$ at the cost of an error term of size $O(q^{-\kappa/2})$, which we can ignore, and convert the first factor into

$$\zeta(2) \sum_n \frac{1}{n^2} \left(\sum_{dn'=n^2} \Theta(\vec{z}, d) \Phi(s, t; n') \nu(s, t, \vec{z}; d, n') \right)$$

(note that the $\zeta(2)$ compensates for the $\zeta(2)^{-1}$ in $1/4\pi(f, f)$). The second factor exactly matches (68), so that it contributes as before. Thus, the main terms for the harmonic and the natural averages are proportional. What remains is to compute the proportionality constant: it will turn out to be 1.

6.1 Computation of the proportionality constant

We proceed exactly as in section 5.4, using formal Euler products. It is enough to calculate

$$\begin{aligned} & \frac{1}{\zeta(2)} \sum_{m, n} \frac{1}{m^2 n^2} \left[\sum_{dn'=n^2} \sum_{\ell} \frac{\Theta(d\ell)}{\ell} \frac{\Phi(n'\ell)}{\zeta(2)} \right] \\ &= \frac{\zeta(2)^2}{\zeta(1)^{10}} \sum_n \frac{1}{n^2} \prod_{p^f | n} \frac{(1-p^{-2})^4}{(1-p^{-1})^{10}} \left[\sum_{i=0}^{2f} \sum_{k \geq 0} \frac{\Theta(p^{k+i}) \Phi(p^{k+2f-i})}{p^k} \right]. \end{aligned}$$

By a formal computation (which we performed using MAPLE), using the table of Θ and (75), one can check that for $f \geq 1$

$$\sum_{i=0}^f \sum_{k \geq 0} \frac{\Theta(p^{k+i}) \Phi(p^{k+f-i})}{p^k} = \sum_{i=0}^{\inf(4, f)} \sum_{k=0}^4 \frac{\Theta(p^{k+i}) \Phi(p^{k+f-i})}{p^k} = 0. \quad (86)$$

Thus the only contribution to the main term comes from $f = 0$ for all primes, that is, from $n = 1$. Thus the natural average is the same as the harmonic average, as desired.

7 Simultaneous non-vanishing

Let χ be a fixed primitive character of conductor D , coprime to q . In this section we use Theorem 1.4 and the techniques of second moment analysis to prove Theorem 1.5.

We borrow from the methods and notation of [KMV]: we let $M := \hat{q}^\Delta$ and $M' := \hat{q}^{\Delta'}$ with $\Delta, \Delta' > 0$, and for P, Q polynomials satisfying $P(0) = Q(0) = Q'(0) = 0$ we define

$$M(f) = M_Q(f) := \sum_{m < M} \frac{\lambda_f(m)\mu(m)}{\psi(m)m^{1/2}} Q\left(\frac{\log M/m}{\log M}\right)$$

with $\psi(m) = \prod_{p|m}(1 + 1/p)$ (we will take $\Delta < 1/48$ and $Q(x) = x^3$, giving the mollifier already considered) and

$$M(f \otimes \chi) = M_P(f \otimes \chi) := \sum_{m < M'} \frac{\chi(m)\lambda_f(m)\mu(m)}{\psi_{\chi^2}(m)m^{1/2}} P\left(\frac{\log M'/m}{\log M'}\right),$$

with $\psi_{\chi^2}(m) := \prod_{p|m}(1 - \chi^2(p)/p)$.

We have

$$\begin{aligned} & \left| \sum_f^* L(f, 1/2)M(f)L(f \otimes \chi, 1/2)M(f \otimes \chi) \right| \\ & \leq \left(\sum_{L(f, 1/2)L(f \otimes \chi, 1/2) \neq 0}^* L(f, 1/2)^2 M(f)^2 \right)^{1/2} \left(\sum_f^* |L(f \otimes \chi, 1/2)M(f \otimes \chi)|^2 \right)^{1/2} \\ & \leq \left(\sum_{L(f, 1/2)L(f \otimes \chi, 1/2) \neq 0}^* 1 \right)^{1/4} \left(\sum_f^* L(f, 1/2)^4 M(f)^4 \right)^{1/4} \left(\sum_f^* |L(f \otimes \chi, 1/2)M(f \otimes \chi)|^2 \right)^{1/2} \end{aligned}$$

so that

$$\sum_{L(f, 1/2)L(f \otimes \chi, 1/2) \neq 0}^* 1 \geq \frac{\left| \sum_f^* L(f, 1/2)M(f)L(f \otimes \chi, 1/2)M(f \otimes \chi) \right|^4}{\left(\sum_f^* L(f, 1/2)^4 M(f)^4 \right) \left(\sum_f^* |L(f \otimes \chi, 1/2)M(f \otimes \chi)|^2 \right)^2}. \quad (87)$$

Let

$$\mathcal{Q}_{f \otimes \chi}(P) := \sum_f^* |L(f \otimes \chi, 1/2)M(f \otimes \chi)|^2$$

and

$$\mathcal{L}_{f, \chi}(P, Q) := \sum_f^* L(f, 1/2)M(f)L(f \otimes \chi, 1/2)M(f \otimes \chi).$$

Theorem 1.5 follows from the evaluation of $\mathcal{Q}_{f \otimes \chi}(P)$ and $\mathcal{L}_{f, \chi}(P, Q)$, since the fourth moment has already been handled in Theorem 1.4. Below we sketch the proof of the following proposition:

Proposition 7.1 *Let P, Q be fixed polynomials with $P(0) = Q(0) = Q'(0) = 0$. For any $0 < \Delta, \Delta' < 1$ we have:*

- for χ^2 non-trivial,

$$\mathcal{Q}_{f \otimes \chi}(P) = 2 \frac{|L(\chi^4, 2)|^2}{|L(\chi^2, 1)|^2} \left(\frac{1}{\Delta'} \int_0^1 P'(x)^2 dx + \frac{1}{2} P(1)^2 \right) (1 + o(1)) \quad (88)$$

$$\mathcal{L}_{f, \chi}(P, Q) = 2\zeta(2) \frac{L(\chi^4, 2)}{L(\chi^2, 1)} \frac{P(1)Q'(1)}{\Delta \log q} (1 + o(1)); \quad (89)$$

- for χ^2 trivial, $\chi(-q) = 1$, $P'(0) = 0$, and $P'(1)Q'(1) \neq 0$,

$$\mathcal{Q}_{f \otimes \chi}(P) = 8 \frac{\zeta(2)^2 \psi(D)^2}{(\log q)^2 \Delta'^2} \left(\frac{1}{\Delta'} \int_0^1 P''(x)^2 dx + P'(1)^2 \right) (1 + o(1)) \quad (90)$$

$$\mathcal{L}_{f, \chi}(P, Q) = 8\zeta(2)^2 \psi(D) \frac{P'(1)Q'(1)}{\Delta \Delta' (\log q)^2} (1 + o(1)). \quad (91)$$

From Proposition 7.1, Theorem 1.4, and (87) we obtain

Proposition 7.2 For $0 < \Delta < 1/48$, $0 < \Delta' < 1$,

- For χ^2 non-trivial

$$\sum_{L(f, 1/2) L(f \otimes \chi, 1/2) \neq 0}^* 1 \geq \frac{4P(1)^4 Q'(1)^4}{\Delta^4 F(\Delta) \left(\Delta'^{-1} \int_0^1 P'(x)^2 dx + \frac{1}{2} P(1)^2 \right)^2} (1 + o(1))$$

- For χ^2 trivial

$$\sum_{L(f, 1/2) L(f \otimes \chi, 1/2) \neq 0}^* 1 \geq \frac{64P'(1)^4 Q'(1)^4}{\Delta^4 F(\Delta) \left(\Delta'^{-1} \int_0^1 P''(x)^2 dx + P'(1)^2 \right)^2} (1 + o(1))$$

where $F(\Delta)$ is defined in (5).

Note in particular that the bounds are independent of the choice of χ , for q sufficiently large. When χ^2 is non-trivial, we can take $Q(x) = x^3$, $P(x) = x$, $\Delta' = 1$, and $\Delta = 1/48$ to get a non-vanishing fraction of

$$6.0 \times 10^{-12}.$$

When χ^2 is trivial, we take $P(x) = x^2$, and get a non-vanishing fraction of

$$2.9 \times 10^{-10}.$$

The proof of Proposition 7.1 follows directly from the techniques of [IS], [KMV], and [V1]. Since there is essentially nothing new to the proofs, we merely touch on the main points. As always, one uses the functional equation for $L(f \otimes \chi_1, s)L(f \otimes \chi_2, s)$ to express the value at $s = 1/2$ as a rapidly converging sum (the requirement that $\chi(-q) = 1$ when χ^2 is trivial arises because otherwise the functional equation implies that $L(f, 1/2)L(f \otimes \chi, 1/2)$ vanishes identically). One then uses the Hecke recursion to separate the diagonal term, which is the only main term for the second moment, from the remainders. One can bound the remainders terms through

large sieve type inequalities, in particular Proposition 1 of [DFI II] and the remark following it, is sufficient to give the $\Delta < 1$ upper limit. This leaves us with the main term, which for $\mathcal{Q}_{f \otimes \chi}(P)$ is

$$2 \sum_{d_1, d_2} \frac{\chi^2(d_1) \bar{\chi}^2(d_2)}{d_1 d_2} \sum_c \frac{1}{c} \sum_n \frac{1}{n} \sum_{(m_1, m_2)=1} \frac{\mu(m_1 d_1 c) \mu(m_2 d_2 c)}{\psi_{\chi^2}(m_1 d_1 c) \psi_{\bar{\chi}^2}(m_2 d_2 c) m_1 m_2} \\ \times P\left(\frac{\log M' / (m_1 d_1 c)}{\log M'}\right) P\left(\frac{\log M' / (m_2 d_2 c)}{\log M'}\right) W\left(\frac{m_1 m_2 d_1 d_2 n^2}{\widehat{q}_D^2}\right), \quad (92)$$

the summation restricted to variables m_1, m_2, d_1, d_2, c, n coprime with D . The main term for $\mathcal{L}_{f \otimes \chi}(P, Q)$ is similar: multiply (92) by $\frac{1}{2} \chi(d_2^2 c n m_1 m_2)$ and replace the second P by Q (and the corresponding M' by M). The \mathcal{L} case also has a second term, involving $\chi(-q)$, but this can be analyzed in much the same way and is smaller than the main term by a factor of $1/\log q$.

To evaluate (92), we proceed exactly as in Section 5 of [KMV] or Section 5 of this paper. We transform W , P , and Q into their integral forms, shift the contours past the origin (where now the end contour, C_γ , depends also on χ), and evaluate the residues. This is far easier for the second moment: the only pole which matters is $z_1 = z_2 = 0$. In particular, $z_1 = -z_2$ contributes negligible terms provided that there are “enough” poles, much as $z_1 = -z_2, z_3 = -z_4$ did for the fourth moment. When the ζ functions are replaced with $L(\chi)$ functions, as is the case here, fewer poles constitute “enough,” which is why we only need $P(0) = 0$. Finally, we calculate the constant by which to multiply the residues. This works much the same way as before, with $L(\chi^4, 2)/L(\chi^2, 1)$ replacing $\zeta(2)/\log M$ (as one would expect) when χ^2 is non-trivial.

8 Appendix: Removing the exponential factor

For many applications (although not in this paper), it is more useful to have Theorem 1.1 with a bound whose integral against $\Gamma(1+\mu)$ is absolutely convergent (recall that the Lindelöf Hypothesis states that the true bound is $(q|\mu|)^\varepsilon$). In particular, one often wants a bound which breaks convexity in q and is polynomial in $|\mu|$, and this can be done through our methods, albeit with more intricate calculations. Instead of the formula we used for $L(f, 1/2 + \mu)^2$, one takes

$$|L(f, 1/2 + \mu)|^2 = \frac{2}{2\pi i} \int_{(2)} G(s + \mu) G(s - \mu) \Gamma(1 + s + \mu) \Gamma(1 + s - \mu) \\ \times L(f, 1/2 + s + \mu) L(f, 1/2 + s - \mu) \hat{q}^{2s} \frac{ds}{s}.$$

The integrand is still even in s , and the pole at $s = 0$ gives the desired value. The advantage of this version is that it is relatively easy to show that the function

$$W(y) = \frac{2}{2\pi i} \int_{(2)} G(s + \mu) G(s - \mu) \Gamma(1 + s + \mu) \Gamma(1 + s - \mu) y^{-s} \frac{ds}{s}$$

is dominated by

$$(1 + |\mu|)^B y^{-C},$$

simply by shifting to $\Re s = C$: since the arguments of the Γ functions differ by 2μ , at least one must be exponentially small in $|\mu|$.

The disadvantage of this method, and the reason we did not use it here, is that one must then replace $\tau(m)$ and $\tau(n)$ everywhere with $\eta_\mu(m)$ and $\eta_\mu(n)$, with

$$\eta_\mu(m) = \sum_{ab=m} \left(\frac{a}{b}\right)^\mu.$$

This in turn means that there will be μ 's throughout the analysis, in almost every Γ function and almost every ζ . In particular, one would have to rederive [DFI I] with η_μ instead of τ . This can be done, since it is all based on Proposition 2.3, which extends to η_μ : provided that f is smooth and vanishes near the origin,

$$\begin{aligned} \sum_{n \geq 1} \eta_\mu(n) e\left(\frac{an}{c}\right) f(n) &= \frac{\zeta(1+2\mu)}{c^{1+2\mu}} \int_0^\infty f(x) x^\mu dx + \frac{\zeta(1-2\mu)}{c^{1-2\mu}} \int_0^\infty f(x) x^{-\mu} dx \\ &+ \frac{1}{c} \sum_{n \geq 1} \eta_\mu(n) e\left(-\frac{\bar{a}n}{c}\right) \int_0^\infty f(x) Y^*\left(\frac{4\pi\sqrt{nx}}{c}\right) dx \\ &+ \frac{1}{c} \sum_{n \geq 1} \eta_\mu(n) e\left(\frac{\bar{a}n}{c}\right) \int_0^\infty f(x) K^*\left(\frac{4\pi\sqrt{nx}}{c}\right) dx. \end{aligned}$$

The Y^* and K^* are closely related to the standard Bessel functions Y_μ and K_μ ; their integral forms are

$$\begin{aligned} Y^*(x) &= \frac{1}{2\pi i} \int_{(*)} (x/2)^{-s} \Gamma(s-\mu) \Gamma(s+\mu) \cos \pi s ds, \\ K^*(x) &= \frac{\cosh \pi |\mu|}{2\pi i} \int_{(2)} (x/2)^{-s} \Gamma(s-\mu) \Gamma(s+\mu) ds. \end{aligned}$$

Note that K^* is polynomially bounded in $|\mu|$ due to Stirling's formula. The reproof of [DFI I] with η_μ instead of τ is relatively straightforward, as is the calculation of the diagonal and off-diagonal terms. The off-off-diagonal terms are treated in the same way, only with many more factors of μ in all terms (nonetheless, everything is still even in s and t). One thus gets the same bounds in q at every step, just without the exponential dependence on μ . Since the exponential bound was sufficient for our application, we chose not to make an already computationally complex argument any worse, and used the τ function instead.

References

- [Br] Brumer A. *The rank of $J_0(N)$* , Asterisque, Vol 228 (1995), 41-68.
- [CF] Conrey B., Farmer D. *Mean values of L -functions and symmetry* Preprint (1999) American Institute of Mathematics. Available at <http://www.aimath.org/>.
- [D] Davenport H., *Multiplicative Number Theory*, Graduate Texts in Mathematics 74, Springer-Verlag, 1980.
- [DI] Deshouillers J.-M., Iwaniec H., *An additive divisor problem*, J. London Math. Soc. (2) 26 (1982), no. 1, 1-14.

- [DFI I] W. Duke , J. Friedlander, H. Iwaniec, *A quadratic divisor problem*, Invent. Math. 115, (1994), 219-239.
- [DFI II] W. Duke , J. Friedlander, H. Iwaniec, *Bounds for automorphic L -functions. II*, Invent. Math. 115, (1994), 209-217.
- [DFI III] W. Duke , J. Friedlander, H. Iwaniec, *Erratum to “Bounds for automorphic L -functions. II”*, Invent. Math. (to appear).
- [E] Estermann, T.: *Über die Darstellung einer Zahl als Differenz von zwei Produkten*, J. Reine Angew. Math. 164, (1931) 173-182.
- [G] Gross, B. Heights and the special values of L -series. Number theory (Montreal, Que., 1985), 115–187, CMS Conf. Proc. 7, Amer. Math. Soc., Providence, RI, 1987.
- [GKZ] Gross, B. Kohlen W. and Zagier, D.: Heegner points and derivatives of L -series II, Math. Ann., 278, (1987), 497-562.
- [GZ] Gross, W. and Zagier, D.: Heegner points and derivatives of L -series. Invent. Math. 84 (1986), no. 2, 225–320.
- [GHL] D. Goldfeld, J. Hoffstein, D. Liemann *An effective zero-free region (Appendix to Coefficients of Maass forms and the Siegel zero)* , Ann. of Math. 140 (1994), 177-181.
- [HB1] Heath-Brown D. R. *The fourth power moment of the Riemann zeta function*, Proc. London Math. Soc. (3) 38 (1979), no. 3, 385–422.
- [HB2] Heath-Brown D. R. *The fourth power mean of Dirichlet’s L -functions*, Analysis 1 (1981), no. 1, 25–32.
- [In] Ingham, A. E. *Mean value theorems in the theory of the Riemann zeta function*, Proc. London Math. Soc., 27 (1926) 273-300.
- [I1] H. Iwaniec *Topics in classical automorphic forms*. Graduate Studies in Mathematics, 17. American Mathematical Society, Providence, RI, 1997.
- [I2] H. Iwaniec : Lecture Notes on Analytic Number Theory (Rutgers University).
- [IS] H. Iwaniec and P. Sarnak: The non-vanishing of central values of automorphic L -functions and Siegel-Landau zeros, Israel Journal of Math. (to appear).
- [IS2] H. Iwaniec and P. Sarnak, *Dirichlet L -functions at the central point*, Number theory in progress, Vol. 2 (Zakopane-Kościełisko, 1997), 941–952, de Gruyter, Berlin, 1999.
- [ISL] H. Iwaniec, P. Sarnak, and W. Z. Luo *Low Lying Zeros of families of L -functions*. Preprint 1998.
- [J] M. Jutila *A method in the theory of exponential sums*, Tata Lect. Notes Math. 80 Bombay (1987).
- [KS] N.M. Katz, P. Sarnak *Zeroes of zeta functions and symmetry*. Bull. Amer. Math. Soc. (N.S.) 36 (1999), no. 1, 1–26.

- [KM1] E. Kowalski and P. Michel *The analytic rank of $J_0(q)$ and zeros of automorphic L -functions*, Duke Math. J. 100, nb 3, 503-542 (1999).
- [KM2] E. Kowalski and P. Michel *A lower bound for the rank of $J_0(q)$* , Acta Arithmetica (to appear).
- [KM3] E. Kowalski and P. Michel: *An explicit upper bound for the rank of $J_0(q)$* , Israel Journal of Math. (to appear).
- [KMV] E. Kowalski, P. Michel and J. M. VanderKam *Non-vanishing of high derivatives of automorphic L -functions at the center of the critical strip*, Journ. Reine Angew. Math. (to appear)
- [MV] P. Michel and J. M. VanderKam *Non-vanishing of high derivatives of Dirichlet L -functions at the center of the critical strip*, Journal of Number Theory (to appear)
- [Mo] Y. Motohashi *The binary additive divisor problem*, Ann. Sci. cole Norm. Sup. (4) 27 (1994), no. 5, 529–572.
- [Ob] F. Oberhettinger *Tables of Mellin Transforms*, Springer Verlag, 1974.
- [PP] A. Perelli et J. Pomykala *Averages over twisted elliptic L -functions*, Acta Arith. 80, No 2, 149-163 (1997).
- [S] A. Selberg *Contributions to the theory of Dirichlet's L -functions*, Skr. utg. av Det Norske Videnskaps-Akademi, Mat. Nat. Kl. (1946) No 3, 1-62.
- [Sh1] G. Shimura *Introduction to the arithmetic of automorphic functions*, Iwanami Shoten and Princeton Univ. Press (1971).
- [Sh2] G. Shimura, *On the periods of modular forms*. Math. Ann. 229 (1977), no. 3, 211–221.
- [So] K. Soundararajan *Non-vanishing of quadratic Dirichlet L -functions at $s = 1/2$* , Annals of math. (to appear).
- [V1] J. VanderKam: *The rank of quotients of $J_0(N)$* , Duke Math. J. 97 (1999), no. 3, 545–577.
- [V2] J. VanderKam: *Linear independance in the homology of $X_0(N)$* , J. London Math. Society, to appear.

P. Michel: michel@darboux.math.univ-montp2.fr
 Université Montpellier II CC 051, 34095 Montpellier Cedex 05, FRANCE.

E. Kowalski: ekowalsk@math.princeton.edu
 Dept. of Math. - Fine Hall, Princeton University, Princeton, NJ 08544-1000, USA.

J. Vanderkam: vanderkm@math.ias.edu
 Dept. of Math, Institute for Advanced Study, Olden Lane, Princeton, NJ 08540, USA.