BILINEAR FORMS WITH GL_3 KLOOSTERMAN SUMS AND THE SPECTRAL LARGE SIEVE

MATTHEW P. YOUNG

Abstract. We analyze certain bilinear forms involving GL_3 Kloosterman sums. As an application, we obtain an improved estimate for the GL_3 spectral large sieve inequality.

1. Introduction

Given a family of L-functions, $\{L(s,f) : f \in \mathcal{F}\}$, one of the most basic questions one can study is its orthogonality properties. More precisely, if $L(s,f) = \sum_{n=1}^{\infty} \lambda_f(n) n^{-s}$, then one wishes to understand $\Delta_{\mathcal{F}}(m,n) := \sum_{f \in \mathcal{F}} \lambda_f(m) \overline{\lambda_f(n)}$. For instance, when the family consists of Dirichlet characters, a formula for $\Delta_{\mathcal{F}}$ is given by orthogonality of characters. For families of GL_2 forms, $\Delta_{\mathcal{F}}$ can be expanded into a sum of Kloosterman sums, by the Petersson/Bruggeman-Kuznetsov trace formula, which has seen extensive applications in number theory.

A large sieve inequality takes this analysis even futher, by bounding

$$\sum_{f \in \mathcal{F}} \Big| \sum_{n \leq N} a_n \lambda_f(n) \Big|^2, \tag{1.1}$$

where a_n are arbitrary complex coefficients. By general principles, the best one may hope for is a bound of the form $(|\mathcal{F}| + N) \sum_{n \leq N} |a_n|^2$. One can view this as a much more robust form of orthogonality, probing the sequence of values of $\lambda_f(n)$ by correlations with arbitrary sequences a_n. Large sieve inequalities are flexible and powerful estimates for bilinear forms having many applications. For instance, the classical large sieve inequality for Dirichlet characters plays a key role in proving the Bombieri-Vinogradov theorem. The GL_2 spectral large sieve has been valuable in understanding mean values of L-functions (in particular, to the fourth moment of the zeta function, which was Iwaniec's original application [I]). The reader is referred to [IK, Chapter 7] for a good introduction to large sieve inequalities.

The corresponding studies of higher rank families are still in their infancy. Bump, Friedberg, and Goldfeld [BFG] developed many of the foundational properties of the GL_3 Poincare series, and in particular discovered the analogous sums to the GL_2 Kloosterman sums. Recently, Blomer [Bl] succeeded in formulating a GL_3 Bruggeman-Kuznetsov formula with smooth bump functions appearing on the spectral side. Blomer also derived a form of the spectral GL_3 large sieve inequality, but without a focus on obtaining a sharp result. In principle, one may also derive a large sieve inequality from Goldfeld-Kontorovich's work [GK], but again this was not the focus of the authors and the result would not be numerically strong.

This material is based upon work supported by the National Science Foundation under agreement No. DMS-1401008. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

One of our main goals here is to obtain a stronger form of the GL_3 spectral large sieve inequality. To state the results, we set up some of the necessary notation as in [BFG] [G] [Bl]. Consider the family of Hecke-Maass cusp forms ϕ_j for $SL_3(\mathbb{Z})\backslash\mathcal{H}$, with spectral parameters ν_1, ν_2. The Langlands parameters associated to ϕ_j are $\alpha_1 = 2\nu_1 + \nu_2$, $\alpha_2 = -\nu_1 + \nu_2$, and $\alpha_3 = -\nu_1 - 2\nu_2$. Blomer has shown that the number of ϕ_j with $\nu_1 = iT_1 + O(1)$, $\nu_2 = iT_2 + O(1)$, weighted by R_j^{-1}, where

$$R_j = \mathrm{Res}_{s=1} L(s, \phi_j \times \overline{\phi_j}), \tag{1.2}$$

is $\asymp T_1T_2(T_1 + T_2)$ (also see [Bl, (1.4)]). This is a natural weighting from the point of view of the Bruggeman-Kuznetsov formula. Let $\lambda_j(m, n)$ denote the Hecke eigenvalues of ϕ_j, with $\lambda_j(1, 1) = 1$. With an appropriate choice of scaling of Whittaker functions, then $\|\phi_j\|^2 \asymp R_j$ (e.g., see [Bl, Lemma 1]).

Theorem 1.1. *For an arbitrary complex sequence a_n, we have*

$$\sum_{\substack{\nu_1 = iT_1 + O(1) \\ \nu_2 = iT_2 + O(1)}} \frac{1}{R_j} \Big| \sum_{n \leq N} a_n \lambda_j(n, 1) \Big|^2 \ll \Big(T_1T_2(T_1 + T_2) + T_1T_2N^{3/2} \Big)^{1+\varepsilon} \sum_{n \leq N} |a_n|^2. \tag{1.3}$$

For comparison, Blomer's proof of the spectral large sieve (implicitly) shows

$$\sum_{\substack{\nu_1 = iT_1 + O(1) \\ \nu_2 = iT_2 + O(1)}} \frac{1}{R_j} \Big| \sum_{n \leq N} a_n \lambda_j(n, 1) \Big|^2 \ll \Big(T_1T_2(T_1 + T_2) + T_1T_2N^{2} \Big)^{1+\varepsilon} \sum_{n \leq N} |a_n|^2, \tag{1.4}$$

so Theorem 1.2 saves a potentially rather large factor $N^{1/2}$. In fact, Blomer shows a dyadic bound:

$$\sum_{\substack{T_1 \leq |\nu_1| \leq 2T_1 \\ T_2 \leq |\nu_2| \leq 2T_2}} \frac{1}{R_j} \Big| \sum_{n \leq N} a_n \lambda_j(n, 1) \Big|^2 \ll \Big(T_1^2T_2^2(T_1 + T_2) + T_1T_2N^{2} \Big)^{1+\varepsilon} \sum_{n \leq N} |a_n|^2, \tag{1.5}$$

which saves a factor T_1T_2 in the second, "off-diagonal," term compared to (1.4), via an oscillatory integral. The proof of Theorem 1.1 also uses an oscillatory integral for an extra savings, but it is a technical challenge to combine these two sources of savings and convert Theorem 1.1 into a dyadic version with a secondary term of the same size. It should be noted that Blomer's estimate arises by applying absolute values to the GL_3 Kloosterman sum, and estimating everything trivially (analogously to applying the Weil bound for Kloosterman sums). One can view the quality of a large sieve inequality for a family $\mathcal{F}$ as a measure of how well one may average with the family. As such, it is desirable to have strong results.

There are also large sieve-type results in higher rank due to Duke and Kowalski [DK], Venkatesh [V], and Blomer-Buttcane-Maga [BBM], but these study the conductor (or level) aspect. By adapting the method of [DK, Theorem 4], one could use duality and the convexity bound for Rankin-Selberg L-functions on $GL_3 \times GL_3$ to attempt to obtain estimates on the left hand side of (1.3). However, this method requires N to be very large compared to $T_1 + T_2$ to give a strong bound.

The GL_3 Bruggeman-Kuznetsov formula relates these spectral sums to a sum of GL_3 Kloosterman sums. The main technical contribution of this paper is to analyze multilinear forms with these Kloosterman sums. We will be using the Bruggeman-Kuznetsov formula in the form of [Bl, Proposition 4][1]. The geometric side of this formula involves the GL_3

[1] A corrected version of the formula can be found in [BBM, Theorem 6]

Kloosterman sums, which we now define. The (long element) Kloosterman sum is

$$S(m_1,m_2,n_1,n_2;D_1,D_2) = \sum\sum_{\substack{B_1,C_1 \,(\mathrm{mod}\ D_1)\\ B_2,C_2\,(\mathrm{mod}\ D_2)\\ (B_1,C_1,D_1)=(B_2,C_2,D_2)=1\\ D_1C_2+B_1B_2+C_1D_2\equiv 0\,(\mathrm{mod}\ D_1D_2)}} e\Big(\frac{m_1B_1+n_1(Y_1D_2-Z_1B_2)}{D_1}\Big) e\Big(\frac{m_2B_2+n_2(Y_2D_1-Z_2B_1)}{D_2}\Big), \tag{1.6}$$

where Y_1, Y_2, Z_1, Z_2 are defined (chosen) so that

$$Y_1B_1+Z_1C_1\equiv 1 \pmod{D_1}, \qquad Y_2B_2+Z_2C_2\equiv 1 \pmod{D_2}. \tag{1.7}$$

Bump, Friedberg, and Goldfeld [BFG, Lemmas 4.1 and 4.2] have shown that the above sum is well-defined, meaning that the value of the sum is independent of the choices of the Y_i and Z_i, and the coset representatives of the B_i and C_i.

Define

$$\mathcal{S} = \mathcal{S}(\alpha,\beta,\gamma) = \sum_{D_1,D_2,m,n} \gamma_{D_1,D_2}\alpha_m\beta_n S(1,m,n,1,D_1,D_2), \tag{1.8}$$

where $\alpha_m, \beta_n, \gamma_{D_1,D_2}$ are finite sequences. For our application to the spectral large sieve, we are most interested in the case where $|\gamma_{D_1,D_2}| \leq 1$. Especially in light of its connections to the large sieve, it is fundamental to estimate $\mathcal{S}$, but it is also of independent interest. Our main result is

Theorem 1.2. *Suppose that α_m, β_n, and γ_{D_1,D_2} are complex sequences supported on $m,n \leq N$, $D_1 \leq X_1$, and $D_2 \leq X_2$. Furthermore suppose that $|\gamma_{D_1,D_2}| \leq 1$. For an arbitrary finitely supported sequence $\beta = (\beta_n)$, let*

$$M(\beta) = \sum_{q\leq \min(X_1,X_2)} \sum_{d_1|q} \frac{d_1}{q} \sum_{\substack{c\leq \frac{X_1}{q}\\ (c,q)=1}} \sideset{}{^*}\sum_{t\,(\mathrm{mod}\ c)} \Big| \sum_{(n,q)=d_1} \beta_n e\Big(\frac{tn}{c}\Big)\Big|^2 \tag{1.9}$$

where Σ^* denotes that t is restricted by $(t,c)=1$. Then

$$|\mathcal{S}| \ll (X_1X_2)^{1+\varepsilon} M(\alpha)^{1/2} M(\beta)^{1/2}. \tag{1.10}$$

For some special choices of coefficients α_m, β_n (e.g. Dirichlet series coefficients of an L-function), one could potentially use alternative techniques to handle small c, which explains why we have stated Theorem 1.2 in this form. For arbitrary coefficients, one cannot do better than the large sieve inequality (see [IK, Theorem 7.11]), which implies

$$M(\beta) \ll (X_1^2+N)X_1^{\varepsilon}\|\beta\|^2, \tag{1.11}$$

where here and throughout the paper we use the notation (for an arbitrary sequence β of finite support)

$$\|\beta\| = \Big(\sum_{n\in\mathbb{Z}} |\beta_n|^2\Big)^{1/2}. \tag{1.12}$$

Hence we immediately derive

Corollary 1.3. *With the same conditions and notation as Theorem 1.2, we have*

$$|\mathcal{S}| \ll (X_1X_2)^{\varepsilon}(X_1X_2)(X_1^2+N)^{1/2}(X_2^2+N)^{1/2}\|\alpha\|\,\|\beta\|. \tag{1.13}$$

For the applications to the GL_3 spectral large sieve inequality, the formulation in Theorem 1.2 is better, because one can obtain additional savings using a hybrid large sieve inequality, which includes an archimedean integral.

For some ranges of the parameters, the following result is superior to Theorem 1.2:

Theorem 1.4. *Let $1 \leq H_1 \leq X_1$, $1 \leq H_2 \leq X_2$. Then*

$$|\mathcal{S}| \ll (X_1H_2+X_2H_1)(X_1X_2)^{\varepsilon}M^*(\alpha)^{1/2}M^*(\beta)^{1/2} + (X_1X_2)^{3/2+\varepsilon}N^{1+\varepsilon}\|\alpha\|\|\beta\|(H_1^{-1}+H_2^{-1}), \tag{1.14}$$

where $M^*(\beta)$ is defined as in (1.9), but with q restricted by $q \leq \min(H_1, H_2)$.

Remarks. In case $H_1 = X_1$, $H_2 = X_2$ the first term in (1.14) reduces to Theorem 1.2 (and the second term may be dropped). For the opposite extreme $H_1 = H_2 = 1$, the latter term corresponds to the "Weil bound" (see (3.2) below) while the first term may be dropped. The restrictions $1 \leq H_i \leq X_i$ may be dropped from the statement of Theorem 1.4, however then the result is worse than Theorem 1.2 or (3.2) below.

Again, the large sieve implies

Corollary 1.5. *Let $1 \leq H_1 \leq X_1$, $1 \leq H_2 \leq X_2$. Then*

$$|\mathcal{S}| \ll \Big[(X_1H_2+X_2H_1)(X_1^2+N)^{1/2}(X_2^2+N)^{1/2} + \frac{(X_1X_2)^{3/2}N}{H_1} + \frac{(X_1X_2)^{3/2}N}{H_2}\Big](X_1X_2N)^{\varepsilon}\|\alpha\|\,\|\beta\|. \tag{1.15}$$

Remarks. For N large, say $N \gg X_1^2+X_2^2$, Corollary 1.5 is optimized with $H_1 = X_1^{3/4}X_2^{1/4}$, $H_2 = X_1^{1/4}X_2^{3/4}$, and reduces to a bound that can be seen to be inferior to Corollary 1.3. On the other hand, if $N \ll \min(X_1^2, X_2^2)$, then the optimal bound occurs with $H_1 = N^{1/2}X_1^{1/4}X_2^{-1/4}$, $H_2 = N^{1/2}X_2^{1/4}X_1^{-1/4}$, and gives

$$|\mathcal{S}| \ll (X_1X_2)^{5/4}N^{1/2}(X_1^{1/2}+X_2^{1/2})(X_1X_2N)^{\varepsilon}\|\alpha\|\,\|\beta\|. \tag{1.16}$$

Recently, Buttcane [Bu1] [Bu2] has developed Mellin-Barnes integral representations for the weight functions occuring on the Kloosterman sum side of the Bruggeman-Kuznetsov formula. Blomer and Buttcane [BB] have used this formulation, with additional ideas, to obtain a subconvexity result for GL_3 Maass forms in the spectral aspect. It could be interesting to investigate if these alternative integral representations lead to additional savings in the spectral large sieve. Our preliminary calculations indicate this could be rather complicated, and since our main focus here is on the arithmetical aspects of the problem (rather than the archimedean integrals), we leave this for another occasion.

2. Acknowledgments

I thank Valentin Blomer, Jack Buttcane, and the referee for numerous suggestions and corrections that improved the quality of the paper.

3. Heuristic remarks

3.1. **On theorem 1.2.** We include a few remarks of an informal nature indicating that Theorem 1.2 is in a somewhat robust form, at least, under the assumption that X_1 and X_2 are not highly asymmetrical in size.

The Weil-type bound of Steven [S] (see [Bl, Lemma 3]) implies

$$\sum_{D_1 \le X_1} \sum_{D_2 \le X_2} |S(1, m, n, 1, D_1, D_2)| \ll (X_1 X_2)^{3/2+\varepsilon} (mn)^{\varepsilon}, \tag{3.1}$$

and therefore the trivial bound applied to $\mathcal{S}$ along with Cauchy's inequality gives

$$|\mathcal{S}| \ll (X_1 X_2)^{3/2+\varepsilon} N^{1+\varepsilon} \|\alpha\| \|\beta\|. \tag{3.2}$$

Therefore, for large N, Corollary 1.3 saves an additional factor $(X_1X_2)^{1/2}$ over (3.2).

In case $(D_1, D_2) = 1$, then from [BFG, Property 4.9], we have

$$S(m_1, m_2, n_1, n_2, D_1, D_2) = S(D_2 m_1, n_1, D_1) S(D_1 m_2, n_2, D_2) \tag{3.3}$$

so the contribution to $\mathcal{S}$ from $(D_1, D_2) = 1$, say $\mathcal{S}'$, is

$$\mathcal{S}' = \sum_{\substack{m,n \\ (D_1,D_2)=1}} \gamma_{D_1,D_2} \alpha_m \beta_n S(D_2, n, D_1) S(D_1, m, D_2). \tag{3.4}$$

It could so happen that γ_{D_1,D_2} always has the same sign as $\sum_{m,n} \alpha_m \beta_n S(D_2, n, D_1) S(D_1, m, D_2)$, so it should be essentially impossible to do better than bounding $\mathcal{S}'$ as follows:

$$|\mathcal{S}'| \le \sum_{(D_1,D_2)=1} \Big| \sum_n \beta_n S(D_2, n, D_1) \sum_m \alpha_m S(D_1, m, D_2) \Big|. \tag{3.5}$$

By an application of Cauchy's inequality, we have

$$|\mathcal{S}'| \le \Big(\sum_{(D_1,D_2)=1} \Big| \sum_n \beta_n S(D_2, n, D_1) \Big|^2 \Big)^{1/2} (\dots)^{1/2}, \tag{3.6}$$

with the dots representing a similar term. Next we drop the condition $(D_2, D_1) = 1$ and extend the sum over D_2 to $D_2 \le MD_1$ where M is the unique integer satisfying $X_2 \le MD_1 < X_2 + D_1$ (this extension is presumably rather wasteful in case X_2 is much smaller than X_1). Then we have

$$\sum_{D_2 \le MD_1} S(n_1, D_2, D_1) S(n_2, D_2, D_1) = MD_1 \sideset{}{^*}\sum_{x \,(\mathrm{mod}\ D_1)} e\Big(\frac{x(n_1 - n_2)}{D_1}\Big), \tag{3.7}$$

so the first expression in parentheses on the right hand side of (3.6) satisfies

$$(\dots) \le (X_1 + X_2) \sum_{D_1 \le X_1} \sideset{}{^*}\sum_{x \,(\mathrm{mod}\ D_1)} \Big| \sum_n \beta_n e\Big(\frac{xn}{D_1}\Big) \Big|^2. \tag{3.8}$$

A similar bound holds for the second factor in (3.6), of course. Therefore, by the large sieve inequality, we have

$$|\mathcal{S}'| \le (X_1 + X_2)(X_1^2 + N)^{1/2} (X_2^2 + N)^{1/2} \|\alpha\| \|\beta\|. \tag{3.9}$$

This gives a limitation to the final estimates we wish to obtain for $\mathcal{S}$. One observes that the bound (3.9) is superior to that of Corollary 1.3 by a factor $\min(X_1, X_2)$, which arises in the proof from considering D_1 and D_2 with a common factor.

The opposite extreme of $(D_1, D_2) = 1$ is $D_1 = D_2$. For simplicity consider $D_1 = D_2 = p$, prime. In this case, we have (see [BFG, Property 4.10] or Lemma 4.2 below)

$$S(1, m, n, 1, p, p) = S(m, 0; p)S(n, 0; p) + p. \tag{3.10}$$

Therefore, if $p|(m, n)$ the Kloosterman sum is of order p^2, while if $p \nmid m$, $p \nmid n$, it is of order p. The term p gives the dominant contribution, because in the situation when the Kloosterman sum has order p^2 (i.e., $p|(m, n)$), the rarity in m and n has relatively frequency p^{-2}, which is a net saving by a factor p. These terms give to $\mathcal{S}$ an amount, say $\mathcal{S}''$, given by

$$\mathcal{S}'' = \sum_{p \leq \min(X_1, X_2)} (p + 1)\gamma_{p,p} \sum_{(m,p)=1} \alpha_m \sum_{(n,p)=1} \beta_n. \tag{3.11}$$

If say $X_1 = X_2 = X$, then

$$\mathcal{S}'' \ll X^2 \Big| \sum_m \alpha_m \Big| \cdot \Big| \sum_n \beta_n \Big|, \tag{3.12}$$

which is best-possible since the sum of $(p+1)\gamma_{p,p}$ may have the same sign as $\sum_m \alpha_m \sum_n \beta_n$. A bound of this magnitude is included with $c = q = d_1 = 1$ in (1.9) and (1.10). Cauchy's inequality applied to $\mathcal{S}''$ gives

$$\mathcal{S}'' \ll X^2 N \|\alpha\| \|\beta\|. \tag{3.13}$$

This matches the bound in Corollary 1.3 for N large and $X_1 = X_2$.

Of course, in actual practice it is necessary to treat all possible values of $\gcd(D_1, D_2)$ that "interpolate" the two extremes $(D_1, D_2) = 1$, and $D_1 = D_2$, and indeed this is accomplished in the proof of Theorem 1.2. In fact, this is the main difficulty in the proof.

The above remarks indicate that the quality of Theorem 1.2 comes largely from terms where (D_1, D_2) is large. This might be surprising in light of the relative rarity of such terms.

3.2. **The GL_2 spectral large sieve.** The spectral large sieve for $SL_2(\mathbb{Z})\backslash\mathbb{H}$ was originally proved by Iwaniec [I], while the case of congruence subgroups was extensively developed by Deshouillers and Iwaniec [DI]. Here we sketch a proof inspired by Jutila [J, Section 3], since we shall use this method as a motivating guide for the more challenging GL_3 case. Recall that the GL_2 spectral large sieve states

$$\sum_{T \leq t_j \leq T+\Delta} \frac{1}{R_j} \Big| \sum_{n \leq N} a_n \lambda_j(n) \Big|^2 \ll (NT)^{\varepsilon} (\Delta T + N) \sum_n |a_n|^2, \tag{3.14}$$

where R_j is given by (1.2) (but for ϕ_j a Hecke-Maass cusp form on $SL_2(\mathbb{Z})$), and $1 \leq \Delta \leq T$.

The GL_2 Bruggeman-Kuznetsov formula gives

$$\sum_{T \leq t_j \leq T+\Delta} \frac{1}{R_j} \Big| \sum_{N/2 < n \leq N} a_n \lambda_j(n) \Big|^2 \ll \Delta T \sum_n |a_n|^2 + \mathcal{K}, \tag{3.15}$$

where

$$\mathcal{K} = \sum_{N/2 < m,n \leq N} a_m \overline{a_n} \sum_{c=1}^{\infty} \frac{S(m, n; c)}{c} B\Big(\frac{\sqrt{mn}}{c}\Big), \tag{3.16}$$

and where $B(x)$ is a certain integral transform of a nonnegative weight function h that is $\gg 1$ for $T \leq t \leq T + \Delta$. For an appropriate smooth choice of h, $B(x)$ is very small unless

$x \gg \Delta T^{1-\varepsilon}$. Then by a Mellin transform, we have approximately that for $x \asymp X$,

$$B(x) \approx \Delta T \int_{|t| \ll X} X^{-1} x^{it} b(t) dt, \tag{3.17}$$

where $b(t) \ll 1$. Here b depends on X, but not on x. Applying this formula to $\mathcal{K}$, we derive

$$\mathcal{K} \lessapprox \sum_{C \text{ dyadic}} \frac{\Delta T}{CX} \int_{|t| \ll X} b(2t) \sum_{c \asymp C} c^{-2it} \sideset{}{^*}\sum_{a \,(\mathrm{mod}\, c)} \Big(\sum_{m \asymp N} a_m m^{it} e\Big(\frac{am}{c}\Big)\Big)\Big(\sum_{n \asymp N} \overline{a_n} n^{it} e\Big(\frac{\overline{a}n}{c}\Big)\Big) dt. \tag{3.18}$$

The hybrid large sieve inequality of Gallagher [Ga] states

$$\int_{|t| \leq X} \sum_{c \leq C} \sideset{}{^*}\sum_{a \,(\mathrm{mod}\, c)} \Big| \sum_{n \leq N} a_n n^{-it} e\Big(\frac{an}{c}\Big)\Big|^2 dt \ll (C^2 X + N) \sum_n |a_n|^2. \tag{3.19}$$

Applying this to $\mathcal{K}$ after a use of Cauchy-Schwarz, and using $X \asymp \frac{N}{C}$, and $C \ll \frac{N}{\Delta T}(NT)^\varepsilon$, we derive

$$\mathcal{K} \lessapprox \Delta T \sum_{C \text{ dyadic}} (CX)^{-1} (C^2 X + N) \sum_n |a_n|^2 \ll N (NT)^\varepsilon \sum_n |a_n|^2. \tag{3.20}$$

The main observation is that the GL_1 hybrid large sieve inequality drives the final estimations, and only rather crude information is required on B, namely its truncation and size of its Mellin transform. The hybrid aspect of the large sieve is able to recover the loss in separation of variables in B.

For later use, we shall require a different version (though morally equivalent) of the hybrid large sieve than that given by Gallagher. The following is a special case of [Y, Lemma 6.1].

Lemma 3.1. *Let b_m be arbitrary complex numbers, and suppose $Y \gg 1$. Then*

$$\int_1^2 \sum_{b \leq B} \sideset{}{^*}\sum_{x \,(\mathrm{mod}\, b)} \Big| \sum_{N \leq m < N+M} b_m e\left(\frac{xm}{b}\right) e\Big(\frac{tm}{Y}\Big)\Big|^2 dt \ll (B^2 + Y) \sum_{N \leq m < N+M} |b_m|^2. \tag{3.21}$$

4. Preliminary arithmetical results

For ease of reference, we collect here some results. First we need an individual "Weil-type" bound. This estimate was proved by Stevens [S] but without explicit dependence on the m_i and n_i, which was subsequently investigated by Buttcane [Bu1, Theorem 4].

Lemma 4.1. *For $m_1, m_2, n_1, n_2 \in \mathbb{Z} \setminus \{0\}$, we have*

$$S(m_1, m_2, n_1, n_2, D_1, D_2) \ll (D_1 D_2)^{1/2+\varepsilon} ((D_1, D_2)(m_1 n_2, [D_1, D_2])(m_2 n_1, [D_1, D_2]))^{1/2}. \tag{4.1}$$

This estimate is not sharp for $(D_1, D_2) > 1$, but it is difficult to extract clean results from the literature (see [DF, Theorem 3.7]). We may obtain some easy improvements by way of explicit computations in some important special cases:

Lemma 4.2 ([BFG])**.** *Suppose $l \geq 1$. Then*

$$S(m_1, m_2, n_1, n_2, p, p^l) = S(n_1, 0; p) S(m_2, n_2 p, p^l) + S(m_1, 0; p) S(n_2, m_2 p; p^l) + \delta_{l=1}(p-1). \tag{4.2}$$

Lemma 4.3. *Suppose $b \geq 1$, $c \geq 2$, and $(\alpha\beta, p) = 1$. Then*

$$S(\alpha, \beta p^b, p^c) = 0. \tag{4.3}$$

Proof. If $b \geq c$, then $S(\alpha, \beta p^b, p^c) = S(1, 0, p^c) = 0$ since $c \geq 2$, so suppose $c \geq b+1$. Opening the Kloosterman sum as a sum over $x \pmod{p^c}$, we change variables $x = x_1(1 + p^{c-b}x_2)$, where x_1 runs modulo p^{c-b} (coprime to p) and x_2 runs modulo p^b. Then $\overline{x} \equiv \overline{x_1} \pmod{p^{c-b}}$, and so

$$S(\alpha, \beta p^b, p^c) = \sideset{}{^*}\sum_{x_1 \,(\mathrm{mod}\ p^{c-b})} \sum_{x_2 \,(\mathrm{mod}\ p^b)} e\Big(\frac{\alpha x_1(1 + p^{c-b}x_2) + \beta p^b \overline{x_1}}{p^c}\Big). \tag{4.4}$$

The sum over x_2 then vanishes since $b \geq 1$ and $(\alpha x_1, p) = 1$. $\square$

Consider $S(a, y, x, b, D_1, D_2)$ with $(a, D_1) = (b, D_2) = 1$. Then define its (partial, middle two-variable) Fourier transform by

$$\widehat{S}(a, u, t, b, D_1, D_2) = \frac{1}{D_1 D_2} \sum_{x \,(\mathrm{mod}\ D_1)} \sum_{y \,(\mathrm{mod}\ D_2)} S(a, y, x, b, D_1, D_2) e\Big(\frac{-xt}{D_1}\Big) e\Big(\frac{-yu}{D_2}\Big), \tag{4.5}$$

so that the Fourier inversion formula reads

$$S(a, m, n, b, D_1, D_2) = \sum_{t \,(\mathrm{mod}\ D_1)} \sum_{u \,(\mathrm{mod}\ D_2)} e\Big(\frac{tn}{D_1} + \frac{um}{D_2}\Big) \widehat{S}(a, u, t, b, D_1, D_2). \tag{4.6}$$

Define

$$\mathcal{R}(t, D_1, D_2) = \max_{(ab, D_1)=1} \sum_{u \,(\mathrm{mod}\ D_2)} |\widehat{S}(a, u, bt, 1, D_1, D_2)|. \tag{4.7}$$

Remark. Using elementary properties of the Kloosterman sums, we may alternatively use the definition

$$\mathcal{R}(t, D_1, D_2) = \max_{(b, D_1)=1} \sum_{u \,(\mathrm{mod}\ D_2)} |\widehat{S}(1, u, bt, 1, D_1, D_2)|. \tag{4.8}$$

This follows by using that $S(a, y, x, 1, D_1, D_2) = S(1, y, ax, 1, D_1, D_2)$ (see [BFG, Property 4.3]), so that after a change of variables we derive

$$\widehat{S}(a, u, bt, 1, D_1, D_2) = \widehat{S}(1, u, \overline{a}bt, 1, D_1, D_2). \tag{4.9}$$

The presence of the maximum in (4.7) is to facilitate the use of the Chinese Remainder Theorem which leads to a more pleasant multiplicative structure for $\mathcal{R}$:

Lemma 4.4. *The function $\mathcal{R}(t, D_1, D_2)$ is jointly multiplicative in t, D_1, D_2.*

Proof. Say $D_1 = C_1 E_1$ and $D_2 = C_2 E_2$ with $(C_1 C_2, E_1 E_2) = 1$. Also write $x = x_C E_1 \overline{E_1} + x_E C_1 \overline{C_1}$, and similarly $y = y_C E_1 \overline{E_1} + y_E C_1 \overline{C_1}$, where x_C, y_C, x_E, y_E run modulo C_1, C_2, E_1, E_2, respectively. Then using [BFG, Property 4.15], we have

(4.10)

$$\widehat{S}(a, u, t, 1, D_1, D_2) = \sum_{x_C, y_C} S(\overline{E_1}^2 E_2 a, \overline{E_2}^2 E_1 y_C, x_C, 1, C_1, C_2) e\Big(\frac{-x_C \overline{E_1} t}{C_1}\Big) e\Big(\frac{-y_C \overline{E_2} u}{C_2}\Big)$$

$$\frac{1}{D_1 D_2} \sum_{x_E, y_E} S(\overline{C_1}^2 C_2 a, \overline{C_2}^2 C_1 y_E, x_E, 1, E_1, E_2) e\Big(\frac{-x_E \overline{C_1} t}{E_1}\Big) e\Big(\frac{-y_E \overline{C_2} u}{E_2}\Big).$$

Changing variables $y_C \to E_2^2\overline{E_1}y_C$, $y_E \to C_2^2\overline{C_1}y_E$, we derive

(4.11)
$$\widehat{S}(a,u,t,1,D_1,D_2) = \frac{1}{C_1C_2}\sum_{x_C,y_C} S(\overline{E_1}^2E_2a, y_C, x_C, 1, C_1, C_2)e\Big(\frac{-x_C\overline{E_1}t}{C_1}\Big)e\Big(\frac{-y_CE_2\overline{E_1}u}{C_2}\Big)$$
$$\frac{1}{E_1E_2}\sum_{x_E,y_E} S(\overline{C_1}^2C_2a, y_E, x_E, 1, E_1, E_2)e\Big(\frac{-x_E\overline{C_1}t}{E_1}\Big)e\Big(\frac{-y_EC_2\overline{C_1}u}{E_2}\Big).$$

Therefore,

(4.12)
$$\widehat{S}(a,u,t,1,C_1E_1,C_2E_2) = \widehat{S}(\overline{E_1}^2E_2a, uE_2\overline{E_1}, t\overline{E_1}, 1, C_1, C_2)\widehat{S}(\overline{C_1}^2C_2a, uC_2\overline{C_1}, t\overline{C_1}, 1, E_1, E_2).$$

Using (4.12), we derive that

$$\mathcal{R}(t,C_1E_1,C_2E_2) = \max_{(ab,C_1E_1)=1}\sum_{u \,(\mathrm{mod}\ C_2E_2)} |\widehat{S}(\overline{E_1}^2E_2a, uE_2\overline{E_1}, bt\overline{E_1}, 1, C_1, C_2)|$$
$$|\widehat{S}(\overline{C_1}^2C_2a, uC_2\overline{C_1}, bt\overline{C_1}, 1, E_1, E_2)|. \tag{4.13}$$

In the right hand side of (4.13), the first line only depends on u modulo C_2, and a, b modulo C_1, while the second line only depends on u modulo E_2, and a, b modulo E_1. Therefore, we have

$$\mathcal{R}(t,C_1E_1,C_2E_2) = \max_{(a_Cb_C,C_1)=1}\sum_{u_C \,(\mathrm{mod}\ C_2)} |\widehat{S}(\overline{E_1}^2E_2a_C, u_CE_2\overline{E_1}, b_Ct\overline{E_1}, 1, C_1, C_2)|$$
$$\max_{(a_Eb_E,E_1)=1}\sum_{u_E \,(\mathrm{mod}\ E_2)} |\widehat{S}(\overline{C_1}^2C_2a_E, u_EC_2\overline{C_1}, b_Et\overline{C_1}, 1, E_1, E_2)|. \tag{4.14}$$

Changing variables $a_C \to E_1^2\overline{E_2}a_C$, $u_C \to E_1\overline{E_2}u_C$, $b_C \to E_1b_C$, and similarly for a_E, u_E, and b_E, we derive $\mathcal{R}(t,C_1E_1,C_2E_2) = \mathcal{R}(t,C_1,C_2)\mathcal{R}(t,E_1,E_2)$, as desired. □

Lemma 4.5. *Let*

$$\mathcal{R}'(u,D_1,D_2) = \max_{\substack{(a,D_1)=1\\(b,D_2)=1}}\sum_{t \,(\mathrm{mod}\ D_1)} |\widehat{S}(a,bu,t,1,D_1,D_2)|. \tag{4.15}$$

Then

$$\mathcal{R}'(u,D_1,D_2) = \mathcal{R}(u,D_2,D_1). \tag{4.16}$$

Proof. Using $S(a,y,x,1,D_1,D_2) = S(x,1,a,y,D_1,D_2) = S(1,x,y,a,D_2,D_1)$ (see Properties 4.5 and 4.4 of [BFG]), along with $S(1,x,y,a,D_2,D_1) = S(1,ax,y,1,D_2,D_1)$ ([BFG, Property 4.3]), we derive that

(4.17)
$$\widehat{S}(a,bu,t,1,D_1,D_2) = \frac{1}{D_1D_2}\sum_{x \,(\mathrm{mod}\ D_1)}\sum_{y \,(\mathrm{mod}\ D_2)} S(1,x,y,1,D_2,D_1)e\Big(\frac{-x\overline{a}t}{D_1}\Big)e\Big(\frac{-ybu}{D_2}\Big)$$
$$= \widehat{S}(1,\overline{a}t,bu,1,D_2,D_1).$$

From this, and using (4.8), we complete the proof. □

Lemma 4.6. *We have*

$$\widehat{S}(a,u,t,b,D_1,D_2) = \widehat{S}(b,t,u,a,D_2,D_1). \tag{4.18}$$

Proof. A minor variation of the proof of Lemma 4.5 gives the result. ☐

Definition 4.7 (Definition of ν)**.** *Suppose p is a prime. If $n \in \mathbb{Z}$, we define $\nu_p(n)$ to be the standard p-adic valuation of n. If $k \geq 1$ and $t \in \mathbb{Z}/p^k\mathbb{Z}$ we define $\nu_p(t)$ to be the largest $j \leq k$ such that $t \equiv 0 \pmod{p^j}$.*

Remark. One may easily check that $\nu_p(t)$ is well-defined for $t \in \mathbb{Z}/p^k\mathbb{Z}$; without the restriction $j \leq k$, two coset representatives may have different p-adic valuations.

Lemma 4.8. *Suppose $(ab,p) = 1$, and set $\nu = \nu_p(t)$. Then*

$$\mathcal{R}(t,p^k,p^l) \leq (k+1)p^l + p^{\nu+l}\delta(\nu \leq \tfrac{2}{3}\min(k,l)). \tag{4.19}$$

Remark. For $k \neq l$, our proof shows that we can replace $p^{\nu+l}$ by $p^{\frac{\nu}{2}+l}$, and restrict $\nu \leq \frac{1}{2}\min(k,l)$. It is plausible one can save this factor $p^{\nu/2}$ for $k = l$, but since this would not improve Theorem 1.2, and since our proof is already quite long, we avoided this line of inquiry. The key point in Lemma 4.8 is that the "loss" from the factor p^ν is countered by the condition $p^\nu | t$. This has the practical effect that large values of ν give essentially the same bound as for $\nu = 0$.

The proof of Lemma 4.8 is given in Section 6.

Corollary 4.9. *We have*

$$\mathcal{R}(t,D_1,D_2) \ll D_2(D_1D_2)^\varepsilon \sum_{\substack{d|t \\ d^3|(D_1,D_2)^2}} d. \tag{4.20}$$

Proof. Since both sides are multiplicative, it suffices to check on prime powers, in which case it follows immediately from Lemma 4.8. ☐

Lemma 4.10. *Let $q \leq X$. The number of integers $n \leq X$ that share the same set of prime divisors as q (that is, such that $\nu_p(n) \geq 1$ iff $\nu_p(q) \geq 1$ for all primes p) is $\ll_\varepsilon X^\varepsilon$, for any $\varepsilon > 0$.*

Proof. This is similar to a divisor-type bound. Suppose that the prime factors occuring in q are $p_1, \dots, p_r$. Then by Rankin's trick, we have

$$\sum_{\substack{n = p_1^{a_1} \dots p_r^{a_r} \leq X \\ a_i \geq 1, \text{ all } i}} 1 \leq \sum_{a_1=1}^{\infty} \dots \sum_{a_r=1}^{\infty} \left(\frac{X}{p_1^{a_1} \dots p_r^{a_r}}\right)^\varepsilon = \frac{X^\varepsilon}{(p_1^\varepsilon - 1)\dots(p_r^\varepsilon - 1)}. \tag{4.21}$$

Given $\varepsilon > 0$, there are finitely many primes such that $p^\varepsilon \leq 2$. Then with $C(\varepsilon) = \prod_{p: p^\varepsilon \leq 2}(p^\varepsilon - 1)^{-1}$, we may bound the right hand side of (4.21) by $C(\varepsilon)X^\varepsilon$. ☐

5. Proof of Theorems 1.2 and 1.4

5.1. **Initial decomposition.** Our first steps involve factoring D_1 and D_2 in appropriate ways and using the Chinese remainder theorem to correspondingly factor the Kloosterman sum.

First we extract the largest divisors of D_1 and D_2 that are coprime to each other. Precisely, write $D_1 = g_1E_1$, $D_2 = g_2E_2$, where $(E_1E_2, g_1g_2) = 1$, $(E_1,E_2) = 1$, and g_1 and g_2 have the same set of prime divisors (meaning, $\nu_p(g_1) \geq 1$ iff $\nu_p(g_2) \geq 1$). Then by [BFG, Property 4.7], we have

$$S(1, m, n, 1, g_1 E_1, g_2 E_2) = S(\overline{g_1}^2 g_2, \overline{g_2}^2 g_1 m, n, 1, E_1, E_2) S(\overline{E_1}^2 E_2, \overline{E_2}^2 E_1 m, n, 1, g_1, g_2). \tag{5.1}$$

By (3.3), we have

$$S(\overline{g_1}^2 g_2, \overline{g_2}^2 g_1 m, n, 1, E_1, E_2) = S(E_2 \overline{g_1}^2 g_2, n, E_1) S(E_1 \overline{g_2}^2 g_1, m, E_2). \tag{5.2}$$

Therefore,

$$|\mathcal{S}| \leq \sideset{}{'}\sum_{g_1, g_2, E_1, E_2} \Big| \sum_{m,n} \alpha_m \beta_n S(E_2 \overline{g_1}^2 g_2, n, E_1) S(E_1 \overline{g_2}^2 g_1, m, E_2) S(\overline{E_1}^2 E_2, \overline{E_2}^2 E_1 m, n, 1, g_1, g_2) \Big|, \tag{5.3}$$

where the prime represents the conditions:

$$g_1 E_1 \leq X_1, \quad g_2 E_2 \leq X_2, \quad (E_1 E_2, g_1 g_2) = 1, \quad (E_1, E_2) = 1, \quad \nu_p(g_1) \geq 1 \text{ iff } \nu_p(g_2) \geq 1. \tag{5.4}$$

We factor the moduli further by extracting the prime factors of g_1 and g_2 such that $\nu_p(g_1) = \nu_p(g_2) = 1$. Precisely, write $g_1 = q h_1$, $g_2 = q h_2$ where q is the product of primes such that $\nu_p(g_1) = \nu_p(g_2) = 1$, so that for all $p | h_1 h_2$, $\nu_p(h_1) \geq 2$ or $\nu_p(h_2) \geq 2$, and $(q, h_1 h_2) = 1$. Then we have

$$\begin{aligned} &S(\overline{E_1}^2 E_2, \overline{E_2}^2 E_1 m, n, 1, q h_1, q h_2) \\ &\qquad = S(\overline{(h_1 E_1)}^2 h_2 E_2, \overline{(h_2 E_2)}^2 h_1 E_1 m, n, 1, q, q) S(\overline{q} \overline{E_1}^2 E_2, \overline{q} \overline{E_2}^2 E_1 m, n, 1, h_1, h_2). \end{aligned} \tag{5.5}$$

By (3.10), and using $(ab, p) = 1$, we have

$$S(a, bm, n, 1, p, p) = S(m, 0, p) S(n, 0, p) + p = \begin{cases} p^2 - p + 1, & p | (m, n), \\ p + 1, & p \nmid m, p \nmid n, \\ 1, & p | m, p \nmid n \\ 1, & p | n, p \nmid m, \end{cases} \tag{5.6}$$

and so by the Chinese remainder theorem, if q is squarefree and $(ab, q) = 1$, then

$$S(a, bm, n, 1, q, q) = \prod_{\substack{p | q \\ p \nmid m, p \nmid n}} (p + 1) \prod_{p | (m,n,q)} (p^2 - p + 1). \tag{5.7}$$

Set $d_1 = (n, q)$ and $d_2 = (m, q)$, and define

$$A(d_1, d_2, q) = \prod_{p | d_1, d_2} (p^2 - p + 1) \prod_{p | q, p \nmid d_1, p \nmid d_2} (p + 1). \tag{5.8}$$

Then the above calculations show

$$S(\overline{(h_1 E_1)}^2 h_2 E_2, \overline{(h_2 E_2)}^2 h_1 E_1 m, n, 1, q, q) = A(d_1, d_2, q). \tag{5.9}$$

One easily checks

$$A(d_1, d_2, q) \ll q^{1+\varepsilon} \frac{(d_1, d_2)^3}{d_1 d_2}. \tag{5.10}$$

Summarizing the above discussion, we have shown

$$|\mathcal{S}| \leq \sideset{}{'}\sum_{h_1,h_2,q,E_1,E_2} \sum_{d_1,d_2|q} A(d_1,d_2,q) \Big| \sum_{(n,q)=d_1} \sum_{(m,q)=d_2} \alpha_m \beta_n S(\overline{q}\overline{E_1}^2 E_2, \overline{q}\overline{E_2}^2 E_1 m, n, 1, h_1, h_2) \\ S(E_2 \overline{q}\overline{h_1}^2 h_2, n, E_1) S(E_1 \overline{q}\overline{h_2}^2 h_1, m, E_2) \Big|, \tag{5.11}$$

where the prime on the sum is updated to represent the conditions:

$$\begin{gathered} qh_1E_1 \leq X_1, \quad qh_2E_2 \leq X_2, \quad (E_1E_2, qh_1h_2) = 1, \quad (E_1, E_2) = 1, \quad \nu_p(q) \in \{0,1\}, \\ (q, h_1h_2) = 1, \quad \nu_p(h_1) = 0 \text{ iff } \nu_p(h_2) = 0, \quad p|h_1h_2 \Rightarrow \nu_p(h_1) \geq 2 \text{ or } \nu_p(h_2) \geq 2. \end{gathered} \tag{5.12}$$

Remark. Heuristically, the sum over h_1 and h_2 is somewhat small since both integers share the same prime divisors, and for each prime $p|h_1h_2$, p^2 divides at least one of h_1, h_2. If we let $\mathcal{S}'''$ denote the terms on the right hand side of (5.11) with $h_1 = h_2 = 1$, then following the arguments of Section 3.1, one can derive

$$\mathcal{S}''' \ll (X_1 + X_2)^{1+\varepsilon} (X_1^2 + \min(X_1, X_2)N)^{1/2} (X_2^2 + \min(X_1, X_2)N)^{1/2} \|\alpha\| \|\beta\|. \tag{5.13}$$

This is better than our final bound on $\mathcal{S}$ given by Corollary 1.3 for large X_1, X_2, so perhaps a more careful analysis of h_1 and h_2 could lead to a modest improvement.

If either h_1 or h_2 is large, then it can be beneficial to estimate the sum with absolute values, exploiting the reduced number of moduli under consideration. Define $\mathcal{S}_{qh_i \leq H_i}$ to be the sum on the right hand side of (5.11) with $qh_1 \leq H_1$ and $qh_2 \leq H_2$, and similarly define $\mathcal{S}_{qh_1 > H_1}$ and $\mathcal{S}_{qh_2 > H_2}$ corresponding to the terms with $qh_1 > H_1$ and $qh_2 > H_2$, respectively. Then we have the decomposition $|\mathcal{S}| \leq \mathcal{S}_{qh_1 > H_1} + \mathcal{S}_{qh_2 > H_2} + \mathcal{S}_{qh_i \leq H_i}$. In the proof of Theorem 1.2, we may set $H_1 = X_1$, $H_2 = X_2$, and then $\mathcal{S}_{qh_i > H_i} = 0$, for $i = 1, 2$, so these terms may be discarded.

5.2. **Large h_i.** In this subsection we estimate $\mathcal{S}_{qh_1 > H_1}$ and $\mathcal{S}_{qh_2 > H_2}$.

Lemma 5.1. *We have*

$$\mathcal{S}_{qh_1 > H_1} \ll H_1^{-1} (X_1X_2)^{3/2+\varepsilon} N \|\alpha\| \|\beta\|, \tag{5.14}$$

and

$$\mathcal{S}_{qh_2 > H_2} \ll H_2^{-1} (X_1X_2)^{3/2+\varepsilon} N \|\alpha\| \|\beta\|. \tag{5.15}$$

Proof. Define

$$T(m, n, D_1, D_2) = \max_{\substack{(a,D_1)=1 \\ (b,D_2)=1}} |S(a, bm, n, 1, D_1, D_2)|. \tag{5.16}$$

By the Weil bound, we have

$$\mathcal{S}_{qh_1 > H_1} \leq \sideset{}{'}\sum_{h_2,E_1,E_2} \sideset{}{'}\sum_{qh_1 > H_1} \sum_{d_1,d_2|q} A(d_1,d_2,q) \sum_{(n,q)=d_1} \sum_{(m,q)=d_2} |\alpha_m \beta_n| \\ \tau(E_1)\tau(E_2)(E_1E_2)^{1/2} T(m, n, h_1, h_2). \tag{5.17}$$

Trivially summing over E_1 and E_2 and using (5.10), we obtain

$$\mathcal{S}_{qh_1>H_1} \ll (X_1X_2)^{3/2+\varepsilon} \sum_{q\leq \min(X_1,X_2)} \sum_{d_1,d_2|q} \frac{(d_1,d_2)}{q^2} \sideset{}{'}\sum_{h_1>q^{-1}H_1} \sideset{}{'}\sum_{h_2} \frac{1}{(h_1h_2)^{3/2}} \sum_{(n,q)=d_1} \sum_{(m,q)=d_2} |\alpha_m\beta_n| T(m,n,h_1,h_2). \tag{5.18}$$

Write the prime factorizations of h_1 and h_2 as follows:

$$\begin{aligned} h_1 &= j_1k_1l_1, \quad j_1 = p_1\dots p_r, \quad k_1 = q_1^{b_1}\dots q_s^{b_s}, \quad l_1 = \rho_1^{c_1}\dots\rho_t^{c_t} \\ h_2 &= j_2k_2l_2, \quad j_2 = p_1^{a_1}\dots p_r^{a_r}, \quad k_2 = q_1\dots q_s, \quad l_2 = \rho_1^{\gamma_1}\dots\rho_t^{\gamma_t}, \end{aligned} \tag{5.19}$$

where $a_i, b_i, c_i, \gamma_i \geq 2$, for all i, $(p_i, q_j\rho_k) = (q_j,\rho_k) = 1$ for all i,j,k, and all p_i, q_j, ρ_k are prime.

We first estimate $T(m,n,j_1,j_2)$. Suppose $l \geq 2$. By Lemmas 4.2 and 4.3, we have

$$|S(a,bm,n,1,p,p^l)| = |S(n,0;p)S(m,p;p^l)|. \tag{5.20}$$

If $p \nmid m$ then this vanishes, while if $p|m$ then we have

$$|S(a,bm,n,1,p,p^l)| = p|S(n,0;p)S(\tfrac{m}{p},1;p^{l-1})| \leq p(n,p)\delta(p|m)lp^{\frac{l-1}{2}}. \tag{5.21}$$

Therefore,

$$T(m,n,j_1,j_2) \ll (j_1j_2)^{1/2+\varepsilon}(n,j_1)\delta(j_1|m), \tag{5.22}$$

and by symmetry,

$$T(m,n,k_1,k_2) \ll (k_1k_2)^{1/2+\varepsilon}(m,k_2)\delta(k_2|n). \tag{5.23}$$

Finally, by Lemma 4.1, we have

$$T(m,n,l_1,l_2) \leq (l_1l_2)^{1/2+\varepsilon}((l_1,l_2)(mn,[l_1,l_2]))^{1/2}. \tag{5.24}$$

Therefore, we have

$$\mathcal{S}_{qh_1>H_1} \ll (X_1X_2)^{3/2+\varepsilon} \sum_{q\leq \min(X_1,X_2)} \sum_{d_1,d_2|q} \frac{(d_1,d_2)}{q^2} \sideset{}{'}\sum_{h_1>q^{-1}H_1} \frac{1}{h_1} \sideset{}{'}\sum_{h_2} \frac{(l_1,l_2)^{1/2}}{h_2} \sum_{\substack{n\equiv 0 \,(\mathrm{mod}\ k_2) \\ (n,q)=d_1}} \sum_{\substack{m\equiv 0 \,(\mathrm{mod}\ j_1) \\ (m,q)=d_2}} |\alpha_m\beta_n|(mn,[l_1,l_2])^{1/2}(n,j_1)(m,k_2). \tag{5.25}$$

We claim

$$\sum_{\substack{n\equiv 0 \,(\mathrm{mod}\ k_2) \\ (n,q)=d_1}} \sum_{\substack{m\equiv 0 \,(\mathrm{mod}\ j_1) \\ (m,q)=d_2}} |\alpha_m\beta_n|(mn,[l_1,l_2])^{1/2}(n,j_1)(m,k_2) \ll \frac{N(X_1X_2)^\varepsilon}{(d_1d_2)^{1/2}}\|\alpha\|\|\beta\|. \tag{5.26}$$

Toward this, we first observe the simple bound

$$\sum_{n\leq N} |\alpha_n|(n,q)^{1/2} \leq \tau(q)N^{1/2}\|\alpha\|. \tag{5.27}$$

Using the trivial inequalities $(mn,[l_1,l_2]) \leq (m,l_1l_2)(n,l_1l_2)$, $(n,j_1) \leq (n,j_1)^{1/2}j_1^{1/2}$, $(m,k_2) \leq (m,k_2)^{1/2}k_2^{1/2}$, and the fact that $(j_1,k_2)=1$, we derive the claim.

Inserting (5.26) into (5.25), we conclude
(5.28)
$$\mathcal{S}_{qh_1>H_1} \ll (X_1X_2)^{3/2+\varepsilon} N\|\alpha\|\|\beta\| \sum_{q\leq \min(X_1,X_2)} \sum_{d_1,d_2|q} \frac{(d_1,d_2)}{(d_1d_2)^{1/2}q^2} \sideset{}{'}\sum_{h_1>q^{-1}H_1} \frac{1}{h_1} \sideset{}{'}\sum_{h_2} \frac{(l_1,l_2)^{1/2}}{h_2},$$

and to complete the proof of Lemma 5.1 it now suffices to show

$$\sideset{}{'}\sum_{h_1>H} \frac{1}{h_1} \sideset{}{'}\sum_{h_2} \frac{(l_1,l_2)^{1/2}}{h_2} \ll H^{-1}(X_1X_2)^{\varepsilon}, \tag{5.29}$$

where $H>0$, using the easy estimate $(d_1,d_2)\leq (d_1d_2)^{1/2}$, and trivially summing over q (it is essentially a harmonic series).

We now prove (5.29). First we examine the inner sum over h_2. Writing the expression in terms of the prime factorizations (5.19), we have

$$\sideset{}{'}\sum_{h_2} \frac{(l_1,l_2)^{1/2}}{h_2} = \sum_{a_1,\dots,a_r,\gamma_1,\dots,\gamma_t\geq 2} \frac{\rho_1^{\frac{\min(c_1,\gamma_1)}{2}}\dots\rho_t^{\frac{\min(c_t,\gamma_t)}{2}}}{p_1^{a_1}\dots p_r^{a_r} q_1\dots q_s \rho_1^{\gamma_1}\dots\rho_t^{\gamma_t}}. \tag{5.30}$$

The reader may recall that once h_1 is fixed, the prime divisors of h_2 are already determined, which explains why the sum is only over the exponents a_i, γ_i. It is easily noted that

$$\sum_{a_1,\dots,a_r\geq 2} \frac{1}{p_1^{a_1}\dots p_r^{a_r}} \leq \frac{2^r}{(p_1\dots p_r)^2}, \qquad \sum_{\gamma_1,\dots,\gamma_t\geq 2} \frac{\rho_1^{\frac{\min(c_1,\gamma_1)}{2}}\dots\rho_t^{\frac{\min(c_t,\gamma_t)}{2}}}{\rho_1^{\gamma_1}\dots\rho_t^{\gamma_t}} \ll \frac{2^t}{\rho_1\dots\rho_t}, \tag{5.31}$$

with an absolute implied constant, and so,

$$\frac{1}{h_1} \sideset{}{'}\sum_{h_2} \frac{(l_1,l_2)^{1/2}}{h_2} \ll \frac{2^{r+t}}{(p_1\dots p_r)^3 q_1^{b_1+1}\dots q_s^{b_s+1}\rho_1^{c_1+1}\dots\rho_t^{c_t+1}}. \tag{5.32}$$

Inserting this into the left hand side of (5.29), and recalling the implicit condition $h_1\leq X_1$, we have

$$\begin{aligned} \sideset{}{'}\sum_{h_1>H} \frac{1}{h_1} \sideset{}{'}\sum_{h_2} \frac{(l_1,l_2)^{1/2}}{h_2} &\ll \sum_{p_1\dots p_r q_1^{b_1}\dots q_s^{b_s}\rho_1^{c_1}\dots\rho_t^{c_t}>H} \frac{2^{r+t}}{(p_1\dots p_r)^3 q_1^{b_1+1}\dots q_s^{b_s+1}\rho_1^{c_1+1}\dots\rho_t^{c_t+1}} \\ &\leq \frac{1}{H} \sum_{p_1\dots p_r q_1^{b_1}\dots q_s^{b_s}\rho_1^{c_1}\dots\rho_t^{c_t}>H} \frac{2^{r+t}}{(p_1\dots p_r)^2 q_1\dots q_s\rho_1\dots\rho_t} \ll H^{-1}X_1^{\varepsilon}. \end{aligned} \tag{5.33}$$

This shows (5.29), and concludes the proof of (5.14). The other estimate (5.15) follows from (5.14) by symmetry. □

5.3. **Small h_i.** The main goal of this subsection is

Lemma 5.2. *We have*

$$\mathcal{S}_{qh_i\leq H_i} \ll (X_1H_2+X_2H_1)(X_1X_2N)^{\varepsilon} M^*(\alpha)^{1/2} M^*(\beta)^{1/2}. \tag{5.34}$$

Choosing $H_1=X_1$, $H_2=X_2$, we have $M^*(\beta)=M(\beta)$ and $M^*(\alpha)=M(\alpha)$, and $\mathcal{S}_{qh_1>X_1} = \mathcal{S}_{qh_2>X_2} = 0$, and we obtain Theorem 1.2. More generally, combining Lemmas 5.1 and 5.2 proves Theorem 1.4.

Proof. We begin by inserting the formula

$$S(\overline{q}\overline{E_1}^2 E_2, \overline{q}\overline{E_2}^2 E_1 m, n, 1, h_1, h_2) = \sum_{t \,(\mathrm{mod}\ h_1)} \sum_{u \,(\mathrm{mod}\ h_2)} \widehat{S}(\overline{q}\overline{E_1}^2 E_2, qE_2^2\overline{E_1}u, t, 1, h_1, h_2) e\Big(\frac{tn}{h_1}\Big) e\Big(\frac{um}{h_2}\Big), \tag{5.35}$$

into (5.11), getting

$$\mathcal{S}_{qh_i \le H_i} \le \sideset{}{'}\sum_{h_1,h_2,q,E_1,E_2} \sum_{d_1,d_2|q} A(d_1,d_2,q) \sum_{t \,(\mathrm{mod}\ h_1)} \sum_{u \,(\mathrm{mod}\ h_2)} |\widehat{S}(\overline{q}\overline{E_1}^2 E_2, qE_2^2\overline{E_1}u, t, 1, h_1, h_2)| \Big| \sum_{(n,q)=d_1} \beta_n S(E_2\overline{q}\overline{h_1}^2 h_2, n, E_1) e\Big(\frac{tn}{h_1}\Big) \sum_{(m,q)=d_2} \alpha_m S(E_1\overline{q}\overline{h_2}^2 h_1, m, E_2) e\Big(\frac{um}{h_2}\Big)\Big|. \tag{5.36}$$

We shall occasionally leave the conditions $qh_1 \le H_1$, $qh_2 \le H_2$ implicit in the notation. By Cauchy's inequality, we write $\mathcal{S}_{qh_i \le H_i} \le S_1^{1/2} S_2^{1/2}$ where

$$S_1 = \sideset{}{'}\sum_{h_1,h_2,q,E_1,E_2} \sum_{d_1,d_2|q} A(d_1,d_2,q) \sum_{t \,(\mathrm{mod}\ h_1)} \sum_{u \,(\mathrm{mod}\ h_2)} |\widehat{S}(\overline{q}\overline{E_1}^2 E_2, qE_2^2\overline{E_1}u, t, 1, h_1, h_2)| \Big| \sum_{(n,q)=d_1} \beta_n S(E_2\overline{q}\overline{h_1}^2 h_2, n, E_1) e\Big(\frac{tn}{h_1}\Big)\Big|^2, \tag{5.37}$$

and S_2 is given by a similar formula. Write this as

$$S_1 = \sideset{}{'}\sum_{q,E_1} \sum_{d_1,d_2|q} A(d_1,d_2,q) S_1', \tag{5.38}$$

where

$$S_1' = \sideset{}{'}\sum_{h_1,h_2,E_2} \sum_{t \,(\mathrm{mod}\ h_1)} \sum_{u \,(\mathrm{mod}\ h_2)} |\widehat{S}(\overline{q}\overline{E_1}^2 E_2, qE_2^2\overline{E_1}u, t, 1, h_1, h_2)| \Big| \sum_{(n,q)=d_1} \beta_n S(E_2\overline{q}\overline{h_1}^2 h_2, n, E_1) e\Big(\frac{tn}{h_1}\Big)\Big|^2. \tag{5.39}$$

Recalling the definition of $\mathcal{R}$ from (4.7), we have

$$S_1' \le \sideset{}{'}\sum_{h_1,h_2,E_2} \sum_{t \,(\mathrm{mod}\ h_1)} \mathcal{R}(t,h_1,h_2) \Big| \sum_{(n,q)=d_1} \beta_n S(E_2\overline{q}\overline{h_1}^2 h_2, n, E_1) e\Big(\frac{tn}{h_1}\Big)\Big|^2. \tag{5.40}$$

Using the trick described surrounding (3.7), we extend the sum over E_2 to a complete sum modulo $E_1 \le \frac{X_1}{qh_1}$ (forgetting the various coprimality conditions on E_2 by positivity), giving

$$S_1' \le \sideset{}{^*}\sum_{x \,(\mathrm{mod}\ E_1)} \sideset{}{'}\sum_{h_1,h_2} \sum_{t \,(\mathrm{mod}\ h_1)} \mathcal{R}(t,h_1,h_2) \Big(\frac{X_1}{qh_1} + \frac{X_2}{qh_2}\Big) \Big| \sum_{(n,q)=d_1} \beta_n e\Big(\frac{xn}{E_1}\Big) e\Big(\frac{tn}{h_1}\Big)\Big|^2. \tag{5.41}$$

Next write $g_1 = (t, h_1)$, and change variables $h_1 = g_1 h_1'$, $t = g_1 t'$, so that $(t', h_1') = 1$. This gives

(5.42)
$$S_1' \leq \sideset{}{^*}\sum_{x \pmod{E_1}} \sideset{}{'}\sum_{g_1, h_1', h_2} \sideset{}{^*}\sum_{t' \pmod{h_1'}} \mathcal{R}(t' g_1, g_1 h_1', h_2)\Big(\frac{X_1}{q g_1 h_1'} + \frac{X_2}{q h_2}\Big) \Big| \sum_{(n,q)=d_1} \beta_n e\Big(\frac{xn}{E_1}\Big) e\Big(\frac{t'n}{h_1'}\Big)\Big|^2.$$

Here the primes on the sums refer to the conditions (5.12) with h_1 replaced by $g_1 h_1'$. Observe $\mathcal{R}(t' g_1, h_1' g_1, h_2) = \mathcal{R}(g_1, h_1' g_1, h_2)$, by definition, and re-arrange this in the form

$$S_1' \leq \sideset{}{^*}\sum_{x \pmod{E_1}} \sideset{}{'}\sum_{h_1' \leq \frac{H_1}{q}} \sideset{}{^*}\sum_{t' \pmod{h_1'}} \Big| \sum_{(n,q)=d_1} \beta_n e\Big(\frac{xn}{E_1}\Big) e\Big(\frac{t'n}{h_1'}\Big)\Big|^2 \sideset{}{'}\sum_{g_1 \leq \frac{H_1}{q h_1'}} \sideset{}{'}\sum_{h_2 \leq \frac{H_2}{q}} \mathcal{R}(g_1, g_1 h_1', h_2)\Big(\frac{X_1}{q g_1 h_1'} + \frac{X_2}{q h_2}\Big). \tag{5.43}$$

We claim

$$\sideset{}{'}\sum_{g_1 \leq \frac{H_1}{q h_1'}} \sideset{}{'}\sum_{h_2 \leq \frac{H_2}{q}} \mathcal{R}(g_1, g_1 h_1', h_2)\Big(\frac{X_1}{q g_1 h_1'} + \frac{X_2}{q h_2}\Big) \ll \frac{(X_1 X_2)^\varepsilon}{q^2 h_1'}(X_1 H_2 + X_2 H_1). \tag{5.44}$$

Proof of claim. We first bound the sum with the factor $\frac{X_1}{q g_1 h_1'}$. By Corollary 4.9, we have

$$\sideset{}{'}\sum_{g_1 \leq \frac{H_1}{q h_1'}} \sideset{}{'}\sum_{h_2 \leq \frac{H_2}{q}} \mathcal{R}(g_1, g_1 h_1', h_2) \frac{X_1}{q g_1 h_1'} \ll (X_1 X_2)^\varepsilon \sideset{}{'}\sum_{g_1 \leq \frac{H_1}{q h_1'}} \sideset{}{'}\sum_{h_2 \leq \frac{H_2}{q}} h_2 \frac{X_1}{q g_1 h_1'} \sum_{\substack{d | g_1 \\ d^3 | (g_1 h_1', h_2)^2}} d. \tag{5.45}$$

Reversing the order of summation, and estimating the sum over h_2 by Lemma 4.10 (one may safely drop the condition $d^3 | h_2^2$ when summing over h_2), this is

$$\ll \frac{(X_1 X_2)^\varepsilon X_1 H_2}{q^2 h_1'} \sum_{d \leq X_1} d \sum_{\substack{g_1 \leq \frac{H_1}{h_1'} \\ d | g_1,\, d^3 | (g_1 h_1')^2}} \frac{1}{g_1}. \tag{5.46}$$

Next we write $g_1 = dr$, where now $d | r^2 h_1'^2$, so we may execute the sum over d first as a divisor sum, and finally the sum over r satisfies $\sum_{r \leq X_1} r^{-1+\varepsilon} \ll X_1^\varepsilon$. This immediately gives a bound consistent with (5.44).

For the second sum with $\frac{X_2}{q h_2}$, we have by Corollary 4.9 that

$$\sideset{}{'}\sum_{g_1 \leq \frac{H_1}{q h_1'}} \sideset{}{'}\sum_{h_2 \leq \frac{H_2}{q}} \mathcal{R}(g_1, g_1 h_1', h_2) \frac{X_2}{q h_2} \ll (X_1 X_2)^\varepsilon \sideset{}{'}\sum_{g_1 \leq \frac{H_1}{q h_1'}} \sideset{}{'}\sum_{h_2 \leq \frac{H_2}{q}} \frac{X_2}{q} \sum_{\substack{d | g_1 \\ d^3 | (g_1 h_1', h_2)^2}} d. \tag{5.47}$$

We shall reverse the order of summation and execute the sum over h_2 first. The sum over h_2 is bounded by $O(X_2^\varepsilon)$ using Lemma 4.10, because one of the summation conditions is that h_2 and $g_1 h_1'$ share the same prime factors. Then the right hand side of (5.47) is

$$\ll (X_1 X_2)^\varepsilon \frac{X_2}{q} \sum_{g_1 \leq \frac{H_1}{q h_1'}} \sum_{\substack{d | g_1 \\ d^3 | (g_1 h_1')^2}} d. \tag{5.48}$$

Write $g_1 = dr$, whence this is

$$\ll (X_1X_2)^{\varepsilon}\frac{X_2}{q}\sum_{d\le X_1} d \sum_{\substack{r\le \frac{H_1}{dqh_1'} \\ r^2h_1'^2\equiv 0 \pmod{d}}} 1. \tag{5.49}$$

Suppose that $d = d'f$ where d' consists of the prime powers corresponding to the primes that divide h_1', so that $f := d/d'$ is then coprime to h_1'. Then we have $r^2 \equiv 0 \pmod{f}$. Let f^* be the integer such that the congruence $r^2 \equiv 0 \pmod{f}$ is equivalent to $r \equiv 0 \pmod{f^*}$. Then the above expression is bounded by

$$\ll (X_1X_2)^{\varepsilon}\frac{X_2H_1}{q^2h_1'}\sum_{d'}\sum_{f}\frac{1}{f^*} \ll (X_1X_2)^{\varepsilon}\frac{X_2H_1}{q^2h_1'}, \tag{5.50}$$

since Lemma 4.10 shows the sum over d' is $\ll (X_1X_2)^{\varepsilon}$, and the sum over f is $\ll (X_1X_2)^{\varepsilon}$ by elementary reasoning (e.g. Rankin's trick). Thus we arrive at a bound consistent with (5.44). $\square$

The claim (5.44) applied to (5.43) implies

$$S_1' \ll \frac{(X_1X_2)^{\varepsilon}}{q^2}(X_1H_2 + X_2H_1)\sideset{}{^*}\sum_{x \pmod{E_1}}\sideset{}{'}\sum_{h_1'\le H_1}\frac{1}{h_1'}\sideset{}{^*}\sum_{t' \pmod{h_1'}}\Big|\sum_{(n,q)=d_1}\beta_n e\Big(\frac{xn}{E_1}\Big)e\Big(\frac{t'n}{h_1'}\Big)\Big|^2. \tag{5.51}$$

Inserting this into S_1, and using the Chinese remainder theorem to combine the sums modulo E_1 and h_1' to the single modulus E_1h_1', we derive

$$\begin{aligned} S_1 \ll (X_1H_2 + X_2H_1)(X_1X_2)^{\varepsilon} &\sum_{q\le \min(H_1,H_2)}\sum_{d_1,d_2|q}\frac{A(d_1,d_2,q)}{q^2} \\ &\sum_{\substack{E_1h_1'\le \frac{X_1}{q} \\ h_1'\le H_1}}\frac{1}{h_1'}\sideset{}{^*}\sum_{x \pmod{E_1h_1'}}\Big|\sum_{(n,q)=d_1}\beta_n e\Big(\frac{xn}{E_1h_1'}\Big)\Big|^2. \end{aligned} \tag{5.52}$$

We group together E_1h_1' into a single variable c, use $h_1'^{-1} \le 1$ and (5.10), obtaining

$$S_1 \ll (X_1H_2 + X_2H_1)(X_1X_2)^{\varepsilon}\sum_{q\le \min(H_1,H_2)}\sum_{d_1,d_2|q}\frac{(d_1,d_2)^3}{qd_1d_2}\sum_{\substack{c\le \frac{X_1}{q} \\ (c,q)=1}}\sideset{}{^*}\sum_{x \pmod{c}}\Big|\sum_{(n,q)=d_1}\beta_n e\Big(\frac{xn}{c}\Big)\Big|^2. \tag{5.53}$$

Using the crude bound $(d_1d_2)^{-1}(d_1,d_2)^3 \le d_1$, and trivially summing over d_2, we obtain

$$S_1 \ll (X_1H_2 + X_2H_1)(X_1X_2)^{\varepsilon}\sum_{q\le \min(H_1,H_2)}\sum_{d_1|q}\frac{d_1}{q}\sum_{\substack{c\le \frac{X_1}{q} \\ (c,q)=1}}\sideset{}{^*}\sum_{x \pmod{c}}\Big|\sum_{(n,q)=d_1}\beta_n e\Big(\frac{xn}{c}\Big)\Big|^2, \tag{5.54}$$

which is $(X_1H_2 + X_2H_1)(X_1X_2)^{\varepsilon}M^*(\beta)$ where recall $M(\beta)$ was defined by (1.9) and $M^*(\beta)$ has the same definition but with $q \le \min(H_1, H_2)$.

We also need to estimate S_2. It is given by a similar formula to S_1, except with h_1 and h_2 switched, E_1 and E_2 switched, β_n replaced by α_m, d_1 and d_2 switched, and we need to work with $\mathcal{R}'(u, h_1, h_2)$ instead of $\mathcal{R}$ (for which see Lemma 4.5). Therefore, by a symmetry argument, we have $S_2 \ll (X_1H_2 + X_2H_1)(X_1X_2)^{\varepsilon}M^*(\alpha)$. $\square$

Remark 5.3. *The proof given above works equally well if we replace* $S(1, m, n, 1, D_1, D_2)$ *by* $S(1, \epsilon_1 m, \epsilon_2 n, 1, D_1, D_2)$ *for* $\epsilon_1, \epsilon_2 \in \{-1, 1\}$.

6. Bounds on $\mathcal{R}$

This section is devoted to the long proof of Lemma 4.8. The overarching idea of the proof is to evaluate $\widehat{S}(a, u, t, b, p^k, p^l)$ in explicit terms (as much as possible), and to trivially sum over u. Lemma 4.6 will allow us to focus almost entirely on the case $k < l$. Except for the cases $k = l \geq 2$, we have evaluated $\widehat{S}$ exactly. It is a pleasant fact that this is much easier than evaluating the Kloosterman sum itself (compare to Theorem 0.3 of [DF]).

In the proof of Theorem 1.2 we only needed estimates on $\mathcal{R}(t, p^k, p^l)$ when $k, l \geq 1$ and $\max(k, l) \geq 2$, but since the small values of k and l are easily treated, we shall cover all the cases as stated in Lemma 4.8.

6.1. **The case $k = 0$, or $l = 0$.** By a direct calculation, and using (3.3), we have

$$\widehat{S}(a, u, t, b, 1, p^l) = e\Big(\frac{\overline{u}b}{p^l}\Big)\delta(p \nmid u), \tag{6.1}$$

and by symmetry (that is, Lemma 4.6),

$$\widehat{S}(a, u, t, b, p^k, 1) = e\Big(\frac{\overline{t}a}{p^k}\Big)\delta(p \nmid t). \tag{6.2}$$

Trivially summing over u, we easily derive $\mathcal{R}(t, p^k, p^l) \leq p^l$ in case $k = 0$ or $l = 0$.

6.2. **The case $k = l = 1$.** By Lemma 4.2,

$$S(a, y, x, b, p, p^l) = S(x, 0; p)S(y, bp; p^l) + S(a, 0; p)S(b, yp; p^l) + (p-1)\delta(l = 1). \tag{6.3}$$

Therefore, recalling $(ab, p) = 1$ and $S(a, 0; p) = -1 = S(b, 0; p)$, we have

$$\begin{aligned}\widehat{S}(a, u, t, b, p, p) &= \frac{1}{p^2} \sum_{x \,(\mathrm{mod}\ p)} \sum_{y \,(\mathrm{mod}\ p)} e\Big(\frac{-xt}{p}\Big)e\Big(\frac{-yu}{p}\Big)[S(x, 0; p)S(y, 0; p) + p] \\ &= \delta(p \nmid t)\delta(p \nmid u) + p\delta(p|t)\delta(p|u).\end{aligned} \tag{6.4}$$

We immediately deduce $\mathcal{R}(t, p, p) \leq p$.

6.3. **The case $k = 1, l \geq 2$.** Using (6.3) and the fact that $S(b, yp; p^l) = 0$ following from Lemma 4.3, we derive

$$\begin{aligned}\widehat{S}(a, u, t, b, p, p^l) &= \frac{1}{p^{l+1}} \sum_{x \,(\mathrm{mod}\ p)} \sum_{y \,(\mathrm{mod}\ p^l)} S(x, 0; p)S(y, bp; p^l)e\Big(\frac{-xt}{p}\Big)e\Big(\frac{-yu}{p^l}\Big) \\ &= \delta(p \nmid t)\delta(p \nmid u)e\Big(\frac{\overline{u}bp}{p^l}\Big).\end{aligned} \tag{6.5}$$

We conclude

$$\mathcal{R}(t, p, p^l) \leq \delta(p \nmid t)p^l. \tag{6.6}$$

6.4. **The case $l = 1, k \geq 2$.** By (6.5) and Lemma 4.6, we have

$$\widehat{S}(a,u,t,b,p^k,p) = \delta(p \nmid t)\delta(p \nmid u) e\Big(\frac{\overline{t}ap}{p^k}\Big), \tag{6.7}$$

and so $\mathcal{R}(t,p^k,p) \leq p$.

Remark. For the remaining cases we do not use direct evaluations of S and instead calculate $\widehat{S}$ from the definition. We have

$$\widehat{S}(a,u,t,b,p^k,p^l) = \frac{1}{p^{k+l}} \sum_{x \,(\mathrm{mod}\, p^k)} \sum_{y \,(\mathrm{mod}\, p^l)} S(a,y,x,b,p^k,p^l) e\Big(\frac{-xt}{p^k}\Big) e\Big(\frac{-yu}{p^l}\Big) \tag{6.8}$$

$$= \Big[\frac{1}{p^{k+l}} \sum_{x \,(\mathrm{mod}\, p^k)} \sum_{y \,(\mathrm{mod}\, p^l)} \sum_{\substack{B_1,C_1 \,(\mathrm{mod}\, p^k) \\ (B_1,C_1,p^k)=1}} \sum_{\substack{B_2,C_2 \,(\mathrm{mod}\, p^l) \\ (B_2,C_2,p^l)=1 \\ p^kC_2+B_1B_2+p^lC_1\equiv 0 \,(\mathrm{mod}\, p^{k+l}) \\ Y_1B_1+Z_1C_1\equiv 1 \,(\mathrm{mod}\, p^k) \\ Y_2B_2+Z_2C_2\equiv 1 \,(\mathrm{mod}\, p^l)}} e\Big(\frac{aB_1 + x(Y_1p^l - Z_1B_2)}{p^k}\Big)$$

$$e\Big(\frac{yB_2 + b(Y_2p^k - Z_2B_1)}{p^l}\Big) e\Big(\frac{-xt}{p^k}\Big) e\Big(\frac{-yu}{p^l}\Big)\Big].$$

This simplifies as

$$\widehat{S}(a,u,t,b,p^k,p^l) = \sum_{\substack{B_1,C_1 \,(\mathrm{mod}\, p^k) \\ (B_1,C_1,p^k)=1}} \sum_{\substack{C_2 \,(\mathrm{mod}\, p^l) \\ (u,C_2,p^l)=1 \\ p^kC_2+B_1u+p^lC_1\equiv 0 \,(\mathrm{mod}\, p^{k+l}) \\ Y_1p^l-Z_1u\equiv t \,(\mathrm{mod}\, p^k) \\ Y_1B_1+Z_1C_1\equiv 1 \,(\mathrm{mod}\, p^k) \\ Y_2u+Z_2C_2\equiv 1 \,(\mathrm{mod}\, p^l)}} e\Big(\frac{aB_1}{p^k}\Big) e\Big(\frac{b(Y_2p^k - Z_2B_1)}{p^l}\Big). \tag{6.9}$$

Although a large expression, we found it helpful to have all the conditions written in the summation sign.

6.5. **The case $l > k \geq 2$.** Suppose that $p^\nu || t$, and write $t = p^\nu t'$. Here we will show

$$\widehat{S}(a,u,t,b,p^k,p^l) = p^\nu e\Big(\frac{b\overline{u'}}{p^{l-k+\nu}}\Big) e\Big(\frac{a\overline{t'}p^{l-k}}{p^\nu}\Big) S(a, b\overline{t'u'}, p^\nu)\delta(\nu \leq k/2), \tag{6.10}$$

where the sum vanishes unless $p^\nu || u$, in which case we write $u = p^\nu u'$.

Using only the trivial bound (not even the Weil bound) for the Kloosterman sum, we conclude

$$\mathcal{R}(t,p^k,p^l) \leq p^\nu \sum_{u \,(\mathrm{mod}\, p^l), p^\nu | u} p^\nu = p^{\nu+l}, \tag{6.11}$$

and in addition we have $\nu \leq k/2$. This estimate is consistent with Lemma 4.8.

The congruence $p^kC_2 + B_1u + p^lC_1 \equiv 0 \pmod{p^{k+l}}$ implies $p^k | B_1u$, and is equivalent to

$$C_2 \equiv -\frac{B_1u}{p^k} - p^{l-k}C_1 \pmod{p^l}. \tag{6.12}$$

Suppose $p^{k_1} || B_1$, and write $\widehat{S} = \sum_{k_1=0}^{k} V_{k_1}$ correspondingly. We first evaluate the terms with $1 \leq k_1 \leq k-1$. We write $B_1 = p^{k_1}R_1$ where R_1 runs mod p^{k_2}, with $k_1 + k_2 = k$. We also have $p^{k_2} | u$. Since $p | B_1$ and $p | u$, the coprimality conditions now require $p \nmid C_1$ and $p \nmid C_2$.

If $p^{k_2+1}|u$ then (6.12) would imply $p|C_2$, a contradiction. So $p^{k_2}||u$, and we write $u = p^{k_2}u'$. We set $Y_1 = Y_2 = 0$, and $Z_i = \overline{C_i}$. Then we have

$$V_{k_1} = \sideset{}{^*}\sum_{R_1 \pmod{p^{k_2}}} \sideset{}{^*}\sum_{C_1 \pmod{p^k}} \sideset{}{^*}\sum_{\substack{C_2 \pmod{p^l} \\ C_2 \equiv -R_1 u' - p^{l-k}C_1 \pmod{p^l} \\ -\overline{C_1}u'p^{k_2} \equiv t \pmod{p^k}}} e\Big(\frac{aR_1}{p^{k_2}}\Big) e\Big(\frac{-b\overline{C_2}R_1}{p^{l-k_1}}\Big). \tag{6.13}$$

Next we observe that $p^{k_2}||t$, so we write $t = p^{k_2}t'$, and then we have $C_1 \equiv -u'\overline{t'} \pmod{p^{k_1}}$. With these evaluations, we have

$$V_{k_1} = \sideset{}{^*}\sum_{R_1 \pmod{p^{k_2}}} \sideset{}{^*}\sum_{\substack{C_1 \pmod{p^k} \\ C_1 \equiv -u'\overline{t'} \pmod{p^{k_1}}}} e\Big(\frac{aR_1}{p^{k_2}}\Big) e\Big(\frac{bR_1\overline{(R_1u' + p^{l-k}C_1)}}{p^{l-k_1}}\Big). \tag{6.14}$$

To help simplify this expression, we expand as follows:

$$e\Big(\frac{bR_1\overline{(R_1u' + p^{l-k}C_1)}}{p^{l-k_1}}\Big) = e\Big(\frac{b\overline{u'}(R_1u' + p^{l-k}C_1 - p^{l-k}C_1)\overline{(R_1u' + p^{l-k}C_1)}}{p^{l-k_1}}\Big). \tag{6.15}$$

After simplification, this gives

$$V_{k_1} = e\Big(\frac{b\overline{u'}}{p^{l-k_1}}\Big) \sideset{}{^*}\sum_{R_1 \pmod{p^{k_2}}} \sideset{}{^*}\sum_{\substack{C_1 \pmod{p^k} \\ C_1 \equiv -u'\overline{t'} \pmod{p^{k_1}}}} e\Big(\frac{aR_1}{p^{k_2}}\Big) e\Big(\frac{-b\overline{u'}C_1\overline{(R_1u' + p^{l-k}C_1)}}{p^{k_2}}\Big). \tag{6.16}$$

We claim that if $1 \leq k_1 < k_2 \leq k-1$, then $V_{k_1} = 0$. For this, write $C_1 = f_1 + p^{k_2-1}f_2$ with f_1 running mod p^{k_2-1} and f_2 mod p^{k_1+1}. Since $k_1 < k_2$, the congruence $C_1 \equiv -u'\overline{t'} \pmod{p^{k_1}}$ gives no condition on f_2. Then note that since $l - k + k_2 - 1 \geq k_2$, we have

$$e\Big(\frac{-b\overline{u'}(f_1 + p^{k_2-1}f_2)\overline{(R_1u' + p^{l-k}(f_1 + p^{k_2-1}f_2))}}{p^{k_2}}\Big) = e\Big(\frac{-b\overline{u'}(f_1 + p^{k_2-1}f_2)\overline{(R_1u' + p^{l-k}f_1)}}{p^{k_2}}\Big), \tag{6.17}$$

and so the sum over f_2 will cause V_{k_2} to vanish.

Now suppose $k_1 \geq k_2$. Then the congruence on C_1 mod p^{k_1} determines C_1 mod p^{k_2}, and hence

$$V_{k_1} = e\Big(\frac{b\overline{u'}}{p^{l-k_1}}\Big) \sideset{}{^*}\sum_{R_1 \pmod{p^{k_2}}} \sideset{}{^*}\sum_{\substack{C_1 \pmod{p^k} \\ C_1 \equiv -u'\overline{t'} \pmod{p^{k_1}}}} e\Big(\frac{aR_1}{p^{k_2}}\Big) e\Big(\frac{b\overline{t'}\overline{(R_1u' + p^{l-k}(-u'\overline{t'}))}}{p^{k_2}}\Big), \tag{6.18}$$

which simplifies as

$$V_{k_1} = p^{k_2} e\Big(\frac{b\overline{u'}}{p^{l-k_1}}\Big) \sideset{}{^*}\sum_{R_1 \pmod{p^{k_2}}} e\Big(\frac{aR_1}{p^{k_2}}\Big) e\Big(\frac{b\overline{t'}\overline{u'}\overline{(R_1 - p^{l-k}\overline{t'})}}{p^{k_2}}\Big). \tag{6.19}$$

Changing variables $R_1 \to R_1 + p^{l-k}\overline{t'}$, we have

$$V_{k_1} = p^{k_2} e\Big(\frac{b\overline{u'}}{p^{l-k_1}}\Big) e\Big(\frac{a\overline{t'}p^{l-k}}{p^{k_2}}\Big) S(a, b\overline{t'}\overline{u'}, p^{k_2}), \tag{6.20}$$

and we recollect that $p^{k_2}||u$, $p^{k_2}||t$, and $1 \leq k_2 \leq k_1 \leq k-1$. If we define ν by $p^{\nu}||t$, then we have

$$\sum_{k_1=1}^{k-1} V_{k_1} = p^{\nu} e\Big(\frac{b\overline{u'}}{p^{l-k+\nu}}\Big) e\Big(\frac{a\overline{t'}p^{l-k}}{p^{\nu}}\Big) S(a, b\overline{t'}\overline{u'}, p^{\nu})\delta(1 \leq \nu \leq k/2). \tag{6.21}$$

So far we have left the cases with $k_1 = 0$ and $k_1 = k$ unevaluated, so we next turn to this.

We claim $V_0 = 0$, which takes some calculation. We have $p \nmid B_1$ so we can set $Y_1 = \overline{B_1}$ and $Z_1 = 0$. If $p^{k+1}|u$ then (6.12) would mean $p|C_2$, a contradiction. So $p^k||u$ and we write $u = p^k u'$. We must have $p \nmid C_2$ so we set $Y_2 = 0$, $Z_2 = \overline{C_2}$. With these evaluations, we have $p^k|t$ and

$$V_0 = \sideset{}{^*}\sum_{B_1 \,(\mathrm{mod}\, p^k)} \sum_{C_1 \,(\mathrm{mod}\, p^k)} \sum_{\substack{C_2 \,(\mathrm{mod}\, p^l) \\ C_2 \equiv -B_1 u' - p^{l-k} C_1 \pmod{p^l}}} e\Big(\frac{aB_1}{p^k}\Big) e\Big(\frac{-b\overline{C_2}B_1}{p^l}\Big). \tag{6.22}$$

Now we can write $C_1 = f_1 + p^{k-1} f_2$, and $C_2 = -B_1 u' - p^{l-k} f_1 - p^{l-1} f_2$, and so

$$V_0 = \sideset{}{^*}\sum_{B_1 \,(\mathrm{mod}\, p^k)} e\Big(\frac{aB_1}{p^k}\Big) \sum_{f_1 \,(\mathrm{mod}\, p^{k-1})} \sum_{f_2 \,(\mathrm{mod}\, p)} e\Big(\frac{-b\overline{(-B_1 u' - p^{l-k} f_1 - p^{l-1} f_2)}B_1}{p^l}\Big). \tag{6.23}$$

Note $\overline{1 + p^{l-1} f_2} \equiv 1 - p^{l-1} f_2 \pmod{p^l}$ since $l \geq 2$. This shows that the sum over f_2 vanishes, as desired.

Finally we evaluate V_k. Then we have $p^k|B_1$ so may set $B_1 = 0$ (that is, we choose the integer 0 for the coset representative of $0 \pmod{p^k}$), $p \nmid C_1$, $p|C_2$ so $p \nmid u$, and we may set $Y_1 = 0$, $Z_1 = \overline{C_1}$, $Y_2 = \overline{u}$, $Z_2 = 0$. Then

$$V_k = e\Big(\frac{bp^k\overline{u}}{p^l}\Big) \sideset{}{^*}\sum_{\substack{C_1 \,(\mathrm{mod}\, p^k) \\ -\overline{C_1}u \equiv t \,(\mathrm{mod}\, p^k)}} \sum_{\substack{C_2 \,(\mathrm{mod}\, p^l) \\ C_2 \equiv -p^{l-k} C_1 \,(\mathrm{mod}\, p^l)}} 1. \tag{6.24}$$

This means $p \nmid t$ and both C_1 and C_2 are uniquely determined. Therefore,

$$V_k = \delta(p \nmid t)\delta(p \nmid u) e\Big(\frac{bp^k\overline{u}}{p^l}\Big), \tag{6.25}$$

which coincidentally agrees with the right hand side of (6.20), except with $k_2 = \nu = 0$. Therefore, by adding (6.21) and (6.25) we obtain (6.10), as desired.

6.6. **The case $k > l \geq 2$.** By (6.10) and Lemma 4.6, we have

$$\widehat{S}(a, u, t, b, p^k, p^l) = p^{\nu} e\Big(\frac{a\overline{t'}}{p^{k-l+\nu}}\Big) e\Big(\frac{b\overline{u'}p^{k-l}}{p^{\nu}}\Big) S(b, a\overline{t'}\overline{u'}, p^{\nu})\delta(\nu \leq l/2), \tag{6.26}$$

where again ν is defined by $p^{\nu}||u$ and $p^{\nu}||t$. Using only the trivial bound for the Kloosterman sum, we derive

$$\mathcal{R}(t, p^k, p^l) \leq p^{\nu} \sum_{u \,(\mathrm{mod}\, p^l), p^{\nu}|u} p^{\nu} \leq p^{l+\nu}, \tag{6.27}$$

and in addition we have $\nu \leq l/2$. Again, this is consistent with Lemma 4.8.

6.7. **The case $k = l \geq 2$.** The case $k = l$ follows somewhat similar lines to the $k \neq l$ case, but there are some significant differences that require careful scrutiny. We do not have a clean formula for $\widehat{S}$ analogous to (6.10).

Performing some mild simplifications in (6.9), we obtain that $p^k | B_1 u$ and then

$$\widehat{S}(a, u, t, b, p^k, p^k) = \sum_{\substack{B_1, C_1 \pmod{p^k} \\ (B_1, C_1, p^k) = 1}} \sum_{\substack{C_2 \pmod{p^k} \\ (u, C_2, p^k) = 1 \\ C_1 + C_2 \equiv -\frac{B_1 u}{p^k} \pmod{p^k} \\ -Z_1 u \equiv t \pmod{p^k} \\ Y_1 B_1 + Z_1 C_1 \equiv 1 \pmod{p^k} \\ Y_2 u + Z_2 C_2 \equiv 1 \pmod{p^k}}} e\Big(\frac{aB_1}{p^k}\Big) e\Big(\frac{-bZ_2 B_1}{p^k}\Big). \tag{6.28}$$

As before, let V_{k_1} denote the subsum with $p^{k_1} || B_1$. We have $p^{k_2} | u$, where $k_1 + k_2 = k$, but unlike the case $l > k \geq 2$, we cannot conclude that $p^{k_2} || u$. We first estimate the cases with $k_1 = 0$ and $k_1 = k$.

We claim $V_0 = 0$. With these terms, we have $p \nmid B_1$, and so we may set $u = 0$ (that is, we choose $u = 0$ as the coset representative of $0 \pmod{p^k}$). Then $p \nmid C_2$, and $C_1 \equiv -C_2 \pmod{p^k}$. We also have $t = 0$. We may set $Y_2 = 0, Z_2 = \overline{C_2}$, and $Y_1 = \overline{B_1}$, $Z_1 = 0$. With these evaluations, we derive

$$V_0 = \sideset{}{^*}\sum_{B_1 \pmod{p^k}} \sideset{}{^*}\sum_{C_2 \pmod{p^k}} e\Big(\frac{aB_1}{p^k}\Big) e\Big(\frac{-b\overline{C_2} B_1}{p^k}\Big). \tag{6.29}$$

The sum over C_2 vanishes since it is a Ramanujan sum with modulus p^k, $k \geq 2$ (see Lemma 4.3).

For V_k, we have $B_1 = 0$. Then $p \nmid C_1$ and we set $Y_1 = 0$, $Z_1 = \overline{C_1}$. Then (6.12) becomes $C_2 \equiv -C_1 \pmod{p^k}$, so we have $p \nmid C_2$ and we are free to set $Y_2 = 0$, $Z_2 = \overline{C_2}$. Thus

$$V_k = \sideset{}{^*}\sum_{\substack{C_1 \pmod{p^k} \\ -\overline{C_1} u \equiv t \pmod{p^k}}} 1. \tag{6.30}$$

Since u is uniquely determined from C_1, we have

$$\sum_{u \pmod{p^k}} |V_k| = \phi(p^k). \tag{6.31}$$

Now consider V_{k_1} with $1 \leq k_1 \leq k-1$. Then $p | B_1$ and $p | u$ so $p \nmid C_1$, $p \nmid C_2$, and we set $Y_1 = Y_2 = 0$, $Z_i = \overline{C_i}$. We write $B_1 = p^{k_1} R_1$. Suppose $\nu_p(t) = \nu$, and write $t = p^\nu t''$. Then we must have $\nu \geq k_2$, from the congruence $-\overline{C_1} u \equiv t \pmod{p^k}$, and we can write $u = p^\nu u''$ instead of $u = p^{k_2} u'$. Then

$$V_{k_1} = \sideset{}{^*}\sum_{R_1 \pmod{p^{k_2}}} \sideset{}{^*}\sum_{\substack{C_1 \pmod{p^k} \\ C_1 \equiv -u'' \overline{t''} \pmod{p^{k-\nu}}}} e\Big(\frac{aR_1}{p^{k_2}}\Big) e\Big(\frac{bR_1 \overline{(C_1 + R_1 p^{\nu - k_2} u'')}}{p^{k_2}}\Big). \tag{6.32}$$

We claim that $V_{k_1} = 0$ if $k_1 < \nu < k$, as we now argue. Observing that $k - \nu < k_2$ (since $k - \nu = k_1 + k_2 - \nu$), we can write $C_1 = -u'' \overline{t''} + p^{k-\nu} f_1$, with f_1 running mod p^ν. Then we can write $f_1 = f_2 + f_3 p^{\nu-k_1-1}$ where f_2 runs mods $p^{\nu-k_1-1}$, and f_3 runs mod p^{k_1+1}. Then we arrive at a sum over f_3 of the form

$$\sum_{f_3 \pmod{p^{k_1+1}}} e\Big(\frac{\alpha\overline{(1+\beta p^{k_2-1} f_3)}}{p^{k_2}}\Big) = \sum_{f_3 \pmod{p^{k_1+1}}} e\Big(\frac{\alpha(1-\beta p^{k_2-1} f_3)}{p^{k_2}}\Big), \tag{6.33}$$

where $(\alpha\beta, p) = 1$, using $k_2 \geq 2$ which follows from $k_2 > k - \nu \geq 1$. Since this sum over f_3 vanishes, this means $V_{k_1} = 0$.

If $\nu = k$, we have from (6.32), after changing variables $C_1 \to C_1 - R_1 p^{\nu-k_2} u''$, that

$$V_{k_1} = \sideset{}{^*}\sum_{R_1 \pmod{p^{k_2}}} \sideset{}{^*}\sum_{C_1 \pmod{p^k}} e\Big(\frac{aR_1}{p^{k_2}}\Big) e\Big(\frac{bR_1\overline{C_1}}{p^{k_2}}\Big) = p^{k_1} S(1,0;p^{k_2})^2 = \begin{cases} 0, & k_2 \geq 2, \\ p^{k_1}, & k_2 = 1. \end{cases} \tag{6.34}$$

For $\nu = k$, we have

$$\sum_{k_1=1}^{k-1} \sum_{u \pmod{p^k}} |V_{k_1}| = p^{k-1}\delta(\nu = k). \tag{6.35}$$

Now suppose $\nu \leq k_1 < k$. This condition implies $k-\nu \geq k_2$, so the congruence $C_1 \equiv -u''\overline{t''} \pmod{p^{k-\nu}}$ determines C_1 modulo p^{k_2}, and so (6.32) simplifies as

$$V_{k_1} = p^{\nu} \sideset{}{^*}\sum_{R_1 \pmod{p^{k_2}}} e\Big(\frac{aR_1}{p^{k_2}}\Big) e\Big(\frac{bR_1\overline{u''}\overline{(-\overline{t''} + R_1 p^{\nu-k_2})}}{p^{k_2}}\Big). \tag{6.36}$$

If, in addition, $\nu - k_2 \geq k_2$, then this is simply given by

$$V_{k_1} = p^{\nu} S(u''a - t''b, 0; p^{k_2})\delta(k_2 \leq \nu/2). \tag{6.37}$$

It follows easily that for these values of k_2 and ν that

$$\sum_{u \pmod{p^k}} |V_{k_1}| \leq 2p^{\nu} p^{k_2} \frac{p^{k-\nu}}{p^{k_2}} = 2p^k, \tag{6.38}$$

and so,

$$\sum_{1 \leq k_2 \leq \nu/2} \sum_{u \pmod{p^k}} |V_{k_1}| \leq \nu p^k. \tag{6.39}$$

On the other hand, if $\nu - k_2 < k_2$ (we continue to assume $\nu \leq k_1 < k$), then we can write $R_1 = f_1 + p^{2k_2-\nu} f_2$, with f_1 mod $p^{2k_2-\nu}$, and f_2 mod $p^{\nu-k_2}$. Then (6.36) becomes

$$V_{k_1} = p^{\nu} \sideset{}{^*}\sum_{f_1 \pmod{p^{2k_2-\nu}}} \sum_{f_2 \pmod{p^{\nu-k_2}}} e\Big(\frac{a(f_1 + p^{2k_2-\nu} f_2)}{p^{k_2}}\Big) e\Big(\frac{b\overline{u''}(f_1 + p^{2k_2-\nu} f_2)\overline{(-\overline{t''} + f_1 p^{\nu-k_2})}}{p^{k_2}}\Big). \tag{6.40}$$

The inner sum over f_2 simplifies, and detects $u''a \equiv bt'' \pmod{p^{\nu-k_2}}$. Hence

$$V_{k_1} = p^{2\nu-k_2} \sideset{}{^*}\sum_{f_1 \pmod{p^{2k_2-\nu}}} e\Big(\frac{af_1}{p^{k_2}}\Big) e\Big(\frac{b\overline{u''} f_1 \overline{(-\overline{t''} + f_1 p^{\nu-k_2})}}{p^{k_2}}\Big) \delta(u''a \equiv t''b \pmod{p^{\nu-k_2}}). \tag{6.41}$$

Therefore, by a trivial bound on the f_1-sum, we have

$$|V_{k_1}| \leq p^{2\nu-k_2}\phi(p^{2k_2-\nu})\delta(u''a \equiv t''b \pmod{p^{\nu-k_2}}), \tag{6.42}$$

which implies

$$\sum_{u \,(\mathrm{mod}\ p^k)} |V_{k_1}| \leq p^{2\nu - k_2} \phi(p^{2k_2-\nu}) \frac{p^{k-\nu}}{p^{\nu-k_2}} \leq \phi(p^{2k_2+k-\nu}). \tag{6.43}$$

Here we certainly have $k_2 \leq \nu$ (from the paragraph following (6.31)), but we also have that ν is restricted by $\nu \leq \min(k-k_2, 2k_2) \leq 2k/3$. Therefore, these values of k_1 give

$$\sum_{k_1} \sum_{u \,(\mathrm{mod}\ p^k)} |V_{k_1}| \leq p^{k+\nu} \delta(\nu \leq 2k/3). \tag{6.44}$$

Combining (6.31), (6.35), (6.39), and (6.44), we derive

$$\mathcal{R}(t, p^k, p^k) \leq (k+1)p^k + p^{k+\nu}\delta(\nu \leq 2k/3). \tag{6.45}$$

7. Spectral summation formula

The remaining sections of the paper contain the proof of Theorems 1.1.

We shall use the GL_3 Bruggeman-Kuznetsov formula in the form given by Blomer [Bl]. Suppose that $T_1, T_2 \gg 1$, and consider the sum

$$\sum_{\substack{\nu_1 = iT_1 + O(1)\\ \nu_2 = iT_2 + O(1)}} \frac{1}{R_j} \Big| \sum_{N/2 < n \leq N} a_n \lambda_j(n,1) \Big|^2. \tag{7.1}$$

As in [Bl, (8.5)], let

$$F(y_1, y_2) = \sqrt{T_1 T_2 (T_1+T_2)}\, y_1^{i(\tau_1 + 2\tau_2)} y_2^{i(2\tau_1+\tau_2)} f(y_1) f(y_2), \tag{7.2}$$

where f is a fixed smooth, nonzero, non-negative function with support on $[1,2]$. Here τ_1, τ_2 are parameters satisfying $\tau_1 \asymp T_1$, $\tau_2 \asymp T_2$.

By following the proof of [Bl, Theorem 3], we have that

$$(7.1) \ll \sum_j \frac{1}{\|\phi_j\|^2} |\langle \widetilde{W}_{\nu_1,\nu_2}, F\rangle|^2 \Big| \sum_{n \leq N} a_n \lambda_j(n,1) \Big|^2 + (\mathrm{cts}) =: S(T_1, T_2, N), \tag{7.3}$$

where $\widetilde{W}_{\nu_1,\nu_2}$ is a completed Whittaker function associated to the Maass form ϕ_j, and (cts) represents the non-negative continuous spectrum contribution.

The Bruggeman-Kuznetsov formula in the form of [Bl, Proposition 4] says

$$S(T_1, T_2, N) = \sum_n \sum_m \overline{a_n} a_m \Big(\Sigma_1 + \Sigma_{2a} + \Sigma_{2b} + \Sigma_3 \Big), \tag{7.4}$$

where

$$\Sigma_1 = \delta_{m=n} \|F\|^2, \tag{7.5}$$

and Σ_{2a}, Σ_{2b}, and Σ_3 are sums involving GL_3 Kloosterman sums. With F defined by (7.2), then $\|F\|^2 \asymp T_1 T_2 (T_1 + T_2)$, which is the mass of the spectral ball to account for the diagonal terms. It turns out that we do not need to analyze Σ_{2a} and Σ_{2b}, so we omit their definitions (the interested reader may find them defined in [Bl, (8.2)]).

Here

$$\Sigma_3 = \sum_{\epsilon_1, \epsilon_2 = \pm 1} \sum_{D_1, D_2} \frac{S(1, \epsilon_1 m, \epsilon_2 n, 1, D_1, D_2)}{D_1 D_2} \mathcal{J}_{\epsilon_1, \epsilon_2}\Big(\frac{\sqrt{m D_1}}{D_2}, \frac{\sqrt{n D_2}}{D_1} \Big), \tag{7.6}$$

where with shorthand $x_3' = x_1x_2 - x_3$,

$$\mathcal{J}_{\epsilon_1,\epsilon_2}(A_1,A_2) = (A_1A_2)^{-2}\int_0^\infty\int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty e(-\epsilon_1A_1x_1y_1 - \epsilon_2A_2x_2y_2)$$
$$e\Big(-\frac{A_2}{y_2}\frac{x_1x_3+x_2}{1+x_2^2+x_3^2}\Big)e\Big(-\frac{A_1}{y_1}\frac{x_2x_3'+x_1}{1+x_1^2+x_3'^2}\Big)$$
$$F\Big(\frac{A_2}{y_2}\frac{\sqrt{1+x_1^2+x_3'^2}}{1+x_2^2+x_3^2},\frac{A_1}{y_1}\frac{\sqrt{1+x_2^2+x_3^2}}{1+x_1^2+x_3'^2}\Big)\overline{F}(A_1y_1,A_2y_2)dx_1dx_2dx_3\frac{dy_1dy_2}{y_1y_2}. \tag{7.7}$$

One pleasant feature of this integral expression is that the variables are practically separated, and the kernel function is easily bounded uniformly in all parameters. For comparison, the formulas of Buttcane [Bu2, Theorem 2] also directly separate the variables and only require a 2-fold integral, which should in principle be more efficient. However the tradeoff is that the kernel function is not as easy to bound, requiring one to work on multiple scales. Furthermore, the weight function depends on ϵ_1, ϵ_2 in a non-trivial way, leading to further case analysis.

Blomer (see [Bl, p.722]) showed that

$$\sum_{m,n} a_m\overline{a_n}(|\Sigma_{2a}|+|\Sigma_{2b}|) \ll N^\varepsilon(T_1+T_2)^{-100}\sum_n |a_n|^2, \tag{7.8}$$

which means these terms are practically negligible. This estimate arises because the weight function on the sum of Kloosterman sums side is very small for these terms. Taken together, this shows

Lemma 7.1. *We have*

$$\sum_{\substack{\nu_1=iT_1+O(1)\\ \nu_2=iT_2+O(1)}}\frac{1}{R_j}\Big|\sum_{N/2<n\le N} a_n\lambda_j(n,1)\Big|^2 \ll T_1T_2(T_1+T_2)\sum_n|a_n|^2 + \sum_{m,n} a_m\overline{a_n}\Sigma_3 + \frac{N^\varepsilon}{(T_1+T_2)^{100}}\sum_n |a_n|^2. \tag{7.9}$$

8. Manipulations of Σ_3

Our next step is to perform some elementary manipulations to Σ_3 in order to prepare it for the use of Theorem 1.2. Note that we can change variables $y_1 \to y_1/A_1$ and $y_2 \to y_2/A_2$ and use the definitions $\xi_1 = 1 + x_1^2 + x_3'^2$, $\xi_2 = 1 + x_2^2 + x_3^2$, to give

$$\mathcal{J}_{\epsilon_1,\epsilon_2}(A_1,A_2) = (A_1A_2)^{-2}\int_0^\infty\int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty e(-\epsilon_1x_1y_1 - \epsilon_2x_2y_2)$$
$$e\Big(-\frac{A_2^2}{y_2}\frac{x_1x_3+x_2}{1+x_2^2+x_3^2}\Big)e\Big(-\frac{A_1^2}{y_1}\frac{x_2x_3'+x_1}{1+x_1^2+x_3'^2}\Big)F\Big(\frac{A_2^2}{y_2}\frac{\xi_1^{1/2}}{\xi_2},\frac{A_1^2}{y_1}\frac{\xi_2^{1/2}}{\xi_1}\Big)\overline{F}(y_1,y_2)dx_1dx_2dx_3\frac{dy_1dy_2}{y_1y_2}. \tag{8.1}$$

Our next step is to insert the definition (7.2), and re-arrange the resulting expression. For later use, it may be helpful to note that for our values of A_1 and A_2 that

$$(A_2^2)^{i(\tau_1+2\tau_2)}(A_1^2)^{i(2\tau_1+\tau_2)} = m^{i(2\tau_1+\tau_2)}n^{i(\tau_1+2\tau_2)}D_1^{-3i\tau_2}D_2^{-3i\tau_1}. \tag{8.2}$$

In this way, we obtain

$$\sum_{n,m} \overline{a_n} a_m \Sigma_3 = \sum_{\epsilon_1,\epsilon_2=\pm 1} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \xi_1^{-\frac{3}{2}i\tau_1} \xi_2^{-\frac{3}{2}i\tau_2} dx_1 dx_2 dx_3 \tag{8.3}$$
$$\sum_{D_1,D_2}\sum_{n,m} \frac{T_1T_2(T_1+T_2)}{\frac{mn}{D_1D_2}} n^{i(\tau_1+2\tau_2)} m^{i(2\tau_1+\tau_2)} \overline{a_n} a_m \frac{S(1,\epsilon_1 m, \epsilon_2 n, 1, D_1, D_2)}{D_1^{1+3i\tau_2} D_2^{1+3i\tau_1}}$$
$$\Big[\int_0^\infty y_2^{-3i(\tau_1+\tau_2)} e(-\epsilon_2 x_2 y_2) e\Big(-\frac{nD_2}{y_2 D_1^2}\frac{x_1x_3+x_2}{1+x_2^2+x_3^2}\Big) f\Big(\frac{nD_2}{y_2D_1^2}\frac{\xi_1^{1/2}}{\xi_2}\Big) f(y_2)\frac{dy_2}{y_2}\Big]$$
$$\Big[\int_0^\infty y_1^{-3i(\tau_1+\tau_2)} e(-\epsilon_1 x_1 y_1) e\Big(-\frac{mD_1}{y_1 D_2^2}\frac{x_2x_3'+x_1}{1+x_1^2+x_3'^2}\Big) f\Big(\frac{mD_1}{y_1D_2^2}\frac{\xi_2^{1/2}}{\xi_1}\Big) f(y_1)\frac{dy_1}{y_1}\Big].$$

Now let us restrict to $D_1 \asymp X_1$, and $D_2 \asymp X_2$, and sum over these dyadic values of X_1, X_2 at the end. Since $m, n \asymp N$, if we let $(N/n)a_n = a_n'$, then $a_n \asymp a_n'$. The support on f constrains the x-variables into a certain region V of $\mathbb{R}^3$ that has measure $\ll (A_1A_2)^{2+\varepsilon} = (\frac{mn}{D_1D_2})^{1+\frac{\varepsilon}{2}}$, by [Bl, Lemma 4]. Furthermore, V is independent of m, n, y_1, y_2, D_1, D_2 (it depends on N, X_1, X_2). Write $|V| = \frac{N^2}{X_1X_2}$, so that the measure of V is at most $|V|N^\varepsilon$.

Blomer's bound [Bl, (8.9)] shows that $\mathcal{J}_{\epsilon_1,\epsilon_2}(A_1, A_2)$ is very small for $A_1^{4/3}A_2^{2/3} \ll (T_1 + T_2)^{1-\varepsilon}$, or $A_1^{2/3}A_2^{4/3} \ll (T_1+T_2)^{1-\varepsilon}$. This means that we may assume

$$X_1, X_2 \ll \frac{N}{(T_1+T_2)^{1-\varepsilon}}. \tag{8.4}$$

Remark. In [BB, Lemma 9], Blomer and Buttcane have shown, using Buttcane's Mellin-Barnes integral representations, that the X_i can be truncated earlier, in the special case $T_1 \asymp T_2 \asymp T$, to $X_i \ll T^{-2+\varepsilon}N$.

Then we have

$$\Big|\sum_{n,m} \overline{a_n} a_m \Sigma_3\Big| \ll \sum_{\epsilon_1,\epsilon_2=\pm1} \sum_{X_1,X_2 \text{ dyadic}} \frac{N^\varepsilon T_1T_2(T_1+T_2)}{X_1X_2} \int_{x_1,x_2,x_3\in V} \frac{dx_1dx_2dx_3}{|V|}|U|, \tag{8.5}$$

where $U = U(\gamma, \epsilon_1, \epsilon_2, X_1, X_2, x_1, x_2, x_3)$ is defined by

$$U = \sum_{\substack{D_1 \asymp X_1 \\ D_2 \asymp X_2}} \sum_{n,m} b_n'' a_m'' S(1, \epsilon_1 m, \epsilon_2 n, 1, D_1, D_2)\gamma_{D_1,D_2} \tag{8.6}$$
$$\Big[\int_0^\infty y_2^{-3i(\tau_1+\tau_2)} e(-\epsilon_2 x_2 y_2) e\Big(-\frac{nD_2}{y_2 D_1^2}\frac{x_1x_3+x_2}{1+x_2^2+x_3^2}\Big) f\Big(\frac{nD_2}{y_2D_1^2}\frac{\xi_1^{1/2}}{\xi_2}\Big) f(y_2)\frac{dy_2}{y_2}\Big]$$
$$\Big[\int_0^\infty y_1^{-3i(\tau_1+\tau_2)} e(-\epsilon_1 x_1 y_1) e\Big(-\frac{mD_1}{y_1 D_2^2}\frac{x_2x_3'+x_1}{1+x_1^2+x_3'^2}\Big) f\Big(\frac{mD_1}{y_1D_2^2}\frac{\xi_2^{1/2}}{\xi_1}\Big) f(y_1)\frac{dy_1}{y_1}\Big],$$

for some sequence γ_{D_1,D_2} with $|\gamma_{D_1,D_2}| \leq 1$. Here we have used the shorthand $a_m'' = a_m' m^{i(2\tau_1+\tau_2)}$, $b_n'' = \overline{a_n'} n^{i(\tau_1+2\tau_2)}$.

Proposition 8.1. *We have*

$$|U| \ll X_1X_2\Big(X_1^2 + \frac{N}{T_1+T_2}\Big)^{1/2}\Big(X_2^2 + \frac{N}{T_1+T_2}\Big)^{1/2}(NT_1T_2)^\varepsilon \sum_{n\leq N}|a_n|^2. \tag{8.7}$$

The estimate is uniform in terms of $\gamma, X_1, X_2, x_1, x_2, x_3$.

Assuming Proposition 8.1, we may quickly show the following variant of Theorem 1.1:

$$\sum_{\substack{\nu_1=iT_1+O(1)\\ \nu_2=iT_2+O(1)}} \frac{1}{R_j}\Big|\sum_{n\le N} a_n\lambda_j(n,1)\Big|^2 \ll \Big(T_1T_2(T_1+T_2)+T_1T_2\frac{N^2}{T_1+T_2}\Big)^{1+\varepsilon}\sum_{n\le N}|a_n|^2, \tag{8.8}$$

as we now explain. By inserting (8.7) into (8.5), we obtain

$$\Big|\sum_{n,m}\overline{a_n}a_m\Sigma_3\Big| \ll \sum_{\substack{X_1,X_2\ll \frac{N}{(T_1+T_2)^{1-\varepsilon}}\\ \text{dyadic}}} \frac{T_1T_2(T_1+T_2)(T_1T_2N)^\varepsilon}{X_1X_2} \tag{8.9}$$
$$X_1X_2\Big(X_1^2+\frac{N}{T_1+T_2}\Big)^{1/2}\Big(X_2^2+\frac{N}{T_1+T_2}\Big)^{1/2}\sum_{n\le N}|a_n|^2,$$

plus a small error term from the truncation on X_1, X_2. By a direct calculation, this gives

$$\Big|\sum_{n,m}\overline{a_n}a_m\Sigma_3\Big| \ll T_1T_2(T_1+T_2)(T_1T_2N)^\varepsilon\Big(\frac{N}{T_1+T_2}\Big)^2\sum_n|a_n|^2, \tag{8.10}$$

which proves (8.8).

Proof of Proposition 8.1. The main difficulty in the proof is exploiting cancellation in the y_1, y_2 integrals. For point of reference, if we apply Corollary 1.3 directly to (8.6), trivially integrating over y_1 and y_2, we obtain

$$|U| \ll X_1X_2(X_1^2+N)^{1/2}(X_2^2+N)^{1/2}(NT_1T_2)^\varepsilon\sum_{n\le N}|a_n|^2. \tag{8.11}$$

One easily observes that if $T_1+T_2 \ll N^\varepsilon$, then (8.11) implies (8.7), so for the rest of the proof we assume

$$T_1+T_2 \gg N^\varepsilon. \tag{8.12}$$

We will first show the bound (8.7) under the assumptions

$$|x_1|, |x_2| \le \delta(T_1+T_2), \tag{8.13}$$

where $\delta > 0$ is some small but fixed number (certainly 1/1000 suffices for the proof). Let

$$Y_1 = \frac{X_2^2}{X_1}\frac{1+x_1^2+x_3'^2}{x_2x_3'+x_1}, \quad \text{and} \quad Y_2 = \frac{X_1^2}{X_2}\frac{1+x_2^2+x_3^2}{x_1x_3+x_2}. \tag{8.14}$$

With this definition, we have that the y_1-integral takes the form

$$\int h(y_1)e^{i\phi_1(y_1)}dy_1, \quad \phi_1(y_1) = c_1(T_1+T_2)\log y_1 + c_2x_1y_1 + c_3\frac{N}{y_1Y_1}, \tag{8.15}$$

where each $c_i \asymp 1$ and h is a weight function with bounded derivatives. Under the assumption (8.13), repeated integration by parts (see [BKY, Lemma 8.1]) shows the integral is smaller than an arbitrarily large negative power of $\max(T_1+T_2, \frac{N}{Y_1})$ (and hence, using (8.12), an arbitrarily large power of T_1T_2N), unless $\frac{N}{Y_1} \asymp (T_1+T_2)$. If (8.13) does not hold then there is potentially cancellation between the first two terms in the phase in which case this argument breaks down.

A similar argument holds for y_2 also. Thus we may assume

$$Y_1, Y_2 \ll \frac{N}{T_1 + T_2}. \tag{8.16}$$

Moving the integrals to the outside, we derive

$$|U| \ll \int_1^2 \int_1^2 \Big| \sum_{\substack{D_1 \asymp X_1 \\ D_2 \asymp X_2}} \sum_{n,m} b_n'' a_m'' S(1, \epsilon_1 m, \epsilon_2 n, 1, D_1, D_2) \gamma_{D_1,D_2} \tag{8.17}$$
$$e\Big(-\frac{nD_2}{y_2 D_1^2} \frac{x_1 x_3 + x_2}{1 + x_2^2 + x_3^2}\Big) f\Big(\frac{nD_2}{y_2 D_1^2} \frac{\xi_1^{1/2}}{\xi_2}\Big) e\Big(-\frac{mD_1}{y_1 D_2^2} \frac{x_2 x_3' + x_1}{1 + x_1^2 + x_3'^2}\Big) f\Big(\frac{mD_1}{y_1 D_2^2} \frac{\xi_2^{1/2}}{\xi_1}\Big) \Big| \frac{dy_1}{y_1} \frac{dy_2}{y_2}.$$

Now we can change variables $y_1 \to y_1^{-1} \frac{D_1}{X_1} \frac{X_2^2}{D_2^2}$, and $y_2 \to y_2^{-1} \frac{D_2}{X_2} \frac{X_1^2}{D_1^2}$, giving that

$$|U| \ll \int_{y_1 \asymp 1} \int_{y_2 \asymp 1} \Big| \sum_{\substack{D_1 \asymp X_1 \\ D_2 \asymp X_2}} \sum_{n,m} b_n'' a_m'' S(1, \epsilon_1 m, \epsilon_2 n, 1, D_1, D_2) \gamma_{D_1,D_2} \tag{8.18}$$
$$e\Big(-\frac{my_1}{Y_1}\Big) e\Big(-\frac{ny_2}{Y_2}\Big) f\Big(\frac{ny_2 X_2}{X_1^2} \frac{\xi_1^{1/2}}{\xi_2}\Big) f\Big(\frac{my_1 X_1}{X_2^2} \frac{\xi_2^{1/2}}{\xi_1}\Big) \Big| \frac{dy_1}{y_1} \frac{dy_2}{y_2}.$$

Now we can apply Mellin inversion to $f(\frac{ny_2 X_2}{X_1^2} \frac{\xi_1^{1/2}}{\xi_2})$ (and the other f), showing now

$$|U| \ll \int_{-\infty}^{\infty} \frac{1}{1 + r_1^2} \int_{-\infty}^{\infty} \frac{1}{1 + r_2^2} \int_{y_1 \asymp 1} \int_{y_2 \asymp 1} \Big| \sum_{\substack{D_1 \asymp X_1 \\ D_2 \asymp X_2}} \gamma_{D_1,D_2} \tag{8.19}$$
$$\sum_{n,m} b_n'' n^{ir_2} a_m'' m^{ir_1} S(1, \epsilon_1 m, \epsilon_2 n, 1, D_1, D_2) e\Big(-\frac{my_1}{Y_1}\Big) e\Big(-\frac{ny_2}{Y_2}\Big) \Big| \frac{dy_1}{y_1} \frac{dy_2}{y_2} dr_1 dr_2.$$

Remark. The r_1 and r_2 integrals are practically harmless because our bound will be in terms of the L^2 norms of the sequences (b_n'') and (a_m''), which are then independent of r_1, r_2.

At this point we can apply Theorem 1.2 (see also Remark 5.3), showing

$$|U| \ll (X_1 X_2)^{1+\varepsilon} \Big[\int_{-\infty}^{\infty} \frac{1}{1 + r_2^2} \tag{8.20}$$
$$\int_{y_2 \asymp 1} \sum_{q \leq \min(X_1, X_2)} \sum_{d_1 | q} \frac{d_1}{q} \sum_{\substack{c \leq \frac{X_1}{q} \\ (c,q)=1}} \sideset{}{^*}\sum_{t \,(\mathrm{mod}\, c)} \Big| \sum_{(n,q)=d_1} b_n'' n^{ir_2} e\Big(\frac{tn}{c}\Big) e\Big(-\frac{ny_2}{Y_2}\Big) \Big|^2 dy_2 dr_2 \Big]^{1/2} [\dots]^{1/2},$$

with $[\dots]$ representing a similar term. Using the hybrid large sieve (Lemma 3.1) shows that the first expression in brackets is bounded by

$$\sum_{q \leq \min(X_1, X_2)} \sum_{d_1 | q} \frac{d_1}{q} \Big(\frac{X_1^2}{q^2} + \frac{Y_2}{d_1}\Big) \sum_{(n,q)=d_1} |\beta_n|^2 \ll (X_1 X_2)^{\varepsilon} \Big(X_1^2 + \frac{N}{T_1 + T_2}\Big) \sum_{n \leq N} |b_n|^2. \tag{8.21}$$

The second expression in brackets in (8.20) is bounded in a similar way, which completes the proof under the assumption (8.13).

Now we show how to modify the proof in case (8.13) does not hold. Say that $|x_1| \geq \delta(T_1+T_2)$, and $|x_2| \leq \delta(T_1+T_2)$. The y_2-analysis is unchanged while in (8.6), we change variables $y_1 \to y_1 \frac{m}{N}$. Following the calculations above, in place of (8.18), we obtain

$$|U| \ll \int_{y_1 \asymp 1} \int_{y_2 \asymp 1} \Big| \sum_{\substack{D_1 \asymp X_1 \\ D_2 \asymp X_2}} \sum_{n,m} b_n'' a_m''' S(1, \epsilon_1 m, \epsilon_2 n, 1, D_1, D_2) \gamma_{D_1,D_2} \\ e\Big(-\frac{m y_1 x_1}{N}\Big) e\Big(-\frac{n y_2}{Y_2}\Big) f\Big(\frac{n y_2 X_2}{X_1^2} \frac{\xi_1^{1/2}}{\xi_2}\Big) f\Big(\frac{m y_1}{N}\Big)\Big| \frac{dy_1}{y_1} \frac{dy_2}{y_2}, \tag{8.22}$$

where $|a_m'''| = |a_m''|$. This has the same essential form as (8.18) but with Y_1 replaced by $\frac{N}{|x_1|} \ll \frac{N}{T_1+T_2}$. Thus we arrive at the same bound in this case. By symmetry, the same bound holds in case $|x_1| \leq \delta(T_1+T_2)$ and $|x_2| \geq \delta(T_1+T_2)$. A simple modification covers the case $|x_1|, |x_2| \geq \delta(T_1+T_2)$, where we apply the change of variables in both y_1, y_2. □

9. Proof of Theorem 1.1

If the X_i are large, then we can obtain an improved version of Proposition 8.1, namely

Proposition 9.1. *We have*

$$|U| \ll \Big[(X_1H_2 + X_2H_1)\Big(X_1^2 + \frac{N}{T_1+T_2}\Big)^{1/2}\Big(X_2^2 + \frac{N}{T_1+T_2}\Big)^{1/2} \\ + \frac{(X_1X_2)^{3/2}N}{H_1} + \frac{(X_1X_2)^{3/2}N}{H_2}\Big](X_1X_2)^{\varepsilon} \sum_{n \leq N} |a_n|^2. \tag{9.1}$$

Here the proof is identical to that of Proposition 8.1 except at (8.20) we apply Theorem 1.4 instead of Theorem 1.2, so we omit the details.

We continue with bounding (8.5). We shall use the bound implied by (8.9) for certain ranges of X_i. Specifically, for the values of X_1 with $X_1^2 \leq \frac{N}{T_1+T_2}$, the bound (8.9) simplifies as

$$\ll \sum_{\substack{X_1^2 \ll \frac{N}{T_1+T_2} \\ X_2 \ll \frac{N}{(T_1+T_2)^{1-\varepsilon}}}} T_1T_2(T_1+T_2)\Big(\frac{N}{T_1+T_2}\Big)^{1/2}\Big(X_2^2 + \frac{N}{T_1+T_2}\Big)^{1/2}(T_1T_2)^{\varepsilon} \sum_{n \leq N} |a_n|^2 \\ \ll T_1T_2 \frac{N^{3/2}}{(T_1+T_2)^{1/2}} (T_1T_2)^{\varepsilon} \sum_{n \leq N} |a_n|^2, \tag{9.2}$$

which is stronger than required for Theorem 1.1. By symmetry, the same bound holds if $X_2^2 \leq \frac{N}{T_1+T_2}$. For the complementary terms with $X_1^2 > \frac{N}{T_1+T_2}$ and $X_2^2 > \frac{N}{T_1+T_2}$, Proposition 9.1 simplifies to give

$$|U| \ll (X_1X_2)\Big[(X_1H_2 + X_2H_1) + \frac{(X_1X_2)^{1/2}N}{H_1} + \frac{(X_1X_2)^{1/2}N}{H_2}\Big](X_1X_2)^{\varepsilon} \sum_{n \leq N} |a_n|^2. \tag{9.3}$$

The optimal choice is $H_1 = N^{1/2}X_1^{1/4}X_2^{-1/4}$, $H_2 = N^{1/2}X_1^{-1/4}X_2^{1/4}$, and gives

$$|U| \ll (X_1X_2)N^{1/2}(X_1^{3/4}X_2^{1/4} + X_1^{1/4}X_2^{3/4})(X_1X_2)^{\varepsilon} \sum_{n \leq N} |a_n|^2. \tag{9.4}$$

The contribution of these terms to (8.9) is then seen to be

$$\ll T_1T_2(T_1+T_2)N^{1/2}\Big(\frac{N}{T_1+T_2}\Big)(T_1T_2)^{\varepsilon}\sum_{n\leq N}|a_n|^2 \ll T_1T_2N^{3/2}(T_1T_2)^{\varepsilon}\sum_{n\leq N}|a_n|^2. \tag{9.5}$$

This is precisely what is required for Theorem 1.1.

References

[Bl] V. Blomer, *Applications of the Kuznetsov formula on $GL(3)$*. Invent. Math. 194 (2013), no. 3, 673–729.

[BB] V. Blomer and J. Buttcane, *On the subconvexity problem for L-functions on $GL(3)$*. Preprint, 2015. http://arxiv.org/abs/1504.02667.

[BBM] V. Blomer, J. Buttcane, and P. Maga, *Applications of the Kuznetsov formula on $GL(3)$: the level aspect.* Preprint, 2014. http://arxiv.org/abs/1410.5106.

[BKY] V. Blomer, R. Khan, and M. Young, *Distribution of mass of holomorphic cusp forms.* Duke Math. J. 162 (2013), no. 14, 2609–2644.

[BFG] D. Bump, S. Friedberg, and D. Goldfeld, *Poincaré series and Kloosterman sums for* $\mathrm{SL}(3,\mathbf{Z})$. Acta Arith. 50 (1988), no. 1, 31–89.

[Bu1] J. Buttcane, *On sums of $SL(3,\mathbb{Z})$ Kloosterman sums.* Ramanujan J. 32 (2013), no. 3, 371–419.

[Bu2] J. Buttcane, *The spectral Kuznetsov formula on $SL(3)$*. Preprint, 2014. http://arxiv.org/abs/1411.7802.

[DF] R. Dąbrowski and B. Fisher, *A stationary phase formula for exponential sums over $\mathbf{Z}/p^m\mathbf{Z}$ and applications to* GL(3)*-Kloosterman sums.* Acta Arith. 80 (1997), no. 1, 1–48.

[DI] J.-M. Deshouillers and H. Iwaniec, *Kloosterman sums and Fourier coefficients of cusp forms.* Invent. Math. 70 (1982/83), no. 2, 219–288.

[DK] W. Duke, and E. Kowalski, *A problem of Linnik for elliptic curves and mean-value estimates for automorphic representations.* With an appendix by Dinakar Ramakrishnan. Invent. Math. 139 (2000), no. 1, 1–39.

[Ga] P. X. Gallagher, *A large sieve density estimate near $\sigma = 1$.* Invent. Math. 11 1970 329–339.

[G] D. Goldfeld, *Automorphic forms and L-functions for the group* $\mathrm{GL}(n,\mathbb{R})$. With an appendix by Kevin A. Broughan. Cambridge Studies in Advanced Mathematics, 99. Cambridge University Press, Cambridge, 2006.

[GK] D. Goldfeld, and A. Kontorovich, *On the* GL(3) *Kuznetsov formula with applications to symmetry types of families of L-functions.* Automorphic representations and L-functions, 263–310, Tata Inst. Fundam. Res. Stud. Math., 22, Tata Inst. Fund. Res., Mumbai, 2013.

[I] H. Iwaniec, *Fourier coefficients of cusp forms and the Riemann zeta-function.* Seminar on Number Theory, 1979–1980, Exp. No. 18, 36 pp., Univ. Bordeaux I, Talence, 1980.

[IK] H. Iwaniec and E. Kowalski, *Analytic number theory.* American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI, 2004.

[J] M. Jutila, *On spectral large sieve inequalities.* Funct. Approx. Comment. Math. 28 (2000), 7–18.

[S] G. Stevens, *Poincaré series on* $\mathrm{GL}(r)$ *and Kloosterman sums.* Math. Ann. 277 (1987), no. 1, 25–51.

[V] A. Venkatesh, *Large sieve inequalities for $GL(n)$-forms in the conductor aspect.* Adv. Math. 200 (2006), no. 2, 336–356.

[Y] M. Young, *The second moment of $GL(3)\times GL(2)$ L-functions at special points.* Math. Ann. 356 (2013), no. 3, 1005–1028.

Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, U.S.A.

E-mail address: myoung@math.tamu.edu