ON THE ERROR TERM IN A MIXED MOMENT OF L-FUNCTIONS

RIZWANUR KHAN AND ZEYUAN ZHANG

ABSTRACT. There has recently been some interest in optimizing the error term in the asymptotic for the fourth moment of Dirichlet L-functions and a closely related mixed moment of L-functions involving automorphic L-functions twisted by Dirichlet characters. We obtain an improvement for the error term of the latter.

1. INTRODUCTION

Heath-Brown [5] was the first to prove an asymptotic formula for the fourth moment of the Riemann Zeta function on the critical line with a power saving error term. The analogous asymptotic for Dirichlet L-functions required new ideas and came much later. Young [14] proved that

$$(1.1) \qquad \frac{1}{\varphi^*(q)} \sum_{\chi \bmod q}^{\star} |L(\tfrac{1}{2}, \chi)|^4 = P_4(\log q) + O(q^{-\frac{5}{512}+\epsilon}),$$

where q is prime, $\sum^{\star}$ restricts the sum to primitive Dirichlet characters, $\varphi^*(q)$ is the number of primitive Dirichlet characters, and P_4 is a degree four polynomial. Wu [13] obtained a power saving asymptotic for general moduli q with an error term of $O(q^{-\frac{11}{448}})$. Blomer, Fouvry, Kowalski, Michel, and Milićević [1] revisited the problem, and introducing new ideas, they significantly sharpened Young's error term for q prime to $O(q^{-\frac{1}{32}+\epsilon})$. Later Blomer et al. [2] returned to this problem once more and used an idea of Shparlinski and T. Zhang [12] to reduce the error term even further to $O(q^{-\frac{1}{20}+\epsilon})$, and to $O(q^{-\frac{1}{14}+\epsilon})$ assuming the Ramanujan Conjecture.[1]

In [1], Blomer et al. also considered a related problem which is one step greater in difficulty. They established the mixed moment asymptotic

$$(1.2) \qquad \frac{1}{\varphi^*(q)} \sum_{\chi \bmod q}^{\star} L(\tfrac{1}{2}, f \otimes \chi)\overline{L(\tfrac{1}{2}, \chi)}^2 = \frac{L(1,f)^2}{\zeta(2)} + O(q^{-\frac{1}{68}+\epsilon}),$$

for prime q, where f is a fixed (holomorphic or non-holomorphic) Hecke eigenform for $SL_2(\mathbb{Z})$. There has been some interest in improving the error term in this asymptotic as well. Shparlinski [11] improved the error term in (1.2) slightly to $O(q^{-\frac{1}{64}+\epsilon})$. The goal of this paper is to make a more significant improvement to this error term. In fact when f is holomorphic, the quality of our error term is almost as good as the error term in the fourth moment of Dirichlet L-functions. This is achieved by an idea which simplifies previous approaches to the problem in a critical range.

Theorem 1.1. *Let q be an odd prime. Let $\eta_f = \frac{1}{22}$ for f a holomorphic Hecke-cusp form, and $\eta_f = \frac{5}{152}$ for f a Hecke Maass-cusp form, for $SL_2(\mathbb{Z})$. We have*

$$(1.3) \qquad \frac{1}{\varphi^*(q)} \sum_{\chi \bmod q}^{\star} L(\tfrac{1}{2}, f \otimes \chi)\overline{L(\tfrac{1}{2}, \chi)}^2 = \frac{L(1,f)^2}{\zeta(2)} + O(q^{-\eta_f+\epsilon}).$$

[1]Actually the weaker bound $O(q^{-\frac{1}{16}+\epsilon})$ is claimed on the Ramanujan conjecture at the end of section 4 in [2], but this seems to be a calculation mistake. This has also been observed in [15, Theorem 1.1] with $\theta = 0$ and $L = 1$.

We also mention that the more difficult related asymptotic

$$(1.4) \qquad \frac{1}{\varphi^*(q)} \sum_{\chi \bmod q}^{\star} L(\tfrac{1}{2}, f \otimes \chi)\overline{L(\tfrac{1}{2}, g \otimes \chi)} = \begin{cases} \frac{2L(1, f\otimes g)}{\zeta(2)} + O(q^{-\delta}) & \text{if } f \neq g, \\ \frac{2L(1, \mathrm{sym}^2 f)}{\zeta(2)} \log q + O(q^{-\delta}) & \text{if } f = g, \end{cases}$$

where f and g are Hecke eigenforms and $\delta > 0$, has been obtained by Kowalski, Michel, and Sawin [10] for prime q, though no serious attempt has been made to optimize the error term. We refer the reader to [3] for applications of this result.

Acknowledgment. We thank the anonymous referee for their valuable comments.

2. Sketch.

We briefly indicate the main idea behind the improved error term in Theorem 1.1 before proceeding to the proof. After using an approximate functional equation and the orthogonality of characters, the error term is roughly

$$(2.1) \qquad \max_{1 \leq NM \leq q^{2+\epsilon}} \left| \frac{1}{\sqrt{MN}} \sum_{\substack{n \equiv m \bmod q \\ n \sim N, m \sim M \\ n \neq m}} \lambda_f(m) d(n) \right|,$$

where $\lambda_f(m)$ are the Hecke eigenvalues associated to f and $d(n)$ is the divisor function. This sum is treated using different techniques depending on the the sizes of N and M.

Suppose that $N \geq M$. When N/M is bounded by a reasonably small power of q, this is treated by solving a 'shifted convolution problem', entailing the spectral theory of automorphic forms. Now suppose N/M is somewhat large, the most difficult case being $N/M \sim q^{\frac{1}{2}}$, or equivalently $N \sim q^{\frac{3}{2}}$ and $M \sim q^{\frac{1}{2}}$ when $NM \sim q^2$. The congruence in (2.1) is detected using additive characters and then Voronoi summation is applied to the n sum (for this step we need a long n sum, hence the assumption $N \geq M$). Writing $N = N_1 N_2$ for $N_1 \leq N_2$, this leads to a sum of the shape (after opening up the divisor function)

$$(2.2) \qquad \sum_{h_1 \sim q/N_1, h_2 \sim q/N_2, m \sim M} \lambda_f(m) S(m, h_1 h_2, q).$$

An effective way to treat the most difficult case $N/M \sim q^{\frac{1}{2}}$ is to use Fouvry, Kowalksi, and Michel's [4] bounds for bilinear sums of Kloosterman sums, based on their work on algebraic trace functions. To use this, one groups together two of the three variables h_1, h_2, m to form a new variable, which gives a bilinear sum in this new variable and the remaining variable. The bound is non-trivial when one variable is longer than q^ϵ and one is longer than $q^{\frac{1}{2}+\epsilon}$ (the so called Pólya-Vinogradov range). However this strategy fails when we have $N_1 \sim q^{\frac{1}{2}}$ and $N_2 \sim q$, for then $h_1 \sim M \sim q^{\frac{1}{2}}, h_2 \sim 1$, and there is no grouping that satisfies the requirements. This troublesome range was treated by a new bound for bilinear sums of Kloosterman sums obtained by Blomer et al. [1, Theorem 5.1, equation (5.3)]. Later, Shparlinski gave a better bound using his own result on sums of Kloosterman sums [11, Theorem 2.1]. The aforementioned bounds for sums of Kloosterman sums are very general because they can be used for arbitrary coefficients in place of $\lambda_f(m)$. On the other hand, we make gains by using the automorphic nature of these coefficients. We transform the sum of Kloosterman sums using Poisson summation to an additively twisted sum of $\lambda_f(m)$, and then employ a Wilton type bound giving square-root cancelation. This strategy yields a bound which essentially forces N_1 to be close to N_2, thus giving a superior treatment of the critical range $N/M \sim q^{\frac{1}{2}}$ and $N_1 \sim q^{\frac{1}{2}}$ and $N_2 \sim q$ discussed above. Moreover, this approach is simple in that it completely avoids the consideration of bilinear sums of Kloosterman sums.

The case $M > N$ is somewhat similar after the 'switching trick' of Blomer et al. (Voronoi summation) which reduces to sums like (2.1) with $N' \geq M'$ typically, but now without the restriction $N'M' \leq q^{2+\epsilon}$. The lack of this restriction in this case weakens the final bound a bit.

3. Proof of Theorem 1.1

3.1. **Preliminaries.** Throughout, q will denote an odd prime number, and f will denote a holomorphic Hecke cusp form or non-holomorphic Hecke-Maass cusp form for $SL_2(\mathbb{Z})$, with Hecke eigenvalues $\lambda_f(n)$. We will follow the ϵ convention: that is, ϵ will always represent an arbitrarily small positive constant, but not necessarily the same one from one occurrence to another. All implied constants may depend implicitly on ϵ and f. By a 'negligible' quantity, we will mean one which is $O(q^{-A})$ for any $A > 0$, where the implied constant depends on A. Throughout, we will use W, possibly with a subscript, to denote a smooth function compactly supported on the interval $[\frac{1}{2}, 2]$, with derivatives satisfying

$$W^{(j)}(x) \ll (q^\epsilon)^j$$

for any $j \geq 0$, where the implied constant depends on j. Such a weight function may differ from one occurrence to the next even when we use the same notation.

We will write $e(x)$ to denote $e^{2\pi i x}$. We will use the following elementary orthogonality relation:

$$(3.1) \qquad \frac{1}{q} \sum_{a \bmod q} e\left(\frac{an}{q}\right) = \begin{cases} 1 & \text{if } q|n, \\ 0 & \text{if } q \nmid n. \end{cases}$$

We have the average bound [6, Lemma 1]

$$\sum_{n \leq x} |\lambda_f(n)|^2 \ll x^{1+\epsilon}.$$

For individual bounds, we have

$$|\lambda_f(n)| \ll n^{\theta_f + \epsilon},$$

where $\theta_f = \frac{7}{64}$ if f is Maass (see [9]) and $\theta_f = 0$ if f is holomorphic.

For smooth sums of Hecke eigenvalues, we note a simple consequence of Voronoi summation: for $M \geq q^\epsilon$, we have

$$(3.2) \qquad \sum_{m \geq 1} \lambda_f(m) W\left(\frac{m}{M}\right) \ll q^{-A}$$

for any $A > 0$. See for example [1, Lemmas 2.3, 2.4], where we can take $\vartheta = 0$ since the Selberg eigenvalue conjecture is known in level 1.

For (not necessarily smooth) sums of Hecke eigenvalues twisted by additive characters, we have

Lemma 3.1. *Let $\alpha \in \mathbb{R}$. We have*

$$\sum_{n \leq N} \lambda_f(n) e(n\alpha) \ll N^{\frac{1}{2}+\epsilon}.$$

The implied constant is uniform in α, but depends on f.

Proof. See [8, Theorem 5.3], [7, Theorem 8.1]. $\qquad\qquad\square$

3.2. **Some known bounds.** For $N, M \geq 1$, define

$$(3.3)$$

$$E^\pm(M, N) := S_1 + S_2$$

$$= \frac{1}{\sqrt{MN}} \sum_{\substack{m \equiv \pm n \bmod q \\ m \neq n}} \lambda_f(m) d(n) W_1\left(\frac{m}{M}\right) W_2\left(\frac{n}{N}\right) - \frac{1}{q\sqrt{MN}} \sum_{m,n \geq 1} \lambda_f(m) d(n) W_1\left(\frac{m}{M}\right) W_2\left(\frac{n}{N}\right).$$

Note that $E^\pm(M, N)$ is the same as the quantity $E_{f,E}^\pm(M, N) = B_{f,E}^\pm(M, N)$ defined on [1, page 771]. To establish Theorem 1.1, we need to prove

$$(3.4) \qquad \max_{MN \leq q^{2+\epsilon}} |E^\pm(M, N)| \ll q^{-\eta_f + \epsilon}.$$

We begin by recalling some established bounds for $E^\pm(M,N)$, but for later use it will be important to not restrict to $MN \leq q^{2+\epsilon}$.

Lemma 3.2. *Let $MN \leq q^{4+\epsilon}$. We have*

$$(3.5) \qquad E^\pm(M,N) \ll \begin{cases} q^\epsilon \frac{\sqrt{MN}}{q} & \text{if } N \geq M \\ q^\epsilon M^{\theta_f} \frac{\sqrt{MN}}{q} & \text{if } M > N, \end{cases}$$

$$(3.6) \qquad E^\pm(M,N) \ll q^\epsilon \left(1 + \left(\frac{NM}{q^2}\right)^{\frac{1}{4}}\right)\left(\frac{\max(N,M)}{q\min(N,M)}\right)^{\frac{1}{4}}\left(1 + \left(\frac{\max(N,M)}{q\min(N,M)}\right)^{\frac{1}{4}}\right).$$

If $\frac{\max(N,M)}{\min(N,M)} > 10$, then

$$(3.7) \qquad E^\pm(M,N) \ll q^\epsilon\left(\frac{q\min(N,M)}{\max(N,M)}\right)^{\frac{1}{2}} + q^\epsilon M^{\theta_f}\frac{(NM)^{\frac{1}{2}}}{q^2}.$$

Proof. The first bound (3.5) is the trivial bound given in [1, Proposition 3.1].

The second bound is obtained in [1, Theorem 3.2] by solving a shifted convolution problem using the spectral theory of automorphic forms, with an additional idea required to avoid dependence on the Ramanujan conjecture. However, just reading the statement of that theorem is not enough because it assumes $NM \leq q^{2+\epsilon}$. We look at the proof of [1, Theorem 3.2] and find that the bound (3.6) is contained in the bounds given by [1, equations (3.12), (3.16)], after dividing by the $(MN)^{\frac{1}{2}}$ factor in [1, equation (3.4)]. This then gives exactly what is seen in the statement of [1, Theorem 3.2], apart from an extra factor $O(1 + (\frac{NM}{q^2})^{\frac{1}{4}})$ which we retain in (3.6) and which would have been redundant had we been assuming $MN \leq q^{2+\epsilon}$, and a missing $O(q^{-\frac{1}{2}+\theta_f+\epsilon})$ term, which is specific to the case of f and g both cuspidal (see the bottom of [1, page 726]) and therefore does not appear in (3.6).

The third bound (3.7) is obtained in [1, section 6.4] by applying Voronoi summation to both n and m sums and then bounding trivially using Weil's bound for the Kloosterman sum. The assumption $\frac{\max(N,M)}{\min(N,M)} > 10$ implies that the condition $n \neq m$ is vacuous in the sum S_1 in the definition (3.3) of $E^\pm(M,N)$. Then $E^\pm(M,N)$ may be expressed as [1, equation (6.21)], for which we have the bound given in [1, equation (6.23)] plus the trivial bound for the expression in [1, equation (6.22)]. The latter bound gives the term $q^\epsilon M^{\theta_f}\frac{(NM)^{\frac{1}{2}}}{q^2}$ in (3.7) which does not appear in [1, equation (6.23)] because it is assumed there that $MN \leq q^{2+\epsilon}$, while we do not make this assumption. $\qquad\square$

3.3. **A new bound.** Define the sum (closely following the notation in [1, equation (6.26)]) for $M, N_1, N_2 \geq 1$ with $N = N_1 N_2$ and $N_1 \leq N_2$:

$$C^\pm(M,N,N_1,N_2) := \frac{1}{\sqrt{NM}} \sum_{\substack{n_1,n_2,m\geq 1 \\ n_1 n_2 \equiv \pm m \bmod q}} \lambda_f(m) W_1\left(\frac{n_1}{N_1}\right) W_2\left(\frac{n_2}{N_2}\right) W_3\left(\frac{m}{M}\right).$$

We will prove the following simple bound which was not observed in [1] or [11].

Lemma 3.3. *Let $MN < \frac{1}{8}q^3$ and $M \geq q^\epsilon$. We have*

$$(3.8) \qquad C^\pm(M,N,N_1,N_2) \ll q^\epsilon\left(\frac{N_1}{N_2}\right)^{\frac{1}{2}} + q^{-A},$$

for any $A > 0$.

Proof. Detecting the congruence $n_1 n_2 \equiv \pm m \bmod q$ using additive characters, we have

$$C^{\pm}(M, N, N_1, N_2) = \frac{1}{q\sqrt{NM}} \sum_{a \bmod q} \sum_{n_1, n_2, m \geq 1} \lambda_f(m) W_1\left(\frac{n_1}{N_1}\right) W_2\left(\frac{n_2}{N_2}\right) W_3\left(\frac{m}{M}\right) e\left(\frac{a(n_1 n_2 \mp m)}{q}\right)$$

$$= \frac{1}{q\sqrt{NM}} \sum_{\substack{a \bmod q \\ b \bmod q}} \sum_{\substack{n_1, n_2, m \geq 1 \\ n_2 \equiv b \bmod q}} \lambda_f(m) W_1\left(\frac{n_1}{N_1}\right) W_2\left(\frac{n_2}{N_2}\right) W_3\left(\frac{m}{M}\right) e\left(\frac{a(n_1 b \mp m)}{q}\right).$$

Now we apply Poisson summation to the n_2 sum (more precisely, we write $n_2 = \ell q + b$ and apply Poisson summation to the ℓ sum) to get

$$(3.9) \quad C^{\pm}(M, N, N_1, N_2)$$

$$= \frac{1}{q\sqrt{NM}} \frac{N_2}{q} \sum_{\substack{a \bmod q \\ b \bmod q}} \sum_{\substack{n_1, m \geq 1 \\ -\infty < k < \infty}} \lambda_f(m) e\left(\frac{bk}{q}\right) e\left(\frac{a(n_1 b \mp m)}{q}\right) W_1\left(\frac{n_1}{N_1}\right) \widehat{W_2}\left(\frac{kN_2}{q}\right) W_3\left(\frac{m}{M}\right),$$

where $\widehat{W_2}(y) = \int_{\mathbb{R}} W_2(x) e(-yx) dx$ is the Fourier transform of W_2. By repeated integration by parts, we may restrict the k sum to $|k| < \frac{q^{1+\epsilon}}{N_2}$, up to a negligible error.

Consider first the contribution to (3.9) of the terms with $q | n_1$, a condition which implies that $N_1 \geq \frac{1}{2}q$. This contribution is

$$(3.10)$$

$$\frac{1}{q\sqrt{NM}} \frac{N_2}{q} \sum_{\substack{a \bmod q \\ b \bmod q}} \sum_{\substack{n_1, m \geq 1 \\ 0 \leq |k| < \frac{q^{1+\epsilon}}{N_2} \\ q | n_1}} \lambda_f(m) e\left(\frac{bk}{q}\right) e\left(\frac{a(\mp m)}{q}\right) W_1\left(\frac{n_1}{N_1}\right) \widehat{W_2}\left(\frac{kN_2}{q}\right) W_3\left(\frac{m}{M}\right) + O(q^{-A})$$

for any $A > 0$. Using (3.1), we see that the a sum vanishes unless q divides m, in which case we must have $M \geq \frac{1}{2}q$. But since we are assuming $MN < \frac{1}{8}q^3$, we cannot have $M \geq \frac{1}{2}q$ and $N_2 \geq N_1 \geq \frac{1}{2}q$, so it follows that (3.10) is $O(q^{-A})$.

It remains to consider the contribution to (3.9) of the terms with $q \nmid n_1$. Evaluating the b sum using (3.1) enforces $a \equiv k\overline{n_1} \bmod q$, where $\overline{n_1}$ denotes the multiplicative inverse of n_1 modulo q. Thus we have that (3.9) equals

$$\frac{N_2}{q\sqrt{NM}} \sum_{\substack{n_1, m \geq 1 \\ 0 \leq |k| < \frac{q^{1+\epsilon}}{N_2} \\ q \nmid n_1}} \lambda_f(m) e\left(\frac{\mp mk\overline{n_1}}{q}\right) W_1\left(\frac{n_1}{N_1}\right) \widehat{W_2}\left(\frac{kN_2}{q}\right) W_3\left(\frac{m}{M}\right) + O(q^{-A}).$$

The contribution to the sum of the term $k = 0$ is negligible by using (3.2). We estimate the contribution of the terms $k \neq 0$ by bounding the m sum using Lemma 3.1 and the rest trivially, to get

$$\frac{N_2}{q\sqrt{NM}} \frac{q^{1+\epsilon}}{N_2} N_1 \sqrt{M} \ll q^{\epsilon} \left(\frac{N_1}{N_2}\right)^{\frac{1}{2}}.$$

$\square$

In order to use the new bound towards our goal (3.4), we need

Lemma 3.4. *If* $\frac{\max(N,M)}{\min(N,M)} > 10$ *and* $M \geq q^{\epsilon}$, *then*

$$(3.11) \qquad |E^{\pm}(M, N)| \ll \max_{\substack{N = N_1 N_2 \\ N_1 \leq N_2}} |C^{\pm}(M, N, N_1, N_2)| + O(q^{-A}).$$

for any $A > 0$.

Proof. The assumption $M \geq q^\epsilon$ implies by (3.2) that the sum S_2 in the definition (3.3) of $E^\pm(M, N)$ is negligible. The assumption $\frac{\max(N,M)}{\min(N,M)} > 10$ implies that the condition $n \neq m$ is vacuous in the sum S_1 in the definition (3.3) of $E^\pm(M, N)$. After removing this condition in S_1, we may open the divisor function and apply smooth partitions of unity, as explained in [1, section 6.4.1], to obtain (3.11). □

3.4. **The case $M > N$.** Recall our goal (3.4). We need to start with the 'switching trick' of [1].

Lemma 3.5. *Let $MN \leq q^{2+\epsilon}$ and $M > N$. Define*

$$M^* = \frac{q^2}{M}, \quad N^* = \frac{q^2}{N}.$$

We have

$$(3.12) \qquad E^\pm(M, N) \ll q^\epsilon \max_{\substack{1 \leq N' \leq q^\epsilon N^* \\ 1 \leq M' \leq q^\epsilon M^*}} |E^\pm(M', N')| + O(q^{-\frac{1}{6}+\epsilon}).$$

Proof. By applying Voronoi summation to the n and m sums in (3.3), Blomer et al. [1, page 765] showed[2]

$$(3.13) \quad E^\pm(M, N) \ll q^\epsilon \max_{\substack{1 \leq N' \leq q^\epsilon N^* \\ 1 \leq M' \leq q^\epsilon M^*}} \left| \frac{1}{\sqrt{M^* N^*}} \sum_{\substack{m,n \geq 1 \\ m \equiv \pm n \bmod q}} \lambda_f(m) d(n) W_1\left(\frac{m}{M'}\right) W_2\left(\frac{n}{N'}\right) \right|$$

$$+ q^\epsilon \max_{\substack{1 \leq N' \leq q^\epsilon N^* \\ 1 \leq M' \leq q^\epsilon M^*}} \left| \frac{1}{q\sqrt{M^* N^*}} \sum_{m,n \geq 1} \lambda_f(m) d(n) W_1\left(\frac{m}{M'}\right) W_2\left(\frac{n}{N'}\right) \right| + O(q^{-1+\epsilon}),$$

where

$$M^* = \frac{q^2}{M}, \quad N^* = \frac{q^2}{N}.$$

The contribution of the diagonal terms $n = m$ in the first sum on the right hand side of (3.13) is bounded absolutely by

$$q^\epsilon \frac{1}{\sqrt{M^* N^*}} M' \ll q^\epsilon \frac{1}{\sqrt{M^* N^*}} M^* \ll q^\epsilon \sqrt{\frac{M^*}{N^*}} = q^\epsilon \sqrt{\frac{N}{M}}.$$

We can assume this is $O(q^{-\frac{1}{6}+\epsilon})$ or else by (3.6) we have $E^\pm(M, N) \ll q^{-\frac{1}{6}+\epsilon}$. Thus at the cost of an admissible error term, the diagonal terms can be removed. By (3.2), the second sum on the right hand side of (3.13) is negligible unless $M' < q^\epsilon$. Thus this sum is trivially bounded by

$$q^\epsilon \frac{1}{q\sqrt{M^* N^*}} N^* \ll \frac{q^\epsilon}{q} \sqrt{\frac{M}{N}}.$$

We can assume this is $O(q^{-\frac{1}{4}+\epsilon})$ or else by (3.7) we have $E^\pm(M, N) \ll q^{-\frac{1}{4}+\epsilon}$. The lemma follows. □

We will apply Lemma 3.3 and Lemma 3.4 to bound $E^\pm(M', N')$. For this, we need to verify the conditions of these lemmas. Suppose that the condition $\frac{\max(N',M')}{\min(N',M')} > 10$ does not hold. Then by

---

[2]They actually have an additional factor $(M^*/M')^{2\theta_f}$, where $\theta_f = 7/64$, which arises from the use of their Lemma 2.4, but this factor is not needed because we work with Hecke eigenforms of level 1. In this setting, the Selberg eigenvalue conjecture is known, so that we may take $\vartheta = 0$ in their Lemma 2.4.

(3.5) and (3.6), we would have

$$E^{\pm}(M', N') \ll q^{\epsilon} \min\left((M')^{\theta_f}\left(\frac{N'M'}{q^2}\right)^{-\frac{1}{2}}, \left(\frac{N'M'}{q^2}\right)^{\frac{1}{4}} q^{-\frac{1}{4}}\right)$$

$$\ll q^{\epsilon} \min\left(q^{\frac{7}{32}}\left(\frac{N'M'}{q^2}\right)^{-\frac{1}{2}}, \left(\frac{N'M'}{q^2}\right)^{\frac{1}{4}} q^{-\frac{1}{4}}\right) \ll q^{-\frac{3}{32}+\epsilon},$$

and we would be done by (3.12). So we can assume that $\frac{\max(N',M')}{\min(N',M')} > 10$ holds. Now suppose that the condition $M' \geq q^{\epsilon}$ does not hold. Then by (3.5) and (3.7), we would have

$$E^{\pm}(M', N') \ll q^{\epsilon} \min\left(\frac{\sqrt{N'}}{q}, \left(\frac{q}{N'}\right)^{\frac{1}{2}} + \frac{\sqrt{N'}}{q^2}\right) \ll \left(\frac{\sqrt{N'}}{q}, \left(\frac{q}{N'}\right)^{\frac{1}{2}} + q^{-1}\right) \ll q^{-\frac{1}{4}+\epsilon},$$

and we would be done by (3.12). So we can assume that $M' \geq q^{\epsilon}$ holds. Finally suppose that the condition $M'N' < \frac{1}{8}q^3$ does not hold. Then by (3.5), we would have $M^*N^* \gg q^3$, which implies $MN \ll q$. Then by (3.5), we would have $E^{\pm}(M', N') \ll (M')^{\theta_f} q^{-\frac{1}{2}+\epsilon} \ll q^{\frac{7}{32}-\frac{1}{2}+\epsilon}$ and we would be done by (3.12). So we can assume the condition $M'N' < \frac{1}{8}q^3$ as well.

Now we are ready to prove

Lemma 3.6. *We have*

$$\max_{\substack{MN \leq q^{2+\epsilon} \\ M > N}} |E^{\pm}(M, N)| \ll q^{-\eta_f + \epsilon}.$$

Proof. We use Lemma 3.3 and Lemma 3.4 to bound $E^{\pm}(M', N')$, which implies a bound for $E^{\pm}(M, N)$ via Lemma 3.5. We add our bound to the collection of bounds given in [1]. We need to find the worst case (maximum) of the minimum of all these bounds for $MN \leq q^{2+\epsilon}$ and $M > N$. This leads to a linear optimization problem to maximize the minimum of the exponents of q from each bound. Such a problem was solved in [1] by a computer search, and we do the same by simply inserting the exponent of q from our bound for $C^{\pm}(M', N', N_1, N_2)$, given in (3.8), into the existing Mathematica code from [1, section 7.4.2]. The underlined portion below indicates the new addition to this code. This returns an exponent of $-\frac{5}{152}$ for f Maass and $-\frac{1}{22}$ for f holomorphic (in the latter case we also omit the term $7m/64$ from the first entry in the existing code, corresponding to the bound (3.5), since $\theta_f = 0$).

```
In[1]  :=   Maximize[{Min[(m + n)/2 - 1 + 7m/64, Max[(m - n - 1)/2, (m - n - 1)/4],
            (1 + n - m)/2, 7/64(mstar-mprime) + Max[(2 mprime - mstar - nstar + 1)/ 2,
            (2 n1 + 2 mprime - 2 - mstar - nstar)/2], (n1 - n2)/2,
            7/64(mstar-mprime) + Max[(mprime + n2 + 1/2 - nstar - mstar)/ 2,
            (2 mprime + n1 - mstar - nstar)/ 2, (2 n1 + 2 mprime - 2 - mstar - nstar)/2],
            7/64(mstar-mprime) + mprime + n1 - (mstar + nstar)/2,
            7/64(mstar-mprime) + Max[Min[2 n1 - (mstar + nstar)/2,
            2/3 nprime + n1 + 1/2 - 1/6 mprime - n2 - (mstar + nstar)/2],
            (2 mprime - mstar - nstar)/ 2, (2 mprime + 2 n1 - 2 - mstar - nstar)/2,
            2 mprime - n2 - (mstar + nstar)/2]], 0 <= n, n <= m, m + n <= 2,
            nstar == 2 - n, mstar == 2 - m, 0 <= nprime, nprime <= nstar,
            0 <= mprime, mprime <= mstar, n1 + n2 == nprime, 0 <= n1, n1 <= n2},
            {m, n, n1, n2, nprime, nstar, mprime, mstar}]

Out[1] :=   {-5/152, {m → 24/19, n → 15/38, n1 → 61/76, n2 → 61/76, nprime → 61/38, nstar → 61/38,
            mprime → 14/19, mstar → 14/19}}
```

```
In[2] :=    Maximize[{Min[(m + n)/2 - 1, Max[(m - n - 1)/2, (m - n - 1)/4],
            (1 + n - m)/2, 7/64(mstar-mprime) + Max[(2 mprime - mstar - nstar + 1)/ 2,
            (2 n1 + 2 mprime - 2 - mstar - nstar)/2], (n1 - n2)/2,
            7/64(mstar-mprime) + Max[(mprime + n2 + 1/2 - nstar - mstar)/ 2,
            (2 mprime + n1 - mstar - nstar)/ 2, (2 n1 + 2 mprime - 2 - mstar - nstar)/2],
            7/64(mstar-mprime) + mprime + n1 - (mstar + nstar)/2,
            7/64(mstar-mprime) + Max[Min[2 n1 - (mstar + nstar)/2,
            2/3 nprime + n1 + 1/2 - 1/6 mprime - n2 - (mstar + nstar)/2],
            (2 mprime - mstar - nstar)/ 2, (2 mprime + 2 n1 - 2 - mstar - nstar)/2,
            2 mprime - n2 - (mstar + nstar)/2]], 0 <= n, n <= m, m + n <= 2,
            nstar == 2 - n, mstar == 2 - m, 0 <= nprime, nprime <= nstar,
            0 <= mprime, mprime <= mstar, n1 + n2 == nprime, 0 <= n1, n1 <= n2},
            {m, n, n1, n2, nprime, nstar, mprime, mstar}]
```

$$\text{Out[2] := }\quad \left\{-\tfrac{1}{22}, \left\{\texttt{m} \to \tfrac{15}{11},\ \texttt{n} \to \tfrac{6}{11},\ \texttt{n1} \to \tfrac{8}{11},\ \texttt{n2} \to \tfrac{8}{11},\ \texttt{nprime} \to \tfrac{16}{11},\ \texttt{nstar} \to \tfrac{16}{11}, \right.\right.$$
$$\left.\left. \texttt{mprime} \to \tfrac{7}{11},\ \texttt{mstar} \to \tfrac{7}{11}\right\}\right\}$$

$\square$

3.5. The case $N \geq M$. Recall our goal (3.4). Assuming $MN \leq q^{2+\epsilon}$ and $N \geq M$, we will apply Lemma 3.3 and Lemma 3.4 to bound $E^{\pm}(M, N)$. For this purpose, we need to verify the conditions $\frac{\max(N,M)}{\min(N,M)} > 10$, $M \geq q^{\epsilon}$, and $MN < \frac{1}{8}q^3$ of these lemmas. This is done just as in section 3.4 and is in fact easier. This is because the factor M^{θ_f} does not occur when using (3.5) because we assume that $N \geq M$, and the condition $MN < \frac{1}{8}q^3$ is automatic for q large enough by the assumption $MN \leq q^{2+\epsilon}$.

Now we prove the following, which together with Lemma 3.6 gives Theorem 1.1.

Lemma 3.7. *We have*
$$\max_{\substack{MN \leq q^{2+\epsilon} \\ N \geq M}} |E^{\pm}(M, N)| \ll q^{-\frac{1}{20}+\epsilon}.$$

Proof. As before, we simply insert our bound from Lemma 3.3 into the existing Mathematica code, this time from [1, section 7.4.1], and compute.

```
In[3] :=    Maximize[{Min[(m + n)/2 - 1, Max[(n - m - 1)/2, (n - m - 1)/4],
            (1 + m - n)/2, Max[(m + 1 - n)/2, (2 n1 + m - 2 - n)/2],
            Max[1/4 - n1/2, (m - n2)/2, (2 n1 + m - n - 2)/2], (m + 2 n1 - n)/ 2,
            Max[Min[2 n1 - (m + n)/2, n/6 + n1 + 1/2 - n2 - 2 m/3], (m - n)/2,
            (m + 2 n1 - 2 - n)/2, 3 m/2 - n2 - n/2], If[m + n1prime <= 1,
            (2 m + 2 n1prime + 2 n2prime - 1 - n1circ - n2circ - m)/2
            + Max[-n2prime/2, 1/4 - (m + n1prime)/2], 10], (n1 - n2)/2,
            If[m + n2prime <= 2 n1prime, (2 m + 2 n1prime + 2 n2prime - 1
            - n1circ - n2circ - m)/2 + 1/4 - 5 n1prime/12 - (m + n2prime)/6, 10]],
            m >= 0, n >= m, m + n <= 2, n1 + n2 == n, n1 <= n2, n1 >= 0,
            n1prime >= 0, n1prime <= n1circ, n1circ == 1 - n1, n2prime >= 0,
            n2prime <= n2circ, n2circ == 1 - n2}, {m, n, n1, n2, n1prime,
            n2prime, n1circ, n2circ}]
```

$$\text{Out[3] := }\quad \left\{-\tfrac{1}{20}, \left\{\texttt{m} \to \tfrac{3}{5},\ \texttt{n} \to \tfrac{7}{5},\ \texttt{n1} \to \tfrac{7}{10},\ \texttt{n2} \to \tfrac{7}{10},\ \texttt{n1prime} \to \tfrac{3}{10},\ \texttt{n1prime} \to \tfrac{3}{10}, \right.\right.$$
$$\left.\left. \texttt{n1circ} \to \tfrac{3}{10},\ \texttt{n1circ} \to \tfrac{3}{10}\right\}\right\}$$

$\square$

References

1. V. Blomer, É. Fouvry, E. Kowalski, P. Michel, and D. Milićević, *On moments of twisted L-functions*, Amer. J. Math. **139** (2017), no. 3, 707–768.
2. _____, *Some applications of smooth bilinear forms with Kloosterman sums*, Tr. Mat. Inst. Steklova **296** (2017), no. Analiticheskaya i Kombinatornaya Teoriya Chisel, 24–35, English version published in Proc. Steklov Inst. Math. **296** (2017), no. 1, 18–29.
3. V. Blomer, É. Fouvry, E. Kowalski, P. Michel, D. Milićević, and W. Sawin, *The Second Moment Theory of Families of L-Functions: The Case of Twisted Hecke L-Functions*, Mem. Amer. Math. Soc. **282** (2023), no. 1394.
4. É. Fouvry, Emmanuel K., and P. Michel, *Algebraic trace functions over the primes*, Duke Math. J. **163** (2014), no. 9, 1683–1736.
5. D. R. Heath-Brown, *The fourth power moment of the Riemann zeta function*, Proc. London Math. Soc. (3) **38** (1979), no. 3, 385–422.
6. H. Iwaniec, *The spectral growth of automorphic L-functions*, J. Reine Angew. Math. **428** (1992), 139–159.
7. _____, *Introduction to the spectral theory of automorphic forms*, Biblioteca de la Revista Matemática Iberoamericana. [Library of the Revista Matemática Iberoamericana], Revista Matemática Iberoamericana, Madrid, 1995.
8. _____, *Topics in classical automorphic forms*, Graduate Studies in Mathematics, vol. 17, American Mathematical Society, Providence, RI, 1997.
9. H. H. Kim, *Functoriality for the exterior square of* GL_4 *and the symmetric fourth of* GL_2, J. Amer. Math. Soc. **16** (2003), no. 1, 139–183, With appendix 1 by D. Ramakrishnan and appendix 2 by H. Kim and P. Sarnak.
10. E. Kowalski, P. Michel, and W. Sawin, *Bilinear forms with Kloosterman sums and applications*, Ann. of Math. (2) **186** (2017), no. 2, 413–500.
11. I. E. Shparlinski, *On sums of Kloosterman and Gauss sums*, Trans. Amer. Math. Soc. **371** (2019), no. 12, 8679–8697.
12. I. E. Shparlinski and T. Zhang, *Cancellations amongst Kloosterman sums*, Acta Arith. **176** (2016), no. 3, 201–210.
13. X. Wu, *The fourth moment of Dirichlet L-functions at the central value*, Math. Ann. (doi:10.1007/s00208-022-02483-9).
14. M. P. Young, *The fourth moment of Dirichlet L-functions*, Ann. of Math. (2) **173** (2011), no. 1, 1–50.
15. R. Zacharias, *Mollification of the fourth moment of Dirichlet L-functions*, Acta Arith. **191** (2019), no. 3, 201–257.

Department of Mathematics, University of Mississippi, University, MS 38677
Email address: rrkhan@olemiss.edu, zzhang8@go.olemiss.edu