COMPOSITIO MATHEMATICA

Smooth numbers in arithmetic progressions to large moduli

Alexandru Pascadi

Compositio Math. **161** (2025), 1923–1974.

doi: 10.1112/S0010437X2500747X

FOUNDATION
COMPOSITIO
MATHEMATICA



LONDON
MATHEMATICAL
SOCIETY
EST. 1865



https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Compositio Math. 161 (2025) 1923–1974
doi:10.1112/S0010437X2500747X

Smooth numbers in arithmetic progressions to large moduli

Alexandru Pascadi

Abstract

We show that smooth numbers are equidistributed in arithmetic progressions to moduli of size $x^{66/107-o(1)}$. This overcomes a longstanding barrier of $x^{3/5-o(1)}$ present in previous works of Bombieri, Friedlander and Iwaniec, Fouvry and Tenenbaum, Drappeau, and Maynard. We build on Drappeau's variation of Linnik's dispersion method and on exponential sum manipulations of Maynard, ultimately relying on optimized Deshouillers–Iwaniec-type estimates for sums of Kloosterman sums.

Contents

1 Introduction **1924**
2 Overview of key ideas **1926**
2.1 First steps and limitations of previous approaches 1927
2.2 Improved exponential sum manipulations for specific ranges 1928
2.3 Completing the argument 1929
3 Notation and preliminaries **1930**
3.1 Sets, sums, estimates, and congruences 1930
3.2 Multiplicative number theory 1931
3.3 Fourier analysis 1932
3.4 Bounds for exponential sums 1933
4 The triple convolution estimate **1934**
5 Dispersion and deamplification **1937**
6 The main terms **1939**
7 The second and third dispersion sums **1943**
8 The first dispersion sum **1945**
9 Bombieri–Friedlander–Iwaniec-style estimates **1956**
10 Deshouillers–Iwaniec-style estimates **1968**
Acknowledgements **1971**
References **1972**

Received 2 September 2023, accepted in final form 26 February 2025, published online 12 September 2025.
2020 Mathematics Subject Classification 11N25 (primary).
Keywords: smooth numbers; exponent of distribution; dispersion method; Kloosterman sums.

© The Author(s), 2025. Published by Cambridge University Press on behalf of *Compositio Mathematica*. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited. *Compositio Mathematica* is © Foundation Compositio Mathematica.

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

1. Introduction

The Bombieri–Vinogradov theorem [Bom65, Vin65] famously states that for any $x > 2$, $A > 0$, and B sufficiently large in terms of A, one has

$$\sum_{q \leqslant x^{1/2}/(\log x)^B} \max_{(a,q)=1} \left| \pi(x;q,a) - \frac{\pi(x)}{\varphi(q)} \right| \ll_A \frac{x}{(\log x)^A}, \tag{1.1}$$

where $\pi(x)$ denotes the number of primes up to x, and $\pi(x;q,a)$ denotes the number of these primes which are congruent to a modulo q. Informally, (1.1) states that the primes are well distributed in arithmetic progressions when averaging over moduli almost as large as $x^{1/2}$. Without the sum over q, the Siegel–Walfisz theorem only controls the summand uniformly in the (much smaller) range $q \leqslant (\log x)^A$, and the Generalized Riemann Hypothesis would improve this to $q \leqslant x^{1/2}/(\log x)^B$. Thus (1.1) provides an unconditional substitute for GRH when some averaging over q is available; this is very often the case in sieve theory, where results like (1.1) have led to multiple major breakthroughs (including, e.g., the existence of infinitely many bounded gaps between primes [Zha14, Pol14a, May15, Pol14b]).

We say that the primes have *exponent of distribution* $\alpha < 1$ iff the analogue of (1.1) holds true when summing over all moduli $q \leqslant x^\alpha$. The Elliott–Halberstam conjecture [EH68] asserts that $\alpha = 1 - \varepsilon$ works for any $\varepsilon > 0$ (the implied constant depending on ε), but it remains open whether (1.1) holds for any $\alpha > 1/2$. Quite remarkably, it *is* possible to go beyond this square-root barrier if one slightly weakens the left-hand side of (1.1), by fixing the residue a, assuming various factorization properties of the moduli q, and/or replacing the absolute values with suitable weights. On this front, we mention the pioneering work of Fouvry [Fou84, Fou87, Fou85, Fou82] and Fouvry and Iwaniec [FI80, FI83], a series of three papers by Bombieri, Friedlander, and Iwaniec [BFI86, BFI87, BFI89], the main estimate in Zhang's work on bounded gaps [Zha14], and three recent papers of Maynard [May25a, May25b, May25c]; in particular, in [May25b], Maynard achieved exponents of distribution as large as $3/5 - \varepsilon$ assuming well-factorable weights.

In this paper, we are concerned with the case of *y-smooth* (or *y-friable*) numbers rather than primes; the objects of study here are the sets

$$S(x,y) := \{n \in \mathbf{Z}_+ : n \leqslant x, \text{ and all prime factors of } n \text{ are } \leqslant y\},$$

defined by two parameters $x, y \geqslant 2$, where y will grow like $x^{o(1)}$. To state our main result, we denote

$$\begin{aligned} \Psi(x,y) &:= \#S(x,y), \\ \Psi_q(x,y) &:= \#\{n \in S(x,y) : (n,q) = 1\}, \\ \Psi(x,y;a,q) &:= \#\{n \in S(x,y) : n \equiv a \pmod q\}. \end{aligned}$$

Theorem 1.1 (Smooth numbers in APs with moduli beyond $x^{3/5}$). *Let $a \in \mathbf{Z} \setminus \{0\}$ and $A, \varepsilon > 0$. Then there exists $C = C(A,\varepsilon) > 0$ such that, in the range $x > 2$, $(\log x)^C \leqslant y \leqslant x^{1/C}$, one has*

$$\sum_{\substack{q \leqslant x^{66/107-\varepsilon} \\ (q,a)=1}} \left| \Psi(x,y;a,q) - \frac{\Psi_q(x,y)}{\varphi(q)} \right| \ll_{a,A,\varepsilon} \frac{\Psi(x,y)}{(\log x)^A}.$$

A similar result with an exponent of $1/2 - \varepsilon$ and uniformity in a (as in (1.1)) is due to Granville [Gra93a, Theorem 2]; see also [Wol73a, Wol73b, FT91, Gra93b]. Virtually all results of this type that go beyond the $x^{1/2}$-barrier rely on equidistribution estimates for convolutions of sequences, but unless a long smooth sequence is involved, the exponents have been limited to $3/5$ or less. Bombieri, Friedlander, and Iwaniec proved a triple convolution estimate handling

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

moduli up to $x^{3/5}$ for sequences of convenient lengths [BFI86, Theorem 4], and Fouvry and Tenenbaum used this result (along with the flexible factorization properties of smooth numbers) to prove an analogue of Theorem 1.1 for $q \leqslant x^{3/5-\varepsilon}$, and with a right-hand side of $x/(\log x)^A$ [FT96, Théorème 2]. Motivated by an application to the Titchmarsh divisor problem for smooth numbers, Drappeau improved the bound to $\Psi(x,y)/(\log x)^A$ in the same range $q \leqslant x^{3/5-\varepsilon}$ [Dra15, Théorème 1]. Unfortunately, the BFI estimates and subsequent arguments seem limited to this range of moduli.

Maynard recently introduced a different arrangement of exponential sums [May25a, Chapter 18], which would, in principle, allow for a triple convolution estimate with moduli up to $x^{5/8}$, if the Selberg eigenvalue conjecture for Maass forms [Sel65, Sar95, Iwa85, IS85, Iwa90, LRS95] held true; but his unconditional estimates were still limited below $x^{3/5}$ [May25a, Proposition 8.3]. We introduce a further variation of Maynard's argument which eliminates certain coefficient dependencies, allowing one to use more efficient estimates for sums of Kloosterman sums in some ranges. More precisely, we rely on an optimized estimate of Deshouillers–Iwaniec type (see Theorem 3.10), which averages over exceptional Maass forms (and their levels) more carefully, ultimately allowing us to go beyond $3/5 = 0.6$ unconditionally. Our exponent of $66/107 \approx 0.617$ uses the best progress towards Selberg's conjecture, due to Kim and Sarnak [Kim03, Appendix 2] (based on the automorphy of symmetric fourth-power L-functions).

Notation 1.2 (Exceptional eigenvalues). For $q \in \mathbf{Z}_+$, define $\theta_q := \sup_\lambda \sqrt{\max(0, 1-4\lambda)}$, where λ runs over all eigenvalues of the hyperbolic Laplacian for the Hecke congruence subgroup $\Gamma_0(q)$ (such λ is called exceptional iff $\lambda < 1/4$). Also, let $\theta_{\max} := \sup_{q \geqslant 1} \theta_q$.

Conjecture 1.3 (Selberg [Sel65]). One has $\theta_{\max} = 0$, i.e., there are no exceptional eigenvalues.

Theorem A (Kim–Sarnak [Kim03]). *One has* $\theta_{\max} \leqslant 7/32$.

Remark 1.4. We warn the reader of another common normalization for the θ-parameters, which differs by a factor of 2 (resulting in a bound of $7/64$ in Theorem A); our normalization follows [DI82, May25a]. We give more details on the role of exceptional eigenvalues in our work in § 10.

We now state a more general version of our main result from Theorem 1.1, which makes the dependency on $\theta_{\max}$ explicit, gives a refined bound on the right-hand side (following [Dra15]), and allows for some small uniformity in the residue parameter a.

Theorem 1.5 (Conditional exponent of distribution). *For any $\varepsilon > 0$, there exist $C, \delta > 0$ such that the following holds. Let $x > 2$, $(\log x)^C \leqslant y \leqslant x^{1/C}$, and denote $u := (\log x)/(\log y)$, $H(u) := \exp(u \log^{-2}(u+1))$. Then with an exponent of*

$$\alpha := \frac{5 - 4\theta_{\max}}{8 - 6\theta_{\max}} - \varepsilon, \tag{1.2}$$

one has

$$\sum_{\substack{q \leqslant x^\alpha \\ (q, a_1 a_2) = 1}} \left| \Psi(x, y; a_1 \overline{a_2}, q) - \frac{\Psi_q(x,y)}{\varphi(q)} \right| \ll_{\varepsilon, A} \Psi(x,y)(H(u)^{-\delta}(\log x)^{-A} + y^{-\delta}), \tag{1.3}$$

for all $a_1, a_2 \in \mathbf{Z}$ with $1 \leqslant |a_1|, |a_2| \leqslant x^\delta$, and all $A \geqslant 0$. The implicit constant is effective if $A < 1$.

Remark 1.6. In (1.3), $\overline{a_2}$ denotes a multiplicative inverse of a_2 modulo q; so the residue $a_1\overline{a_2}$ corresponds to congruences of the form $a_2 n \equiv a_1 \pmod q$. The right-hand side of (1.3) is the same as in Drappeau's result [Dra15, Théorème 1], and ultimately comes from a result of Harper (see Lemma 3.2).

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

In particular, Conjecture 1.3 would imply an exponent of distribution of $5/8-o(1)$, while Theorem A leads to the unconditional exponent of $66/107-o(1)$ from Theorem 1.1. As in previous approaches, our main technical result leading to Theorem 1.5 is a triple convolution estimate, given in Theorem 4.2; this improves on [BFI86, Théorème 4], [Dra15, Théorème 3], [DGS17, Lemma 2.3], and [May25a, Proposition 8.3]. We expect all such approaches to face a significant barrier at the exponent $2/3=0.\overline{6}$ (see the remark after (4.4)), so we may view the exponents of $66/107\approx 0.617$ and $5/8=0.625$ as progress towards this limit.

In fact, Theorem 4.2 is already in a suitable form to improve the analogous results of Drappeau, Granville, and Shao about smooth-supported multiplicative functions [DGS17]. More precisely, using Theorem 4.2 instead of [DGS17, Lemma 2.3], one can improve the exponent of 3/5 in [DGS17, Theorem 1.2] to the same value as in (1.2). We state a particular case of this result below, borrowing the notation

$$\Delta(f,x;q,a):=\sum_{\substack{n\leqslant x\\ n\equiv a \pmod q}} f(n)-\frac{1}{\varphi(q)}\sum_{\substack{n\leqslant x\\ (n,q)=1}} f(n),$$

from [DGS17]; we also say that an arithmetic function f satisfies the *Siegel–Walfisz criterion* if and only if

$$\forall\ A>0,\ (a,q)=1,\ x\geqslant 2,\quad \Delta(f,x;q,a)\ll_A \frac{1}{(\log x)^A}\sum_{n\leqslant x}|f(n)|.$$

Theorem 1.7 (Smooth-supported multiplicative functions in APs). *For any $\varepsilon, A>0$, there exists $\delta>0$ such that the following holds. Let $x>2$, $x^\delta\geqslant y\geqslant \exp(\sqrt{\log x}\log\log x)$, and f be a 1-bounded completely multiplicative function supported on y-smooth integers, satisfying the Siegel–Walfisz criterion. Then for $\alpha:=(5-4\theta_{\max})/(8-6\theta_{\max})-\varepsilon$, and all $a_1,a_2\in\mathbf{Z}$ with $1\leqslant |a_1|,|a_2|\leqslant x^\delta$,*

$$\sum_{\substack{q\leqslant x^\alpha\\ (q,a_1a_2)=1}}|\Delta(f,x;q,a_1\overline{a_2})|\ll_{\varepsilon,A}\frac{\Psi(x,y)}{(\log x)^A}.$$

Remark 1.8. The improvement from $3/5-\varepsilon$ to the exponent in (1.2) follows through in most applications of Drappeau's result [Dra15, Théorème 1], such as [Dra15, Corollaire 1]. Following [dLBD20, §§ 2 and 4], our triple convolution estimate also implies a version of Theorem 1.5 restricted to smooth moduli, which can be used to deduce refined upper bounds for the number of smooth values assumed by a factorable quadratic polynomial. For instance, one should obtain

$$\#\{n\sim x : n(n+1)\text{ is }y\text{-smooth}\}\ll_\varepsilon x\varrho(u)^{1+66/107-\varepsilon},$$

for $(\log x)^{O_\varepsilon(1)}\leqslant y\leqslant x$, where $u=(\log x)/(\log y)$ and $\varrho(u)$ is the Dickman function (satisfying $\Psi(x,y)=x\varrho(u)e^{O_\varepsilon(u)}$ in this range [dLBD20, Hil86]).

2. Overview of key ideas

Let us give a very rough sketch of our argument, for a simpler case of Theorem 1.1. Consider the residue $a=1$, the smoothness parameter $y=x^{1/\sqrt{\log\log x}}$, and a sum over moduli just above the 3/5 threshold, say

$$r\leqslant R:=x^{3/5+\sigma},$$

for some small $\sigma\gg 1$ (we switch the variable q to r, following [Dra15, DGS17]). Using the factorization properties of smooth numbers, it suffices to prove a triple convolution estimate

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

roughly of the form

$$\sum_{r\sim R}\rho_r\sum_{m\sim M}\alpha_m\sum_{n\sim N}\beta_n\sum_{\ell\sim L}\gamma_\ell\left(1_{mn\ell\equiv 1 \pmod r}-\frac{\mathbb{1}_{(mn\ell,r)=1}}{\varphi(r)}\right)\ll_A\frac{x}{(\log x)^A},\tag{2.1}$$

where (ρ_r), (α_m), (β_n), and (γ_ℓ) are arbitrary 1-bounded complex sequences, but we are free to choose the parameters $M,N,L\gg 1$ subject to $MNL\asymp x$. We pick

$$M:=\frac{x^{1-\delta}}{R},\quad N:=\frac{x^{1-2\delta}}{R},\quad L:=\frac{R^2}{x^{1-3\delta}},\tag{2.2}$$

for some small $\delta=o(1)$; thus $M,N\approx x^{2/5-\sigma}$ and $L\approx x^{1/5+2\sigma}$. Note additionally that $NL=x^\delta R$.

2.1 First steps and limitations of previous approaches

Following previous works [BFI86, Dra15, May25a] based on Linnik's *dispersion method* [Lin63], we apply Cauchy and Schwarz in the r,m variables, expand the square, and Fourier complete the resulting sums in m to sums over h. Ignoring GCD constraints, the key resulting exponential sum is a smoothed variant of

$$\sum_{r\sim R}\sum_{k\sim NL}u_k\sum_{n\sim N}\beta_n\sum_{\substack{\ell\sim L\\ n\ell\equiv k \pmod r}}\gamma_\ell\sum_{h\sim R/M}\mathrm{e}\left(h\frac{\overline{k}}{r}\right),\tag{2.3}$$

where (u_k) is the convolution of the original sequences (β_n), (γ_ℓ). We then flip moduli in the exponential via Bézout's identity, and substitute $t:=(k-n\ell)/r$; this leads to the sum

$$\sum_{t\sim x^\delta}\sum_{k\sim NL}u_k\sum_{n\sim N}\beta_n\sum_{\substack{\ell\sim L\\ n\ell\equiv k \pmod t}}\gamma_\ell\sum_{h\sim R/M}\mathrm{e}\left(ht\frac{\overline{n\ell}}{k}\right).\tag{2.4}$$

Following Maynard [May25a], we apply Cauchy and Schwarz in the t,n,k variables (keeping the congruence modulo t inside), and expand the square to reach the sum

$$\sum_{t\sim x^\delta}\sum_{\substack{\ell_1,\ell_2\sim L\\ \ell_1\equiv\ell_2 \pmod t}}\gamma_{\ell_1}\overline{\gamma_{\ell_2}}\sum_{h_1,h_2\sim R/M}\sum_{k\sim NL}\sum_{\substack{n\sim N\\ n\equiv k\overline{\ell_1} \pmod t}}\mathrm{e}\left(t(h_1\ell_2-h_2\ell_1)\frac{\overline{n\ell_1\ell_2}}{k}\right),\tag{2.5}$$

which we ultimately need to bound by $\ll_A RNL^2(\log x)^{-A}$; note that the trivial bound is larger by about $(R/M)^2$, due to the sum over h_1,h_2 introduced by Poisson summation. We bound the contribution of the diagonal terms (with $h_1\ell_2=h_2\ell_1$) by

$$\sum_{t\sim x^\delta}\sum_{\ell_1\sim L}\sum_{h_2\sim R/M}x^{o(1)}\sum_{k\sim NL}\sum_{\substack{n\sim N\\ n\equiv k\overline{\ell_1} \pmod t}}1\ll x^{o(1)}RNL^2\frac{N}{M},\tag{2.6}$$

which is acceptable since $N/M=x^{-\delta}$. We then introduce Kloosterman sums $S(i,j;k)$ by completing the sum in n to a sum over j, and find acceptable contributions from the zeroth Fourier coefficient ($j=0$), as well as from the terms with $\ell_1=\ell_2$; here, it suffices to use the Ramanujan bound, respectively an estimate of Deshouillers and Iwaniec [DI82, Theorem 9]. It ultimately remains to bound a variant of

$$\sum_{t\sim x^\delta}\sum_{\substack{\ell_1,\ell_2\sim L\\ \ell_1\equiv\ell_2 \pmod t\\ \ell_1\neq\ell_2}}\sum_{\substack{h_1,h_2\sim R/M\\ h_1\ell_2\neq h_2\ell_1}}\left|\sum_{j\sim x^\delta L}\mathrm{e}\left(\frac{j\overline{\ell_1}}{t}\right)\sum_{k\sim NL}S((h_1\ell_2-h_2\ell_1)\overline{\ell_1\ell_2},j;k)\right|,\tag{2.7}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

by $\ll_A N^2L^4(\log x)^{-A}$. Inserting 1-bounded coefficients ξ_{h_1,h_2} (also depending on t, ℓ_1, ℓ_2), and letting $a_d := \sum_{h_1\ell_2 - h_2\ell_1 = d} \xi_{h_1,h_2}$, the sum over h_1, h_2 in (2.7) roughly reduces to

$$\sum_{d\sim RL/M} a_d \sum_{j\sim x^\delta L} \mathrm{e}\left(\frac{j\overline{\ell_1}}{t}\right) \sum_{k\sim NL} S(d\overline{\ell_1}\ell_2, j; k). \tag{2.8}$$

Such sums can be bounded using the spectral theory of automorphic forms, specifically through the aforementioned work of Deshouillers and Iwaniec [DI82] (based on the Kuznetsov trace formula and the Weil bound); the relevant *level* of the congruence group in Notation 1.2 is $Q = \ell_1\ell_2$. Indeed, Maynard [May25a] uses [DI82, Theorem 9] to bound (a smoothed variant of) the sum in (2.8) by

$$x^{\theta_Q/2+o(1)} NL^{5/2} \Bigg(\sum_{d\sim RL/M} |a_d|^2 \Bigg)^{1/2},$$

and consequently the sum in (2.7) by

$$x^{\theta_{\max}/2+o(1)} NL^{5/2} \frac{R^{3/2}L^{3/2}}{M^{3/2}} \approx x^{(\theta_{\max}+3)/2+10\sigma}. \tag{2.9}$$

Unfortunately, this falls short of the desired bound of $N^2L^4(\log x)^{-A} \approx x^{8/5+6\sigma}$, unless

$$4\sigma < \frac{1}{10} - \frac{\theta_{\max}}{2}. \tag{2.10}$$

This is (barely) impossible with the currently best-known bound of $\theta_{\max}/2 \leqslant 7/64 \approx 0.109$.

2.2 Improved exponential sum manipulations for specific ranges

Starting from the work of Deshouillers and Iwaniec [DI82], better bounds (in the θ-aspect) for sums like (2.8) have been available when one additionally averages over the level $\ell_1\ell_2$, and at least one of the sequences of coefficients is independent of the level. Indeed, Drappeau's triple convolution estimate [Dra15] and prior works rely on [DI82, Theorem 12], which gives such a result for incomplete Kloosterman sums.

Following Maynard's argument (which is in turn based on Bombieri–Friedlander–Iwaniec's work in [BFI87, § 10]), we prefer to complete our Kloosterman sums and bound the contribution from the zeroth Fourier coefficient by hand, and separate into terms with $\ell_1 = \ell_2$ and $\ell_1 \neq \ell_2$, all before invoking Deshouillers–Iwaniec-style bounds. We then aim to apply an optimized bound for sums of complete Kloosterman sums with averaging over the level (given in Theorem 3.10), which improves [DI82, Theorem 11] by making the dependency on the $\theta_{\max}$ parameter explicit. But for this strategy to work out, we would need:

(1) the range of (ℓ_1, ℓ_2) in (2.7) to be (discretely) dense inside $[L, 2L]^2$; and
(2) crucially, the coefficients $e(j\overline{\ell_1}/t)$ to not depend on ℓ_1, ℓ_2.

While (2) is obviously false in our case, it is only barely false for the specific ranges in (2.2), due to the smallness of the parameter $t \sim x^\delta$. In particular, losing a factor of at most $x^{O(\delta)} = x^{o(1)}$ in (2.7), we may fix t and the values of ℓ_1 and $\ell_2 \pmod t$, turning (2) into a true statement at the expense of (1). The number of pairs (ℓ_1, ℓ_2) now becomes $\asymp L^2/t^2$, which ends up costing us another acceptable factor of $x^{o(1)}$. Overall, it remains to bound a sum of the form

$$\sum_{\ell_1,\ell_2\sim L} \sum_{d\sim RL/M} a_{d,\ell_1,\ell_2} \sum_{j\sim x^\delta L} \mathrm{e}(j\omega) \sum_{k\sim NL} S(d\overline{\ell_1}\ell_2, j; k), \tag{2.11}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

for some fixed $\omega \in \mathbf{R}/\mathbf{Z}$ (independent of ℓ_1, ℓ_2). Using Theorem 3.10, we obtain a bound like (2.9) where the factor depending on $\theta_{\max}$ is

$$\left(\frac{NL}{L^2}\right)^{\theta_{\max}} \approx x^{(\frac{1}{5}-3\sigma)\theta_{\max}},$$

rather than $x^{\theta_{\max}/2}$. Thus instead of (2.10), we now reach the desired bound provided that

$$(4-3\theta_{\max})\sigma < \frac{1}{10} - \frac{\theta_{\max}}{5},$$

which is possible since $(1/5)\cdot(7/32) = 0.04375 < 0.1$. In fact, this handles all values

$$\sigma < \frac{1-2\theta_{\max}}{40-30\theta_{\max}} \quad \Longleftrightarrow \quad \frac{3}{5} + \sigma < \frac{5-4\theta_{\max}}{8-6\theta_{\max}},$$

reaching the exponent of distribution in (1.2). Plugging in Kim–Sarnak's bound of $\theta_{\max} \leqslant 7/32$ (Theorem A) yields the unconditional exponent of $66/107 \approx 0.617$ from Theorem 1.1.

Remark 2.1. It is likely that optimized Deshouillers–Iwaniec-style bounds like Theorem 10.3 could also improve Drappeau's argument [Dra15], leading to a triple convolution estimate with different ranges than in our Theorem 4.2. In terms of the final exponent of distribution of smooth numbers, all such methods currently seem limited below 66/107 unconditionally (and 5/8 conditionally).

2.3 Completing the argument

To increase the range of uniformity in y, we adapt Drappeau's version of the dispersion method [Dra15]: we aim for a triple convolution estimate with a power saving in Theorem 4.2, after separating the contribution of small-conductor Dirichlet characters

$$\frac{1}{\varphi(r)} \sum_{\substack{\chi \pmod r \\ \mathrm{cond}(\chi) \leqslant x^{\varepsilon}}} \chi(mn\ell)$$

from (2.1); this can be handled via Lemma 3.2. As a result, the simpler two dispersion sums $\mathcal{S}_2$, $\mathcal{S}_3$ and their main terms involve Dirichlet characters (see Propositions 5.1 and 6.1), which ultimately bring in the classical Gauss sum bound (Lemma 3.7) and the multiplicative large sieve (Lemma 3.4). The difficulties in working with a general residue $a_1\overline{a_2}$ for $a_1, a_2 \ll x^{\delta}$, and in obtaining power savings throughout the computations in § 2.1, are quite tedious but purely technical (following [Dra15]).

We also adapt a 'deamplification' argument of Maynard [May25a], which introduces an artificial sum over $e \sim E = x^{o(1)}$ into the dispersion sums (by averaging over the residue of $n\ell \pmod e$ before applying Cauchy and Schwarz); for instance, the sum in (2.3) becomes

$$\sum_{e\sim E} \sum_{r\sim R} \sum_{k\sim NL} u_k \sum_{n\sim N} \beta_n \sum_{\substack{\ell\sim L \\ n\ell\equiv k \pmod{re}}} \gamma_\ell \sum_{h\sim R/M} \mathrm{e}\left(h\frac{\overline{k}}{r}\right).$$

Keeping e inside the second application of Cauchy and Schwarz, this essentially reduces the contribution of the diagonal terms in (2.6) by a factor of E, allowing us to cover wider ranges of sequence lengths in Theorem 4.2 (including the case $M = N$). This is generally convenient, and critical when one has less control over the sizes of the sequence lengths (which is the case in applications to the primes, but not to smooth numbers).

Figure 1 gives a visual summary of our formal argument, outlining the logical dependencies between our main lemmas, propositions, and theorems.

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

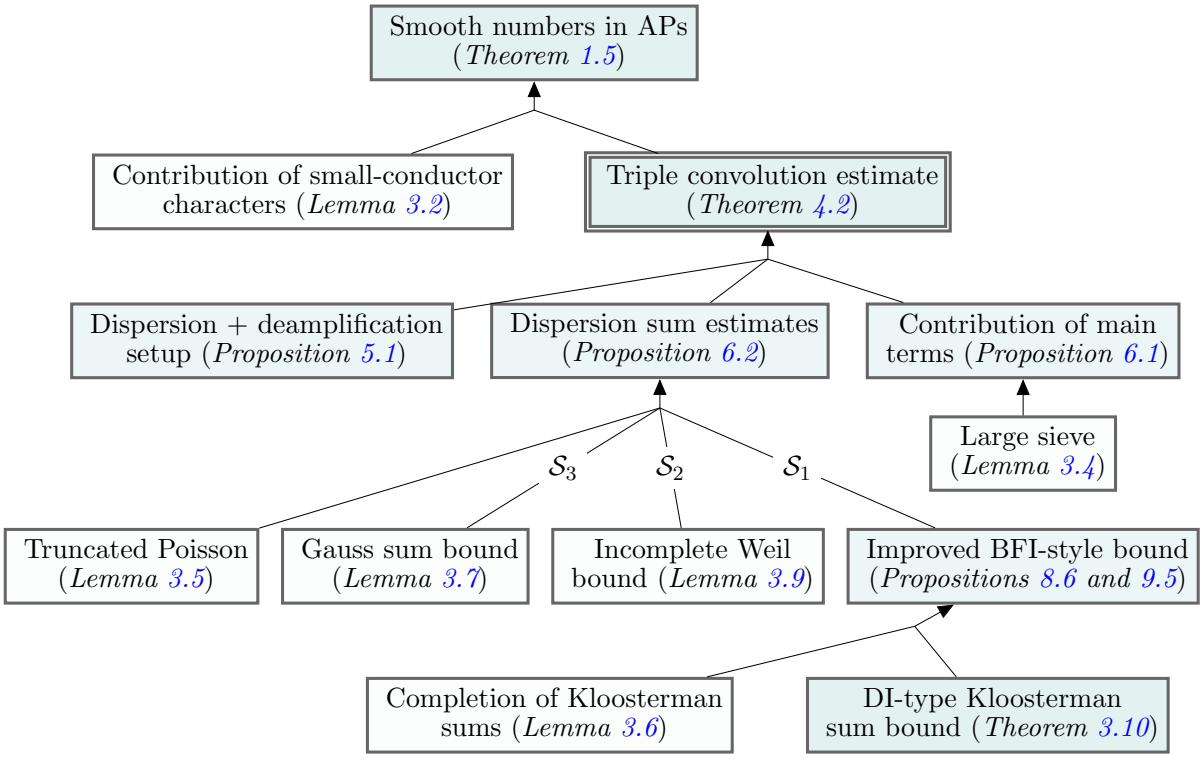


Figure 1 (color online). Structure of argument (*arrows show logical implications*).

3. Notation and preliminaries

3.1 Sets, sums, estimates, and congruences

We use the standard asymptotic notation in analytic number theory, with $f = O(g)$ (or $f \ll g$) meaning that there exists some constant $C > 0$ such that $|f| \leqslant Cg$ globally. We write $f \asymp g$ when $f \ll g$ and $g \ll f$, and indicate that the implied constants may depend on a parameter ε by placing it in the subscript (e.g., $f = O_\varepsilon(g)$, $f \ll_\varepsilon g$, and $f \asymp_\varepsilon g$). When $g \geqslant 0$, we also say that $f = o(g) = o_{x\to\infty}(g)$ if and only if $f(x)/g(x) \to 0$ as $x \to \infty$. Given $q \in [1, \infty]$, we write $\|f\|_q$ for the L^q norm of a measurable function $f : \mathbf{R} \to \mathbf{C}$ (using the Lebesgue measure), and $\|a_n\|_q$ for the ℓ^q norm of a complex sequence (a_n).

We denote by $\mathbf{Z}_+, \mathbf{Z}, \mathbf{R}, \mathbf{C}$, and $\mathbf{H}$ the sets of positive integers, integers, real numbers, complex numbers, and complex numbers with positive imaginary part, and set $\mathrm{e}(x) := \exp(2\pi i x)$ for $x \in \mathbf{R}$ (or $x \in \mathbf{R}/\mathbf{Z}$). We write $\mathbf{Z}/n\mathbf{Z}$ and $(\mathbf{Z}/n\mathbf{Z})^\times$ for the additive and multiplicative groups modulo a positive integer n, and denote the inverse of $c \in (\mathbf{Z}/n\mathbf{Z})^\times$ by $\overline{c}$. We may abuse notation slightly by identifying integers a, b, c with their residue classes modulo n where this is appropriate (e.g., in congruences $a \equiv b\overline{c} \pmod{\pm n}$, $x \equiv b\overline{c}/n \pmod 1$, or in exponentials $\mathrm{e}(b\overline{c}/n)$); the following simple lemma is an example of this.

Lemma 3.1 (Bézout's identity). *For any relatively prime integers a, b, one has*

$$\frac{1}{ab} \equiv \frac{\overline{a}}{b} + \frac{\overline{b}}{a} \pmod 1.$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Proof. Note that, here, $\overline{a}$ and $\overline{b}$ denote the inverses of a and b modulo b and a, respectively, so we have $a\overline{a} \equiv 1 \pmod{b}$ and $b\overline{b} \equiv 1 \pmod{a}$. The conclusion follows from the Chinese remainder theorem, once we multiply the congruence by ab and verify it modulo a and b separately. □

Given $N > 0$, we write $n \sim N$ for the statement that $N < n \leqslant 2N$, usually in the subscripts of sums. Given a statement S, we write $\mathbb{1}_S$ for its truth value (e.g., $\mathbb{1}_{2|n}$ equals 1 when n is even and 0 otherwise); we may use the same notation for the indicator function of a set S (i.e., $\mathbb{1}_S(x) = \mathbb{1}_{x \in S}$).

Given $a_1, \dots, a_k \in \mathbf{Z}$, we write $(a_1, \dots, a_k)$ (if not all a_i are 0) and $[a_1, \dots, a_k]$ (if none of the a_i are 0) for their greatest common divisor and lowest common multiple, among the positive integers. Given $a \in \mathbf{Z} \setminus \{0\}$, we write $\mathrm{rad}(a)$ for the largest square-free positive integer dividing a; for $b \in \mathbf{Z}$, we also write $a \mid b^\infty$ if and only if $\mathrm{rad}(a) \mid b$ (i.e., a divides a large enough power of b), and (a, b^∞) for the greatest divisor of a whose prime factors divide b. If $x > 0$ and $m \in \mathbf{Z} \setminus \{0\}$, sums like $\sum_{n \leqslant x}$, $\sum_{n \sim x}$, $\sum_{d \mid m}$, $\sum_{d \mid m^\infty}$, $\sum_{(a,m)=1}$, $\sum_{(a,m^\infty)=1}$ and $\sum_{ab=m}$ are understood to range over all positive integers n, d, a, b with the respective properties.

We also keep the notations specific to smooth numbers from the introduction, for $S(x, y)$, $\Psi(x, y) = \#S(x, y)$, $\Psi_q(x, y)$, $\Psi(x, y; a, q)$, and $H(u)$.

3.2 Multiplicative number theory

We denote by μ, τ, and φ the Möbius function, the divisor-counting function ($\tau(n) := \sum_{d \mid n} 1$), and the Euler totient function ($\varphi(n) := \sum_{1 \leqslant a \leqslant n} \mathbb{1}_{(a,n)=1}$). We may use various classical bounds involving these functions implicitly, including the divisor bound $\tau(n) \ll_\varepsilon n^\varepsilon$ (valid for all $\varepsilon > 0$), the lower bound $\varphi(n) \gg n/(\log\log n)$, and the upper bounds

$$\sum_{n \sim N} \frac{1}{\varphi(n)} \ll 1, \quad \sum_{n \leqslant x} \frac{\tau(n)}{\varphi(n)} \ll (\log x)^2.$$

(The latter follows from the former, using that $\varphi(ab) \gg \varphi(a)\varphi(b)$ for positive integers a, b.)

We write $\chi \pmod q$ to indicate that χ is a Dirichlet character with period q (of which there are $\varphi(q)$), and denote by $\mathrm{cond}(\chi)$ the conductor of χ (which divides q; this is the smallest positive integer d such that there exists a Dirichlet character $\chi' \pmod d$ with $\chi(n) = \chi'(n)\mathbb{1}_{(n,q)=1}$ for all $n \in \mathbf{Z}$); we say that $\chi \pmod q$ is *primitive* when $\mathrm{cond}(\chi) = q$. We will require a couple of results involving Dirichlet characters, the first being essentially due to Harper.

Lemma 3.2 (Contribution of small-conductor characters). *There exist constants $\varepsilon, \delta, C > 0$ such that, for $(\log x)^C \leqslant y \leqslant x$, $Q \leqslant x$ and $A > 0$, one has*

$$\sum_{q \leqslant Q} \frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod q \\ 1 < \mathrm{cond}(\chi) \leqslant x^\varepsilon}} \Big| \sum_{n \in S(x,y)} \chi(n) \Big| \ll_A \Psi(x, y)\big(H(u)^{-\delta}(\log x)^{-A} + y^{-\delta}\big),$$

with $u := (\log x)/(\log y)$, $H(u) := \exp(u \log^{-2}(u+1))$. The implicit constant is effective if $A < 1$.

Proof. This is the same as [Dra15, Lemme 5], and follows from the work of Harper in [Har12, § 3.3]. □

Remark 3.3. The condition $\mathrm{cond}(\chi) > 1$ in Lemma 3.2 leaves out the trivial character χ_0.

The second result is the classical multiplicative large sieve, as stated in [Dra15, Lemme 6].

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Lemma 3.4 (Multiplicative large sieve). *For $Q, M, N \geqslant 1$ and any sequence (a_n) of complex numbers, one has*

$$\sum_{q \leqslant Q} \frac{q}{\varphi(q)} \sum_{\substack{\chi \pmod q \\ \chi \text{ primitive}}} \left| \sum_{M < n \leqslant M+N} a_n \chi(n) \right|^2 \leqslant (N + Q^2 - 1) \sum_{M < n \leqslant M+N} |a_n|^2.$$

Proof. See, for example, [IK21, Theorem 7.13]. □

3.3 Fourier analysis

Given an integrable function $f : \mathbf{R} \to \mathbf{C}$, we write

$$\widehat{f}(\xi) := \int_{-\infty}^{\infty} f(t)\, \mathrm{e}(-\xi t)\, dt$$

for its Fourier transform. We will need the truncated version of Poisson summation stated below.

Lemma 3.5 (Truncated Poisson/Fourier completion). *Let $C > 0$, $x > 1$, $1 < M \ll x$, and $\Phi : \mathbf{R} \to \mathbf{R}$ be a smooth function supported in $[1/10, 10]$ such that $\|\Phi^{(j)}\|_\infty \ll_j (\log x)^{jC}$ for $j \geqslant 0$. Then for all positive integers $q \ll x$, any $a \in \mathbf{Z}/q\mathbf{Z}$, and any $\varepsilon > 0$, $H \geqslant x^\varepsilon q M^{-1}$, one has*

$$\sum_{m \equiv a \pmod q} \Phi\left(\frac{m}{M}\right) = \frac{M}{q} \sum_{|h| \leqslant H} \widehat{\Phi}\left(\frac{hM}{q}\right) \mathrm{e}\left(\frac{ah}{q}\right) + O_{\varepsilon, C}(x^{-100}).$$

Proof. This is the same as [May25a, Lemma 13.4] (see also [Dra15, Lemme 2]), following directly from the Poisson summation formula. □

While Lemma 3.5 will introduce exponential sums into our estimates, we will need an additional corollary (and generalization) of it to obtain sums of complete *Kloosterman sums*, defined by

$$S(m, n; c) := \sum_{b \in (\mathbf{Z}/c\mathbf{Z})^\times} \mathrm{e}\left(\frac{mb + n\overline{b}}{c}\right), \quad c \in \mathbf{Z}_+,\ m, n \in \mathbf{Z} \text{ (or } \mathbf{Z}/c\mathbf{Z}).$$

The following is the same as [May25a, Lemma 13.5], and can be quickly deduced from Lemma 3.5.

Lemma 3.6 (Kloosterman completion). *Let C, x, M, Φ be as in Lemma 3.5. Then for all positive integers $c, q \ll x$ with $(c, q) = 1$, any $a \in \mathbf{Z}/q\mathbf{Z}$, $n \in \mathbf{Z}/c\mathbf{Z}$, and any $\varepsilon > 0$, $H \geqslant x^\varepsilon cqM^{-1}$, one has*

$$\sum_{\substack{m \equiv a \pmod q \\ (m,c)=1}} \Phi\left(\frac{m}{M}\right) \mathrm{e}\left(\frac{\overline{m}n}{c}\right) = \frac{M}{cq} \sum_{|h| \leqslant H} \widehat{\Phi}\left(\frac{hM}{cq}\right) \mathrm{e}\left(\frac{ah\overline{c}}{q}\right) S(h\overline{q}, n; c) + O_{\varepsilon, C}\left(x^{-99}\right).$$

Proof. Rewrite the left-hand side as

$$\sum_{b \in (\mathbf{Z}/c\mathbf{Z})^\times} \mathrm{e}\left(\frac{\overline{b}n}{c}\right) \sum_{\substack{m \equiv a \pmod q \\ m \equiv b \pmod c}} \Phi\left(\frac{m}{M}\right),$$

and apply Lemma 3.5 to expand the inner summation, for the unique residue class $r \in \mathbf{Z}/cq\mathbf{Z}$ which is congruent to $a \pmod q$ and to $b \pmod c$ (invoking the Chinese remainder theorem). Noting that

$$\mathrm{e}\left(\frac{rh}{cq}\right) = \mathrm{e}\left(\frac{rh\overline{c}}{q} + \frac{rh\overline{q}}{c}\right) = \mathrm{e}\left(\frac{ah\overline{c}}{q} + \frac{bh\overline{q}}{c}\right)$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

by Lemma 3.1, swapping sums and taking out the factor depending on a, the conclusion follows. □

3.4 Bounds for exponential sums

The simpler two of the three dispersion sums arising in our computations (see (5.6)) will be estimated using the classical bounds for Gauss and Kloosterman sums.

LEMMA 3.7 (Gauss sum bound). *For any $a \in \mathbf{Z}$, $q \in \mathbf{Z}_+$, and Dirichlet character $\chi \pmod q$, one has*

$$\left| \sum_{b \in \mathbf{Z}/q\mathbf{Z}} \chi(b)\, \mathrm{e}\left(\frac{ab}{q}\right) \right| \leqslant \mathrm{cond}(\chi)^{1/2} \sum_{d \mid (a,q)} d.$$

Proof. This follows from [IK21, Lemmas 3.2 and 3.1], and is also used in [Dra15, § 3.2]. □

LEMMA 3.8 (Weil and Ramanujan bounds). *For $c \in \mathbf{Z}_+$ and $m, n \in \mathbf{Z}$ (or $\mathbf{Z}/c\mathbf{Z}$), one has*

$$S(m, n; c) \ll \tau(c)\, (m, n, c)^{1/2} c^{1/2}.$$

For $m = 0$, we have in fact $|S(0, n; c)| \leqslant (n, c)$.

Proof. The first (Weil) bound is [IK21, Corollary 11.12], while the second (Ramanujan) bound can be deduced by Möbius inversion. □

LEMMA 3.9 (Incomplete Weil bound). *Let $x > 1$, $1 < M \ll x$, and let $n, c, k, \ell \ll x$ be positive integers. Then for any $\varepsilon > 0$, one has*

$$\sum_{\substack{(m, ck) = 1 \\ \ell \mid m \\ m \leqslant M}} \frac{m}{\varphi(m)} \mathrm{e}\left(\frac{\overline{m} n}{c}\right) \ll_\varepsilon x^\varepsilon \left((n, c)^{1/2} c^{1/2} + \frac{(n, c)}{\ell c} M \right).$$

Proof. This follows immediately from [Dra15, (2.5)] and the divisor bound (see also [Fou82, Lemme 4] and [May25a, Lemma 16.1]). It can be proven by expanding $m\varphi(m)^{-1} = \sum_{v \mid m^\infty} v^{-1}$ and $\mathbb{1}_{(m,k)} = \sum_{d \mid k} \mu(d) \mathbb{1}_{d \mid m}$, changing variables $m \leftarrow m[\ell, \mathrm{rad}(v), d]$, completing sums via a result like Lemma 3.6, and finally applying Lemma 3.8 for the terms $h = 0$ and $h \neq 0$ separately. □

To estimate the first dispersion sum, will crucially need the following bound for sums of Kloosterman sums, which is an optimization of [DI82, Theorem 11] (see also [BFI87, Lemma 5]).

THEOREM 3.10 (The DI-type Kloosterman bound). *Let $1 \ll M, N, R, S, C \ll x^{O(1)}$, $(a_{m,r,s})$ be a complex sequence supported in $m \sim M$, $r \sim R$, $s \sim S$, and $\omega \in \mathbf{R}/\mathbf{Z}$. Also, let $g(t)$ be a smooth function supported on $t \asymp 1$, with bounded derivatives $\|g^{(j)}\|_\infty \ll_j 1$ for $j \geqslant 0$. Then, for any $\eta > 0$, one has*

$$\sum_{\substack{r \sim R \\ s \sim S \\ (r,s) = 1}} \sum_{m \sim M} a_{m,r,s} \sum_{n \sim N} \mathrm{e}(n\omega) \sum_{(c,r) = 1} g\left(\frac{c}{C}\right) S(m\overline{r}, \pm n; sc)$$

$$\ll_\eta x^\eta \left(1 + \frac{C}{R\sqrt{S}}\right)^{\theta_{\max}} \|a_{m,r,s}\|_2 \times \sqrt{NRS} \left(\frac{C^2}{R} (M + RS)(N + RS) + MN \right)^{1/2},$$

where we recall that $\theta_{\max} \leqslant 7/32$ by Theorem A. (The '$\pm n$' notation indicates that either consistent choice of sign is allowable.)

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Theorem 3.10 makes use of the spectral theory of automorphic forms, and follows from a variation of the landmark arguments of Deshouillers and Iwaniec (all of the necessary ingredients being already present in [DI82]). We leave its proof, which requires much additional notation, to § 10.

4. The triple convolution estimate

Here, we state our main technical result, Theorem 4.2, which concerns the distribution in arithmetic progressions of convolutions of three bounded sequences (we point the reader to similar results in [BFI86, Theorem 4], [Dra15, Théorème 3], [DGS17, Lemma 2.3], and [May25a, Proposition 8.3]). We then deduce Theorem 1.5 from Theorem 4.2.

Remark 4.1. One can apply the most efficient convolution estimates directly to the setting of smooth numbers (and smooth-supported multiplicative functions), since these can essentially be factorized into any number of factors of pre-specified sizes. By contrast, in the case of primes, combinatorial decompositions of the von Mangoldt function produce more types of convolution sums, requiring different estimates for different ranges (typically organized into 'type I' and 'type II' information).

To achieve a power saving in Theorem 4.2, appropriate for the application to smooth numbers, one needs a better approximation for indicator functions of the form $\mathbb{1}_{k\equiv 1 \bmod r}$ than $(1/\varphi(r))\mathbb{1}_{(k,r)=1}$ (given $r\in\mathbf{Z}_+$ and k mod r). Drappeau [Dra15, DGS17] noticed that since

$$\mathbb{1}_{k\equiv 1 \bmod r}=\frac{1}{\varphi(r)}\sum_{\chi \pmod r}\chi(k)=\frac{\mathbb{1}_{(k,r)=1}}{\varphi(r)}\sum_{\substack{\chi \text{ primitive}\\ \mathrm{cond}(\chi)\mid r}}\chi(k),$$

one can instead consider the partial sum

$$\omega_D(k;r):=\sum_{\substack{\chi\in\mathcal{X}_D\\ \mathrm{cond}(\chi)\mid r}}\chi(k)\quad\text{where }\mathcal{X}_D:=\{\chi\text{ primitive}:\mathrm{cond}(\chi)\leqslant D\},\tag{4.1}$$

and work with the error term

$$\mathcal{E}_D(k;r):=\mathbb{1}_{k\equiv 1 \pmod r}-\frac{\mathbb{1}_{(k,r)=1}}{\varphi(r)}\omega_D(k;r)=\frac{\mathbb{1}_{(k,r)=1}}{\varphi(r)}\sum_{\substack{\chi \pmod r\\ \mathrm{cond}(\chi)>D}}\chi(k).$$

One should then expect to obtain better bounds for $\mathcal{E}_D(k;r)$ when D is moderately large (i.e., a small power of x) than when $D=1$ (and $\omega_1(k;r)=1$). We also note the crude bound

$$|\omega_D(k;r)|\leqslant|\mathcal{X}(D)|\leqslant D^2,$$

which may be used implicitly in our proofs.

Theorem 4.2 (Triple convolution estimate). *For all sufficiently small $\varepsilon>0$, there exists $\delta>0$ such that the following holds. Let a_1,a_2 be coprime nonzero integers, and let $M,N,L,R,x>2$ satisfy*

$$\begin{aligned}&a_1,a_2\ll x^{\delta},\quad MNL\asymp x,\quad x^{(1-\varepsilon)/2}\ll R\ll x^{-5\varepsilon}NL\ll x^{(2/3)-11\varepsilon},\quad N^9L^8\ll x^3R^4,\\ &N\ll\frac{x^{1-2\varepsilon}}{R},\qquad N^4L^7\max(1,N/L)^{2\theta}\ll x^{2-17\varepsilon}R^2,\quad N^{12-6\theta}L^{11-6\theta}\ll x^{6-4\theta-5\varepsilon}R^2,\end{aligned}\tag{4.2}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

for $\theta = \theta_{\max}$. Then for any 1-bounded complex sequences (α_m), (β_n), (γ_ℓ), one has

$$\sum_{\substack{r\sim R\\(r,a_1a_2)=1}} \Big|\sum_{m\sim M}\sum_{n\sim N}\sum_{\ell\sim L}\alpha_m\beta_n\gamma_\ell\,\mathcal{E}_D(mn\ell\overline{a_1}a_2;r)\Big| \ll_\varepsilon \frac{x(\log x)^4}{\min\big(x^\delta,\sqrt{D}\big)}, \tag{4.3}$$

for all $1\leqslant D\leqslant x^\varepsilon$.

Remark 4.3. Error terms of the form $O_\varepsilon(x^{1-\delta})$ are dominated by the right-hand side of (4.3), and will be available throughout most of our proof. If $x^{2\delta}\leqslant D\leqslant x^\varepsilon$, then the right-hand side of (4.3) becomes $x^{1-\delta}(\log x)^4$, i.e., a power saving; having an explicit dependence on the conductor bound D is required for the application to smooth-supported multiplicative functions, as in [DGS17].

Remark 4.4. If one is free to choose the parameters M, N, and L subject only to the constraints in (4.2) and $MNL\asymp x$, then in order to maximize the range R, it is optimal to pick (up to $x^{o(1)}$ factors)

$$R\approx x^{(5-4\theta)/(8-6\theta)},\quad M\approx\frac{x}{R},\quad N\approx\frac{x}{R},\quad L\approx\frac{R^2}{x}. \tag{4.4}$$

This improves on the conditions from Drappeau's triple convolution estimate [Dra15, (3.2)], which can handle moduli up to $R\approx x^{3/5}$.

Remark 4.5. Although our (conditional) results hit a barrier at $R=x^{5/8-o(1)}$, a more essential limitation of triple convolution estimates proven via the dispersion method lies at $R\leqslant x^{2/3-o(1)}$, corresponding to the case of three equal parameters $M=N=L\asymp x^{1/3}$ in (4.4). Indeed, the diagonal terms in the first Cauchy–Schwarz step require $R<NL$, and already our bounds for the second dispersion sum will use $NL<x^{2/3}$ (moreover, it is natural to Fourier complete in the largest variable m, leaving $NL\leqslant x^{2/3}$). We note again that Theorem 4.2 allows for the case of two equal parameters $M=N\approx x/R$, for any $x^{1/2-o(1)}\ll R\ll x^{(5-4\theta)/(8-6\theta)-o(1)}$; this is possible due to Maynard's deamplification argument [May25a]. In particular, the ranges $M=N=x^{2/5}$, and $L=x^{1/5}$, a limiting case in Drappeau's work [Dra15, Théorème 3]), are now admissible (this is analogous to the infamous case of convolving five sequences of equal sizes).

Given Theorem 4.2 and Lemma 3.2, deducing Theorem 1.5 is now a routine modification of Drappeau's argument in [Dra15, § 3.7] (we follow the same reasoning, using Theorem 4.2 instead of [Dra15, Théorème 3], and with the choice of parameters in (4.4)).

Proof of Theorem 1.5 assuming Theorem 4.2. Let $\varepsilon>0$ be sufficiently small, C be the maximum between $\varepsilon^{-1}(1-\varepsilon)^{-1}$ and the constant C given by Lemma 3.2, δ be the minimum between $\varepsilon/100$ and the δ values of Lemma 3.2 and Theorem 4.2, and let $(\log x)^C\leqslant y\leqslant x^{1/C}$, $D:=x^{\varepsilon/10}$, $\theta:=\theta_{\max}$. It suffices to show (up to a rescaling of ε at the end) that (1.3) holds for the range of moduli

$$r\leqslant x^\alpha,\quad \alpha:=\frac{5-4\theta}{8-6\theta}-1000\varepsilon,$$

and we note that error terms of the form $O_\varepsilon(x^{1-\delta})$ are acceptable in (1.3) (up to slightly modifying the value of δ), due to the inequality $x^{1-\delta}\ll\Psi(x,y)y^{-\delta/2}$ (as in [Dra15, § 3.7]). We may of course assume that a_1 and a_2 are relatively prime, by reducing any common factors.

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

We split the left-hand side of (1.3) into

$$\sum_{\substack{r\leqslant x^{\alpha}\\(r,a_1a_2)=1}}\left|\sum_{n\in S(x,y)}\left(\mathbb{1}_{n\overline{a_1}a_2\equiv 1 \pmod r}-\frac{\mathbb{1}_{(n,r)=1}}{\varphi(r)}\right)\right|$$

$$\leqslant \sum_{\substack{r\leqslant x^{\alpha}\\(r,a_1a_2)=1}}\left|\sum_{n\in S(x,y)}\mathcal{E}_D(n\overline{a_1}a_2;r)\right|+\sum_{\substack{r\leqslant x^{\alpha}\\(r,a_1a_2)=1}}\left|\sum_{n\in S(x,y)}\frac{\mathbb{1}_{(n,r)=1}}{\varphi(r)}\sum_{\substack{\chi \text{ prim., } \mathrm{cond}(\chi)|r\\1<\mathrm{cond}(\chi)\leqslant D}}\chi(n\overline{a_1}a_2)\right|.$$

The second sum is at most

$$\sum_{\substack{r\leqslant x^{\alpha}\\(r,a_1a_2)=1}}\frac{1}{\varphi(r)}\sum_{\substack{\chi \pmod r\\1<\mathrm{cond}(\chi)\leqslant D}}\left|\sum_{n\in S(x,y)}\chi(n)\right|,$$

which is appropriately bounded by Lemma 3.2 and the triangle inequality. It remains to bound the first sum.

Recall that the range $n\in S(x,y)$ means $n\leqslant x$ and $P^+(n)\leqslant y$, where $P^+(n)$ denotes the greatest prime factor of n. We bound the contribution of $n\leqslant x^{1-\varepsilon}$ trivially, as in [Dra15, § 3.7]

$$\sum_{\substack{r\leqslant x^{\alpha}\\(r,a_1a_2)=1}}\left|\sum_{\substack{n\leqslant x^{1-\varepsilon}\\P^+(n)\leqslant y}}\mathcal{E}_D(n\overline{a_1}a_2;r)\right|\leqslant\sum_{\substack{r\leqslant x^{\alpha}\\(r,a_1a_2)=1}}\sum_{n\leqslant x^{1-\varepsilon}}\left(\mathbb{1}_{a_2n\equiv a_1 \pmod r}+\frac{D^2}{\varphi(r)}\right)$$

$$\ll_{\varepsilon} x^{\alpha}+\sum_{\substack{n\leqslant x^{1-\varepsilon}\\a_2n\neq a_1}}\tau(|a_2n-a_1|)+x^{\varepsilon/2}\sum_{n\leqslant x^{1-\varepsilon}}1\ll_{\varepsilon} x^{1-\varepsilon/2},$$

and put the other n values into $O(\log x)$ dyadic ranges $n\sim X$, with $x^{1-\varepsilon}\leqslant X\ll x$. We also extend the range of r in these sums to $r\leqslant X^{\alpha+10\varepsilon}$, noting that $x^{\alpha}\leqslant x^{(1-\varepsilon)(\alpha+10\varepsilon)}$. Putting r into $O(\log x)$ dyadic ranges $r\sim R$, it remains to bound sums of the form

$$\sum_{\substack{r\sim R\\(r,a_1a_2)=1}}\left|\sum_{\substack{n\sim X\\P^+(n)\leqslant y}}\mathcal{E}_D(n\overline{a_1}a_2;r)\right|, \tag{4.5}$$

for $R\ll X^{\alpha+10\varepsilon}$. The contribution of the Bombieri–Vinogradov range $R\ll X^{(1/2)-(\varepsilon/10)}$ is handled by classical methods (e.g., using [IK21, Theorem 17.4]; see [DGS17, Proof of Proposition 2.4] and [Dra15, Proof of Proposition 2]). For any R in the remaining range $X^{(1/2)-(\varepsilon/10)}\leqslant R\ll X^{\alpha+10\varepsilon}$, we set

$$M_0:=\frac{X^{1-10\varepsilon}}{R},\quad N_0:=\frac{X^{1-10\varepsilon}}{R},\quad L_0:=\frac{R^2}{X^{1-20\varepsilon}}, \tag{4.6}$$

and factorize smooth numbers as in [Dra15, Lemme 7] (or [FT96]) to rewrite the sum in (4.5) as

$$\sum_{\substack{r\sim R\\(r,a_1a_2)=1}}\left|\sum_{\substack{L_0<\ell\leqslant L_0P^-(\ell)\\P^+(\ell)\leqslant y}}\sum_{\substack{M_0<m\leqslant M_0P^-(m)\\P^+(m)\leqslant P^-(\ell)}}\sum_{\substack{mn\ell\sim X\\P^+(n)\leqslant P^-(m)}}\mathcal{E}_D(mn\ell\overline{a_1}a_2;r)\right|,$$

where $P^-(n)$ denotes the smallest prime factor of n.

We also put m,n,ℓ into $O((\log y)^3)$ dyadic ranges $m\sim M$, $n\sim N$, $\ell\sim L$, with $M\in[M_0,yM_0]$, $N\in[y^{-2}N_0,N_0]$, $L\in[L_0,yL_0]$, and $MNL\asymp X$. Recalling that $y\leqslant x^{\varepsilon(1-\varepsilon)}\leqslant X^{\varepsilon}$, it is easily checked that for such M,N,L and $X^{(1/2)-(\varepsilon/10)}\leqslant R\ll X^{\alpha+10\varepsilon}=X^{(5-4\theta)/(8-6\theta)-990\varepsilon}$, and small enough ε, the conditions in (4.2) are satisfied (with respect to X instead of x).

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

At this point, our sums are almost in the right form to apply the triple convolution estimate in Theorem 4.2, except for a few joint constraints on the variables m, n, ℓ (these are $P^+(m) \leqslant P^-(\ell)$, $P^+(n) \leqslant P^-(m)$, respectively, $mn\ell \sim X$). The last step of analytically separating these constraints is identical to that in [Dra15, § 3.7], except that in the end we apply Theorem 4.2 instead of [Dra15, Théorème 3]. Overall, the contribution of the range $X^{(1/2)-(\varepsilon/10)} \leqslant R \ll X^{\alpha+10\varepsilon}$ is $O_\varepsilon((\log x)^{O(1)} x^{1-\delta})$, which is acceptable; this completes our proof. □

We only briefly note that the result for smooth-supported multiplicative functions in Theorem 1.7 follows by an analogous modification to the arguments in [DGS17], using the parameters in (4.6), and Theorem 4.2 instead of [DGS17, Lemma 2.3]. The main additional difficulty in [DGS17] lies in the contribution of the small-conductor characters, since Lemma 3.2 is no longer applicable; as a replacement, Drappeau, Granville, and Shao developed a large sieve inequality for smooth-supported sequences [DGS17, Theorem 5.1]. (We also point the reader to the follow-up work of Shparlinski in [Shp18].)

5. Dispersion and deamplification

Our goal for the rest of this paper is to prove Theorem 4.2, proceeding by Linnik's dispersion method. For the reader following the outline in § 2.1, the exponential sum from (2.3) will ultimately arise in the first dispersion sum, after Poisson summation (see Proposition 8.2).

Assume the set-up of Theorem 4.2. We may take x larger than an absolute constant, since the conclusion of Theorem 4.2 is trivial otherwise, and (α_m), (β_n), and (γ_ℓ) to be supported on $m \sim M$, $n \sim N$, and $\ell \sim L$, without loss of generality. We first combine the sequences β_n and γ_ℓ into one sequence

$$u_k := \sum_{n\ell=k} \beta_n \gamma_\ell, \tag{5.1}$$

supported in $(K, 4K]$ where $K := NL$, $|u_k| \leqslant \tau(k) \ll_\varepsilon x^{\varepsilon/2}$, and $\sum_k |u_k| \ll K$. Denoting the left-hand side of (4.3) by $\Delta = \Delta_D(M, N, L, R)$, we can introduce coefficients ρ_r of absolute value 1, supported in $(R, 2R]$, to rewrite

$$\Delta = \sum_{(r, a_1a_2)=1} \rho_r \sum_{(m,r)=1} \alpha_m \sum_{(k,r)=1} u_k \left(\mathbb{1}_{mk \equiv a_1\overline{a_2} \pmod r} - \frac{1}{\varphi(r)} \omega_D(mk\overline{a_1}a_2; r) \right),$$

where we recall that ω_D was defined in (4.1). Normally, at this point we would apply Cauchy and Schwarz in the r, m variables, but we first perform a 'deamplification' step (following Maynard [May25a] with minor modifications), as anticipated in § 2.3. The idea is to split the inner sum according to the residue class of $k \pmod{re}$ for some $e \geqslant 1$, and then to average over a convenient set of $e \sim E$; this artificially introduces a new parameter E (to be chosen later), which will help reduce the contribution of a certain diagonal sum by a small power of x (at the expense of increasing the corresponding off-diagonal terms, which already had a power-saving bound). For now, we require that

$$x^{4\varepsilon} \leqslant E \ll x^{-\varepsilon} \frac{K}{R}, \tag{5.2}$$

which is compatible with (4.2) since $R \ll x^{-5\varepsilon} NL$. For multiple reasons of convenience throughout our proof, we will actually average over the set

$$\mathscr{E} := \{e \sim E : e \text{ prime}\}. \tag{5.3}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Proposition 5.1 (Dispersion set-up with deamplification). *Let Φ be a smooth function satisfying*

$$\mathbb{1}_{[1,2]} \leqslant \Phi \leqslant \mathbb{1}_{(0.5,3)}, \tag{5.4}$$

and $\|\Phi^{(j)}\|_\infty \ll_j 1$ for $j \geqslant 0$. Then, for any $\varepsilon > 0$ and $1 \gg \delta > 0$, under the parameter conditions in (4.2) and (5.2), one has

$$\Delta \ll_\varepsilon x\sqrt{\frac{R\log x}{K^2}(\mathcal{S}_1 - 2\mathrm{Re}\mathcal{S}_2 + \mathcal{S}_3)} + x^{1-\varepsilon}, \tag{5.5}$$

where

$$\begin{aligned}
\mathcal{S}_1 &:= \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\sum_{(m,r)=1}\frac{1}{M}\Phi\left(\frac{m}{M}\right)\sum_{\substack{(k_1k_2,re)=1\\ k_1\equiv k_2 \pmod{re}\\ k_i\equiv a_1\overline{a_2m} \pmod r}} u_{k_1}\overline{u_{k_2}},\\
\mathcal{S}_2 &:= \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\frac{1}{\varphi(re)}\sum_{(m,r)=1}\frac{1}{M}\Phi\left(\frac{m}{M}\right)\sum_{\substack{(k_1k_2,re)=1\\ k_2\equiv a_1\overline{a_2m} \pmod r}} u_{k_1}\omega_D(mk_1\overline{a_1}a_2;r)\overline{u_{k_2}},\\
\mathcal{S}_3 &:= \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\frac{1}{\varphi(r)\varphi(re)}\sum_{(m,r)=1}\frac{1}{M}\Phi\left(\frac{m}{M}\right)\left|\sum_{(k,re)=1} u_k\omega_D(mk\overline{a_1}a_2;r)\right|^2.
\end{aligned} \tag{5.6}$$

Proof. For a fixed prime $e \sim E$, we wish to eliminate the contribution to Δ of the terms k with $(e,k) \neq 1$, i.e., $e \mid k$. This contribution is

$$\begin{aligned}
&\sum_{(r,a_1a_2)=1}\rho_r\sum_{(m,r)=1}\alpha_m\sum_{\substack{k\in(K,4K]\\ (k,r)=1, e\mid k}} u_k\left(\mathbb{1}_{mk\equiv a_1\overline{a_2} \pmod r} - \frac{1}{\varphi(r)}\omega_D(mk\overline{a_1}a_2;r)\right)\\
&\ll_\varepsilon \sum_{\substack{r\sim R\\ (r,a_1a_2e)=1}}\sum_{\substack{s\in(MK,8MK]\\ e\mid s}} x^{\varepsilon/2}\left(\mathbb{1}_{s\equiv a_1\overline{a_2} \pmod r} + \frac{1}{\varphi(r)}D^2\right)\\
&\ll_\varepsilon x^\varepsilon\sum_{r\sim R}\left(\frac{MK}{RE}+1+\left(\frac{MK}{E}+1\right)\frac{D^2}{R}\right)\\
&\ll x^\varepsilon\left(\frac{xD^2}{E}+R+D^2\right) \ll x^\varepsilon\left(\frac{x^{1+2\varepsilon}}{x^{4\varepsilon}}+R+x^{2\varepsilon}\right),
\end{aligned}$$

which is $\ll x^{1-\varepsilon}$ since $R \ll x^{(2/3)-11\varepsilon}$ by (4.2). It follows that, for any $e \sim E$,

$$\Delta = \sum_{(r,a_1a_2)=1}\rho_r\sum_{(m,r)=1}\alpha_m\sum_{(k,re)=1} u_k\left(\mathbb{1}_{mk\equiv a_1\overline{a_2} \pmod r} - \frac{1}{\varphi(r)}\omega_D(mk\overline{a_1}a_2;r)\right) + O_\varepsilon(x^{1-\varepsilon}).$$

Now for fixed m and e we have

$$\sum_{(k,re)=1} u_k\mathbb{1}_{mk\equiv a_1\overline{a_2} \pmod r} = \sum_{\substack{b\in(\mathbf{Z}/re\mathbf{Z})^\times\\ b\equiv a_1\overline{a_2m} \pmod r}}\sum_{(k,re)=1} u_k\mathbb{1}_{k\equiv b \pmod{re}},$$

and there are precisely $\varphi(re)/\varphi(r)$ choices of b in the summation; thus

$$\Delta = \sum_{(r,a_1a_2)=1}\rho_r\sum_{(m,r)=1}\alpha_m\sum_{\substack{b\in(\mathbf{Z}/re\mathbf{Z})^\times\\ b\equiv a_1\overline{a_2m} \pmod r}}\sum_{(k,re)=1} u_k\left(\mathbb{1}_{k\equiv b \pmod{re}} - \frac{\omega_D(mk\overline{a_1}a_2;r)}{\varphi(re)}\right) + O_\varepsilon(x^{1-\varepsilon}).$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

(For a complete deamplification set-up, one could also try to split the term $\omega_D(mk\overline{a_1}a_2;r)$ according to the residue b of $k \pmod{re}$, but we do not need to do this in our proof.)

We then average over e in the set $\mathcal{E}$ from (5.3), which has size $|\mathscr{E}| \asymp_\varepsilon E/\log E$ (recalling that $E \gg x^\varepsilon$ and $|a_2| \leqslant x^\delta$). Thus up to an error of $O_\varepsilon(x^{1-\varepsilon})$, we can rewrite Δ as

$$\frac{1}{|\mathscr{E}|}\sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\rho_r\sum_{(m,r)=1}\alpha_m\sum_{\substack{b\in(\mathbf{Z}/re\mathbf{Z})^\times\\ b\equiv a_1\overline{a_2m} \pmod r}}\sum_{(k,re)=1}u_k\left(\mathbb{1}_{k\equiv b \pmod{re}}-\frac{\omega_D(mk\overline{a_1}a_2;r)}{\varphi(re)}\right).$$

We now apply Cauchy and Schwarz in the e, r, m, b variables, allowing us to eliminate the ρ_r and α_m coefficients; using that $\varphi(re)\leqslant\varphi(r)e$, this gives

$$\begin{aligned}|\Delta|^2 &\ll \frac{1}{|\mathscr{E}|^2}\left(\sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\sum_{(m,r)=1}\sum_{\substack{b\in(\mathbf{Z}/re\mathbf{Z})^\times\\ b\equiv a_1\overline{a_2m} \pmod r}}|\rho_r|^2|\alpha_m|^2\right)\Delta' + O_\varepsilon(x^{2(1-\varepsilon)})\\ &\ll \frac{1}{|\mathscr{E}|^2}\left(\sum_{e\in\mathscr{E}}\sum_{\substack{r\sim R\\ (r,a_1a_2)=1}}\sum_{\substack{m\sim M\\ (m,r)=1}}\frac{\varphi(re)}{\varphi(r)}\right)\Delta' + O_\varepsilon(x^{2(1-\varepsilon)})\\ &\ll_\varepsilon MR(\log E)\Delta' + x^{2(1-\varepsilon)},\end{aligned} \tag{5.7}$$

where

$$\Delta' := \sum_{e\in\mathscr{E}}\sum_{\substack{r\sim R\\ (r,a_1a_2)=1}}\sum_{\substack{m\sim M\\ (m,r)=1}}\sum_{\substack{b\in(\mathbf{Z}/re\mathbf{Z})^\times\\ b\equiv a_1\overline{a_2m} \pmod r}}\left|\sum_{(k,re)=1}u_k\left(\mathbb{1}_{k\equiv b \pmod{re}}-\frac{\omega_D(mk\overline{a_1}a_2;r)}{\varphi(re)}\right)\right|^2.$$

Anticipating a later application of Poisson summation, we bound the indicator functions $\mathbb{1}_{m\sim M}$ and $\mathbb{1}_{r\sim R}$ from above by $\Phi(m/M)$ and $\Phi(r/R)$. Then we expand the square and perform the b-summation to obtain

$$\begin{aligned}\Delta' &\leqslant \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\sum_{(m,r)=1}\Phi\left(\frac{m}{M}\right)\sum_{\substack{b\in(\mathbf{Z}/re\mathbf{Z})^\times\\ b\equiv a_1\overline{a_2m} \pmod r}}\left|\sum_{(k,re)=1}u_k\left(\mathbb{1}_{k\equiv b \pmod{re}}-\frac{\omega_D(mk\overline{a_1}a_2;r)}{\varphi(re)}\right)\right|^2\\ &= M\left(\mathcal{S}_1 - 2\mathrm{Re}\mathcal{S}_2 + \mathcal{S}_3\right).\end{aligned} \tag{5.8}$$

Combining (5.7) with (5.8) and recalling that $M \asymp x/K$, we recover (5.5). □

Since the error term of $O_\varepsilon(x^{1-\varepsilon})$ in Proposition 5.1 is admissible for Theorem 4.2, it remains to estimate the dispersion sums $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$.

6. The main terms

We note that, except for the coefficients $\Phi(m/M)$, only the residue of m modulo r matters in the inner summations from (5.6). Thus if we define

$$\mathcal{E}_r(c) := \left(\frac{1}{M}\sum_{m\equiv c \pmod r}\Phi\left(\frac{m}{M}\right)\right) - \frac{\widehat{\Phi}(0)}{r}, \tag{6.1}$$

which can be estimated via the truncated Poisson summation in Lemma 3.5, we can rewrite

$$\mathcal{S}_1 = \widehat{\Phi}(0)X_1 + \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}\mathcal{E}_r(c)\sum_{\substack{(k_1k_2,re)=1\\ k_1\equiv k_2 \pmod{re}\\ k_i\equiv a_1\overline{a_2c} \pmod r}}u_{k_1}\overline{u_{k_2}},$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

$$\mathcal{S}_2 = \widehat{\Phi}(0)X_2 + \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\frac{1}{\varphi(re)}\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}\mathcal{E}_r(c)\sum_{\substack{(k_1k_2,re)=1\\ k_2\equiv a_1\overline{a_2}c \pmod r}} u_{k_1}\omega_D(k_1\overline{a_1}a_2c;r)\overline{u_{k_2}},$$

$$\mathcal{S}_3 = \widehat{\Phi}(0)X_3 + \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\frac{1}{\varphi(r)\varphi(re)}\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}\mathcal{E}_r(c)\left|\sum_{(k,re)=1}u_k\omega_D(k\overline{a_1}a_2c;r)\right|^2, \quad (6.2)$$

where

$$\begin{aligned}
X_1 &:= \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\frac{1}{r}\sum_{\substack{(k_1k_2,re)=1\\ k_1\equiv k_2 \pmod{re}}} u_{k_1}\overline{u_{k_2}},\\
X_2 &:= \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\frac{1}{r\varphi(re)}\sum_{(k_1k_2,re)=1} u_{k_1}\overline{u_{k_2}}\omega_D(k_1\overline{k_2};r), \qquad (6.3)\\
X_3 &:= \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\frac{1}{r\varphi(r)\varphi(re)}\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}\left|\sum_{(k,re)=1}u_k\omega_D(kc;r)\right|^2.
\end{aligned}$$

Intuitively, these main terms reflect what would happen if, in the summations from (5.6), the variable m (weighted by $\Phi(m/M)$) were uniformly distributed modulo r. Thus for $j\in\{1,2,3\}$, $\widehat{\Phi}(0)X_j$ is essentially the best approximation to $\mathcal{S}_j$ which does not depend on M. We now bound the contribution to (5.5) of $X_1 - 2\mathrm{Re}X_2 + X_3$, using the multiplicative large sieve as in [Dra15, DGS17].

Proposition 6.1 (Contribution of main terms). *With the notation above, one has*

$$0 \leqslant X_1 - 2\mathrm{Re}X_2 + X_3 \ll \frac{K^2}{RD}(\log x)^6,$$

under the parameter conditions in (4.2) and (5.2).

Proof. In analogy with (5.8), we can write

$$\begin{aligned}
&X_1 - 2\mathrm{Re}X_2 + X_3\\
&= \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\frac{1}{r}\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}\sum_{\substack{b\in(\mathbf{Z}/re\mathbf{Z})^\times\\ b\equiv c \pmod r}}\left|\sum_{(k,re)=1}u_k\left(\mathbb{1}_{k\equiv b \pmod{re}} - \frac{\omega_D(k\overline{c};r)}{\varphi(re)}\right)\right|^2\\
&\ll \frac{1}{R}\sum_{e\in\mathscr{E}}\sum_{r}\Phi\left(\frac{r}{R}\right)\sum_{b\in(\mathbf{Z}/re\mathbf{Z})^\times}\left|\sum_{(k,re)=1}u_k\left(\mathbb{1}_{k\equiv b \pmod{re}} - \frac{\omega_D(k\overline{b};r)}{\varphi(re)}\right)\right|^2\\
&= \frac{1}{R}\sum_{e\in\mathscr{E}}\sum_{r}\Phi\left(\frac{r}{R}\right)\frac{1}{\varphi(re)^2}\sum_{b\in(\mathbf{Z}/re\mathbf{Z})^\times}\left|\sum_{(k,re)=1}u_k(\omega_\infty(k\overline{b};re) - \omega_D(k\overline{b};r))\right|^2, \qquad (6.4)
\end{aligned}$$

where the first equality shows that $X_1 - 2\mathrm{Re}X_2 + X_3 \geqslant 0$. Note that

$$\begin{aligned}
\omega_\infty(k\overline{b};re) - \omega_D(k\overline{b};r) &= (\omega_\infty(k\overline{b};re) - \omega_\infty(k\overline{b};r)) + (\omega_\infty(k\overline{b};r) - \omega_D(k\overline{b};r))\\
&= \sum_{\chi\in S}\chi(k)\overline{\chi(b)},
\end{aligned}$$

where $S = S(r,e,\varepsilon) = S_1\cup S_2$ and

$$\begin{aligned}
S_1 &:= \{\chi \text{ primitive} : \mathrm{cond}(\chi)\nmid r, \text{ but } \mathrm{cond}(\chi)\mid re\},\\
S_2 &:= \{\chi \text{ primitive} : \mathrm{cond}(\chi) > D \text{ and } \mathrm{cond}(\chi)\mid r\}.
\end{aligned}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Since all the characters in S are primitive, any distinct $\chi_1, \chi_2 \in S$ must induce different characters modulo re. Thus $\chi_1\overline{\chi_2}\mathbb{1}_{(re,\cdot)=1}$ is not the principal character modulo re, so it must have average 0. But then

$$\begin{aligned}
&\sum_{b\in(\mathbf{Z}/re\mathbf{Z})^\times}\Big|\sum_{(k,re)=1}u_k(\omega_\infty(k\bar{b};re)-\omega_D(k\bar{b};r))\Big|^2\\
&=\sum_{b\in(\mathbf{Z}/re\mathbf{Z})^\times}\Big|\sum_{\chi\in S}\overline{\chi}(b)\sum_{(k,re)=1}u_k\chi(k)\Big|^2\\
&=\sum_{\chi_1,\chi_2\in S}\Big(\sum_{(k,re)=1}u_k\chi_1(k)\Big)\overline{\Big(\sum_{(k,re)=1}u_k\chi_2(k)\Big)}\sum_{b\in(\mathbf{Z}/re\mathbf{Z})^\times}\overline{\chi_1}(b)\chi_2(b)\\
&=\varphi(re)\sum_{\chi\in S}\Big|\sum_{(k,re)=1}u_k\chi(k)\Big|^2. \qquad (6.5)
\end{aligned}$$

From (6.4), (6.5), and the fact that all characters $\chi\in S_1$ also have $\mathrm{cond}(\chi)>D$ (due to $D\leqslant x^\varepsilon<E\leqslant e\,|\,\mathrm{cond}(\chi)$), we conclude that

$$\begin{aligned}
X_1-2\mathrm{Re}X_2+X_3&\ll\frac{1}{R}\sum_{e\in\mathscr{E}}\sum_r\Phi\Big(\frac{r}{R}\Big)\frac{1}{\varphi(re)}\sum_{\chi\in S}\Big|\sum_{(k,re)=1}u_k\chi(k)\Big|^2\\
&\leqslant\frac{1}{R}\sum_{e\in\mathscr{E}}\sum_r\Phi\Big(\frac{r}{R}\Big)\frac{1}{\varphi(re)}\sum_{\substack{\chi\text{ primitive}\\ \mathrm{cond}(\chi)>D\\ \mathrm{cond}(\chi)|re}}\Big|\sum_{(k,re)=1}u_k\chi(k)\Big|^2.
\end{aligned}$$

Now letting $Q:=RE$, substituting q for re, using that q has $O(\log q)$ different prime factors, and decomposing $\mathbb{1}_{(a,b)=1}=\sum_{d|(a,b)}\mu(d)$ to get rid of the coprimality restriction, we can bound the sum above by

$$\begin{aligned}
X_1-2\mathrm{Re}X_2+X_3&\ll\frac{\log x}{R}\sum_{Q/2\leqslant q\leqslant 6Q}\frac{1}{\varphi(q)}\sum_{\substack{\chi\text{ primitive}\\ \mathrm{cond}(\chi)>D\\ \mathrm{cond}(\chi)|q}}\Big|\sum_{(k,q)=1}u_k\chi(k)\Big|^2\\
&\leqslant\frac{\log x}{R}\sum_{D<s\leqslant 6Q}\sum_{\substack{\chi\text{ primitive}\\ \mathrm{cond}(\chi)=s}}\sum_{\substack{q\leqslant 6Q\\ s|q}}\frac{1}{\varphi(q)}\Big|\sum_{(k,q/s)=1}u_k\chi(k)\Big|^2\\
&\leqslant\frac{\log x}{R}\sum_{D<s\leqslant 6Q}\sum_{\substack{\chi\text{ primitive}\\ \mathrm{cond}(\chi)=s}}\sum_{\substack{q\leqslant 6Q\\ s|q}}\frac{\tau(q/s)}{\varphi(q)}\sum_{d|q/s}\Big|\sum_{k'}u_{dk'}\chi(dk')\Big|^2\\
&\leqslant\frac{\log x}{R}\sum_{d\leqslant 6Q}\sum_{D<s\leqslant 6Q}\sum_{\substack{\chi\text{ primitive}\\ \mathrm{cond}(\chi)=s}}\Big|\sum_{k'}u_{dk'}\chi(k')\Big|^2\sum_{\substack{q\leqslant 6Q\\ ds|q}}\frac{\tau(q/s)}{\varphi(q)}.
\end{aligned}$$

Noting that

$$\sum_{\substack{q\leqslant 6Q\\ sd|q}}\frac{\tau(q/s)}{\varphi(q)}\leqslant\sum_{q'\leqslant 6Q}\frac{\tau(q'd)}{\varphi(q'ds)}\leqslant\frac{\tau(d)}{\varphi(d)\varphi(s)}\sum_{q'\leqslant 6Q}\frac{\tau(q')}{\varphi(q')}\ll\frac{\tau(d)}{\varphi(d)\varphi(s)}(\log x)^2,$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

we further have

$$X_1 - 2\mathrm{Re}X_2 + X_3 \ll \frac{(\log x)^3}{R} \sum_{d \leqslant 6Q} \frac{\tau(d)}{\varphi(d)} \sum_{D < s \leqslant 6Q} \frac{1}{\varphi(s)} \sum_{\substack{\chi \text{ primitive} \\ \mathrm{cond}(\chi) = s}} \left| \sum_{k'} u_{dk'} \chi(k') \right|^2$$
$$= \frac{(\log x)^3}{R} \sum_{d \leqslant 6Q} \frac{\tau(d)}{\varphi(d)} \int_D^\infty \sum_{D < s \leqslant \min(6Q, t)} \frac{s}{\varphi(s)} \sum_{\substack{\chi \text{ primitive} \\ \mathrm{cond}(\chi) = s}} \left| \sum_{k'} u_{dk'} \chi(k') \right|^2 \frac{dt}{t^2}.$$

Finally, applying the multiplicative large sieve from Lemma 3.4 as in [DGS17, (2.6)], we conclude that

$$X_1 - 2\mathrm{Re}X_2 + X_3 \ll \frac{(\log x)^3}{R} \sum_{d \leqslant 6Q} \frac{\tau(d)}{\varphi(d)} \int_D^\infty \left(\frac{K}{d} + \min(6Q, t)^2 \right) \sum_{k' \asymp K/d} |u_{dk'}|^2 \frac{dt}{t^2}$$
$$\ll \frac{(\log x)^3}{R} K (\log K)^3 \sum_{d \leqslant 6Q} \frac{\tau(d)^3}{d\varphi(d)} \left(\frac{K}{dD} + Q \right)$$
$$\ll (\log x)^6 \frac{K}{R} \left(\frac{K}{D} + Q \right).$$

Using the condition $RE \ll x^{-\varepsilon} K$ from (5.2) to bound $Q = RE \ll K/D$, we conclude that

$$X_1 - 2\mathrm{Re}X_2 + X_3 \ll (\log x)^6 K^2 (RD)^{-1},$$

as we wanted. □

To bridge Propositions 5.1 and 6.1, it remains to compare the dispersion sums $\mathcal{S}_j$ with their main terms $\widehat{\Phi}(0) X_j$; we make the following claim.

Proposition 6.2 (Estimates for dispersion sums). *For all sufficiently small $\varepsilon > 0$, there exists $\delta > 0$ such that, with the notation in (6.2), the following hold.*

(i) *Assuming the ranges in (4.2), there exists a choice of E satisfying (5.2) such that*

$$\mathcal{S}_1 - \widehat{\Phi}(0) X_1 \ll_\varepsilon x^{-2\delta} \frac{K^2}{R}. \tag{6.6}$$

(ii) *Assuming both (4.2) and (5.2), one has*

$$\mathcal{S}_2 - \widehat{\Phi}(0) X_2 \ll_\varepsilon x^{-2\delta} \frac{K^2}{R}, \tag{6.7}$$

$$\mathcal{S}_3 - \widehat{\Phi}(0) X_3 \ll_\varepsilon x^{-2\delta} \frac{K^2}{R}. \tag{6.8}$$

Proof of Theorem 4.2 assuming Proposition 6.2. The hypothesis of Theorem 4.2 assumes (4.2), so we can pick E as in Proposition 6.2(i), subject to (5.2). Then, by combining Propositions 5.1, 6.1 and 6.2, we obtain

$$\Delta \ll_\varepsilon x \sqrt{\frac{R \log x}{K^2} \left(\widehat{\Phi}(0) \frac{K^2}{RD} (\log x)^6 + x^{-2\delta} \frac{K^2}{R} \right)} + x^{1-\varepsilon}$$
$$\ll x \sqrt{(\log x)^7 \left(\frac{1}{D} + \frac{1}{x^{2\delta}} \right)} + x^{1-\varepsilon}.$$

The conclusion of Theorem 4.2 follows after replacing δ with $\min(\delta, \varepsilon)$. □

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Our remaining task is to prove Proposition 6.2; the truncated Poisson expansion of the coefficients $\mathcal{E}_r(c)$ from (6.2) will ultimately reduce our problem to that of bounding various exponential sums. We note that we have not chosen the value of δ in terms of ε yet; the condition $\delta \leqslant \varepsilon/2$ will suffice for estimating $\mathcal{S}_2$ and $\mathcal{S}_3$, but more will be needed for the (much more involved) study of $\mathcal{S}_1$.

7. The second and third dispersion sums

Here, we prove Proposition 6.2(ii), adapting Drappeau's arguments in [Dra15, §§ 3.2 and 3.3]. We assume all the parameter conditions in (4.2) and (5.2).

Proof of (6.8), estimating $\mathcal{S}_3$. Recall from (6.2) that

$$\begin{aligned}\mathcal{S}_3 - \widehat{\Phi}(0)X_3 &= \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\frac{1}{\varphi(r)\varphi(re)}\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}\mathcal{E}_r(c)\Big|\sum_{(k,re)=1}u_k\sum_{\substack{\chi\in\mathcal{X}_D\\ \mathrm{cond}(\chi)|r}}\chi(k\overline{a_1}a_2c)\Big|^2\\ &= \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\frac{1}{\varphi(r)\varphi(re)}\sum_{\substack{\chi_1,\chi_2\in\mathcal{X}_D\\ \mathrm{cond}(\chi_i)|r}}\overline{\chi_1}\chi_2(a_1\overline{a_2})\sum_{\substack{k_1,k_2\\ (k_1k_2,re)=1}}\chi_1(k_1)u_{k_1}\overline{\chi_2(k_2)u_{k_2}}\\ &\quad\times\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}\chi_1\overline{\chi_2}(c)\mathcal{E}_r(c),\end{aligned}$$

where $\mathcal{E}_r(c)$ is given by (6.1). Expanding $\mathcal{E}_r(c)$ according to Lemma 3.5 with $H := x^\varepsilon RM^{-1}$, we obtain

$$\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}\chi_1\overline{\chi_2}(c)\mathcal{E}_r(c) = \frac{1}{r}\sum_{0<|h|<H}\widehat{\Phi}\left(\frac{hM}{r}\right)\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}\chi_1\overline{\chi_2}(c)\,\mathrm{e}\left(\frac{ch}{r}\right)+O_\varepsilon(x^{-99}). \tag{7.1}$$

(In such manipulations, we warn the reader of the potential confusion between the integer variable $e\in\mathscr{E}$ and the function $\mathrm{e}(\cdot)$; the difference should be clear from context.)

The inner sum (over c) in (7.1) is a Gauss sum, which we can bound using Lemma 3.7 for the Dirichlet character $\chi_1\overline{\chi_2}\mathbb{1}_{(\cdot,r)=1} \pmod r$ (whose conductor divides r and is at most equal to $D^2\leqslant x^{2\varepsilon}$). This yields

$$\begin{aligned}\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}\chi_1\overline{\chi_2}(c)\mathcal{E}_r(c) &\ll_\varepsilon x^{-100}+\frac{1}{r}\sum_{0<|h|<H}x^\varepsilon\sum_{d|(h,r)}d\\ &\ll x^{-100}+\frac{x^\varepsilon}{r}\sum_{d|r}d\sum_{0<|h|<H/d}1\ll\frac{x^\varepsilon}{R}\tau(r)H=x^{2\varepsilon}\frac{\tau(r)}{M},\end{aligned}$$

which leads to

$$\begin{aligned}\mathcal{S}_3-\widehat{\Phi}(0)X_3\ll_\varepsilon x^{2\varepsilon}\sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\frac{\tau(r)}{\varphi(r)\varphi(re)M}|\mathcal{X}_D|^2K^2&\ll x^{6\varepsilon}\frac{K^2}{M}\sum_r\Phi\left(\frac{r}{R}\right)\frac{\tau(r)}{\varphi(r)^2}\\ &\ll x^{6\varepsilon}\frac{K^2}{MR}(\log R)^{O(1)}.\end{aligned}$$

Since $x/M\asymp K\ll x^{(2/3)-6\varepsilon}$ by (4.2), we have $M\gg x^{1/3+6\varepsilon}\gg x^{7\varepsilon}$ for small enough ε, and in particular $\mathcal{S}_3-\widehat{\Phi}(0)X_3\ll_\varepsilon x^{-\varepsilon}K^2/R$, proving the easiest third of Proposition 6.2. □

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Proof of (6.7), estimating $\mathcal{S}_2$. Recall from (6.2) that

$$\begin{aligned}\mathcal{S}_2 - \widehat{\Phi}(0)X_2 &= \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\frac{1}{\varphi(re)}\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}\mathcal{E}_r(c)\sum_{\substack{(k_1k_2,re)=1\\ k_2\equiv a_1\overline{a_2c}\ (\mathrm{mod}\ r)}}u_{k_1}\overline{u_{k_2}}\sum_{\substack{\chi\in\mathcal{X}_D\\ \mathrm{cond}(\chi)|r}}\chi(k_1\overline{a_1}a_2c)\\ &= \sum_{e\in\mathscr{E}}\sum_{\chi\in\mathcal{X}_D}\sum_{(k_1k_2,e)=1}\chi(k_1)u_{k_1}\overline{\chi(k_2)u_{k_2}}\sum_{\substack{(r,a_ik_i)=1\\ \mathrm{cond}(\chi)|r}}\Phi\left(\frac{r}{R}\right)\frac{1}{\varphi(re)}\mathcal{E}_r(a_1\overline{a_2k_2}).\end{aligned}$$

Applying Lemma 3.5 with $H := x^{\varepsilon}RM^{-1}$ to expand $\mathcal{E}_r(a_1\overline{a_2k_2})$ (as given in (6.1)), we obtain

$$\begin{aligned}\mathcal{S}_2 - \widehat{\Phi}(0)X_2 &= \sum_{e\in\mathscr{E}}\sum_{\chi\in\mathcal{X}_D}\sum_{(k_1k_2,e)=1}\chi(k_1)u_{k_1}\overline{\chi(k_2)u_{k_2}}\sum_{\substack{(r,a_ik_i)=1\\ \mathrm{cond}(\chi)|r}}\frac{1}{r\varphi(re)}\Phi\left(\frac{r}{R}\right)\sum_{1\leqslant|h|\leqslant H}\widehat{\Phi}\left(\frac{hM}{r}\right)\mathrm{e}\left(\frac{a_1h\overline{a_2k_2}}{r}\right)\\ &\quad + O_{\varepsilon}\left(x^{-90}\right),\end{aligned}$$

where we used that $\varphi(re)\gg\varphi(r)e$ as before (since e is prime). The error term is acceptable, so let us focus on the main term on the right-hand side (denote this by Y_2). By Lemma 3.1, we have

$$\frac{\overline{a_2k_2}}{r}+\frac{\overline{r}}{a_2k_2}\equiv\frac{1}{a_2k_2r}\pmod 1,$$

so that

$$Y_2\ll\sum_{e\in\mathscr{E}}\sum_{\chi\in\mathcal{X}_D}\sum_{k_1,k_2}|u_{k_1}u_{k_2}|\sum_{1\leqslant|h|\leqslant H}\left|\sum_{\substack{(r,a_ik_i)=1\\ \mathrm{cond}(\chi)|r}}\frac{1}{r\varphi(re)}\Phi\left(\frac{r}{R}\right)\widehat{\Phi}\left(\frac{hM}{r}\right)\mathrm{e}\left(\frac{a_1h}{a_2k_2r}-\frac{a_1h\overline{r}}{a_2k_2}\right)\right|. \tag{7.2}$$

At this point we decompose

$$\frac{1}{\varphi(re)}=\frac{1}{\varphi(r)}\left(\frac{1}{e-1}-\frac{1}{e(e-1)}\mathbb{1}_{e|r}\right),$$

aiming to apply the exponential sum bound in Lemma 3.9. Fixing e, a_i, k_i, h, this lets us rewrite the sum over r on the right-hand side of (7.2) as

$$\frac{1}{e-1}Z_2(\mathrm{cond}(\chi))-\frac{1}{e(e-1)}Z_2([\mathrm{cond}(\chi),e]),$$

where

$$Z_2(\ell):=\sum_{\substack{(r,a_ik_i)=1\\ \ell|r}}u(r)\frac{r}{\varphi(r)}\,\mathrm{e}\left(-\frac{a_1h\overline{r}}{a_2k_2}\right),$$

and

$$u(r):=\frac{1}{r^2}\Phi\left(\frac{r}{R}\right)\widehat{\Phi}\left(\frac{hM}{r}\right)\mathrm{e}\left(\frac{a_1h}{a_2k_2r}\right).$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Note that u extends to a differentiable function of a real variable ξ, supported in $[R/2, 3R]$, and with derivative bounds

$$|u'(\xi)| \ll \frac{1}{R^3} + \frac{1}{R^2}\left(\frac{1}{R} + \frac{HM}{R^2} + \frac{|a_1|H}{KR^2}\right)$$
$$\ll \frac{1}{R^2}\left(\frac{1}{R} + \frac{x^\varepsilon R}{R^2} + \frac{|a_1|x^\varepsilon R}{xR^2}\right) \ll \frac{x^{2\varepsilon}}{R^3},$$

in this region (we used that $H = x^\varepsilon RM^{-1}$, $MK \asymp x$, and the very crude bound $|a_1| \ll x^{1+\varepsilon}$). So we may use integration by parts to estimate Z_2; letting

$$v_\ell(\xi) := \sum_{\substack{(r, a_i k_i)=1 \\ \ell \mid r \\ r \leqslant \xi}} \frac{r}{\varphi(r)}\, \mathrm{e}\left(-\frac{a_1 h \overline{r}}{a_2 k_2}\right),$$

which can be bounded via Lemma 3.9 (with $n = a_1 h$, $c = a_2 k_2$, $m = r$, and $k = a_1 k_1$), we obtain

$$Z_2(\ell) = \int u(\xi) dv_\ell(\xi) = -\int v_\ell(\xi) du(\xi) \ll \frac{x^{2\varepsilon}}{R^2} \sup_{\xi \in [R/2, 3R]} |v_\ell(\xi)|$$
$$\ll_\varepsilon \frac{x^{3\varepsilon}}{R^2}\Big((a_1 h, a_2 k_2)^{1/2} K^{1/2} + (a_1 h, a_2 k_2)\frac{R}{K}\Big),$$

uniformly in $\ell \geqslant 1$. Returning to (7.2) and summing over h and k_2, the GCD terms contribute at most $O_\varepsilon(x^\varepsilon)$ on average (since $(a_1, a_2) = 1$). Thus

$$Y_2 \ll_\varepsilon x^{4\varepsilon} |\mathscr{E}||\chi_D| K^2 H \frac{1}{E}\frac{1}{R^2}\Big(\sqrt{K} + \frac{R}{K}\Big) \leqslant x^{7\varepsilon}\frac{K^2}{MR}\Big(\sqrt{K} + \frac{R}{K}\Big)$$
$$\ll x^{7\varepsilon - 1}\Big(\frac{K^{7/2}}{R} + K^2\Big).$$

By (4.2), we have $K^{3/2} \ll x^{1-9\varepsilon}$ and $R \ll x^{1-11\varepsilon}$, so we get a final bound of $Y_2 \ll_\varepsilon x^{-\varepsilon} K^2/R$. This completes our proof of Proposition 6.2(ii). □

8. The first dispersion sum

Finally, we work towards establishing Proposition 6.2(i) (for a suitable choice of δ in terms of ε); the first part of this section is very similar to [Dra15, § 3.4]. Recall from (6.2) that

$$\mathcal{S}_1 - \widehat{\Phi}(0) X_1 = \sum_{e \in \mathscr{E}} \sum_{(r, a_1 a_2)=1} \Phi\left(\frac{r}{R}\right) \sum_{c \in (\mathbf{Z}/r\mathbf{Z})^\times} \mathcal{E}_r(c) \sum_{\substack{(k_1 k_2, re)=1 \\ k_1 \equiv k_2 \pmod{re} \\ k_i \equiv a_1 \overline{a_2 c} \pmod{r}}} u_{k_1} \overline{u_{k_2}},$$

where $\mathcal{E}_r(c)$ is given by (6.1). We wish to bound this by $O_\varepsilon(x^{-2\delta} K^2/R)$, as in (6.6).

We still aim to apply Poisson summation for the sums $\mathcal{E}_r(c)$, and reduce our problem to bounding certain exponential sums. But due to issues that would arise later in manipulating these exponential sums, we first need to eliminate the contribution of certain 'bad' pairs (k_1, k_2), in terms of a small parameter η (to be chosen later in terms of ε, as an intermediary step to choosing δ).

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Proposition 8.1 (Eliminating bad index pairs). *For $\varepsilon \geqslant \eta > 0$, under the parameter conditions in (4.2) and (5.2), one has*

$$\mathcal{S}_1 - \widehat{\Phi}(0)X_1 = \sum_{e\in\mathscr{E}} \sum_{(r,a_1a_2)=1} \Phi\left(\frac{r}{R}\right) \sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times} \mathcal{E}_r(c) \sum_{\substack{(k_1,k_2)\in\mathcal{K}(\eta)\\(k_1k_2,re)=1\\k_1\equiv k_2 \pmod{re}\\k_i\equiv a_1\overline{a_2c} \pmod{r}}} u_{k_1}\overline{u_{k_2}} + O_\eta\left(x^{-\eta/4}\frac{K^2}{R}\right),$$

where

$$\mathcal{K}(\eta) := \left\{(k_1,k_2)\in\mathbf{N}^2 \;\middle|\; \begin{array}{c}(k_1,(a_2k_2)^\infty)\leqslant x^\eta,\ (k_2,(a_2k_1)^\infty)\leqslant x^\eta,\\ (k_1-k_2,(a_2k_1k_2)^\infty)\leqslant x^\eta,\ |k_1-k_2|>K/x^\eta\end{array}\right\}. \tag{8.1}$$

Proof. We eliminate the contribution of several sets of pairs (k_1,k_2), putting absolute values on all the coefficients involved; thus, it does not matter if some of the 'eliminated sets' have nonempty intersections. First, we consider the almost-diagonal pairs with $|k_1-k_2|\leqslant K/x^\eta$; using that $\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}|\mathcal{E}_r(c)|\leqslant (1/M)\sum_m \Phi(m/M)+\widehat{\Phi}(0)\ll 1$ and $x^\eta RE\ll K$ (which follows from (5.2) and $\eta\leqslant\varepsilon$), these contribute to $\mathcal{S}_1-\widehat{\Phi}(0)X_1$ at most

$$\begin{aligned}
&\ll \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}|\mathcal{E}_r(c)| \sum_{\substack{k_1\equiv a_1\overline{a_2c} \pmod r\\(k_1,e)=1}}|u_{k_1}| \sum_{\substack{k_2\equiv k_1 \pmod{re}\\|k_1-k_2|\leqslant K/x^\eta}}|u_{k_2}|\\
&\ll_\eta x^{\eta/2}ER\Big(\frac{K}{R}+1\Big)\Big(\frac{K}{x^\eta RE}+1\Big)\\
&\ll_\eta x^{\eta/2}ER\frac{K}{R}\frac{K}{x^\eta RE}\\
&\asymp x^{-\eta/2}\frac{K^2}{R}.
\end{aligned}$$

Then, we consider those pairs with $v:=(k_1,k_2)>x^{\eta/2}$. Their contribution to $\mathcal{S}_1-\widehat{\Phi}(0)X_1$ is at most

$$\ll \sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}|\mathcal{E}_r(c)| \sum_{\substack{b\in(\mathbf{Z}/re\mathbf{Z})^\times\\ b\equiv a_1\overline{a_2c} \pmod r}} \sum_{\substack{v>x^{\eta/2}\\(v,re)=1}} \Bigg(\sum_{\substack{k\equiv b \pmod{re}\\ v|k}}|u_k|\Bigg)^2.$$

Using that $(v,re)=1$, we can bound one inner sum over k by $\ll_\eta x^{\eta/8}(K(vRE)^{-1}+1)\ll x^{-3\eta/8}K(RE)^{-1}$ (recall that $x^\eta RE\ll K$ by (5.2)). This yields a total contribution of

$$\begin{aligned}
&\ll_\eta \frac{K}{REx^{3\eta/8}}\sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}|\mathcal{E}_r(c)| \sum_{\substack{v>x^\eta\\(v,re)=1}}\sum_{\substack{k\equiv a_1\overline{a_2c} \pmod r\\ v|k}}|u_k|\\
&\leqslant \frac{K}{Rx^{3\eta/8}}\sum_{(r,a_1a_2)=1}\Phi\left(\frac{r}{R}\right)\sum_{c\in(\mathbf{Z}/r\mathbf{Z})^\times}|\mathcal{E}_r(c)| \sum_{k\equiv a_1\overline{a_2c} \pmod r}\tau(k)|u_k|\\
&\ll_\eta \frac{K}{Rx^{2\eta/8}}R\Big(\frac{K}{R}+1\Big)\\
&\ll x^{-\eta/4}\frac{K^2}{R},
\end{aligned}$$

which is also acceptable. Keeping the notation $v=(k_1,k_2)$, which we may now assume is at most $x^{\eta/2}$, note that

$$d_1 := (k_1,(a_2k_2)^\infty) = (k_1,(a_2v)^\infty),$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

and let us consider those pairs (k_1, k_2) with $d_1 > x^\eta$. Using that $x^{\eta/2}RE \ll K$ and swapping sums, these contribute at most

$$\sum_{v\leqslant x^{\eta/2}}\sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\Big(\frac{r}{R}\Big)\sum_{(m,r)=1}\frac{1}{M}\Phi\Big(\frac{m}{M}\Big)\sum_{\substack{d_1\mid(a_2v)^\infty\\ d_1>x^\eta}}\sum_{\substack{k_1\equiv a_1\overline{a_2m}\pmod r\\ (k_1,re)=1,\ d_1\mid k_1}}|u_{k_1}|\sum_{\substack{k_2\equiv k_1\pmod{re}\\ (k_1,k_2)=v}}|u_{k_2}|$$

$$\ll_\eta x^{\eta/8}\sum_{v\leqslant x^{\eta/2}}\sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\Big(\frac{r}{R}\Big)\sum_{(m,r)=1}\frac{1}{M}\Phi\Big(\frac{m}{M}\Big)\sum_{\substack{d_1\mid(a_2v)^\infty\\ d_1>x^\eta}}\sum_{\substack{k_1\in(K,4K]\\ (k_1,re)=1,\ d_1\mid k_1\\ r\mid a_2mk_1-a_1}}\Big(\frac{K}{REv}+1\Big)$$

$$\ll x^{\eta/8}\frac{K}{RE}\sum_{v\leqslant x^{\eta/2}}\frac{1}{v}\sum_{e\in\mathscr{E}}\sum_m\frac{1}{M}\Phi\Big(\frac{m}{M}\Big)\sum_{\substack{d_1\mid(a_2v)^\infty\\ d_1>x^\eta}}\sum_{\substack{k_1\in(K,4K]\\ d_1\mid k_1}}\sum_{\substack{(r,ma_1a_2)=1\\ r\mid a_2mk_1-a_1}}\Phi\Big(\frac{r}{R}\Big).$$

Considering the cases $a_2mk_1 = a_1$ and $a_2mk_1 - a_1 \neq 0$ separately, and using $R \ll x^{1-\eta/2} \asymp KMx^{-\eta/2}$ (by (4.2)), this is further bounded by

$$\ll x^{\eta/8}\frac{K}{RE}\sum_{v\leqslant x^{\eta/2}}\frac{1}{v}\sum_{e\in\mathscr{E}}\bigg(\tau(|a_1|)^3\frac{R}{M}+\sum_m\frac{1}{M}\Phi\Big(\frac{m}{M}\Big)\sum_{\substack{d_1\mid(a_2v)^\infty\\ x^\eta<d_1\leqslant 4K}}\sum_{\substack{k_1\in(K,4K]\\ d_1\mid k_1\\ a_2mk_1\neq a_1}}\tau(|a_2mk_1-a_1|)\bigg)$$

$$\ll_\eta x^{\eta/4}\frac{K}{M}+x^{\eta/4}\frac{K}{RE}\max_{v\leqslant x^{\eta/2}}\sum_{e\in\mathscr{E}}\sum_m\frac{1}{M}\Phi\Big(\frac{m}{M}\Big)\sum_{\substack{d_1\mid(a_2v)^\infty\\ x^\eta<d_1\leqslant 4K}}\Big(\frac{K}{d_1}+1\Big)$$

$$\ll x^{-\eta/4}\frac{K^2}{R}+x^{\eta/4}\frac{K}{R}\max_{v\leqslant x^{\eta/2}}\sum_{\substack{d_1\mid(a_2v)^\infty\\ x^\eta<d}}\frac{K}{d_1}.$$

Now since the number of distinct prime factors of a positive integer b is $O(\log b/\log\log b)$, for $b \ll x$ we have the majorization

$$\sum_{\substack{d\mid b^\infty\\ x^\eta<d}}d^{-1}\leqslant x^{-2\eta/3}\sum_{d\mid b^\infty}d^{-1/3}$$

$$=x^{-2\eta/3}\prod_{\text{prime }p\mid b}\frac{1}{1-p^{-1/3}}\ll x^{-2\eta/3}O(1)^{O(\log b/\log\log b)}\ll_\eta x^{-\eta/2}. \tag{8.2}$$

Using this, we find that the previous sum contributes an acceptable $O_\eta(x^{-\eta/4}K^2/R)$.

The contribution of the pairs with $d_1 > x^\eta$ to $\widehat{\Phi}(0)X_1$ is simpler and similarly bounded, and the contribution of the pairs with $d_2 := (k_2, (a_2k_1)^\infty) = (k_2, (a_2v)^\infty)$ is bounded symmetrically. All that is left is to eliminate the contribution of the pairs with large values of $(k_1 - k_2, (a_2k_1k_2)^\infty)$; since $(k_1 - k_2, k_1k_2) = (k_1 - k_2, v^2)$, we have

$$d_\Delta := (k_1 - k_2, (a_2k_1k_2)^\infty) = (k_1 - k_2, (a_2v)^\infty).$$

Using that $Rx^{\eta/2} \ll K$, the pairs with $d_\Delta > x^\eta$ (as well as $v \leqslant x^{\eta/2}$ and $|k_1 - k_2| > K/x^\eta$) contribute to $\mathcal{S}_1$ at most

$$\ll\sum_{v\leqslant x^{\eta/2}}\sum_{e\in\mathscr{E}}\sum_{(r,a_1a_2)=1}\Phi\Big(\frac{r}{R}\Big)\sum_{(m,r)=1}\frac{1}{M}\Phi\Big(\frac{m}{M}\Big)\sum_{\substack{d_\Delta\mid(a_2v)^\infty\\ d_\Delta>x^\eta}}\sum_{\substack{k_1\equiv a_1\overline{a_2m}\pmod r\\ (k_1,re)=1,\ (k_1,k)=v\\ 0<|k|\leqslant 8K\\ [d_\Delta,re]\mid k}}|u_{k_1}u_{(k_1+k)}|$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

$$\ll_\eta x^{\eta/8} \sum_{v\leqslant x^{\eta/2}} \sum_{e\in\mathscr{E}} \sum_{(r,a_1a_2)=1} \Phi\left(\frac{r}{R}\right) \sum_{(m,r)=1} \frac{1}{M}\Phi\left(\frac{m}{M}\right) \sum_{\substack{d_\Delta \mid (a_2v)^\infty \\ d_\Delta > x^\eta}} \sum_{\substack{0<|k|\leqslant 8K \\ [d_\Delta, re]\mid k}} \sum_{\substack{k_1\in(k,4K] \\ k_1\equiv a_1\overline{a_2m} \pmod r \\ (k_1,re)=1,\ (k_1,k)=v}} 1$$

$$\ll x^{\eta/8} \sum_{v\leqslant x^{\eta/2}} \sum_{e\in\mathscr{E}} \sum_{(r,a_1a_2)=1} \Phi\left(\frac{r}{R}\right) \sum_{(m,r)=1} \frac{1}{M}\Phi\left(\frac{m}{M}\right) \sum_{\substack{d_\Delta \mid (a_2v)^\infty \\ d_\Delta > x^\eta}} \sum_{\substack{0<|k|\leqslant 8K \\ [d_\Delta, re]\mid k}} \left(\frac{K}{Rv}+1\right)$$

$$\ll x^{\eta/8}\frac{K}{R} \sum_{v\leqslant x^{\eta/2}} \frac{1}{v} \sum_{m} \frac{1}{M}\Phi\left(\frac{m}{M}\right) \sum_{\substack{d_\Delta \mid (a_2v)^\infty \\ d_\Delta > x^\eta}} \sum_{\substack{0<|k|\leqslant 8K \\ d_\Delta \mid k}} \sum_{\substack{e\in\mathscr{E} \\ (r,a_1a_2)=1 \\ re\mid k}} \Phi\left(\frac{r}{R}\right).$$

Bounding the inner sum by $\tau(k)^2 \ll_\eta x^{\eta/9}$, this further becomes

$$\ll_\eta x^{\eta/4}\frac{K}{R} \max_{v\leqslant x^{\eta/2}} \sum_{m} \frac{1}{M}\Phi\left(\frac{m}{M}\right) \sum_{\substack{d_\Delta \mid (a_2v)^\infty \\ x^\eta < d_\Delta \leqslant 8K}} \sum_{\substack{|k|\leqslant 8K \\ d_\Delta \mid k}} 1$$

$$\ll x^{\eta/4}\frac{K}{R} \max_{v\leqslant x^{\eta/2}} \sum_{\substack{d_\Delta \mid (a_2v)^\infty \\ x^\eta < d_\Delta \leqslant 8K}} \left(\frac{K}{d_\Delta}+1\right)$$

$$\ll x^{\eta/4}\frac{K}{R} \max_{v\leqslant x^{\eta/2}} \sum_{\substack{d_\Delta \mid (a_2v)^\infty \\ x^\eta < d_\Delta \leqslant 8K}} \frac{K}{d_\Delta}.$$

Using the majorization from (8.2), this is again $O_\eta(x^{-\eta/4}K^2/R)$.

The contribution of the pairs with $d_\Delta > x^\eta$ to $\widehat{\Phi}(0)X_1$ is simpler and similarly bounded by $O_\eta(x^{-\eta/4}K^2/R)$. Having eliminated the absolute contribution of all pairs in $\mathcal{K}(\eta)$ at least once, while incurring only admissible errors, we can conclude our proof of Proposition 8.1. □

We can now apply Poisson summation to prove the following.

Proposition 8.2 (Reduction to exponential sum). *For $\varepsilon \geqslant \eta > 0$ and $H := x^\eta RM^{-1}$, under the parameter conditions in (4.2) and (5.2), one has*

$$\mathcal{S}_1 - \widehat{\Phi}(0)X_1 = \sum_{e\in\mathscr{E}} \sum_{(r,a_1a_2)=1} \Phi\left(\frac{r}{R}\right) \frac{1}{r} \sum_{\substack{(k_1,k_2)\in\mathcal{K}(\eta) \\ (k_1k_2,re)=1 \\ k_1\equiv k_2 \pmod{re}}} u_{k_1}\overline{u_{k_2}} \sum_{1\leqslant|h|\leqslant H} \widehat{\Phi}\Big(\frac{hM}{r}\Big)\mathrm{e}\Big(a_1h\frac{\overline{a_2k_1}}{r}\Big)$$

$$+ O_\eta\Big(x^{-\eta/4}\frac{K^2}{R}\Big).$$

Proof. Rewrite the sum in Proposition 8.1 as

$$\sum_{e\in\mathscr{E}} \sum_{(r,a_1a_2)=1} \Phi\left(\frac{r}{R}\right) \sum_{\substack{(k_1,k_2)\in\mathcal{K}(\eta) \\ (k_1k_2,re)=1 \\ k_1\equiv k_2 \pmod{re}}} u_{k_1}\overline{u_{k_2}}\, \mathcal{E}_r(a_1\overline{a_2k_1}),$$

and apply Lemma 3.5 to expand $\mathcal{E}_r(a_1\overline{a_2k_1})$. The resulting main term is precisely the sum in Proposition 8.2, while the error terms are acceptable. □

Remark 8.3. The trivial bound for the right-hand side of Proposition 8.2 is H times worse than for the right-hand side of Proposition 8.1, due to the additional sum over h. This is relevant because H is a nontrivial power of x for the choice of parameters in (4.4) (where $H \approx RM^{-1} \approx R^2/x$),

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

since we are working with moduli r well beyond the $\sqrt{x}$ barrier. This is why we needed to eliminate the bad pairs (k_1, k_2) (via Proposition 8.1) *before* applying Poisson summation.

We now go through a series of fairly technical manipulations, following [Dra15] and [May25a], to reduce the sum in Proposition 8.2 to a variation of the exponential sum considered by Bombieri, Friedlander, and Iwaniec in [BFI87, § 10]; the goal is to prove the remaining dispersion estimate for $\mathcal{S}_1$ in Proposition 6.2. We do this in two steps (after the statements of Propositions 8.4 and 8.6); first, we assume the following exponential sum bound, which can be compared with Drappeau's [Dra15, Proposition 1].

PROPOSITION 8.4 (Improved Drappeau-style exponential sum bound). *For all sufficiently small $\varepsilon > 0$ and all $\eta \in (0,1)$, under the parameter conditions in (4.2), there exists E satisfying (5.2) (with $K := NL$) such that the following holds. For any nonzero integer $a \ll x^{O(\eta)}$ and positive integers $b, d_a, d_1, d_2, d_\Delta, v, \delta_1, \delta_2 \ll x^{O(\eta)}$ with $d_2 = \delta_1\delta_2$ and*

$$d_a \mid a, \quad (d_1, d_2) = (d_1, d_\Delta) = (d_2, d_\Delta) = v, \quad d_1 d_2 d_\Delta \mid b^\infty,$$

one has

$$\sum_{\substack{e \sim E \text{ prime} \\ (e,b)=1}} \sum_{\substack{(r,b)=1 \\ r \equiv 0 \pmod{d_a}}} \Phi\left(\frac{r}{R}\right) \frac{1}{r} \sum_{\substack{(k,n\ell)=1 \\ (kn\ell, reb)=1 \\ d_1 k - d_2 n\ell = red_\Delta t \\ (t,b)=1,\ red_\Delta |t| > K/x^\eta}} u'_k \beta'_n \gamma'_\ell \sum_{1 \leqslant |h| \leqslant H} \widehat{\Phi}\left(\frac{hM}{r}\right) \mathrm{e}\left(ah\frac{\overline{bk}}{r}\right)$$

$$\ll_{\varepsilon,\eta} x^{O(\eta)-\varepsilon/4} \frac{K^2}{R},$$

where $H := x^\eta RM^{-1}$, and $|u'_k| \leqslant \tau(d_1 k)$, $|\beta'_n| \leqslant 1$, $|\gamma'_\ell| \leqslant 1$ are sequences supported in $k \asymp K/d_1$, $n \sim N/\delta_1$, and $\ell \sim L/\delta_2$.

Remark 8.5. The exponential sum from Proposition 8.4 is essentially that anticipated in (2.3).

Proof of Proposition 6.2(i), assuming Proposition 8.4. Let $\varepsilon \in (0,1)$ be sufficiently small, and let us pick E as in Proposition 8.4. By Proposition 8.2, it remains to establish the bound

$$\mathcal{D}_1 := \sum_{e \in \mathscr{E}} \sum_{(r, a_1 a_2) = 1} \Phi\left(\frac{r}{R}\right) \frac{1}{r} \sum_{\substack{(k_1,k_2) \in \mathcal{K}(\eta) \\ (k_1 k_2, re) = 1 \\ k_1 \equiv k_2 \pmod{re}}} u_{k_1} \overline{u_{k_2}} \sum_{1 \leqslant |h| \leqslant H} \widehat{\Phi}\left(\frac{hM}{r}\right) \mathrm{e}\left(a_1 h \frac{\overline{a_2 k_1}}{r}\right)$$

$$\ll_{\varepsilon,\eta} x^{-2\delta} \frac{K^2}{R},$$

for some choice of $0 < 8\delta \leqslant \eta \leqslant \varepsilon$ in terms of ε (since $\delta \leqslant \eta/8$, the error term of $x^{-\eta/4} K^2/R$ from Proposition 8.2 is acceptable). For now, let us fix δ and η such that $8\delta \leqslant \eta$; we will give explicit choices at the end of this proof.

By the definition of $\mathcal{K}(\eta)$ from (8.1), we may consider the x^η-bounded variables

$$\begin{gathered} d_1 := (k_1, (a_2 k_2)^\infty), \quad d_2 := (k_2, (a_2 k_1)^\infty), \quad d_\Delta := (k_1 - k_2, (a_2 k_1 k_2)^\infty), \\ v := (k_1, k_2) = (d_1, d_2) = (d_1, d_\Delta) = (d_2, d_\Delta). \end{gathered} \tag{8.3}$$

Noting that d_1, d_2, and d_Δ all divide $(a_2 v)^\infty$, we may then expand

$$\mathcal{D}_1 = \sum_{\substack{d_1, d_2, d_\Delta \leqslant x^\eta \\ v = (d_1,d_2) = (d_1,d_\Delta) = (d_2,d_\Delta) \\ d_1 d_2 d_\Delta \mid (a_2 v)^\infty}} \mathcal{D}_2(d_1, d_2, d_\Delta, v), \tag{8.4}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

where

$$\mathcal{D}_2 := \sum_{e\in\mathscr{E}} \sum_{(r,a_1a_2)=1} \Phi\left(\frac{r}{R}\right)\frac{1}{r} \sum_{\substack{(k_1,(a_2k_2)^\infty)=d_1\\(k_2,(a_2k_1)^\infty)=d_2\\(k_1-k_2,(a_2k_1k_2)^\infty)=d_\Delta\\|k_1-k_2|>K/x^\eta\\(k_1k_2,re)=1\\k_1\equiv k_2 \pmod{re}}} u_{k_1}\overline{u_{k_2}} \sum_{1\leqslant|h|\leqslant H} \widehat{\Phi}\left(\frac{hM}{r}\right) \mathrm{e}\left(a_1h\frac{\overline{a_2k_1}}{r}\right).$$

Changing variables $k_i \mapsto d_ik_i$ and adjusting coprimality conditions accordingly (e.g., we now have $(k_1, a_2d_2k_2) = 1$ as well as $(k_1, d_1) = 1$, and $(d_1k_1d_2k_2, re) = 1$), we get

$$\begin{aligned}\mathcal{D}_2 &= \sum_{e\in\mathscr{E}} \sum_{(r,a_1a_2)=1} \Phi\Big(\frac{r}{R}\Big)\frac{1}{r} \sum_{\substack{(k_1,a_2d_1d_2k_2)=1\\(k_2,a_2d_1d_2k_1)=1\\(d_1k_1-d_2k_2,(a_2d_1d_2k_1k_2)^\infty)=d_\Delta\\|d_1k_1-d_2k_2|>K/x^\eta\\(d_1k_1d_2k_2,re)=1\\d_1k_1\equiv d_2k_2 \pmod{re}}} u_{d_1k_1}\overline{u_{d_2k_2}} \sum_{1\leqslant|h|\leqslant H} \widehat{\Phi}\Big(\frac{hM}{r}\Big) \mathrm{e}\Big(a_1h\frac{\overline{a_2d_1k_1}}{r}\Big)\\ &= \sum_{\substack{e\in\mathscr{E}\\(r,a_1a_2)=1\\(re,d_1d_2)=1}} \Phi\left(\frac{r}{R}\right)\frac{1}{r} \sum_{\substack{(k_1,k_2)=1\\(k_1k_2,rea_2d_1d_2)=1\\(d_1k_1-d_2k_2,(a_2d_1d_2)^\infty)=d_\Delta\\|d_1k_1-d_2k_2|>K/x^\eta\\d_1k_1\equiv d_2k_2 \pmod{re}}} u_{d_1k_1}\overline{u_{d_2k_2}} \sum_{1\leqslant|h|\leqslant H} \widehat{\Phi}\left(\frac{hM}{r}\right) \mathrm{e}\left(a_1h\frac{\overline{a_2d_1k_1}}{r}\right).\end{aligned}$$

Let us denote

$$a := a_1\mathrm{sgn}(a_2) \quad \text{and} \quad b := |a_2|d_1$$

for convenience. At this point we record that, since $d_1d_2d_\Delta \mid (a_2v)^\infty$, $v \mid d_1$, and $a_1, a_2 \ll x^\delta \leqslant x^\eta$ by (4.2), we have

$$a_2d_1d_2d_\Delta \mid b^\infty \quad \text{and} \quad abd_1d_2d_\Delta v \ll x^{O(\eta)},$$

as needed in Proposition 8.4 (in particular, b will act as a bookkeeper for the prime factors of d_1, d_2, d_Δ, a_2 inside coprimality constraints). Recalling that we chose $\mathcal{E} = \{e \sim E : e \text{ prime}\}$ in (5.3), we can ensure that $(e, a_2d_1) = (e, b) = 1$ for $e \in \mathcal{E}$ by enforcing $\delta < 4\varepsilon$ (since then $|a_2| \leqslant x^\delta \leqslant x^{4\varepsilon} \leqslant E$). Writing $d_1k_1 - d_2k_2 = red_\Delta t$, where $(t, a_2d_1d_2k_1k_2) = 1$ (which is further absorbed by the conditions $(t, b) = (k_1, k_2) = (k_1k_2, b) = 1$), we further get

$$\mathcal{D}_2 = \sum_{\substack{e\sim E \text{ prime}\\(e,b)=1}} \sum_{(r,ab)=1} \Phi\left(\frac{r}{R}\right)\frac{1}{r} \sum_{\substack{(k_1,k_2)=1\\(k_1k_2,reb)=1\\d_1k_1-d_2k_2=red_\Delta t\\(t,b)=1,\ red_\Delta|t|>K/x^\eta}} u_{d_1k_1}\overline{u_{d_2k_2}} \sum_{1\leqslant|h|\leqslant H} \widehat{\Phi}\left(\frac{hM}{r}\right) \mathrm{e}\left(ah\frac{\overline{bk_1}}{r}\right).$$

We can also get rid of the restriction $(r, a) = 1$ using Möbius inversion, by writing $\mathbb{1}_{(r,a)=1} = \sum_{d_a|a} \mu(d_a)\mathbb{1}_{d_a|r}$ and expanding

$$\mathcal{D}_2(d_1, d_2, d_\Delta, v) = \sum_{d_a|a} \mu(d_a)\, \mathcal{D}_3(d_1, d_2, d_\Delta, v, d_a), \tag{8.5}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

where

$$\mathcal{D}_3 := \sum_{\substack{e\sim E \text{ prime}\\(e,b)=1}} \sum_{\substack{(r,b)=1\\ r\equiv 0 \pmod{d_a}}} \Phi\left(\frac{r}{R}\right)\frac{1}{r} \sum_{\substack{(k_1,k_2)=1\\(k_1k_2,reb)=1\\ d_1k_1-d_2k_2=red_\Delta t\\ (t,b)=1,\ red_\Delta|t|>K/x^\eta}} u_{d_1k_1}\overline{u_{d_2k_2}} \sum_{1\leqslant|h|\leqslant H} \widehat{\Phi}\left(\frac{hM}{r}\right)\mathrm{e}\left(ah\frac{\overline{bk_1}}{r}\right).$$

Finally, using the definition of (u_k) from (5.1) and the fact that $(d_2,k_2)\,|\,(b,k_2)=1$, we can expand

$$u_{d_2k_2} = \sum_{\delta_1\delta_2=d_2}\sum_{n\ell=k_2} \beta_{\delta_1 n}\gamma_{\delta_2\ell},$$

and thus

$$\mathcal{D}_3(d_1,d_2,d_\Delta,v,d_a) = \sum_{\delta_1\delta_2=d_2} \mathcal{D}_4(d_1,d_2,d_\Delta,v,d_a,\delta_1), \tag{8.6}$$

where

$$\mathcal{D}_4 := \sum_{\substack{e\sim E \text{ prime}\\(e,b)=1}} \sum_{\substack{(r,b)=1\\ r\equiv 0 \pmod{d_a}}} \Phi\left(\frac{r}{R}\right)\frac{1}{r} \sum_{\substack{(k,n\ell)=1\\(kn\ell,reb)=1\\ d_1k-d_2n\ell=red_\Delta t\\ (t,b)=1,\ red_\Delta|t|>K/x^\eta}} u_{d_1k}\overline{\beta_{\delta_1 n}\gamma_{\delta_2\ell}} \sum_{1\leqslant|h|\leqslant H} \widehat{\Phi}\left(\frac{hM}{r}\right)\mathrm{e}\left(ah\frac{\overline{bk}}{r}\right).$$

We can now apply Proposition 8.4 for the sequences

$$u'_k := u_{d_1k}, \quad \beta'_n := \overline{\beta_{\delta_1 n}}, \quad \gamma'_\ell := \overline{\gamma_{\delta_2\ell}},$$

supported in $k\asymp K/d_1$, $n\sim N/\delta_1$, and $\ell\sim L/\delta_2$, respectively, to get

$$\mathcal{D}_4 \ll_{\eta,\varepsilon} x^{O(\eta)-\varepsilon/4}\frac{K^2}{R}. \tag{8.7}$$

Putting together (8.4) to (8.7), we obtain

$$\mathcal{D}_1 \ll_\eta x^{O(\eta)} \max_{d_1,d_2,d_\Delta,v,d_a,\delta_1} |\mathcal{D}_4| \ll_{\varepsilon,\eta} x^{O(\eta)-\varepsilon/4}\frac{K^2}{R}, \tag{8.8}$$

where the maximum includes all applicable restrictions on the tuple $(d_1,d_2,d_\Delta,v,d_a,\delta_1)$ (which takes at most $O_\eta(x^{O(\eta)})$ values).

Now let $C>0$ denote the absolute constant in the final exponent of $O(\eta)$ from (8.8), and let us pick

$$\eta=\eta(\varepsilon) := \min(\varepsilon/(8C),\varepsilon), \quad \delta=\delta(\varepsilon) := \min(\varepsilon/16,\eta/8).$$

Then we have $0<8\delta\leqslant\eta\leqslant\varepsilon$ as desired, and the bound in (8.8) implies

$$\mathcal{D}_1 \ll_{\varepsilon,\eta} x^{\varepsilon/8-\varepsilon/4}\frac{K^2}{R} \leqslant x^{-2\delta}\frac{K^2}{R},$$

completing our proof. □

Finally, we prove Proposition 8.4 assuming the following BFI-style bound, the proof of which is left to the later sections. This should be compared with Maynard's [May25a, Lemma 18.5].

PROPOSITION 8.6 (Improved BFI/Maynard exponential sum bound). *For all sufficiently small $\varepsilon>0$ and all $\eta\in(0,1)$, the following holds. Under the conditions in (4.2), there exists E satisfying (5.2), such that for any positive integers b,d with $b\ll x^{O(\eta)}$ and $d\ll x^{O(1)}$, and for any parameters*

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

$K' \ll NLx^{O(\eta)}$, $N' \asymp Nx^{O(\eta)}$, $L' \asymp Lx^{O(\eta)}$, $T' \ll NL(RE)^{-1}x^{O(\eta)}$, and $H' \ll RM^{-1}x^{O(\eta)}$, one has

$$\sum_{\substack{k\sim K'\\(k,b)=1}}\sum_{\substack{n\sim N'\\(n,k)=1}}\sum_{\substack{t\sim T'\\(t,dk)=1}} t\left|\sum_{\substack{e\sim E\\(e,dk)=1}}\sum_{h\sim H'}\sum_{\substack{\ell\sim L'\\(\ell,ket)=1\\ n\ell\equiv dk \pmod{et}}}\beta(e,h,\ell)\,\mathrm{e}\left(het\frac{\overline{bn\ell}}{k}\right)\right|^2$$
$$\ll_{\varepsilon,\eta} x^{O(\eta)-\varepsilon/2}N^2L^3, \tag{8.9}$$

for any 1-bounded complex coefficients $\beta(e,h,\ell)$ (independent of k,n,t).

Remark 8.7. The exponential sum from Proposition 8.6 is essentially anticipated in (2.4) and (2.5).

Proof of Proposition 8.4, assuming Proposition 8.6. Let us denote the sum in Proposition 8.4 by $\mathcal{D}_4$, and assume without loss of generality that $(d_a,b)=1$ (since otherwise $\mathcal{D}_4$ vanishes). We choose E as in Proposition 8.6, and take a closer look at the exponential: by iterating Lemma 3.1, since b,k,r are pairwise coprime we have

$$\frac{\overline{bk}}{r}+\frac{\overline{rb}}{k}+\frac{\overline{rk}}{b}\equiv\frac{1}{bkr}\pmod 1,$$

and thus, using that $red_\Delta t\equiv -d_2n\ell \pmod k$ and $d_1d_2d_\Delta \mid b^\infty$ (so in particular $(d_2,k)=1$),

$$ah\frac{\overline{bk}}{r}\equiv ahed_\Delta t\frac{\overline{bd_2n\ell}}{k}-ah\frac{\overline{rk}}{b}+\frac{ah}{bkr}\pmod 1.$$

Since $|a|\ll x^{O(\eta)}$, $h\ll x^\eta RM^{-1}$, $kr\gg KR/x^{O(\eta)}$ and $KM\asymp x$, we obtain

$$\mathrm{e}\left(ah\frac{\overline{bk}}{r}\right)=\mathrm{e}\left(ahed_\Delta t\frac{\overline{bd_2n\ell}}{k}-ah\frac{\overline{rk}}{b}\right)+O(x^{O(\eta)-1}),$$

and thus

$$\mathcal{D}_4=\sum_{\substack{e\sim E\text{ prime}\\(e,b)=1}}\sum_{\substack{(r,b)=1\\ r\equiv 0 \pmod{d_a}}}\Phi\left(\frac{r}{R}\right)\frac{1}{r}\sum_{\substack{(k,n\ell)=1\\(kn\ell,reb)=1\\ d_1k-d_2n\ell=red_\Delta t\\ (t,b)=1,\ red_\Delta|t|>K/x^\eta}}u'_k\beta'_n\gamma'_\ell$$
$$\times\sum_{1\leqslant|h|\leqslant H}\widehat{\Phi}\left(\frac{hM}{r}\right)\mathrm{e}\left(ahed_\Delta t\frac{\overline{bd_2n\ell}}{k}-ah\frac{\overline{rk}}{b}\right)+O_\eta\left(E\frac{R}{R}NL\left(\frac{K}{RE}+1\right)Hx^{O(\eta)-1}\right).$$

Recalling that $K\gg RE$ (due to (5.2)), $H=x^\eta RM^{-1}$, $MK\asymp x$, and $KR\ll x^{2-\varepsilon}$ (again by (4.2)), the error term gives an acceptable contribution of

$$\ll_\eta\frac{EK^2Hx^{O(\eta)-1}}{RE}\ll K^3x^{O(\eta)-2}\ll x^{O(\eta)-\varepsilon}\frac{K^2}{R}.$$

We now change variables in the main term by replacing the r-summation with a summation over

$$t:=\frac{d_1k-d_2n\ell}{red_\Delta},$$

noting that the condition $red_\Delta|t|>K/x^\eta$ implies $|t|>K/(REx^{O(\eta)})$. We also put $|t|$ into dyadic intervals $|t|\sim T$ to obtain

$$\mathcal{D}_4\ll(\log x)\sup_{T\asymp x^{O(\eta)}K/(RE)}|\mathcal{D}_5(T)|+O_\eta\left(x^{O(\eta)-\varepsilon}\frac{K^2}{R}\right), \tag{8.10}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

where, after adjusting coprimality conditions as explained below,

$$\mathcal{D}_5 := \sum_{\substack{e\sim E \text{ prime}\\(e,b)=1}} \sum_{\substack{|t|\sim T\\(t,b)=1}} \sum_{\substack{(k,n\ell)=(kn\ell,d_a ebt)=1\\ d_1k\equiv d_2n\ell \pmod{d_a ed_\Delta t}\\ |d_1k-d_2n\ell|>K/x^\eta\\ r:=(d_1k-d_2n\ell)/(ed_\Delta t)\\ (r,b)=1}} \Phi\left(\frac{r}{R}\right)\frac{1}{r}u'_k\beta'_n\gamma'_\ell$$
$$\times \sum_{1\leqslant|h|\leqslant H} \widehat{\Phi}\left(\frac{hM}{r}\right) \mathrm{e}\left(ahed_\Delta t\frac{\overline{bd_2n\ell}}{k} - ah\frac{\overline{rk}}{b}\right).$$

(We inserted the condition $(kn\ell,t)=1$; this must happen since $t\,|\,d_1k-d_2n\ell$ and $(t,d_1d_2)\,|\,(t,b^\infty)=1$; if a prime divides both t and one of k and $n\ell$, then it must also divide the other, contradicting $(k,n\ell)=1$. Moreover, the conditions in the sum over k,n,ℓ are enough to imply $(kn\ell,r)=1$, since $(d_1k-d_2n\ell,k)=(d_2n\ell,k)=(d_2,k)\,|\,(b^\infty,k)=1$ and similarly $(d_1k-d_2n\ell,n\ell)=1$.)

We now aim to simplify the term $ah\overline{rk}/b$ from the exponential, by fixing all relevant residues modulo b. With this goal, we denote the residues of e,t modulo b by $\widehat{e},\widehat{t}$, and those of k,n,ℓ modulo $d_\Delta b$ by $\widehat{k},\widehat{n},\widehat{\ell}$. Since $(d_1k-d_2n\ell)/d_\Delta=ret$ is coprime with b, we must have $d_1\widehat{k}-d_2\widehat{n}\widehat{\ell}\in d_\Delta(\mathbf{Z}/b\mathbf{Z})^\times = \{d_\Delta(n+b\mathbf{Z}):(n,b)=1\}\subseteq \mathbf{Z}/d_\Delta b\mathbf{Z}$. This allows us to expand $\mathcal{D}_5$ as

$$\mathcal{D}_5(T) = \sum_{\widehat{e},\widehat{t}\in(\mathbf{Z}/b\mathbf{Z})^\times} \sum_{\substack{\widehat{k},\widehat{n},\widehat{\ell}\in\mathbf{Z}/d_\Delta b\mathbf{Z}\\(\widehat{k}\widehat{n}\widehat{\ell},b)=1\\ d_1\widehat{k}-d_2\widehat{n}\widehat{\ell}\in d_\Delta(\mathbf{Z}/b\mathbf{Z})^\times}} \mathcal{D}_6(T,\widehat{e},\widehat{t},\widehat{k},\widehat{n},\widehat{\ell}), \tag{8.11}$$

with

$$\mathcal{D}_6 := \sum_{\substack{e\sim E \text{ prime}\\ e\equiv\widehat{e} \pmod b}} \sum_{\substack{|t|\sim T\\ t\equiv\widehat{t} \pmod b}} \sum_{\substack{(k,n,\ell)\equiv(\widehat{k},\widehat{n},\widehat{\ell}) \pmod{d_\Delta b}\\ (k,n\ell)=(kn\ell,d_aet)=1\\ d_1k\equiv d_2n\ell \pmod{d_aed_\Delta t}\\ |d_1k-d_2n\ell|>K/x^\eta\\ r:=(d_1k-d_2n\ell)/(ed_\Delta t)}} \Phi\left(\frac{r}{R}\right)\frac{1}{r}u'_k\beta'_n\gamma'_\ell$$
$$\times \sum_{1\leqslant|h|\leqslant H} \widehat{\Phi}\left(\frac{hM}{r}\right) \mathrm{e}\left(ahed_\Delta t\frac{\overline{bd_2n\ell}}{k} - ah\frac{\overline{\widehat{r}\widehat{k}}}{b}\right),$$

where $\widehat{r}=\widehat{r}(\widehat{e},\widehat{t},\widehat{k},\widehat{n},\widehat{\ell})\in(\mathbf{Z}/b\mathbf{Z})^\times$ is the unique residue mod b such that $d_\Delta\widehat{r}\widehat{e}\widehat{t}=d_1\widehat{k}-d_2\widehat{n}\widehat{\ell}\in d_\Delta(\mathbf{Z}/b\mathbf{Z})^\times$ (this $\widehat{r}$ is the residue of r mod b, and it is fixed inside each $\mathcal{D}_6$). Denoting $y(h):=\mathrm{e}(-ah\overline{\widehat{r}\widehat{k}}/b)$ and suppressing the congruences to $\widehat{e},\widehat{t},\widehat{k},\widehat{n},\widehat{\ell}$ through the notation $\sum^*$, we obtain

$$\mathcal{D}_6 = \sum_{\substack{|t|\sim T\\(t,b)=1}}^* \sum_{\substack{(k,n)=1\\(kn,d_abt)=1}}^* u'_k\beta'_n \sum_{\substack{e\sim E \text{ prime}\\(e,knb)=1}}^* \sum_{\substack{(\ell,d_aekbt)=1\\ d_2n\ell\equiv d_1k \pmod{d_aed_\Delta t}\\ |d_2n\ell-d_1k|>K/x^\eta}}^* \sum_{1\leqslant|h|\leqslant H} \gamma'_\ell\, y(h)\, \mathrm{e}\left(ahed_\Delta t\frac{\overline{bd_2n\ell}}{k}\right)$$
$$\times\left[\Phi\left(\frac{d_1k-d_2n\ell}{ed_\Delta tR}\right)\frac{ed_\Delta t}{d_1k-d_2n\ell}\widehat{\Phi}\left(\frac{hed_\Delta tM}{d_1k-d_2n\ell}\right)\right]. \tag{8.12}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

We now remove some of the dependencies between the variables t, k, n and e, ℓ, h, as in the proof of [May25a, Lemma 18.4]. Consider the function

$$\Psi(e,\ell,h) := \Phi\left(\frac{d_1k-d_2n\ell}{ed_\Delta tR}\right)\frac{ed_\Delta t}{d_1k-d_2n\ell}\widehat{\Phi}\left(\frac{hed_\Delta tM}{d_1k-d_2n\ell}\right) = \frac{1}{R}\alpha\,\Phi\left(\frac{1}{\alpha}\right)\widehat{\Phi}\left(\frac{Mh}{R}\alpha\right),$$

where $\alpha := ed_\Delta tR/(d_1k-d_2n\ell)$; note that Ψ is smooth in e,ℓ,h, and nonzero only if $\alpha \asymp 1$. Since $Mh/R \ll x^\eta$ and Φ, $\widehat{\Phi}$ have bounded derivatives, the chain rule and the bounds $d_2n\ell \asymp K$, $|d_2n\ell - d_1k| > K/x^\eta$ imply

$$\begin{aligned}\frac{\partial^{j_1+j_2+j_3}}{(\partial e)^{j_1}(\partial(d_1k-d_2n\ell))^{j_2}(\partial h)^{j_3}}\Psi(e,\ell,h) &\ll_{j_1,j_2,j_3} \frac{x^{\eta(j_1+j_2+j_3)}}{R}|e|^{-j_1}|d_1k-d_2n\ell|^{-j_2}|h|^{-j_3}\\ &\ll \frac{x^{\eta(j_1+2j_2+j_3)}}{R}|e|^{-j_1}|d_2n\ell|^{-j_2}|h|^{-j_3}.\end{aligned}$$

We thus have

$$\frac{\partial^{j_1+j_2+j_3}}{(\partial e)^{j_1}(\partial \ell)^{j_2}(\partial h)^{j_3}}\Psi(e,\ell,h) \ll_{j_1,j_2,j_3} \frac{x^{\eta(j_1+2j_2+j_3)}}{R}|e|^{-j_1}|\ell|^{-j_2}|h|^{-j_3},$$

and then by partial summation, (8.12) implies that

$$\mathcal{D}_6 \ll_\eta \frac{x^{O(\eta)}}{R}\sup_{\substack{H''\leqslant H\\ E''\leqslant 2E,\ L''\leqslant 2L}}\mathcal{D}_7, \tag{8.13}$$

where, after removing the residue constraints in the outer sums over t,k,n for an upper bound and putting $|h|$ in a dyadic interval,

$$\mathcal{D}_7 := \sum_{\substack{|t|\sim T\\(t,b)=1}}\sum_{\substack{(k,n)=1\\(kn,d_abt)=1}}|u'_k\beta'_n|\left|\sideset{}{^*}\sum_{\substack{e\sim E \text{ prime}\\ e\leqslant E''\\(e,knb)=1}}\sideset{}{^*}\sum_{\substack{(\ell,d_aekbt)=1,\ \ell\leqslant L''\\ d_2n\ell\equiv d_1k \pmod{d_aed_\Delta t}\\ |d_2n\ell-d_1k|>K/x^\eta\\ (d_1k-d_2n\ell)/t>0}}\gamma'_\ell\sum_{|h|\sim H''}y(h)\,\mathrm{e}\left(ahed_\Delta t\frac{\overline{bd_2n\ell}}{k}\right)\right|.$$

According to the desired bound in Proposition 8.4, and in light of (8.10), (8.11) and (8.13), it remains to show that

$$\mathcal{D}_7 \ll_{\varepsilon,\eta} x^{O(\eta)-\varepsilon/4}K^2, \tag{8.14}$$

for all $\eta\in(0,1)$. Now let $\mathcal{I}(t,k,n)$ be the subinterval of $[L/\delta_2, 2L/\delta_2]$ (which is the support of γ'_ℓ) containing those ℓ values such that

$$\ell\leqslant L'', \quad |d_2n\ell - d_1k| > K/x^\eta, \quad \text{and} \quad (d_1k-d_2n\ell)/t>0.$$

As in the proof of [May25a, Lemma 18.4], we can remove the constraint $\ell\in\mathcal{I}(t,k,n)$ using the identity

$$\mathbb{1}_{\ell\in\mathcal{I}(t,k,n)} = \int_0^1 \mathrm{e}(\ell\omega)\sum_{j\in\mathcal{I}(t,k,n)}\mathrm{e}(-j\omega)\,d\omega = \int_0^1 \mathrm{e}(\ell\omega)c(t,k,n,\omega)\min(L,\omega^{-1}(1-\omega)^{-1})\,d\omega,$$

for some coefficients $c(t,k,n,\omega)\ll 1$, and the L^1 bound $\int_0^{1/2}\min(L,2\omega^{-1})d\omega \ll \log x$. Together with the divisor bound $|u'_k|\ll_\eta x^\eta$, this shows that

$$\mathcal{D}_7 \ll_\eta x^{O(\eta)}\sup_{\omega\in\mathbf{R}/\mathbf{Z}}\mathcal{D}_8,$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

where

$$\mathcal{D}_8 := \sum_{\substack{|t|\sim T\\(t,b)=1}} \sum_{\substack{k\asymp K/d_1,\ n\sim N/\delta_1\\(k,n)=1\\(kn,d_abt)=1}} \Bigg| \sideset{}{^*}\sum_{\substack{e\sim E \text{ prime}\\ e\leqslant E''\\(e,knb)=1}} \sideset{}{^*}\sum_{\substack{(\ell,d_aekbt)=1\\ d_2n\ell\equiv d_1k \pmod{d_aed_\Delta t}}} \gamma'_\ell\, \mathrm{e}(\ell\omega) \sum_{|h|\sim H''} y(h)\, \mathrm{e}\left(ahed_\Delta t\frac{\overline{bd_2n\ell}}{k}\right)\Bigg|.$$

We denote $d'_i := d_i/v$ for $i\in\{1,2,\Delta\}$, so that the exponential term and the congruence in the summation over ℓ may be rewritten as

$$\mathrm{e}\left(ahed'_\Delta t\frac{\overline{bd'_2n\ell}}{k}\right), \quad d'_2n\ell\equiv d'_1k \pmod{ed_ad'_\Delta t},$$

where $(d'_1,d'_2)=(d'_1,d'_\Delta)=(d'_2,d'_\Delta)=1$. At this point, it also makes sense to denote

$$h' := \frac{a}{d_a}h, \quad t' := d_ad'_\Delta t, \quad n' := d'_2n,$$

$$H' := \frac{a}{d_a}H'', \quad T' := d_ad'_\Delta T, \quad N' := \frac{d'_2}{\delta_1}N, \quad L' := \frac{L}{\delta_2}, \quad K' := \frac{K}{d_1},$$

to bound (by dropping some divisibility constraints on n', t')

$$\mathcal{D}_8 \ll \sum_{\substack{k\asymp K'\\(k,b)=1}} \sum_{\substack{n'\sim N'\\(n',k)=1}} \sum_{\substack{|t'|\sim T'\\(t',d'_1k)=1}} \Bigg| \sideset{}{^*}\sum_{\substack{e\sim E \text{ prime}\\ e\leqslant E''\\(e,kn'b)=1}} \sideset{}{^*}\sum_{\substack{\ell\sim L'\\(\ell,bket')=1\\ n'\ell\equiv d'_1k \pmod{et'}}} \gamma'_\ell\, \mathrm{e}(\ell\omega) \sum_{\substack{|h'|\sim H'\\ \frac{a}{d_a}|h'}} y(h'd_a/a)\, \mathrm{e}\left(h'et'\frac{\overline{bn'\ell}}{k}\right)\Bigg|.$$

(To verify the new coprimality constraints, recall that $(d_a,b)=1$ and $d'_1d'_2d'_\Delta \mid b^\infty$.) We may replace the restriction that $(e,n')=1$ with $(e,d'_1)=1$, since each follows from the other and the congruence $n'\ell\equiv d'_1k \pmod e$, where $(\ell k,e)=1$. Moreover, by inserting 1-bounded coefficients $\beta(e,h',\ell)$, we can get rid of the coefficients $\gamma'_\ell\,\mathrm{e}(\ell\omega)y(h'd_a/a)$, the residue constraints (modulo b) in the summations over e and ℓ, as well as of the constraints that e is a prime and $e\leqslant E'$, and that $a/d_a \mid h'$. Overall, this yields

$$\mathcal{D}_8 \ll \sum_{\substack{k\asymp K'\\(k,b)=1}} \sum_{\substack{n'\sim N'\\(n',k)=1}} \sum_{\substack{|t'|\sim T'\\(t',d'_1k)=1}} \Bigg| \sum_{\substack{e\sim E\\(e,d'_1k)=1}} \sum_{|h'|\sim H'} \sum_{\substack{\ell\sim L'\\(\ell,ket')=1\\ n'\ell\equiv d'_1k \pmod{et'}}} \beta(e,h',\ell)\, \mathrm{e}\left(h'et'\frac{\overline{bn'\ell}}{k}\right)\Bigg|.$$

Finally, we insert a factor of $\sqrt{|t'|/T'}$ into the sum, and apply Cauchy and Schwarz in the outer variables k, n', t' to bound

$$\mathcal{D}_8^2 \ll K'N' \sum_{\substack{k\asymp K'\\(k,b)=1}} \sum_{\substack{n'\sim N'\\(n',k)=1}} \sum_{\substack{|t'|\sim T'\\(t',d'_1k)=1}} |t'| \Bigg| \sum_{\substack{e\sim E\\(e,d'_1k)=1}} \sum_{|h'|\sim H'} \sum_{\substack{\ell\sim L'\\(\ell,ket')=1\\ n'\ell\equiv d'_1k \pmod{et'}}} \beta(e,h',\ell)\, \mathrm{e}\left(h'et'\frac{\overline{bn'\ell}}{k}\right)\Bigg|^2.$$

Conjugating if necessary when $h'<0$ or $t'<0$, Proposition 8.6 implies that

$$\mathcal{D}_8^2 \ll_{\varepsilon,\eta} K'N'\cdot x^{O(\eta)-\varepsilon/2}N^2L^3 \ll x^{O(\eta)-\varepsilon/2}K^4,$$

for all $\eta\in(0,1)$. Putting things together, we conclude that

$$\mathcal{D}_7 \ll_{\varepsilon,\eta} x^{O(\eta)} \sup_{\omega\in\mathbf{R}/\mathbf{Z}} \mathcal{D}_8 \ll_\eta x^{O(\eta)-\varepsilon/4}\sqrt{K^4} \asymp x^{O(\eta)-\varepsilon/4}K^2,$$

as we required in (8.14). □

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

9. Bombieri–Friedlander–Iwaniec-style estimates

In this section, we establish Proposition 8.6, thus completing the proof of Proposition 6.2 and Theorem 4.2. We build on Maynard's work in [May25a, Chapter 18] (in a slightly more general setting, and using Theorem 3.10 instead of [DI82, Theorem 9]), which is in turn based on Bombieri–Friedlander–Iwaniec's work in [BFI87, § 10]. To aid future research, we shall consider a general sum

$$\mathcal{B}(K,N,T,E,H,L) := \sum_{\substack{k\sim K\\(k,b)=1}}\sum_{\substack{n\sim N\\(n,k)=1}}\sum_{\substack{t\sim T\\(t,dk)=1}} t\left|\sum_{\substack{e\sim E\\(e,dk)=1}}\sum_{h\sim H}\sum_{\substack{\ell\sim L\\(\ell,ket)=1\\ n\ell\equiv dk \pmod{et}}}\beta(e,h,\ell)\,\mathrm{e}\left(het\frac{\overline{bn\ell}}{k}\right)\right|^2,$$

where b,d are given positive integers with $b\ll x^{O(\eta)}$ and $d\ll x$, $\beta(e,h,\ell)$ are arbitrary 1-bounded coefficients, and the parameters K,N,T,E,H,L are almost arbitrary.

Remark 9.1. The trivial bound for $\mathcal{B}$ is $KN\,(TEH((L/ET)+1))^2\ll KN(HL)^2+KN(TEH)^2$, but we need more than a power saving over this (note that the desired bound in (8.9) is of the order of KNL^2, since we need to make up for the factors of H introduced during Poisson summation). So the relative sizes of K,N,T,E,H,L (as given by Proposition 8.6 and (4.2)) will ultimately be crucial, although we only take them into account after proving a general bound in Proposition 9.5.

After expanding the square inside $\mathcal{B}$, we reach a more complicated version of the sum anticipated in (2.5). The 'diagonal terms' with $h_1e_1\ell_2=h_2e_2\ell_1$ bring a contribution of roughly $O(KNTHL+KEHT^2L)$, similarly to (2.6); our deamplification set-up will be helpful here. In the off-diagonal terms, we complete Kloosterman sums via Lemma 3.6, and the principal frequency will contribute $O(NH^2L^2+NH^2TL)$. The remaining terms are ultimately separated into $\mathcal{B}_=$ and $\mathcal{B}_{\neq}$ (the latter corresponding to (2.7)), depending on whether $\ell_1=\ell_2$ or $\ell_1\neq\ell_2$.

Lemma 9.2 (Splitting the BFI-style sum). *For $\eta\in(0,1)$, $1\ll K,N,T,E,H,L\ll x$, and any positive integers b,d with $b\ll x^{O(\eta)}$ and $d\ll x$, one has*

$$\mathcal{B}(K,N,T,E,H,L)\ll_\eta x^{O(\eta)}\Bigg(KNTHL+KEHT^2L+NH^2L^2+NH^2TL+\frac{N}{KE^2}\sup_{\substack{E_0\ll E,\ S\ll TE_0,\\ J\ll (KTE^2x^\eta)/(NE_0)\\ w\in\mathbf{R}/\mathbf{Z},\,g_0}} E_0\,(\mathcal{B}_=+\mathcal{B}_{\neq})\Bigg),$$

where

$$\mathcal{B}_= := \sum_{t\sim T}\sum_{e_0\sim E_0}\sum_{\substack{e_1',e_2'\sim E/e_0\\(e_1',e_2')=1}}\sum_{\substack{s\sim S\\ s\mid te_0\\(s,e_1'e_2'b)=1}}\sum_{\substack{\ell\sim L\\(\ell,te_0e_1'e_2')=1}}\sum_{\substack{h_1,h_2\sim H\\ h_1e_1'\neq h_2e_2'}}$$
$$\times\left|\sum_{\substack{k'\\(k',e_1'e_2'\ell b)=1}} g_0\left(\frac{k'}{K/S}\right)\sum_{|j|\sim J}\mathrm{e}\left(j\left(\frac{d\bar{\ell}}{te_0e_1'e_2'}-w\right)\right)S((h_1e_1'-h_2e_2')\,\overline{b\ell e_1'e_2'},j;k's)\right|,$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

$$\mathcal{B}_{\neq} := \sum_{t\sim T}\sum_{e_0\sim E_0}\sum_{\substack{e_1',e_2'\sim E/e_0\\(e_1',e_2')=1}}\sum_{\substack{s\sim S\\ s\mid te_0\\(s,e_1'e_2'b)=1}}\sum_{\substack{\ell_1,\ell_2\sim L\\ \ell_1\equiv\ell_2 \pmod{te_0}\\(\ell_1\ell_2,te_0)=1\\(\ell_1,e_1')=(\ell_2,e_2')=1\\ \ell_1\neq\ell_2}}\sum_{\substack{h_1,h_2\sim H\\ h_1e_1'\ell_2\neq h_2e_2'\ell_1}}\Bigg|\sum_{\substack{k'\\(k',e_1'e_2'\ell_1\ell_2b)=1}} g_0\left(\frac{k'}{K/S}\right)$$
$$\times\sum_{|j|\sim J}\mathrm{e}\left(j\left(\frac{\mu}{te_0e_1'e_2'}-w\right)\right)S((h_1e_1'\ell_2-h_2e_2'\ell_1)\,\overline{b\ell_1\ell_2e_1'e_2'},j;k's)\Bigg|,$$

and $g_0(t)$ runs over smooth functions supported on $t\asymp 1$, satisfying $\|g_0^{(j)}\|_\infty\ll_j 1$ for each $j\geqslant 0$ (with fixed implicit constants). Here, $\mu=\mu(\ell_1,\ell_2,t,e_0,e_1',e_2',d)$ is the unique solution $\pmod{te_0e_1'e_2'}$ to the congruences $\mu\equiv d\overline{\ell_1}\pmod{te_0e_1'}$ and $\mu\equiv d\overline{\ell_2}\pmod{te_0e_2'}$.

Proof. This is essentially the same as the proof of [May25a, Lemma 18.5], but in a slightly more general setting (the main difference being the additional parameters b,d). We first replace the indicator functions of $k\sim K$ and $n\sim N$ with smooth majorants, using a suitable smooth compactly supported function f_0 (we choose this as in the proof of [May25a, Lemma 18.5]). Expanding out the square in $\mathcal{B}$ and swapping sums, then using that $(n,et)=1$ to deduce a congruence between the resulting variables ℓ_1 and ℓ_2 (indeed, if a prime p divided both n and et, then it would divide dk, but $(et,dk)=1$), we obtain

$$\mathcal{B}\leqslant\sum_{\substack{k\\(k,b)=1}}f_0\left(\frac{k}{K}\right)\sum_{\substack{n\\(n,k)=1}}f_0\left(\frac{n}{N}\right)\sum_{\substack{t\sim T\\(t,dk)=1}}t\Bigg|\sum_{\substack{e\sim E\\(e,dk)=1}}\sum_{h\sim H}\sum_{\substack{\ell\sim L\\(\ell,ket)=1\\ n\ell\equiv dk \pmod{et}}}\beta(e,h,\ell)\,\mathrm{e}\left(het\frac{\overline{bn\ell}}{k}\right)\Bigg|^2$$
$$=\sum_{\substack{t\sim T\\(t,d)=1}}t\sum_{\substack{e_1,e_2\sim E\\(e_1e_2,d)=1}}\sum_{\substack{\ell_1,\ell_2\sim L\\ \ell_1\equiv\ell_2 \pmod{t(e_1,e_2)}\\(\ell_1,te_1)=1\\(\ell_2,te_2)=1}}\sum_{h_1,h_2\sim H}\beta(e_1,h_1,\ell_1)\,\overline{\beta(e_2,h_2,\ell_2)}$$
$$\times\sum_{\substack{k\\(k,te_1e_2\ell_1\ell_2b)=1}}f_0\left(\frac{k}{K}\right)\sum_{\substack{(n,k)=1\\ n\equiv dk\overline{\ell_1} \pmod{te_1}\\ n\equiv dk\overline{\ell_2} \pmod{te_2}}}f_0\left(\frac{n}{N}\right)\mathrm{e}\left(t\frac{(h_1e_1\ell_2-h_2e_2\ell_1)\overline{bn\ell_1\ell_2}}{k}\right). \tag{9.1}$$

Let $\mathcal{B}_1$ denote the contribution of the 'diagonal' terms with $h_1e_1\ell_2=h_2e_2\ell_1$, and $\mathcal{B}_2$ contain the other terms; thus we have $\mathcal{B}\leqslant\mathcal{B}_1+\mathcal{B}_2$. As in (2.6), we first bound $\mathcal{B}_1$ trivially (using the divisor bound), by

$$\mathcal{B}_1\ll\sum_{\substack{t\sim T\\(t,d)=1}}t\sum_{\substack{e_1\sim E\\ \ell_2\sim L\\ h_1\sim H}}\sum_{\substack{e_2\sim E\\ \ell_1\sim L\\ e_2\ell_1\mid h_1e_1\ell_2\\ \ell_1\equiv\ell_2 \pmod{t(e_1,e_2)}\\(\ell_1,te_1)=(\ell_2,te_2)=1}}\sum_{\substack{h_2\sim H\\ h_1e_1\ell_2=h_2e_2\ell_1}}\sum_k f_0\left(\frac{k}{K}\right)\sum_{\substack{n\\ n\equiv dk\overline{\ell_1} \pmod{te_1}\\ n\equiv dk\overline{\ell_2} \pmod{te_2}}}f_0\left(\frac{n}{N}\right)$$
$$\ll\sum_{t\sim T}t\sum_{\substack{e_1\sim E\\ \ell_2\sim L\\ h_1\sim H}}\sum_{\substack{e_2\sim E,\ \ell_1\sim L\\ h_2\sim H\\ h_1e_1\ell_2=h_2e_2\ell_1\\(\ell_1,te_1)=1}}\sum_{k\ll K}\sum_{\substack{n\ll N\\ n\equiv dk\overline{\ell_1} \pmod{te_1}}}1$$
$$\ll_\eta x^{O(\eta)}\,T^2ELHK\left(\frac{N}{TE}+1\right), \tag{9.2}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

recovering the first two terms in the desired bound. Next, we consider $\mathcal{B}_2$, containing the terms with $h_1e_1\ell_2 \neq h_2e_2\ell_1$. We let $e_0 := (e_1, e_2)$, $e_1' := e_1/e_0$ and $e_2' := e_2/e_0$ and put e_0 in dyadic ranges $e_0 \sim E_0$ to write

$$\begin{aligned}\mathcal{B}_2 \ll (\log x) &\\ \times \sup_{E_0 \ll E} \Bigg| \sum_{\substack{t\sim T\\(t,d)=1}} t \sum_{\substack{e_0\sim E_0\\(e_0,d)=1}} \sum_{\substack{e_1',e_2'\sim E/e_0\\(e_1',e_2')=1\\(e_1'e_2',d)=1}} \sum_{\substack{\ell_1,\ell_2\sim L\\ \ell_1\equiv\ell_2 \pmod{te_0}\\(\ell_1\ell_2,te_0)=1\\(\ell_1,e_1')=(\ell_2,e_2')=1}} \sum_{\substack{h_1,h_2\sim H\\h_1e_1'\ell_2\neq h_2e_2'\ell_1}} \beta(e_0e_1',h_1,\ell_1)\,\overline{\beta(e_0e_2',h_2,\ell_2)} &\\ \times \sum_{\substack{k\\(k,te_0e_1'e_2'\ell_1\ell_2b)=1}} f_0\left(\frac{k}{K}\right) \sum_{\substack{(n,k)=1\\n\equiv dk\overline{\ell_1} \pmod{te_0e_1'}\\n\equiv dk\overline{\ell_2} \pmod{te_0e_2'}}} f_0\left(\frac{n}{N}\right) \mathrm{e}\left(te_0\frac{(h_1e_1'\ell_2-h_2e_2'\ell_1)\overline{bn\ell_1\ell_2}}{k}\right)\Bigg|.&\end{aligned}$$

Note that the inner sum over n can be rewritten as

$$\sum_{\substack{n\equiv k\mu \pmod{te_0e_1'e_2'}\\(n,k)=1}} f_0\left(\frac{n}{N}\right)\mathrm{e}\left(\frac{\overline{n}r_0}{k}\right),$$

where $r_0 := te_0(h_1e_1'\ell_2 - h_2e_2'\ell_1)\overline{b\ell_1\ell_2}$ (defined mod k), and $\mu = \mu(\ell_1,\ell_2,t,e_0,e_1',e_2',d)$ is the unique solution (mod $te_0e_1'e_2'$) to the congruences $\mu \equiv d\overline{\ell_1}$ (mod te_0e_1') and $\mu \equiv d\overline{\ell_2}$ (mod te_0e_2'); the latter is well-defined by the Chinese remainder theorem, since $(te_0e_1', te_0e_2') = te_0(e_1',e_2') = te_0$, $[te_0e_1', te_0e_2'] = te_0e_1'e_2'$, and $d\overline{\ell_1} \equiv d\overline{\ell_2}$ (mod te_0). Crucially, note that μ does not depend on k.

We can thus complete Kloosterman sums using Lemma 3.6, with

$$\begin{aligned} q &:= te_0e_1'e_2',\\ J_0 &:= 32\,x^{\eta}\frac{KTE^2}{NE_0} \geqslant x^{\eta}\frac{kq}{N},\end{aligned}$$

giving us

$$\begin{aligned}\sum_{\substack{n\equiv k\mu \pmod q\\(n,k)=1}} f_0\left(\frac{n}{N}\right)\mathrm{e}\left(\frac{\overline{n}r_0}{k}\right) &= \frac{N}{kq}\sum_{|j|\leqslant J_0}\widehat{f_0}\left(\frac{jN}{kq}\right)\mathrm{e}\left(\frac{j\mu}{q}\right)S(j\overline{q},r_0;k)+O_\eta(x^{-99})\\ &= \frac{N}{kq}\sum_{|j|\leqslant J_0}\widehat{f_0}\left(\frac{jN}{kq}\right)\mathrm{e}\left(\frac{j\mu}{q}\right)S(r_1,j;k)+O_\eta(x^{-99}),\end{aligned}$$

where

$$r_1 := r_0\overline{q} = (h_1e_1'\ell_2 - h_2e_2'\ell_1)\,\overline{b\ell_1\ell_2e_1'e_2'}.$$

We now plug this bound into our estimate for $\mathcal{B}_2$, isolate the contribution of $j=0$ into $\mathcal{B}_3$, and let $\mathcal{B}_4$ contain the terms with $j \neq 0$. This yields

$$\mathcal{B}_2 \ll_\eta x^{O(\eta)} \sup_{E_0\ll E}(|\mathcal{B}_3|+|\mathcal{B}_4|) + O_\eta(x^{-50}), \tag{9.3}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

where

$$\mathcal{B}_3 := \sum_{\substack{t\sim T\\(t,d)=1}} t \sum_{\substack{e_0\sim E_0\\(e_0,d)=1}} \sum_{\substack{e_1',e_2'\sim E/e_0\\(e_1',e_2')=1\\(e_1'e_2',d)=1}} \sum_{\substack{\ell_1,\ell_2\sim L\\\ell_1\equiv\ell_2 \pmod{te_0}\\(\ell_1\ell_2,te_0)=1\\(\ell_1,e_1')=(\ell_2,e_2')=1}} \sum_{\substack{h_1,h_2\sim H\\h_1e_1'\ell_2\neq h_2e_2'\ell_1}} \beta(e_0e_1',h_1,\ell_1)\,\overline{\beta(e_0e_2',h_2,\ell_2)}$$
$$\times \sum_{\substack{k\\(k,te_0e_1'e_2'\ell_1\ell_2b)=1}} f_0\left(\frac{k}{K}\right)\frac{N}{kq}\widehat{f}_0\left(0\right)S(r_1,0;k),$$

$$\mathcal{B}_4 := \sum_{\substack{t\sim T\\(t,d)=1}} t \sum_{\substack{e_0\sim E_0\\(e_0,d)=1}} \sum_{\substack{e_1',e_2'\sim E/e_0\\(e_1',e_2')=1\\(e_1'e_2',d)=1}} \sum_{\substack{\ell_1,\ell_2\sim L\\\ell_1\equiv\ell_2 \pmod{te_0}\\(\ell_1\ell_2,te_0)=1\\(\ell_1,e_1')=(\ell_2,e_2')=1}} \sum_{\substack{h_1,h_2\sim H\\h_1e_1'\ell_2\neq h_2e_2'\ell_1}} \beta(e_0e_1',h_1,\ell_1)\,\overline{\beta(e_0e_2',h_2,\ell_2)}$$
$$\times \sum_{\substack{k\\(k,te_0e_1'e_2'\ell_1\ell_2b)=1}} f_0\left(\frac{k}{K}\right)\frac{N}{kq}\sum_{\substack{|j|\leqslant J_0\\j\neq 0}}\widehat{f}_0\left(\frac{jN}{kq}\right)\mathrm{e}\left(\frac{j\mu}{q}\right)S(r_1,j;k).$$

We bound $\mathcal{B}_3$ trivially using the Ramanujan bound (Lemma 3.8)

$$\begin{aligned}\mathcal{B}_3 &\ll \sum_{t\sim T} t \sum_{e_0\sim E_0} \sum_{e_1',e_2'\sim E/e_0} \sum_{\substack{\ell_1,\ell_2\sim L\\\ell_1\equiv\ell_2 \pmod{te_0}}} \sum_{h_1,h_2\sim H} \sum_{k\ll K} \frac{N}{kq}(r_1,k)\\ &\ll_\eta x^{O(\eta)} \sum_{t\sim T} t \sum_{e_0\sim E_0} \sum_{e_1',e_2'\sim E/e_0} \sum_{\substack{\ell_1,\ell_2\sim L\\\ell_1\equiv\ell_2 \pmod{te_0}}} \frac{H^2N}{q}\\ &\ll x^{O(\eta)}\, T^2\frac{E^2}{E_0}L\left(1+\frac{L}{TE_0}\right)\frac{H^2NE_0}{TE^2}\\ &\ll x^{O(\eta)}\left(TLH^2N+L^2H^2N\right),\end{aligned} \tag{9.4}$$

giving the third and fourth terms in the desired bound. We finally turn to estimating $\mathcal{B}_4$, and start by removing the coprimality constraint $(k,te_0)=1$, via Möbius inversion. We write $\mathbb{1}_{(k,te_0)=1} = \sum_{s|(k,te_0)}\mu(s)$ and $k=k's$, and put j,s into dyadic ranges $j\sim J, s\sim S$ to obtain

$$\mathcal{B}_4 \ll_\eta x^\eta T \sup_{\substack{S\ll TE_0\\J\ll J_0}} \mathcal{B}_5, \tag{9.5}$$

where

$$\begin{aligned}\mathcal{B}_5 &:= \sum_{\substack{t\sim T\\(t,d)=1}} \sum_{\substack{e_0\sim E_0\\(e_0,d)=1}} \sum_{\substack{e_1',e_2'\sim E/e_0\\(e_1',e_2')=1\\(e_1'e_2',d)=1}} \sum_{\substack{s\sim S\\s|te_0\\(s,e_1'e_2'b)=1}} \sum_{\substack{\ell_1,\ell_2\sim L\\\ell_1\equiv\ell_2 \pmod{te_0}\\(\ell_1\ell_2,te_0)=1\\(\ell_1,e_1')=(\ell_2,e_2')=1}} \mathcal{B}_6,\\ \mathcal{B}_6 &:= \sum_{\substack{h_1,h_2\sim H\\h_1e_1'\ell_2\neq h_2e_2'\ell_1}} \left|\sum_{\substack{k'\\(k',e_1'e_2'\ell_1\ell_2b)=1}} f_0\left(\frac{k's}{K}\right)\frac{N}{k'sq}\sum_{j\sim J}\widehat{f}_0\left(\frac{jN}{k'sq}\right)\mathrm{e}\left(\frac{j\mu}{q}\right)S(r_1,j;k's)\right|.\end{aligned} \tag{9.6}$$

We now wish to separate the j,k' variables in $\mathcal{B}_6$ from the others, in the factors of f_0, $\widehat{f}_0$, and the exponential term; note that s, $q=te_0e_1'e_2'$ and $\mu=\mu(\ell_1,\ell_2,c,t,e_0,e_1',e_2',d)$ do not depend on j and k'. As in the proof of [May25a, Lemma 18.5], we make use of the special choice of

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

the smooth function $f_0(t) := \int_0^\infty \psi_0(y)\psi_0(t/y)\,dy/y$ (which is a multiplicative convolution of a bounded smooth function ψ_0 supported on $[1/2, 5/2]$ with itself) to find that

$$f_0\left(\frac{k's}{K}\right) = \int_U^{20U} \psi_0(su)\,\psi_0\left(\frac{k'}{Ku}\right)\frac{du}{u},$$

where $U \asymp 1/S$, and also

$$\frac{N}{k'sq}\widehat{f}_0\left(\frac{jN}{k'sq}\right) = \int_V^{20V}\int_W^{20W} \psi_0(k'v)\,\psi_0\left(\frac{wsq}{Nv}\right)\mathrm{e}(-jw)\,dw\frac{dv}{v},$$

where $V \asymp S/K$ and $W \asymp NVE_0/(STE^2) \asymp NE_0/(KTE^2)$. Plugging this into our expression for $\mathcal{B}_6$, taking the integrals over u, v, w outside the absolute value by the triangle inequality, and swapping them with the sum over h_1, h_2, we get

$$\mathcal{B}_6 \ll \int_U^{20U}\int_V^{20V}\int_W^{20W} |\mathcal{B}_7|\frac{du\,dv\,dw}{uv}, \tag{9.7}$$

where

$$\mathcal{B}_7 := \sum_{\substack{h_1,h_2\sim H\\ h_1e_1'\ell_2\neq h_2e_2'\ell_1}}\left|\sum_{\substack{k'\\ (k',e_1'e_2'\ell_1\ell_2b)=1}} g_{u,v}\left(\frac{k'}{K/S}\right)\sum_{|j|\sim J}\mathrm{e}\left(j\left(\frac{\mu}{q}-w\right)\right)S(r_1,j;k's)\right|,$$

and the smooth function

$$g_{u,v}(t) := \psi_0\left(\frac{t}{uS}\right)\psi_0\left(\frac{vtK}{S}\right)$$

is supported on $t \asymp 1$. Combining this with (9.5) to (9.7), moving the integrals in u, v, w to the front and taking an L^∞ bound, we find that

$$\mathcal{B}_4 \ll_\eta x^\eta TW \sup_{\substack{S\ll TE_0\\ J\ll J_0}}\sup_{\substack{u\asymp 1/S\\ v\asymp S/K\\ w\asymp NE_0/KE^2}}\sum_{t\sim T}\sum_{e_0\sim E_0}\sum_{\substack{e_1',e_2'\sim E/e_0\\ (e_1',e_2')=1}}\sum_{\substack{s\sim S\\ s|te_0\\ (s,e_1'e_2'b)=1}}\sum_{\substack{\ell_1,\ell_2\sim L\\ \ell_1\equiv\ell_2 \pmod{te_0}\\ (\ell_1\ell_2,te_0)=1\\ (\ell_1,e_1')=(\ell_2,e_2')=1}}\mathcal{B}_7,$$

where $TW \asymp NE_0/KE^2$. Letting $\mathcal{B}_=$ be the contribution of the terms $\ell_1 = \ell_2$ and $\mathcal{B}_\neq$ contain the terms with $\ell_1 \neq \ell_2$, and combining this with (9.1) to (9.4), we recover the desired bound for $\mathcal{B}$ (note that when $\ell_1 = \ell_2 = \ell$, one can take $\mu = d\bar{\ell}$). □

Lemma 9.3 (Contribution of $\ell_1 = \ell_2$). *With the notation of Lemma 9.2, assuming that $EHT \ll x^{O(\eta)}KNL$, one has*

$$\mathcal{B}_= \ll_\eta \frac{x^{O(\eta)}KE^2}{NE_0}\left(1+\frac{K}{E^2L}\right)^{\theta_{\max}}(KT^5E^8H^2L^3N)^{1/2}\left(1+\frac{H}{E}+\frac{H^2}{E^2L}\right)^{1/2}\left(1+\frac{K}{NL}\right)^{1/2}.$$

Proof of Lemma 9.3 assuming Theorem 3.10. Here, we adapt the proof of [May25a, Lemma 18.7], using Theorem 3.10 instead of [DI82, Theorem 9]. To do this, we need to eliminate the dependency of the inner exponential coefficients on ℓ, so we write

$$\mathcal{B}_= = \sum_{t\sim T}\sum_{e_0\sim E_0}\sum_{\substack{e_1',e_2'\sim E/e_0\\ (e_1',e_2')=1}}\sum_{\substack{s\sim S\\ s|te_0\\ (s,e_1'e_2'b)=1}}\sum_{\widehat{\ell}\in(\mathbf{Z}/te_0e_1'e_2'\mathbf{Z})^\times}\sum_{\substack{\ell\sim L\\ \ell\equiv\widehat{\ell} \pmod{te_0e_1'e_2'}}}$$
$$\times\sum_{\substack{h_1,h_2\sim H\\ h_1e_1'\neq h_2e_2'}}\left|\sum_{\substack{k'\\ (k',e_1'e_2'\ell b)=1}} g_0\left(\frac{k'}{K/S}\right)\sum_{|j|\sim J}\mathrm{e}(j\omega)\,S((h_1e_1'-h_2e_2')\,\overline{b\ell e_1'e_2'},j;k's)\right|,$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

where

$$\omega = \omega(t, e_0, e_1', e_2', \widehat{\ell}, d, w) := \frac{d\overline{\widehat{\ell}}}{te_0e_1'e_2'} - w \in \mathbf{R}/\mathbf{Z}.$$

We now denote

$$m := h_1e_1' - h_2e_2' \ll \frac{HE}{E_0}, \quad r := b\ell e_1'e_2' \asymp x^{O(\eta)}\frac{LE^2}{E_0^2},$$

split into dyadic ranges $m \sim M_=$, $r \sim R_=$, and change variables from ℓ to r to obtain

$$\mathcal{B}_= \ll_\eta x^{O(\eta)} \sup_{\substack{M_= \ll HE/E_0 \\ R_= \asymp x^{O(\eta)}LE^2/E_0^2}} \sum_{t\sim T}\sum_{e_0\sim E_0}\sum_{\substack{e_1',e_2'\sim E/e_0 \\ (e_1',e_2')=1}}\sum_{\widehat{\ell}\in(\mathbf{Z}/te_0e_1'e_2'\mathbf{Z})^\times}$$
$$\times \sum_{\substack{r\sim R_= \\ be_1'e_2'\mid r \\ r/(be_1'e_2')\equiv\widehat{\ell} \\ (\mathrm{mod}\ te_0e_1'e_2')}}\sum_{\substack{s\sim S \\ (s,r)=1 \\ s\mid te_0}}\sum_{m\sim M_=}\sum_{\substack{h_1,h_2\sim H \\ h_1e_1'-h_2e_2'=m}}\left|\sum_{\substack{k' \\ (k',r)=1}} g_0\left(\frac{k'}{K/S}\right)\sum_{|j|\sim J}\mathrm{e}(j\omega)\, S(m\overline{r}, j; k's)\right|.$$

Crucially, once the variables $t, e_0, e_1', e_2', \widehat{\ell}$ are fixed, ω does not depend on r, s, m, h_1, h_2, k', or j. Finally, we remove the absolute values by inserting 1-bounded coefficients ξ_{h_1,h_2} (also depending on $t, e_0, e_1', e_2', \widehat{\ell}, r, s, m$), and denote

$$a_{m,r,s} = a_{m,r,s}(t, e_0, e_1', e_2', \widehat{\ell}) := \mathbb{1}_{be_1'e_2'\mid r}\mathbb{1}_{\substack{r/(be_1'e_2')\equiv\widehat{\ell} \\ (\mathrm{mod}\ te_0e_1'e_2')}}\mathbb{1}_{s\mid te_0}\sum_{\substack{h_1,h_2\sim H \\ h_1e_1'-h_2e_2'=m}}\xi_{h_1,h_2},$$

to get

$$\mathcal{B}_= \ll_\eta x^{O(\eta)} \sup_{\substack{M_= \ll HE/E_0 \\ R_= \asymp x^{O(\eta)}LE^2/E_0^2}} \sum_{t\sim T}\sum_{e_0\sim E_0}\sum_{\substack{e_1',e_2'\sim E/e_0 \\ (e_1',e_2')=1}}\sum_{\widehat{\ell}\in(\mathbf{Z}/te_0e_1'e_2'\mathbf{Z})^\times}|\mathcal{K}_=|, \tag{9.8}$$

with

$$\mathcal{K}_= := \sum_{\substack{r\sim R_= \\ s\sim S \\ (r,s)=1}}\sum_{m\sim M_=} a_{m,r,s}\sum_{|j|\sim J}\mathrm{e}(j\omega)\sum_{\substack{k' \\ (k',r)=1}} g_0\left(\frac{k'}{K/S}\right) S(m\overline{r}, j; k's).$$

At this point we apply our Deshouillers–Iwaniec-style bound from Theorem 3.10, finding that

$$\mathcal{K}_= \ll_\eta x^{O(\eta)}\left(1 + \frac{K/S}{R_=\sqrt{S}}\right)^{\theta_{\max}}\|a_{m,r,s}\|_2\sqrt{JR_=S}$$
$$\times\left(\frac{K^2/S^2}{R_=}(M_= + R_=S)(J + R_=S) + M_=J\right)^{1/2},$$

where, by Cauchy and Schwarz,

$$\sum_{t\sim T}\sum_{e_0\sim E_0}\sum_{\substack{e_1',e_2'\sim E/e_0 \\ (e_1',e_2')=1}}\sum_{\widehat{\ell}\in(\mathbf{Z}/te_0e_1'e_2'\mathbf{Z})^\times}\|a_{m,r,s}\|_2$$
$$\ll \frac{TE^2}{E_0}\sqrt{\sum_{t\sim T}\sum_{e_0\sim E_0}\sum_{\substack{e_1',e_2'\sim E/e_0 \\ (e_1',e_2')=1}}\sum_{\substack{r\sim R_= \\ be_1'e_2'\mid r}}\sum_{\substack{s\sim S \\ (s,r)=1 \\ s\mid te_0}}\sum_{m\sim M_=}\left(\sum_{\substack{h_1,h_2\sim H \\ h_1e_1'-h_2e_2'=m}}1\right)^2}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

$$\ll_\eta x^{O(\eta)}\frac{TE^2}{E_0}\sqrt{TE_0\frac{R_=E_0^2}{E^2}\sum_{m\sim M_=}\sum_{h_2\sim H}\sum_{e_2'\ll E/E_0}\sum_{\substack{h_1,e_1'\\(e_1',e_2')=1\\h_1e_1'=m+h_2e_2'}}\sum_{\substack{h_1'\sim H\\h_1'e_1'\equiv m \pmod{e_2'}}}\sum_{\substack{h_2'\sim H\\h_2'e_2'=h_1'e_1'-m}}1}$$

$$\ll_\eta x^{O(\eta)}\frac{TE^2}{E_0}\left(\frac{TE_0^3R_=}{E^2}M_=H\frac{E}{E_0}\left(1+\frac{HE_0}{E}\right)\right)^{1/2}$$

$$=x^{O(\eta)}\left(T^3E^3R_=M_=H\left(1+\frac{HE_0}{E}\right)\right)^{1/2}.$$

Plugging these bounds into (9.8), we obtain

$$\mathcal{B}_=\ll_\eta x^{O(\eta)}\sup_{\substack{M_=\ll HE/E_0\\R_=\asymp x^{O(\eta)}LE^2/E_0^2}}\left(1+\frac{K/S}{R_=\sqrt{S}}\right)^{\theta_{\max}}\left(T^3E^3R_=M_=H\left(1+\frac{HE_0}{E}\right)\right)^{1/2}(JR_=S)^{1/2}$$

$$\times\left(\frac{K^2/S^2}{R_=}(M_=+R_=S)(J+R_=S)+M_=J\right)^{1/2}.$$

Using $M_=\ll HE/E_0$ and $R_=\asymp x^{O(\eta)}LE^2/E_0^2$, and recalling that $J\ll(KTE^2x^\eta)/(NE_0)$ (from Lemma 9.2), we conclude that

$$\mathcal{B}_=\ll_\eta x^{O(\eta)}\left(1+\frac{KE_0^2}{LE^2S^{3/2}}\right)^{\theta_{\max}}\left(T^3E^3\frac{LE^2}{E_0^2}\frac{HE}{E_0}H\left(1+\frac{HE_0}{E}\right)\right)^{1/2}\left(\frac{KTE^2}{NE_0}\frac{LE^2}{E_0^2}S\right)^{1/2}$$

$$\times\left(\frac{K^2E_0^2}{LE^2S^2}\left(\frac{HE}{E_0}+\frac{LE^2}{E_0^2}S\right)\left(\frac{KTE^2}{NE_0}+\frac{LE^2}{E_0^2}S\right)+\frac{HKTE^3}{NE_0^2}\right)^{1/2}$$

$$\ll x^{O(\eta)}\left(1+\frac{KE_0^2}{LE^2}\right)^{\theta_{\max}}\left(T^3E^3\frac{LE^2}{E_0^2}\frac{HE}{E_0}H\left(1+\frac{HE_0}{E}\right)\right)^{1/2}\left(\frac{KTE^2}{NE_0}\frac{LE^2}{E_0^2}\right)^{1/2}$$

$$\times\left(\frac{K^2E_0^2}{LE^2}\left(\frac{HE}{E_0S}+\frac{LE^2}{E_0^2}\right)\left(\frac{KTE^2}{NE_0}+\frac{LE^2}{E_0^2}S\right)+\frac{HKTE^3}{NE_0^2}S\right)^{1/2},$$

where we lower-bounded $S\gg1$ in the factor raised to $\theta_{\max}$. Since $1\ll S\ll TE_0$ (from Lemma 9.2), our bound becomes

$$\mathcal{B}_=\ll_\eta x^{O(\eta)}\left(1+\frac{KE_0^2}{LE^2}\right)^{\theta_{\max}}\left(T^3E^3\frac{LE^2}{E_0^2}\frac{HE}{E_0}H\left(1+\frac{HE_0}{E}\right)\right)^{1/2}\left(\frac{KTE^2}{NE_0}\frac{LE^2}{E_0^2}\right)^{1/2}$$

$$\times\left(\frac{K^2E_0^2}{LE^2}\left(\frac{HE}{E_0}+\frac{LE^2}{E_0^2}\right)\left(\frac{KTE^2}{NE_0}+\frac{LE^2T}{E_0}\right)+\frac{HKT^2E^3}{NE_0}\right)^{1/2}.$$

This expression is nonincreasing in E_0, even after extracting a factor of E_0^{-1} (since $\theta_{\max}<3/4$); thus lower-bounding $E_0\gg1$ we obtain

$$\mathcal{B}_=\ll_\eta\frac{x^{O(\eta)}}{E_0}\left(1+\frac{K}{LE^2}\right)^{\theta_{\max}}\left(T^3E^6LH^2\left(1+\frac{H}{E}\right)\right)^{1/2}\left(\frac{KTLE^4}{N}\right)^{1/2}$$

$$\times\left(\frac{K^2}{LE^2}(HE+LE^2)\left(\frac{KTE^2}{N}+LE^2T\right)+\frac{HKT^2E^3}{N}\right)^{1/2}.$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

We can simplify this further using our assumption that $EHT \ll x^{O(\eta)}KNL$, which implies

$$x^{-O(\eta)}\frac{HKT^2E^3}{N} \ll K^2LE^2T = \frac{K^2}{LE^2}LE^2LE^2T,$$

allowing us to discard the term of HKT^2E^3/N in the second line. Thus

$$\begin{aligned}\mathcal{B}_= &\ll_\eta \frac{x^{O(\eta)}}{E_0}\left(1+\frac{K}{LE^2}\right)^{\theta_{\max}}\left(1+\frac{H}{E}\right)^{1/2}\left(T^3E^6LH^2\frac{KTLE^4}{N}\right)^{1/2}\\ &\quad\times K\left(\frac{H}{LE}+1\right)^{1/2}(LE^2T)^{1/2}\left(\frac{K}{NL}+1\right)^{1/2}\\ &\ll \frac{x^{O(\eta)}}{E_0}\left(1+\frac{K}{LE^2}\right)^{\theta_{\max}}\left(\frac{K^3T^5L^3E^{12}H^2}{N}\right)^{1/2}\left(1+\frac{H}{E}+\frac{H^2}{LE^2}\right)^{1/2}\left(1+\frac{K}{NL}\right)^{1/2}\\ &\ll \frac{x^{O(\eta)}}{E_0}\frac{KE^2}{N}\left(1+\frac{K}{LE^2}\right)^{\theta_{\max}}(KT^5L^3E^8H^2N)^{1/2}\left(1+\frac{H}{E}+\frac{H^2}{LE^2}\right)^{1/2}\left(1+\frac{K}{NL}\right)^{1/2}.\end{aligned}$$

After slightly rearranging factors, this yields the desired bound. □

LEMMA 9.4 (Contribution of $\ell_1 \neq \ell_2$). *With the notation of Lemma 9.2, assuming that $EHT \ll x^{O(\eta)}KNL$, one has*

$$\mathcal{B}_{\neq} \ll_\eta \frac{x^{O(\eta)}KE^2}{NE_0}\left(1+\frac{K}{E^2L^2}\right)^{\theta_{\max}}(KT^4E^8H^2L^6N)^{1/2}\left(1+\frac{H}{EL}\right)\left(1+\frac{K}{NL^2}\right)^{1/2}.$$

Proof of Lemma 9.4 assuming Theorem 3.10. Here, we follow the proof of [May25a, Lemma 18.6], using Theorem 3.10 instead of [DI82, Theorem 9]. As in Lemma 9.3, we need to eliminate the dependency of the inner exponential coefficients on ℓ_1 and ℓ_2, so we write

$$\begin{aligned}\mathcal{B}_{\neq} = \sum_{t\sim T}\sum_{e_0\sim E_0}\sum_{\substack{e_1',e_2'\sim E/e_0\\(e_1',e_2')=1}}\sum_{\substack{s\sim S\\ s\mid te_0\\(s,e_1'e_2'b)=1}}\sum_{\widehat{\mu}\in\mathbf{Z}/te_0e_1'e_2'\mathbf{Z}}\sum_{\substack{\ell_1,\ell_2\sim L\\ \ell_1\equiv\ell_2 \pmod{te_0}\\(\ell_1\ell_2,te_0)=1\\(\ell_1,e_1')=(\ell_2,e_2')=1\\ \ell_1\neq\ell_2\\ \mu(\ell_1,\ell_2)\equiv\widehat{\mu} \pmod{te_0e_1'e_2'}}}\\ \times\sum_{\substack{h_1,h_2\sim H\\ h_1e_1'\ell_2\neq h_2e_2'\ell_1}}\left|\sum_{\substack{k'\\(k',e_1'e_2'\ell_1\ell_2b)=1}}g_0\left(\frac{k'}{K/S}\right)\sum_{|j|\sim J}\mathrm{e}\left(j\omega\right)\,S((h_1e_1'\ell_2-h_2e_2'\ell_1)\,\overline{b\ell_1\ell_2e_1'e_2'},j;k's)\right|,\end{aligned}$$

where

$$\omega=\omega(t,e_0,e_1',e_2',\widehat{\mu},d,w):=\frac{\widehat{\mu}}{te_0e_1'e_2'}-w\in\mathbf{R}/\mathbf{Z}.$$

This is essentially the exponential sum anticipated in (2.7).

We then let $\ell_0:=(\ell_1,\ell_2)$, $\ell_1':=\ell_1/\ell_0$, $\ell_2':=\ell_2/\ell_0$, and put the variables

$$m:=h_1e_1'\ell_2'-h_2e_2'\ell_1'\ll\frac{HEL}{E_0\ell_0},\quad r:=b\ell_0\ell_1'\ell_2'e_1'e_2'\asymp x^{O(\eta)}\frac{L^2E^2}{\ell_0E_0^2},$$

and ℓ_0 into dyadic ranges $m\sim M_{\neq}$, $r\sim R_{\neq}$, $\ell_0\sim L_0$ to obtain

$$\mathcal{B}_{\neq}\ll_\eta x^{O(\eta)}\sup_{\substack{L_0\ll L\\ M_{\neq}\ll HEL/(E_0L_0)\\ R_{\neq}\asymp x^{O(\eta)}L^2E^2/(L_0E_0^2)}}|\mathcal{B}_{\neq}'|,\tag{9.9}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

where

$$\mathcal{B}'_{\neq} := \sum_{t\sim T}\sum_{e_0\sim E_0}\sum_{\substack{e_1',e_2'\sim E/e_0\\(e_1',e_2')=1}}\sum_{\widehat{\mu}\in\mathbf{Z}/te_0e_1'e_2'\mathbf{Z}} \times \sum_{\substack{r\sim R_{\neq}\\ be_1'e_2'|r\\ s\sim S\\(s,r)=1\\ s|te_0}}\sum_{\substack{\ell_0\sim L_0\\ \ell_1',\ell_2'\sim L/\ell_0\\(\ell_1',\ell_2')=1,\ell_1'\neq\ell_2'\\ b\ell_0\ell_1'\ell_2'e_1'e_2'=r\\ \ell_1'\equiv\ell_2' \pmod{te_0}\\(\ell_1',e_1')=(\ell_2',e_2')=1\\ \mu(\ell_0\ell_1',\ell_0\ell_2')\equiv\widehat{\mu}\\ \pmod{te_0e_1'e_2'}}}\sum_{m\sim M_{\neq}}\sum_{\substack{h_1,h_2\sim H\\ h_1e_1'\ell_2-h_2e_2'\ell_1=m}}\Bigg|\sum_{\substack{k'\\(k',r)=1}} g_0\left(\frac{k'}{K/S}\right)\sum_{|j|\sim J}\mathrm{e}(j\omega)\,S(m\overline{r},j;k's)\Bigg|.$$

As before, once the variables $t, e_0, e_1', e_2', \widehat{\mu}$ are fixed, ω does not depend on r, s, m, h_1, h_2, k'. We remove the absolute values by inserting 1-bounded coefficients ξ_{h_1,h_2} (also depending on $t, e_0, e_1', e_2', \widehat{\mu}$ and r, s, m), and denote

$$a_{m,r,s} = a_{m,r,s}(t, e_0, e_1', e_2', \widehat{\mu}) := \mathbb{1}_{be_1'e_2'|r}\mathbb{1}_{s|te_0}\sum_{\substack{\ell_0\sim L_0\\ \ell_1',\ell_2'\sim L/\ell_0\\(\ell_1',\ell_2')=1,\ell_1'\neq\ell_2'\\ b\ell_0\ell_1'\ell_2'e_1'e_2'=r\\ \ell_1'\equiv\ell_2' \pmod{te_0}\\(\ell_1',e_1')=(\ell_2',e_2')=1\\ \mu(\ell_0\ell_1',\ell_0\ell_2')\equiv\widehat{\mu}\\ \pmod{te_0e_1'e_2'}}}\sum_{\substack{h_1,h_2\sim H\\ h_1e_1'\ell_2'-h_2e_2'\ell_1'=m}}\xi_{h_1,h_2},$$

to obtain

$$\mathcal{B}'_{\neq} \ll_{\eta} x^{O(\eta)}\sum_{t\sim T}\sum_{e_0\sim E_0}\sum_{\substack{e_1',e_2'\sim E/e_0\\(e_1',e_2')=1}}\sum_{\widehat{\mu}\in\mathbf{Z}/te_0e_1'e_2'\mathbf{Z}}|\mathcal{K}_{\neq}|, \tag{9.10}$$

where

$$\mathcal{K}_{\neq} := \sum_{\substack{r\sim R_{\neq}\\ s\sim S\\(r,s)=1}}\sum_{m\sim M_{\neq}} a_{m,r,s}\sum_{|j|\sim J}\mathrm{e}(j\omega)\sum_{\substack{k'\\(k',r)=1}} g_0\left(\frac{k'}{K/S}\right)S(m\overline{r},j;k's).$$

This is roughly the sum of Kloosterman sums anticipated in (2.11). By Theorem 3.10, we have

$$\mathcal{K}_{\neq} \ll_{\eta} x^{O(\eta)}\left(1+\frac{K/S}{R_{\neq}\sqrt{S}}\right)^{\theta_{\max}}\|a_{m,r,s}\|_2\sqrt{JR_{\neq}S}\times\left(\frac{K^2/S^2}{R_{\neq}}(M_{\neq}+R_{\neq}S)(J+R_{\neq}S)+M_{\neq}J\right)^{1/2},$$

where, by Cauchy and Schwarz,

$$\sum_{t\sim T}\sum_{e_0\sim E_0}\sum_{\substack{e_1',e_2'\sim E/e_0\\(e_1',e_2')=1}}\sum_{\widehat{\mu}\in\mathbf{Z}/(te_0e_1'e_2'\mathbf{Z})}\|a_{m,r,s}\|_2$$
$$\ll\frac{TE^2}{E_0}\sqrt{\sum_{t\sim T}\sum_{e_0\sim E_0}\sum_{\substack{e_1',e_2'\sim E/e_0\\(e_1',e_2')=1}}\sum_{\substack{s\sim S\\ s|te_0}}\sum_{\substack{\ell_0\sim L_0\\ \ell_1',\ell_2'\sim L/\ell_0\\(\ell_1',\ell_2')=1,\ell_1'\neq\ell_2'\\ \ell_1'\equiv\ell_2' \pmod{te_0}\\(\ell_1',e_1')=(\ell_2',e_2')=1}}\sum_{m\sim M_{\neq}}\left(\sum_{\substack{h_1,h_2\sim H\\ h_1e_1'\ell_2'-h_2e_2'\ell_1'=m}}1\right)^2}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

$$\ll \frac{TE^2}{E_0}\sqrt{L_0\sum_{e_1',e_2'\ll E/E_0}\sum_{\substack{\ell_1',\ell_2'\asymp L/L_0\\(\ell_1'e_2',\ell_2'e_1')=1\\\ell_1'\neq\ell_2'}}\tau(\ell_1'-\ell_2')^3\sum_{m\sim M_{\neq}}\sum_{\substack{h_1,h_2\sim H\\h_1e_1'\ell_2'-h_2e_2'\ell_1'=m}}\sum_{\substack{h_1',h_2'\sim H\\h_1'e_1'\ell_2'-h_2'e_2'\ell_1'=m}}1}$$

$$\ll_\eta x^{O(\eta)}\frac{TE^2}{E_0}\sqrt{L_0\sum_{m\sim M_{\neq}}\sum_{\substack{h_1\sim H\\e_1'\ll E/E_0\\\ell_2'\ll L/L_0}}\sum_{\substack{h_2,e_2',\ell_1'\\h_2e_2'\ell_1'=h_1e_1'\ell_2'-m\\(e_2'\ell_1',e_1'\ell_2')=1}}\sum_{h_1'\sim H\atop h_1'e_1'\ell_2'\equiv m\pmod{e_2'\ell_1'}}\sum_{\substack{h_2'\sim H\\h_2'e_2'\ell_1'=h_1'\ell_2'e_1'-m}}1}$$

$$\ll_\eta x^{O(\eta)}\frac{TE^2}{E_0}\left(L_0M_{\neq}H\frac{E}{E_0}\frac{L}{L_0}\left(1+\frac{HE_0L_0}{EL}\right)\right)^{1/2}$$

$$\ll x^{O(\eta)}\left(\frac{T^2E^5M_{\neq}HL}{E_0^2}\left(1+\frac{HL_0}{EL}\right)\right)^{1/2}.$$

Plugging these bounds into (9.10), we find that

$$\mathcal{B}_{\neq}'\ll_\eta x^{O(\eta)}\Big(1+\frac{K/S}{R_{\neq}\sqrt{S}}\Big)^{\theta_{\max}}\left(\frac{T^2E^5M_{\neq}HL}{E_0^2}\left(1+\frac{HL_0}{EL}\right)\right)^{1/2}(JR_{\neq}S)^{1/2}$$
$$\times\left(\frac{K^2/S^2}{R_{\neq}}(M_{\neq}+R_{\neq}S)(J+R_{\neq}S)+M_{\neq}J\right)^{1/2}.$$

Recalling that $J\ll(KTE^2x^\eta)/(NE_0)$ (from Lemma 9.2), $M_{\neq}\ll HEL/(E_0L_0)$ and $R_{\neq}\asymp x^{O(\eta)}L^2E^2/(L_0E_0^2)$ (from (9.9)), this yields

$$\mathcal{B}_{\neq}'\ll_\eta x^{O(\eta)}\left(1+\frac{KL_0E_0^2}{L^2E^2S^{3/2}}\right)^{\theta_{\max}}\left(\frac{T^2E^5HELHL}{E_0^2E_0L_0}\left(1+\frac{HL_0}{EL}\right)\right)^{1/2}\left(\frac{KTE^2}{NE_0}\frac{L^2E^2}{L_0E_0^2}S\right)^{1/2}$$
$$\times\left(\frac{K^2L_0E_0^2}{L^2E^2S^2}\left(\frac{HEL}{E_0L_0}+\frac{L^2E^2}{L_0E_0^2}S\right)\left(\frac{KTE^2}{NE_0}+\frac{L^2E^2}{L_0E_0^2}S\right)+\frac{HE^3LKT}{E_0^2L_0N}\right)^{1/2}.$$

Note that this expression is nonincreasing in the GCD parameter L_0, since $\theta_{\max}\leqslant 1/2$; thus lower-bounding $L_0\gg 1$, and then using that $1\ll S\ll TE_0$ (from Lemma 9.2), we get

$$\mathcal{B}_{\neq}'\ll_\eta x^{O(\eta)}\left(1+\frac{KE_0^2}{L^2E^2S^{3/2}}\right)^{\theta_{\max}}\left(\frac{T^2E^6H^2L^2}{E_0^3}\left(1+\frac{H}{EL}\right)\right)^{1/2}\left(\frac{KTE^2}{NE_0}\frac{L^2E^2}{E_0^2}\right)^{1/2}$$
$$\times\left(\frac{K^2E_0^2}{L^2E^2}\left(\frac{HEL}{E_0S}+\frac{L^2E^2}{E_0^2}\right)\left(\frac{KTE^2}{NE_0}+\frac{L^2E^2}{E_0^2}S\right)+\frac{HE^3LKT}{E_0^2N}S\right)^{1/2}$$
$$\ll x^{O(\eta)}\left(1+\frac{KE_0^2}{L^2E^2}\right)^{\theta_{\max}}\left(\frac{T^2E^6H^2L^2}{E_0^3}\left(1+\frac{H}{EL}\right)\right)^{1/2}\left(\frac{KTE^2}{NE_0}\frac{L^2E^2}{E_0^2}\right)^{1/2}$$
$$\times\left(\frac{K^2E_0^2}{L^2E^2}\left(\frac{HEL}{E_0}+\frac{L^2E^2}{E_0^2}\right)\left(\frac{KTE^2}{NE_0}+\frac{L^2E^2T}{E_0}\right)+\frac{HE^3LKT^2}{E_0N}\right)^{1/2}.$$

Finally, this expression is nonincreasing in the E_0 parameter even after extracting a factor of E_0^{-1}, so lower-bounding $E_0\gg 1$ yields

$$\mathcal{B}_{\neq}'\ll_\eta\frac{x^{O(\eta)}}{E_0}\left(1+\frac{K}{L^2E^2}\right)^{\theta_{\max}}\left(1+\frac{H}{EL}\right)^{1/2}\left(T^2E^6H^2L^2\frac{KTE^2}{N}L^2E^2\right)^{1/2}$$
$$\times\left(\frac{K^2}{L^2E^2}(HEL+L^2E^2)\left(\frac{KTE^2}{N}+L^2E^2T\right)+\frac{HE^3LKT^2}{N}\right)^{1/2}.$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Due to our assumption that $EHT \ll x^{O(\eta)}KNL$, we have

$$x^{-O(\eta)}\frac{HE^3LKT^2}{N} \ll K^2L^2E^2T = \frac{K^2}{L^2E^2}L^2E^2L^2E^2T,$$

so we may ignore the term of HE^3LKT^2/N on the second line to obtain

$$\begin{aligned}\mathcal{B}'_{\neq} &\ll_\eta \frac{x^{O(\eta)}}{E_0}\left(1+\frac{K}{L^2E^2}\right)^{\theta_{\max}}\left(1+\frac{H}{EL}\right)^{1/2}\left(\frac{KT^3E^{10}H^2L^4}{N}\right)^{1/2}\\ &\quad\times K\left(\frac{H}{LE}+1\right)^{1/2}(L^2E^2T)^{1/2}\left(\frac{K}{NL^2}+1\right)^{1/2}\\ &\ll \frac{x^{O(\eta)}}{E_0}\left(1+\frac{K}{L^2E^2}\right)^{\theta_{\max}}\left(\frac{K^3T^4E^{12}H^2L^6}{N}\right)^{1/2}\left(1+\frac{H}{LE}\right)\left(1+\frac{K}{NL^2}\right)^{1/2}\\ &\ll \frac{x^{O(\eta)}}{E_0}\frac{KE^2}{N}\left(1+\frac{K}{L^2E^2}\right)^{\theta_{\max}}(KT^4E^8H^2L^6N)^{1/2}\left(1+\frac{H}{LE}\right)\left(1+\frac{K}{NL^2}\right)^{1/2}.\end{aligned}$$

Rearranging factors (and combining this with (9.9)), we obtain the desired bound. □

Combining our results so far, we obtain the following general estimate.

Proposition 9.5 (The BFI-style bound with general parameters). *For $\eta \in (0,1)$, $K, N, T, E, H, L \ll x$, and any positive integers $b \ll x^{O(\eta)}$ and $d \ll x$, assuming that $EHT \ll x^{O(\eta)}KNL$, one has*

$$\begin{aligned}\mathcal{B}(K,N,T,E,H,L)^2 \ll_\eta x^{O(\eta)}\Bigg(&N^2K^2H^2T^2L^2 + K^2E^2H^2T^4L^2 + N^2H^4L^4 + N^2H^4T^2L^2\\ &+\left(1+\frac{K^2}{E^4L^4}\right)^{\theta_{\max}}KT^4E^8H^2L^6N\left(1+\frac{H^2}{E^2L^2}\right)\left(1+\frac{K}{NL^2}\right)\\ &+\left(1+\frac{K^2}{E^4L^2}\right)^{\theta_{\max}}KT^5E^8H^2L^3N\left(1+\frac{H}{E}+\frac{H^2}{E^2L}\right)\left(1+\frac{K}{NL}\right)\Bigg).\end{aligned}$$

Proof of Proposition 9.5 assuming Theorem 3.10. This follows by putting together Lemmas 9.2 to 9.4 and squaring (the second line comes from Lemma 9.4, and the third line from Lemma 9.3). □

Finally, we use Proposition 9.5 and the conditions from (4.2) to prove Proposition 8.6.

Proof of Proposition 8.6 assuming Theorem 3.10. Let $\theta := \theta_{\max}$; we will soon pick a value for E such that (5.2) holds. We can assume without loss of generality that $K', N', T', E, H', L' \gg 1$, since otherwise the sum in Proposition 8.6 is void. We now apply the bound in Proposition 9.5 (which is increasing in all six parameters) for the parameters K', N', T', E, H', L' from Proposition 8.6, noting that

$$\begin{aligned}EH'T' &\ll x^{O(\eta)}ERM^{-1}NL(ER)^{-1}\\ &= x^{O(\eta)}NLM^{-1} \ll x^{O(\eta)}NL \ll x^{O(\eta)}K'N'L'.\end{aligned}$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Plugging in the bounds $K' \ll NLx^{O(\eta)}$, $N' \asymp Nx^{O(\eta)}$, $L' \asymp Lx^{O(\eta)}$, $T' \ll NL(RE)^{-1}x^{O(\eta)}$, $H' \ll RNLx^{O(\eta)-1}$, we obtain

$$\begin{aligned}
&\mathcal{B}(K', N', T', E, H', L')^2 \\
&\ll_\eta x^{O(\eta)}\Bigg(N^2(NL)^2\left(\frac{RNL}{x}\right)^2\left(\frac{NL}{RE}\right)^2L^2 + (NL)^2E^2\left(\frac{RNL}{x}\right)^2\left(\frac{NL}{RE}\right)^4L^2 \\
&\quad + N^2\left(\frac{RNL}{x}\right)^4L^4 + N^2\left(\frac{RNL}{x}\right)^4\left(\frac{NL}{RE}\right)^2L^2 \\
&\quad + \left(1+\frac{(NL)^2}{E^4L^4}\right)^\theta NL\left(\frac{NL}{RE}\right)^4E^8\left(\frac{RNL}{x}\right)^2L^6N\left(1+\frac{(RNL/x)^2}{E^2L^2}\right) \\
&\quad + \left(1+\frac{(NL)^2}{E^4L^2}\right)^\theta NL\left(\frac{NL}{RE}\right)^5E^8\left(\frac{RNL}{x}\right)^2L^3N\left(1+\frac{(RNL/x)}{E}+\frac{(RNL/x)^2}{E^2L}\right)\Bigg).
\end{aligned}$$

Simplifying terms and dividing both sides by N^4L^6, we further get

$$\begin{aligned}
\frac{\mathcal{B}(K', N', T', E, H', L')^2}{N^4L^6} \ll_\eta x^{O(\eta)}\Bigg(&\frac{N^4L^2}{x^2E^2} + \frac{N^4L^4}{R^2x^2E^2} + \frac{N^2L^2R^4}{x^4} + \frac{N^4L^2R^2}{x^4E^2} \\
&+ \left(1+\frac{N^2}{E^4L^2}\right)^\theta\frac{N^4L^7E^4}{x^2R^2}\left(1+\frac{N^2R^2}{x^2E^2}\right) \\
&+ \left(1+\frac{N^2}{E^4}\right)^\theta\frac{N^5L^5E^3}{x^2R^3}\left(1+\frac{NLR}{xE}+\frac{N^2LR^2}{x^2E^2}\right)\Bigg),
\end{aligned}$$

and we wish to show that the right-hand side is $\ll x^{O(\eta)-\varepsilon}$. To handle the term of $N^4L^2x^{-2}E^{-2}$, we require that $N^2L \ll x^{1-\varepsilon}E$; thus we pick

$$E := \max\left(x^{4\varepsilon}, \frac{N^2L}{x^{1-\varepsilon}}\right). \tag{9.11}$$

For (5.2) to hold, we also need to have $E \ll x^{-\varepsilon}NLR^{-1}$, so we impose the restrictions

$$R \ll x^{-5\varepsilon}NL \quad \text{and} \quad N \ll \frac{x^{1-2\varepsilon}}{R}$$

(which are part of (4.2)). The fact that $NR \ll x$ simplifies our expression a bit; combined with the fact that $E \ll x^{-\varepsilon}NLR^{-1} \ll NLRx^{-1}$ (due to $x^{1-\varepsilon} \ll R^2$ from (4.2)), this shows that

$$\frac{N^2R^2}{x^2E^2} \ll 1 \quad \text{and} \quad \max\left(\frac{N^2LR^2}{x^2E^2}, 1\right) \ll \frac{NLR}{xE}.$$

Moreover, since $x^{(1-\varepsilon)/2} \ll R$ by (4.2), we have $NR \ll x^{1-2\varepsilon} \ll R^2$, so $N \ll R$, which implies

$$\frac{N^4L^2R^2}{x^4E^2} \ll \frac{N^2L^2R^4}{x^4}.$$

Overall, it remains to bound the expression

$$\frac{N^4L^4}{R^2E^2x^2} + \frac{N^2L^2R^4}{x^4} + \left(1+\frac{N^2}{E^4L^2}\right)^\theta\frac{N^4L^7E^4}{x^2R^2} + \left(1+\frac{N^2}{E^4}\right)^\theta\frac{N^6L^6E^2}{x^3R^2} \tag{9.12}$$

by $O(x^{-\varepsilon})$.

Using $x^{(1-\varepsilon)/2} \ll R \ll NL \ll x^{2/3-6\varepsilon}$ (from (4.2)) and $E \geqslant x^{4\varepsilon}$, the first term is admissible since

$$\frac{N^2L^2}{REx} \ll \frac{x^{4/3}}{x^{3\varepsilon}x^{3/2}} = x^{-1/6-3\varepsilon}.$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

The (square root of the) second term is similarly bounded:

$$\frac{NLR^2}{x^2} \ll x^{(2/3)+(4/3)-2-6\varepsilon} = x^{-6\varepsilon}.$$

For the third term in (9.12), we use our choice of E from (9.11) to obtain

$$\left(1+\frac{N^2}{E^4L^2}\right)^{\theta}\frac{N^4L^7E^4}{x^2R^2} \ll \left(1+\frac{N^2}{L^2}\right)^{\theta}\frac{N^4L^7}{x^{2-16\varepsilon}R^2} + \left(1+\frac{x^4}{N^6L^6}\right)^{\theta}\frac{N^{12}L^{11}}{x^{6-4\varepsilon}R^2}.$$

Since $NL \ll x^{2/3}$ by (4.2), we can ignore the 1-term in the last parenthesis. For the terms above to be admissible, we require the restrictions

$$N^4L^7\max(1,N/L)^{2\theta} \ll x^{2-17\varepsilon}R^2, \quad N^{12-6\theta}L^{11-6\theta} \ll x^{6-4\theta-5\varepsilon}R^2$$

(which are part of (4.2)). Finally, using that $1 \ll E \ll x^{-\varepsilon}NLR^{-1}$ (from (5.2)), we crudely bound the fourth and last term in (9.12) by

$$\left(1+\frac{N^2}{E^4}\right)^{\theta}\frac{N^6L^6E^2}{x^3R^2} \ll N^{2\theta}\frac{N^6L^6(x^{-\varepsilon}NLR^{-1})^2}{x^3R^2} \ll x^{-2\varepsilon}\frac{N^9L^8}{x^3R^4},$$

which is at most $O(x^{-\varepsilon})$ by the last condition in the first line of (4.2). This completes our proof. □

10. Deshouillers–Iwaniec-style estimates

The seminal work [DI82] of Deshouillers and Iwaniec on sums of Kloosterman sums makes repeated use of the Kuznetsov trace formula [Kuz80, Mot97], which is in turn based on the spectral decomposition of $L^2(\Gamma_0(q)\backslash\mathbf{H})$ with respect to the hyperbolic Laplacian (where q is a positive integer and $\Gamma_0(q)$ is its associated Hecke congruence subgroup). Here, we prove Theorem 3.10, which is an optimization of [DI82, Theorem 11] in the θ-aspect, using the same technology. We note that such optimizations of Deshouillers–Iwaniec bounds (specifically of [DI82, Theorem 12]) have also been used in [DPR23].

We will use all of the notation (and normalization) from [DI82], with the exception of making some dependencies on the level q explicit. In particular, we consider an orthonormal basis of Maass cusp forms $(u_{j,q})_{j\geqslant 1}$ such that $u_{j,q}$ has eigenvalue $\lambda_{j,q}$ (which increases to ∞ as $j\to\infty$), and Fourier coefficients $\rho_{j,\mathfrak{a}}(n)$ when expanding around the cusp $\mathfrak{a}$ of $\Gamma_0(q)$, via an implicit scaling matrix $\sigma_{\mathfrak{a}} \in \mathrm{PSL}_2(\mathbf{R})$. We denote

$$\mu(\mathfrak{a}) := \frac{(w, q/w)}{q},$$

whenever $\mathfrak{a}$ is equivalent to u/w, for some relatively prime $u, w \in \mathbf{Z}_+$ such that $w \mid q$; in particular, one has $\mu(\infty) = q^{-1}$. We also write

$$\theta_{j,q} := 2i\kappa_{j,q}, \quad \kappa_{j,q}^2 = \lambda_{j,q} - \tfrac{1}{4} \quad \Longleftrightarrow \quad \theta_{j,q}^2 = 1 - 4\lambda_{j,q},$$

where $\kappa_{j,q}$ is chosen such that either $\kappa_{j,q} \geqslant 0$ (when $\lambda_{j,q} \geqslant 1/4$), or $i\kappa_{j,q} > 0$ (when $\lambda_{j,q}$ is exceptional). Recall from Notation 1.2 that $\theta_q := \max_{\lambda_j<1/4}\theta_{j,q}$ (with $\theta_q := 0$ if there are no exceptional eigenvalues), and that $\theta_{\max} := \sup_q \theta_q$. Also, recall that all exceptional eigenvalues lie in the interval $[3/16, 1/4)$ by [DI82, Theorem 4] (in fact, the best currently known lower bound is $975/4096$, due to Kim and Sarnak [Kim03, Appendix 2]; this is equivalent to Theorem A).

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

The contribution of the exceptional Maass forms to the spectral side of the Kuznetsov trace formula would vanish if Selberg's eigenvalue conjecture (Conjecture 1.3) were true, but would be dominating in most applications otherwise. To deduce better bounds for the geometric side (which consists of weighted sums of Kloosterman sums), Deshouillers and Iwaniec [DI82] proved a series of large sieve inequalities for the Fourier coefficients of Maass cusp forms, which temper this exceptional contribution in bilinear sums. Remarkably, these results make further use of the Kuznetsov formula, applying it back and forth and ultimately reducing to the Weil bound.

LEMMA 10.1 (Large sieve inequalities from [DI82]). *Given $\varepsilon > 0$, $q \in \mathbf{Z}_+$, $N \gg 1$, a complex sequence $(a_n)_{n\sim N}$, a cusp $\mathfrak{a}$ of $\Gamma_0(q)$, and an associated scaling matrix $\sigma_{\mathfrak{a}}$, one has*

$$\sum_{\substack{j\geqslant 1\\ \lambda_{j,q}<1/4}} X^{\theta_{j,q}} \Big| \sum_{n\sim N} a_n\, \rho_{j,\mathfrak{a}}(n)\Big|^2 \ll_\varepsilon (QN)^\varepsilon\, (1+\mu(\mathfrak{a})N)\, \|a_n\|_2^2, \tag{10.1}$$

for any $0 < X \ll \max(1, \mu(\mathfrak{a})^{-1}N^{-1})$. Moreover, if $(\mathfrak{a}, \sigma_{\mathfrak{a}}) = (\infty, \mathrm{Id})$, then given $Q \gg 1$ and $\alpha \in \mathbf{R}/\mathbf{Z}$, one has

$$\frac{1}{Q}\sum_{q\sim Q}\sum_{\substack{j\geqslant 1\\ \lambda_{j,q}<1/4}} X^{\theta_{j,q}} \Big| \sum_{n\sim N} \mathrm{e}(n\omega)\, \rho_{j,\infty}(n)\Big|^2 \ll_\varepsilon (QN)^\varepsilon (1+Q^{-1}N)N, \tag{10.2}$$

in the larger range $0 < X \ll \max(N, Q^2N^{-1})$.

Proof. The bounds in (10.1) and (10.2) follow immediately from [DI82, Theorems 5 and 7], respectively. We note that changing the choice of the scaling matrix $\sigma_{\mathfrak{a}}$ results in multiplying the Fourier coefficients $\rho_{j,\mathfrak{a}}(n)$ by an exponential phase $\mathrm{e}(n\omega)$; thus in (10.2), using an arbitrary value of ω is equivalent to using an arbitrary (but consistent) choice of the scaling matrix σ_∞.

We also remark that the proof of [DI82, Theorem 7] from [DI82, § 8.3] only considers the case $\omega = 0$ (and $\sigma_\infty = \mathrm{Id}$), but the same proof extends to any $\omega \in \mathbf{R}/\mathbf{Z}$ (or equivalently, to any valid scaling matrix σ_∞); this was already noted, for instance, in [BFI87, Lemma 5]. Ultimately, this is because the proof of [DI82, Theorem 14] also extends to sums with additional weights of $\mathrm{e}(m\omega_1)\,\mathrm{e}(n\omega_2)$. □

Remark 10.2. The large sieve inequalities in [DI82] are stated for general values of X on the left-hand sides (resulting in right-hand sides that depend on X), and are equivalent to those given in Lemma 10.1. Indeed, to recover large sieve inequalities with an arbitrary $X > 0$ on the left-hand sides, it suffices to multiply the right-hand sides by $(1 + (X/X_0)^{\theta_q})$, where X_0 is the best allowable value in Lemma 10.1.

We find the versions stated above easier to apply optimally in the θ-aspect, and also easier to compare, by contrasting the maximal permitted values of X (recalling that $\mu(\infty) = q^{-1}$).

We now adapt the proof of [DI82, Theorem 11], making the dependence on $\theta_{\max}$ explicit.

THEOREM 10.3 ([DI82]-type multilinear Kloosterman bound). *Let $C, M, N, R, S \gg 1$, $(b_{n,r,s})$ be a complex sequence, and $\omega \in \mathbf{R}/\mathbf{Z}$. Then given a five-variable smooth function $g(t_1, \dots, t_5)$ with compact support in $t_1 \asymp 1$, and bounded derivatives $\|(\partial^{\sum j_i}/\prod(\partial t_i)^{j_i})g\|_\infty \ll_{j_1,\dots,j_5} 1$, one has*

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

$$\sum_{\substack{r\sim R\\ s\sim S\\ (r,s)=1}}\sum_{\substack{m\sim M\\ n\sim N}}\mathrm{e}(m\omega)\,b_{n,r,s}\sum_{(c,r)=1}g\left(\frac{c}{C},\frac{m}{M},\frac{n}{N},\frac{r}{R},\frac{s}{S}\right)\,S(m\overline{r},\pm n;sc)$$

$$\ll_{\varepsilon}(CMNRS)^{\varepsilon}\left(1+\frac{CS\sqrt{R}}{\max(M,RS)\sqrt{\max(N,RS)}}\right)^{\theta_{\max}}\sqrt{MRS}\,\|b_{n,r,s}\|_2$$

$$\times\frac{\left(CS\sqrt{R}+\sqrt{MN}+C\sqrt{SM}\right)\left(CS\sqrt{R}+\sqrt{MN}+C\sqrt{SN}\right)}{CS\sqrt{R}+\sqrt{MN}}. \tag{10.3}$$

Proof. We follow the proof of [DI82, Theorem 11] in [DI82, § 9.1], reducing to the case of smooth functions of the form $(CS\sqrt{R}/cs\sqrt{r})f(4\pi\sqrt{mn}/cs\sqrt{r})$ (up to using slightly different values of ω and $b_{n,r,s}$); here, $f(t)$ is a smooth function supported in $t\asymp X^{-1}$, for $X:=CS\sqrt{R}/\sqrt{MN}$. After applying the Kuznetsov formula, we bound the contribution of the exceptional spectrum more carefully; as in [DI82, § 9.1], this is given by

$$\mathcal{S}_{\mathrm{exc}}:=CS\sqrt{R}\sum_{\substack{r\sim R\\ s\sim S\\ (r,s)=1}}\sum_{\substack{j\geqslant 1\\ \lambda_{j,rs}<1/4}}\frac{\widehat{f}(\kappa_{j,rs})}{\mathrm{ch}(\pi\kappa_{j,rs})}\left(\sum_{m\sim M}\mathrm{e}(m\omega)\,\overline{\rho_{j,\infty}}(m)\right)\left(\sum_{n\sim N}b'_{n,r,s}\,\rho_{j,1/s}(n)\right),$$

where

$$b'_{n,r,s}:=\mathrm{e}\Big(-n\frac{\overline{s}}{r}\Big)b_{n,r,s}.$$

Using the bounds $\mathrm{ch}(\pi\kappa_{j,rs})\asymp 1$ and

$$|\widehat{f}(\kappa_{j,rs})|\ll\frac{1+X^{2|\kappa_{j,rs}|}}{1+X^{-1}}\ll\frac{(1+X)^{\theta_{j,rs}}}{1+X^{-1}}$$

(see [DI82, (7.1)]), and denoting

$$X_0:=\frac{1+X}{\sqrt{X_1X_2}}\leqslant 1+\frac{X}{\sqrt{X_1X_2}},$$

for some $X_1,X_2\geqslant 1$ to be chosen shortly, we obtain

$$\mathcal{S}_{\mathrm{exc}}\ll CS\sqrt{R}\sum_{\substack{r\sim R\\ s\sim S\\ (r,s)=1}}\sum_{\substack{j\geqslant 1\\ \lambda_{j,rs}<1/4}}\frac{\left(X_0\sqrt{X_1X_2}\right)^{\theta_{j,rs}}}{1+X^{-1}}\left|\sum_{m\sim M}\overline{\mathrm{e}(m\omega)}\,\rho_{j,\infty}(m)\right|\left|\sum_{n\sim N}b'_{n,r,s}\,\rho_{j,1/s}(n)\right|$$

$$\ll CS\sqrt{R}\,\frac{(1+X_0)^{\theta_{\max}}}{1+X^{-1}}\left(\sum_{\substack{r\sim R\\ s\sim S}}\sum_{\substack{j\geqslant 1\\ \lambda_{j,rs}<1/4}}X_1^{\theta_{j,rs}}\left|\sum_{m\sim M}\mathrm{e}(-m\omega)\,\rho_{j,\infty}(m)\right|^2\right)^{1/2}$$

$$\times\left(\sum_{\substack{r\sim R\\ s\sim S\\ (r,s)=1}}\sum_{\substack{j\geqslant 1\\ \lambda_{j,rs}<1/4}}X_2^{\theta_{j,rs}}\left|\sum_{n\sim N}b'_{n,r,s}\,\rho_{j,1/s}(n)\right|^2\right)^{1/2},$$

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

by Cauchy and Schwarz. Recall that $\mu(1/s)=\mu(\infty)=(rs)^{-1}$ since $(r,s)=1$; thus, using the divisor bound and Lemma 10.1, we conclude that

$$\begin{aligned}\mathcal{S}_{\text{exc}} &\ll_\varepsilon (MNRS)^\varepsilon CS\sqrt{R}\,\frac{(1+X_0)^{\theta_{\max}}}{1+X^{-1}}\sqrt{RS}\left(1+\sqrt{\frac{M}{RS}}\right)\sqrt{M}\left(1+\sqrt{\frac{N}{RS}}\right)\|b_{n,r,s}\|_2\\ &\ll (MNRS)^\varepsilon\left(1+\frac{CS\sqrt{R}}{\sqrt{MNX_1X_2}}\right)^{\theta_{\max}}\sqrt{MRS}\,\|b_{n,r,s}\|_2\\ &\quad\times\frac{(CS\sqrt{R}+C\sqrt{SM})(CS\sqrt{R}+C\sqrt{SN})}{CS\sqrt{R}+\sqrt{MN}},\end{aligned}$$

for $X_1=\max(M,R^2S^2M^{-1})$ (coming from (10.2)), and $X_2=\max(1,RSN^{-1})$ (from (10.1)), which gives the desired bound up to minor rearrangements. As in [DI82, (9.4)], the non-exceptional spectrum contributes a similar amount of

$$\ll_\varepsilon (MNRS)^\varepsilon\sqrt{MRS}\,\|b_{n,r,s}\|_2\frac{(CS\sqrt{R}+\sqrt{MN}+C\sqrt{SM})(CS\sqrt{R}+\sqrt{MN}+C\sqrt{SN})}{CS\sqrt{R}+\sqrt{MN}},$$

and putting these together completes our proof. □

Finally, Theorem 3.10 follows almost immediately from Theorem 10.3.

Proof of Theorem 3.10. We swap the m and n variables, and pick the second term in each maximum from (10.3) for an upper bound, resulting in a θ-factor of

$$\left(1+\frac{C}{R\sqrt{S}}\right)^{\theta_{\max}}.$$

We also rewrite the last fraction in (10.3) as

$$CS\sqrt{R}+\sqrt{MN}+C\sqrt{SM}+C\sqrt{SN}+\frac{C^2S\sqrt{MN}}{CS\sqrt{R}+\sqrt{MN}},$$

and we use the lower bound $CS\sqrt{R}+\sqrt{MN}\geqslant CS\sqrt{R}$ in the final term.

To reduce to a smooth function depending only on c, we can take

$$g(t_1,t_2,t_3,t_4,t_5)=g_1(t_1)\,g_2(t_2)\,g_3(t_3)\,g_4(t_4)\,g_5(t_5),$$

for some smooth compactly supported functions g_i, where g_2,g_3,g_4,g_5 are equal to 1 on $[1,2]$. □

Acknowledgements

The author wishes to thank his advisor, James Maynard, for his kind support and guidance, as well as Sary Drappeau, Lasse Grimmelt, Régis de la Bretèche, and the referees, for many helpful comments and suggestions.

Conflicts of interest

None.

Financial support

The author is supported by EPSRC.

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Journal information

Compositio Mathematica is owned by the Foundation Compositio Mathematica and published by the London Mathematical Society in partnership with Cambridge University Press. All surplus income from the publication of *Compositio Mathematica* is returned to mathematics and higher education through the charitable activities of the Foundation, the London Mathematical Society and Cambridge University Press.

References

Bom65 E. Bombieri, *On the large sieve*, Mathematika **12** (1965), 201–225.

BFI86 E. Bombieri, J. B. Friedlander and H. Iwaniec, *Primes in arithmetic progressions to large moduli*, Acta Math. **156** (1986), 203–251.

BFI87 E. Bombieri, J. B. Friedlander and H. Iwaniec, *Primes in arithmetic progressions to large moduli. II*, Math. Ann. **277** (1987), 361–393.

BFI89 E. Bombieri, J. B. Friedlander and H. Iwaniec, *Primes in arithmetic progressions to large moduli. III*, J. Amer. Math. Soc. **2** (1989), 215–224.

dLBD20 R. de La Bretèche and S. Drappeau, *Niveau de répartition des polynômes quadratiques et crible majorant pour les entiers friables*, J. Eur. Math. Soc. (JEMS) **22** (2020), 1577–1624.

DI82 J.-M. Deshouillers and H. Iwaniec, *Kloosterman sums and Fourier coefficients of cusp forms*, Invent. Math. **70** (1982), 219–288.

Dra15 S. Drappeau, *Théorèmes de type Fouvry-Iwaniec pour les entiers friables*, Compositio Math. **151** (2015), 828–862.

DGS17 S. Drappeau, A. Granville and X. Shao, *Smooth-supported multiplicative functions in arithmetic progressions beyond the $x^{1/2}$-barrier*, Mathematika **63** (2017), 895–918.

DPR23 S. Drappeau, K. Pratt and M. Radziwiłł, *One-level density estimates for Dirichlet L-functions with extended support*, Algebra Number Theory **17** (2023), 805–830.

EH68 P. D. T. A. Elliott and H. Halberstam, *A conjecture in prime number theory*, in *Symposia mathematica*, vol. IV (INDAM, Rome, 1968/69) (Academic Press, London–New York, 1968), 59–72.

Fou82 É. Fouvry, *Répartition des suites dans les progressions arithmétiques*, Acta Arith. **41** (1982), 359–382.

Fou84 É. Fouvry, *Autour du théorème de Bombieri-Vinogradov*, Acta Math. **152** (1984), 219–244.

Fou85 É. Fouvry, *Sur le problème des diviseurs de Titchmarsh*, J. Reine Angew. Math. **357** (1985), 51–76.

Fou87 É. Fouvry, *Autour du théorème de Bombieri-Vinogradov. II*, Ann. Sci. École Norm. Supér. (4) **20** (1987), 617–640.

FI80 É. Fouvry and H. Iwaniec, *On a theorem of Bombieri-Vinogradov type*, Mathematika **27** (1980), 135–152.

FI83 É. Fouvry and H. Iwaniec, *Primes in arithmetic progressions*, Acta Arith. **42** (1983), 197–218.

FT91 É. Fouvry and G. Tenenbaum, *Entiers sans grand facteur premier en progressions arithmetiques*, Proc. Lond. Math. Soc. (3) **63** (1991), 449–494.

FT96 É. Fouvry and G. Tenenbaum, *Répartition statistique des entiers sans grand facteur premier dans les progressions arithmétiques*, Proc. Lond. Math. Soc. (3) **72** (1996), 481–514.

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Gra93a A. Granville, *Integers, without large prime factors, in arithmetic progressions. I*, Acta Math. **170** (1993), 255–273.

Gra93b A. Granville, *Integers, without large prime factors, in arithmetic progressions. II*, Philos. Trans. Roy. Soc. A **345** (1993), 349–362.

Har12 A. J. Harper, *Bombieri–Vinogradov and Barban–Davenport–Halberstam type theorems for smooth numbers*, Preprint (2012), arXiv:1208.5992.

Hil86 A. Hildebrand, *On the number of positive integers $\leqslant x$ and free of prime factors $> y$*, J. Number Theory **22** (1986), 289–307.

Iwa85 H. Iwaniec, *Character sums and small eigenvalues for $\Gamma_0(p)$*, Glasgow Math. J. **27** (1985), 99–116.

Iwa90 H. Iwaniec, *Small eigenvalues of Laplacian for $\Gamma_0(N)$*, Acta Arith. **56** (1990), 65–82.

IK21 H. Iwaniec and E. Kowalski, *Analytic number theory*, vol. 53 (American Mathematical Society, Providence, RI, 2021).

IS85 H. Iwaniec and J. Szmidt, *Density theorems for exceptional eigenvalues of the Laplacian for congruence groups*, in *Elementary and analytic theory of numbers (Warsaw, 1982)*, Banach Center Publications, vol. 17 (PWN, Warsaw, 1985), 317–331.

Kim03 H. H. Kim, *Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2*, J. Amer. Math. Soc. **16** (2003), 139–183, with appendix 1 by D. Ramakrishnan and appendix 2 by Kim and P. Sarnak.

Kuz80 N. V. Kuznetsov, *The Petersson conjecture for cusp forms of weight zero and the Linnik conjecture. Sums of Kloosterman sums*, Mat. Sb. **111** (1980), 334–383, 479.

Lin63 J. Ṽ. Linnik, *The dispersion method in binary additive problems* (American Mathematical Society, Providence, RI, 1963), translated by S. Schuur.

LRS95 W. Luo, Z. Rudnick and P. Sarnak, *On Selberg's eigenvalue conjecture*, Geom. Funct. Anal. **5** (1995), 387–401.

May15 J. Maynard, *Small gaps between primes*, Ann. of Math. (2) **181** (2015), 383–413.

May25a J. Maynard, *Small gaps between primes Primes in Arithmetic Progressions to Large Moduli I: Fixed Residue Classes*, Mem. Amer. Math. Soc. **306** (2025), 1542.

May25b J. Maynard, *Small gaps between primes*, *Primes in Arithmetic Progressions to Large Moduli II: Well-Factorable Estimates*, Mem. Amer. Math. Soc. **306** (2025), 1543.

May25c J. Maynard, *Primes in Arithmetic Progressions to Large Moduli III: Uniform Residue Classes*, Mem. Amer. Math. Soc. **306** (2025), 1544.

Mot97 Y. Motohashi, *Spectral theory of the Riemann zeta-function*, Cambridge Tracts in Mathematics, vol. 127 (Cambridge University Press, Cambridge, 1997).

Pol14a D. H. J. Polymath, *New equidistribution estimates of Zhang type*, Algebra Number Theory **8** (2014), 2067–2199.

Pol14b D. H. J. Polymath, *Variants of the Selberg sieve, and bounded intervals containing many primes*, Res. Math. Sci. **1** (2014), 83.

Sar95 P. Sarnak, *Selberg's eigenvalue conjecture*, Notices Amer. Math. Soc. **42** (1995), 1272–1277.

Sel65 A. Selberg, *On the estimation of Fourier coefficients of modular forms*, Proceedings of Symposia in Pure Mathematics, vol. VIII (American Mathematical Society, Providence, RI, 1965), 1–15.

Shp18 I. E. Shparlinski, *Character sums with smooth numbers*, Arch. Math. (Basel) **110** (2018), 467–476.

Vin65 A. I. Vinogradov, *The density hypothesis for Dirichet L-series*, Izv. Math. **29** (1965), 903–934.

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press

Wol73a D. Wolke, *Über die mittlere Verteilung der Werte zahlentheoretischer Funktionen auf Restklassen. I*, Math. Ann. **202** (1973), 1–25.
Wol73b D. Wolke, *Über die mittlere Verteilung der Werte zahlentheoretischer Funktionen auf Restklassen. II*, Math. Ann. **204** (1973), 145–153.
Zha14 Y. Zhang, *Bounded gaps between primes*, Ann. of Math. (2) **179** (2014), 1121–1174.

Alexandru Pascadi alexpascadi@gmail.com
Mathematical Institute, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK

https://doi.org/10.1112/S0010437X2500747X Published online by Cambridge University Press