Subconvex bounds on GL3 via degeneration to frequency zero

Math. Ann. (2018) 372:299–319
https://doi.org/10.1007/s00208-018-1711-y

Mathematische Annalen

Subconvex bounds on GL_3 via degeneration to frequency zero

Roman Holowinsky[1] · **Paul D. Nelson**[2]

Received: 7 February 2018 / Revised: 31 May 2018 / Published online: 13 June 2018

Abstract For a fixed cusp form π on $GL_3(\mathbb{Z})$ and a varying Dirichlet character χ of prime conductor q, we prove that the subconvex bound

$$L\left(\pi \otimes \chi, \tfrac{1}{2}\right) \ll q^{3/4-\delta}$$

holds for any $\delta < 1/36$. This improves upon the earlier bounds $\delta < 1/1612$ and $\delta < 1/308$ obtained by Munshi using his GL_2 variant of the δ-method. The method developed here is more direct. We first express χ as the degenerate zero-frequency contribution of a carefully chosen summation formula à la Poisson. After an elementary "amplification" step exploiting the multiplicativity of χ, we then apply a sequence of standard manipulations (reciprocity, Voronoi, Cauchy–Schwarz and the Weil bound) to bound the contributions of the nonzero frequencies and of the dual side of that formula.

Mathematics Subject Classification 11F66 · 11M41

Communicated by Kannan Soundararajan.

✉ Paul D. Nelson
paul.nelson@math.ethz.ch

Roman Holowinsky
holowinsky.1@osu.edu

1 Department of Mathematics, The Ohio State University, 100 Math Tower, 231 West 18th Ave, Columbus, OH 43210, USA

2 Department of Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland

Contents

1 Introduction

We consider the problem of bounding $L(\pi \otimes \chi, \frac{1}{2})$, where

- π is a fixed cusp form on $\mathrm{GL}_3(\mathbb{Z})$, not necessarily self-dual, and
- χ traverses a sequence of Dirichlet characters χ of (say) prime conductor q tending off to ∞.

Munshi [21] recently established the first subconvex bound in this setting by showing that if π satisfies the Ramanujan–Selberg conjecture, then for any fixed $\delta < 1/1612$, the estimate

$$|L\left(\pi \otimes \chi, \tfrac{1}{2}\right)| \leqslant Cq^{3/4-\delta} \tag{1.1}$$

holds for some positive quantity C that may depend upon δ and π, but not upon χ. In the preprint [17], he improves the exponent range to $\delta < 1/308$ and removes the Ramanujan–Selberg assumption.

A striking feature of his work is the introduction of a novel "GL_2 δ-symbol method," whereby one detects an equality of integers $n_1 = n_2$ by averaging several instances of the Petersson trace formula. We summarize this approach in Appendix B, referring to [21] and [17] for details, to [19] and [18] for other recent applications of the GL_2 δ-symbol method, and to [10, Sect. 5.5] for general discussion of the spectral decomposition of the δ-symbol.

It is natural to ask about the true strength of the GL_2 δ-symbol method. How does it compare to the classical δ-symbol method of Duke–Friedlander–Iwaniec [6] and Heath-Brown [8]? For which problems does one fail and the other succeed? For which problems are the two methods "identical" or "equivalent"? Can the GL_2 δ-symbol method be simplified or removed in certain applications?

In pondering such questions, we were able to better understand the arithmetical structure and mechanisms underlying Munshi's argument and construct a more direct proof of the following quantitative strengthening of Munshi's bound.

Theorem 1 *The subconvex bound* (1.1) *holds for any* $\delta < \delta_0 := 1/36$.

The proof is surprisingly short compared to earlier proofs of related estimates. Indeed, we regard the primary novelty of this work as not in the numerical improvement of the exponent δ but rather in the drastic simplification obtained for the proof of any subconvex bound (1.1).

Our point of departure is a formula (see Sect. 3.2), derived via Poisson summation, that expresses χ in terms of additive characters and twisted Kloosterman sums. We insert this into an approximate functional equation for $L(\pi \otimes \chi, 1/2)$. After an elementary "amplification" step exploiting the multiplicativity of χ, we then conclude via standard manipulations. We discuss in Appendix B how we arrived at this approach through a careful study of Munshi's arguments.

We hope that the technique described here may be applied to many other problems. For instance, it seems natural to ask whether it allows a simplification or generalization of the arguments of [19] for bounding symmetric square L-functions.

The works [3,9,12,14,20,22–26] bound twisted L-functions on GL_3 in other aspects. In the preprint [13], Yongxiao Lin has generalized our method to incorporate the t-aspect. The preprint [1] applies a simpler technique to the corresponding problem for GL_2.

2 Preliminaries

2.1 Asymptotic notation

We work throughout this article with a cusp form π on $\mathrm{GL}_3(\mathbb{Z})$ and a sequence of primitive Dirichlet characters χ_{j} to prime moduli q_{j}, indexed by $\mathrm{j} \in \mathbb{Z}_{\geqslant 1}$, with $q_{\mathrm{j}} \rightarrow \infty$. To simplify notation, we drop the subscripts and write simply $\chi := \chi_{\mathrm{j}}$ and $q := q_{\mathrm{j}}$. Our convention is that any object (number, set, function,...) considered below may depend implicitly upon j unless we designate it as *fixed*; it must then be independent of j. Thus π is understood as fixed, while χ is not. All assertions are to be understood as holding after possibly passing to some subsequence $q_{\mathbf{j}_k}$ of the original sequence $q_{\mathbf{j}}$, and in particular, for j sufficiently large.

We define standard asymptotic notation accordingly: $A = \mathrm{O}(B)$ or $A \ll B$ or $B \gg A$ means that $|A| \leqslant c|B|$ for some fixed $c \geqslant 0$, while $A = o(B)$ means $|A| \leqslant c|B|$ for every fixed $c > 0$ (for j large enough, by convention). We write $A \asymp B$ for $A \ll B \ll A$. We write $A = \mathrm{O}(q^{-\infty})$ to denote that $A = \mathrm{O}(q^{-c})$ for each fixed $c \geqslant 0$. Less standardly, we write $A \prec B$ or $B \succ A$ as shorthand for $A \ll q^{o(1)} B$, or equivalently, $|A| \leqslant q^{o(1)}|B|$. Our goal is then to show that

$$L\left(\pi \otimes \chi, \tfrac{1}{2}\right) \prec q^{3/4-\delta_0}. \tag{2.1}$$

We say that $V \in C_c^\infty(\mathbb{R}_+^\times)$ is *inert* if it satisfies the support condition

$$V(x) \neq 0 \implies x \asymp 1$$

and the value and derivative bounds

$$(x\partial_x)^j V(x) \prec 1 \text{ for each fixed } j \geqslant 0.$$

2.2 General notation

We write $e(x) := e^{2\pi i x}$, and denote by $\sum_n$ a sum over integers n. Let $c \in \mathbb{Z}_{\geqslant 1}$. We write $\sum_{a(c)}$ and $\sum_{a(c)^*}$ to denote sums over $a \in \mathbb{Z}/c$ and $a \in (\mathbb{Z}/c)^*$, respectively. We denote the inverse of $x \in (\mathbb{Z}/c)^*$ by x^{-1} or $1/x$. We denote by $e_c : \mathbb{Z}/c \to \mathbb{C}^\times$ the additive character given by $e_c(a) := e^{2\pi i a/c}$, by $S(a, b; c) := \sum_{x(c)^*} e_c(ax + bx^{-1})$ the Kloosterman sum, by $K_c(a) := c^{-1/2} S(a, 1; c)$ the normalized Kloosterman sum, by $S_\chi(a, b; q) := \sum_{x(q)^*} \chi(x) e_q(ax + bx^{-1})$ the twisted Kloosterman sum, and by $\varepsilon(\overline{\chi}) := q^{-1/2} \sum_{a(q)^*} \overline{\chi}(a) e_q(a)$ the normalized Gauss sum (of magnitude one).

We define the Fourier coefficients $\lambda(m, n)$ of π as in [7], so that $L(\pi \otimes \chi, s) = \sum_{n \in \mathbb{Z}_{\geqslant 1}} \lambda(1, n) \chi(n) n^{-s}$ for complex numbers s with large enough real part, and $\lambda(n, m) = \overline{\lambda(m, n)}$.

For a condition C, we define 1_C to be 1 if C holds and 0 otherwise. For instance, $1_{a=b}$ is 1 if $a = b$ and 0 if $a \neq b$.

We denote by $\hat{V}(\xi) := \int_{x \in \mathbb{R}} V(x) e(-\xi x)\, dx$ the Fourier transform of a Schwartz function V on $\mathbb{R}$.

For a pair of integers a, b, we denote by (a, b) and $[a, b]$ their the greatest common divisor and least common multiple, respectively.

2.3 Voronoi summation formula

By [15] (cf. [2, Sect. 4] for the formulation used here), we have for $V \in C_c^\infty(\mathbb{R}_+^\times)$, $m, c \in \mathbb{Z}_{\geqslant 1}$, $a \in (\mathbb{Z}/c)^*$, and $X > 0$ that

$$\sum_n V\left(\frac{n}{X}\right) \lambda(m, n) e_c(an) = c \sum_{\substack{\pm, n \\ d | cm}} \mathcal{I}_\pm V\left(\frac{nd^2}{c^3 m / X}\right) \frac{\lambda(n, d)}{nd} S\left(\frac{m}{a}, \pm n; \frac{mc}{d}\right) \tag{2.2}$$

for integral transforms $V \mapsto \mathcal{I}_\pm V \in C^\infty(\mathbb{R}_+^\times)$ of the shape

$$\mathcal{I}_\pm V(x) = \int_{\mathrm{Re}(s)=1} x^{-s} \mathcal{G}^\pm(s+1) \left(\int_{y \in \mathbb{R}_+^\times} V(y) y^{-s} \frac{dy}{y} \right) \frac{ds}{2\pi i},$$

where $\mathcal{G}^\pm$ is meromorphic on $\mathbb{C}$ and holomorphic in the domain $\mathrm{Re}(s) > 5/14$, where it satisfies $\mathcal{G}^\pm(s) \ll (1 + |s|)^{\mathrm{O}(1)}$ for fixed $\mathrm{Re}(s)$. (The indices n and d in (2.2) are implicitly restricted to be positive integers.) Set $\theta = 5/14 + \varepsilon$ for some sufficiently small fixed $\varepsilon > 0$. By shifting the contour to $\mathrm{Re}(s) = \theta - 1$ and to $\mathrm{Re}(s) = A$, we see that if V is inert, then

$$(x \partial_x)^j \mathcal{I}_\pm V(x) \ll \min(x^{1-\theta}, x^{-A}) \tag{2.3}$$

for all fixed $j, A \geqslant 0$.

In the special case $m = 1$, we have $S(m/a, \pm n; mc/d) = (c/d)^{1/2} K_{c/d}(\pm n/a)$, and so

$$\sum_n V\left(\frac{n}{X}\right)\lambda(1,n)e_c(an) = c^{3/2}\sum_{\substack{\pm,n\\ d|c}} \mathcal{I}_\pm V\left(\frac{nd^2}{c^3/X}\right)\frac{\lambda(n,d)}{nd^{3/2}} K_{c/d}\left(\frac{\pm n}{a}\right).$$

2.4 Rankin–Selberg bounds

By [16], we have for each fixed $\varepsilon > 0$ and all $X \geqslant 1$ that

$$\sum_{n\leqslant X}|\lambda(n,1)|^2 = \sum_{n\leqslant X}|\lambda(1,n)|^2 \leqslant \sum_{m,n:m^2n\leqslant X}|\lambda(m,n)|^2 \ll X^{1+\varepsilon}. \tag{2.4}$$

Using the Hecke relations as in the proof of [17, Lem 2], we deduce that for all $M, N \geqslant 1$,

$$\sum_{m\leqslant M, n\leqslant N}|\lambda(m,n)|^2 \ll (MN)^{1+\varepsilon}. \tag{2.5}$$

(Indeed, we may reduce to considering the dyadic sums over $M/2 < m \leqslant M$, $N/2 < n \leqslant N$, and then to establishing that $\sum_{X/8<m^2n\leqslant X} m|\lambda(m,n)|^2 \ll X^{1+\varepsilon}$, which is shown in *loc. cit.*)

3 Division of the proof

3.1 Approximate functional equation

Recall our main goal (2.1). By [11, Sect. 5.2], we may write

$$L(\pi\otimes\chi,\tfrac{1}{2}) = \sum_n \frac{\lambda(1,n)\chi(n)}{\sqrt{n}} V_1\left(\frac{n}{q^{3/2}}\right) + \eta\sum_n \frac{\overline{\lambda(1,n)\chi(n)}}{\sqrt{n}} V_2\left(\frac{n}{q^{3/2}}\right),$$

for some $\eta \in \mathbb{C}$ with $|\eta| = 1$ and some smooth functions $V_1, V_2 : \mathbb{R}_+^\times \to \mathbb{C}$ satisfying $(x\partial_x)^j V_i(x) \ll \min(1, x^{-A})$ for all fixed $j, A \in \mathbb{Z}_{\geqslant 0}$. By a smooth dyadic partition of unity and the Rankin–Selberg estimate (2.4), it will suffice to show for each $0 < N \prec q^{3/2}$ and each inert $V \in C_c^\infty(\mathbb{R}_+^\times)$ that the normalized sum

$$\Sigma := \sum_n \frac{V(n/N)}{N}\lambda(1,n)\chi(n)$$

satisfies the estimate

$$\Sigma \prec N^{-1/2}q^{3/4-\delta_0}. \tag{3.1}$$

By further application of (2.4), we may and shall assume further that

$$q^{3/2-2\delta_0} \leqslant N \prec q^{3/2}. \tag{3.2}$$

The proof of (3.1) will involve positive parameters R, S, T satisfying

$$q^{\varepsilon} \ll R, S, T \ll q^{1-\varepsilon} \text{ for some fixed } \varepsilon > 0. \tag{3.3}$$

Thus every integer in $[R, 2R] \cup [S, 2S] \cup [T, 2T]$ is coprime to q.

3.2 A formula for χ

Fix a smooth function W on $\mathbb{R}$ supported in the interval $[1, 2]$ with $\int W(x)\,dx = 1$. Then $\hat{W}(0) = 1$. Observe that $1/r \in \mathbb{Z}/q$ is defined for all integers r for which $W(r/R) \neq 0$. Set

$$H := q/R.$$

By Poisson summation, we have

$$\frac{\sqrt{q}}{R} \sum_r W\left(\frac{r}{R}\right) \chi(r) e_q\left(\frac{u}{r}\right) = \sum_h \hat{W}\left(\frac{h}{H}\right) \frac{1}{\sqrt{q}} \underbrace{\sum_{r(q)^*} \chi(r) e_q\left(\frac{u}{r}\right) e_q(hr)}_{=S_\chi(h,u;q)}. \tag{3.4}$$

For $h \equiv 0 \pmod q$, we have $S_\chi(h, u; q) = \sqrt{q}\varepsilon(\overline{\chi})\chi(u)$. Setting $\alpha_r := \varepsilon(\overline{\chi})^{-1} R^{-1} W(r/R)\chi(r)$, we deduce by rearranging (3.4) that

$$\chi(u) = q^{1/2} \sum_r \alpha_r e_q\left(\frac{u}{r}\right) - \varepsilon(\overline{\chi})^{-1} \sum_{h \neq 0} \hat{W}\left(\frac{h}{H}\right) \frac{S_\chi(h, u; q)}{\sqrt{q}}. \tag{3.5}$$

The properties of the sequence α to be used in what follows are that it is supported on $[R, 2R]$ and satisfies the estimates $\alpha_r \prec R^{-1}$ and $\sum_r |\alpha_r| \asymp 1$.

3.3 "Amplification"

We choose sequences of complex numbers β_s and γ_t supported on (say) primes in the intervals $[S, 2S]$ and $[T, 2T]$, respectively, so that

$$\beta_s \prec S^{-1}, \quad \gamma_t \prec T^{-1}, \quad \sum_s \beta_s \overline{\chi}(s) = \sum_t \gamma_t \chi(t) = 1. \tag{3.6}$$

Then

$$\Sigma = \sum_{n,s,t} \frac{V(n/N)}{N} \lambda(1,n) \beta_s \gamma_t \chi\left(\frac{tn}{s}\right). \tag{3.7}$$

The properties of β_s and γ_t just enunciated, rather than an explicit choice, are all that will be used; one could take, for instance $\beta_s := \chi(s)|\mathcal{P} \cap [S, 2S]|^{-1} 1_{s \in \mathcal{P} \cap [S, 2S]}$, where $\mathcal{P}$ denotes the set of primes, and similarly for γ_t.

3.4 A formula for Σ

Substituting (3.5) with $u = tn/s$ into (3.7) gives $\Sigma = \mathcal{F} - \varepsilon(\overline{\chi})^{-1}\mathcal{O}$, where

$$\mathcal{F} = q^{1/2} \sum_{r,s,t} \alpha_r \beta_s \gamma_t \sum_n \frac{V(n/N)}{N} \lambda(1,n) e_q\left(\frac{tn}{rs}\right),$$

$$\mathcal{O} = \sum_n \frac{V(n/N)}{N} \lambda(1,n) \sum_{s,t} \beta_s \gamma_t \sum_{h \neq 0} \hat{W}\left(\frac{h}{H}\right) \frac{S_\chi(h, tn/s; q)}{\sqrt{q}}.$$

3.5 Main estimates

We prove these in the next two sections.

Proposition 1 *Assume that*

$$qRS \succ TN. \tag{3.8}$$

Then

$$|\mathcal{F}|^2 \prec \frac{q}{N}\left(\frac{qRS}{TN}\right)^3 + q\frac{(RS)^3}{N^2}\left(\frac{1}{ST} + \frac{1 + N/R^2S}{R^{1/2}S}\right). \tag{3.9}$$

Remark As explained in the remark of Sect. 4.3, the first term on the RHS of (3.9) is unnecessary. Including it simplifies slightly our proofs without affecting our final estimates.

Proposition 2 *Assume that*

$$ST \leqslant q^{-\varepsilon} R \tag{3.10}$$

for some fixed $\varepsilon > 0$. Then

$$|\mathcal{O}|^2 \prec H^2 \frac{1}{STH}. \tag{3.11}$$

3.6 Optimization

Our goal reduces to establishing that $\mathcal{F}, \mathcal{O} \prec N^{-1/2} q^{3/4-\delta_0}$. (By comparison, we note the trivial bounds $\mathcal{F} \prec q^{1/2}$ and $\mathcal{O} \prec H$.) We achieve this by applying the above estimates with

$$R := \frac{TN}{qS}, \quad S := q^{2/18}, \quad T := q^{5/18}.$$

Then (3.8) is clear, while (3.10) follows from (3.2). The required bound for $\mathcal{O}$ follows readily from (3.11). We now deduce the required bound for $\mathcal{F}$. Note that the first term on the RHS of (3.9) is acceptable thanks to our choice of R. Note also from (3.2) that $qST \leqslant N \prec q^{3/2}$; from our choice of R, it follows that $1/ST \gg (1+N/R^2 S)/R^{1/2} S$. The bound for $|\mathcal{F}|^2$ then readily simplifies to $|\mathcal{F}|^2 \prec q^{-2\delta_0} \prec N^{-1} q^{3/2-2\delta_0}$. (By solving a linear programming problem, we see moreover that these choices give the optimal bound for $L(\pi \otimes \chi, 1/2)$ derivable from the above propositions.)

4 Estimates for $\mathcal{F}$

We now prove Proposition 1.

4.1 Reciprocity

Our assumption (3.8) implies that for all r, s, t with $\alpha_r \beta_s \gamma_t \neq 0$, the function $V'_{r,s,t}(x) := V(x) e(tNx/qrs)$ is inert. By the Chinese remainder theorem, we have $e_q(tn/rs) = e_{qrs}(tn) e_{rs}(-tn/q)$ for $(rs, q) = 1$. We may thus rewrite

$$\mathcal{F} = \sum_{r,s,t} \alpha_r \beta_s \gamma_t \mathcal{S}(r, s, t),$$

where

$$\mathcal{S}(r, s, t) := q^{1/2} \sum_n \frac{V'_{r,s,t}(n/N)}{N} \lambda(1, n) e_{rs}\left(-\frac{tn}{q}\right).$$

4.2 Voronoi

We introduce the notation

$$c := c(rs, t) := \frac{rs}{(rs, t)}, \quad a := a(rs, t) := \frac{-t}{(rs, t)},$$

so that $e_{rs}(-tn/q) = e_c(an/q)$ and $(a, c) = 1$. Applying Voronoi summation (Sect. 2.3), we obtain

$$\mathcal{S}(r,s,t) = \frac{q^{1/2}c^{3/2}}{N} \sum_{\substack{\pm,n \\ d|c}} V''_{\pm,r,s,t}\left(\frac{nd^2}{c^3/N}\right) \frac{\lambda(n,d)}{nd^{3/2}} K_{c/d}\left(\frac{\pm qn}{a}\right)$$

for some smooth functions $V''_{\pm,r,s,t}$ satisfying $(x\partial_x)^j V''_{\pm,r,s,t}(x) \prec \min(x^{1-\theta}, x^{-A})$ for fixed $j, A \in \mathbb{Z}_{\geqslant 0}$.

4.3 Cleaning up

The Weil bound, the Rankin–Selberg bound (2.4) and the condition $N \prec q^{3/2}$ give

$$\mathcal{S}(r,s,t) \prec q^{1/2}c^{3/2}/N \prec N^{-1/2}q^{1/2}(qc/N)^{3/2}. \tag{4.1}$$

If $(rs,t) \neq 1$, then (because t is prime) $c = rs/t$, hence by (4.1),

$$\sum_{r,s,t:(rs,t)\neq 1} \alpha_r \beta_s \gamma_t \mathcal{S}(r,s,t) \prec N^{-1/2}q^{1/2}\left(\frac{qRS}{TN}\right)^{3/2}.$$

Since the square of the latter is the first term on the RHS of (3.9), the proof of Proposition 1 reduces to that of an adequate bound for the sum

$$\mathcal{F}_1 := \sum_{r,s,t:(rs,t)=1} \alpha_r \beta_s \gamma_t \mathcal{S}(r,s,t).$$

If $(rs,t) = 1$, then $c = rs$ and $a = -t$, hence

$$\mathcal{F}_1 = \sum_{\pm,n,r,d} \alpha_r \frac{\lambda(n,d)}{\sqrt{nd}} \sum_{\substack{s,t: \\ d|rs,(rs,t)=1}} \beta_s \gamma_t \Phi(n,d,r,s,t) \tag{4.2}$$

with

$$\Phi(n,d,r,s,t) := \frac{q^{1/2}(rs)^{3/2}}{N\sqrt{nd}} V''_{\pm,r,s,t}\left(\frac{nd^2}{(rs)^3/N}\right) K_{rs/d}\left(\frac{\mp qn}{t}\right).$$

Remark With slightly more case-by-case analysis in the arguments to follow, one can verify that the reduction performed here to the case $(rs,t) = 1$ is unnecessary, hence that the bound (3.9) remains valid in the stated generality even after deleting the first term on its RHS.

4.4 Cauchy–Schwarz

Let $\varepsilon > 0$ be fixed and small. The rapid decay of $V''_{\pm,r,s,t}$ implies that truncating (4.2) to $nd^2 \leqslant q^{\varepsilon}(RS)^3/N$ introduces the negligible error $\mathrm{O}(q^{-\infty})$. By the Rankin–Selberg bound (2.5), we have

$$\sum_{\substack{\pm,n,r,d:\\ nd^2 \leqslant q^{\varepsilon}(RS)^3/N}} |\alpha_r| \frac{|\lambda(n,d)|^2}{nd} \prec 1.$$

It follows by Cauchy–Schwarz that

$$|\mathcal{F}_1|^2 \prec \sum_{\pm,n,r,d} |\alpha_r| \left| \sum_{\substack{s,t:\\ d|rs,(rs,t)=1}} \beta_s \gamma_t \Phi(n,d,r,s,t) \right|^2 + \mathrm{O}(q^{-\infty}).$$

4.5 Application of exponential sum bounds

Opening the square, expanding the definition of Φ and wastefully discarding some summation conditions, we obtain

$$|\mathcal{F}_1|^2 \prec \frac{q(RS)^3}{N^2} \sum_{\substack{\pm,r,d,s_1,s_2,t_1,t_2:\\ d|(rs_1,rs_2)}} \frac{|\alpha_r \beta_{s_1} \beta_{s_2} \gamma_{t_1} \gamma_{t_2}|}{d^2} |\mathcal{C}| + \mathrm{O}(q^{-\infty}), \tag{4.3}$$

where $\mathcal{C}$ is defined for (r, s_1, s_2, t_1, t_2) in the support of $\alpha_r \beta_{s_1} \beta_{s_2} \gamma_{t_1} \gamma_{t_2}$ by

$$\mathcal{C} := \frac{1}{X} \sum_n U\left(\frac{n}{X}\right) K_{rs_1/d}\left(\frac{\mp qn}{t_1}\right) \overline{K_{rs_2/d}\left(\frac{\mp qn}{t_2}\right)} \tag{4.4}$$

with

$$X := \frac{(rs_1)^{3/2}(rs_2)^{3/2}}{d^2 N} \asymp \frac{(RS)^3}{d^2 N}$$

and

$$U(x) := \frac{1}{x} V''_{\pm,r,s_1,t_1}\left(\frac{Xxd^2}{(rs_1)^3/N}\right) \overline{V''_{\pm,r,s_2,t_2}\left(\frac{Xxd^2}{(rs_2)^3/N}\right)}.$$

We have $(x\partial_x)^j U(x) \prec \min(x^{1-2\theta}, x^{-A})$ for fixed $j, A \in \mathbb{Z}_{\geqslant 0}$. By a smooth dyadic partition of unity, we may write

$$U(x) = \sum_{Y \in \exp(\mathbb{Z})} \min(Y^{1-2\theta}, Y^{-10}) U_Y\left(\frac{x}{Y}\right), \tag{4.5}$$

where each function U_Y is inert. Substituting (4.5) into (4.4) and applying the incomplete exponential sum estimates recorded in Appendix A, we obtain with

$$\Delta := q\frac{(rs_2/d)^2 t_2 - (rs_1/d)^2 t_1}{(rs_1/d, rs_2/d)^2} = q\frac{s_2^2 t_2 - s_1^2 t_1}{(s_1,s_2)^2}$$

that

$$\mathcal{C} \prec \frac{1}{X} \sum_{Y \in \exp(\mathbb{Z})} \min\left(Y^{1-2\theta}, Y^{-10}\right) \left(XY \frac{\left(\Delta, \frac{rs_1}{d}, \frac{rs_2}{d}\right)^{1/2}}{\left[\frac{rs_1}{d}, \frac{rs_2}{d}\right]^{1/2}} + \left[\frac{rs_1}{d}, \frac{rs_2}{d}\right]^{1/2} \right).$$

Since $\theta < 1/2$, the above sum is dominated by the contribution from $Y = 1$; estimating that contribution a bit crudely with respect to d, we obtain

$$\mathcal{C} \prec d^{1/2} \frac{(\Delta, rs_1, rs_2)^{1/2}}{r^{1/2}[s_1, s_2]^{1/2}} + \frac{r^{1/2}[s_1, s_2]^{1/2}}{d^{1/2} X}. \tag{4.6}$$

4.6 Diagonal and off-diagonal

To state the estimates to be obtained shortly, we introduce the notation

$$\mathbb{E}_{r,s_1,s_2,t_1,t_2} := \frac{1}{RS^2T^2} \sum_{\substack{r: \\ R \leqslant s \leqslant 2R}} \sum_{\substack{s_1,s_2: \\ S \leqslant s_1,s_2 \leqslant 2S}} \sum_{\substack{t_1,t_2: \\ T \leqslant t_1,t_2 \leqslant 2T}} .$$

We estimate separately the contribution of each term on the RHS of (4.6) to $\mathcal{F}_1$ via (4.3), splitting off the contribution to the first from terms with $\Delta = 0$. We obtain in this way that

$$|\mathcal{F}_1|^2 \prec \frac{q(RS)^3}{N^2} \sum_{i=0,1,2} \mathcal{B}_i + \mathrm{O}(q^{-\infty}),$$

where

$$\mathcal{B}_0 := \mathbb{E}_{r,s_1,s_2,t_1,t_2} 1_{\Delta=0} \frac{(s_1, s_2)^{1/2}}{[s_1, s_2]^{1/2}},$$

$$\mathcal{B}_1 := \mathbb{E}_{r,s_1,s_2,t_1,t_2} 1_{\Delta\neq 0} \frac{(\Delta, rs_1, rs_2)^{1/2}}{r^{1/2}[s_1, s_2]^{1/2}},$$

$$\mathcal{B}_2 := \frac{N}{(RS)^3} \mathbb{E}_{r,s_1,s_2,t_1,t_2} r^{1/2}[s_1, s_2]^{1/2}.$$

(In deriving the estimate involving $\mathcal{B}_2$, we used the slightly wasteful bound $\frac{1}{d^2}\frac{1}{d^{1/2}X} \ll \frac{N}{(RS)^3}$.) Noting that $\Delta = 0$ iff $s_2^2 t_2 = s_1^2 t_1$, we verify using the divisor bound that

$$\mathcal{B}_0 \prec \frac{1}{ST},$$

$$\mathcal{B}_1 \prec \frac{1}{R^{1/2}S},$$

$$\mathcal{B}_2 \prec \frac{N}{(RS)^3} R^{1/2} S.$$

These estimates combine to give an adequate estimate for $\mathcal{F}_1$.

5 Estimates for $\mathcal{O}$

We now prove Proposition 2.

5.1 Cauchy–Schwarz

Using again the Rankin–Selberg bound (2.4), we obtain

$$|\mathcal{O}|^2 \prec \sum_n \frac{|V(n/N)|^2}{N} \left| \sum_{s,t,h:h\neq 0} \beta_s \gamma_t \hat{W}\left(\frac{h}{H}\right) \frac{S_\chi(h, tn/s; q)}{\sqrt{q}} \right|^2 .$$

5.2 Elementary exponential sum bounds

Let $\varepsilon > 0$ be fixed but sufficiently small. Since q is prime and R satisfies the lower bound in (3.3), we know that the integers h and q are coprime whenever $0 \neq |h| \leqslant q^\varepsilon H$. By the rapid decay of $\hat{W}$, we may truncate the h-sum to $|h| \leqslant q^\varepsilon H$ with negligible error $\mathrm{O}(q^{-\infty})$. We then open the square and apply Cauchy–Schwarz, leading us to consider for $s_1, t_1, h_1, s_2, t_2, h_2$ with

$$S \leqslant s_i \leqslant 2S, \quad T \leqslant t_i \leqslant 2T, \quad 0 \neq |h_i| \leqslant q^\varepsilon H \tag{5.1}$$

the sums

$$\Pi := \sum_n \frac{|V(n/N)|^2}{N} \frac{S_\chi(h_1, t_1 n/s_1; q)}{\sqrt{q}} \overline{\frac{S_\chi(h_2, t_2 n/s_2; q)}{\sqrt{q}}}. \tag{5.2}$$

We apply Poisson summation. By the lower bound on N in (3.2) and the assumption $\delta_0 = 1/36 < 1/4$, we have $N \gg q^{1+\varepsilon}$ for some fixed $\varepsilon > 0$. Thus only the zero frequency $\xi = 0$ after Poisson contributes non-negligibly, and so $\Pi \prec q^{-1}\Pi_0 + \mathrm{O}(q^{-\infty})$ with

$$\Pi_0 := \sum_{n(q)} \frac{S_\chi(h_1, t_1 n/s_1; q)}{\sqrt{q}} \overline{\frac{S_\chi(h_2, t_2 n/s_2; q)}{\sqrt{q}}}.$$

Opening the Kloosterman sums and executing the n-sum gives

$$\Pi_0 = \sum_{x,y(q)^*} 1_{t_1/s_1 x = t_2/s_2 y} \chi(x/y) e_q(h_1 x - h_2 y).$$

Our assumptions imply that the quantities s_i, t_i, h_i are all coprime to q, so after a change of variables we arrive at

$$|\Pi_0| = |\sum_{x(q)^*} e_q((s_1t_2h_1 - s_2t_1h_2)x)| \leqslant (t_1s_2h_2 - t_2s_1h_1, q).$$

5.3 Diagonal vs. off-diagonal

We have shown thus far that

$$|\mathcal{O}|^2 \prec H^2 \frac{1}{(STH)^2} \sum_{s_1,t_1,h_1,s_2,t_2,h_2} q^{-1}(t_1s_2h_2 - t_2s_1h_1, q) + \mathrm{O}(q^{-\infty}),$$

where the sum is restricted by the condition (5.1). By our assumption (3.10), the quantities $t_1s_2h_2$ and $t_2s_1h_1$ are congruent modulo q precisely when they are equal. By the divisor bound, the number of tuples for which $t_1s_2h_2 = t_2s_1h_1$ is $\prec q^{2\varepsilon}STH$. Since $\varepsilon > 0$ was arbitrary, we obtain

$$|\mathcal{O}|^2 \prec H^2 \left(\frac{1}{STH} + \frac{1}{q} \right). \tag{5.3}$$

By another application of our assumption (3.10), the first term in the latter bound dominates, giving the required bound for $\mathcal{O}$.

The proof of our main result (Theorem 1) is now complete.

Acknowledgements This work was initiated during a visit of PN to RH at The Ohio State University in June 2017. RH thanks PN for taking the time to schedule that visit to Columbus on his return to ETH from MSRI. RH also thanks the Department of Mathematics at The Ohio State University for giving him the opportunity to teach a topics course on R. Munshi's delta method during the Fall 2016 semester. PN thanks The Ohio State University, the STEAM Factory, and the Erdős Institute for their hospitality. We thank Ritabrata Munshi for encouragement and Yongxiao Lin for helpful corrections and feedback on an earlier draft. We thank the referee for helpful feedback which has led to corrections and improvements to the exposition.

Appendix A: Correlations of Kloosterman sums

The estimates recorded here are unsurprising, but we were unable to find references containing all cases that we require (compare with e.g. [4,5,17]).

Lemma 1 *Let s be a natural number. Let $a, b, c, d \in \mathbb{Z}/s$ be congruence classes for which $(d, s) = 1$. For each prime $p \mid s$, let $\mathcal{X}_0(p) \subseteq \mathbb{Z}/p$ be a subset of cardinality $p - \mathrm{O}(1)$. Let $\mathcal{X}$ denote the set of elements $x \in \mathbb{Z}/s$ for which*

- *the class of x modulo p belongs to $\mathcal{X}_0(p)$ for each $p \mid s$, and*
- *$(cx + d, s) = 1$.*

Define $\phi : \mathcal{X} \to \mathbb{Z}/s$ *by*

$$\phi(x) := x \frac{ax+b}{cx+d}.$$

Then the exponential sum $\Sigma := s^{-1} \sum_{x \in \mathcal{X}} e_s(\phi(x))$ *satisfies*

$$|\Sigma| \leqslant 2^{\mathrm{O}(\omega(s))} \frac{(a,b,s)}{s^{1/2}(a,s)^{1/2}},$$

where $\omega(s)$ *denotes the number of prime divisors of* s*, without multiplicity.*

Proof We may assume that $s = p^n$ for some prime p. For $n = 0$, there is nothing to show. For $n = 1$, we appeal either to the Weil bound, to bounds for Ramanujan sums, or to the trivial bound according as $(a, p) = 1$, or $(a, p) = p$ and $(a, b, p) = 1$, or $(a, b, p) = p$. We treat the remaining cases by induction on $n \geqslant 2$.

If $(a, b, p) > 1$, then the conclusion follows by our inductive hypothesis applied to $s/p, a/p, b/p, c, d$. We may thus assume that $(a, b, p) = 1$.

A short calculation gives the identities of rational functions

$$\phi'(x) = \frac{acx^2 + 2adx + bd}{(cx+d)^2}, \quad \phi''(x) = \frac{2a + 2c\phi'(x)}{cx+d}. \tag{A.1}$$

Write $n = 2\alpha$ or $2\alpha + 1$, and set $\mathcal{R} := \{x \in \mathcal{X}/p^\alpha : \phi'(x) \equiv 0 \;\; (p^\alpha)\}$. Then by p-adic stationary phase [11, Sect. 12.3],

$$\Sigma \ll s^{-1/2} \sum_{x \in \mathcal{R}} (\phi''(x), p)^{1/2}.$$

If $(a, p) > 1$, then $(b, p) = 1$ and $\phi'(x) \equiv bd/(cx+d)^2 \;\; (p)$, so $(\phi'(x), p) = 1$. Thus $\mathcal{R} = \emptyset$ and $\Sigma = 0$. Assume otherwise that $(a, p) = 1$. For $x \in \mathcal{R}$, we have $\phi''(x) \equiv 2a/(cx+d) \;\; (p^\alpha)$, so that

$$x \in \mathcal{R} \implies (\phi''(x), p) = (2a, p^\alpha) = (2, p^\alpha) \ll 1. \tag{A.2}$$

Thus $\Sigma \ll s^{-1/2} \# \mathcal{R}$ and, by Hensel's lemma, $\# \mathcal{R} \ll 1$. The proof of the required bound is then complete. □

Lemma 2 *Let* s_1, s_2 *be natural numbers. Let* a_1, a_2, b_1, b_2 *be integers with* $(b_1, s_1) = (b_2, s_2) = 1$*. Set* $\ell_i := a_i/b_i \in \mathbb{Z}/s_i$*. Set*

$$\Delta := \frac{s_2^2 b_2 a_1 - s_1^2 b_1 a_2}{(s_1, s_2)^2}.$$

(i) Let ξ be an integer. Set

$$\Sigma := \frac{1}{[s_1, s_2]} \sum_{x([s_1,s_2])} K_{s_1}(\ell_1 x)\overline{K_{s_2}(\ell_2 x)} e_{[s_1,s_2]}(\xi x)$$

Then

$$|\Sigma| \leqslant 2^{\mathrm{O}(\omega([s_1,s_2]))} \frac{(\Delta, \xi, s_1, s_2)}{[s_1, s_2]^{1/2}(\xi, s_1, s_2)^{1/2}}. \tag{A.3}$$

In particular,

$$|\Sigma| \leqslant 2^{\mathrm{O}(\omega([s_1,s_2]))} \frac{(\Delta, \xi, s_1, s_2)^{1/2}}{[s_1, s_2]^{1/2}}. \tag{A.4}$$

(ii) Let $V : \mathbb{R} \to \mathbb{C}$ be a smooth function satisfying $x^m \partial_x^n V(x) \prec 1$ for all fixed $m, n \in \mathbb{Z}_{\geqslant 0}$. Let $X > 0$.

Assume that $s_1, s_2 = \mathrm{O}(q^{\mathrm{O}(1)})$. Then

$$\sum_n V\left(\frac{n}{X}\right) K_{s_1}(\ell_1 n)\overline{K_{s_2}(\ell_2 n)} \prec X \frac{(\Delta, s_1, s_2)^{1/2}}{[s_1, s_2]^{1/2}} + [s_1, s_2]^{1/2}. \tag{A.5}$$

Remark These estimates are not sharp if either (a_1, s_1) or (a_2, s_2) is large, but that case is unimportant for us. In fact, we have recorded (A.3) only for completeness; the slightly weaker bound (A.4) is the relevant one for our applications. We note finally that K_s is real-valued.

Proof We begin with (i). Each side of (A.3) factors naturally as a product over primes, so we may assume that $s_i = p^{n_i}$ for some prime p. By the change of variables $x \mapsto b_1 b_2 x$, we may reduce further to the case $b_1 = b_2 = 1$, so that $\ell_i = a_i$.

In the case that some ℓ_i is divisible by p, the quantity $K_{s_i}(\ell_i x)$ is independent of x, has magnitude at most $s_i^{-1/2}$, and vanishes if $n_i > 1$. The required estimate then follows in the stronger form $\Sigma \ll (s_1 s_2)^{-1/2}$ by opening the other Kloosterman sum and executing the sum over x. We will thus assume henceforth that ℓ_1 and ℓ_2 are coprime to p.

Write $w_i := s_i/(s_1, s_2)$, so that $w_1 s_2 = s_1 w_2 = [s_1, s_2]$ and $\Delta = w_2^2 b_2 a_1 - w_1^2 b_1 a_2$. By opening the Kloosterman sums and summing over x, we obtain

$$\Sigma = \frac{1}{\sqrt{s_1 s_2}} \sum_{x_1(s_1)^*} \sum_{\substack{x_2(s_2)^* \\ w_1 \ell_2 x_2^{-1} \equiv w_2 \ell_1 x_1^{-1} + \xi \ ([s_1,s_2])}} e_{[s_1,s_2]}(w_2 x_1 - w_1 x_2). \tag{A.6}$$

Consider first the case $s_1 = s_2 =: s$, so that $w_1 = w_2 = 1$ and $\Delta = \ell_1 - \ell_2$ and $[s_1, s_2] = (s_1, s_2) = s$. The subscripted identity in (A.6) then shows that x_2 is determined uniquely by $x_1 =: x$ and, after a short calculation, that

$$\Sigma = \frac{1}{s} \sum_{x(s)^*} e_s \left(x \frac{\xi x + \Delta x}{\xi x + \ell_1} \right).$$

By the previous lemma, it follows that

$$\Sigma \ll \frac{(\Delta, \xi, s)}{s^{1/2}(\xi, s)^{1/2}},$$

as required.

Suppose now that $s_1 \neq s_2$. Without loss of generality, $s_1 < s_2$. Then $w_1 = 1$ and $w_2 = s_2/s_1$; in particular, w_2 is divisible by p. The summation condition in (A.6) shows that $\Sigma = 0$ unless $(\xi, p) = 1$, as we henceforth assume. Since $(\ell_1 \ell_2, p) = 1$, we have $(\Delta, p) = 1$, so our goal is to show that $\Sigma \ll s_2^{-1/2}$. We introduce the variable

$$y := \xi x_1 + w_2 \ell_1.$$

Then

$$x_1 = \frac{y - w_2 \ell_1}{\xi}, \quad x_2 = \frac{w_1 \ell_2}{y} \frac{y - w_2 \ell_1}{\xi},$$

and as y runs over $(\mathbb{Z}/s_1)^*$, the pair (x_1, x_2) traverses the set indicated in (A.6). A short calculation gives

$$w_2 x_1 - w_1 x_2 = -\frac{\Delta}{\xi} + \frac{w_2}{\xi} \left(y + \frac{\ell_2 \ell_1}{y} \right),$$

hence

$$\Sigma = \frac{1}{\sqrt{s_1 s_2}} e_{s_2} \left(-\frac{\Delta}{\xi} \right) \underbrace{\sum_{y(s_1)^*} e_{s_1} \left(\frac{1}{\xi} \left(y + \frac{\ell_2 \ell_1}{y} \right) \right)}_{\sqrt{s_1} K_{s_1}(\ell_2 \ell_1/\xi^2)}.$$

The required conclusion then follows from the Weil bound.

To prove (ii), we first apply Poisson summation to write the LHS of (A.5) as

$$X \sum_{\xi} \hat{V} \left(\frac{\xi}{[s_1, s_2]/X} \right) \frac{1}{[s_1, s_2]} \sum_{x([s_1, s_2])} K_{s_1}(\ell_1 x) \overline{K_{s_2}(\ell_2 x)} e_{[s_1, s_2]}(\xi x), \quad \text{(A.7)}$$

where $\hat{V}$ satisfies estimates analogous to those assumed for V. We then apply (A.4). The $\xi = 0$ term in (A.7) then contributes the first term on the RHS of (A.5), while an adequate estimate for the remaining terms follows from the consequence

$$\sum_{\xi \neq 0} |\hat{V}| \left(\frac{\xi}{[s_1, s_2]/X} \right) (\Delta, \xi, s_1, s_2)^{1/2} \prec [s_1, s_2]/X$$

of the divisor bound. □

Appendix B: Comparison with Munshi's approach

We outline Munshi's approach [17,21] to the sums Σ arising as in Sect. 3.1 after a standard application of the approximate functional equation, and compare with our own treatment. For simplicity we focus on the most difficult range $N \approx q^{3/2}$.

B.1. Averaged Petersson formula

Munshi employs the following decomposition of the diagonal symbol:

$$\delta(m, n) = \frac{1}{B^\star} \sum_{b \in \mathcal{B}} \sum_{\psi(b)} (1 - \psi(-1)) \sum_{f \in S_k(b,\psi)} w_f{}^{-1} \overline{\lambda_f(m)} \lambda_f(n)$$
$$- 2\pi i^{-k} \frac{1}{B^\star} \sum_{b \in \mathcal{B}} \sum_{\psi(b)} (1 - \psi(-1)) \sum_{c \equiv 0(b)} \frac{S_\psi(m, n, c)}{c} J_{k-1} \left(\frac{4\pi \sqrt{mn}}{c} \right). \tag{B.1}$$

Here $\mathcal{B}$ is a suitable set of natural numbers, ψ runs over a suitable collection of odd Dirichlet characters modulo $b \in \mathcal{B}$, and $B^\star$ denotes the appropriate normalizing factor.

B.2 Munshi's initial transformations

Set $A(n) := \lambda(1, n)$. Munshi writes[1]

$$\sum_{n \sim N} A(n) \chi(n) \approx \frac{1}{S} \sum_{s \sim S} \sum_{n \sim N} A(n) \sum_{r \sim NS} \chi\left(\frac{r}{s}\right) \delta(r, ns) \tag{B.2}$$

where s runs over primes of size S. Munshi applies (B.1) to $\delta(r, n\ell)$ with $\mathcal{B} = \{tq : t \sim T\}$, where t runs over primes of size T, and the characters ψ are taken to be trivial modulo q. The use of (B.1) produces two main contributing terms, $\mathcal{F}^M$ from the sum of Fourier coefficients and $\mathcal{O}^M$ from the sum of Kloosterman sums, given roughly by

$$\mathcal{F}^M \approx \frac{1}{T^2 S} \sum_s \sum_t \sum_{\psi(t)} \sum_{n \sim N} \sum_{r \sim NS} A(n) \chi\left(\frac{r}{s}\right) \sum_{f \in S_k(tq,\psi)} \omega_f{}^{-1} \overline{\lambda_f(r)} \lambda_f(ns) \tag{B.3}$$

[1] For the sake of comparison, we note that Munshi used the notation R, L, P, M corresponding to our R, S, T, q.

and

$$\mathcal{O}^M \approx \frac{1}{T^2 S} \sum_s \sum_t \sum_{\psi(t)} \sum_{n \sim N} \sum_{r \sim NS} A(n) \chi\left(\frac{r}{s}\right) \sum_{c \ll \sqrt{q} S/T} \frac{1}{ctq} S_\psi(r, ns; ctq) \quad \text{(B.4)}$$

which Munshi then works to balance with the appropriate choices of S and T. (The superscripted M has been included to disambiguate from the closely related expressions defined in Sect. 3.4 of this paper.) In (B.4) we sum over moduli c up to the transition range of the resulting J-Bessel function, which we do not display for notational simplicity. (For the analogous problem in spectral or t-aspects, the J-Bessel function plays an important analytic role; cf. forthcoming work of Yongxiao Lin.)

B.3 Outline of Munshi's method

We now present a brief outline of Munshi's treatment of $\mathcal{F}^M$ and $\mathcal{O}^M$ (see [17] for details).

B.3.1 Treatment of $\mathcal{F}^M$

(1) Dualize the n-sum via the $\mathrm{GL}_3 \times \mathrm{GL}_2$ functional equation.
(2) Dualize the r-sum via the $\mathrm{GL}_2 \times \mathrm{GL}_1$ functional equation.
(3) Sum over f via the Petersson trace formula. The diagonal contribution is negligible. The off-diagonal contribution is a c-sum over Kloosterman sums of the form $S_\psi(t^2 qn, rs; ctq)$ with $c \ll \sqrt{q} T^2$.
(4) Factor the Kloosterman sums modulo t and modulo cq. This yields Gauss sums modulo t; evaluate them. Sum over ψ modulo t. Factor the remaining Kloosterman sum modulo c and modulo q. The mod q contribution gives a Ramanujan sum equal to -1.
(5) The n-sum now oscillates only modulo c. Apply GL_3 Voronoi and reciprocity.
(6) Dualize the c-sum modulo r via Poisson. Only the zero dual frequency contributes. It remains to estimate sums of the form

$$\frac{\sqrt{q}}{T^4} \sum_{t \sim T} \sum_{s \sim S} \sum_{r \sim \sqrt{q} T/S} \sum_{n \sim T^3} A(n) \overline{\chi}\left(\frac{rs}{t}\right) S\left(-\frac{nq}{t}, 1; rs\right). \quad \text{(B.5)}$$

(7) Pull the n, r sums outside and apply Cauchy-Schwarz.
(8) Conclude via Poisson in n.

Such a treatment produces the following bound

$$\mathcal{F} \ll N \left[\frac{T}{q^{1/4} S^{1/2}} + \left(\frac{TS}{q^{1/2}}\right)^{1/4} + \mathrm{noise}_{\mathcal{F}} \right], \quad \text{(B.6)}$$

where $\text{noise}_{\mathcal{F}}$ comes from all of the other technical aspects resulting from working outside of the transition ranges and appropriately setting up the remaining object for each step of the above proof.

B.3.2. Treatment of $\mathcal{O}^M$

(1) Factor the Kloosterman sums modulo t and cq. Evaluate the sum over ψ; this simplifies the Kloosterman sums modulo t to additive characters. Apply reciprocity. One now has oscillations only modulo cq.
(2) Apply Poisson to the r sum. Only the zero frequency contributes non-negligibly to the dual sum. One is now left with estimating sums of the form

$$\frac{1}{TS\sqrt{q}}\sum_{t\sim T}\sum_{s\sim S}\sum_{c\ll\sqrt{q}S/T}\sum_{n\sim N}A(n)\chi\left(\frac{tc}{s}\right)\mathcal{D}\left(\frac{ns}{tc};q\right), \tag{B.7}$$

where

$$\mathcal{D}(u;q) := \sum_{\substack{b(q)\\(b(b-1),q)=1}}\overline{\chi}(b-1)e_q((b^{-1}-1)u) \tag{B.8}$$

(3) Apply Cauchy–Schwarz with the n-sum outside.
(4) Conclude via Poisson in n.

Such a treatment produces the following bound

$$\mathcal{O}\ll N\left[\frac{q^{1/4}}{T}+\frac{S}{T}+\text{noise}_{\mathcal{O}}\right] \tag{B.9}$$

where $\text{noise}_{\mathcal{O}}$ comes from all of the other technical aspects resulting from working outside of the transition ranges and appropriately setting up the remaining object for each step of the above proof.

B.3.3. Optimization

Ignoring the contributions from $\text{noise}_{\mathcal{F}}$ and $\text{noise}_{\mathcal{O}}$ in (B.6) and (B.9), one first restricts $S<q^{1/4}$, sets

$$\frac{T}{q^{1/4}S^{1/2}}=\left(\frac{TS}{q^{1/2}}\right)^{1/4}$$

to get that $S=Tq^{-1/6}$, and then sets

$$\frac{T}{q^{1/4}S^{1/2}}=\frac{q^{1/4}}{T}$$

to get that $T = q^{5/18}$ and $S = q^{2/18}$ which would produce a combined bound of

$$\sum_{n \sim N} A(n)\chi(n) \ll N\left[q^{-1/36} + \text{noise}_{\mathcal{F}+\mathcal{O}}\right]. \tag{B.10}$$

Therefore, the best possible bound that one could hope to achieve is a saving over the convexity bound of size $q^{-1/36}$. However, due to all of the technical obstacles that present themselves in the course of the proof, Munshi's original approach [21] produced a saving of $q^{-1/1612}$, improved in the preprint [17] to $q^{-1/308}$.

B.4 Discovering the key identity (3.5)

After a topics course taught by the first author in the Fall of 2016 and subsequent discussions with the second author in June 2017, the key identity in this paper was discovered hidden within Munshi's work. Indeed, starting from (B.5) in the treatment of $\mathcal{F}^M$, if one were to now apply Voronoi summation in the n sum followed by an application of reciprocity for the resulting additive characters, then one would need to instead analyze sums of the form

$$\frac{1}{T^2}\sum_{t\sim T}\sum_{s\sim S}\sum_{r\sim\sqrt{q}T/S}\sum_{n\sim q^{3/2}} \overline{A}(n)\overline{\chi}\left(\frac{rs}{t}\right)e_q\left(-\frac{nt}{rs}\right). \tag{B.11}$$

Viewing $-t/rs$ as the u in (3.5), we see that an application of Poisson summation in r returns us to the dual of our original object of interest (from the $h = 0$ frequency of the dual) plus a sum which is the "GL_3 dual" of $\mathcal{O}^M$ (from the dual non-zero h frequencies) as expressed in (B.7)

$$\frac{1}{TS\sqrt{q}}\sum_{t\sim T}\sum_{s\sim S}\sum_{h\ll\sqrt{q}S/T}\sum_{n\sim q^{3/2}} \overline{A}(n)S_{\overline{\chi}}\left(\frac{ht}{s}, n, q\right). \tag{B.12}$$

By "GL_3 dual," we mean that Voronoi summation in n applied to (B.12) returns one to objects of the form (B.7). This observation led to the simplification presented in this paper whereby many of the initial steps of Munshi's argument, as outlined above, are eliminated.

References

1. Aggarwal, K., Holowinsky, R., Lin, Y., Sun, Q.: The Burgess bound via a trivial delta method. ArXiv e-prints (2018)
2. Blomer, V., Khan, R.: Twisted moments of L-functions and spectral reciprocity. ArXiv e-prints (2017)
3. Blomer, V.: Subconvexity for twisted L-functions on GL(3). Am. J. Math. **134**(5), 1385–1421 (2012)
4. Blomer, V., Milićević, D.: p-adic analytic twists and strong subconvexity. Ann. Sci. Éc. Norm. Supér. (4) **48**(3), 561–605 (2015)
5. Castryck, W., Fouvry, É., Harcos, G., Kowalski, E., Michel, P., Nelson, P., Paldi, E., Pintz, J., Sutherland, A.V., Tao, T., Xie, X.-F.: New equidistribution estimates of Zhang type. Algebra Number Theory **8**(9), 2067–2199 (2014)

6. Duke, W., Friedlander, J., Iwaniec, H.: Bounds for automorphic L-functions. Invent. Math. **112**(1), 1–8 (1993)
7. Goldfeld, D.: Automorphic forms and L-functions for the group GL(n, R), volume 99 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2015. With an appendix by Kevin A. Broughan, Paperback edition of the (2006) original [MR2254662]
8. Heath-Brown, D.R.: A new form of the circle method, and its application to quadratic forms. J. Reine Angew. Math. **481**, 149–206 (1996)
9. Huang, B.: Hybrid subconvexity bounds for twisted L-functions on $GL(3)$. ArXiv e-prints (2016)
10. Iwaniec, H.: Topics in classical automorphic forms, volume 17 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI (1997)
11. Iwaniec, H., Kowalski, E.: Analytic number theory, volume 53 of American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI (2004)
12. Li, Xiaoqing: Bounds for GL(3) $\times$ GL(2) L-functions and GL(3) L-functions. Ann. of Math. (2) **173**(1), 301–336 (2011)
13. Lin, Y.: Bounds for twists of GL(3) L-functions. ArXiv e-prints (2018)
14. McKee, Mark, Sun, Haiwei, Ye, Yangbo: Improved subconvexity bounds for $GL(2) \times GL(3)$ and $GL(3)$ L-functions by weighted stationary phase. Trans. Am. Math. Soc. **370**(5), 3745–3769 (2018)
15. Miller, Stephen D., Schmid, Wilfried: Automorphic distributions, L-functions, and Voronoi summation for GL(3). Ann. Math. (2) **164**(2), 423–488 (2006)
16. Molteni, Giuseppe: Upper and lower bounds at $s = 1$ for certain Dirichlet series with Euler product. Duke Math. J. **111**(1), 133–158 (2002)
17. Munshi, R.:. Twists of $GL(3)$ L-functions. ArXiv e-prints (2016)
18. Munshi, R.: A note on Burgess bound. ArXiv e-prints (2017)
19. Munshi, R.: Subconvexity for symmetric square L-functions. ArXiv e-prints (2017)
20. Munshi, R.: The circle method and bounds for L-functions–III: t-aspect subconvexity for $GL(3)$ L-functions. J. Am. Math. Soc. **28**(4), 913–938 (2015)
21. Munshi, R.: The circle method and bounds for L-functions–IV: subconvexity for twists of GL(3) L-functions. Ann. Math. (2) **182**(2), 617–672 (2015)
22. Munshi, R.: The circle method and bounds for L-functions, II: subconvexity for twists of GL(3) L-functions. Am. J. Math. **137**(3), 791–812 (2015)
23. Nunes, R.M.: Subconvexity for GL(3) L-functions. ArXiv e-prints (2017)
24. Qi, Z.: Subconvexity for twisted L-functions on GL_3 over the Gaussian number field. ArXiv e-prints (2018)
25. Sun, Q.: Hybrid bounds for twists of $GL(3)$ L-functions. ArXiv e-prints (2017)
26. Sun, Q., Zhao, R.: Bounds for GL_3 L-functions in depth aspect. ArXiv e-prints (2018)