ACTA ARITHMETICA
Online First version

The large sieve for self-dual Eisenstein series of varying levels

by

MATTHEW P. YOUNG (College Station, TX)

*To Henryk Iwaniec, with admiration and gratitude,
on the occasion of his 75th birthday*

Abstract. We prove an essentially optimal large sieve inequality for self-dual Eisenstein series of varying levels. This bound can alternatively be interpreted as a large sieve inequality for rationals ordered by height. The method of proof is recursive, and has some elements in common with Heath-Brown's quadratic large sieve, and the asymptotic large sieve of Conrey, Iwaniec, and Soundararajan.

1. Introduction

1.1. Setting up the problem. A general large sieve inequality is an upper bound on the operator norm of an arithmetically defined matrix $\varLambda = (\lambda_{m,n})$, with $\lambda_{m,n} \in \mathbb{C}$. Define the norm of $\varLambda$, denoted $\|\varLambda\|$, by

$$\|\varLambda\| = \max_{|\boldsymbol{\alpha}|=1} \sum_m \Big| \sum_n \alpha_n \lambda_{m,n} \Big|^2, \qquad \boldsymbol{\alpha} = (\alpha_n).$$

The duality principle implies that $\|\varLambda\| = \|\varLambda^t\|$, where $\varLambda^t = (\lambda_{n,m})$.

A particularly interesting choice of $\lambda_{m,n}$ is $\lambda_f(n)$, where f ranges over a family $\mathcal{F}$ of automorphic forms or L-functions, n ranges over an interval of positive integers, say $N/2 < n \le N$, and $\lambda_f(n)$ is the nth Dirichlet series coefficient of the L-function $L(f,s)$. In this case, we write $\varDelta(\mathcal{F}, N)$ for the norm of this large sieve matrix, namely

$$\varDelta(\mathcal{F}, N) = \max_{|\boldsymbol{\alpha}|=1} \sum_{f \in \mathcal{F}} \Big| \sum_{N/2 < n \le N} \alpha_n \lambda_f(n) \Big|^2. \tag{1.1}$$

2020 *Mathematics Subject Classification*: Primary 11M06; Secondary 11N75.
Key words and phrases: large sieve inequality, Eisenstein series.
Received 13 February 2023; revised 13 December 2023.
Published online 21 February 2024.

DOI: 10.4064/aa230213-1-1

The dual norm $\Delta^*(\mathcal{F}, N)$ is given by

$$\Delta^*(\mathcal{F}, N) = \max_{|\boldsymbol{\beta}|=1} \sum_{N/2<n\leq N} \Big|\sum_{f\in\mathcal{F}} \beta_f \lambda_f(n)\Big|^2. \tag{1.2}$$

The classical multiplicative large sieve inequality concerns the case where $\lambda_f(n) = \chi(n)$, and where the family runs over primitive Dirichlet characters χ modulo q, with $q \leq Q$. Applications include the Bombieri–Vinogradov theorem, estimates for moments of L-functions, zero density estimates, and a variety of sieving problems. See [M] for details.

There are many works on large sieve inequalities for other families. For instance, Deshouillers and Iwaniec [DI] obtained a sharp bound for cusp forms on GL_2, which in turn has been a powerful tool in studying statistical properties of the Riemann zeta function on the critical line. Heath-Brown [H-B] showed an essentially optimal upper bound on the sparse subfamily of quadratic Dirichlet characters. Many state of the art works on quadratic twists of modular forms, with elliptic curves being of particular interest, have relied on Heath-Brown's bound.

In this paper, we are interested in the following family $\mathcal{F}$. For any Dirichlet character ψ modulo r and real number t, define

$$\lambda_{\psi,t}(a,b) = \psi(a)\overline{\psi}(b)(a/b)^{it}.$$

Here $\sum_{ab=n} \lambda_{\psi,t}(a,b) =: \lambda_{\psi,t}(n)$ is the nth Hecke eigenvalue of a self-dual Eisenstein series on $\Gamma_0(r^2)$, and when ψ is primitive, the Eisenstein series is a newform. Let k be a positive integer, and let θ run over all Dirichlet characters modulo k. Let $Q \geq 1$ be a real number, and for each $Q/2 < q \leq Q$ with $(q,k) = 1$, let χ run over primitive Dirichlet characters modulo q. Finally, let $T \geq 1$ be a real number, and let $|t| \leq T$. Then define $\mathcal{F}$ to consist of the characters $\chi\theta$, with corresponding data $\lambda_{\chi\theta,t}(a,b)$, with $N/2 < ab \leq N$ and $(a,b) = 1$. We write

$$\Delta(Q,k,T,N) = \max_{|\boldsymbol{\alpha}|=1} \int_{T/2\leq t\leq T} \sum_{\substack{Q/2<q\leq Q \\ (q,k)=1}} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)} \sum_{\theta \,(\mathrm{mod}\, k)} \Big| \sum_{\substack{N/2<ab\leq N \\ (a,b)=1}} \alpha_{a,b} \lambda_{\chi\theta,t}(a,b) \Big|^2 dt, \tag{1.3}$$

which agrees with $\Delta(\mathcal{F}, N)$ for this family $\mathcal{F}$. The dual norm $\Delta^*(Q,k,T,N)$ is given by

$$\Delta^*(Q,k,T,N) = \max_{|\boldsymbol{\beta}|=1} \sum_{\substack{N/2<ab\leq N \\ (a,b)=1}} \Big| \int_{T/2\leq t\leq T} \sum_{\substack{Q/2<q\leq Q \\ (q,k)=1}} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)} \sum_{\theta \,(\mathrm{mod}\, k)} \beta_{\chi,\theta,t} \lambda_{\chi\theta,t}(a,b) \, dt \Big|^2. \tag{1.4}$$

As a "trivial" bound, which we mainly state for reference, one may deduce from the classical large sieve inequality the bound

$$\mathit{\Delta}(Q,k,T,N) \ll (Q^2 kT\sqrt{N} + N\log N). \tag{1.5}$$

Deducing the estimate (1.5) uses the idea of the Dirichlet hyperbola method, by summing over $a \leq \sqrt{N}$ trivially, and applying the classical large sieve to the sum over $b \ll N/a$.

The condition $(a,b) = 1$ may be easily overlooked, yet it is vital. The above sketch shows that the trivial bound (1.5) holds even without this condition. In fact, if the condition $(a,b) = 1$ were to be omitted in (1.3), then the term of size $Q^2 kT\sqrt{N}$ in (1.5) would not be removable, because one could choose $\alpha_{a,b}$ in (1.3) to be the indicator function of $a = b$. For this, note $\lambda_{\chi,t}(a,a) = 1$ for a coprime to the modulus of χ. Therefore, any substantial improvement over this trivial bound must use the condition $(a,b) = 1$. The restriction $(a,b) = 1$ is similar in spirit to the (necessary) square-free restriction when studying quadratic characters, as in [H-B]; for more on this point, see Section 1.4.1. We also observe that choosing $\alpha_{a,b} = \alpha_{ab}$ to depend only on the product ab would give rise to sums of the form $\sum_n \alpha_n \lambda_{\psi,t}(n)$ appearing in (1.3). Then considering $n = p^2$ would lead to a large term as discussed above.

1.2. Main results, and discussion

THEOREM 1.1. *We have*

$$\mathit{\Delta}(Q,k,T,N) \ll_\varepsilon (QkTN)^\varepsilon (Q^2 kT + N). \tag{1.6}$$

This estimate is optimal (up to the ε-aspect) by general principles (see [IK, Chapter 7]). We may interpret this as a spectral large sieve inequality for the family of trivial nebentypus newform Eisenstein series on $\mathit{\Gamma}_0(q^2k^2)$, with varying level q. Theorem 1.1 appears to be the first sharp large sieve inequality for a GL_2 family with varying levels. The classical large sieve inequality can be interpreted as a GL_1 large sieve inequality, while Heath-Brown's celebrated quadratic large sieve can be viewed as an estimate for the subfamily of self-dual GL_1 forms. The GL_2 families of varying nebentypus do not seem to have strong orthogonality properties, as shown by Iwaniec and Li [IL].

We also have an additive character variant of Theorem 1.1.

THEOREM 1.2. *Define a norm*

$$\mathit{\Delta}_{\mathrm{add.}}(Q,N) = \max_{|\boldsymbol{\alpha}|=1} \sum_{\substack{Q/2<q\leq Q \\ (q,k)=1}} \sideset{}{^*}\sum_{t \,(\mathrm{mod}\, q)} \Big| \sum_{\substack{N/2<ab\leq N \\ (a,b)=1 \\ (ab,q)=1}} \alpha_{a,b} e_q(ta\bar{b}) \Big|^2 \, dt.$$

Then $\mathit{\Delta}_{\mathrm{add.}}(Q,N) \ll (Q^2+N)^{1+\varepsilon}$.

Theorem 1.2 follows quickly from Theorem 1.1, by the method in [IK, Section 7.5]. We have omitted the T- and k-aspects solely to simplify the expressions; hybrid bounds analogous to (1.6) hold for the additive characters as well.

We may interpet Theorem 1.1 as a large sieve inequality for *rationals*, which we now explain. Let v_p be the usual p-adic valuation. For $q \geq 1$, let $\mathbb{Q}_{(q)} = \{x \in \mathbb{Q} : v_p(x) \geq 0 \text{ for all } p \mid q\}$, which is a ring. Indeed, with the multiplicative set S defined by $S = \{n \in \mathbb{Z} : (n, q) = 1\}$, we have $\mathbb{Q}_{(q)} = S^{-1}\mathbb{Z}$, the localization of $\mathbb{Z}$ by S. There exists a natural reduction map $\mathrm{red}_q : \mathbb{Q}_q \to \mathbb{Z}/q\mathbb{Z}$. The reduction map may be restricted to $\mathbb{Q}_{(q)}^{\times} = \{x \in \mathbb{Q} : v_p(x) = 0 \text{ for all } p \mid q\}$, which is a multiplicative subgroup of $\mathbb{Q}^{\times}$. If χ is a Dirichlet character modulo q, and $n \in \mathbb{Q}_{(q)}^{\times}$, then define $\chi(n)$ by $\chi(\mathrm{red}_q(n))$. That is, if $n = a/b \in \mathbb{Q}_{(q)}^{\times}$, then $\chi(n) = \chi(a\overline{b})$. For $n = a/b \in \mathbb{Q}^{\times}$ in lowest terms, define $\mathrm{ht}(n) = |ab|$, which is a cousin of a height function. Note that $|\{n \in \mathbb{Q}^{\times} : \mathrm{ht}(n) \leq X\}| = X^{1+o(1)}$.

Theorem 1.3. *We have*

$$\sum_{q \leq Q} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)} \Big| \sum_{\substack{n \in \mathbb{Q}_{(q)}^{\times} \\ \mathrm{ht}(n) \leq N}} \alpha_n \chi(n) \Big|^2 \ll (Q^2 + N)^{1+\varepsilon} \sum_{\substack{n \in \mathbb{Q}^{\times} \\ \mathrm{ht}(n) \leq N}} |\alpha_n|^2. \tag{1.7}$$

This is simply a restatement of Theorem 1.1 in this notation, with $k = 1$ and the omission of T. These specializations are not necessary, and are only in place to de-clutter the statement.

From Theorem 1.3 one can also easily derive results about rationals ordered by the more standard height function. For $n = a/b \in \mathbb{Q}^{\times}$ in lowest terms, let $\mathrm{Ht}(n) = \max(|a|, |b|)$. Note that $\mathrm{ht}(n) \leq \mathrm{Ht}(n)^2$, from which we immediately deduce:

Corollary 1.4. *We have*

$$\sum_{q \leq Q} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)} \Big| \sum_{\substack{n \in \mathbb{Q}_{(q)}^{\times} \\ \mathrm{Ht}(n) \leq N}} \alpha_n \chi(n) \Big|^2 \ll (Q^2 + N^2)^{1+\varepsilon} \sum_{\substack{n \in \mathbb{Q}^{\times} \\ \mathrm{Ht}(n) \leq N}} |\alpha_n|^2.$$

This is sharp, since $|\{n \in \mathbb{Q}^{\times} : \mathrm{Ht}(n) \leq X\}| = X^{2+o(1)}$. Since Theorem 1.3 easily implies Corollary 1.4, but not vice-versa, this supports our usage of ht in place of Ht.

For $n \in \mathbb{Q}^{\times}$, one may define $\alpha_n = e(n)$, or $\alpha_{a/b} = e_b(\overline{a})$, etc. These examples illustrating Theorem 1.3 are somewhat similar to the quantities studied in [DFI].

The proof of Theorem 1.1 attacks the problem from both sides, via Δ and Δ^*. In this sense, the proof has new features not seen in previous large sieve inequality bounds. Very briefly, the strategy of proof is as follows. If

$N \gg Q^2kT$, then we study the dual norm Δ^* and apply the functional equation of Dirichlet L-functions. The dual side is effective in this range of parameters because the functional equation will shorten the lengths of summation. On the other hand, if $N \ll Q^2kT$, then we more directly study the family average. The main tool on this side is the divisor-switching method used by Conrey–Iwaniec–Soundararajan on the asymptotic large sieve [CIS] (see also [H, p. 210]). On both sides, we derive a recursive bound which relates the norm to itself, but with different (smaller) parameters.

When $N \approx Q^2kT$, then both methods are essentially circular. The key to breaking out of this deadlock is to use monotonicity, lengthening one of the sums. The use of the functional equation and monotonicity were both crucial tools in Heath-Brown's quadratic large sieve. A major difference between our method and Heath-Brown's is that in the quadratic case, the norm was almost self-dual by quadratic reciprocity. This property completely fails in our situation.

We now discuss the two main workhorse results used to prove Theorem 1.1, both of which require defining some variants on Δ. Let

$$\Delta'(Q,k,T,N) = \max_{\substack{X,R,U,C\in\mathbb{R}_{\geq 1},\, \ell\in\mathbb{Z}_{>0} \\ XR^2\ell U\leq Q^2kT \\ X\leq C}} X\Delta\Big(R,\ell,U,\frac{N}{C}\Big). \tag{1.8}$$

Note that trivially $\Delta(Q,k,T,N) \leq \Delta'(Q,k,T,N)$, by taking $X=1$, $R=Q$, $\ell=k$, $U=T$, $C=1$. Theorem 1.1 will show these norms are essentially of the same order of magnitude. On a first pass, the reader is encouraged to think of $\Delta'(Q,k,T,N)$ as $\Delta(Q,k,T,N)$ itself. Another notational convenience is to write

$$\overline{\Delta}(Q,k,T,N) = \max_{\substack{Q\leq R\leq Q(Q^2kTN)^\varepsilon \\ T\leq U\leq T(Q^2kTN)^\varepsilon \\ N\leq M\leq N(Q^2kTN)^\varepsilon}} \Delta(R,k,U,M), \tag{1.9}$$

and similarly for other norms, such as $\overline{\Delta'}$. In practice, the choices of ε will be either unimportant, or apparent from the context, and no confusion should arise from suppressing them on the left hand side of (1.9).

Theorem 1.5 (Recursive functional equation). *If $N \gg Q^2kT(QkT)^{-\varepsilon}$, then*

$$\Delta(Q,k,T,N) \ll (QkTN)^\varepsilon\Big[N + \frac{N}{Q^2kT}\,\overline{\Delta'}\Big(Q,k,T,\frac{Q^4k^2T^2}{N}\Big)\Big]. \tag{1.10}$$

We also derive a recursive bound on Δ by the family average approach.

THEOREM 1.6 (Recursive family average). *If $Q^2kT \gg N(QkT)^{-\varepsilon}$, then*

$$\Delta(Q,k,T,N) \ll (QkTN)^{\varepsilon}\Big[Q^2kT + \frac{Q^2kT}{N}\,\overline{\Delta'}\Big(\frac{N}{kQT},k,T,N\Big)\Big]. \tag{1.11}$$

The proofs of Theorems 1.5 and 1.6, appearing in Sections 4 and 5, respectively, are logically independent, and can be read in either order. Although very different in the fine details, the two proofs have important structural similarities. Because of the logical independence of these two sections, and due to the strong analogies, we have deliberately chosen to "refresh" notation when passing from Section 4 to Section 5. Even more, we have structured the proofs in a similar way, and chosen notation to help draw the reader's attention to analogous quantities in the two proofs.

Our main interest in Theorem 1.1 is with $k = T = 1$. However, the recursive nature of the proofs and the appearance of the generalized norm Δ' in Theorems 1.5 and 1.6 force us to consider more general values of k and T.

1.3. Applications. The classical large sieve has a wealth of important applications, and we consider some variants for the new rational large sieve (Theorem 1.1). The literature in analytic number theory on sieving problems for the rational numbers is relatively sparse. The authors of [EEHK, Z] give versions of Gallagher's larger sieve for rationals, and deduce some impressive algebraic applications. More applications could be of great interest.

Consider the following sieving problem. Let $\mathcal{N} = \{n \in \mathbb{Q}_{>0} : \mathrm{ht}(n) \le N\}$. Let $\mathcal{P}$ be a finite set of prime numbers. For each $p \in \mathcal{P}$, let $\Omega_p \subset \mathbb{Z}/p\mathbb{Z}$. Define the sifted set

$$\mathcal{S}(\mathcal{N},\mathcal{P},\Omega) = \{n \in \mathcal{N} : \text{for all } p \text{ with } v_p(n) = 0,\ \mathrm{red}_p(n) \notin \Omega_p\}.$$

Note that if $v_p(n) \neq 0$, then $\mathrm{red}_p(n)$ may not be defined. Let $\omega(p) = |\Omega_p|$, and suppose that $\omega(p) < p$ for all $p \in \mathcal{P}$. Let $h(p) = \frac{\omega(p)}{p-\omega(p)}$ for $p \in \mathcal{P}$, and $h(p) = 0$ for $p \notin \mathcal{P}$, and extend h multiplicatively on the square-free integers. Define $H = \sum_{q \le Q} h(q)$.

PROPOSITION 1.7. *With the above notation, we have*

$$|\mathcal{S}(\mathcal{N},\mathcal{P},\Omega)| \ll \frac{(N+Q^2)^{1+\varepsilon}}{H}.$$

One can prove this following the method of [IK, Theorem 7.14]. Alternatively, see [K, Proposition 2.3] for a proof in much greater generality. For a nontrivial result, one needs $H \gg N^{\varepsilon}$, which is more restrictive than in the classical arithmetic large sieve.

A standard application of the classical large sieve is to let Ω_p consist of $\frac{p-1}{2}$ residue classes chosen arbitrarily, for all $p \le Q$. Then $H \gg Q$, and taking $Q = \sqrt{N}$ gives $|\mathcal{S}(\mathcal{N},\mathcal{P},\Omega)| \ll N^{1/2+\varepsilon}$.

We also present a Barban–Davenport–Halberstam type theorem. Suppose that α_n is a sequence supported on $\mathbb{Q}_{>0}$, with $\mathrm{ht}(n) \leq X$. We assume a weak Siegel–Walfisz (S–W) type condition for the sequence, as follows. Define

$$S(X,\chi) = \sum_{\mathrm{ht}(n)\leq X} \alpha_n \chi(n).$$

For $\chi = \chi'\chi_0$ with χ' of conductor $r > 1$, and χ_0 trivial modulo s, we assume

$$|S(X,\chi)| \ll_{B,k} |\alpha| \frac{X^{1/2}\tau_k(s)}{(\log X)^B} \tag{1.12}$$

for some k-fold divisor function τ_k and all $r \leq (\log X)^B$.

PROPOSITION 1.8. *Suppose that α satisfies the S–W condition* (1.12) *for any $B > 0$. Then*

$$\sum_{q\leq Q} \sideset{}{^*}\sum_{a \,(\mathrm{mod}\, q)} \Bigg| \sum_{\substack{\mathrm{ht}(n)\leq X \\ n\equiv a \,(\mathrm{mod}\, q)}} \alpha_n - \frac{1}{\varphi(q)} \sum_{\substack{\mathrm{ht}(n)\leq X \\ (n,q)=1}} \alpha_n \Bigg|^2 \ll \frac{X|\alpha|^2}{(\log X)^A}$$

for any $A > 0$, provided $Q \ll X^{1-\varepsilon}$.

We prove Proposition 1.8 in Section 3.

1.4. Proof sketches. Here we present some overly-simplified outlines of the proofs. In this section we freely drop factors of size $(Q^2kTN)^\varepsilon$, as if they were 1.

1.4.1. *Theorem 1.5.* For simplicity, we omit the t-aspect, and we write $\Delta(Q,k,N)$ for the norm. For a bump function w supported on $[1/2, 2]$, consider

$$S = \sum_{(a,b)=1} w\left(\frac{ab}{N}\right) |T(a,b)|^2, \quad \text{where} \quad T(a,b) = \sum_{q,\chi,\theta} \beta_{\chi,\theta} \chi\theta(a\bar{b}).$$

The condition $(a,b) = 1$ is necessary but difficult to use. In comparison to the quadratic large sieve, this condition is analogous to the restriction to fundamental discriminants. Inspired by this similarity, and following [H-B], let $1 \leq Y < N/10$ to be chosen later, and note $S \leq S_{>Y}$, where

$$S_{>Y} = \sum_{ab/(a,b)^2 > Y} w\left(\frac{ab}{N}\right) |T(a,b)|^2.$$

We then write $S_{>Y} = S_\infty - S_{\leq Y}$, where $S_{\leq Y}$ has $ab/(a,b)^2 \leq Y$, and S_∞ has a and b unconstrained. These two sums are treated in completely different ways. For $S_{\leq Y}$, let $g = (a,b)$ and change variables $a \mapsto ga$ and $b \mapsto gb$.

Ignoring coprimality issues, we get $T(ga,gb) \approx T(a,b)$, and so

$$\begin{aligned} S_{\leq Y} &\approx \sum_{\substack{ab \leq Y \\ (a,b)=1}} \sum_g w\left(\frac{g^2 ab}{N}\right) |T(a,b)|^2 \\ &= \int_{(2)} \widetilde{w}(s) \zeta(2s) \sum_{\substack{ab \leq Y \\ (a,b)=1}} \left(\frac{N}{ab}\right)^s |T(a,b)|^2 \, \frac{ds}{2\pi i}. \end{aligned}$$

Next, shift contours to the line ε, passing a pole at $s = 1/2$. The contribution to $S_{\leq Y}$ from the new contour is essentially $\ll N^\varepsilon \Delta(Q,k,Y)|\beta|^2$. The pole at $s = 1/2$ gives

$$\frac{1}{2} \widetilde{w}(1/2) \sum_{\substack{ab \leq Y \\ (a,b)=1}} \left(\frac{N}{ab}\right)^{1/2} |T(a,b)|^2. \tag{1.13}$$

This term is *not* satisfactorily bounded on its own. Indeed, even if we accept Theorem 1.1, then by breaking up into dyadic segments $M/2 < ab \leq M$, with $1 \leq M \leq Y$, we can at best bound (1.13) by

$$\max_{1 \leq M \leq Y} \left(\frac{N}{M}\right)^{1/2} (Q^2 k + M)|\beta|^2 \ll (Q^2 k \sqrt{N} + N^{1/2} Y^{1/2})|\beta|^2.$$

The former term of size $Q^2 k \sqrt{N}$ is the culprit, and matches with (1.5). Luckily, and crucially, the term (1.13) will partially cancel with another term from S_∞. This cancellation property also appeared in [H-B].

Next, consider S_∞. Opening $|T(a,b)|$ and applying the Mellin inversion formula gives

$$S_\infty = \sum_{q_1,q_2,\chi_1,\chi_2,\theta_1,\theta_2} \beta_1 \overline{\beta_2} \int_{(2)} \widetilde{w}(s) N^s L(s,\Phi) L(s,\overline{\Phi}) \, \frac{ds}{2\pi i},$$

where $\Phi = \chi_1 \overline{\chi_2} \theta_1 \overline{\theta_2}$. Unfortunately, Φ may not be primitive, and this complicates the application of the functional equation. For this sketch, we consider the two extremes, where either Φ is primitive of conductor $q_1 q_2 k$, or Φ is trivial. The trivial case is easy to control, since this means $\chi_1 = \chi_2$ (whence $q_1 = q_2$) and $\theta_1 = \theta_2$. This gives rise to a diagonal term of acceptable size $O(N|\beta|^2)$. For the primitive characters, we shift contours to the line -1, change variables $s \mapsto 1 - s$, and apply the functional equation. This gives (roughly)

$$\sum_{q_1,q_2,\chi_1,\chi_2,\theta_1,\theta_2} \beta_1 \overline{\beta_2} \int_{(2)} \widetilde{w}(1-s) \frac{(q_1 q_2 k)^{2s-1}}{N^{s-1}} \, \frac{\gamma(s)}{\gamma(1-s)} L(s,\Phi) L(s,\overline{\Phi}) \, \frac{ds}{2\pi i},$$

where $\gamma(s)$ is the product of gamma factors in the completed L-function of

$L(s,\Phi)L(s,\overline{\Phi})$. Next re-open the Dirichlet series and rearrange, which gives

$$\sum_{a,b}\int_{(2)}\widetilde{w}(1-s)\frac{\gamma(s)}{\gamma(1-s)}\sum_{q_1,q_2,\chi_1,\chi_2,\theta_1,\theta_2}\beta_1\overline{\beta_2}\frac{(q_1q_2k)^{2s-1}}{(ab)^s N^{s-1}}\chi_1\overline{\chi_2}\theta_1\overline{\theta_2}(a\bar{b})\,\frac{ds}{2\pi i}.$$

Letting $g=(a,b)$, replacing a by ga and b by gb, and summing over g, we obtain

$$\sum_{\substack{ab\le Q^4k^2/N\\(a,b)=1}}\int_{(2)}\widetilde{w}(1-s)\frac{\gamma(s)}{\gamma(1-s)}\zeta(2s)$$
$$\times\sum_{q_1,q_2,\chi_1,\chi_2,\theta_1,\theta_2}\beta_1\overline{\beta_2}\frac{(q_1q_2k)^{2s-1}}{(ab)^s N^{s-1}}\chi_1\overline{\chi_2}\theta_1\overline{\theta_2}(a\bar{b})\,\frac{ds}{2\pi i},$$

as the sum can be truncated at $ab\le Q^4k^2/N$ (by shifting the contour far to the right). Next, we shift contours back to the line ε, crossing a pole at $s=1/2$. This polar term has a nice simplification, and takes the same form as (1.13), but with ab truncated at Q^4k^2/N instead of Y. Taking $Y=Q^4k^2/N$ then causes these two polar terms to cancel! The contribution on the line ε essentially becomes bounded by $\frac{N}{Q^2k}\Delta(Q,k,Q^4k^2/N)$, which agrees with Theorem 1.5.

1.4.2. *Theorem 1.6.* For simplicity, take $k=1$ and omit t, and write $\Delta(Q,N)$ for the norm. For a bump function w, let

$$S=\sum_q w(q/Q)\sideset{}{^*}\sum_{\chi\,(\mathrm{mod}\,q)}|T(\chi)|^2,\qquad T(\chi)=\sum_{\substack{N/2<ab\le N\\(a,b)=1}}\alpha_{a,b}\chi(a\bar{b}).$$

The condition that χ is primitive is necessary but difficult to use. In analogy with the proof of Theorem 1.5, let $Y<Q/10$, and define

$$S_{>Y}=\sum_q w(q/Q)\sideset{}{^*}\sum_{\substack{\chi\,(\mathrm{mod}\,q)\\\mathrm{cond}(\chi)>Y}}|T(\chi)|^2.$$

Then $S\le S_{>Y}$, by positivity. Again, write $S=S_\infty-S_{\le Y}$ where $S_{\le Y}$ has characters modulo q with $\mathrm{cond}(\chi)\le Y$ and S_∞ has χ ranging over all characters of modulus q.

For $S_{\le Y}$, replace q by qq_0 and χ by $\chi\chi_0$ where (the new) χ has conductor q, and χ_0 is trivial. Ignoring coprimality, we have $T(\chi\chi_0,t)\approx T(\chi,t)$. Applying Mellin inversion, and summing over q_0 to form a zeta function, we obtain

$$S_{\le Y}\approx\sum_{q\le Y}\int_{(2)}\widetilde{w}(s)\Big(\frac{Q}{q}\Big)^s\zeta(s)\sideset{}{^*}\sum_{\chi\,(\mathrm{mod}\,q)}|T(\chi)|^2\,\frac{ds}{2\pi i}.$$

We shift contours to the line ε, passing a pole at $s=1$ only. This polar term takes the form

$$Q\widetilde{w}(1)\sum_{q\le Y} q^{-1} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)} |T(\chi)|^2. \tag{1.14}$$

On the new line ε, we essentially obtain an expression of size $\Delta(Y,N)|\beta|^2$. This polar term is the analog of (1.13), and as before, it is not satisfactorily bounded on its own. Indeed, Theorem 1.1 would imply that at best (1.14) is bounded by

$$Q\max_{R\le Y} R^{-1}(R^2+N)|\alpha|^2 = (QY+QN)|\alpha|^2.$$

Here the term QN is the culprit, and as before, we will cancel this polar term with one arising within S_∞.

Now consider S_∞. Opening the square and applying orthogonality of characters gives

$$S_\infty \approx Q\sum_q w_1(q/Q) \sum_{\substack{(a_1,b_1)=(a_2,b_2)=1\\ a_1b_2\equiv a_2b_1 \,(\mathrm{mod}\, q)}} \alpha_{a_1,b_1}\overline{\alpha_{a_2,b_2}},$$

where $w_1(x)=xw(x)$. The range of possible values of $\gcd(a_1b_2,a_2b_1)$ causes some arithmetical difficulties. For this sketch, we consider the two extreme cases, where either they are coprime, or $a_1b_2=a_2b_1$, which we call the diagonal case. Since $(a_1,b_1)=(a_2,b_2)=1$, the diagonal reduces to $a_1=a_2$ and $b_1=b_2$, giving a term of size $O(Q^2|\alpha|^2)$, which is acceptable.

We now focus on the case $(a_1b_2,a_2b_1)=1$. Write $a_1b_2=a_2b_1+qr$, which we now interpret as $a_1b_2\equiv a_2b_1 \pmod r$, with $q=(a_1b_2-a_2b_1)/r$. Note that typically $r\ll N/Q$, so this reduces the modulus when $Q^2\gg N$. This leads to

$$S_\infty \approx Q\sum_r \sum_{\substack{(a_1,b_1)=(a_2,b_2)=1\\ a_1b_2\equiv a_2b_1 \,(\mathrm{mod}\, r)}} w_1\Big(\frac{a_1b_2-a_2b_1}{Qr}\Big)\alpha_{a_1,b_1}\overline{\alpha_{a_2,b_2}}.$$

Next, we detect the congruence with characters modulo r, as in [CIS], giving

$$\begin{aligned} S_\infty \approx Q\sum_r \sum_{\chi\,(\mathrm{mod}\, r)} r^{-1} \\ \times \sum_{(a_1,b_1)=(a_2,b_2)=1} w_1\Big(\frac{a_1b_2-a_2b_1}{Qr}\Big)\alpha_{a_1,b_1}\overline{\alpha_{a_2,b_2}}\chi(a_1b_2\overline{a_2b_1}).\end{aligned}$$

Since the characters are not primitive, replace χ by $\chi\chi_0$ and r by rr_0 where the new χ has conductor r, and χ_0 is trivial modulo r_0. Applying Mellin inversion, and evaluating the r_0-sum in terms of a zeta function, we deduce that S_∞ is roughly

$$Q \int_{(1)} \widetilde{w_1}(-s) \sum_{r \leq N/Q} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, r)} r^{-1} \\ \times \sum_{\substack{(a_1,b_1)=1 \\ (a_2,b_2)=1}} \left(\frac{a_1 b_2 - a_2 b_1}{Qr}\right)^s \zeta(s+1) \alpha_{a_1,b_1} \overline{\alpha_{a_2,b_2}} \chi(a_1 b_2 \overline{a_2 b_1}) \,\frac{ds}{2\pi i}.$$

Next, we shift contours to the line $-1+\varepsilon$, passing a pole at $s=0$ only. Note that $\widetilde{w_1}(0) = \widetilde{w}(1)$. This polar term nicely simplifies, and takes the same form as (1.14), but with r truncated at N/Q instead of Y. Taking $Y = N/Q$ causes the two polar terms to cancel. Next consider the integral along the line $-1+\varepsilon$. The variables a_i, b_i are not separated, but one might hope that this is only a technical issue solvable with integral transform techniques (indeed, see Lemma 5.2). We might then expect the contribution from the new line of integration to be bounded by $\frac{Q^2}{N} \Delta(N/Q, N)|\alpha|^2$, which is consistent with Theorem 1.6.

The wealth of extra parameters in the definition of Δ' in (1.8) are there to account for the overlooked conditions (both arithmetical and archimedean).

1.4.3. *Reflections.* The similarities between the proofs are remarkable, even if the fine details are different. We also observe that the divisor-switching method used in the proof of Theorem 1.6 is analogous to the functional equation of the Dirichlet L-functions used for Theorem 1.5. At the cost of some exaggeration, one might call the divisor switch itself a functional equation. In support of this, consider the family of functions $\tau_s(n) = \sum_{ab=n} (a/b)^s$, which does indeed satisfy the functional equation $\tau_{-s}(n) = \tau_s(n)$, by the divisor switch. Moreover, they appear as Fourier coefficients of the level 1 Eisenstein series, and the functional equation of the Eisenstein series is entwined with the functional equation of its Fourier coefficients.

1.4.4. *Theorem 1.1.* Theorem 1.1 is deduced from Theorems 1.5 and 1.6 in Section 2. The proof uses the fact that the norm Δ is monotonic, and applies the two self-referential theorems in a recursive manner. In retrospect, some of these ideas have similarities to elements used in [BI1, BI2].

1.5. Notation and conventions. Let $\Gamma_{\mathbb{R}}(s) = \pi^{-s/2}\Gamma(s/2)$. If χ is a Dirichlet character and $a/b \in \mathbb{Q}$ in lowest terms, we may interchangeably write

$$\chi(a)\overline{\chi}(b) = \chi(a\overline{b}) = \chi(a/b). \tag{1.15}$$

We use the notation $A \lesssim B$ as a synonym for

$$A \leq C(\varepsilon)(Q^2 kTN)^{\varepsilon} B. \tag{1.16}$$

2. Deduction of Theorem 1.1. In this section, we use Theorems 1.5 and 1.6 to prove Theorem 1.1.

2.1. Monotonicity. As in the quadratic large sieve [H-B], it is vital that the norm $\Delta(Q,k,T,N)$ is essentially monotonic in the N- and Q-components. The proofs differ a bit depending on the case, but the overall theme is similar, and based on an idea of Forti and Viola [FV].

LEMMA 2.1. *Suppose $P \gg \log(QN)$ with a large (but absolute) implied constant. Then there exists a prime $p \in [P, 2P]$ such that*

$$\Delta(Q,k,T,N) \le 8\Delta(Q,k,T,Np).$$

Proof. Since k and T are frozen, we suppress them from the discussion, writing $\Delta(Q,N)$ in place of $\Delta(Q,k,T,N)$. Let $\gamma_{a,b}$ be complex numbers supported on $N/2 < ab \le N$, and $(a,b) = 1$. Let $P \ge 1$ be a parameter to be chosen, and let P^* denote the number of primes $p \in [P, 2P]$. The prime number theorem implies $P^* \sim \frac{P}{\log P}$. Now we have

$$\sum_{q,\chi}\Big|\sum_{\substack{(a,b)=1\\ N/2<ab\le N}}\gamma_{a,b}\chi(a)\overline{\chi}(b)\Big|^2 = \sum_{q,\chi}\frac{1}{P^*}\sum_{P\le p\le 2P}\Big|\sum_{\substack{(a,b)=1\\ N/2<ab\le N}}\gamma_{a,b}\chi(a)\overline{\chi}(b)\Big|^2$$
$$= \sum_{q,\chi}\frac{1}{P^*}\Big(\sum_{\substack{P\le p\le 2P\\ p\nmid q}} + \sum_{\substack{P\le p\le 2P\\ p\mid q}}\Big)\Big|\sum_{\substack{(a,b)=1\\ N/2<ab\le N}}\gamma_{a,b}\chi(a)\overline{\chi}(b)\Big|^2.$$

For the terms with $p \mid q$, we simply use

$$\frac{1}{P^*}\sum_{\substack{P\le p\le 2P\\ p\mid q}} 1 \le \frac{\log Q}{P^*\log P}.$$

Taking $P \gg \log Q$ large enough so that $P^*\log P \ge 2\log Q$, and rearranging, we obtain

$$\Delta(Q,N) \le \max_{\gamma\ne 0}\frac{2}{|\gamma|^2}\sum_{q,\chi}\frac{1}{P^*}\sum_{\substack{P\le p\le 2P\\ p\nmid q}}\Big|\sum_{\substack{(a,b)=1\\ N/2<ab\le N}}\gamma_{a,b}\chi(a)\overline{\chi}(b)\Big|^2.$$

Next we separate the values of a and b to make two subsums corresponding to $(p,ab) = 1$ and $p \mid ab$. This gives

$$\Delta(Q,N) \le \max_{\gamma\ne 0}\frac{4}{|\gamma|^2}\sum_{q,\chi}\frac{1}{P^*}\sum_{\substack{P\le p\le 2P\\ p\nmid q}}\Big(\Big|\sum_{(ab,p)=1}\Big|^2 + \Big|\sum_{p\mid ab}\Big|^2\Big).$$

We bound the terms with $p \mid ab$ similarly to the treatment of $p \mid q$, which gives

$$\max_{\gamma\ne 0}\frac{4}{|\gamma|^2}\sum_{q,\chi}\frac{1}{P^*}\sum_{\substack{P\le p\le 2P\\ p\nmid q}}\Big|\sum_{p\mid ab}\Big|^2 \le \max_{\gamma\ne 0}\frac{4}{|\gamma|^2 P^*}\sum_{P\le p\le 2P}\Delta(Q,N)\sum_{p\mid ab}|\gamma_{a,b}|^2$$
$$\le \frac{4\log N}{P^*\log P}\Delta(Q,N).$$

We choose $P \gg \log N$ large enough so that $\frac{4\log N}{P^* \log P} \le \frac{1}{2}$, whence

$$\Delta(Q,N) \le \max_{\gamma \neq 0} \frac{8}{|\gamma|^2} \sum_{q,\chi} \frac{1}{P^*} \sum_{\substack{P \le p \le 2P \\ p \nmid q}} \Big| \sum_{\substack{(a,b)=1 \\ (ab,p)=1}} \gamma_{a,b} \chi(a\bar{b}) \Big|^2.$$

Now we freely multiply by $|\chi(p)|^2$, which has absolute value 1 since $p \nmid q$. In addition, we change variables $A = ap$, let $\delta_{A,b} = \gamma_{A/p,b}$, make note that $Np/2 < Ab \le Np$, $|\delta| = |\gamma|$, and $(A,b) = 1$. Thus

$$\Delta(Q,N) \le \frac{8}{P^*} \sum_{P \le p \le 2P} \Delta(Q,Np) \le 8 \max_{P \le p \le 2P} \Delta(Q,Np). \ \blacksquare$$

Lemma 2.2. *Suppose $P \gg \log(NQ)$ with a large (but absolute) implied constant. Then there exists a prime $p \in [P,2P]$ such that*

$$\Delta(Q,k,T,N) \le 8\Delta(Qp,k,T,N).$$

Proof. Since k and T are frozen, we suppress them in the notation. Let $P \ge 10$ to be chosen, and let $P^{**} = \sum_{P \le p \le 2P} \sum^*_{\psi \,(\mathrm{mod}\, p)} \sum 1$, so $P^{**} \asymp \frac{P^2}{\log P}$. We have

$$\sum_{\substack{(a,b)=1 \\ N/2 < ab \le N}} \Big| \sum_{q,\chi} \beta_\chi \chi(a)\overline{\chi}(b) \Big|^2 = \sum_{\substack{(a,b)=1 \\ N/2 < ab \le N}} \frac{1}{P^{**}} \sum_{P \le p \le 2P} \sideset{}{^*}\sum_{\psi \,(\mathrm{mod}\, p)} \Big| \sum_{q,\chi} \beta_\chi \chi(a)\overline{\chi}(b) \Big|^2$$
$$= \sum_{\substack{(a,b)=1 \\ N/2 < ab \le N}} \frac{1}{P^{**}} \Big(\sum_{\substack{p,\psi \\ (p,ab)=1}} + \sum_{\substack{p,\psi \\ p \mid ab}} \Big) \Big| \sum_{q,\chi} \beta_\chi \chi(a)\overline{\chi}(b) \Big|^2.$$

For the terms with $p \mid ab$, we simply use

$$\frac{1}{P^{**}} \sum_{\substack{p,\psi \\ p \mid ab}} 1 \le \frac{2P \log N}{P^{**} \log P},$$

and choose $P \gg \log N$ large enough so that $\frac{2P\log N}{P^{**}\log P} \le \frac{1}{2}$. For the terms with $p \nmid ab$, we freely multiply by $|\psi(a)\overline{\psi}(b)|^2$, which is 1 for such primes. This gives

$$\Delta(Q,N) \le \max_{\beta \neq 0} \frac{2}{|\beta|^2} \sum_{\substack{(a,b)=1 \\ N/2 < ab \le N}} \frac{1}{P^{**}} \sum_{p,\psi} \Big| \sum_{q,\chi} \beta_\chi \chi\psi(a)\overline{\chi\psi}(b) \Big|^2.$$

Next we separate the values of q to make two subsums corresponding to $(p,q) = 1$ and $p \mid q$. This gives

$$\Delta(Q,N) \le \max_{\beta \neq 0} \frac{4}{|\beta|^2} \sum_{\substack{(a,b)=1 \\ N/2 < ab \le N}} \frac{1}{P^{**}} \sum_{p,\psi} \Big(\Big| \sum_{\substack{q,\chi \\ (q,p)=1}} \Big|^2 + \Big| \sum_{\substack{q,\chi \\ p \mid q}} \Big|^2 \Big).$$

We upper bound the terms with $p \mid q$, which gives

$$\sum_{\substack{(a,b)=1 \\ N/2<ab\le N}} \frac{4}{P^{**}} \sum_{p,\psi} \Big| \sum_{\substack{q,\chi \\ p|q}} \Big|^2 \le \frac{4}{P^{**}} \sum_{p,\psi} \Delta(Q,N) \sum_{\substack{q,\chi \\ p|q}} |\beta_\chi|^2$$

$$\le \frac{4P\log Q}{P^{**}\log P} \Delta(Q,N)|\beta|^2.$$

We choose $P \gg \log Q$ large enough so that $\frac{4P\log Q}{P^{**}\log P} \le \frac{1}{2}$, whence

$$\Delta(Q,N) \le \max_{\beta\ne 0} \frac{8}{|\beta|^2} \sum_{\substack{(a,b)=1 \\ N/2<ab\le N}} \frac{1}{P^{**}} \sum_{p,\psi} \Big| \sum_{\substack{q,\chi \\ (q,p)=1}} \beta_\chi \chi\psi(a\bar{b}) \Big|^2.$$

Now $\chi\psi$ is a character of conductor pq, with $pQ/2 \le pq \le pQ$, so we obtain

$$\Delta(Q,N) \le \frac{8}{P^{**}} \sum_{p,\psi} \Delta(pQ,N) \le 8 \max_{P\le p\le 2P} \Delta(pQ,N). \ \blacksquare$$

Remark. The norm Δ is also monotonic in the k- and T-aspects, but this property is not needed in this work, so we do not give proofs.

2.2. Relations between norms. To simplify the recursive steps in the proof of Theorem 1.1, it is convenient to have the following relations. Proofs follow from the definitions (1.8) and (1.9).

Lemma 2.3. *Suppose that there exists $e > 1$ such that*

$$\Delta(Q,k,T,N) \lesssim Q^2kT + N^e$$

for all Q, k, T, N. Then for all Q, k, T, N we have

$$\overline{\Delta'}(Q,k,T,N) \lesssim Q^2kT + N^e.$$

Lemma 2.4. *Suppose that there exists $e > 1$ such that*

$$\Delta(Q,k,T,N) \lesssim (Q^2kT)^e + N$$

for all Q, k, T, N. Then for all Q, k, T, N we have

$$\overline{\Delta'}(Q,k,T,N) \lesssim (Q^2kT)^e + N.$$

2.3. The recursions

Proposition 2.5. *Suppose that there exists $e > 1$ such that*

$$\Delta(Q,k,T,N) \lesssim Q^2kT + N^e \tag{2.1}$$

for all Q,k,T,N. Then, with $e' = 2 - \frac{1}{e}$, for all Q,k,T,N we have

$$\Delta(Q,k,T,N) \lesssim (Q^2kT)^{e'} + N.$$

Proof. Let $F = Q^2kT$, which is the size of the family. By monotonicity (Lemma 2.1), we have $\Delta(Q,k,T,N) \ll \Delta(Q,k,T,N_1)$ for $N_1 \gg N\log(FN)$.

Let $N_1 \asymp N \log N + F^\alpha$ for some $\alpha > 1$, so that $F \ll N_1$. By Theorem 1.5,

$$\Delta(Q,k,T,N) \ll \Delta(Q,k,T,N_1) \lesssim N_1 + \frac{N_1}{F}\,\overline{\Delta'}\Big(Q,k,T,\frac{F^2}{N_1}\Big).$$

By Lemma 2.3, we can use the assumption (2.1) to obtain

$$\begin{aligned}\Delta(Q,k,T,N) &\lesssim N_1 + \frac{N_1}{F}\bigg(F + \Big(\frac{F^2}{N_1}\Big)^e\bigg) \ll N_1 + \frac{F^{2e-1}}{N_1^{e-1}} \\ &\lesssim N + F^\alpha + F^{2e-1-\alpha(e-1)}.\end{aligned}$$

We choose α optimally so that $\alpha = 2e-1-\alpha(e-1)$, which simplifies as $\alpha = 2 - 1/e$. Since $e > 1$ by assumption, this means $\alpha > 1$. ■

We also have a complementary version:

PROPOSITION 2.6. *Suppose that there exists $e > 1$ such that*

$$\Delta(Q,k,T,N) \lesssim (Q^2kT)^e + N \tag{2.2}$$

for all Q,k,T,N. Then, with $e' = 2-1/e$, for all Q,k,T,N we have

$$\Delta(Q,k,T,N) \lesssim Q^2kT + N^{e'}.$$

Proof. Let $F = Q^2kT$. By monotonicity (Lemma 2.2), we deduce that $\Delta(Q,k,T,N) \ll \Delta(Q_1,k,T,N)$ for $Q_1 \gg Q\log(FN)$. We take $F_1 := Q_1^2kT \asymp Q^2kT\log^2(FN) + N^\alpha$ for some $\alpha > 1$, so that $N \ll Q_1^2kT$. By Theorem 1.6, we have

$$\Delta(Q,k,T,N) \ll \Delta(Q_1,k,T,N) \lesssim F_1 + \frac{F_1}{N}\,\overline{\Delta'}\Big(\frac{N}{kQ_1T},k,T,N\Big).$$

By Lemma 2.4, we can use the assumption (2.2) to obtain

$$\begin{aligned}\Delta(Q,k,T,N) &\lesssim F_1 + \frac{F_1}{N}\bigg(\Big(\frac{N^2}{F_1}\Big)^e + N\bigg) \ll F_1 + \frac{N^{2e-1}}{F_1^{e-1}} \\ &\lesssim F + N^\alpha + N^{2e-1-\alpha(e-1)}.\end{aligned}$$

Choosing $\alpha = 2-1/e$ completes the proof. ■

2.4. Proof of Theorem 1.1. Using the trivial bound (1.5), we have

$$\Delta(Q,k,T,N) \lesssim Q^2kT\sqrt{N} + N \le (\sqrt{N} + Q^2kT)^2 \ll N + (Q^2kT)^2,$$

which is (2.2) with exponent $e = e_0 = 2$. Applying Proposition 2.6 gives (2.1) with $e_1 = 2 - 1/e_0 = 3/2$. Continuing this process, we obtain a sequence of exponents e_i, with $e_{i+1} = 2 - 1/e_i$, for which either (2.2) or (2.1) holds (in an alternating fashion). It is easy to check that the e_i are decreasing, with limit 1, whence Theorem 1.1 holds.

3. Proof of Proposition 1.8. The following proof is based on [IK, Section 17.2]. Decomposing with Dirichlet characters and applying orthogonality gives

(3.1)
$$\sum_{q \le Q} \sideset{}{^*}\sum_{a \,(\mathrm{mod}\, q)} \Big| \sum_{\substack{\mathrm{ht}(n) \le X \\ n \equiv a \,(\mathrm{mod}\, q)}} \alpha_n - \frac{1}{\varphi(q)} \sum_{\substack{\mathrm{ht}(n) \le X \\ (n,q)=1}} \alpha_n \Big|^2 = \sum_{q \le Q} \sum_{\substack{\chi \,(\mathrm{mod}\, q) \\ \chi \neq \chi_0}} \frac{1}{\varphi(q)} |S(X,\chi)|^2.$$

Write $q = q_0 q'$ and $\chi = \chi_0 \chi'$, where χ has conductor q'. Then (3.1) is at most
$$\sum_{\substack{q_0 q' \le Q \\ q' > 1}} \sideset{}{^*}\sum_{\chi' \,(\mathrm{mod}\, q')} \frac{1}{\varphi(q_0)\varphi(q')} |S(X, \chi' \chi_0)|^2.$$

We break up this sum according as $q' \le Q_0 = (\log X)^B$ or $q' > Q_0$. For $q' \le Q_0$, we apply the S–W condition (1.12), which gives a bound of the form
$$\sum_{q_0 \le Q} \frac{\tau_k(q_0)^2}{\varphi(q_0)} \sum_{1 < q' \le Q_0} \frac{X|\alpha|^2}{(\log X)^{2B}} \ll (\log Q)^{(k+1)^2} \frac{X|\alpha|^2}{(\log X)^B}.$$

The terms with $Q_0 < q' \le Q/q_0$ are bounded by
$$\ll \sum_{q_0 \le Q} \frac{1}{\varphi(q_0)} \sum_{\substack{Q_0 \le R \le Q/q_0 \\ \text{dyadic}}} R^{-1+\varepsilon} \varDelta(R,X) \sum_{\substack{\mathrm{ht}(n) \le X \\ (n,q_0)=1}} |\alpha_n|^2.$$

For $R \le (XQ)^{1/10}$, we use the "ε-free" bound $\varDelta(R,X) \ll (R^4 + X\log X)$ (see (1.5)), while for $R > (XQ)^{1/10}$, we use Theorem 1.1. In total, we obtain the following bound for the terms with $q' > Q_0$:
$$|\alpha|^2 \sum_{q_0 \le Q} \frac{1}{\varphi(q_0)} \Big(Q^{1+\varepsilon} + \frac{X}{(\log X)^{B(1-\varepsilon)-1}} \Big) \ll \Big(Q^{1+\varepsilon} + \frac{X}{(\log X)^{B(1-\varepsilon)-2}} \Big) |\alpha|^2.$$

Choosing $B(1-\varepsilon) - 2 > A$ completes the proof of Proposition 1.8.

4. Proof of Theorem 1.5

4.1. Miscellany. We begin with some miscellaneous results that will be useful later.

DEFINITION 4.1 (A partition of unity). Let $T \ge 1$, $\varepsilon > 0$. Choose smooth and even functions ω_0 and $\omega_{T'}(r) = \omega(r/T')$ so that for all $|r| \ll T$ we have

$$\omega_0(r) + \sum_{T' \text{ dyadic}} \omega_{T'}(r) = 1, \tag{4.1}$$

where $\omega_0(r)$ is supported on $r \ll T^\varepsilon$, ω is supported on $[1,2] \cup [-2,-1]$, and T' runs over $O(\log T)$ real numbers with $T^\varepsilon \ll T' \ll T$.

It is convenient to re-write the left hand side of (4.1) as $\sum_{T'} \omega_{T'}$, where T' runs over the dyadic numbers from Definition 4.1, along with an additional value $T' = 1$ giving rise to ω_0.

LEMMA 4.2. *Let w be an integrable function supported on $[U, 2U]$, with $1 \le U \le 2T$. Suppose $\beta_t \in L^2(\mathbb{R})$, supported on $[T/2, T]$. Then*

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \beta_{t_1}\overline{\beta_{t_2}}w(t_1 - t_2)\,dt_1\,dt_2 \tag{4.2}$$
$$= \sum_{\substack{0\le j_1,j_2\le 10T/U\\ |j_1-j_2|\le 1}} \int_U^{2U}\int_U^{2U} \beta_{T-U+Uj_1+v_1}\overline{\beta_{T-U+Uj_2+v_2}}w(U(j_1-j_2)+v_1-v_2)\,dv_1\,dv_2.$$

Proof. We cover the interval $[T/2, T]$ without overlaps by smaller intervals $[T/2, T/2 + U]$, $[T/2 + U, T/2 + 2U]$, ..., which gives

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \beta_{t_1}\overline{\beta_{t_2}}w(t_1 - t_2)\,dt_1\,dt_2 \tag{4.3}$$
$$= \sum_{0\le j_1,j_2\le 10T/U} \int_{T/2+Uj_1}^{T/2+Uj_1+U} \beta_{t_1} \int_{T/2+Uj_2}^{T/2+Uj_2+U} \overline{\beta_{t_2}}w(t_1 - t_2)\,dt_1\,dt_2.$$

Next, change variables $t_i \mapsto T/2-U+Uj_i+v_i$ for $i = 1, 2$, where $U \le v_i \le 2U$. Note that the integrand vanishes unless $|j_1 - j_2| \le 1$. The result follows. ∎

LEMMA 4.3 (Archimedean separation of variables). *For $s = \sigma + iy$ with $\sigma > 0$ fixed, $|r| \le T$, and $|y| \le |r|^{1/2}$, let*

$$\gamma(r) = \gamma_s(r) = \frac{\Gamma_{\mathbb{R}}(\sigma + iy + ir)\Gamma_{\mathbb{R}}(\sigma + iy - ir)}{\Gamma_{\mathbb{R}}(1 - \sigma - iy + ir)\Gamma_{\mathbb{R}}(1 - \sigma - iy - ir)}. \tag{4.4}$$

Let ω and ω_0 be as in Definition 4.1. Then for T' with $1 + |s|^2 \ll T' \le T$, there exists a function $\eta = \eta_{T'}$ satisfying

$$\eta_{T'}(u) \ll (T')^{2\sigma}(1 + |u|T')^{-A} \quad \text{and} \quad \int_{-\infty}^{\infty} |\eta_{T'}(u)|\,du \ll (T')^{2\sigma-1}, \tag{4.5}$$

so that

$$\gamma(r)\omega_{T'}(r) = \int_{-\infty}^{\infty} \eta_{T'}(u)e(ur)\,du. \tag{4.6}$$

If $|s| \ll T^{\varepsilon}$ and $T' = 1$ (that is, $\omega_{T'} = \omega_0$), then (4.6) holds with

$$\eta_1(u) \ll T^{\varepsilon}\left(1 + \frac{|u|}{T^{\varepsilon}}\right)^{-A}. \tag{4.7}$$

Proof. A tedious but straightforward calculation with Stirling's approximation gives

$$\gamma(r) = \left(\frac{|r|}{2}\right)^{2s-1}\left(c_0 + \frac{c_1}{r^2} + \cdots\right),$$

where the c_i are some polynomials in s, of degree at most $2i+1$. This provides an asymptotic expansion as $r \to \infty$ provided $s \ll |r|^{1/2}$, say. From this, one may derive

$$\gamma^{(j)}(r) \ll |r|^{2\sigma-1-j} \quad \text{for } |r| \gg |s|^2 + 1. \tag{4.8}$$

By Fourier inversion, we have

$$\gamma(r)\omega(r/T') = \int_{-\infty}^{\infty} \eta_{T'}(u)e(ur)\,du, \qquad \eta_{T'}(u) = \int_{-\infty}^{\infty} \gamma(r)\omega(r/T')e(-ur)\,dr.$$

Integration by parts, aided with (4.8), gives (4.5). For $T' = 1$ and $|s| \ll T^{\varepsilon}$, the asymptotic Stirling formula does not hold, yet we can claim a crude but uniform upper bound of the form $\gamma^{(j)}(r) \ll (T^{\varepsilon})^j$, which suffices to obtain (4.7). ∎

Corollary 4.4. *Let $\gamma = \gamma_s$ be as in* (4.4), *and suppose $b_t \in L^2(\mathbb{R})$, supported on $[T/2, T]$. Suppose $s \ll T^{o(1)}$. Suppose $\omega_{T'}$ is as in Definition* 4.1 *for some $1 \ll T' \ll T$. Then*

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \beta_{t_1}\overline{\beta_{t_2}}\gamma(t_1-t_2)\omega_{T'}(t_1-t_2)\,dt_1\,dt_2 &= \sum_{\substack{0\le j_1,j_2\le 10T/T' \\ |j_1-j_2|\le 1}} \int_{-\infty}^{\infty} \eta_{T'}(u)e(uT'(j_1-j_2)) \\ \times \Big(\int_{T'}^{2T'} \beta_{T/2-T'+T'j_1+v_1} e(v_1u)\,dv_1\Big)&\Big(\int_{T'}^{2T'} \overline{\beta_{T/2-T'+T'j_2+v_2}} e(-v_1u)\,dv_2\Big)\,du,\end{aligned} \tag{4.9}$$

with $\eta_{T'}$ as in Lemma 4.3.

Proof. This follows from Lemma 4.2 followed by (4.6). ∎

4.2. Preparation. Here we begin the proof of Theorem 1.5. Choose a nonnegative smooth weight function w, with $w(x) \ge 1$ for $1/2 \le x \le 1$, and $w(x) = 0$ for $x < 1/4$ and for $x \ge 2$. From (1.4), we have $\Delta^*(Q,k,T,N) \le \max_{|\beta|=1} S$, where

$$S = \sum_{(a,b)=1} w(ab/N)\Big|\int_{T/2\le t\le T} \sum_{\substack{Q/2<q\le Q \\ (q,k)=1}} \sideset{}{^*}\sum_{\chi\,(\mathrm{mod}\,q)} \sum_{\theta\,(\mathrm{mod}\,k)} \beta_{\chi,\theta,t}\lambda_{\chi\theta,t}(a,b)\Big|^2. \tag{4.10}$$

We will assume that $\beta_{\chi,\theta,t}$ is supported on

$$\begin{aligned} &\operatorname{cond}(\chi)=q, \quad Q/2<q\le Q, \quad (q,k)=1,\\ &\theta \pmod{k}, \qquad T/2\le t\le T, \end{aligned} \tag{4.11}$$

and that an otherwise un-labeled integral/sum over t,q,χ,θ is implied to run over this domain. In particular, we will often suppress these conditions and recall them only when needed. To prove Theorem 1.5, it suffices to prove the bound for χ and θ of fixed parities, so for convenience we also assume that this condition is enforced by the support of $\beta_{\chi,\theta,t}$.

Let $1\le Y\le \frac{N}{100}$ be a parameter to be chosen later. Then $S\le S_{>Y}$, where

(4.12)
$$S_{>Y}=\sum_{\frac{ab}{(a,b)^2}>Y} w(ab/N)\Big|\int_{T/2\le t\le T}\sum_{\substack{Q/2<q\le Q\\(q,k)=1}}\sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)}\sum_{\theta\,(\mathrm{mod}\,k)}\beta_{\chi,\theta,t}\lambda_{\chi\theta,t}(a,b)\Big|^2$$

by positivity, since if $(a,b)=1$, then the condition $ab>Y$ is redundant on the support of $w(ab/N)$. By simple inclusion-exclusion, we have

$$S_{>Y}=S_{\le\infty}-S_{\le Y},$$

where for $*\in\{Y,\infty\}$, $S_{\le *}$ corresponds to the sum over $ab/(a,b)^2\le *$. We will often write S_∞ as an alias for $S_{\le\infty}$.

One of the main issues with applying the functional equation is that, after opening the square, we obtain a character of the form $\chi_1\overline{\chi_2}\theta_1\overline{\theta_2}$ which may be imprimitive. In order to facilitate the problem of controlling the conductor, we will apply some combinatorial type decompositions. These preparatory results are bookended by Lemmas 4.5 and 4.11.

Lemma 4.5 (Detecting primitivity). *Let $q\ge 1$ be an integer. There exist complex numbers $c_\ell=c_\ell(q)$ supported on a finite set of integers with the following two properties:*

- *For each $\psi \pmod{q}$, the sum $\sum_\ell c_\ell\psi(\ell)$ is 1 if ψ is primitive, and is 0 if ψ is imprimitive.*
- *We have $\sum_\ell |c_\ell|\le\tau(q)$, where $\tau(q)$ denotes the number of divisors of q.*

Proof. Suppose ψ has conductor q^*. Consider the expression

$$\sum_{d|q}\mu(d)\Big(\frac{1}{d}\sum_{y\,(\mathrm{mod}\,d)}\psi\Big(1+\frac{q}{d}y\Big)\Big).$$

The inner sum is 1 if q^* divides q/d (equivalently, d divides q/q^*), and 0 otherwise. Hence the above sum evaluates as $\sum_{d|q/q^*}\mu(d)$, which by Möbius inversion is the indicator function of $q^*=q$, i.e., ψ is primitive. To finish the proof, we can let c_ℓ be supported on $1 \le \ell \le q+1$, and let

$$c_\ell = \sum_{d|q} \frac{\mu(d)}{d} \sum_{\substack{1 \le y \le d \\ 1+qy/d=\ell}} 1 = \sum_{e|(q,\ell-1)} \frac{\mu(q/e)}{q/e}, \tag{4.13}$$

so that $\sum_\ell |c_\ell| \le \tau(q)$. ∎

Suppose $q, r \ge 1$ are integers with $r \mid q$. Let G_q (resp. G_r) be the group of Dirichlet characters modulo q (resp. r). By a slight abuse of notation, we can view G_r as a subgroup of G_q, by multiplying every element of G_r by the trivial character modulo q.

The following lemma is analogous to Lemma 4.2.

Lemma 4.6. *Let q, r, G_q, and G_r be as above. Let $F(\chi_1, \chi_2)$ be a function defined on pairs of Dirichlet characters modulo q. Then*

$$\sum_{\substack{\chi_1, \chi_2 \pmod q \\ \chi_1\overline{\chi_2} \text{ of modulus } r}} F(\chi_1, \chi_2) = \sum_{\gamma \in G_q/G_r} \sum_{\psi_1, \psi_2 \pmod r} F(\gamma\psi_1, \gamma\psi_2).$$

Proof. The condition that $\chi_1\overline{\chi_2}$ has modulus r means that $\chi_1\overline{\chi_2} \in G_r$. Now say $G_q = \bigcup_\gamma \gamma G_r$, where γ runs over G_q/G_r. By basic group theory, we can write uniquely $\chi_1 = \gamma\psi_1$ and $\chi_2 = \gamma\psi_2$ with $\gamma \in G_q/G_r$ and with $\psi_1, \psi_2 \in G_r$. ∎

Corollary 4.7 (Separation of variables). *Let notation be as in Lemma 4.6. Then*

$$\sum_{\substack{\chi_1, \chi_2 \pmod q \\ \chi_1\overline{\chi_2} \text{ of conductor } r}} F(\chi_1, \chi_2) = \sum_\ell c_\ell(r) \sum_{\gamma \in G_q/G_r} \sum_{\psi_1, \psi_2 \pmod r} (\psi_1\overline{\psi_2})(\ell) F(\gamma\psi_1, \gamma\psi_2).$$

Proof. We first apply Lemma 4.6 to detect that $\chi_1\overline{\chi_2}$ has modulus r, and then use Lemma 4.5 to detect that $\psi_1\overline{\psi_2}$ is primitive. ∎

Definition 4.8. Let $k \ge 1$ be an integer. Define the set D_k to consist of tuples $\mathbf{k} = (k_0, k_1, k', \delta)$, where k_0, k_1, k' run over divisors of k satisfying $k_0k_1k' = k$, $(k_0, k') = 1$, and $k_1 \mid (k')^\infty$, and where δ runs over coset representatives of $G_k/G_{k'}$.

Lemma 4.9. *Let $k \ge 1$ be an integer, and let b_θ be any sequence of complex numbers indexed by Dirichlet characters θ modulo k. Then we have a decomposition of the form*

$$\Big| \sum_{\theta \pmod k} b_\theta \Big|^2 = \sum_{\mathbf{k} \in D_k} \sum_\ell c_\ell(k') \Big| \sum_{\theta' \pmod{k'}} b_{\delta\theta'} \theta'(\ell) \Big|^2, \tag{4.14}$$

which can alternatively be written as

$$\Big|\sum_{\theta\,(\mathrm{mod}\,k)} b_\theta\Big|^2 = \sum_{\mathbf{k}\in D_k}\ \sum_{\substack{\theta_1',\theta_2'\,(\mathrm{mod}\,k')\\ \mathrm{cond}(\theta_1'\overline{\theta_2'})=k'}} b_{\delta\theta_1'}\overline{b_{\delta\theta_2'}}. \tag{4.15}$$

Proof. We begin by opening the square, obtaining a double sum $\sum_{\theta_1,\theta_2\,(\mathrm{mod}\,k)} b_{\theta_1}\overline{b_{\theta_2}}$. Parameterizing the sum according to the conductor (say k') of $\theta_1\overline{\theta_2}$, we obtain

$$\Big|\sum_{\theta\,(\mathrm{mod}\,k)} b_\theta\Big|^2 = \sum_{k'|k}\ \sum_{\substack{\theta_1,\theta_2\,(\mathrm{mod}\,k)\\ \mathrm{cond}(\theta_1\overline{\theta_2})=k'}} b_{\theta_1}\overline{b_{\theta_2}}.$$

Next we apply Corollary 4.7 with $F(\theta_1,\theta_2)=b_{\theta_1}\overline{b_{\theta_2}}$, which gives

$$\Big|\sum_{\theta\,(\mathrm{mod}\,k)} b_\theta\Big|^2 = \sum_{k'|k}\sum_{\ell} c_\ell(k') \sum_{\delta\in G_k/G_{k'}}\ \sum_{\theta_1',\theta_2'\,(\mathrm{mod}\,k')} (\theta_1'\overline{\theta_2'})(\ell) b_{\delta\theta_1'}\overline{b_{\delta\theta_2'}}.$$

With a further factorization $k_0k_1=\frac{k}{k'}$ with $(k_0,k')=1$ and $k_0\,|\,(k')^\infty$, we obtain (4.14). The variant (4.15) is similar. ∎

We also need more elaborate versions of Definition 4.8 and Lemma 4.9 to handle χ of varying modulus.

Definition 4.10. For $i=1,2$, suppose χ_i is primitive of conductor q_i. Factor

$$q_i = q_i'q_i^+q_i^-r \quad\text{and}\quad \chi_i=\chi_i'\chi_i^+\chi_i^-\chi_i^{(r)}, \tag{4.16}$$

where χ_i' has conductor q_i', χ_i^+ has conductor q_i^+, and so on, and the factorization is defined in terms of local information as follows:

- (i) The primes making up q_1' are those that divide q_1 but do not divide q_2, and likewise the primes in q_2' are those that divide q_2 but not q_1.
- (ii) The factors q_1^+ and q_2^- are characterized by $1\le v_p(q_2^-)<v_p(q_1^+)$ for all $p\,|\,q_1^+$. Similarly, q_2^+ and q_1^- are characterized by $1\le v_p(q_1^-)<v_p(q_2^+)$ for all $p\,|\,q_2^+$.
- (iii) The remaining factor r corresponds to the primes where $v_p(q_1)=v_p(q_2)$.

Definition 4.10 is motivated by the fact that

$$\chi_1\overline{\chi_2} = \underbrace{\chi_1'}_{q_1'}\ \underbrace{\overline{\chi_2'}}_{q_2'}\ \underbrace{(\chi_1^+\overline{\chi_2^-})}_{q_1^+}\ \underbrace{(\chi_1^-\overline{\chi_2^+})}_{q_2^+}\ \chi_1^{(r)}\overline{\chi_2^{(r)}}, \tag{4.17}$$

which has conductor $q_1'q_2'q_1^+q_2^+\,\mathrm{cond}(\chi_1^{(r)}\overline{\chi_2^{(r)}})$.

Let b_χ be any sequence of complex numbers indexed by primitive Dirichlet characters χ modulo q, with q varying over a finite set of positive integers.

Consider the sum $|\sum_{q,\chi} b_\chi|^2$. Opening the square gives a sum of the form $\sum_{q_1,q_2,\chi_1,\chi_2} b_{\chi_1}\overline{b_{\chi_2}}$. Definition 4.10 shows that the parameters q_i', q_i^+, etc., are uniquely determined. We can then arrange the sum according to the values of these parameters, giving

$$\left|\sum_{q,\chi} b_\chi\right|^2 = \sum_{\substack{q_1^+,q_1^-,q_2^+,q_2^-,r \\ \text{(Def. 4.10)}}} \Bigg(\sum_{\substack{q_1',\chi_1',\chi_1^+,\chi_1^-,\chi_1^{(r)} \\ \text{(Def. 4.10)}}} b_{\chi_1'\chi_1^+\chi_1^-\chi_1^{(r)}} \Bigg) \Bigg(\sum_{\substack{q_2',\chi_2',\chi_2^+,\chi_2^-,\chi_2^{(r)} \\ \text{(Def. 4.10)}}} \overline{b_{\chi_2'\chi_2^+\chi_2^-\chi_2^{(r)}}} \Bigg), \tag{4.18}$$

where "(Def. 4.10)" in the summation conditions indicates the conditions translated into appropriate summation form.

We further develop the sums over $\chi_1^{(r)}$ and $\chi_2^{(r)}$, using (4.15). Specifically, write

$$r = r_0 r_1 r', \tag{4.19}$$

where $\chi_1^{(r)}\overline{\chi_2^{(r)}}$ has conductor r', $(r_0, r') = 1$, and $r_1 | (r')^\infty$. We then write $\chi_i^{(r)} = \gamma\psi_i$, where γ runs over $G_r/G_{r'}$ and ψ_i run over characters modulo r'. The property that $\chi_1^{(r)}\overline{\chi_2^{(r)}}$ has conductor r' is equivalent to $\psi_1\overline{\psi_2}$ being primitive (of modulus r'). Applying this to (4.18), we find that $\sum_{q,\chi} |b_\chi|^2$ equals

$$\sum_{\substack{q_1^+,q_1^-,q_2^+,q_2^-,r \\ (r_0,r_1,r',\gamma)\in D_r \\ \text{(Def. 4.10)}}} \Bigg(\sum_{\substack{q_1',\chi_1',\chi_1^+,\chi_1^-,\psi_1 \\ \text{(Def. 4.10)}}} b_{\chi_1'\chi_1^+\chi_1^-\gamma\psi_1} \Bigg) \times \Bigg(\sum_{\substack{q_2',\chi_2',\chi_2^+,\chi_2^-,\psi_2 \\ \text{(Def. 4.10)}}} \overline{b_{\chi_2'\chi_2^+\chi_2^-\gamma\psi_2}} \Bigg) \delta(\mathrm{cond}(\psi_1\overline{\psi_2}) = r'). \tag{4.20}$$

Now let

$$\mathbf{q} = (q_1^+, q_1^-, q_2^+, q_2^-, r_0, r_1, r', \gamma),$$

where the integers $q_i^\pm$ satisfy Definition 4.10(ii), r is coprime to the $q_i^\pm$, and $(r_0, r_1, r', \gamma) \in D_r$ (as in Definition 4.8). The two sums in parentheses in (4.20) have only the following conditions *between each other*: q_1' and q_2' are coprime, and the conductor of $\psi_1\overline{\psi_2}$ is r'. We have thus derived the following.

Lemma 4.11. *Let b_χ be any sequence of complex numbers indexed by primitive Dirichlet characters χ modulo q, with q varying over a finite set of positive integers. Then*

$$\Big|\sum_{q,\chi} b_\chi\Big|^2 = \sum_{\mathbf{q}} \sum_{\substack{q_i',\chi_i',\chi_i^+,\chi_i^-,\psi_i \\ (q_1',q_2')=1,\psi_1\overline{\psi_2}\ prim. \\ (Def.\ 4.10)}} b_{\chi_1'\chi_1^+\chi_1^-\gamma\psi_1}\overline{b_{\chi_2'\chi_2^+\chi_2^-\gamma\psi_2}}. \tag{4.21}$$

In reference to (4.17), now $\chi_1^{(r)}\overline{\chi_2^{(r)}} = \psi_1\overline{\psi_2}|\gamma|^2$, which has conductor r', so $\chi_1\overline{\chi_2}$ has conductor $q_1'q_2'q_1^+q_2^+r'$.

We are now ready to apply the preceding decompositions to $S_{\leq *}$ (see (4.12) for the definition). Specifically, we apply Lemmas 4.9 (in the form (4.15)) and 4.11, which gives

$$S_{\leq *} = \sum_{\substack{\mathbf{k} \\ (\text{Def. } 4.8)}} \sum_{\substack{\mathbf{q} \\ (\text{Def. } 4.10)}} S_{\leq *}(\mathbf{k},\mathbf{q}), \tag{4.22}$$

where

$$\begin{aligned} &S_{\leq *}(\mathbf{k},\mathbf{q}) \\ &= \sum_{\substack{\theta_1',\theta_2' \,(\mathrm{mod}\, k') \\ \theta_1'\overline{\theta_2'}\ \text{prim.}}} \sum_{\substack{q_i',\chi_i',\chi_i^+,\chi_i^-,\psi_i \\ (q_1',q_2')=1,\psi_1\overline{\psi_2}\ \text{prim.} \\ (\text{Def. } 4.10)}} \int_{t_1,t_2} \beta_1\overline{\beta_2} \sum_{\substack{\frac{ab}{(a,b)^2}\leq * \\ (ab,k_0r_0)=1}} w\Big(\frac{ab}{N}\Big) \varPhi(a\bar{b})\, dt_1\, dt_2 \end{aligned} \tag{4.23}$$

with

$$\beta_i = \beta_{\chi_i'\chi_i^+\chi_i^-\gamma\psi_i,\delta\theta_i',t_i}, \tag{4.24}$$

and where $\varPhi = \varPhi_1\overline{\varPhi_2}$ with

$$\varPhi_i(m) = (\chi_i'\chi_i^+\chi_i^-\psi_i\theta_i')(m)m^{it_i}.$$

We remind the reader that there are additional conditions encoded in the support of the coefficients, as recorded in (4.11), which will be recalled as needed. Observe that the finite part of $\varPhi$ (i.e., omitting $m^{it_1-it_2}$) is primitive of modulus $q_1'q_2'q_1^+q_2^+r'k'$. It is convenient to record here for later purposes that for $i=1,2$,

$$\sum_{\mathbf{k},\mathbf{q}} |\beta_i|^2 := \sum_{\mathbf{k},\mathbf{q}} \int_{t_i} \sum_{q_i',\chi_i',\chi_i^+,\chi_i^-,\psi_i,\theta_i'} |\beta_{\chi_i'\chi_i^+\chi_i^-\gamma\psi_i,\delta\theta_i',t_i}|^2\, dt_i \ll (kQ)^\varepsilon|\beta|^2. \tag{4.25}$$

At this point our treatments of $S_{\leq *}$ for $*=Y$ and $*=\infty$ diverge.

4.3. Elementary side. In this section we develop $S_{\leq Y}(\mathbf{k},\mathbf{q})$.

Proposition 4.12. *We have $S_{\leq Y}(\mathbf{k},\mathbf{q}) = S_{\leq Y}^{(0)}(\mathbf{k},\mathbf{q}) + S_{\leq Y}'(\mathbf{k},\mathbf{q})$, where $S_{\leq Y}^{(0)}(\mathbf{k},\mathbf{q})$ is given by (4.30) below, and where*

$$|S_{\leq Y}'(\mathbf{k},\mathbf{q})| \lesssim \prod_{i=1}^{2} \overline{\varDelta}\Big(\frac{Q}{q_i^+q_i^-r'r_0r_1}, q_i^+q_i^-r'k', T, Y\Big)^{1/2}|\beta_i|. \tag{4.26}$$

Proof. Let $g = (a, b)$, and change variables $a \mapsto ga$ and $b \mapsto gb$, getting

$$S_{\leq Y}(\mathbf{k}, \mathbf{q}) = \sum_{(g,k_0r_0)=1} \sum_{\substack{\theta_1',\theta_2' \pmod{k'} \\ \theta_1'\overline{\theta_2'} \text{ prim.}}} \sum_{\substack{q_i',\chi_i',\chi_i^+,\chi_i^-,\psi_i \\ (q_1',q_2')=1,\psi_1\overline{\psi_2} \text{ prim.} \\ \text{(Def. 4.10)}}} \times \int_{t_1,t_2} \beta_1\overline{\beta_2} \sum_{\substack{ab\leq Y \\ (a,b)=1 \\ (ab,k_0r_0)=1}} w\left(\frac{g^2ab}{N}\right) \varPhi(a\bar{b}g\bar{g})\, dt_1\, dt_2.$$

Next we apply the Mellin inversion formula and evaluate the g-sum as a Dirichlet L-function of principal character to modulus $q_1'q_2'q_1^+q_2^+k'r'k_0r_0$. We further write

$$L(2s, \chi_{0,q_1'q_2'q_1^+q_2^+r'k'r_0k_0}) = \zeta(2s)\rho_{q_1'}\rho_{q_2'}\rho_{q_1^+}\rho_{q_2^+}\rho_{r'r_0}\rho_{k'k_0}, \tag{4.27}$$

where $\rho_n = \rho_n(s) = \prod_{p|n}(1 - p^{-2s})$. This gives

$$S_{\leq Y}(\mathbf{k}, \mathbf{q}) = \sum_{\substack{\theta_1',\theta_2' \pmod{k'} \\ \theta_1'\overline{\theta_2'} \text{ prim.}}} \sum_{\substack{q_i',\chi_i',\chi_i^+,\chi_i^-,\psi_i \\ (q_1',q_2')=1,\psi_1\overline{\psi_2} \text{ prim.} \\ \text{(Def. 4.10)}}} \times \int_{(2)} \rho_{r'r_0}\rho_{k'k_0} \int_{t_1,t_2} \beta_1'\overline{\beta_2'} \sum_{\substack{ab\leq Y \\ (a,b)=1 \\ (ab,k_0r_0)=1}} \left(\frac{N}{ab}\right)^s \frac{\widetilde{w}(s)}{2\pi i}\zeta(2s)\varPhi(a\bar{b})\, dt_1\, dt_2\, ds,$$

with $\beta_1' = \beta_1\rho_{q_1'}\rho_{q_1^+}$ and $\overline{\beta_2'} = \overline{\beta_2}\rho_{q_2'}\rho_{q_2^+}$.

Next we use Lemma 4.5 to detect the condition that $\theta_1'\overline{\theta_2'}$ is primitive, and again to detect that $\psi_1\overline{\psi_2}$ is primitive (modulo r'). We additionally use Möbius inversion to detect $(q_1', q_2') = 1$, via $\sum_{g'|(q_1',q_2')} \mu(g')$. Altogether, this gives

$$S_{\leq Y}(\mathbf{k}, \mathbf{q}) = \sum_{g'} \mu(g') \sum_{\ell_1,\ell_2} c_{\ell_1}(k')c_{\ell_2}(r') \times \int_{(2)} N^s \frac{\widetilde{w}(s)}{2\pi i}\zeta(2s)\rho_{r'r_0}\rho_{k'k_0} \sum_{\substack{(a,b)=1 \\ ab\leq Y \\ (ab,k_0r_0)=1}} \mathcal{A}_1\overline{\mathcal{A}_2}\, \frac{ds}{(ab)^s}, \tag{4.28}$$

where

$$\mathcal{A}_1 = \int_{t_1} \sum_{\substack{q_1',\chi_1',\chi_1^+,\chi_1^-,\psi_1,\theta_1' \\ q_1'\equiv 0 \pmod{g'} \\ \text{(Def. 4.10)}}} \beta_1\rho_{q_1'}\rho_{q_1^+}\theta_1'(\ell_1)\psi_1(\ell_2)\varPhi_1(a\bar{b})\, dt_1,$$

and $\mathcal{A}_2$ is similarly defined.

Now we shift the s-contour of integration to $\mathrm{Re}(s) = \varepsilon$, crossing a pole at $s = 1/2$ only. Write

$$S_{\leq Y}(\mathbf{k},\mathbf{q}) = S^{(0)}_{\leq Y}(\mathbf{k},\mathbf{q}) + S'_{\leq Y}(\mathbf{k},\mathbf{q}),$$

where $S^{(0)}_{\leq Y}$ denotes the polar term, and $S'_{\leq Y}$ denotes the new line of integration. Note that $\mathcal{A}_i|_{s=1/2} = \mathcal{A}_i^{(0)}$, where

$$\mathcal{A}_i^{(0)} = \int_{t_i} \sum_{\substack{q_i',\chi_i',\chi_i^+,\chi_i^-,\psi_i,\theta_i' \\ q_i'\equiv 0 \,(\mathrm{mod}\, g') \\ (\text{Def. } 4.10)}} \beta_i \frac{\varphi(q_i'q_i^+)}{q_i'q_i^+} \theta_i'(\ell_1)\psi_i(\ell_2)\Phi_i(a\bar{b})\, dt_i, \tag{4.29}$$

since $\rho_n(1/2) = \varphi(n)/n$. Therefore, using $(k'k_0, r'r_0) = 1$ for a slight simplification (recalling (4.11)), we have

$$\begin{aligned} &S^{(0)}_{\leq Y}(\mathbf{k},\mathbf{q}) \\ &= \sum_{g'} \mu(g') \sum_{\ell_1,\ell_2} c_{\ell_1}(k')c_{\ell_2}(r')\widetilde{w}(1/2)\frac{\varphi(k'k_0r'r_0)}{2k'k_0r'r_0} \sum_{\substack{(a,b)=1\\ ab\leq Y\\ (ab,k_0r_0)=1}} \left(\frac{N}{ab}\right)^{1/2} \mathcal{A}_1^{(0)}\overline{\mathcal{A}_2^{(0)}}. \end{aligned} \tag{4.30}$$

Now we estimate $S'_{\leq Y}(\mathbf{k},\mathbf{q})$. We arrange the expression to most closely resemble (4.10), specifically

$$|S'_{\leq Y}(\mathbf{k},\mathbf{q})| \ll (QkN)^{\varepsilon} \sum_{g'} \sum_{\ell_1,\ell_2} |c_{\ell_1}(k')c_{\ell_2}(r')| \max_{\mathrm{Re}(s)=\varepsilon} \sum_{\substack{(a,b)=1\\ ab\leq Y}} |\mathcal{A}_1\mathcal{A}_2|. \tag{4.31}$$

Referring back to (1.4), and noting that our new family has varying modulus q_i' of size $Q/q_i^+q_i^-r'r_0r_1$, and fixed modulus $q_i^+q_i^-r'k'$, we see

$$\begin{aligned} &\sum_{g'} \sum_{\substack{(a,b)=1\\ ab\leq Y}} |\mathcal{A}_i|^2 \\ &\qquad \ll (QkN)^{\varepsilon} \max_{1\leq Y'\leq Y} \Delta\left(\frac{Q}{q_i^+q_i^-r'r_0r_1}, q_i^+q_i^-r'k', T, Y'\right)|\beta_i|^2. \end{aligned} \tag{4.32}$$

Using Cauchy's inequality and monotonicity (Lemma 2.1) leads quickly to (4.26). ∎

4.4. Functional equation side. In this section we will apply the functional equation of Dirichlet L-functions to $S_\infty(\mathbf{k},\mathbf{q})$, picking up from the expression (4.23). To facilitate this, we first apply Möbius inversion, in the form

$$\sum_{(ab,k_0r_0)=1} w\left(\frac{ab}{N}\right)\Phi(a\overline{b}) \tag{4.33}$$
$$= \sum_{\substack{g_1|k_0 \\ g_2|k_0}}\sum_{\substack{g_3|r_0 \\ g_4|r_0}} \mu(g_1)\mu(g_2)\mu(g_3)\mu(g_4)\Phi(g_1g_3\overline{g_2g_4})\sum_{a,b} w\left(\frac{g_1g_2g_3g_4ab}{N}\right)\Phi(a\overline{b}).$$

To continue the theme of concise notation, let $\mathbf{g} = (g_1, g_2, g_3, g_4)$, $\mu(\mathbf{g}) = \mu(g_1)\mu(g_2)\mu(g_3)\mu(g_4)$, $\Phi(\mathbf{g}) = \Phi(g_1g_3\overline{g_2g_4})$, and $|\mathbf{g}| = g_1g_2g_3g_4$. The summation condition on $\mathbf{g}$ is that

$$g_1 \mid k_0, \quad g_2 \mid k_0, \quad g_3 \mid r_0, \quad g_4 \mid r_0, \tag{4.34}$$

though we will usually suppress this and only recall it as needed. Then $S_\infty(\mathbf{k}, \mathbf{q})$ equals

$$\sum_{\substack{\mathbf{g} \\ (4.34) \text{ holds}}} \mu(\mathbf{g}) \sum_{\substack{\theta_1',\theta_2' \,(\mathrm{mod}\, k') \\ \theta_1'\overline{\theta_2'} \text{ prim.}}} \sum_{\substack{q_i',\chi_i',\chi_i^+,\chi_i^-,\psi_i \\ (q_1',q_2')=1,\psi_1\overline{\psi_2} \text{ prim.} \\ (\text{Def. } 4.10)}}$$
$$\times \int_{t_1,t_2} \beta_1\overline{\beta_2}\sum_{a,b} w\left(\frac{ab|\mathbf{g}|}{N}\right)\Phi(\mathbf{g}a\overline{b})\, dt_1\, dt_2.$$

We also have need to decompose the t_i-integrals to help pin down the archimedean conductor. Applying the partition from Definition 4.1, we deduce that $S_\infty(\mathbf{k}, \mathbf{q})$ equals

$$\sum_{\mathbf{g},T'} \mu(\mathbf{g}) \sum_{\substack{\theta_1',\theta_2' \,(\mathrm{mod}\, k') \\ \theta_1'\overline{\theta_2'} \text{ prim.}}} \sum_{\substack{q_i',\chi_i',\chi_i^+,\chi_i^-,\psi_i \\ (q_1',q_2')=1,\psi_1\overline{\psi_2} \text{ prim.} \\ (\text{Def. } 4.10)}} \tag{4.35}$$
$$\times \int_{t_1,t_2} \beta_1\overline{\beta_2}\omega_{T'}(t_1-t_2)\sum_{a,b} w\left(\frac{ab|\mathbf{g}|}{N}\right)\Phi(\mathbf{g}a\overline{b})\, dt_1\, dt_2.$$

Define quantities

(4.36)
$$Q^* = \frac{Q^2kT'}{q_1^- q_2^- r' r_0^2 r_1^2 k_0 k_1}, \qquad N^* = \frac{Q^4k^2(T')^2|\mathbf{g}|(QkTN)^\varepsilon}{N(q_1^- q_2^- r' r_0^2 r_1^2 k_0 k_1)^2} = (QkTN)^\varepsilon \frac{(Q^*)^2|\mathbf{g}|}{N},$$

and note that among the variables of summation, Q^* depends only on the outer variables $\mathbf{q}$, $\mathbf{k}$, and T', while N^* depends only on $\mathbf{q}$, $\mathbf{k}$, T', and $\mathbf{g}$.

Proposition 4.13. *We have a decomposition*

$$S_\infty(\mathbf{k},\mathbf{q}) = S_\infty^{(0)}(\mathbf{k},\mathbf{q}) + S_\infty'(\mathbf{k},\mathbf{q}) + S_\infty^{\mathrm{diag}}(\mathbf{k},\mathbf{q}) + \mathcal{E}_\infty, \tag{4.37}$$

with the following properties. The term $S_\infty^{(0)}(\mathbf{k},\mathbf{q})$ is given by (4.43) *below,* *and* $S'_\infty(\mathbf{k},\mathbf{q})$ *satisfies*

$$(4.38)\quad |S'_\infty(\mathbf{k},\mathbf{q})| \lesssim \sum_{\mathbf{g},T'} \frac{N}{Q^*|\mathbf{g}|} \prod_{i=1}^{2} \overline{\Delta}\Big(\frac{Q}{q_i^+ q_i^- r' r_0 r_1}, q_i^+ q_i^- r' k', T', N^*\Big)^{1/2} |\beta_i|.$$

The diagonal term satisfies the bound

$$(4.39)\qquad \sum_{\mathbf{k},\mathbf{q}} |S_\infty^{\mathrm{diag}}(\mathbf{k},\mathbf{q})| \lesssim N|\beta|^2,$$

and the term $\mathcal{E}_\infty$ *is negligibly small.*

Proof. Applying the Mellin inversion formula to w and writing the sum over a and b as a product of Dirichlet L-functions in (4.35) gives

$$S_\infty(\mathbf{k},\mathbf{q}) = \sum_{\mathbf{g},T'} \mu(\mathbf{g}) \sum_{\substack{\theta_1',\theta_2' \,(\mathrm{mod}\, k') \\ \theta_1'\overline{\theta_2'}\ \mathrm{prim.}}} \sum_{\substack{q_i',\chi_i',\chi_i^+,\chi_i^-,\psi_i \\ (q_1',q_2')=1, \psi_1\overline{\psi_2}\ \mathrm{prim.} \\ (\mathrm{Def.}\ 4.10)}}$$
$$\times \int_{t_1,t_2} \omega_{T'}(t_1-t_2)\Phi(\mathbf{g})\beta_1\overline{\beta_2} \int_{(2)} \Big(\frac{N}{|\mathbf{g}|}\Big)^s \widetilde{w}(s) L(s,\Phi)L(s,\overline{\Phi}) \frac{ds}{2\pi i}\, dt_1\, dt_2.$$

We shift contours to the line $-\varepsilon$, crossing a pair of poles at $s = 1 \pm i(t_1-t_2)$, which exist only when Φ is trivial, and let $S'_\infty(\mathbf{k},\mathbf{q})$ be the new integral on the line $-\varepsilon$. Recall that the finite part of Φ is primitive of modulus

$$(4.40)\qquad \mathfrak{q} = q_1' q_2' q_1^+ q_2^+ r' k'.$$

In particular, Φ being trivial forces $q_1' = q_2' = q_1^+ = q_2^+ = q_1^- = q_2^- = r' = k' = 1$, and the rapid decay of $\widetilde{w}(s)$ practically forces $|t_1 - t_2| \ll T^\varepsilon$. It is easy to see that the contribution of this diagonal polar term is consistent with (4.39).

On the line $-\varepsilon$ we change variables $s \mapsto 1-s$. Note that $L(s,\Phi)L(s,\overline{\Phi})$ satisfies the asymmetric functional equation

$$(4.41)\qquad L(1-s,\Phi)L(1-s,\overline{\Phi}) = \mathfrak{q}^{2s-1}\gamma_s L(s,\Phi)L(s,\overline{\Phi}),$$

where $\gamma_s = \gamma_s(t_1-t_2)$ (recall (4.4) for the definition), which is holomorphic for $\mathrm{Re}(s) > 0$. Recall that the parity of the χ_i and θ_i was assumed to be fixed, so that $\chi_1\overline{\chi_2}\theta_1\overline{\theta_2}$ is even, and hence the gamma factor is as stated in (4.4). For later use, note that $\gamma_s|_{s=1/2} = 1$. In addition, recall the bound (4.8), which in the present context means $\gamma_s(r) \ll (T')^{2\sigma-1}$. We then obtain

$$S'_\infty(\mathbf{k},\mathbf{q}) = \sum_{\mathbf{g},T'} \mu(\mathbf{g}) \sum_{\substack{\theta_1',\theta_2' \,(\mathrm{mod}\, k') \\ \theta_1'\overline{\theta_2'} \text{ prim.}}} \sum_{\substack{q_i',\chi_i',\chi_i^+,\chi_i^-,\psi_i \\ (q_1',q_2')=1,\, \psi_1\overline{\psi_2} \text{ prim.} \\ \text{(Def. 4.10)}}} \int_{t_1,t_2} \omega_{T'}(t_1-t_2)\Phi(\mathbf{g})\beta_1\overline{\beta_2}$$
$$\times \int_{(1+\varepsilon)} \widetilde{w}(1-s)\left(\frac{N}{|\mathbf{g}|}\right)^{1-s} \mathfrak{q}^{2s-1}\gamma_s L(s,\Phi)L(s,\overline{\Phi})\,\frac{ds}{2\pi i}\,dt_1\,dt_2.$$

Next we will re-open the Dirichlet series expansions of the Dirichlet L-functions. A small modification is that we write

$$L(s,\Phi) = \rho_{\Phi,k_0r_0} \sum_{(a,k_0r_0)=1} a^{-s}\Phi(a), \quad \text{where } \rho_{\Phi,k_0r_0} = \prod_{p\mid k_0r_0} (1-\Phi(p)p^{-s})^{-1},$$

and likewise for $L(s,\overline{\Phi})$. This gives

$$S'_\infty(\mathbf{k},\mathbf{q}) = \sum_{\mathbf{g},T'} \mu(\mathbf{g}) \sum_{\substack{\theta_1',\theta_2' \,(\mathrm{mod}\, k') \\ \theta_1'\overline{\theta_2'} \text{ prim.}}} \sum_{\substack{q_i',\chi_i',\chi_i^+,\chi_i^-,\psi_i \\ (q_1',q_2')=1,\, \psi_1\overline{\psi_2} \text{ prim.} \\ \text{(Def. 4.10)}}} \int_{t_1,t_2} \omega_{T'}(t_1-t_2)\Phi(\mathbf{g})\beta_1\overline{\beta_2}$$
$$\times \frac{N}{|\mathbf{g}|\mathfrak{q}} \sum_{(ab,k_0r_0)=1} \int_{(1+\varepsilon)} \widetilde{w}(1-s)\left(\frac{\mathfrak{q}^2|\mathbf{g}|}{Nab}\right)^{s} \gamma_s\Phi(a\overline{b})\rho_{\Phi,k_0r_0}\rho_{\overline{\Phi},k_0r_0}\,\frac{ds}{2\pi i}\,dt_1\,dt_2.$$

We then factor out the gcd of a and b, by writing $g' = (a,b)$ and changing variables $a \mapsto g'a$ and $b \mapsto g'b$. The sum over g' forms a Dirichlet L-function of principal character of modulus $\mathfrak{q}k_0r_0$, which is given by (4.27). Then $S'_\infty(\mathbf{k},\mathbf{q})$ equals

$$\sum_{\mathbf{g},T'} \mu(\mathbf{g}) \sum_{\substack{\theta_1',\theta_2' \,(\mathrm{mod}\, k') \\ \theta_1'\overline{\theta_2'} \text{ prim.}}} \sum_{\substack{q_i',\chi_i',\chi_i^+,\chi_i^-,\psi_i \\ (q_1',q_2')=1,\, \psi_1\overline{\psi_2} \text{ prim.} \\ \text{(Def. 4.10)}}} \int_{t_1,t_2} \omega_{T'}(t_1-t_2)\Phi(\mathbf{g})\beta_1\overline{\beta_2}$$
$$\times \int_{(1+\varepsilon)} \widetilde{w}(1-s)\frac{N}{|\mathbf{g}|\mathfrak{q}} \sum_{\substack{(a,b)=1 \\ (ab,k_0r_0)=1}} \left(\frac{\mathfrak{q}^2|\mathbf{g}|}{Nab}\right)^{s} \zeta(2s)$$
$$\times\, \rho_{q_1'}\rho_{q_2'}\rho_{q_1^+}\rho_{q_2^+}\rho_{k'r'k_0r_0}\rho_{\Phi,k_0r_0}\rho_{\overline{\Phi},k_0r_0}\gamma_s\Phi(a\overline{b})\,\frac{ds}{2\pi i}\,dt_1\,dt_2.$$

Shifting the integral far to the right shows that the portion of the sum with $ab \gg \frac{\mathfrak{q}^2(T')^2|\mathbf{g}|}{N}(QkTN)^\varepsilon$ is very small. Note

$$\mathfrak{q} = \frac{q_1'q_1^+q_1^-r'r_0r_1}{q_1^-\sqrt{r'}\,r_0r_1}\,\frac{q_2'q_2^+q_2^-r'r_0r_1}{q_2^-\sqrt{r'}\,r_0r_1}\,\frac{k'k_0k_1}{k_0k_1} \asymp \frac{Q^2k}{q_1^-q_2^-r'r_0^2r_1^2k_0k_1} = \frac{Q^*}{T'}, \tag{4.42}$$

and hence

$$\frac{\mathfrak{q}^2|\mathbf{g}|(T')^2}{N} \asymp \frac{(Q^*)^2|\mathbf{g}|}{N}.$$

Thus we can truncate the sum at $ab \leq N^*$. Let $S''_\infty(\mathbf{k},\mathbf{q})$ denote the contribution to $S'_\infty(\mathbf{k},\mathbf{q})$ from the terms with $ab \leq N^*$. Let $\mathfrak{q} = \mathfrak{q}_1\mathfrak{q}_2$, where $\mathfrak{q}_i = q_i' q_i^+ q_i^- \sqrt{r'k'}$.

Next we apply Lemma 4.5 to detect the condition that $\theta_1'\overline{\theta_2'}$ is primitive of modulus k', and likewise for $\psi_1\overline{\psi_2}$ of modulus r'. We also apply Möbius inversion to detect $(q_1', q_2') = 1$, as before (4.28). Our final arithmetical separation of variables step is to write

$$\rho_{\Phi,k_0r_0} = \sum_{d_1|(k_0r_0)^\infty} d^{-s}\Phi_1\overline{\Phi_2}(d_1),$$

and likewise for $\rho_{\overline{\Phi},k_0r_0}$ (indexing the sum with the letter d_2). We need an archimedean separation of variables too, and this is provided by Corollary 4.4. With this, and rearranging, we then obtain

$$\begin{aligned}
&S''_\infty(\mathbf{k},\mathbf{q})\\
&\quad= \sum_{\substack{\mathbf{g},T',g'\\|j_1-j_2|\leq 1}} \mu(\mathbf{g})\mu(g')\sum_{\ell_1,\ell_2} c_{\ell_1}(k')c_{\ell_2}(r') \sum_{d_1,d_2|(k_0r_0)^\infty} \int_{-\infty}^{\infty} \eta_{T'}(u)e(uT'(j_1-j_2))\\
&\qquad\times \sum_{\substack{(a,b)=1\\ab\leq N^*\\(ab,k_0r_0)=1}} \int_{(1+\varepsilon)} \frac{\widetilde{w}(1-s)}{(abd_1d_2)^s}\left(\frac{N}{|\mathbf{g}|}\right)^{1-s}\zeta(2s)\rho_{k'r'k_0r_0}\mathcal{B}_1\overline{\mathcal{B}_2}\,\frac{ds}{2\pi i}\,du,
\end{aligned}$$

where

$$\begin{aligned}
&\mathcal{B}_1 = \mathcal{B}_{1,s}\\
&= \int_U^{2U} \sum_{\substack{q_1',\chi_1',\chi_1^+,\chi_1^-,\psi_1,\theta_1'\\q_1'\equiv 0 \,(\mathrm{mod}\, g')\\ \text{(Def. 4.10)}}} \beta_{1,j_1}\theta_1'(\ell_1)\psi_1(\ell_2)\Phi_1(\mathbf{g}d_1\overline{d_2})\mathfrak{q}_1^{2s-1}\rho_{q_1'}\rho_{q_1^+}\Phi_1(a\overline{b})e(ut_1)\,dt_1,
\end{aligned}$$

with β_{1,j_1} taking the form $\beta_{*,T-T'/2+T'j_1+t_1}$ (i.e., with a linear change of variables as in Corollary 4.4), and where $\mathcal{B}_2$ is given by a similar definition.

We next shift the contour of integration back to the line $\mathrm{Re}(s) = \varepsilon$, crossing a pole at $s = 1/2$ only. Let $S_\infty^{(0)}(\mathbf{k},\mathbf{q})$ denote this polar term, and let $S'''_\infty(\mathbf{k},\mathbf{q})$ be the new integral. We record the polar term:

$$\begin{aligned}
&S_\infty^{(0)}(\mathbf{k},\mathbf{q}) \qquad (4.43)\\
&\quad= \sum_{\substack{\mathbf{g},T',g'\\|j_1-j_2|\leq 1}} \mu(\mathbf{g})\mu(g')\sum_{\ell_1,\ell_2} c_{\ell_1}(k')c_{\ell_2}(r') \sum_{d_1,d_2|(k_0r_0)^\infty} \frac{\widetilde{w}(1/2)}{\sqrt{d_1d_2}}\,\frac{\varphi(k'r'k_0r_0)}{2k'r'k_0r_0}
\end{aligned}$$

$$\times \int_{-\infty}^{\infty} \eta_{T'}(u) e(uT'(j_1 - j_2)) \sum_{\substack{(a,b)=1 \\ (ab,k_0 r_0)=1 \\ ab \leq N^*}} \left(\frac{N}{ab|\mathbf{g}|} \right)^{1/2} \mathcal{B}_1^{(0)} \overline{\mathcal{B}_2^{(0)}} \, du,$$

where $\mathcal{B}_i^{(0)} = \mathcal{B}_i|_{s=1/2}$ is given by

$$\mathcal{B}_i^{(0)} = \int_{t_i} \sum_{\substack{q_i', \chi_i', \chi_i^+, \chi_i^-, \psi_i, \theta_i' \\ q_i' \equiv 0 \,(\mathrm{mod}\, g') \\ \text{(Def. 4.10)}}} \beta_{i,j_i} \frac{\varphi(q_i' q_i^+)}{q_i' q_i^+} \theta_i'(\ell_1) \psi_i(\ell_2) \Phi_i(\mathbf{g} d_1 \overline{d_2} a \overline{b}) e(u t_i) \, dt_i. \tag{4.44}$$

Now we turn to $S_\infty'''(\mathbf{k}, \mathbf{q})$. By the triangle inequality, and using (4.5) to bound the L^1 norm of $\eta_{T'}$, we obtain

$$|S_\infty'''(\mathbf{k}, \mathbf{q})| \lesssim \sum_{\substack{\mathbf{g}, T', g' \\ |j_1 - j_2| \leq 1}} \frac{N}{|\mathbf{g}| Q^*} \max_{\substack{\mathrm{Re}(s) = \varepsilon \\ u \in \mathbb{R} \\ \ell_1, \ell_2}} \sum_{\substack{(a,b)=1 \\ ab \leq N^*}} |\mathfrak{q}_1^{-2s+1} \mathcal{B}_{1,s}| \, |\mathfrak{q}_2^{-2s+1} \mathcal{B}_{2,s}|. \tag{4.45}$$

Analogously to (4.32), on the line $\mathrm{Re}(s) = \varepsilon$, we obtain the bound

$$\sum_{\substack{(a,b)=1 \\ ab \leq N^*}} |\mathfrak{q}_i^{-2s+1} \mathcal{B}_{i,s}|^2 \lesssim \overline{\Delta}\left(\frac{Q}{q_i^+ q_i^- r' r_0}, q_i^+ q_i^- r' k', 2T', N^* \right) |\beta_{i,j_i}|^2. \tag{4.46}$$

We note that $\sum_{j_1} |\beta_{1,j_1}|^2 = |\beta_1|^2$, since this simply re-assembles the integral to all of $[T/2, T]$ (also, for each j_1, the number of j_2 with $|j_1 - j_2| \leq 1$ is at most three). Applying (4.46) to (4.45) via Cauchy's inequality and using (4.25) (and the sentence preceding it to handle the sum over the j_i) completes the proof of Proposition 4.13. ∎

4.5. Conclusion. Now we use Propositions 4.12 and 4.13 to prove Theorem 1.5. We have a decomposition

$$\begin{aligned} S(\mathbf{k}, \mathbf{q}) = S_\infty^{\mathrm{diag}}(\mathbf{k}, \mathbf{q}) + S_\infty'(\mathbf{k}, \mathbf{q}) - S_{\leq Y}'(\mathbf{k}, \mathbf{q}) \\ + (S_\infty^{(0)}(\mathbf{k}, \mathbf{q}) - S_{\leq Y}^{(0)}(\mathbf{k}, \mathbf{q})) + \mathcal{E}_\infty. \end{aligned} \tag{4.47}$$

The diagonal term is acceptable for Theorem 1.5, as also is the small error term $\mathcal{E}_\infty$.

Next we turn to the terms $S_*'(\mathbf{k}, \mathbf{q})$, where $*$ refers to $\leq Y$ or ∞. We choose

$$Y = (QkTN)^\varepsilon \frac{Q^4 k^2 T^2}{N}, \tag{4.48}$$

with the same value of ε as in the definition of N^* (see (4.36)). First consider $S'_{\leq Y}$, where Cauchy's inequality implies

$$\sum_{\mathbf{k},\mathbf{q}} |S'_{\leq Y}(\mathbf{k},\mathbf{q})| \precsim \prod_{i=1}^{2} \bigg(\sum_{\mathbf{k},\mathbf{q}} \overline{\Delta}\Big(\frac{Q}{q_i^+ q_i^- r' r_0 r_1}, q_i^+ q_i^- r' k', T, Y\Big) |\beta_i|^2 \bigg)^{1/2}.$$

Recall from (4.25) that $\sum_{\mathbf{k},\mathbf{q}} |\beta_i|^2 \ll (kQ)^{\varepsilon} |\beta|^2$. Hence

$$\sum_{\mathbf{k},\mathbf{q}} |S'_{\leq Y}(\mathbf{k},\mathbf{q})| \precsim \prod_{i=1}^{2} \bigg(\max_{\mathbf{k},\mathbf{q}} \overline{\Delta}\Big(\frac{Q}{q_i^+ q_i^- r' r_0 r_1}, q_i^+ q_i^- r' k', T, Y\Big) \bigg)^{1/2} |\beta|^2.$$

Recalling the definition (1.8), it is easy to see that

$$\max_{\mathbf{k},\mathbf{q}} \overline{\Delta}\Big(\frac{Q}{q_i^+ q_i^- r' r_0 r_1}, q_i^+ q_i^- r' k', T, Y\Big) \leq \overline{\Delta'}(Q, k, T, Y).$$

In summary, we have shown

$$\sum_{\mathbf{k},\mathbf{q}} |S'_{\leq Y}(\mathbf{k},\mathbf{q})| \precsim \overline{\Delta'}\Big(Q, k, T, \frac{Q^4 k^2 T^2}{N}\Big) |\beta|^2,$$

which is consistent with Theorem 1.5.

The case of S'_{∞} is fairly similar to that of $S'_{\leq Y}$, though the details are more complicated. Following similar steps to the case of $S'_{\leq Y}$, and using the AM-GM inequality, we derive

$$\sum_{\mathbf{k},\mathbf{q}} |S'_{\infty}(\mathbf{k},\mathbf{q})| \precsim |\beta|^2 \max_{\mathbf{k},\mathbf{q},\mathbf{g},T'} \frac{N}{Q^* |\mathbf{g}|} \overline{\Delta}\Big(\frac{Q}{q_1^+ q_1^- r' r_0 r_1}, q_1^+ q_1^- r' k', T', N^*\Big),$$

plus a similar term with the $i = 2$ variables (q_2^+, q_2^-, etc.). By symmetry, this latter term will give the same bound as the displayed one. Substituting the values of Q^* and N^* from (4.36), we obtain

$$\sum_{\mathbf{k},\mathbf{q}} |S'_{\infty}(\mathbf{k},\mathbf{q})| \precsim \frac{N}{Q^2 k T} |\beta|^2 \max_{\mathbf{k},\mathbf{q},\mathbf{g},T'} \frac{q_1^- q_2^- r' r_0^2 r_1^2 k_0 k_1 T}{|\mathbf{g}| T'} \times \overline{\Delta}\Big(\frac{Q}{q_1^+ q_1^- r' r_0 r_1}, q_1^+ q_1^- r' k', T', \frac{Q^4 k^2 (T')^2 |\mathbf{g}| (QkN)^{\varepsilon}}{N (q_1^- q_2^- r' r_0^2 r_1^2 k_0 k_1)^2}\Big). \tag{4.49}$$

A bit of checking, recalling $q_2^- \leq q_1^+$, shows that this is consistent with Theorem 1.5.

Finally, we consider $S_{\infty}^{(0)}(\mathbf{k},\mathbf{q}) - S_{\leq Y}^{(0)}(\mathbf{k},\mathbf{q})$, that is, the polar terms from $s = 1/2$. We need to show there is substantial cancellation between these two terms. To aid in this, we first simplify $S_{\infty}^{(0)}(\mathbf{k},\mathbf{q})$, which is defined in (4.43). Observe that

$$N^* = Y \frac{|\mathbf{g}| (T')^2}{(q_1^- q_2^- r' r_0^2 r_1^2 k_0 k_1)^2 T^2}, \tag{4.50}$$

and since $|\mathbf{g}|$ divides $k_0^2 r_0^2$ (recall (4.34)), we have $N^* \leq Y$. Then in the definition of $S_\infty^{(0)}$, we extend the sum over $ab \leq N^*$ to $ab \leq Y$, and subtract back the terms between N^* and Y. Write $S_{\infty,Y}^{(0)}$ for the terms with $ab \leq Y$, and let $S_{\infty,Y^*}^{(0)} = S_{\infty,Y}^{(0)} - S_\infty^{(0)}$ (which represents the terms with $N^* < ab \leq Y$). We claim that $S_{\infty,Y}^{(0)} = S_{\leq Y}^{(0)}$. To see this, we sum over $\mathbf{g}$ and d_1 and d_2 in (4.43) (though modified to read $ab \leq Y$ in place of $ab \leq N^*$). The sum over $\mathbf{g}$ is not constrained, and we have

$$\sum_{\mathbf{g}} \frac{\mu(\mathbf{g})\Phi(\mathbf{g})}{\sqrt{|\mathbf{g}|}} = \prod_{p|k_0 r_0} \left(1 - \frac{\Phi(p)}{\sqrt{p}}\right)\left(1 - \frac{\overline{\Phi}(p)}{\sqrt{p}}\right).$$

For d_1 and d_2, we have

$$\sum_{d_1,d_2|(k_0 r_0)^\infty} \frac{\Phi(d_1\overline{d_2})}{\sqrt{d_1 d_2}} = \prod_{p|k_0 r_0} \left(1 - \frac{\Phi(p)}{\sqrt{p}}\right)^{-1}\left(1 - \frac{\overline{\Phi}(p)}{\sqrt{p}}\right)^{-1}.$$

Therefore, these two evaluations perfectly cancel. The sums over j_1 and j_2 can be simplified by using Lemma 4.2 in the reverse order. Moreover, since $\gamma_s(t_1 - t_2) = 1$ at $s = 1/2$, we can write $\sum_{T'} \omega_{T'}(t_1 - t_2) = 1$. Hence, the partition of unity is fully re-assembled.

Comparing (4.29) and (4.44), it is not hard to see that $\mathcal{B}_i^{(0)}$ agrees with $\mathcal{A}_i^{(0)}$ after removal of $\Phi_i(\mathbf{g} d_1 \overline{d_2}) e(u t_i)$. This shows the claim that $S_{\infty,Y}^{(0)} = S_{\leq Y}^{(0)}$. Hence $S_\infty^{(0)} - S_{\leq Y}^{(0)} = -S_{\infty,Y^*}^{(0)}$, which for ease of reference we write directly as follows:

$$\begin{aligned} &S_{\infty,Y^*}^{(0)}(\mathbf{k},\mathbf{q}) \\ &\quad = \sum_{\substack{\mathbf{g},T',g' \\ |j_1-j_2|\leq 1}} \mu(\mathbf{g})\mu(g') \sum_{\ell_1,\ell_2} c_{\ell_1}(k') c_{\ell_2}(r') \sum_{d_1,d_2|(k_0 r_0)^\infty} \frac{\widetilde{w}(1/2)}{\sqrt{d_1 d_2}} \frac{\varphi(k' r' k_0 r_0)}{2k' r' k_0 r_0} \\ &\qquad \times \int_{-\infty}^{\infty} \eta_{T'}(u) e(uT'(j_1 - j_2)) \sum_{\substack{(a,b)=1 \\ N^* < ab \leq Y \\ (ab, k_0 r_0)=1}} \left(\frac{N}{ab|\mathbf{g}|}\right)^{1/2} \mathcal{B}_1^{(0)} \overline{\mathcal{B}_2^{(0)}}\, du. \end{aligned}$$

Now the estimations are similar to those of $S'_{\leq Y}$ and S'_∞, though the details are a little different. Following the same initial steps as in $S'_{\leq Y}$, we obtain

$$\begin{aligned} \text{(4.51)} \qquad &\sum_{\mathbf{k},\mathbf{q}} |S_{\infty,Y^*}^{(0)}(\mathbf{k},\mathbf{q})| \\ &\quad \lesssim |\beta|^2 \max_{\mathbf{k},\mathbf{q},\mathbf{g},T'} \max_{N^* \ll M \ll Y} \frac{N^{1/2}}{(|\mathbf{g}|M)^{1/2}} \Delta\left(\frac{Q}{q_1^+ q_1^- r' r_0 r_1}, q_1^+ q_1^- r' k', T', M\right). \end{aligned}$$

We claim this is bounded consistently with Theorem 1.5. To see this, first note $\frac{N^{1/2}}{Y^{1/2}} \leq \frac{N}{Q^2kT}$. Then the condition "$XR^2\ell U \leq Q^2kT$" from (1.8) is deduced from

$$\frac{Q^2kT}{N}\frac{N^{1/2}}{(|\mathbf{g}|M)^{1/2}}\left(\frac{Q^2k'T'}{q_1^+q_1^-r'r_0^2r_1^2}\right) \leq \frac{Q^2kT}{|\mathbf{g}|} \leq Q^2kT.$$

The condition "$X \leq C$" from (1.8) is easy to check, by setting $M = Y/C$. This completes the proof of Theorem 1.5.

5. Proof of Theorem 1.6

5.1. Miscellany. Here we present a couple tools with self-contained proofs.

Lemma 5.1. *Let c, d be positive integers, and define the Dirichlet series*

$$Z_{c,d}(s) = \sum_{\substack{(n,cd)=1\\ m|c^\infty}} \frac{n}{\varphi(n)}\frac{1}{(mn)^s}, \qquad \mathrm{Re}(s) > 1. \tag{5.1}$$

Then

$$Z_{c,d}(s) = Z_{1,1}(s)\nu_c(s)\delta_d(s), \tag{5.2}$$

where $Z_{1,1}(s)$ has meromorphic continuation to $\mathrm{Re}(s) > 0$ *with a simple pole at $s = 1$ only, and where*

$$\nu_c(s) = \prod_{p|c}\left(1 + \frac{p^{-s-1}}{1-p^{-1}}\right)^{-1}, \qquad \delta_d(s) = \prod_{p|d}\left(1 + (1-p^{-1})\frac{p^{-s}}{1-p^{-s}}\right)^{-1}.$$

Proof. A routine calculation gives

$$Z_{c,d}(s) = \prod_{p|c}(1-p^{-s})^{-1}\prod_{p\nmid cd}\left(1 + (1-p^{-1})^{-1}\frac{p^{-s}}{1-p^{-s}}\right),$$

from which the lemma follows with a bit of calculation. ∎

Lemma 5.2 (Separation of variables). *Let $\omega = \omega_V$ be a smooth, even function supported on $[-2V, 2V]$, where $V > 0$, satisfying $\omega_V^{(j)}(x) \ll V^{-j}$ for all $j = 0, 1, \ldots$. Let $w(x, y, z, w)$ be smooth of compact support on $\mathbb{R}^4_{>0}$. Let g be a Schwartz-class function. Define $F : \mathbb{R}^4_{>0} \to \mathbb{R}$ by*

$$F(x_1, y_1, x_2, y_2) = \omega_V(x_1y_2 - x_2y_1)g\left(T\log\frac{x_1y_2}{x_2y_1}\right)w\left(\frac{x_1}{X}, \frac{y_1}{Y}, \frac{x_2}{X}, \frac{y_2}{Y}\right),$$

where T, X, Y are positive parameters. Let $R = \frac{V}{XY}$ and $U = \max(T, R^{-1})$. Then

$$F(x_1, y_1, x_2, y_2) = \int_{\mathbb{R}^4} G(u_1, u_2, u_3, t)\left(\frac{x_1y_2}{x_2y_1}\right)^{it}\frac{du_1du_2du_3}{y_1^{iu_1}y_2^{iu_2}x_2^{iu_3}}\,dt,$$

where G (depending on T, V, X, Y) satisfies for any $A > 0$ the bound

$$|G(u_1, u_2, u_3, t)| \ll_A U^{-1}\left(1 + \frac{|t|}{U}\right)^{-A} \prod_{i=1}^{3} (1 + |u_i|)^{-A}. \tag{5.3}$$

REMARK. If $s \in \mathbb{C}$ and $\omega(x, s) = x^{s-1}\omega_V(x)$, then one may apply the lemma to $\omega(x, s)$, giving rise to a family of functions $G = G_s$. The proof shows that G_s satisfies (5.3) with an implied constant depending polynomially on s.

Proof of Lemma 5.2. By Mellin inversion,

$$F(x_1, y_1, x_2, y_2) = \int \widetilde{F}(s_1, u_1, s_2, u_2) x_1^{-s_1} y_1^{-u_1} x_2^{-s_2} y_2^{-u_2} \frac{ds_1\, du_1\, ds_2\, du_2}{(2\pi i)^4}, \tag{5.4}$$

where $\widetilde{F}(s_1, u_1, s_2, u_2)$ is defined by

$$\int_{\mathbb{R}^4_{>0}} \omega_V(x_1 y_2 - x_2 y_1) g\left(T \log \frac{x_1 y_2}{x_2 y_1}\right) w\left(\frac{x_1}{X}, \frac{y_1}{Y}, \frac{x_2}{X}, \frac{y_2}{Y}\right) \times x_1^{s_1} y_1^{u_1} x_2^{s_2} y_2^{u_2} \frac{dx_1\, dy_1\, dx_2\, dy_2}{x_1 y_1 x_2 y_2}. \tag{5.5}$$

In (5.5), change variables $x_1 \mapsto \frac{x_2 y_1}{y_2} x_1$ to find that $\widetilde{F}(s_1, s_2, s_2, s_4)$ equals

$$\int_{\mathbb{R}^4_{>0}} \omega_V\left(\frac{x_2 y_1}{XY} \frac{(x_1 - 1)}{R/V}\right) g(T \log x_1) w\left(\frac{x_1 x_2 y_1}{X y_2}, \frac{y_1}{Y}, \frac{x_2}{X}, \frac{y_2}{Y}\right) \times x_1^{s_1} y_1^{s_1 + u_1} x_2^{s_1 + s_2} y_2^{-s_1 + u_2} \frac{dx_1\, dy_1\, dx_2\, dy_2}{x_1 y_1 x_2 y_2}.$$

Now in (5.4), change variables $u_1 \mapsto u_1 - s_1$, $s_2 \mapsto s_2 - s_1$, and $u_2 \mapsto u_2 + s_1$ to get

$$F(x_1, y_1, x_2, y_2) = \int \widetilde{F}(s_1, u_1 - s_1, s_2 - s_1, u_2 + s_1) \left(\frac{x_1 y_2}{x_2 y_1}\right)^{-s_1} \frac{ds_1\, du_1\, ds_2\, du_2}{y_1^{u_1} x_2^{s_2} y_2^{u_2} (2\pi i)^4}, \tag{5.6}$$

where now $\widetilde{F}(s_1, u_1 - s_1, s_2 - s_1, u_2 + s_1)$ takes the form of $\widetilde{H}(s_1, u_1, s_2, u_2)$, with

$$H(x_1, y_1, x_2, y_2) = \omega_V\left(\frac{x_2 y_1}{XY} \frac{(x_1 - 1)}{R/V}\right) g(T \log x_1) w\left(\frac{x_1 x_2 y_1}{X y_2}, \frac{y_1}{Y}, \frac{x_2}{X}, \frac{y_2}{Y}\right).$$

It is easy to check that

$$H^{(j_1, k_1, j_2, k_2)}(x_1, y_1, x_2, y_2) \ll U^{j_1} X^{-j_2} Y^{-k_1 - k_2},$$

and that x_1 is concentrated on $x_1 = 1+O(\min(R, T^{-1}))$, whence integration by parts gives

$$\widetilde{H}(-it, u_1, u_3, u_2) \ll_A U^{-1}\left(1+\frac{|t|}{U}\right)^{-A} Y^{\mathrm{Re}(u_1+u_2)} X^{\mathrm{Re}(u_3)} \prod_{j=1}^{3}(1+|u_j|)^{-A}.$$

Taking $\mathrm{Re}(u_i) = 0$ and defining G on $\mathbb{R}^4$ appropriately completes the proof. ∎

5.2. Preparation. It is convenient to work with a couple modified norms that are closely related to (1.3). Define

$$\begin{aligned}(5.7)\qquad &\Delta_1(Q,k,T,N)\\ &= \max_{|\boldsymbol{\alpha}|=1} \int\limits_{T/2\le t\le T} \sum_{\substack{Q/2<q\le Q\\ (q,k)=1}} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)} \sideset{}{^*}\sum_{\theta \,(\mathrm{mod}\, k)} \Big| \sum_{\substack{N/2<ab\le N\\ (a,b)=1}} \alpha_{a,b}\lambda_{\chi\theta,t}(a,b)\Big|^2 \, dt.\end{aligned}$$

Clearly, $\Delta_1(Q,k,T,N) \le \Delta(Q,k,T,N)$, and in the other direction, we have

$$\Delta(Q,k,T,N) \le \sum_{j|k} \Delta_1(Q,j,T,N).$$

Secondly, define

$$\begin{aligned}(5.8)\qquad &\Delta_2(Q,k,T,N)\\ &= \max_{|\boldsymbol{\alpha}|=1} \int\limits_{T/2\le t\le T} \sum_{Q/2<q\le Q} \sideset{}{^*}\sum_{\psi \,(\mathrm{mod}\, qk)} \Big| \sum_{\substack{N/2<ab\le N\\ (a,b)=1}} \alpha_{a,b}\lambda_{\psi,t}(a,b)\Big|^2 \, dt.\end{aligned}$$

It is easy to see that $\Delta_1(Q,k,T,N) \le \Delta_2(Q,k,T,N)$, since when $(q,k)=1$, the map $(\chi,\theta)\mapsto \chi\theta$ is a bijection onto the set of primitive characters modulo qk. After having done this, we arrive at (5.8) by dropping the condition $(q,k)=1$, by positivity. For the proof of Theorem 1.6, we will bound the norm Δ_2. Indeed, we can deduce Theorem 1.6 from the bound

$$(5.9)\qquad \Delta_2(Q,k,T,N) \lesssim Q^2kT + \frac{Q^2kT}{N}\,\overline{\Delta'}\left(\frac{N}{kQT}, k, T, N\right).$$

Let w be a nonnegative smooth weight function with $w(x)\ge 1$ for $1/2 \le x \le 1$, and $w(x) = 0$ for $x < 1/4$ and for $x \ge 2$. Then $\Delta_2(Q,k,T,N) \le \max_{|\alpha|=1} S$, where

$$S = \int_{-\infty}^{\infty} w\left(\frac{t}{T}\right) \sum_q w\left(\frac{q}{Q}\right) \sideset{}{^*}\sum_{\psi \,(\mathrm{mod}\, qk)} \frac{qk}{\varphi(qk)}\Big| \sum_{\substack{(a,b)=1\\ N/2<ab\le N}} \alpha_{a,b}\psi(a\overline{b})(a/b)^{it}\Big|^2 \, dt.$$

We will assume that $\alpha_{a,b}$ is supported on

$$(5.10)\qquad N/2 < ab \le N, \qquad \text{where } (ab,k)=1 \text{ and } (a,b)=1.$$

A simple argument with a dyadic partition of unity and Cauchy's inequality shows that

$$\Big|\sum_{a,b} \alpha_{a,b}\Big|^2 = \Big|\sum_{\substack{N_1 N_2 \asymp N \\ \text{dyadic}}} \sum_{\substack{a \asymp N_1 \\ b \asymp N_2}} \alpha_{a,b}\Big|^2 \ll (\log N) \cdot \sum_{\substack{N_1 N_2 \asymp N \\ \text{dyadic}}} \Big|\sum_{\substack{a \asymp N_1 \\ b \asymp N_2}} \alpha_{a,b}\Big|^2.$$

Hence, in the proof of Theorem 1.6, we may assume that a and b are each supported in dyadic ranges, say $a \asymp N_1$ and $b \asymp N_2$.

Let $1 \leq Y \leq \frac{Q}{100}$ be a parameter to be chosen later. For $\psi \pmod{qk}$, write $qk = q'(dk)$, where $d \mid k^\infty$ and $(q', k) = 1$, and write $\psi = \psi_k \psi'$, where ψ_k has modulus dk and ψ' has modulus q'. Let $m_k(\psi) = dk$ denote the modulus of the k-part of ψ, and $\mathrm{cond}_{q'}(\psi)$ denote the conductor of ψ', i.e., the coprime-to-k part of ψ. Then $S \leq S_{>Y}$, where

$$S_{>Y} = \int_{-\infty}^{\infty} w\Big(\frac{t}{T}\Big) \sum_q w\Big(\frac{q}{Q}\Big) \sum_{\substack{\psi \,(\mathrm{mod}\, qk) \\ \mathrm{cond}_{q'}(\psi) m_k(\psi) > Yk}} \frac{qk}{\varphi(qk)} \Big|\sum_{a,b} \alpha_{a,b} \psi(a\overline{b})(a/b)^{it}\Big|^2 dt,$$

by positivity, since if ψ is primitive modulo qk, then $\mathrm{cond}_{q'}(\psi) m_k(\psi) = \mathrm{cond}(\psi) = qk$. This uses the fact that the condition $qk > Yk$ is redundant on the support of $w(q/Q)$.

By inclusion-exclusion, we have $S_{>Y} = S_{\leq \infty} - S_{\leq Y}$, where for $* \in \{Y, \infty\}$, $S_{\leq *}$ corresponds to the sum over $\mathrm{cond}_{q'}(\psi) m_k(\psi)/k \leq *$. We will write S_∞ as an alias for $S_{\leq \infty}$.

We begin with some arithmetic manipulations that are in common between S_∞ and $S_{\leq Y}$. Opening the square, we have

$$S_{\leq *} = \int_{-\infty}^{\infty} w\Big(\frac{t}{T}\Big) \sum_q w\Big(\frac{q}{Q}\Big)$$
$$\times \sum_{\substack{\psi \,(\mathrm{mod}\, qk) \\ \mathrm{cond}_{q'}(\psi) m_k(\psi)/k \leq *}} \frac{qk}{\varphi(qk)} \sum_{\substack{a_1, b_1 \\ a_2, b_2}} \alpha_{a_1,b_1} \overline{\alpha_{a_2,b_2}} \psi(a_1 b_2 \overline{b_1 a_2}) \Big(\frac{a_1 b_2}{b_1 a_2}\Big)^{it} dt.$$

Define

$$g_1 = (a_1, a_2), \qquad g_2 = (b_1, b_2), \qquad g_3 = (a_1, b_2), \qquad g_4 = (b_1, a_2), \tag{5.11}$$

and note that the g_i are pairwise coprime since $(a_1, b_1) = (a_2, b_2) = 1$ by the support of α (recall (5.10)). Then change variables

$$\begin{aligned} a_1 &\mapsto g_1 g_3 h_{11} h_{13} a_1, && \text{where } (a_1, g_1 g_3) = 1, \\ a_2 &\mapsto g_1 g_4 h_{21} h_{24} a_2, && \text{where } (a_2, g_1 g_4) = 1, \\ b_1 &\mapsto g_2 g_4 h_{32} h_{34} b_1, && \text{where } (b_1, g_2 g_4) = 1, \\ b_2 &\mapsto g_2 g_3 h_{42} h_{43} b_2, && \text{where } (b_2, g_2 g_3) = 1, \end{aligned} \tag{5.12}$$

and where

$$h_{ij} \mid g_j^\infty \quad \text{for all } i, j, \quad \text{and} \quad (h_{ij}, h_{kj}) = 1 \quad \text{for } i \neq k. \tag{5.13}$$

The conditions (5.11) translate into

$$(a_1 b_1, a_2 b_2) = 1.$$

Moreover, the conditions $(a_1, g_1 g_3) = 1, \dots, (b_2, g_2 g_3) = 1$ in (5.12) may be expressed succinctly as $(a_1 a_2 b_1 b_2, g_1 g_2 g_3 g_4) = 1$, since prior to (5.12) we had $(a_i, b_i) = 1$ from (5.10). Let

$$\mathbf{g} = (g_1, g_2, g_3, g_4, h_{11}, h_{13}, h_{21}, h_{24}, h_{32}, h_{34}, h_{42}, h_{43}), \tag{5.14}$$

where the h_{ij} satisfy (5.13). In addition, let

$$\beta_{13} = g_1 g_3 h_{11} h_{13}, \quad \beta_{23} = g_2 g_3 h_{42} h_{43},$$
$$\beta_{14} = g_1 g_4 h_{21} h_{24}, \quad \beta_{24} = g_2 g_4 h_{32} h_{34},$$

and

$$\gamma_1 = g_3^2 h_{11} h_{42} h_{13} h_{43} = \frac{\beta_{13}\beta_{23}}{g_1 g_2} \quad \text{and} \quad \gamma_2 = g_4^2 h_{21} h_{32} h_{24} h_{34} = \frac{\beta_{14}\beta_{24}}{g_1 g_2}.$$

Observe that $(\gamma_1, \gamma_2) = 1$ since the g_i are pairwise coprime, and by (5.13). With these substitutions, we obtain

$$S_{\leq *} = \sum_{\mathbf{g}} \int_{-\infty}^{\infty} w\left(\frac{t}{T}\right) \sum_{(q, g_1 g_2)=1} w\left(\frac{q}{Q}\right) \sum_{\substack{\psi \,(\mathrm{mod}\, qk) \\ \mathrm{cond}_{q'}(\psi) m_k(\psi)/k \leq *}} \frac{qk}{\varphi(qk)} \tag{5.15}$$
$$\times \sum_{\substack{(a_1 b_1, a_2 b_2)=1 \\ (\mathbf{a}, \mathbf{g})=1}} \alpha^{(1,\mathbf{g})}_{a_1, b_1} \overline{\alpha}^{(2,\mathbf{g})}_{a_2, b_2} \psi\left(\frac{\gamma_1 a_1 b_2}{\gamma_2 b_1 a_2}\right) \left(\frac{\gamma_1 a_1 b_2}{\gamma_2 b_1 a_2}\right)^{it} dt,$$

where

$$\alpha^{(1,\mathbf{g})}_{a_1, b_1} \overline{\alpha}^{(2,\mathbf{g})}_{a_2, b_2} = \alpha_{\beta_{13} a_1, \beta_{24} b_1} \overline{\alpha}_{\beta_{14} a_2, \beta_{23} b_2}, \tag{5.16}$$

and where the condition $(\mathbf{a}, \mathbf{g}) = 1$ is shorthand for $(a_1 a_2 b_1 b_2, g_1 g_2 g_3 g_4) = 1$. There are additional conditions that are implicitly enforced by (5.10), which we will recall only as needed. For later use, note

$$\gamma_1 a_1 b_2 \asymp \gamma_2 a_2 b_1 \asymp \frac{N}{g_1 g_2}. \tag{5.17}$$

Moreover, we claim that

$$\sum_{\mathbf{g}, a_1, b_1} |\alpha^{(1,\mathbf{g})}_{a_1, b_1}|^2 \lesssim |\alpha|^2, \tag{5.18}$$

and similarly for $\alpha^{(2,\mathbf{g})}_{a_2, b_2}$. To see this, note that the variables g_1, g_2, g_3, g_4 appear as divisors of β_{13} or β_{24}, and similarly for half of the h_{ij} variables (namely, h_{11}, h_{13}, h_{32}, and h_{34}). For the remaining h_{ij} variables, we recall from (5.13) that $h_{12} \mid g_2$, etc., so these variables range over a set of cardinality $\ll N^{\varepsilon}$. Then (5.18) follows easily.

5.3. Direct method. In this section we estimate $S_{\leq Y}$ by reducing to an instance of the original norm, but with smaller parameters.

PROPOSITION 5.3. *We have* $S_{\leq Y} = S_{\leq Y}^{(0)} + S'_{\leq Y}$, *where* $S_{\leq Y}^{(0)}$ *is given by* (5.22) *below, and where*

$$S'_{\leq Y} \lesssim \max_{\substack{Y' \leq Y \\ r_k \mid k^{\infty}}} \Delta(Y'/r_k, r_k k, 2T, N)|\alpha|^2. \tag{5.19}$$

Proof. We pick up from (5.15). Write $q = r_k q'$ where $r_k \mid k^{\infty}$ and $(q', k) = 1$, and write $\psi = \chi\theta$ where θ runs modulo $r_k k$ and χ runs modulo q'. Then

$$\begin{aligned} S_{\leq Y} &= \sum_{\mathbf{g}} \sum_{r_k \mid k^{\infty}} \int_{-\infty}^{\infty} w\Big(\frac{t}{T}\Big) \sum_{(q', k g_1 g_2)=1} w\Big(\frac{q' r_k}{Q}\Big) \sum_{\theta \,(\mathrm{mod}\, r_k k)} \sum_{\substack{\chi \,(\mathrm{mod}\, q') \\ \mathrm{cond}(\chi) \leq Y/r_k}} \frac{q' k}{\varphi(q')\varphi(k)} \\ &\quad \times \sum_{\substack{(a_1 b_1, a_2 b_2)=1 \\ (\mathbf{a}, \mathbf{g})=1}} \alpha_{a_1,b_1}^{(1,\mathbf{g})} \overline{\alpha}_{a_2,b_2}^{(2,\mathbf{g})} \chi\theta\Big(\frac{\gamma_1 a_1 b_2}{\gamma_2 b_1 a_2}\Big) \Big(\frac{\gamma_1 a_1 b_2}{\gamma_2 b_1 a_2}\Big)^{it} dt. \end{aligned} \tag{5.20}$$

We next replace q' by $q' q_0 q_1$ where q' is the conductor of χ, $(q_0, q') = 1$, and $q_1 \mid (q')^{\infty}$, and correspondingly write $\chi = \chi' \chi_0$ where χ' is primitive modulo q', and χ_0 is trivial modulo q_0. Applying this substitution in (5.20), we obtain

$$\begin{aligned} S_{\leq Y} &= \sum_{\mathbf{g}} \sum_{r_k \mid k^{\infty}} \int_{-\infty}^{\infty} w\Big(\frac{t}{T}\Big) \sum_{\substack{(q', k g_1 g_2)=1 \\ q' \leq Y/r_k}} \sum_{\theta \,(\mathrm{mod}\, r_k k)} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q')} \frac{q' k}{\varphi(q')\varphi(k)} \\ &\times \sum_{\substack{(q_0, q' k \mathbf{g})=1 \\ q_1 \mid (q')^{\infty}}} w\Big(\frac{q' q_0 q_1 r_k}{Q}\Big) \frac{q_0}{\varphi(q_0)} \sum_{\substack{(a_1 b_1, a_2 b_2)=1 \\ (\mathbf{a}, q_0 \mathbf{g})=1}} \alpha_{a_1,b_1}^{(1,\mathbf{g})} \overline{\alpha}_{a_2,b_2}^{(2,\mathbf{g})} \chi'\theta\Big(\frac{\gamma_1 a_1 b_2}{\gamma_2 b_1 a_2}\Big) \Big(\frac{\gamma_1 a_1 b_2}{\gamma_2 b_1 a_2}\Big)^{it} dt. \end{aligned}$$

By Mellin inversion, and evaluating the sums over q_0 and q_1 with Lemma 5.1, the second line above equals

$$\begin{aligned} &\sum_{\substack{(a_1 b_1, a_2 b_2)=1 \\ (\mathbf{a}, \mathbf{g})=1}} \alpha_{a_1,b_1}^{(\mathbf{g})} \overline{\alpha}_{a_2,b_2}^{(2,\mathbf{g})} \chi'\theta\Big(\frac{\gamma_1 a_1 b_2}{\gamma_2 b_1 a_2}\Big) \Big(\frac{\gamma_1 a_1 b_2}{\gamma_2 b_1 a_2}\Big)^{it} \\ &\qquad\qquad \times \frac{1}{2\pi i} \int_{(2)} \Big(\frac{Q}{r_k q'}\Big)^s \widetilde{w}(s) Z_{q', k\mathbf{g}\mathbf{a}}(s)\, ds\, dt. \end{aligned}$$

Since $k, g_1, g_2, g_3, g_4, a_1, a_2, b_1, b_2$ are pairwise coprime, we have

$$Z_{q',k\mathbf{ga}}(s) = Z_{1,1}(s)\nu_{q'}(s)\delta_{k\mathbf{g}}(s)\delta_{a_1b_1}(s)\delta_{a_2b_2}(s), \tag{5.21}$$

which is an important separation of variables.

Using the meromorphic continuation of $Z_{1,1}(s)$ provided by Lemma 5.1, we shift the contour of integration to the line $\mathrm{Re}(s) = \varepsilon$, passing a pole at $s = 1$. Let $S^{(0)}_{\le Y}$ denote the residue term, which is given by

(5.22)
$$S^{(0)}_{\le Y} = \sum_{r_k \mid k^\infty} \sum_{\mathbf{g}} \delta_{k\mathbf{g}} \int_{-\infty}^{\infty} w\left(\frac{t}{T}\right) \sum_{\theta\,(\mathrm{mod}\,r_k k)} \sum_{\substack{(q',kg_1g_2)=1\\ q' \le Y/r_k}} \sideset{}{^*}\sum_{\chi'\,(\mathrm{mod}\,q')} \frac{Q\widetilde{w}(1)Z_{1,1}\nu_{q'}k}{r_k\varphi(q')\varphi(k)}$$
$$\times \sum_{\substack{(a_1b_1,a_2b_2)=1\\ (\mathbf{a},\mathbf{g})=1}} \delta_{a_1b_1}\alpha^{(1,\mathbf{g})}_{a_1,b_1}\delta_{a_2b_2}\overline{\alpha}^{(2,\mathbf{g})}_{a_2,b_2}\chi'\theta\left(\frac{\gamma_1a_1b_2}{\gamma_2b_1a_2}\right)\left(\frac{\gamma_1a_1b_2}{\gamma_2b_1a_2}\right)^{it} dt,$$

where $Z_{1,1}$ denotes $\mathrm{Res}_{s=1} Z_{1,1}(s)$, $\nu_{q'}$ denotes $\nu_{q'}(1)$, and $\delta_n = \delta_n(1)$.

By the triangle inequality, and some simple bounds, we have

$$|S'_{\le Y}| \lesssim \sum_{r_k \mid k^\infty} r_k^{-\varepsilon} \max_{\mathrm{Re}(s)=\varepsilon} \sum_{\mathbf{g}} \int_{-2T}^{2T} \sum_{\theta\,(\mathrm{mod}\,r_k k)} \sum_{\substack{(q',kg_1g_2)=1\\ q' \le Y/r_k}} \sideset{}{^*}\sum_{\chi'\,(\mathrm{mod}\,q')} \tag{5.23}$$
$$\times \Bigg| \sum_{\substack{(a_1b_1,a_2b_2)=1\\ (\mathbf{a},\mathbf{g})=1}} \delta_{a_1b_1}(s)\alpha^{(1,\mathbf{g})}_{a_1,b_1}\delta_{a_2b_2}(s)\overline{\alpha}^{(2,\mathbf{g})}_{a_2,b_2}\chi'\theta\left(\frac{\gamma_1a_1b_2}{\gamma_2b_1a_2}\right)\left(\frac{\gamma_1a_1b_2}{\gamma_2b_1a_2}\right)^{it}\Bigg| \, dt.$$

Note $|\chi'\theta(\gamma_1\overline{\gamma_2})(\gamma_1/\gamma_2)^{it}| \le 1$, which may be used to simplify this bound. To show the desired bound (5.19), we state and prove Lemma 5.4 below, as it will be useful later as well. ∎

LEMMA 5.4. *Let $\gamma^{(1)}_{a,b}$ and $\gamma^{(2)}_{a,b}$ be sequences of complex numbers supported on $ab \asymp M$, $(a,b) = 1$. Consider an expression of the form $\mathcal{S}_\gamma(Q,k,T,M)$ defined by*

$$\int_{-T}^{T} \sum_{\theta\,(\mathrm{mod}\,k)} \sum_{\substack{(q,k)=1\\ Q/2<q\le Q}} \sideset{}{^*}\sum_{\chi\,(\mathrm{mod}\,q)} \Bigg| \sum_{(a_1b_1,a_2b_2)=1} \gamma^{(1)}_{a_1,b_1}\overline{\gamma}^{(2)}_{a_2,b_2}\chi\theta\left(\frac{a_1b_2}{b_1a_2}\right)\left(\frac{a_1b_2}{b_1a_2}\right)^{it}\Bigg| \, dt.$$

Then

$$\mathcal{S}_\gamma(Q,k,T,M) \lesssim \overline{\Delta}(Q,k,T,M)\max_{i=1,2} |\gamma^{(i)}|^2.$$

Proof. To separate the inner variables, we use Möbius inversion in the form

(5.24)
$$\delta((a_1b_1, a_2b_2) = 1) = \sum_{e_1 \mid (a_1,a_2)} \sum_{e_2 \mid (a_1,b_2)} \sum_{e_3 \mid (b_1,a_2)} \sum_{e_4 \mid (b_1,b_2)} \mu(e_1)\mu(e_2)\mu(e_3)\mu(e_4).$$

The e_i are pairwise coprime, by the support of γ. Thus

$$\mathcal{S}_\gamma(Q,k,T,M) \ll \sum_{e_1,e_2,e_3,e_4} \int_{-T}^{T} \sum_{\theta \,(\mathrm{mod}\, k)} \sum_{\substack{(q,k)=1 \\ Q/2<q\le Q}} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)} |\mathcal{A}_1\mathcal{A}_2| \, dt,$$

where

$$\mathcal{A}_1 = \sum_{\substack{a_1 \equiv 0 \,(\mathrm{mod}\, e_1e_2) \\ b_1 \equiv 0 \,(\mathrm{mod}\, e_3e_4)}} \gamma_{a_1,b_1} \chi\theta(a_1\overline{b_1}) \left(\frac{a_1}{b_1}\right)^{it},$$

and $\mathcal{A}_2$ has a similar definition. Lemma 5.4 follows by using $|\mathcal{A}_1\mathcal{A}_2| \ll |\mathcal{A}_1|^2 + |\mathcal{A}_2|^2$ and monotonicity (Lemma 2.1). ∎

5.4. Divisor switching method

PROPOSITION 5.5. *We have a decomposition*

$$S_\infty = S_\infty^{(0)} + S_\infty' + S_\infty^{\mathrm{diag}} + \mathcal{E}_\infty$$

with the following properties. The term $S_\infty^{(0)}$ is given by (5.34) *below, and S_∞' satisfies the bound*

$$|S_\infty'| \lesssim \frac{Q^2kT}{N} \overline{\Delta'}\left(\frac{N}{kQT}, k, T, N\right)|\alpha|^2. \tag{5.25}$$

The diagonal term satisfies the bound

$$|S_\infty^{\mathrm{diag}}| \ll Q^2kT|\alpha|^2, \tag{5.26}$$

and the term $\mathcal{E}_\infty$ is negligibly small.

Proof. We carry on with (5.15) and apply orthogonality of characters to the sum over ψ. This picks out the congruence $\gamma_1a_1b_2 \equiv \gamma_2a_2b_1 \pmod{kq}$, but with a side condition $(\gamma_1\gamma_2a_1a_2b_1b_2, kq) = 1$. This side condition can be dropped, since the congruence $\gamma_1a_1b_2 \equiv \gamma_2a_2b_1 \pmod{kq}$, combined with $(\gamma_1a_1b_2, \gamma_2a_2b_1) = 1$, implies that $(\gamma_1\gamma_2a_1b_2a_2b_1, kq) = 1$. Additionally evaluating the t-integral, in all we obtain

$$S_\infty = QkT \sum_{\mathbf{g}} \sum_{(q,g_1g_2)=1} w_1\left(\frac{q}{Q}\right) \sum_{\substack{(a_1b_1,a_2b_2)=(\mathbf{a},\mathbf{g})=1 \\ \gamma_1a_1b_2 \equiv \gamma_2a_2b_1 \,(\mathrm{mod}\, kq)}} \alpha_{a_1,b_1}^{(1,\mathbf{g})} \overline{\alpha}_{a_2,b_2}^{(2,\mathbf{g})} \widehat{w}\left(T \log \frac{\gamma_1a_1b_2}{\gamma_2a_2b_1}\right),$$

where $w_1(x) = xw(x)$ and $\widehat{w}(x) = \int_{-\infty}^{\infty} w(t)e^{ixt}\, dt$.

Let S_∞^{diag} be the contribution to S_∞ from the diagonal $\gamma_1a_1b_2 = \gamma_2a_2b_1$. Since $(\gamma_1a_1b_2, \gamma_2a_2b_1) = 1$, this forces $\gamma_i = a_i = b_i = 1$ for $i = 1, 2$. Hence, recalling (5.16), we obtain

$$S_\infty^{\mathrm{diag}} \ll Q^2 kT \sum_{g_1,g_2} |\alpha_{g_1,g_2}|^2 = Q^2 kT |\alpha|^2. \tag{5.27}$$

Let $S''_\infty = S_\infty - S_\infty^{\mathrm{diag}}$ be the nondiagonal portion of S_∞. Write $\gamma_1 a_1 b_2 = \gamma_2 a_2 b_1 + qkr$, where $r \neq 0$. Additionally, we detect the condition $(q, g_1 g_2) = 1$ by Möbius inversion in the form $\sum_{d|(q,g_1g_2)} \mu(d)$, and substitute $q = de$. This gives

$$\begin{aligned} S''_\infty = QkT \sum_{\mathbf{g}} \sum_{d|g_1g_2} \mu(d) \sum_e w_1\left(\frac{de}{Q}\right) \\ \times \sum_{r\in\mathbb{Z}\setminus\{0\}} \sum_{\substack{(a_1b_1,a_2b_2)=(\mathbf{a},\mathbf{g})=1 \\ \gamma_1a_1b_2-\gamma_2a_2b_1=dekr}} \alpha^{(1,\mathbf{g})}_{a_1,b_1} \overline{\alpha}^{(2,\mathbf{g})}_{a_2,b_2} \widehat{w}\left(T \log \frac{\gamma_1 a_1 b_2}{\gamma_2 a_2 b_1}\right). \end{aligned}$$

Now we perform the divisor switch: re-write $\gamma_1 a_1 b_2 - \gamma_2 a_2 b_1 = dekr$ as

$$\gamma_1 a_1 b_2 \equiv \gamma_2 a_2 b_1 \pmod{dk|r|}, \qquad e = \frac{\gamma_1 a_1 b_2 - \gamma_2 a_2 b_1}{dkr}. \tag{5.28}$$

It is convenient to record that the side condition

$$(\gamma_1\gamma_2 a_1 a_2 b_1 b_2, dkr) = 1 \tag{5.29}$$

follows from the congruence (5.28) together with the coprimality

$$(\gamma_1 a_1 b_2, \gamma_2 a_2 b_1) = 1.$$

We also factor r as

$$r = r_0 r_1, \qquad r_0 \mid (kg_1g_2)^\infty,\ (r_1, kg_1g_2) = 1.$$

With these substitutions, we obtain

$$\begin{aligned} S''_\infty = QkT \sum_{\mathbf{g}} \sum_{\substack{d|g_1g_2 \\ r_0|(kg_1g_2)^\infty}} \mu(d) \sum_{\substack{r_1\in\mathbb{Z}\setminus\{0\} \\ (r_1,kg_1g_2)=1}} w_1\left(\frac{\gamma_1 a_1 b_2 - \gamma_2 a_2 b_1}{kr_0r_1Q}\right) \\ \times \sum_{\substack{(a_1b_1,a_2b_2)=(\mathbf{a},\mathbf{g})=1 \\ \gamma_1a_1b_2\equiv\gamma_2a_2b_1 \,(\mathrm{mod}\, dkr_0) \\ \gamma_1a_1b_2\equiv\gamma_2a_2b_1 \,(\mathrm{mod}\, |r_1|)}} \alpha^{(1,\mathbf{g})}_{a_1,b_1} \overline{\alpha}^{(2,\mathbf{g})}_{a_2,b_2} \widehat{w}\left(T \log \frac{\gamma_1 a_1 b_2}{\gamma_2 a_2 b_1}\right). \end{aligned}$$

Next we express the congruences using Dirichlet characters modulo dkr_0 and $|r_1|$; this is enabled by the side condition (5.29). This leads to

$$S''_\infty = QkT \sum_{\mathbf{g}} \sum_{\substack{d|g_1g_2 \\ r_0|(kg_1g_2)^\infty}} \frac{\mu(d)}{\varphi(dkr_0)} \sum_{\theta \,(\mathrm{mod}\, dkr_0)} \sum_{\substack{r_1 \in \mathbb{Z}\setminus\{0\} \\ (r_1, kg_1g_2)=1}} \frac{1}{\varphi(|r_1|)} \sum_{\chi \,(\mathrm{mod}\, |r_1|)}$$
$$\times \sum_{\substack{(a_1b_1, a_2b_2)=1 \\ (\mathbf{a},\mathbf{g})=1}} \alpha^{(1,\mathbf{g})}_{a_1,b_1} \overline{\alpha}^{(2,\mathbf{g})}_{a_2,b_2} \chi\theta\Big(\frac{\gamma_1 a_1 b_2}{\gamma_2 a_2 b_1}\Big) w_1\Big(\frac{\gamma_1 a_1 b_2 - \gamma_2 a_2 b_1}{kr_0 r_1 Q}\Big) \widehat{w}\Big(T \log \frac{\gamma_1 a_1 b_2}{\gamma_2 a_2 b_1}\Big).$$

The characters of varying modulus need to be primitive, so we substitute

$$r_1 \mapsto r_1 r_2 q', \qquad \chi \mapsto \chi_0 \chi,$$

where $r_1 \,|\, (q')^\infty$, $(r_2, q') = 1$, χ is primitive of modulus $|q'|$, and χ_0 is trivial modulo r_2. With this, we obtain

$$S''_\infty = QkT \sum_{\mathbf{g}} \sum_{\substack{d|g_1g_2 \\ r_0|(kg_1g_2)^\infty}} \frac{\mu(d)}{\varphi(dkr_0)}$$
$$\times \sum_{\theta \,(\mathrm{mod}\, dkr_0)} \sum_{\substack{q' \neq 0 \\ (q', kg_1g_2)=1}} \sum_{\substack{r_1|(q')^\infty \\ (r_2, q'k\mathbf{g})=1}} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, |q'|)} \sum_{\substack{(a_1b_1, a_2b_2)=1 \\ (a_1a_2b_1b_2, \mathbf{g}r_2)=1}} \alpha^{(1,\mathbf{g})}_{a_1,b_1} \overline{\alpha}^{(2,\mathbf{g})}_{a_2,b_2} \chi\theta\Big(\frac{\gamma_1 a_1 b_2}{\gamma_2 a_2 b_1}\Big)$$
$$\times \frac{w_1\big(\frac{\gamma_1 a_1 b_2 - \gamma_2 a_2 b_1}{kr_0 r_1 r_2 q' Q}\big)}{\varphi(r_1 r_2 |q'|)} \widehat{w}\Big(T \log \frac{\gamma_1 a_1 b_2}{\gamma_2 a_2 b_1}\Big).$$

Let $w_1(x) = x^{-1} w_2(x)$, so $w_2(x) = x^2 w(x)$, and $\widetilde{w_2}(-s) = \widetilde{w}(2-s)$. In addition, apply the Mellin inversion formula to w_2. Then we deduce that S''_∞ equals

$$Q^2 kT \sum_{\mathbf{g}} \sum_{\substack{d|g_1g_2 \\ r_0|(kg_1g_2)^\infty}} \frac{\mu(d) kr_0}{\varphi(dkr_0)} \sum_{\theta \,(\mathrm{mod}\, dkr_0)} \sum_{(q', kg_1g_2)=1} \sum_{\substack{r_1|(q')^\infty \\ (r_2, q'k\mathbf{g})=1}} \frac{r_2 |q'|}{\varphi(r_2)\varphi(|q'|)}$$
$$\times \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, |q'|)} \int_{(2)} \widetilde{w}(2-s) \sum_{\substack{(a_1b_1, a_2b_2)=1 \\ (a_1a_2b_1b_2, \mathbf{g}r_2)=1}} \Big(\frac{\gamma_1 a_1 b_2 - \gamma_2 a_2 b_1}{kr_0 r_1 r_2 q' Q}\Big)^s \frac{(\mathrm{sgn}) \alpha^{(1,\mathbf{g})}_{a_1,b_1} \overline{\alpha}^{(2,\mathbf{g})}_{a_2,b_2}}{|\gamma_1 a_1 b_2 - \gamma_2 a_2 b_1|}$$
$$\times \chi\theta\Big(\frac{\gamma_1 a_1 b_2}{\gamma_2 a_2 b_1}\Big) \widehat{w}\Big(T \log \frac{\gamma_1 a_1 b_2}{\gamma_2 a_2 b_1}\Big) \frac{ds}{2\pi i},$$

where (sgn) is shorthand for the indicator function of

$$\mathrm{sgn}(q') = \mathrm{sgn}(\gamma_1 a_1 b_2 - \gamma_2 a_2 b_1). \tag{5.30}$$

Prior to the Mellin inversion formula, (5.30) was enforced by the support of w_2.

The sums over r_1 and r_2 evaluate exactly as in (5.21). Thus

$$S''_\infty = Q^2kT \sum_{\mathbf{g}} \sum_{\substack{d|g_1g_2 \\ r_0|(kg_1g_2)^\infty}} \frac{\mu(d)kr_0}{\varphi(dkr_0)} \sum_{\theta \,(\mathrm{mod}\, dkr_0)} \sum_{\substack{q'\neq 0 \\ (q',kg_1g_2)=1}} \frac{|q'|}{\varphi(|q'|)}$$
$$\times \sideset{}{^*}\sum_{\chi\,(\mathrm{mod}\,|q'|)} \int_{(2)} \widetilde{w}(2-s)Z_{1,1}(s)\nu_{q'}(s)\delta_{\mathbf{g}k}(s) \sum_{\substack{(a_1b_1,a_2b_2)=1 \\ (\mathbf{a},\mathbf{g})=1}} \left(\frac{\gamma_1a_1b_2-\gamma_2a_2b_1}{kr_0q'Q}\right)^s$$
$$\times \frac{(\mathrm{sgn})\delta_{a_1b_1}(s)\delta_{a_2b_2}(s)\alpha^{(1,\mathbf{g})}_{a_1,b_1}\overline{\alpha}^{(2,\mathbf{g})}_{a_2,b_2}}{|\gamma_1a_1b_2-\gamma_2a_2b_1|}\chi\theta\left(\frac{\gamma_1a_1b_2}{\gamma_2a_2b_1}\right)\widehat{w}\left(T\log\frac{\gamma_1a_1b_2}{\gamma_2a_2b_1}\right)\frac{ds}{2\pi i}.$$

Now we apply a dyadic partition of unity of the form

$$1 = \sum_{V \text{ dyadic}} \omega\left(\frac{\gamma_1a_1b_2-\gamma_2a_2b_1}{V}\right),$$

where ω is smooth, even, and supported on $[1,2]\cup[-2,-1]$. By the rapid decay of $\widehat{w}$, and recalling (5.17), note that

$$\widehat{w}\left(T\log\frac{\gamma_1a_1b_2}{\gamma_2a_2b_1}\right) \ll \left(1+T\frac{|\gamma_1a_1b_2-\gamma_2a_2b_1|}{\gamma_2a_2b_1}\right)^{-A}$$
$$\ll \left(1+T\frac{|\gamma_1a_1b_2-\gamma_2a_2b_1|}{N/(g_1g_2)}\right)^{-A}.$$

Therefore, we may assume that

$$1 \ll V \le V_{\max} = \frac{N}{g_1g_2T}(QkTN)^\varepsilon, \tag{5.31}$$

absorbing $V > V_{\max}$ into the error term $\mathcal{E}_\infty$.

With this partition, we obtain

$$S''_\infty = Q^2kT \sum_{\mathbf{g}} \sum_{1\ll V\le V_{\max}} V^{-1} \sum_{\substack{d|g_1g_2 \\ r_0|(kg_1g_2)^\infty}} \frac{\mu(d)kr_0}{\varphi(dkr_0)} \sum_{\theta\,(\mathrm{mod}\,dkr_0)} \sum_{\substack{q'\neq 0 \\ (q',kg_1g_2)=1}} \frac{|q'|}{\varphi(|q'|)}$$
$$\times \sideset{}{^*}\sum_{\chi\,(\mathrm{mod}\,|q'|)} \frac{1}{2\pi i}\int_{(2)} \left(\frac{V}{kr_0|q'|Q}\right)^s \widetilde{w}(2-s)Z_{1,1}(s)\nu_{q'}(s)\delta_{\mathbf{g}k}(s)$$
$$\times \sum_{\substack{(a_1b_1,a_2b_2)=1 \\ (\mathbf{a},\mathbf{g})=1}} (\mathrm{sgn})\delta_{a_1b_1}(s)\delta_{a_2b_2}(s)\alpha^{(1,\mathbf{g})}_{a_1,b_1}\overline{\alpha}^{(2,\mathbf{g})}_{a_2,b_2}\chi\theta\left(\frac{\gamma_1a_1b_2}{\gamma_2a_2b_1}\right)$$
$$\times\, \omega_s\left(\frac{\gamma_1a_1b_2-\gamma_2a_2b_1}{V}\right)\widehat{w}\left(T\log\frac{\gamma_1a_1b_2}{\gamma_2a_2b_1}\right)ds,$$

where $\omega_s(x) = x^{s-1}\omega(x)$. By shifting the contour far to the right, q' may be truncated at

$$|q'| \leq Q^* := \frac{V}{kr_0 Q}(QkTN)^{\varepsilon}. \tag{5.32}$$

We next want to apply Lemma 5.2. Note that

$$\gamma_1 a_1 b_2 = \frac{\beta_{13}a_1}{g_1}\frac{\beta_{23}b_2}{g_2}, \qquad \gamma_2 a_2 b_1 = \frac{\beta_{14}a_2}{g_1}\frac{\beta_{24}b_1}{g_2},$$

where the support of α implies $\beta_{13}a_1 \asymp \beta_{14}a_2 \asymp N_1$ and $\beta_{23}b_2 \asymp \beta_{24}b_1 \asymp N_2$. We may then freely attach a redundant weight function of the form

$$w\Big(\frac{\beta_{13}a_1}{N_1}, \frac{\beta_{24}b_1}{N_2}, \frac{\beta_{14}a_2}{N_1}, \frac{\beta_{23}b_2}{N_2}\Big).$$

Now this is set up to apply Lemma 5.2 with $x_1 = g_1^{-1}\beta_{13}a_1$, $y_1 = g_1^{-1}\beta_{24}b_1$, $x_2 = g_1^{-1}\beta_{14}a_2$, $y_2 = g_2^{-1}\beta_{23}b_2$, $X = N_1/g_1$, $Y = N_2/g_2$, and with $\omega = \omega_s$. Observe that with this substitution, $\gamma_1 a_1 b_2 - \gamma_2 a_2 b_1 = x_1 y_2 - x_2 y_1$, as desired. This gives

$$\begin{aligned} S''_\infty = Q^2kT \sum_{\mathbf{g}} \sum_{1 \ll V \leq V_{\max}} V^{-1} \int_{\mathbb{R}^4} G_s(u_1,u_2,u_3,t) \sum_{\substack{d \mid g_1 g_2 \\ r_0 \mid (kg_1g_2)^\infty}} \frac{\mu(d)kr_0}{\varphi(dkr_0)} \\ \times \sum_{\theta \,(\mathrm{mod}\, dkr_0)} \sum_{\substack{|q'| \leq Q^* \\ (q', kg_1g_2)=1}} \frac{|q'|}{\varphi(|q'|)} \sideset{}{^*}\sum_{\chi\,(\mathrm{mod}\,|q'|)} \frac{1}{2\pi i}\int_{(2)} \Big(\frac{V}{kr_0|q'|Q}\Big)^s \\ \times \widetilde{w}(2-s) Z_{1,1}(s)\nu_{q'}(s)\delta_{k\mathbf{g}}(s) \\ \times \sum_{\substack{(a_1b_1, a_2b_2)=1 \\ (\mathbf{a},\mathbf{g})=1 \\ (\mathrm{sgn})}} \delta_{a_1b_1}(s)\delta_{a_2b_2}(s)\alpha^{(1,\mathbf{g})}_{a_1,b_1}\overline{\alpha}^{(2,\mathbf{g})}_{a_2,b_2} \\ \times \chi\theta\Big(\frac{\gamma_1 a_1 b_2}{\gamma_2 a_2 b_1}\Big)\Big(\frac{\gamma_1 a_1 b_2}{\gamma_2 a_2 b_1}\Big)^{it} dt \frac{du_1 du_2 du_3}{y_1^{iu_1} y_2^{iu_2} x_2^{iu_3}}\, ds, \end{aligned}$$

plus a small error term. Here $G = G_s$ depends on s, via $\omega_s(x) = x^{s-1}\omega(x)$. We also record

$$R = \frac{Vg_1g_2}{N} \quad \text{and} \quad U = \frac{N}{g_1g_2V}(QkTN)^{o(1)}. \tag{5.33}$$

Now we shift the contour to the line $\mathrm{Re}(s) = \varepsilon$. In doing so we cross a pole at $s = 1$, and we denote its residue by $S_\infty^{(0)}$. There is a small but convenient simplification with the sign condition (5.30), namely that all the summands are independent of $\mathrm{sgn}(\gamma_1 a_1 b_2 - \gamma_2 a_2 b_1)$ and $\mathrm{sgn}(q')$, except for the indicator function of the set where these signs agree. We may therefore take $q' > 0$. We also make a small modification by factoring $r_0 = r_g r_k$ where

$r_g \mid (g_1g_2)^\infty$ and $r_k \mid k^\infty$. With this simplification and others, we obtain

$$S_\infty^{(0)} = QkT \sum_{r_k|k^\infty} \sum_{\mathbf{g}} \sum_{1 \ll V \leq V_{\max}} \int_{\mathbb{R}^4} G_1(u_1,u_2,u_3,t) \times \sum_{\substack{d|g_1g_2 \\ r_g|(g_1g_2)^\infty}} \frac{\mu(d)}{\varphi(dkr_gr_k)} \sum_{\theta \,(\mathrm{mod}\, dkr_gr_k)} \sum_{\substack{q' \leq Q^* \\ (q',kg_1g_2)=1}} \frac{\widetilde{w}(1)Z_{1,1}\nu_{q'}\delta_{k\mathbf{g}}}{\varphi(q')} \times \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q')} \sum_{\substack{(a_1b_1,a_2b_2)=1 \\ (\mathbf{a},\mathbf{g})=1}} \delta_{a_1b_1}\delta_{a_2b_2}\alpha_{a_1,b_1}^{(1,\mathbf{g})}\overline{\alpha}_{a_2,b_2}^{(2,\mathbf{g})} \times \chi\theta\left(\frac{\gamma_1a_1b_2}{\gamma_2a_2b_1}\right)\left(\frac{\gamma_1a_1b_2}{\gamma_2a_2b_1}\right)^{it} dt \,\frac{du_1\,du_2\,du_3}{y_1^{iu_1}y_2^{iu_2}x_2^{iu_3}}. \tag{5.34}$$

Let S'_∞ denote the remaining contour integral along $\mathrm{Re}(s) = \varepsilon$. Here we obtain

$$|S'_\infty| \lesssim Q^2kT \int_{(\varepsilon)} |\widetilde{w}(2-s)| \sum_{\mathbf{g}} \sum_{1 \ll V \leq V_{\max}} V^{-1} \sum_{\substack{d|g_1g_2 \\ r_0|(kg_1g_2)^\infty}} d^{-1} \times \int_{\mathbb{R}^4} |G_s(u_1,u_2,u_3,t)| du_1\,du_2\,du_3 \sum_{\theta \,(\mathrm{mod}\, dkr_0)} \sum_{\substack{q' \leq Q^* \\ (q',kg_1g_2)=1}} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q')} \left| \sum_{\substack{(a_1b_1,a_2b_2)=1 \\ (\mathbf{a},\mathbf{g})=1}} \delta_{a_1b_1}(s)\alpha_{a_1,b_1}^{(1,\mathbf{g})}\chi\theta(a_1\overline{b_1})\left(\frac{a_1}{b_1}\right)^{it} \delta_{a_2b_2}(s)\overline{\alpha}_{a_2,b_2}^{(2,\mathbf{g})}\chi\theta(b_2\overline{a_2})\left(\frac{b_2}{a_2}\right)^{it} \right| dt\,|ds|.$$

A small issue here concerns the dependence of G_s on s. By the rapid decay of $|\widetilde{w}(2-s)|$, we may truncate the s-integral at $|s| \lesssim 1$. The remark following Lemma 5.2 shows that the family of functions G_s have a good uniform bound. We may then truncate the t-integral at $U(QkTN)^{o(1)}$. Lemma 5.4 allows us to essentially remove the coprimality condition $(a_1b_1, a_2b_2) = 1$; we apply this lemma with $M \ll \frac{N}{g_1g_2}$ and $\gamma_{a,b}^{(i)} = \delta_{ab}(s)\alpha_{a,b}^{(i,\mathbf{g})}$. With these steps, we may then estimate S'_∞ in terms of the original norm (1.3), which gives

$$|S'_\infty| \lesssim Q^2kT \sum_{\mathbf{g}} \sum_{1 \ll V \leq V_{\max}} V^{-1} \sum_{\substack{d|g_1g_2 \\ r_0|(kg_1g_2)^\infty}} d^{-1} \times U^{-1}\overline{\Delta}\left(Q^*, dkr_0, U, \frac{N}{g_1g_2}\right)|\alpha_{a_1,b_1}^{(1,\mathbf{g})}\alpha_{a_2,b_2}^{(2,\mathbf{g})}|, \tag{5.35}$$

where U is given by (5.33) and Q^* was defined by (5.32). Note $UV = \frac{N}{g_1g_2}(QkTN)^{o(1)}$. It is convenient to write $V = V_{\max}/P$, where $1 \ll P \ll V_{\max}$, in which case (5.35) simplifies as

$$|S'_\infty| \lesssim \frac{Q^2kT}{N} \sum_{\mathbf{g}} \sum_{1 \ll P \ll V_{\max}} \sum_{\substack{d|g_1g_2 \\ r_0|(kg_1g_2)^\infty}} \frac{1}{d} \times g_1g_2 \overline{\Delta}\Big(\frac{N}{QkTg_1g_2r_gr_kP}, dkr_gr_k, PT, \frac{N}{g_1g_2}\Big) |\alpha^{(1,\mathbf{g})}_{a_1,b_1}\alpha^{(2,\mathbf{g})}_{a_2,b_2}|. \tag{5.36}$$

Recalling the definition (1.8) completes the proof of Proposition 5.5. ∎

5.5. Conclusion. Now we use Propositions 5.3 and 5.5 to prove Theorem 1.6. Recall that we need to show that $S_{>Y}$ satisfies (5.9), that is,

$$S_{>Y} \lesssim |\alpha|^2 \Big(Q^2kT + \frac{Q^2kT}{N}\overline{\Delta'}\Big(\frac{N}{kQT}, k, T, N\Big)\Big),$$

where for convenience of the reader we recall the definition (1.8):

$$\Delta'(Q,k,T,N) = \max_{\substack{X,R,U,C \in \mathbb{R}_{\geq 1}, \ell \in \mathbb{Z}_{>0} \\ XR^2\ell U \leq Q^2kT \\ X \leq C}} X\Delta\Big(R, \ell, U, \frac{N}{C}\Big).$$

We have a decomposition

$$S_{>Y} = S_\infty - S_{\leq Y} = S_\infty^{\mathrm{diag}} + S'_\infty - S'_{\leq Y} + (S_\infty^{(0)} - S_{\leq Y}^{(0)}) + \mathcal{E}_\infty.$$

The diagonal term S_∞^{diag} is acceptable for Theorem 1.6, as is $\mathcal{E}_\infty$.

Now we turn to the terms $S'_{\leq *}$. Recall the definitions (5.32) and (5.31). We choose

$$Y = (QkTN)^\varepsilon \frac{N}{QkT}, \tag{5.37}$$

with a value of ε such that when $V = V_{\max}$, then $Q^* = \frac{Y}{g_1g_2r_gr_k}$. Using the assumption $Q^2kT \gg N^{1-\varepsilon}$, it is easy to check that (5.19) is acceptable for Theorem 1.6, and also that $Y \leq Q/100$, so this is a valid choice of Y. Moreover, (5.25) directly shows that S'_∞ is bounded in accord with the theorem.

Finally, consider the polar terms from $s = 1$, namely $S_\infty^{(0)}$ and $S_{\leq Y}^{(0)}$ given by (5.34) and (5.22). We simplify $S_\infty^{(0)}$, continuing with (5.34). We reverse the order of summation between V and q'; the condition $q' \leq Q^* = \frac{CV}{Qkr_gr_k}$ (where C is shorthand for $(QkTN)^\varepsilon$) becomes instead $V > C^{-1}q'Qkr_kr_g$ (on the inside) and $q' \leq \frac{Y}{r_gr_kg_1g_2}$ (on the outside). We then write $S_\infty^{(0)} = S_{\infty,1}^{(0)} - S_{\infty,2}^{(0)}$, where $S_{\infty,1}^{(0)}$ has V unconstrained, and $S_{\infty,2}^{(0)}$ has $V \leq C^{-1}q'Qkr_kr_g$. A pleasant feature of $S_{\infty,1}^{(0)}$ is that the sum over V re-assembles the partition of unity, since G_1 corresponds to $\omega_s(x)|_{s=1} = \omega(x)$. We also re-open the definition of $\widehat{w}$. Together, these steps give

$$S_{\infty,1}^{(0)} = Qk \sum_{r_k|k^\infty} \sum_{\mathbf{g}} \int_{-\infty}^{\infty} w\left(\frac{t}{T}\right) \sum_{\substack{d|g_1g_2 \\ r_g|(g_1g_2)^\infty}} \frac{\mu(d)}{\varphi(dkr_g r_k)} \times \sum_{\theta \,(\mathrm{mod}\, dkr_g r_k)} \sum_{\substack{q' \le \frac{Y}{r_k r_g g_1 g_2} \\ (q', kg_1g_2)=1}} \frac{1}{\varphi(q')} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q')} \widetilde{w}(1) Z_{1,1} \nu_{q'} \delta_{k\mathbf{g}} \times \sum_{\substack{(a_1b_1, a_2b_2)=1 \\ (\mathbf{a},\mathbf{g})=1}} \delta_{a_1b_1}\delta_{a_2b_2} \alpha_{a_1,b_1}^{(1,\mathbf{g})} \overline{\alpha}_{a_2,b_2}^{(2,\mathbf{g})} \chi\theta\left(\frac{\gamma_1 a_1 b_2}{\gamma_2 a_2 b_1}\right) \left(\frac{\gamma_1 a_1 b_2}{\gamma_2 a_2 b_1}\right)^{it} dt. \tag{5.38}$$

Next we further cut this sum into four pieces, via

$$\sum_{q' \le \frac{Y}{r_k r_g g_1 g_2}} = \sum_{q' \le \frac{Y}{r_k}} - \sum_{\frac{Y}{r_k g_1 g_2} < q' \le \frac{Y}{r_k}} - \sum_{\frac{Y}{r_k r_g d g_1 g_2} < q' \le \frac{Y}{r_k g_1 g_2}} + \sum_{\frac{Y}{r_k r_g d g_1 g_2} \le q' \le \frac{Y}{r_k r_g g_1 g_2}} . \tag{5.39}$$

Call the corresponding sums S_i for $i = 1, 2, 3, 4$. There is a pleasant simplification available for S_1, S_2, and S_3. In these three sums, both the summation conditions in (5.39), as well as all the summands in (5.38), depend only on the *product* $dr_g = D$ (say), with the exception of the presence of $\mu(d)$. Möbius inversion means that the sum over $d \,|\, D$ detects $D = 1$. This immediately implies $S_3 = 0$. Moreover, we see that $S_1 = S_{\le Y}^{(0)}$, which is a crucial cancellation. The sum S_2 becomes

$$S_2 = -Qk \sum_{r_k|k^\infty} \sum_{\mathbf{g}} \int_{-\infty}^{\infty} w\left(\frac{t}{T}\right) \frac{1}{\varphi(kr_k)} \times \sum_{\theta \,(\mathrm{mod}\, kr_k)} \sum_{\substack{\frac{Y}{r_k g_1 g_2} < q' \le \frac{Y}{r_k} \\ (q', kg_1g_2)=1}} \frac{1}{\varphi(q')} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q')} \widetilde{w}(1) Z_{1,1} \nu_{q'} \delta_{k\mathbf{g}} \times \sum_{\substack{(a_1b_1, a_2b_2)=1 \\ (\mathbf{a},\mathbf{g})=1}} \delta_{a_1b_1}\delta_{a_2b_2} \alpha_{a_1,b_1}^{(1,\mathbf{g})} \overline{\alpha}_{a_2,b_2}^{(2,\mathbf{g})} \chi\theta\left(\frac{\gamma_1 a_1 b_2}{\gamma_2 a_2 b_1}\right) \left(\frac{\gamma_1 a_1 b_2}{\gamma_2 a_2 b_1}\right)^{it} dt.$$

Similarly to the estimation of S'_∞, using Lemma 5.4 we obtain

$$|S_2| \lesssim |\alpha|^2 Qk \max_{\mathbf{g}, r_k|k^\infty} \frac{1}{kr_k} \max_{\frac{Y}{g_1 g_2 r_k} \le Q' \le \frac{Y}{r_k}} \frac{1}{Q'} \overline{\Delta}\left(Q', kr_k, T, \frac{N}{g_1 g_2}\right).$$

Write $Q' = \frac{Y}{r_k P}$, where $1 \le P \le g_1 g_2$, which gives

$$S_2 \lesssim |\alpha|^2 \frac{Q^2 kT}{N} \max_{\substack{\mathbf{g}, r_k|k^\infty \\ 1 \le P \le g_1 g_2}} P\overline{\Delta}\left(\frac{N}{QkTr_k P}, kr_k, T, \frac{N}{g_1 g_2}\right).$$

This is consistent with Theorem 1.6. The sum S_4 is similar in shape, and we obtain

$$S_4 \lesssim |\alpha|^2 \frac{Q^2kT}{N} \max_{\substack{\mathbf{g}, r_k | k^\infty \\ r_g | (g_1g_2)^\infty}} \max_{1 \leq P \leq d} \frac{Pg_1g_2}{d} \overline{\Delta}\bigg(\frac{N}{QkTr_kr_gg_1g_2P}, dkr_kr_g, T, \frac{N}{g_1g_2}\bigg),$$

which is acceptable.

Next we turn to estimating $S^{(0)}_{\infty,2}$. Our expression for this is identical to (5.38), except we have an additional weight function of the form

$$\Omega(x) := \sum_{1 \ll V \lesssim q'Qkr_kr_g} \omega\Big(\frac{x}{V}\Big), \qquad \text{with } x = \gamma_1a_1b_2 - \gamma_2a_2b_1. \tag{5.40}$$

The function $\Omega(x)$ is identically 1 for $1 \leq |x| \lesssim q'Qkr_kr_g$, but it vanishes at $x = 0$. Let $\Omega_0(x) = 1 - \Omega(x)$ for $|x| \leq 1$, and $\Omega_0(x) = 0$ for $|x| \geq 1$. Let $S'_{\infty,2}$ denote the same expression as $S^{(0)}_{\infty,2}$ but with Ω replaced by $\Omega_1 := \Omega + \Omega_0$, and let $S^{\mathrm{diag}}_{\infty,2} = S'_{\infty,2} - S^{(0)}_{\infty,2}$. Indeed, $S^{\mathrm{diag}}_{\infty,2}$ is supported only on $\gamma_1a_1b_2 = \gamma_2a_2b_1$. By similar reasoning to that in (5.27), we obtain

$$|S^{\mathrm{diag}}_{\infty,2}| \lesssim QkTY|\alpha|^2 \lesssim N|\alpha|^2.$$

Since $N \lesssim Q^2kT$, this is no worse than (5.27).

Finally, consider $S'_{\infty,2}$. The function Ω_1 meets the conditions of Lemma 5.2, with V taking the value $C^{-1}q'Qkr_kr_g$. Hence we obtain an expression of the form

$$S'_{\infty,2} = QkT \sum_{r_k|k^\infty} \sum_{\mathbf{g}} \int_{\mathbb{R}^4} \sum_{\substack{d | g_1g_2 \\ r_g | (g_1g_2)^\infty}} \frac{\mu(d)}{\varphi(dkr_gr_k)}$$

$$\times \sum_{\theta \,(\mathrm{mod}\, dkr_gr_k)} \sum_{\substack{q' \leq \frac{Y}{r_kr_gg_1g_2} \\ (q', kg_1g_2)=1}} \frac{1}{\varphi(q')} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q')} \widetilde{w}(1) Z_{1,1} \nu_{q'} \delta_{k\mathbf{g}} \sum_{\substack{(a_1b_1, a_2b_2)=1 \\ (\mathbf{a},\mathbf{g})=1}} G(t, u_1, u_2, u_3)$$

$$\times \delta_{a_1b_1} \delta_{a_2b_2} \alpha^{(1,\mathbf{g})}_{a_1,b_1} \overline{\alpha}^{(2,\mathbf{g})}_{a_2,b_2} \chi\theta\Big(\frac{\gamma_1a_1b_2}{\gamma_2a_2b_1}\Big)\Big(\frac{\gamma_1a_1b_2}{\gamma_2a_2b_1}\Big)^{it} dt\, \frac{du_1\, du_2\, du_3}{y_1^{u_1} y_2^{u_2} x_2^{u_3}}.$$

The bound on G is given by (5.3), with now

$$U = \frac{N}{q'Qkr_kr_gg_1g_2}(QkTN)^{o(1)}.$$

The estimations are similar to those of S'_∞, S_2, and S_4, and we obtain

$$|S'_{\infty,2}| \lesssim |\alpha|^2 QkT \max_{\substack{\mathbf{g}, r_k | k^\infty \\ r_g | (g_1g_2)^\infty}} \frac{1}{dkr_gr_k} \max_{Q' \leq \frac{Y}{r_kr_gg_1g_2}} \frac{1}{UQ'} \overline{\Delta}\Big(Q', dkr_kr_g, U, \frac{N}{g_1g_2}\Big).$$

This simplifies as

$$|S'_{\infty,2}| \lesssim |\alpha|^2 \frac{Q^2kT}{N} \times \max_{\mathbf{g},r_k|k^\infty} \max_{1\ll P\lesssim \frac{N}{QkTr_kr_gg_1g_2}} \frac{g_1g_2}{d} \overline{\Delta}\left(\frac{N}{QkTr_kr_gg_1g_2P}, dkr_kr_g, TP, \frac{N}{g_1g_2}\right).$$

One checks this is consistent with Theorem 1.6, which completes its proof.

Acknowledgments. I thank Henryk Iwaniec and Emmanuel Kowalski for valuable comments. I am also grateful to the referee for a careful reading which uncovered several inaccuracies.

This material is based upon work supported by the National Science Foundation under agreement No. DMS-2001306. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

[BI1] E. Bombieri and H. Iwaniec, *On the order of* $\zeta(\frac{1}{2}+it)$, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 13 (1986), 449–472.

[BI2] E. Bombieri and H. Iwaniec, *Some mean-value theorems for exponential sums*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 13 (1986), 473–486.

[CIS] J. B. Conrey, H. Iwaniec, and K. Soundararajan, *Asymptotic large sieve*, arXiv: 1105.1176 (2011).

[DI] J.-M. Deshouillers and H. Iwaniec, *Kloosterman sums and Fourier coefficients of cusp forms*, Invent. Math. 70 (1982), 219–288.

[DFI] W. Duke, J. B. Friedlander, and H. Iwaniec, *Bilinear forms with Kloosterman fractions*, Invent. Math. 128 (1997), 23–43.

[EEHK] J. S. Ellenberg, C. Elsholtz, C. Hall, and E. Kowalski, *Non-simple abelian varieties in a family: geometric and analytic approaches*, J. London Math. Soc. (2) 80 (2009), 135–154.

[FV] M. Forti and C. Viola, *On large sieve type estimates for the Dirichlet series operator*, in: Analytic Number Theory (St. Louis, MO, 1972), Proc. Sympos. Pure Math. 24, Amer. Math. Soc., Providence, RI, 1973, 31–49.

[H-B] D. R. Heath-Brown, *A mean value estimate for real character sums*, Acta Arith. 72 (1995), 235–275.

[H] C. Hooley, *On the Barban–Davenport–Halberstam theorem. I*, J. Reine Angew. Math. 274/275 (1975), 206–223.

[IK] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, Amer. Math. Soc. Colloq. Publ. 53, Amer. Math. Soc., Providence, RI, 2004.

[IL] H. Iwaniec and X. Li, *The orthogonality of Hecke eigenvalues*, Compos. Math. 143 (2007), 541–565.

[K] E. Kowalski, *The Large Sieve and Its Applications, Arithmetic Geometry, Random Walks and Discrete Groups*, Cambridge Tracts in Math. 175, Cambridge Univ. Press, Cambridge, 2008.

[M] H. L. Montgomery, *Topics in Multiplicative Number Theory*, Lecture Notes in Math. 227, Springer, Berlin, 1971.

[Z] D. Zywina, *Hilbert's irreducibility theorem and the larger sieve*, arXiv:1011.6465 (2010).

Matthew P. Young
Department of Mathematics
Texas A&M University
College Station, TX 77843-3368, USA
E-mail: myoung@math.tamu.edu