COMPOSITIO MATHEMATICA

Dense clusters of primes in subsets

James Maynard

Compositio Math. **152** (2016), 1517–1554.

doi:10.1112/S0010437X16007296

FOUNDATION COMPOSITIO MATHEMATICA

LONDON MATHEMATICAL SOCIETY
EST. 1865

Compositio Math. 152 (2016) 1517–1554
doi:10.1112/S0010437X16007296

Dense clusters of primes in subsets

James Maynard

Abstract

We prove a generalization of the author's work to show that any subset of the primes which is 'well distributed' in arithmetic progressions contains many primes which are close together. Moreover, our bounds hold with some uniformity in the parameters. As applications, we show there are infinitely many intervals of length $(\log x)^\epsilon$ containing $\gg_\epsilon \log\log x$ primes, and show lower bounds of the correct order of magnitude for the number of strings of m congruent primes with $p_{n+m} - p_n \leqslant \epsilon \log x$.

1. Introduction

Let $\mathcal{L} = \{L_1, \dots, L_k\}$ be a set of distinct linear functions $L_i(n) = a_i n + b_i$ $(1 \leqslant i \leqslant k)$ with coefficients in the positive integers. We say such a set is *admissible* if $\prod_{i=1}^k L_i(n)$ has no fixed prime divisor (that is, for every prime p there is an integer n_p such that $\prod_{i=1}^k L_i(n_p)$ is coprime to p). Dickson made the following conjecture.

Conjecture (Prime k-tuples conjecture). Let $\mathcal{L} = \{L_1, \dots, L_k\}$ be admissible. Then there are infinitely many integers n such that all $L_i(n)$ $(1 \leqslant i \leqslant k)$ are prime.

Although such a conjecture appears well beyond the current techniques, recent progress ([Zha14, May15], and unpublished work of Tao) has enabled us to prove weak forms of this conjecture, where instead we show that there are infinitely many integers n such that *several* (rather than *all*) of the $L_i(n)$ are primes.

As noted in [May15], the method of Maynard and Tao can also prove such weak versions of Dickson's conjecture in various more general settings. This has been demonstrated in various recent works [Tho14, CHLPT15, BFT15, Pol14, LP00]. In this paper we consider generalized versions of Dickson's conjecture, and prove corresponding weak versions of them.

Based on heuristics from the Hardy–Littlewood circle method, it has been conjectured that the number of $n \leqslant x$ such that all the $L_i(n)$ are prime should have an asymptotic formula $(\mathfrak{S}(\mathcal{L}) + o(1))x/(\log x)^k$, where $\mathfrak{S}(\mathcal{L})$ is a constant depending only on $\mathcal{L}$ (with $\mathfrak{S}(\mathcal{L}) > 0$ if and only if $\mathcal{L}$ is admissible). Moreover, these heuristics would suggest that the formulae should hold even if we allow the coefficients a_i, b_i and the number k of functions in $\mathcal{L}$ to vary slightly with x.

One can also speculate that Dickson's conjecture might hold for more general sets, where we ask for infinitely many integers $n \in \mathcal{A}$ such that all of $L_i(n)$ are primes in $\mathcal{P}$, for some 'nice' sets of integers $\mathcal{A}$ and of primes $\mathcal{P}$, and provided $\mathcal{L}$ satisfies some simple properties in terms of $\mathcal{A}$ and $\mathcal{P}$. For example, Schinzel's Hypothesis H would imply this if either $\mathcal{A}$ or $\mathcal{P}$ were restricted to the values given by an irreducible polynomial, and a uniform version of Dickson's conjecture would give this if $\mathcal{A}$ or $\mathcal{P}$ were restricted to the union of short intervals.

Received 31 July 2014, accepted in final form 15 October 2015, published online 1 April 2016.
2010 Mathematics Subject Classification 11N05, 11N35 (primary).
Keywords: prime numbers, sieve methods.

The aim of this paper is to show that the flexibility of the method introduced in [May15] allows us to prove weak analogues of these generalizations of Dickson's conjecture. In particular, if $\mathcal{A}$ and $\mathcal{P}\cap L(\mathcal{A})$ are well distributed in arithmetic progressions, then we can obtain a lower bound close to the expected truth for the number of $n\in\mathcal{A}, n\leqslant x$ such that several of the $L_i(n)$ are primes in $\mathcal{P}$, and we can show this estimate holds with some uniformity in the size of a_i, b_i and k.

2. Well-distributed sets

Given a set of integers $\mathcal{A}$, a set of primes $\mathcal{P}$, and a linear function $L(n)=l_1n+l_2$, we define

$$\begin{gathered}\mathcal{A}(x)=\{n\in\mathcal{A}: x\leqslant n<2x\},\quad \mathcal{A}(x;q,a)=\{n\in\mathcal{A}(x), n\equiv a \pmod q\},\\ L(\mathcal{A})=\{L(n):n\in\mathcal{A}\},\quad \varphi_L(q)=\varphi(|l_1|q)/\varphi(|l_1|),\\ \mathcal{P}_{L,\mathcal{A}}(x)=L(\mathcal{A}(x))\cap\mathcal{P},\quad \mathcal{P}_{L,\mathcal{A}}(x;q,a)=L(\mathcal{A}(x;q,a))\cap\mathcal{P}.\end{gathered}\tag{2.1}$$

This paper will focus on sets which satisfy the following hypothesis, which is given in terms of $(\mathcal{A},\mathcal{L},\mathcal{P},B,x,\theta)$ for $\mathcal{L}$ an admissible set of linear functions, $B\in\mathbb{N}$, x a large real number, and $0<\theta<1$.

Hypothesis 1. $(\mathcal{A},\mathcal{L},\mathcal{P},B,x,\theta)$. Let $k=\#\mathcal{L}$.

(1) $\mathcal{A}$ is well distributed in arithmetic progressions: we have

$$\sum_{q\leqslant x^\theta}\max_a\left|\#\mathcal{A}(x;q,a)-\frac{\#\mathcal{A}(x)}{q}\right|\ll\frac{\#\mathcal{A}(x)}{(\log x)^{100k^2}}.$$

(2) Primes in $L(\mathcal{A})\cap\mathcal{P}$ are well distributed in most arithmetic progressions: for any $L\in\mathcal{L}$ we have

$$\sum_{\substack{q\leqslant x^\theta\\ (q,B)=1}}\max_{(L(a),q)=1}\left|\#\mathcal{P}_{L,\mathcal{A}}(x;q,a)-\frac{\#\mathcal{P}_{L,\mathcal{A}}(x)}{\varphi_L(q)}\right|\ll\frac{\#\mathcal{P}_{L,\mathcal{A}}(x)}{(\log x)^{100k^2}}.$$

(3) $\mathcal{A}$ is not too concentrated in any arithmetic progression: for any $q<x^\theta$ we have

$$\#\mathcal{A}(x;q,a)\ll\frac{\#\mathcal{A}(x)}{q}.$$

We expect to be able to show this hypothesis holds (for all large x, some fixed $\theta>0$ and some $B<x^{O(1)}$ with few prime factors) for sets $\mathcal{A},\mathcal{P}$ where we can establish 'Siegel–Walfisz' type asymptotics for arithmetic progressions to small moduli, and a large sieve estimate to handle larger moduli.

We note that the recent work of Benatar [Ben00] showed the existence of small gaps between primes for sets which satisfy similar properties to those considered here.

3. Main results

Theorem 3.1. *Let $\alpha>0$ and $0<\theta<1$. Let $\mathcal{A}$ be a set of integers, $\mathcal{P}$ a set of primes, $\mathcal{L}=\{L_1,\dots,L_k\}$ an admissible set of k linear functions, and B,x integers. Let the coefficients $L_i(n)=a_in+b_i\in\mathcal{L}$ satisfy $0\leqslant a_i,b_i\leqslant x^\alpha$ for all $1\leqslant i\leqslant k$, and let $k\leqslant(\log x)^\alpha$ and $1\leqslant B\leqslant x^\alpha$.*

There is a constant C depending only on α and θ such that the following holds. If $k \geqslant C$ and $(\mathcal{A}, \mathcal{L}, \mathcal{P}, B, x, \theta)$ satisfy Hypothesis 1, and if $\delta > (\log k)^{-1}$ is such that

$$\frac{1}{k}\frac{\varphi(B)}{B}\sum_{L\in\mathcal{L}}\frac{\varphi(a_i)}{a_i}\#\mathcal{P}_{L,\mathcal{A}}(x) \geqslant \delta\frac{\#\mathcal{A}(x)}{\log x},$$

then

$$\#\{n\in\mathcal{A}(x) : \#(\{L_1(n),\dots,L_k(n)\}\cap\mathcal{P}) \geqslant C^{-1}\delta\log k\} \gg \frac{\#\mathcal{A}(x)}{(\log x)^k\exp(Ck)}.$$

Moreover, if $\mathcal{P}=\mathbb{P}$, $k \leqslant (\log x)^{1/5}$ and all $L\in\mathcal{L}$ have the form $an+b_i$ with $|b_i| \leqslant (\log x)k^{-2}$ and $a \ll 1$, then the primes counted above can be restricted to be consecutive, at the cost of replacing $\exp(Ck)$ with $\exp(Ck^5)$ in the bound.

All implied constants in Theorem 3.1 are effectively computable if the implied constants in Hypothesis 1 for $(\mathcal{A}, \mathcal{L}, \mathcal{P}, B, x, \theta)$ are.

We note that Theorem 3.1 can show that several of the $L_i(n)$ are primes for sets $\mathcal{A}$, $\mathcal{P}$ where it is *not* the case that there are infinitely many $n\in\mathcal{A}$ such that *all* of the $L_i(n)$ are primes in $\mathcal{P}$. For example, if $\mathcal{P}=\{p_{2n} : n\in\mathbb{N}\}$ is the set of primes of even index and $\mathcal{A}=\mathbb{N}$, then we would expect $\mathcal{P}$ to be equidistributed in the sense of Hypothesis 1. However, there are clearly no integers n such that $n, n+2\in\mathcal{P}$, and so the analogue of the twin prime conjecture does not hold in this case. Similarly if $\mathcal{P}$ is restricted to the union of arithmetic progressions in short intervals.[1] Therefore without extra assumptions on our sets $\mathcal{A}$, $\mathcal{P}$ we cannot hope for a much stronger statement than that several of the $L_i(n)$ are primes in $\mathcal{P}$.

We also note that Theorem 3.1 can apply to very sparse sets $\mathcal{A}$, and no density assumptions are required beyond the estimates of Hypothesis 1. Of course, for such sets the major obstacle is in establishing Hypothesis 1.

We give some applications of this result.

THEOREM 3.2. *For any $x, y \geqslant 1$ there are $\gg x\exp(-\sqrt{\log x})$ integers $x_0\in[x,2x]$ such that*

$$\pi(x_0+y)-\pi(x_0) \gg \log y.$$

Theorem 3.2 is non-trivial in the region $y=o(\log x)$ (and y sufficiently large), when typically there are no primes in the interval $[x, x+y]$. For such values of y, it shows that there are many intervals of length y containing considerably more than the typical number of primes. By comparison, a uniform version of the prime k-tuples conjecture would suggest that for small y there are intervals $[x, x+y]$ containing $\gg y/\log y$ primes. For large fixed y, we recover the main result of [May15], that $\liminf_n(p_{n+m}-p_n) \ll_m 1$ for all m.

THEOREM 3.3. *Fix $\epsilon > 0$ and let $x > x_0(\epsilon, q)$. There is a constant $c_\epsilon > 0$ (depending only on ϵ) such that uniformly for $m \leqslant c_\epsilon\log\log x$, $q \leqslant (\log x)^{1-\epsilon}$ and $(a,q)=1$ we have*

$$\#\{p_n \leqslant x : p_n \equiv \cdots \equiv p_{n+m} \equiv a \pmod q, p_{n+m}-p_n \leqslant \epsilon\log x\} \gg_\epsilon \frac{\pi(x)}{(2q)^{\exp(Cm)}}.$$

Here $C > 0$ is a fixed constant.

[1] For example, one could take $\mathcal{P}=\bigcup_{x=2^j}\bigcup_{i\leqslant x^{1/4}/2}\{x+(2i-1)x^{3/4} < p \leqslant x+2ix^{3/4} : n\equiv i \pmod 5\}$. This set is equidistributed in the sense of Hypothesis 1, but also has no gaps of size 2.

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

Theorem 3.3 extends a result of Shiu [Shi00] who showed the same result but with a lower bound $\gg x^{1-\varepsilon(x)}$ for $\varepsilon(x) \approx C_q m(\log\log x)^{-1/\varphi(q)}$ in the shorter range $m \ll (\log\log x)^{1/\varphi(q)-\epsilon}$ and without the constraint $p_{n+m} - p_n \leqslant \epsilon \log x$.

We see that for fixed m, q, Theorem 3.3 shows a positive proportion of primes p_n are counted (and so our lower bound is of the correct order of magnitude). In particular, for a positive proportion of primes p_n we have $p_n \equiv p_{n+1} \equiv \cdots \equiv p_{n+m} \equiv a \pmod q$ and $p_{n+m} - p_n \leqslant \epsilon \log p_n$. This disproves the conjecture of Knapowski and Turán [KT77] that $\#\{p_n \leqslant x : p_n \equiv p_{n+1} \equiv 1 \pmod 4\} = o(\pi(x))$. It also extends a result of Goldston *et al.* [GPY11], who showed a positive proportion of primes p_n have $p_{n+1} - p_n \leqslant \epsilon \log p_n$, and of Freiberg [Fre11] who showed at least $x^{1-c/\log\log x}$ primes $p_n \leqslant x$ have $p_{n+1} - p_n \leqslant \epsilon \log p_n$ and $p_{n+1} \equiv p_n \equiv a \pmod q$.

Theorem 3.4. *Fix $m \in \mathbb{N}$ and $\epsilon > 0$. There exists a $k = k_m \leqslant \exp(Cm)$ such that for $x > x_0(\epsilon, m)$ and $x^{7/12+\epsilon} \leqslant y \leqslant x$ and for any admissible set $\mathcal{L} = \{L_1, \dots, L_k\}$ where $L_i(n) = a_i n + b_i$ with $a_i \ll (\log x)^{1/\epsilon}$ and $b_i \ll x$, we have*

$$\#\{n \in [x, x+y] : \text{at least } m \text{ of } L_i(n) \text{ are prime}\} \gg \frac{y}{(\log x)^k}.$$

Here $C > 0$ is a fixed constant.

Theorem 3.4 relies on a Bombieri–Vinogradov type theorem for primes in intervals of length $x^{7/12+\epsilon}$, the best such result being due to Timofeev [Tim87]. By adapting Hypothesis 1 to allow for weighted sums instead of $\#\mathcal{P}_{L,\mathcal{A}}(x)$, we could presumably use the results of [HWW04] and [Kum02] to extend this to the wider range $x^{0.525} \leqslant y \leqslant x$.

Theorem 3.4 explicitly demonstrates the claim from [May15] that the method also shows the existence of bounded gaps between primes in short intervals, and for linear functions. We note that we would expect the lower bound to be of size $y/(\log x)^m$, and so our bound is smaller than the expected truth by a factor of a fixed power of $\log x$. It appears such a loss is an unavoidable feature of the method when looking at bounded length intervals.

Our final application uses Theorem 3.1 to apply to a subset $\mathcal{P}$ of the primes. This extends the result of Thorner [Tho14] to sets of linear functions, and with an explicit lower bound.

Theorem 3.5. *Let $K/\mathbb{Q}$ be a Galois extension of $\mathbb{Q}$ with discriminant Δ_K. There exists a constant C_K depending only on K such that the following holds. Let $\mathcal{C} \subseteq \mathrm{Gal}(K/\mathbb{Q})$ be a conjugacy class in the Galois group of $K/\mathbb{Q}$, and let*

$$\mathcal{P} = \left\{p \text{ prime} : p \nmid \Delta_K, \left[\frac{K/\mathbb{Q}}{p}\right] = \mathcal{C}\right\},$$

where $\left[\frac{K/\mathbb{Q}}{\cdot}\right]$ denotes the Artin symbol. Let $m \in \mathbb{N}$ and $k = \exp(C_K m)$. For any fixed admissible set $\mathcal{L} = \{L_1, \dots, L_k\}$ of k linear functions $L_i(n) = a_i n + b_i$ with $(a_i, \Delta_K) = 1$ for each $1 \leqslant i \leqslant k$, we have

$$\#\{x \leqslant n \leqslant 2x : \text{at least } m \text{ of } L_1(n), \dots, L_k(n) \text{ are in } \mathcal{P}\} \gg \frac{x}{(\log x)^{\exp(C_K m)}},$$

provided $x \geqslant x_0(K, \mathcal{L})$.

Thorner gives several arithmetic consequences of finding such primes of a given splitting type; we refer the reader to the paper [Tho14] for such applications.

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

As with Theorem 3.4, we only state the result for fixed m, because it relies on other work which establishes the Bombieri–Vinogradov type estimates of Hypothesis 1, and these results only save an arbitrary power of $\log x$. One would presume these results can be extended to save $\exp(-c\sqrt{\log x})$ or similar (having excluded some possible bad moduli), which would allow uniformity for $m \leqslant \epsilon \log\log x$, but we do not pursue this here. Similarly, the implied constant in the lower bounds of Theorems 3.4 and 3.5 is not effective as stated, but presumably a small modification to the underlying results would allow us to obtain an effective bound.

4. Notation

We shall view $0 < \theta < 1$ and $\alpha > 0$ as fixed real constants. All asymptotic notation such as $O(\cdot), o(\cdot), \ll, \gg$ should be interpreted as referring to the limit $x \to \infty$, and any constants (implied by $O(\cdot)$ or denoted by c, C with subscripts) may depend on θ, α but no other variable, unless otherwise noted. We will adopt the main assumptions of Theorem 3.1 throughout. In particular, we will view $\mathcal{A}$, $\mathcal{P}$ as given sets of integers and primes respectively and $k = \#\mathcal{L}$ will be the size of $\mathcal{L} = \{L_1, \dots, L_k\}$ an admissible set of integer linear functions, and the coefficients $a_i, b_i \in \mathbb{Z}$ of $L_i(n) = a_i n + b_i$ will satisfy $|a_i|, |b_i| \leqslant x^\alpha$ and $a_i \neq 0$. $B \leqslant x^\alpha$ will be an integer, and x, k will always be assumed sufficiently large (in terms of θ, α).

All sums, products and suprema will be assumed to be taken over variables lying in the natural numbers $\mathbb{N} = \{1, 2, \dots\}$ unless specified otherwise. The exception to this is when sums or products are over a variable p (or p'), which instead will be assumed to lie in the prime numbers $\mathbb{P} = \{2, 3, \dots\}$.

Throughout the paper, φ will denote the Euler totient function, $\tau_r(n)$ the number of ways of writing n as a product of r natural numbers, and μ the Möbius function. We let $\#\mathcal{A}$ denote the number of elements of a finite set $\mathcal{A}$, and $\mathbf{1}_{\mathcal{A}}(x)$ the indicator function of $\mathcal{A}$ (so $\mathbf{1}_{\mathcal{A}}(x) = 1$ if $x \in \mathcal{A}$, and 0 otherwise). We let (a, b) be the greatest common divisor of integers a and b, and $[a, b]$ the least common multiple of integers a and b. (For real numbers x, y we also use $[x, y]$ to denote the closed interval. The usage of $[\cdot, \cdot]$ should be clear from the context.)

To simplify notation we will use vectors in a way which is somewhat non-standard. $\mathbf{d}$ will denote a vector $(d_1, \dots, d_k) \in \mathbb{N}^k$. Given a vector $\mathbf{d}$, when it does not cause confusion, we write $d = \prod_{i=1}^k d_i$. Given $\mathbf{d}, \mathbf{e}$, we will let $[\mathbf{d}, \mathbf{e}] = \prod_{i=1}^k [d_i, e_i]$ be the product of least common multiples of the components of $\mathbf{d}, \mathbf{e}$, and similarly let $(\mathbf{d}, \mathbf{e}) = \prod_{i=1}^k (d_i, e_i)$ be the product of greatest common divisors of the components, and $\mathbf{d}|\mathbf{e}$ denote the k conditions $d_i | e_i$ for each $1 \leqslant i \leqslant k$. An unlabelled sum $\sum_{\mathbf{d}}$ should be interpreted as being over all $\mathbf{d} \in \mathbb{N}^k$.

Finally, for reference we list here the key quantities used in the proof (apart from the previously introduced $\mathcal{A}$, $\mathcal{L}$, $\mathcal{P}$, B, x, θ) and where they are first introduced.

ω, φ_ω, $\mathfrak{S}_D(\mathcal{L})$:	equations (7.1)–(7.3),
W, w_n:	equations (7.4), (7.5),
$\mathcal{D}_k$, W_j:	equation (7.7),
$\lambda_{\mathbf{d}}$, $y_{\mathbf{r}}$:	equation (7.8),
R:	inequality (7.9),
F, T_k, U_k:	equation (7.10),
F_1, F_2:	equation (7.12),
$Y_{\mathbf{r}}$:	Lemma 8.2,
$I_k(\cdot), J_k(\cdot)$:	Lemma 8.6.

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

5. Outline

The methods of this paper are based on the 'GPY method' for detecting primes. The GPY method works by considering a weighted sum associated to an admissible set $\mathcal{L} = \{L_1, \dots, L_k\}$,

$$S = \sum_{x \leqslant n \leqslant 2x} \bigg(\sum_{i=1}^{k} \mathbf{1}_{\mathbb{P}}(L_i(n)) - m \bigg) w_n, \tag{5.1}$$

where m and k are fixed integers, x is a large positive number and w_n are some non-negative weights (typically chosen to be of the form of the weights in Selberg's Λ^2 sieve).

If $S > 0$, then at least one integer n must make a positive contribution to S. Since the weights w_n are non-negative, if n makes a positive contribution then the term in parentheses in (5.1) must be positive at n, and so at least $m + 1$ of the $L_i(n)$ must be prime. Thus to show at least $m + 1$ of the $L_i(n)$ are simultaneously prime infinitely often, it suffices to show that $S > 0$ for all large x.

The shape of S means that one can consider the terms weighted by $\mathbf{1}_{\mathbb{P}}(L_i(n))$ separately for each $L_i \in \mathcal{L}$, which makes these terms feasible to estimate accurately using current techniques. In particular, the only knowledge about the joint behaviour of the prime values of the L_i is derived from the pigeonhole principle described above.

The method only succeeds if the weights w_n are suitably concentrated on integers n when many of the $L_i(n)$ are prime. To enable an unconditional asymptotic estimate for S, the w_n are typically chosen to mimic sieve weights, and in particular Selberg sieve weights (which tend to be the best performing weights when the 'dimension' k of the sieve is large). One can then hope to estimate a quantity involving such sieve weights provided one can prove suitable equidistribution results in arithmetic progressions. The strength of concentration of the weights w_n on primes depends directly on the strength of equidistribution results available.

The original work of Goldston, Pintz and Yıldırım [GPY09] showed that one could construct weights w_n which would show that $S > 0$ for $m = 1$ (and for k sufficiently large) if one could prove a suitable extension of the Bombieri–Vinogradov theorem. Zhang [Zha14] succeeded in proving such an extension,[2] and as a consequence showed the existence of bounded gaps between primes.

The author's work [May15] introduced a modification to the choices of the sieve weights w_n (this modification was also independently discovered by Terence Tao at the same time). This modification enables w_n to be rather more concentrated on n for which many of the $L_i(n)$ are prime. This allows one to show $S > 0$ for any $m \in \mathbb{N}$, and moreover the method works even if one has much more limited knowledge about primes in arithmetic progressions.

As remarked in [May15], the fact that the method now works even with only a limited amount of knowledge about primes in arithmetic progressions makes it rather flexible, and in particular applicable to counting primes in subsets, where we have more limited equidistribution results. Moreover, it is possible to exploit the flexibility of the pigeonhole principle set-up in (5.1) to consider slightly more exotic combinations, which can ensure that the n making a positive contribution to S also satisfy 'typical' properties.

Therefore we can consider modified sums of the form

$$S = \sum_{n \in \mathcal{A}(x)} \bigg(\sum_{i=1}^{k} \mathbf{1}_{\mathcal{P}}(L_i(n)) - m - k\mathbf{1}_{\mathcal{B}}(n) \bigg) w_n$$

[2] The actual form of Zhang's extension is slightly weaker than that considered in the original conditional result of Goldston, Pintz and Yıldırım, although it is sufficient for the argument.

for some set of integers $\mathcal{A}$, set of primes $\mathcal{P}$ and set of 'atypical' integers $\mathcal{B}$. Provided we have some weak distribution results available (such as those asserted by Hypothesis 1), we can estimate all the terms involved in this sum. Again, by the pigeonhole principle, we see that if $n \in \mathcal{A}(x)$ makes a positive contribution to S, then at least $m+1$ of the $L_i(n)$ are primes in $\mathcal{P}$, and that $n \notin \mathcal{B}$. We expect that if $\mathcal{B}$ represents an 'atypical' set, and $\mathcal{P}$ is not too sparse (relative to $\mathcal{A}$), then we can choose w_n similarly to before and show that $S > 0$ for k sufficiently large. Moreover, by modifying some of the technical aspects of the method in [May15], we can obtain suitable uniform estimates for such sums S even when we allow the coefficients a_i, b_i of $L_i(n) = a_i n + b_i$, the number k of functions and the number m of primes we find to vary with x in certain ranges.

Our work necessarily builds on previous work in [May15], and a certain degree of familiarity with [May15] is assumed.

6. Proof of Theorems 3.1–3.5

The proof of Theorems 3.1–3.5 relies on the following key proposition.

Proposition 6.1. *Let $\alpha > 0$ and $0 < \theta < 1$. Let $\mathcal{A}$ be a set of integers, $\mathcal{P}$ a set of primes, $\mathcal{L} = \{L_1, \dots, L_k\}$ an admissible set of k linear functions, and B, x integers. Assume that the coefficients $L_i(n) = a_i n + b_i \in \mathcal{L}$ satisfy $1 \leqslant a_i, b_i \leqslant x^\alpha$ for all $1 \leqslant i \leqslant k$, and that $k \leqslant (\log x)^{1/5}$ and $1 \leqslant B \leqslant x^\alpha$. Let $x^{\theta/10} \leqslant R \leqslant x^{\theta/3}$. Let ρ, ξ satisfy $k(\log\log x)^2/(\log x) \leqslant \rho, \xi \leqslant \theta/10$, and define*

$$\mathcal{S}(\xi; D) = \{n \in \mathbb{N} : p|n \implies (p > x^\xi \text{ or } p|D)\}.$$

There is a constant C depending only on α and θ such that the following holds. If $k \geqslant C$ and $(\mathcal{A}, \mathcal{L}, \mathcal{P}, B, x, \theta)$ satisfy Hypothesis 1, then there is a choice of non-negative weights $w_n = w_n(\mathcal{L})$ satisfying

$$w_n \ll (\log R)^{2k} \prod_{i=1}^{k} \prod_{p|L_i(n), p\nmid B} 4$$

such that the following statements hold.

(1) *We have*

$$\sum_{n \in \mathcal{A}(x)} w_n = \left(1 + O\left(\frac{1}{(\log x)^{1/10}}\right)\right) \frac{B^k}{\varphi(B)^k} \mathfrak{S}_B(\mathcal{L}) \#\mathcal{A}(x) (\log R)^k I_k.$$

(2) *For $L(n) = a_L n + b_L \in \mathcal{L}$ we have*

$$\sum_{n \in \mathcal{A}(x)} \mathbf{1}_{\mathcal{P}}(L(n)) w_n \geqslant \left(1 + O\left(\frac{1}{(\log x)^{1/10}}\right)\right) \frac{B^{k-1}}{\varphi(B)^{k-1}} \mathfrak{S}_B(\mathcal{L}) \frac{\varphi(a_L)}{a_L} \#\mathcal{P}_{L,\mathcal{A}}(x) (\log R)^{k+1} J_k$$
$$+ O\left(\frac{B^k}{\varphi(B)^k} \mathfrak{S}_B(\mathcal{L}) \#\mathcal{A}(x) (\log R)^{k-1} I_k\right).$$

(3) *For $L = a_0 n + b_0 \notin \mathcal{L}$ and $D \ll x^{O(1)}$, let $\Delta_L = a_0 \prod_{j=1}^{k} |a_0 b_j - b_0 a_j|$. If $\Delta_L \neq 0$ we have*

$$\sum_{n \in \mathcal{A}(x)} \mathbf{1}_{\mathcal{S}(\xi;D)}(L(n)) w_n \ll \xi^{-1} \frac{\Delta_L}{\varphi(\Delta_L)} \frac{D}{\varphi(D)} \frac{B^k}{\varphi(B)^k} \mathfrak{S}_B(\mathcal{L}) \#\mathcal{A}(x) (\log R)^{k-1} I_k.$$

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

(4) *For $L \in \mathcal{L}$ we have*

$$\sum_{n\in\mathcal{A}(x)}\Big(\sum_{\substack{p|L(n)\\ p<x^\rho\\ p\nmid B}}1\Big)w_n \ll \rho^2k^4(\log k)^2\frac{B^k}{\varphi(B)^k}\mathfrak{S}_B(\mathcal{L})\#\mathcal{A}(x)(\log R)^kI_k.$$

Here I_k, J_k are quantities depending only on k, and $\mathfrak{S}_B(\mathcal{L})$ is a quantity depending only on $\mathcal{L}$, and these satisfy

$$\begin{aligned}\mathfrak{S}_B(\mathcal{L}) &= \prod_{p\nmid B}\Big(1-\frac{\#\{1\leqslant n\leqslant p : p|\prod_{i=1}^k L_i(n)\}}{p}\Big)\Big(1-\frac{1}{p}\Big)^{-k} \gg \frac{1}{\exp(O(k))},\\ J_k &\gg \frac{\log k}{k}I_k \gg (2k\log k)^{-k-1}.\end{aligned}$$

Moreover, if all functions $L\in\mathcal{L}$ are of the form $L=an+b_L$, for some fixed a and $b_L \ll \log x/(k\log k)$, then for $\eta \geqslant (\log x)^{-9/10}$, we have

$$\sum_{\substack{b\ll\eta\log x\\ L(n)=an+b}}\frac{\Delta_L}{\varphi(\Delta_L)} \ll \eta(\log x)(\log k).$$

Here the implied constants depend only on θ,α, and the implied constants from Hypothesis 1.

Assuming Proposition 6.1, we now establish Theorems 3.1–3.5 in turn.

Proof of Theorem 3.1. We first note that by passing to a subset of $\mathcal{L}$, it is sufficient to show that in the restricted range $C\leqslant k\leqslant(\log x)^{1/5}$ we have the weaker bound

$$\#\{n\in\mathcal{A}(x) : \#(\{L_1(n),\dots,L_k(n)\}\cap\mathcal{P})\geqslant C^{-1}\delta\log k\} \gg \frac{\#\mathcal{A}(x)}{(\log x)^k\exp(Ck^5)}. \tag{6.1}$$

The main result then follows with a suitably adjusted value of C.

For $m\in\mathbb{N}$, we consider the sum

$$S=\sum_{n\in\mathcal{A}(x)}\Big(\sum_{i=1}^k\mathbf{1}_{\mathcal{P}}(L_i(n))-m-k\sum_{i=1}^k\sum_{\substack{p|L_i(n)\\ p<x^\rho\\ p\nmid B}}1\Big)w_n = S_1-S_2-S_3, \tag{6.2}$$

where w_n are the weights whose existence is guaranteed by Proposition 6.1. We note that for any $n\in\mathcal{A}(x)$, the term in parentheses in (6.2) is positive only if at least $m+1$ of the $L_i(n)$ are primes in $\mathcal{P}$, and none of the $L_i(n)$ have any prime factors $p\nmid B$ less than x^ρ. Moreover, we see that if this is the case then since $a_i,b_i\leqslant x^\alpha$, each $L_i(n)$ can have at most $O(1/\rho)$ prime factors $p\nmid B$, and so

$$w_n \ll (\log x)^{2k}\prod_{i=1}^k\prod_{\substack{p|L_i(n)\\ p\nmid B}}4 \ll (\log x)^{2k}\exp(O(k/\rho)). \tag{6.3}$$

Since the term in parentheses in (6.2) can be at most k, we have that

$$\#\{n\in\mathcal{A}(x) : \#(\{L_1(n),\dots,L_k(n)\}\cap\mathcal{P})\geqslant m\} \gg \frac{S}{k(\log x)^{2k}\exp(O(k/\rho))}. \tag{6.4}$$

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

Thus it is sufficient to obtain a suitable lower bound for S. (Essentially the same idea has been used by Goldston *et al.* [GPY11].) Using Proposition 6.1, we have

$$S_1=\sum_{n\in\mathcal{A}(x)}\sum_{i=1}^k \mathbf{1}_{\mathcal{P}}(L_i(n))w_n \geqslant (1+o(1))\frac{B^{k-1}}{\varphi(B)^{k-1}}\mathfrak{S}_B(\mathcal{L})(\log R)^{k+1}J_k\sum_{i=1}^k\frac{\varphi(a_i)}{a_i}\#\mathcal{P}_{L_i,\mathcal{A}} + o\bigg(\frac{B^k}{\varphi(B)^k}\mathfrak{S}_B(\mathcal{L})\#\mathcal{A}(x)(\log R)^kI_k\bigg), \tag{6.5}$$

$$S_2=m\sum_{n\in\mathcal{A}(x)}w_n=m(1+o(1))\frac{B^k}{\varphi(B)^k}\mathfrak{S}_B(\mathcal{L})\#\mathcal{A}(x)(\log R)^kI_k, \tag{6.6}$$

$$S_3=k\sum_{n\in\mathcal{A}(x)}\sum_{i=1}^k\sum_{\substack{p\mid L_i(n)\\ p<x^\rho\\ p\nmid B}}w_n \ll \rho^2k^6(\log k)^2\frac{B^k}{\varphi(B)^k}\mathfrak{S}_B(\mathcal{L})\#\mathcal{A}(x)(\log R)^kI_k. \tag{6.7}$$

We choose $\rho=c_0k^{-3}(\log k)^{-1}$ with c_0 a small absolute constant such that $S_3\leqslant(1/3+o(1))S_2$. (This choice satisfies the bounds of Proposition 6.1 since $k\leqslant(\log x)^{1/5}$ and k is taken to be sufficiently large in terms of θ.) Thus, for x sufficiently large, we have

$$S\geqslant\frac{B^k}{\varphi(B)^k}\mathfrak{S}_B(\mathcal{L})(\log R)^k\bigg(\frac{J_k}{2}\log R\sum_{i=1}^k\frac{\varphi(a_i)\varphi(B)}{a_iB}\#\mathcal{P}_{L_i,\mathcal{A}}(x)-2mI_k\#\mathcal{A}(x)\bigg). \tag{6.8}$$

By the assumption of Theorem 3.1, we have

$$\frac{1}{k}\sum_{i=1}^k\frac{\varphi(a_i)\varphi(B)}{a_iB}\#\mathcal{P}_{L_i,\mathcal{A}}(x)\geqslant\delta\frac{\#\mathcal{A}(x)}{\log x}. \tag{6.9}$$

From Proposition 6.1, we have $J_k/I_k\gg(\log k)/k$. Combining this with (6.8) and (6.9), we have (for x sufficiently large)

$$S\geqslant(\theta/3)^k\frac{B^k}{\varphi(B)^k}\mathfrak{S}_B(\mathcal{L})\#\mathcal{A}(x)(\log x)^kI_k(3c_1\delta\log k-2m), \tag{6.10}$$

for some constant c_1 depending only on θ. In particular, if $m=c_1\delta\log k$, then $m\gg1$ (since $\delta\geqslant(\log k)^{-1}$ by assumption), and $S>0$. Using the bounds $I_k\gg(2k\log k)^{-k}$ and $\mathfrak{S}_B(\mathcal{L})\geqslant\exp(-Ck)$ from Proposition 6.1, along with the trivial bound $B/\varphi(B)\geqslant1$, we obtain

$$S\gg(\theta/3)^k\frac{B^k}{\varphi(B)^k}\mathfrak{S}_B(\mathcal{L})\#\mathcal{A}(x)(\log x)^kI_k\gg\#\mathcal{A}(x)(\log x)^k\exp(-C_2k^2), \tag{6.11}$$

for a suitable constant C_2 depending only on θ. Combining this with (6.4) and recalling our choices of m,ρ gives for $x\geqslant C_3$,

$$\#\{n\in\mathcal{A}(x):\#(\{L_1(n),\dots,L_k(n)\}\cap\mathcal{P})\geqslant c_1\delta\log k\}\geqslant\frac{\#\mathcal{A}(x)\exp(-C_3k^5)}{(\log x)^k}, \tag{6.12}$$

provided C_3 is chosen sufficiently large in terms of θ and α. This gives (6.1), and so the first claim of the theorem.

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

For the second claim, we have $L_i = an + b_i$ for all $1 \leqslant i \leqslant k$, with $a \ll 1$ and $b_i \leqslant \eta(\log x)$. (We will eventually take $\eta = c_4(k \log k)^{-1}$, for some fixed c_4 which implies the bound in the statement of Theorem 3.1.) In place of S we consider

$$S' = \sum_{n \in \mathcal{A}(x)} \bigg(\sum_{i=1}^{k} \mathbf{1}_{\mathcal{P}}(L_i(n)) - m - k \sum_{i=1}^{k} \sum_{\substack{p \mid L_i(n) \\ p < x^\rho, p \nmid B}} 1 - k \sum_{\substack{b \leqslant \eta \log x \\ L = an+b \notin \mathcal{L}}} \mathbf{1}_{\mathcal{S}(\theta/10;1)}(L(n)) \bigg) w_n$$
$$= S_1 - S_2 - S_3 - S_4. \tag{6.13}$$

The term in parentheses in (6.13) is positive only if at least m of the $L_i(n)$ are primes, none of the $L_i(n)$ have a prime factor $p \nmid B$ smaller than x^ρ, and all integers not in $\{L_1(n), \dots, L_k(n)\}$ of the form $an + b$ with $b \leqslant \eta \log x$ have a prime factor less than $x^{\theta/10}$. In particular, there can be no primes in the interval $[an, an + \eta(\log x)]$ apart from possibly $\{L_1(n), \dots, L_k(n)\}$, and so the primes counted in this way must be consecutive.

For S_4, we notice that $\Delta_L \neq 0$ for all L we consider since any L has the same lead coefficient as the L_i (and so cannot be a multiple of one of them). By Proposition 6.1, we have

$$S_4 \ll k \frac{B^k}{\varphi(B)^k} \# \mathcal{A}(x) (\log R)^{k-1} \mathfrak{S}_B(\mathcal{L}) I_k \sum_{\substack{b \leqslant \eta \log x \\ L = an+b \notin \mathcal{L}}} \frac{\Delta_L}{\varphi(\Delta_L)} \ll \eta k (\log k) S_2. \tag{6.14}$$

We choose $\eta = c_4/(k \log k)$ for some sufficiently small constant c_4 (this satisfies the requirements of Proposition 6.1). We then see that the bound (6.8) holds for S' in place of S provided x, k are sufficiently large. The whole argument then goes through as before. □

Proof of Theorem 3.2. We note that the result is trivial if $y \gg (\log x)^2$, $y = O(1)$ or $x = O(1)$ by the pigeonhole principle, Bertrand's postulate and the prime number theorem. Therefore, by changing the implied constant if necessary, it is sufficient to establish the result for $y \leqslant (\log x)^{1/5}$ with y sufficiently large.

We take $\mathcal{P} = \mathbb{P}$, $\mathcal{A} = \mathbb{N}$, $\mathcal{L} = \{L_1, \dots, L_k\}$, with $L_i(n) = n + h_i$, where h_i is the ith prime larger than k. By the prime number theorem, $h_i \leqslant 2k \log k$ for all i (provided k is sufficiently large). This is an admissible set.

By the Landau–Page theorem (see, for example, [Dav00, ch. 14]) there is at most one modulus $q_0 \leqslant \exp(c_1 \sqrt{\log x})$ such that there exists a primitive character χ modulo q_0 for which $L(s, \chi)$ has a real zero larger than $1 - c_2(\log x)^{-1/2}$ (for suitable fixed constants c_1, c_2). If this exceptional modulus q_0 exists, we take B to be the largest prime factor of q_0, and otherwise we take $B = 1$. If q_0 exists, it must be square-free apart from a possible factor of at most 4, and must satisfy $q_0 \gg \log x$ (from the class number formula). Therefore if q_0 exists, $\log \log x \ll B \ll \exp(c_1 \sqrt{\log x})$. Thus, whether or not q_0 exists, we have

$$\frac{B}{\varphi(B)} = 1 + O\bigg(\frac{1}{\log \log x} \bigg). \tag{6.15}$$

With this choice of parameters, we have error terms for parts (1) and (2) of Hypothesis 1 of size $x \exp(-c_3 \sqrt{\log x}) = \# \mathcal{A}(x) \exp(-c_3 \sqrt{\log x})$ for $\theta = 1/3$ by variants of the Bombieri–Vinogradov theorem avoiding an exceptional character (see, for example, [Dav00, ch. 28]). Thus Hypothesis 1 holds for $(\mathcal{A}, \mathcal{L}, \mathcal{P}, B, x, 1/3)$ for any $k \leqslant (\log x)^{1/5}$ provided k is sufficiently large. We have

$$\# \mathcal{P}_{L, \mathcal{A}}(x) = \frac{(1 + o(1))x}{\log x} = \frac{(1 + o(1)) \# \mathcal{A}(x)}{\log x}, \tag{6.16}$$

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

and so we may take $\delta = (1+o(1))$ in Theorem 3.1. Theorem 3.1 then gives

$$\#\{x \leqslant n \leqslant 2x : \pi(n+2k\log k) - \pi(n) \gg \log k\} \gg \frac{x}{(\log x)^k \exp(Ck)}. \tag{6.17}$$

Thus, given any x, y suitably large with $y \leqslant (\log x)^{1/5}$ we can take $k = \lfloor y/(2\log y)\rfloor$, and see that the above gives the result. All constants we have used are effectively computable. $\square$

Proof of Theorem 3.3. To get lower bounds of the correct order of magnitude, we average over admissible sets. We assume without loss of generality that a is reduced modulo q, so $1 \leqslant a < q$. We then adopt the same set-up as in the proof of Theorem 3.2 for our choice of $\mathcal{A}$, $\mathcal{P}$, θ, R. If an exceptional modulus q_0 exists (as defined in the proof of Theorem 3.2), then we take B to be the largest prime factor of q_0 coprime to q. Since $q \leqslant (\log x)^{1-\epsilon}$ and $q_0 \gg \log x$ (with q_0 essentially square-free) we have $\log\log x \ll_\epsilon B \ll x$ if q_0 exists. Thus $B/\varphi(B) = 1 + o(1)$ regardless of whether q_0 exists.

Instead of our individual choice of $\mathcal{L}$, we will average over all admissible choices of $\mathcal{L}$ with $\#\mathcal{L} = k$ and where $\mathcal{L} = \{L_1, \dots, L_k\}$ contains functions of the form $L_i(n) = qn + a + qb_i$ with $qb_i \leqslant \eta \log x$. We write $\mathcal{L}(\mathbf{b})$ for such a set given by $b_1, \dots, b_k$. We consider

$$S'' = \sum_{\substack{b_1 < \cdots < b_k \\ qb_k \leqslant \eta \log x \\ \mathcal{L}=\mathcal{L}(\mathbf{b}) \text{ admissible}}} \sum_{n \in \mathcal{A}(x)} \Bigg(\sum_{i=1}^k \mathbf{1}_{\mathcal{P}}(L_i(n)) - m - k \sum_{i=1}^k \sum_{\substack{p \mid L_i(n) \\ p < x^\rho, p \nmid B}} 1 - k \sum_{\substack{b \leqslant 2\eta \log x \\ L = qn+b \notin \mathcal{L}}} \mathbf{1}_{\mathcal{S}(\rho;B)}(L(n)) \Bigg) w_n(\mathcal{L}). \tag{6.18}$$

Here $w_n(\mathcal{L})$ are the weights given by Proposition 6.1 for the admissible set $\mathcal{L} = \mathcal{L}(\mathbf{b})$. For a given admissible set $\mathcal{L}$, the sum over n is then essentially the same quantity as S' from (6.13), except in the final term in parentheses we are considering elements with no prime factor less than x^ρ instead of $x^{\theta/10}$.

We see the term in parentheses in (6.18) is positive only if at least m of the $L_i(n)$ are primes, all the remaining $L_i(n)$ have no prime factors $p \nmid B$ less than x^ρ, and all other $qn + b$ with $b \leqslant 2\eta \log n$ have a prime factor $p \nmid B$ less than x^ρ. We see from this than no n can make a positive contribution from two different admissible sets (since if n makes a positive contribution for some admissible set, the $L_i(n)$ are uniquely determined as the integers in $[qn, qn + \eta \log x]$ with no prime factors $p \nmid B$ less than x^ρ). By (6.3), we see that if n makes a positive contribution then $w_n \ll (\log x)^{2k} \exp(O(k/\rho))$, with the implied bound uniform in $\mathcal{L}(\mathbf{b})$.

As before, we choose $\rho = c_0 k^{-3} (\log k)^{-1}$, which makes the contribution of the third of the terms in parentheses small compared to the second one. Following the argument of the proof of Theorem 3.1, using $\mathcal{S}(\rho; B)$ in place of $\mathcal{S}(\theta/10; 1)$ increases the size of the contribution of the final term by a factor $O(\rho^{-1}) = O(k^3 \log k)$. Thus to show the final term is suitably small, we take $\eta \leqslant \epsilon$ to be a small multiple of $k^{-4}(\log k)^{-2}$ instead of $1/(k \log k)$ (which is acceptable for Proposition 6.1). With these choices, we find that for a suitable constant c_1 we have

$$S'' \geqslant (\theta/3)^k \frac{B^k}{\varphi(B)^k} \mathfrak{S}_B(\mathcal{L}) \# \mathcal{A}(x) (\log x)^k I_k \sum_{\substack{b_1 < \cdots < b_k \\ qb_k \leqslant \eta \log x \\ \mathcal{L}(\mathbf{b}) \text{ admissible}}} (3c_1 \log k - 2m). \tag{6.19}$$

Therefore, given $m \in \mathbb{N}$ we choose $k = \lceil \exp(m/c_1) \rceil$. With this choice we see that $S'' > 0$. Using the bounds $I_k \gg (k \log k)^{-k}$ and $\mathfrak{S}_B(\mathcal{L}) \gg \exp(-Ck)$ from Proposition 6.1 and $B^k/\varphi(B)^k \geqslant 1$,

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

we see that for a suitable constant C_2 we have

$$S'' \gg x(\log x)^k \exp(-C_2k^2) \sum_{\substack{b_1<\cdots<b_k\\ qb_k\leqslant \eta\log x\\ \mathcal{L}(\mathbf{b}) \text{ admissible}}} 1. \tag{6.20}$$

Thus we are left with obtaining a lower bound for the inner sum of (6.20). We see all the b_i lie between 0 and $\eta(\log x)/q$. We greedily sieve this interval by removing for each prime $p \leqslant k$ in turn any elements from the residue class modulo p which contains the fewest elements. The resulting set has size at least

$$\frac{\eta\log x}{q}\prod_{p\leqslant k}\Big(1-\frac{1}{p}\Big) \gg \frac{\log x}{qk^4(\log k)^3}. \tag{6.21}$$

Any choice of k distinct b_i from this set will cause the resulting $\mathcal{L}(\mathbf{b})$ to be admissible. We now recall from the theorem that we are only considering $q \leqslant (\log x)^{1-\epsilon}$ and $m \leqslant c_\epsilon \log\log x$. For a suitably small choice of c_ϵ, we see that $k = \lceil \exp(m/c_1)\rceil \leqslant (\log x)^{\epsilon/10}$. Therefore from (6.21) we see the length of the interval is at least k^2 if x is sufficiently large in terms of ϵ. In this case, we obtain the bound

$$\sum_{\substack{b_1<\cdots<b_k\\ qb_k\leqslant \eta\log x\\ \mathcal{L}(\mathbf{b}) \text{ admissible}}} 1 \geqslant k^{-k}\Big(\frac{c_3\log x}{qk^4\log^3 k}-k\Big)^k \gg \Big(\frac{\log x}{q}\Big)^k \exp(-C_4k^2), \tag{6.22}$$

for some constants $c_3, C_4 > 0$. Thus, substituting (6.22) into (6.20), we obtain

$$S'' \gg x(\log x)^{2k}\exp(-C_5k^2)q^{-k}. \tag{6.23}$$

We recall that every pair $n, \mathcal{L}$ which makes a positive contribution to S'' is counted with weight at most $kw_n(\mathcal{L}) \ll k(\log x)^{2k}\exp(O(k/\rho))$ (uniformly over all choices of $\mathcal{L}$). Putting this all together, we obtain the number N of integers n with $x \leqslant n \leqslant 2x$ such that there are $\gg \log k$ consecutive primes all congruent to $a \pmod q$ in the interval $[qn, qn+\eta\log x]$ satisfies

$$N \gg \frac{x}{q^k\exp(C_6k^5)}. \tag{6.24}$$

We see that the initial prime in each such interval is counted by at most $\log x$ values of n. Therefore, changing the count to be over the initial prime, recalling $k = \lceil\exp(m/c_1)\rceil$, recalling that $\eta \leqslant \epsilon$, and replacing x with $x/3q$ gives

$$\#\{p_n \leqslant x : p_n \equiv \cdots \equiv p_{n+m} \equiv a \pmod q, p_{n+m}-p_n \leqslant \epsilon\log x\} \gg_\epsilon \frac{\pi(x)}{(2q)^{\exp(Cm)}}, \tag{6.25}$$

for a suitable constant $C > 0$, as required. □

Proof of Theorem 3.4. We take $\mathcal{P} = \mathbb{P}$, $\mathcal{A} = [x, x+y]$, $B = 1$, $\theta = 1/30-\epsilon$. Given m, we choose $k = \exp(Cm)$ for some suitable constant $C > 0$.

Timofeev [Tim87] (improving earlier work of Huxley and Iwaniec [HI75] and Perelli *et al.* [PPS85]) has shown that, for $\theta = 1/30 - \epsilon/2$, for any $x^{7/12+\epsilon/2} \leqslant y \leqslant x$ and any fixed $C' > 0$ we have

$$\sum_{q<x^\theta}\sup_{(a,q)=1}\Big|\pi(x+y;q,a)-\pi(x;q,a)-\frac{\pi(x+y)-\pi(x)}{\varphi(q)}\Big| \ll_{C',\epsilon} \frac{y}{(\log x)^{C'}}. \tag{6.26}$$

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

By taking C' sufficiently large in terms of k, we see that (6.26) implies Hypothesis 1 holds for our choice of $\theta = 1/30 - \epsilon$ provided x is sufficiently large in terms of m and ϵ. Theorem 3.1 then automatically gives Theorem 3.4. $\square$

Proof of Theorem 3.5. We take $\mathcal{A} = \mathbb{N}$, $B = \Delta_K$ and $\mathcal{P}$, $\mathcal{L}$ the sets given by the statement of the theorem. To avoid confusion, we note that Δ_K here is the discriminant of $K/\mathbb{Q}$, and unrelated to Δ_L from Proposition 6.1. Murty and Murty [MM87] have then established the key estimate (2) of Hypothesis 1 with any $\theta < \min(1/2, 2/\#G)$, where $G = \mathrm{Gal}(K/\mathbb{Q})$ (the other estimates being trivial). Finally, we have

$$\frac{1}{k}\frac{B}{\varphi(B)}\sum_{i=1}^{k}\frac{\varphi(a_i)}{a_i}\#\mathcal{P}_{L_i,\mathcal{A}}(x) \geqslant (1+o(1))\frac{\Delta_K \#\mathcal{C}}{\varphi(\Delta_K)\#G}\frac{x}{\log x}, \tag{6.27}$$

and so for x sufficiently large, we may take δ to be a constant depending only on K. The result now follows directly from Theorem 3.1. $\square$

7. Initial considerations

We recall that we are given a set $\mathcal{A}$ of integers, a set $\mathcal{P}$ of primes, an admissible set $\mathcal{L} = \{L_1, \dots, L_k\}$ of integer linear functions, an integer B and quantities R, x. We assume that the coefficients of $L_i(n) = a_i n + b_i \in \mathcal{L}$ satisfy $|a_i|, |b_i| \leqslant x^{\alpha}$, $a_i \neq 0$, and $k = \#\mathcal{L}$ is sufficiently large in terms of the fixed quantities θ, α and satisfies $k \leqslant (\log x)^{1/5}$. B, R satisfy $1 \leqslant B \leqslant x^{\alpha}$, and $x^{\theta/10} \leqslant R \leqslant x^{\theta/3}$. Finally, we assume from now on that the set $\mathcal{A}$ satisfies

$$\sum_{q \leqslant x^{\theta}} \max_{a}\left|\#\mathcal{A}(x;q,a) - \frac{\#\mathcal{A}(x)}{q}\right| \ll \frac{\#\mathcal{A}(x)}{(\log x)^{100k^2}},$$

and

$$\#\mathcal{A}(x;q,a) \ll \frac{\#\mathcal{A}(x)}{q}$$

for any $q < x^{\theta}$. Together these assumptions are a slight generalization of the assumptions of Proposition 6.1.

We define the multiplicative functions $\omega = \omega_{\mathcal{L}}$ and $\varphi_\omega = \varphi_{\omega,\mathcal{L}}$ and the singular series $\mathfrak{S}_D(\mathcal{L})$ for an integer D by

$$\omega(p) = \begin{cases} \#\left\{1 \leqslant n \leqslant p : \prod_{i=1}^{k} L_i(n) \equiv 0 \pmod p\right\}, & p \nmid B, \\ 0, & p | B, \end{cases} \tag{7.1}$$

$$\varphi_\omega(d) = \prod_{p|d}(p - \omega(p)), \tag{7.2}$$

$$\mathfrak{S}_D(\mathcal{L}) = \prod_{p \nmid D}\left(1 - \frac{\omega(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}. \tag{7.3}$$

Since $\mathcal{L}$ is admissible, we have $\omega(p) < p$ for all p and so $\varphi_\omega(n) > 0$ and $\mathfrak{S}_D(\mathcal{L}) > 0$ for any integer D. Since $\omega(p) = k$ for all $p \nmid \prod_{i=1}^{k} a_i \prod_{i \neq j}(a_i b_j - b_i a_j)$ we see the product $\mathfrak{S}_D(\mathcal{L})$ converges.

The main innovation in [May15] was a different choice of the sieve weights used in the GPY method to detect small gaps between primes. In order to adapt the argument of [May15] to the

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

more general situation considered here, we need to modify the choice of these weights further to produce a choice more amenable to obtaining uniform estimates. In particular, in [May15] the 'W-trick' was used to eliminate the need for consideration of the singular series which would naturally arise. In our situation, however, in order to obtain suitable uniform estimates without stronger assumptions on the error terms in Hypothesis 1, we need to take these singular series into account.

We will consider sieve weights $w_n = w_n(\mathcal{L})$, which are defined to be 0 if $\prod_{i=1}^k L_i(n)$ is a multiple of any prime $p \leqslant 2k^2$ with $p \nmid B$. Specifically, we define

$$W = \prod_{p \leqslant 2k^2, p\nmid B} p \tag{7.4}$$

$$w_n = \begin{cases} \Bigl(\sum_{d_i | L_i(n)\ \forall i} \lambda_{\mathbf{d}}\Bigr)^2 & \text{if } \Bigl(W, \prod_{i=1}^k L_i(n)\Bigr) = 1, \\ 0 & \text{if } \Bigl(W, \prod_{i=1}^k L_i(n)\Bigr) \neq 1, \end{cases} \tag{7.5}$$

for some real variables $\lambda_{\mathbf{d}}$ depending on $\mathbf{d} = (d_1, \dots, d_k)$. We first restrict our $\lambda_{\mathbf{d}}$ to be supported on $\mathbf{d}$ with $d = \prod_{i=1}^k d_i$ square-free and coprime to WB.

Given a prime $p \nmid WB$, let $r_{p,1}, \dots, r_{p,\omega(p)}$ be the $\omega(p)$ residue classes for which $\prod_{i=1}^k L_i(n)$ vanishes modulo p. For each such prime p, we fix a choice of indices $j_{p,1}, \dots, j_{p,\omega(p)} \in \{1, \dots, k\}$ such that for each $i \in \{1, \dots, \omega(p)\}$ we have that $j_{p,i}$ is the smallest index such that

$$L_{j_{p,i}}(r_{p,i}) \equiv 0 \pmod{p}. \tag{7.6}$$

(We could choose any index satisfying the above condition; we choose the smallest index purely for concreteness.) All the functions L_i are linear and, since $\mathcal{L}$ is admissible, none of the L_i are a multiple of p. This means that for any $L \in \mathcal{L}$ there is at most one residue class for which L vanishes modulo p. Thus the indices $j_{p,1}, \dots, j_{p,\omega(p)}$ we have chosen must be distinct. We now restrict the support of $\lambda_{\mathbf{d}}$ to $(d_j, p) = 1$ for all $j \notin \{j_{p,1}, \dots, j_{p,\omega(p)}\}$.

We see these restrictions are equivalent to the restriction that the support of $\lambda_{\mathbf{d}}$ must lie the set

$$\mathcal{D}_k = \mathcal{D}_k(\mathcal{L}) = \{\mathbf{d} \in \mathbb{N}^k : \mu^2(d) = 1, (d_j, W_j) = 1\ \forall 1 \leqslant j \leqslant k\}, \tag{7.7}$$

where W_j are square-free integers each a multiple of WB, and any prime $p \nmid WB$ divides exactly $k - \omega(p)$ of the W_j (such a prime p divides W_j if $j \notin \{j_{p,1}, \dots, j_{p,\omega(p)}\}$). We recall that in our notation $\mu^2(d) = \mu^2(\prod_{i=1}^k d_i)$.

The key point of these restrictions is so that different components of different $\mathbf{d}$ occurring in our sieve weights will be relatively prime. Indeed, let $\mathbf{d}$ and $\mathbf{e}$ both occur in the sum (7.5). If $p|d_i$ then $p|L_i(n)$, and so i must be the chosen index for the residue class $n \pmod{p}$. But if we also have $p|e_j$ then similarly j must be the chosen index for this residue class, and so we must have $i = j$. Hence $(d_i, e_j) = 1$ for all $i \neq j$.

Similar to [May15], we define $\lambda_{\mathbf{d}}$ in terms of variables $y_{\mathbf{r}}$ supported on $\mathbf{r} \in \mathcal{D}_k$ by

$$\lambda_{\mathbf{d}} = \mu(d) d \sum_{\mathbf{d}|\mathbf{r}} \frac{y_{\mathbf{r}}}{\varphi_\omega(r)}, \quad y_{\mathbf{r}} = \frac{\mathbf{1}_{\mathcal{D}_k}(\mathbf{r}) W^k B^k}{\varphi(WB)^k} \mathfrak{S}_{WB}(\mathcal{L}) F\Bigl(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\Bigr) \tag{7.8}$$

(again, we recall $d = \prod_{i=1}^k d_i$), where R is a quantity with

$$x^{\theta/10} \leqslant R \leqslant x^{\theta/3}, \tag{7.9}$$

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

and $F:\mathbb{R}^k\to\mathbb{R}$ is a smooth function given by

$$F(t_1,\dots,t_k)=\psi\Big(\sum_{i=1}^k t_i\Big)\prod_{i=1}^k\frac{\psi(t_i/U_k)}{1+T_kt_i},\quad T_k=k\log k,\quad U_k=k^{-1/2}. \tag{7.10}$$

Here $\psi:[0,\infty)\to[0,1]$ is a fixed smooth non-increasing function supported on $[0,1]$ which is 1 on $[0,9/10]$. In particular, we note that this choice of F is non-negative, and that the support of ψ implies that

$$\lambda_{\mathbf{d}}=0\quad\text{if } d=\prod_{i=1}^k d_i>R. \tag{7.11}$$

We will find it useful to also consider the closely related functions F_1 and F_2 which will appear in our error estimates, defined by

$$F_1(t_1,\dots,t_k)=\prod_{i=1}^k\frac{\psi(t_i/U_k)}{1+T_kt_i},\quad F_2(t_1,\dots,t_k)=\sum_{1\leqslant j\leqslant k}\Bigg(\frac{\psi(t_j/2)}{1+T_kt_j}\prod_{\substack{1\leqslant i\leqslant k\\ i\neq j}}\frac{\psi(t_i/U_k)}{1+T_kt_i}\Bigg). \tag{7.12}$$

Finally, by Möbius inversion, we see that (7.8) implies that for $\mathbf{r}\in\mathcal{D}_k$,

$$y_{\mathbf{r}}=\mu(r)\varphi_\omega(r)\sum_{\mathbf{r}|\mathbf{f}}\frac{y_{\mathbf{f}}}{\varphi_\omega(f)}\sum_{\substack{\mathbf{d}\\ \mathbf{r}|\mathbf{d},\,\mathbf{d}|\mathbf{f}}}\mu(d)=\mu(r)\varphi_\omega(r)\sum_{\mathbf{r}|\mathbf{d}}\frac{\lambda_{\mathbf{d}}}{d}. \tag{7.13}$$

8. Preparatory lemmas

Lemma 8.1. (i) *There is a constant C such that, for any admissible set $\mathcal{L}$ of size k, we have*

$$\mathfrak{S}_B(\mathcal{L})\geqslant\exp(-Ck).$$

(ii) *Let all functions $L_i\in\mathcal{L}$ be of the form $L_i=an+b_i$, for some integers $|a|\ll 1$ and $|b_i|\ll\log x$. Let $\Delta_L=|a|^{k+1}\prod_{i=1}^k|b_i-b|$ and $\eta\geqslant(\log x)^{-9/10}$. Then we have*

$$\sum_{\substack{|b|\leqslant\eta\log x\\ L(n)=an+b\notin\mathcal{L}}}\frac{\Delta_L}{\varphi(\Delta_L)}\ll\eta(\log x)(\log k).$$

Proof. Since $\omega(p)\leqslant\min(k,p-1)$ for any admissible $\mathcal{L}$ of size k, we have

$$\mathfrak{S}_B(\mathcal{L})=\prod_{p\nmid B}\Big(1-\frac{\omega(p)}{p}\Big)\Big(1-\frac{1}{p}\Big)^{-k}\geqslant\prod_{p\leqslant 2k,p\nmid B}\frac{1}{p}\prod_{p>2k,p\nmid B}\Big(1-\frac{k}{p}\Big)\Big(1-\frac{1}{p}\Big)^{-k}. \tag{8.1}$$

Since all terms in the products on the right-hand side are less than 1, we can drop the restriction $p\nmid B$ for a lower bound. This gives

$$\mathfrak{S}_B(\mathcal{L})\geqslant\prod_{p\leqslant 2k}\frac{1}{p}\prod_{p>2k}\exp(O(k^2/p^2))\geqslant\exp(-Ck). \tag{8.2}$$

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

We now consider the second statement. We have $L_i(n)=an+b_i$ with $|b_i|\ll\log x$, and consider $L=an+b\notin\mathcal{L}$ with $|b|\leqslant\eta\log x$. If $k\gg\log\log x$ then we use the bound $\Delta_L/\varphi(\Delta_L)\ll\log\log\Delta_L\ll\log k$ to give

$$\sum_{\substack{|b|\leqslant\eta\log x\\ L=an+b\notin\mathcal{L}}}\frac{\Delta_L}{\varphi(\Delta_L)}\ll\eta(\log k)(\log x). \tag{8.3}$$

We now establish (8.3) in the case $k\ll\log\log x$. Using the identity $e/\varphi(e)=\sum_{d|e}\mu^2(d)/\varphi(d)$, and splitting the terms depending on the size of divisors, we have

$$\begin{aligned}\sum_{\substack{|b|\leqslant\eta\log x\\ L=an+b\notin\mathcal{L}}}\frac{\Delta_L}{\varphi(\Delta_L)}&=\frac{a}{\varphi(a)}\sum_{\substack{|b|\leqslant\eta\log x\\ L=an+b\notin\mathcal{L}}}\sum_{\substack{d|\Delta_L\\ (d,a)=1}}\frac{\mu^2(d)}{\varphi(d)}\\ &\ll\sum_{\substack{|b|\leqslant\eta\log x\\ L=an+b\notin\mathcal{L}}}\Bigg(\sum_{\substack{1\leqslant d\leqslant\eta\log x\\ d|\Delta_L,(d,a)=1}}\frac{\mu^2(d)}{\varphi(d)}+\sum_{\substack{d>\eta\log x\\ d|\Delta_L}}\frac{\mu^2(d)\sum_{p|d}\log p}{\varphi(d)\log(\eta\log x)}\Bigg)\\ &\ll\sum_{\substack{1\leqslant d\leqslant\eta\log x\\ (d,a)=1}}\frac{\mu^2(d)}{\varphi(d)}\sum_{\substack{|b|\leqslant\eta\log x\\ L=an+b\notin\mathcal{L}\\ d|\Delta_L}}1+\sum_{\substack{|b|\leqslant\eta\log x\\ L=an+b\notin\mathcal{L}}}\sum_{p|\Delta_L}\frac{\log p}{p\log(\eta\log x)}\frac{\Delta_L}{\varphi(\Delta_L)}.\end{aligned} \tag{8.4}$$

We first consider the second term on the right-hand side of (8.4). We have $\sum_{p|\Delta_L}p^{-1}\log p\ll\log\log\Delta_L$ and $\Delta_L/\varphi(\Delta_L)\ll\log\log\Delta_L$. But we are only considering $k\ll\log\log x$ and $\eta\geqslant(\log x)^{-9/10}$, and so $(\log\log\Delta_L)^2\ll(\log\log\log x)^2=o(\log(\eta\log x))$. Therefore we see that the total contribution from the second term in (8.4) is $o(\eta\log x)$.

We now consider the first sum in (8.4). For every prime $p|d$, there are at most k choices for the residue class $b \pmod p$ such that $p|\Delta_L$, and trivially there are also at most p choices. Thus the inner sum can be written as $\prod_{p|d}\min(p,k)$ sums over b in a fixed arithmetic progression modulo d. For each such sum there are $\ll\eta(\log x)/d$ possible values of b. Thus we have

$$\begin{aligned}\sum_{\substack{1\leqslant d\leqslant\eta\log x\\ (d,a)=1}}\frac{\mu^2(d)}{\varphi(d)}\sum_{\substack{|b|\leqslant\eta\log x\\ L=an+b\notin\mathcal{L}\\ d|\Delta_L}}1&\ll\sum_{d\leqslant\eta\log x}\frac{\mu^2(d)\prod_{p|d}\min(p,k)}{\varphi(d)}\left(\frac{\eta\log x}{d}\right)\\ &\ll\eta\log x\prod_{p\leqslant k}\left(1+\frac{1}{p-1}\right)\prod_{p>k}\left(1+\frac{k}{p(p-1)}\right)\\ &\ll\eta(\log x)(\log k).\end{aligned} \tag{8.5}$$

This gives the result. □

Lemma 8.2. *Let*

$$Y_{\mathbf{r}}=\frac{W^kB^k\mathfrak{S}_{WB}(\mathcal{L})}{\varphi(WB)^k}F_2\left(\frac{\log r_1}{\log R},\ldots,\frac{\log r_k}{\log R}\right),$$

where F_2 is given by (7.12).

(i) *Let $\mathbf{r},\mathbf{s}\in\mathcal{D}_k$ with $s_i=r_i$ for all $i\neq j$, and $s_j=Ar_j$ for some $A\in\mathbb{N}$. Then*

$$y_{\mathbf{s}}=y_{\mathbf{r}}+O\left(T_kY_{\mathbf{r}}\frac{\log A}{\log R}\right).$$

(ii) *Let* $\mathbf{r}$, $\mathbf{s} \in \mathcal{D}_k$ *with* $r = s$ *and let* A *be the product of primes dividing* r *but not* $(\mathbf{r}, \mathbf{s})$. *Then*

$$y_{\mathbf{s}} = y_{\mathbf{r}} + O\bigg(T_k(Y_{\mathbf{r}} + Y_{\mathbf{s}})\frac{\log A}{\log R}\bigg).$$

Proof. We recall the definitions of ψ, F_2, $U_k = k^{-1/2}$ and $T_k = k \log k$ from § 7. Given any two reals $u, v \geqslant 0$ with $|u - v| \leqslant \epsilon$, we have

$$\frac{1}{1 + T_k u} = \frac{1 + O(T_k \epsilon)}{1 + T_k v}, \quad \psi(u) = \psi(v) + O(\epsilon). \tag{8.6}$$

Given $\mathbf{r}, \mathbf{s}$ as in the lemma, we define $u_1, \dots, u_k, v_1, \dots, v_k, \epsilon_1, \dots, \epsilon_k$ by $u_i = \log r_i / \log R$, $v_i = \log s_i / \log R$ and $\epsilon_i = v_i - u_i$. For part (i) we have $\epsilon_i = 0$ for $i \neq j$ and $\epsilon_j = \log A / \log R$. We may assume $\epsilon_j \leqslant 1, u_j \leqslant U_k$ since otherwise the result is trivial. By (8.6) we have

$$\begin{aligned}\psi\bigg(\sum_{i=1}^k v_i\bigg)\frac{\psi(v_j/U_k)}{1 + T_k v_j} &= \bigg(\psi\bigg(\sum_{i=1}^k u_i\bigg) + O\bigg(\frac{\log A}{\log R}\bigg)\bigg)\bigg(\psi\bigg(\frac{u_j}{U_k}\bigg) + O\bigg(\frac{\log A}{U_k \log R}\bigg)\bigg) \\ &\quad \times \frac{1 + O(T_k(\log A / \log R))}{1 + T_k u_j}.\end{aligned} \tag{8.7}$$

Since $1 + U_k^{-1} \ll T_k$, $0 \leqslant \psi \leqslant 1$ and $\psi(v_j/2) = 1$ (since $v_j = u_j + \epsilon_j \leqslant 1 + U_k < 9/5$), expanding the terms and multiplying by $\prod_{i \neq j} \psi(u_i/U_k)/(1 + T_k u_i)$ gives the result for (i).

We now consider part (ii). We let $\mathbf{t}$ be the vector with $t_i = [r_i, s_i]$. By applying part (i) to each component in turn, and using the fact that $Y_{\mathbf{r}}$ is decreasing, we find that

$$y_{\mathbf{s}} = y_{\mathbf{t}} + O\bigg(T_k Y_{\mathbf{s}} \sum_{i=1}^k \frac{\log{[r_i, s_i]}/s_i}{\log R}\bigg) = y_{\mathbf{t}} + O\bigg(T_k Y_{\mathbf{s}} \frac{\log A}{\log R}\bigg). \tag{8.8}$$

We obtain the same expression for $\mathbf{r}$ in place of $\mathbf{s}$, and hence the result follows. □

We use the following lemma to estimate the various smoothed sums of multiplicative functions which we will encounter.

Lemma 8.3. *Let* $A_1, A_2, L > 0$. *Let* γ *be a multiplicative function satisfying*

$$0 \leqslant \frac{\gamma(p)}{p} \leqslant 1 - A_1 \quad \text{and} \quad -L \leqslant \sum_{w \leqslant p \leqslant z} \frac{\gamma(p) \log p}{p} - \log z/w \leqslant A_2$$

for any $2 \leqslant w \leqslant z$. *Let* g *be the totally multiplicative function defined on primes by* $g(p) = \gamma(p)/(p - \gamma(p))$. *Finally, let* $G : [0, 1] \to \mathbb{R}$ *be smooth, and let* $G_{\max} = \sup_{t \in [0,1]}(|G(t)| + |G'(t)|)$. *Then*

$$\sum_{d < z} \mu(d)^2 g(d) G\bigg(\frac{\log d}{\log z}\bigg) = c_\gamma \log z \int_0^1 G(x)\, dx + O_{A_1, A_2}(c_\gamma (1 + L) G_{\max}),$$

where

$$c_\gamma = \prod_p \bigg(1 - \frac{\gamma(p)}{p}\bigg)^{-1}\bigg(1 - \frac{1}{p}\bigg).$$

Proof. This is [GGPY09, Lemma 4], with $\kappa = 1$ and slight changes to the notation. We note that in the general formulation, an additional term $c_\gamma(L + 1)^\kappa$ should be included in the error term, but this is not necessary in our situation with $\kappa = 1$. We thank Kevin Ford for this observation. □

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

Lemma 8.4. *Let $r \leqslant k \ll (\log R)^{1/5}$. Let $W_1, \dots, W_r \leqslant R^{O(k)}$ all be a multiple of $\prod_{p\leqslant 2k^2} p$. Let g be a multiplicative function with $g(p) = p + O(k)$. Let $G : \mathbb{R} \to \mathbb{R}$ be a smooth function supported on the interval $[0, 1]$ such that*

$$\sup_{t\in[0,1]} (|G(t)| + |G'(t)|) \leqslant \Omega_G \int_0^\infty G(t)\,dt,$$

for some quantity Ω_G which satisfies $r\Omega_G = o((\log R)/(\log\log R))$. Let $\Phi : \mathbb{R} \to \mathbb{R}$ be smooth with $\Phi(t), \Phi'(t) \ll 1$ uniformly for all t.

Then for k sufficiently large we have

$$\begin{aligned}\sum_{\substack{\mathbf{e}\in\mathbb{N}^r\\(e_i,W_i)=1\ \forall i}} &\frac{\mu^2(e)}{g(e)}\Phi\Big(\sum_{i=1}^k \frac{\log e_i}{\log R}\Big)\prod_{i=1}^k G\Big(\frac{\log e_i}{\log R}\Big)\\ &= \Pi_g(\log R)^r \int\cdots\int_{t_1,\dots,t_r\geqslant 0} \Phi\Big(\sum_{i=1}^r t_i\Big)\prod_{i=1}^r G(t_i)\,dt_i\\ &\quad + O\Big(r\Omega_G\Pi_g(\log R)^{r-1}\log\log R\int\cdots\int_{t_1,\dots,t_r\geqslant 0}\prod_{i=1}^r G(t_i)\,dt_i\Big),\end{aligned}$$

where

$$\Pi_g = \prod_p \Big(1 + \frac{n(p)}{g(p)}\Big)\Big(1 - \frac{1}{p}\Big)^r, \quad n(p) = \#\{i \in \{1, \dots, r\} : p\nmid W_i\}.$$

Proof. We let Σ denote the sum in the statement of the lemma. We estimate the Σ by applying Lemma 8.3 r times to each variable $e_1, \dots, e_r$ in turn. We use induction to establish that, having applied the lemma j times, we obtain the estimate

$$\begin{aligned}\Sigma &= c_j(\log R)^j \sum_{\substack{e_{j+1},\dots,e_r\\(e_i,W_i)=1\ \forall i}} \frac{\mu(e_{j+1}\dots e_r)^2}{g_j(e_{j+1}\dots e_k)}\prod_{i=j+1}^r G(u_i)\int\cdots\int_{t_1,\dots,t_j\geqslant 0}\Phi\Big(\sum_{i=1}^j t_i + \sum_{i=j+1}^r u_i\Big)\prod_{i=1}^j G(t_i)\,dt_i\\ &\quad + c_j(\log R)^j\Big(\int_0^\infty G(t)\,dt\Big)^j\Big(\sum_{\ell=1}^j\binom{j}{\ell}O\Big(\frac{\Omega_G\log\log R}{\log R}\Big)^\ell\Big)\\ &\quad \times \sum_{\substack{e_{j+1},\dots,e_r\\(e_i,W_i)=1\ \forall i}} \frac{\mu(e_{j+1}\dots e_r)^2}{g_j(e_{j+1}\dots e_k)}\prod_{i=j+1}^r G(u_i),\end{aligned}\tag{8.9}$$

where

$$\begin{gathered}u_i = \frac{\log e_i}{\log R}, \quad n_j(p) = \#\{i \in \{1, \dots, j\} : p\nmid W_i\},\\ g_j(d) = \prod_{p|d}(g(p) + n_j(p)), \quad c_j = \prod_p\Big(1 + \frac{n_j(p)}{g(p)}\Big)\Big(1 - \frac{1}{p}\Big)^j.\end{gathered}\tag{8.10}$$

We see that (8.9) clearly holds when $j = 0$. We now assume that (8.9) holds for some $j < r$, and apply Lemma 8.3 to the sum over e_{j+1}. In the notation of Lemma 8.3, we have

$$\gamma(p) = \begin{cases} 0, & p|W_{j+1}\prod_{i=j+2}^r e_i,\\ p(1 + n_j(p) + g(p))^{-1} = 1 + O(k/p), & p\nmid W_{j+1}\prod_{i=j+2}^r e_i.\end{cases}\tag{8.11}$$

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

Since W_{j+1} is a multiple of all primes $p \leqslant 2k^2$ (by assumption of the lemma), we see that we can take A_1 and A_2 to be fixed constants (independent of j, k, r, x) provided k is sufficiently large. With this choice of $\gamma(p)$, we see that

$$L \ll 1 + \sum_{p | W_{j+1} \prod_{i=j+2}^r e_i} \frac{\log p}{p} + \sum_{p > 2k^2} \frac{k \log p}{p^2} \ll \log\log R. \tag{8.12}$$

Here we used the fact that the first sum is over prime divisors of an integer which is $\ll R^{O(k^2)}$, and this sum is largest when all the prime divisors are smallest, and that $k \ll (\log R)^{1/5} \ll \log R$.

We apply Lemma 8.3 to the main term with the smooth function G_1, and to the error term with the smooth function G_2, defined by

$$G_1(t) = \underset{t_1,\dots,t_j \geqslant 0}{\int \cdots \int} G(t) \Phi\Bigl(\sum_{i=1}^j t_i + t + \sum_{i=j+2}^r u_i\Bigr)\Bigl(\prod_{i=1}^j G(t_i)\, dt_i\Bigr)\Bigl(\prod_{i=j+2}^r G(u_i)\Bigr), \tag{8.13}$$

$$G_2(t) = G(t)\Bigl(\int_{t' \geqslant 0} G(t')\, dt'\Bigr)^j \Bigl(\prod_{i=j+2}^r G(u_i)\Bigr), \tag{8.14}$$

where we recall $u_i = (\log e_i)/\log R$ for $i > j+1$. With this choice, we see that from the bounds on Φ, G given in the lemma, we have

$$\begin{aligned} \sup_{t \in [0,1]} (|G_1(t)| + |G_1'(t)| + |G_2(t)| + |G_2'(t)|) &\ll \Omega_G \Bigl(\int_{t \geqslant 0} G(t)\, dt\Bigr)^{j+1} \Bigl(\prod_{i=j+2}^r G(u_i)\Bigr) \\ &= \Omega_G \int_{t \geqslant 0} G_2(t)\, dt. \end{aligned} \tag{8.15}$$

Thus Lemma 8.3 gives

$$\begin{aligned} &\sum_{\substack{e_{j+1} \\ (e_{j+1}, W_{j+1} \prod_{i=j+2}^r e_i) = 1}} \frac{\mu^2(e_{j+1})}{g_j(e_{j+1})} G_1\Bigl(\frac{\log e_{j+1}}{\log R}\Bigr) \\ &\quad = \log R \prod_p \Bigl(1 - \frac{\gamma(p)}{p}\Bigr)^{-1} \Bigl(1 - \frac{1}{p}\Bigr) \int_0^\infty G_1(t)\, dt \\ &\quad\quad + O\Bigl(\Omega_G \log\log R \prod_p \Bigl(1 - \frac{\gamma(p)}{p}\Bigr)^{-1} \Bigl(1 - \frac{1}{p}\Bigr) \int_0^\infty G_2(t)\, dt\Bigr), \end{aligned} \tag{8.16}$$

and we obtain the same expression when summing with G_2 instead of G_1, except $\int_0^\infty G_1(t)\, dt$ is replaced by $\int_0^\infty G_2(t)\, dt$ in the main term. The implied constant in the error term is independent of j. We note that

$$\begin{aligned} c_j \prod_p \Bigl(1 - \frac{\gamma(p)}{p}\Bigr)^{-1} \Bigl(1 - \frac{1}{p}\Bigr) &= \prod_{p | W_{j+1}} \Bigl(1 + \frac{n_j(p)}{g(p)}\Bigr)\Bigl(1 - \frac{1}{p}\Bigr)^{j+1} \\ &\quad \times \prod_{p \nmid W_{j+1}} \Bigl(1 + \frac{n_j(p)+1}{g(p)}\Bigr)\Bigl(1 - \frac{1}{p}\Bigr)^{j+1} \\ &\quad \times \prod_{\substack{p | e_{j+2} \dots e_r \\ p \nmid W_{j+1}}} \Bigl(\frac{n_j(p) + g(p)}{n_j(p) + g(p) + 1}\Bigr) \\ &= \frac{c_{j+1} g_j(e_{j+2} \dots e_r)}{g_{j+1}(e_{j+2} \dots e_r)}. \end{aligned} \tag{8.17}$$

Therefore substituting (8.16) and (8.17) into (8.9) gives the result for $j+1$. We conclude that (8.9) holds for all $j \leqslant r$.

Finally, let $\varepsilon = (\Omega_G \log\log R)/\log R$. By assumption of the lemma, we have $\varepsilon = o(1/r)$. We see the sum over ℓ in (8.9) is $(1+O(\varepsilon))^j - 1 = O(j\varepsilon)$ where, by our bound on ε, the implied constant is independent of $j \leqslant r$. Substituting this into (8.9) with $j = r$ gives the result. $\square$

Lemma 8.5. *Let $k \leqslant (\log x)^{1/5}$ be sufficiently large in terms of θ. Then we have:*

(i) $|\lambda_{\mathbf{d}}| \ll k^{-k}(\log R)^k$;

(ii) $w_n \ll k^{-2k}(\log x)^{2k} \prod_{i=1}^k \prod_{p|L_i(n), p\nmid B} 4$;

(iii) $w_n \ll R^{2+o(1)}$.

Proof. Substituting in our choice of $y_{\mathbf{r}}$, we have for $\mathbf{d} \in \mathcal{D}_k$,

$$|\lambda_{\mathbf{d}}| = d \sum_{\mathbf{d}|\mathbf{r}} \frac{y_{\mathbf{r}}}{\varphi_\omega(r)} = \frac{dW^k B^k \mathfrak{S}_{WB}(\mathcal{L})}{\varphi_\omega(d)\varphi(WB)^k} \sum_{\mathbf{d}|\mathbf{r}\in\mathcal{D}_k} \frac{1}{\varphi_\omega(r/d)} F\bigg(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\bigg). \tag{8.18}$$

We obtain an upper bound for (8.18) by replacing the $\log r_i/\log R$ in the argument of F with $\sigma_i = (\log r_i/d_i)/\log R$, since F is decreasing in each argument.

We now estimate the sum using Lemma 8.4. We see from (7.10) that F is of the form $\Phi(\sum_{i=1}^k t_i)\prod_{i=1}^k G(t_i)$, and we have a bound on G, Φ which corresponds to $\Omega_G = O(kT_k)$ (where $T_k = k\log k$ is the constant given by (7.10)). Since $k \leqslant (\log x)^{1/5}$, we see that $k^2 T_k = o(\log R/\log\log R)$. Finally, we note that the condition $\mathbf{r} \in \mathcal{D}_k$ forces $(r_j, dW_j) = 1$ for integers $W_1, \dots, W_j \leqslant x^{O(k)}$ which are all a multiple of WB. Thus we can apply Lemma 8.4, which gives

$$\sum_{\mathbf{d}|\mathbf{r}\in\mathcal{D}_k} \frac{F(\sigma_1,\dots,\sigma_k)}{\varphi_\omega(r/d)} \leqslant \frac{\varphi(WB)^k}{W^k B^k} \prod_{p\nmid WB}\bigg(1 + \frac{\omega(p)}{p-\omega(p)}\bigg)\bigg(1-\frac{1}{p}\bigg)^k \int\!\!\cdots\!\!\int_{t_1,\dots,t_k \geqslant 0} H(t_1,\dots,t_k)\,dt_1\dots dt_k, \tag{8.19}$$

where

$$H(t_1,\dots,t_k) = F(t_1,\dots,t_k) + O\bigg(\frac{k^2 T_k \log\log R}{\log R} F_1(t_1,\dots,t_k)\bigg). \tag{8.20}$$

Substituting (8.19) into (8.18), noting that the singular series cancel and that $H \leqslant (1+o(1))F_1$, we have

$$\begin{aligned} |\lambda_{\mathbf{d}}| &\leqslant (1+o(1))(\log R)^k \int\!\!\cdots\!\!\int_{t_1,\dots,t_k\geqslant 0} F_1(t_1,\dots,t_k)\,dt_1\dots dt_k \\ &\ll (\log R)^k \bigg(\int_0^{U_k} \frac{dt}{1+T_k t}\bigg)^k \ll \bigg(\frac{\log R}{k}\bigg)^k. \end{aligned} \tag{8.21}$$

This gives claim (i). Claim (ii) now follows from this bound and the definition (7.5) of w_n, recalling that $\lambda_{\mathbf{d}} = 0$ unless $d_1, \dots, d_k$ are all square-free and coprime to B. Finally, for claim (iii), the fact that $\lambda_{\mathbf{d}}$ is supported on $d = d_1\cdots d_k < R$ gives

$$w_n \ll \frac{(\log R)^{2k}}{k^{2k}} \bigg(\sum_{d_1\cdots d_k < R} 1\bigg)^2 \ll R^{2+o(1)} \bigg(\sum_{d_1\cdots d_k<R} \frac{1}{d_1\dots d_k}\bigg)^2 \ll R^{2+o(1)}. \qquad \square$$

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

We will eventually be interested in the quantities I_k, J_k considered in the following lemma.

Lemma 8.6. *Given a square-integrable function $G:\mathbb{R}^k\to\mathbb{R}$, let*

$$I_k(G)=\int_0^\infty\cdots\int_0^\infty G^2\,dt_1\dots dt_k,\quad J_k(G)=\int_0^\infty\cdots\int_0^\infty\Big(\int_0^\infty G\,dt_k\Big)^2 dt_1\dots dt_{k-1}.$$

Let F, F_1, F_2 be as given by (7.10) and (7.12). Then

$$\frac{1}{(2k\log k)^k}\ll I_k(F)\leqslant\frac{1}{(k\log k)^k},\quad\frac{\log k}{k}\ll\frac{J_k(F)}{I_k(F)}\ll\frac{\log k}{k},$$
$$I_k(F)\leqslant I_k(F_1)\leqslant I_k(F_2)/k^2\ll I_k(F),\quad J_k(F)\leqslant J_k(F_1)\leqslant J_k(F_2)/k^2\ll J_k(F).$$

Proof. A minor adaption of the argument of [May15, § 7] to account for the slightly different definition of F shows

$$\begin{aligned}J_k(F)&\geqslant\int\cdots\int_{\sum_{i=1}^{k-1}t_i<9/10-U_k}\Big(\int_0^\infty F_1\,dt_k\Big)^2 dt_1\dots dt_{k-1}\\&\gg\int\cdots\int_{t_1,\dots,t_{k-1}\geqslant 0}\Big(\int_0^\infty F_1\,dt_k\Big)^2 dt_1\dots dt_{k-1}=J_k(F_1).\end{aligned}\tag{8.22}$$

Applying the same concentration of measure argument to $I_k(F)$ yields

$$I_k(F)\geqslant\int\cdots\int_{\sum_{i=1}^{k}t_i<9/10}F_1^2\,dt_1\dots dt_k\gg\int\cdots\int_{t_1,\dots,t_k\geqslant 0}F_1^2\,dt_1\dots dt_k=I_k(F_1).\tag{8.23}$$

We also have the trivial bounds $I_k(F)\leqslant I_k(F_1)\leqslant k^{-2}I_k(F_2)$ and $J_k(F)\leqslant J_k(F_1)\leqslant k^{-2}J_k(F_2)$. For our choice of ψ, T_k, U_k from (7.10), we see that

$$\int_0^\infty\frac{\psi(t/U_k)\,dt}{1+T_kt}=\int_0^{9U_k/10}\frac{dt}{1+T_kt}+O\Big(\int_{9U_k/10}^{U_k}\frac{dt}{1+T_kt}\Big)=\frac{1}{2k}+O\Big(\frac{1}{k\log k}\Big),\tag{8.24}$$
$$\int_0^\infty\frac{\psi(t/U_k)^2\,dt}{(1+T_kt)^2}=\int_0^{9U_k/10}\frac{dt}{(1+T_kt)^2}+O\Big(\int_{9U_k/10}^{U_k}\frac{dt}{(1+T_kt)^2}\Big)=\frac{1+O(k^{-1/2})}{k\log k},\tag{8.25}$$
$$\int_0^\infty\frac{\psi(t/2)^2\,dt}{(1+T_kt)^2}=\int_0^{9/5}\frac{dt}{(1+T_kt)^2}+O\Big(\int_{9/5}^{2}\frac{dt}{(1+T_kt)^2}\Big)=\frac{1+O(k^{-1})}{k\log k}.\tag{8.26}$$

From these bounds it follows immediately that $k^{-2}J_k(F_2)\ll J_k(F_1)$, $k^{-2}I_k(F_2)\ll I_k(F_1)$ and

$$\frac{J_k(F_1)}{I_k(F_1)}=\frac{\log k}{4k}\Big(1+O\Big(\frac{1}{\log k}\Big)\Big).\tag{8.27}$$

Combining these statements gives the bounds of the lemma. □

9. Proof of propositions

We see that Lemmas 8.1, 8.5 and 8.6 verify the claims at the end of Proposition 6.1 for w_n given by (7.5), $I_k=I_k(F)$ and $J_k=J_K(F)$. It therefore remains to establish the four main claims of Proposition 6.1, which we now do in turn. To obtain results with the desired uniformity in k, we need to perform calculations in a slightly different manner from the corresponding ones in [May15].

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

PROPOSITION 9.1. *Let w_n be as described in § 7. Then we have*

$$\sum_{n\in\mathcal{A}(x)} w_n = \left(1+O\left(\frac{1}{(\log x)^{1/10}}\right)\right)\frac{B^k}{\varphi(B)^k}\mathfrak{S}_B(\mathcal{L})\#\mathcal{A}(x)(\log R)^k I_k(F).$$

Proof. We recall $W=\prod_{p\leqslant 2k^2, p\nmid B} p < \exp((\log x)^{2/5})$, and consider the summation over n in the residue class $v_0 \pmod W$. If $(\prod_{i=1}^k L_i(v_0), W)\neq 1$ then we have $w_n=0$, and so we restrict our attention to v_0 with $(\prod_{i=1}^k L_i(v_0), W)=1$. We substitute the definition (7.5) of w_n, expand the square and swap the order of summation. This gives

$$\sum_{\substack{n\in\mathcal{A}(x)\\ n\equiv v_0 \ (\mathrm{mod}\ W)}} w_n = \sum_{\mathbf{d},\mathbf{e}\in\mathcal{D}_k}\lambda_{\mathbf{d}}\lambda_{\mathbf{e}}\sum_{\substack{n\in\mathcal{A}(x)\\ n\equiv v_0\ (\mathrm{mod}\ W)\\ [d_i,e_i]|L_i(n)\ \forall i}} 1. \tag{9.1}$$

By our choice of support of the $\lambda_{\mathbf{d}}$, there is no contribution unless $(d_ie_i, d_je_j)=1$ for all $i\neq j$. In this case, given $\mathbf{d},\mathbf{e}\in\mathcal{D}_k$ (so, in particular, $(d_je_j, a_jW)=1$ for $1\leqslant j\leqslant k$), we can combine the congruence conditions by the Chinese remainder theorem, and see that the inner sum is $\mathcal{A}(x;q,a)$ for some a and for $q=W[\mathbf{d},\mathbf{e}]$. We let $E_q^{(1)}=\max_a|\#\mathcal{A}(x;q,a)-\#\mathcal{A}(x)/q|$, and substitute $\#\mathcal{A}(x;q,a)=\#\mathcal{A}(x)/q+O(E_q^{(1)})$ into (9.1).

We first show the contribution from the errors $E_q^{(1)}$ is small. There are $O(\tau_{3k}(q))$ ways of writing $q=W[\mathbf{d},\mathbf{e}]$ and all such q are square-free, coprime to B and less than $R^2W<x^\theta$ (since $\lambda_{\mathbf{d}}$ is supported on $d<R\leqslant x^{\theta/3}$). Since $|\lambda_{\mathbf{d}}|\ll(\log x)^k$ by Lemma 8.5, these contribute

$$\begin{aligned}\sum_{\mathbf{d},\mathbf{e}\in\mathcal{D}_k}|\lambda_{\mathbf{d}}\lambda_{\mathbf{e}}|E_q^{(1)} &\ll (\log x)^{2k}\sum_{q<R^2W,(q,B)=1}\mu^2(q)\tau_{3k}(q)E_q^{(1)}\\ &\ll(\log x)^{2k}\left(\sum_{q<R^2W,(q,B)=1}\mu^2(q)\tau_{3k}(q)^2E_q^{(1)}\right)^{1/2}\left(\sum_{q<R^2W,(q,B)=1}\mu^2(q)E_q^{(1)}\right)^{1/2}.\end{aligned} \tag{9.2}$$

We apply Hypothesis 1 to estimate these terms. Using $E_q^{(1)}\ll\#\mathcal{A}(x)/q$ for the first sum, and the average of $E_q^{(1)}$ for the second sum, we see the contribution is

$$\ll(\log x)^{2k}\left(\#\mathcal{A}(x)\sum_{q<R}\frac{\tau_{9k^2}(q)}{q}\right)^{1/2}\left(\frac{\#\mathcal{A}(x)}{(\log x)^{100k^2}}\right)^{1/2}\ll\frac{\#\mathcal{A}(x)}{W(\log x)^{2k^2}}. \tag{9.3}$$

By Lemmas 8.1 and 8.6, we see that this is $o(\#\mathcal{A}(x)\mathfrak{S}_B(\mathcal{L})I_k(F)/W)$, and so will be negligible compared with our main term.

We now consider the main term. We substitute our expression (7.8) for $\lambda_{\mathbf{d}}$ in terms of $y_{\mathbf{r}}$ to give

$$\frac{\#\mathcal{A}(x)}{W}{\sum_{\mathbf{d},\mathbf{e}\in\mathcal{D}_k}}'\frac{\lambda_{\mathbf{d}}\lambda_{\mathbf{e}}}{[\mathbf{d},\mathbf{e}]}=\frac{\#\mathcal{A}(x)}{W}\sum_{\mathbf{r},\mathbf{s}\in\mathcal{D}_k}\frac{y_{\mathbf{r}}y_{\mathbf{s}}}{\varphi_\omega(r)\varphi_\omega(s)}{\sum_{\mathbf{d}|\mathbf{r},\mathbf{e}|\mathbf{s}}}'\frac{\mu(d)\mu(e)de}{[\mathbf{d},\mathbf{e}]}, \tag{9.4}$$

where $\sum'$ indicates the restriction that $(d_ie_i,d_je_j)=1$ for all $i\neq j$. By multiplicativity, we can write the inner sum as $\prod_{p|rs}S_p(\mathbf{r},\mathbf{s})$, where, for $\mathbf{r},\mathbf{s}$ such that $p|rs$ and $y_{\mathbf{r}}y_{\mathbf{s}}\neq 0$, we have

$$S_p(\mathbf{r},\mathbf{s})={\sum_{\substack{\mathbf{d}|\mathbf{r},\mathbf{e}|\mathbf{s}\\ d_i,e_i|p\ \forall i}}}'\frac{\mu(d)\mu(e)de}{[\mathbf{d},\mathbf{e}]}=\begin{cases}p-1, & p|(\mathbf{r},\mathbf{s}),\\ -1, & p|r, p|s, p\nmid(\mathbf{r},\mathbf{s}),\\ 0, & (p|r \text{ and } p\nmid s) \text{ or } (p|s \text{ and } p\nmid r).\end{cases} \tag{9.5}$$

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

(We remind the reader that in our notation $r=\prod_{i=1}^k r_i$, that $(\mathbf{r},\mathbf{s})=\prod_{i=1}^k (r_i,s_i)$, that $[\mathbf{r},\mathbf{s}]=\prod_{i=1}^k [r_i,s_i]$, and similarly for e,d,s.)

Since $\prod_{p|rs} S_p(\mathbf{r},\mathbf{s})=0$ if there is a prime p which divides one of r,s but not the other, we can restrict to $r=s$. We let $A=A(\mathbf{r},\mathbf{s})=r/(\mathbf{r},\mathbf{s})$ be the product of primes dividing r but not $(\mathbf{r},\mathbf{s})$, so that $\prod_{p|rs} S_p(\mathbf{r},\mathbf{s})=\mu(A)\varphi(r)/\varphi(A)$. Given a choice of $\mathbf{r}\in\mathcal{D}_k$ and $A|r$, for each prime $p|A$ there are $\omega(p)-1$ possible choices of which components of $\mathbf{s}$ can be a multiple of p (since there are $\omega(p)$ indices j for which $p\nmid W_j$, but for one of these we have $p|r_j$), and so $\prod_{p|A}(\omega(p)-1)$ choices of $\mathbf{s}$. By Lemma 8.2, for each such choice we have

$$y_{\mathbf{s}}=y_{\mathbf{r}}+O\Big(T_k(Y_{\mathbf{r}}+Y_{\mathbf{s}})\frac{\log A}{\log R}\Big). \tag{9.6}$$

Thus our main term becomes

$$\frac{\#\mathcal{A}(x)}{W}\sum_{\mathbf{r}\in\mathcal{D}_k}\frac{y_{\mathbf{r}}\varphi(r)}{\varphi_\omega(r)^2}\sum_{A|r}\Big(\prod_{p|A}\frac{-(\omega(p)-1)}{p-1}\Big)\Big(y_{\mathbf{r}}+O\Big(T_k(Y_{\mathbf{r}}+Y_{\mathbf{s}})\frac{\log A}{\log R}\Big)\Big). \tag{9.7}$$

Since $y_{\mathbf{r}}\leqslant Y_{\mathbf{r}}/k$ and $Y_{\mathbf{r}}Y_{\mathbf{s}}\ll Y_{\mathbf{r}}^2+Y_{\mathbf{s}}^2$, the contribution of the error here is

$$\begin{aligned}&\ll\frac{T_k}{k}\frac{\#\mathcal{A}(x)}{W}\sum_{\mathbf{r}\in\mathcal{D}_k}\frac{\varphi(r)Y_{\mathbf{r}}^2}{\varphi_\omega(r)^2}\sum_{A|r}\frac{\omega(A)}{\varphi(A)}\sum_{p|A}\frac{\log p}{\log R}\\ &\ll\frac{T_k}{k}\frac{\#\mathcal{A}(x)}{W\log R}\sum_{p>2k^2}\log p\sum_{\substack{(A,WB)=1\\ p|A}}\frac{\omega(A)}{\varphi(A)}\sum_{\substack{\mathbf{r}\in\mathcal{D}_k\\ A|r}}\frac{\varphi(r)Y_{\mathbf{r}}^2}{\varphi_\omega(r)^2}.\end{aligned} \tag{9.8}$$

We let $\mathbf{r}'$ be the vector formed by removing from $\mathbf{r}$ any factors of A, so $r_i'=r_i/(r_i,A)$. Since $Y_{\mathbf{r}}$ is decreasing, we have $Y_{\mathbf{r}'}\geqslant Y_{\mathbf{r}}$. Given $\mathbf{r}'$, there are $O(\omega(A))$ possible choices of $\mathbf{r}$. Thus, swapping the summation to $\mathbf{r}'$, and letting $A=pA'$, we obtain the bound

$$\ll\frac{T_k}{k}\frac{\#\mathcal{A}(x)}{W\log R}\Big(\sum_{p>2k^2}\frac{\omega(p)^2\log p}{\varphi_\omega(p)^2}\Big)\Big(\sum_{(A',WB)=1}\frac{\omega(A')^2}{\varphi_\omega(A')^2}\Big)\Big(\sum_{\mathbf{r}'\in\mathcal{D}_k}\frac{\varphi(r')Y_{\mathbf{r}'}^2}{\varphi_\omega(r')^2}\Big). \tag{9.9}$$

The first two terms in parentheses can both be seen to be $O(1)$, since all prime factors are greater than $2k^2$. We estimate the final term by Lemma 8.4 (taking $\Omega_G=O(T_k^2)$). This gives a bound for (9.9) of size

$$\ll\frac{T_kW^{k-1}B^k(\log R)^{k-1}\mathfrak{S}_{WB}(\mathcal{L})^2\#\mathcal{A}(x)}{k\varphi(WB)^k}\prod_{p\nmid WB}\Big(1+\frac{\omega(p)(p-1)}{(p-\omega(p))^2}\Big)\Big(1-\frac{1}{p}\Big)^kI_k(F_2). \tag{9.10}$$

We note that

$$\prod_{p\nmid WB}\Big(1+\frac{\omega(p)(p-1)}{(p-\omega(p))^2}\Big)\Big(1-\frac{1}{p}\Big)^k=\prod_{p\nmid WB}\Big(1+\frac{\omega(p)}{p-\omega(p)}\Big)\Big(1+O\Big(\frac{k^2}{p^2}\Big)\Big)\Big(1-\frac{1}{p}\Big)^k\ll\mathfrak{S}_{WB}(\mathcal{L})^{-1}, \tag{9.11}$$

since the product is only over primes $p>2k^2$. Using $I_k(F_2)\ll k^2I_k(F)$ from Lemma 8.6, we see that (9.10) is

$$\ll\frac{kT_kW^{k-1}B^k\mathfrak{S}_{WB}(\mathcal{L})\#\mathcal{A}(x)(\log R)^{k-1}}{\varphi(WB)^k}I_k(F). \tag{9.12}$$

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

This is negligible, and can be absorbed into the error term in the statement of the lemma. We now consider the main term. We have

$$\frac{\#\mathcal{A}(x)}{W}\sum_{\mathbf{r}\in\mathcal{D}_k}\frac{y_{\mathbf{r}}^2\varphi(r)}{\varphi_\omega(r)^2}\sum_{A|r}\prod_{p|A}\frac{-(\omega(p)-1)}{p-1}=\frac{\#\mathcal{A}(x)}{W}\sum_{\mathbf{r}\in\mathcal{D}_k}\frac{y_{\mathbf{r}}^2}{\varphi_\omega(r)}. \tag{9.13}$$

We estimate the inner sum here by applying Lemma 8.4 (again with $\Omega_G=T_k^2$). This gives

$$\begin{aligned}\sum_{\mathbf{r}\in\mathcal{D}_k}\frac{y_{\mathbf{r}}^2}{\varphi_\omega(r)}&=\frac{W^kB^k\mathfrak{S}_{WB}(\mathcal{L})^2}{\varphi(WB)^k}(\log R)^k\prod_{p\nmid WB}\left(1+\frac{\omega(p)}{p-\omega(p)}\right)\left(1-\frac{1}{p}\right)^kI_k(F)\\&\quad+O\left(\frac{W^kB^k\mathfrak{S}_{WB}(\mathcal{L})^2}{\varphi(WB)^k}(\log R)^k\prod_{p\nmid WB}\left(1+\frac{\omega(p)}{p-\omega(p)}\right)\left(1-\frac{1}{p}\right)^k\frac{kT_k^2\log\log R}{\log R}I_k(F_1)\right)\\&=\frac{W^kB^k\mathfrak{S}_{WB}(\mathcal{L})}{\varphi(WB)^k}(\log R)^k\left(1+O\left(\frac{kT_k^2\log\log R}{\log R}\right)\right)I_k(F).\end{aligned} \tag{9.14}$$

In the last line we have used the fact that $I_k(F_1)\ll I_k(F)$ given by Lemma 8.6. Putting this all together (and recalling $k\leqslant(\log x)^{1/5}$ and $T_k=k\log k$), we have shown that

$$\sum_{\substack{n\in\mathcal{A}(x)\\ n\equiv v_0\ (\mathrm{mod}\ W)}}w_n=\left(1+O\left(\frac{1}{(\log x)^{1/10}}\right)\right)\frac{W^{k-1}B^k\mathfrak{S}_{WB}(\mathcal{L})\#\mathcal{A}(x)}{\varphi(WB)^k}(\log R)^kI_k(F). \tag{9.15}$$

Summing this over the $\varphi_\omega(W)$ residue classes v_0 (mod W) such that $(\prod_{i=1}^k L(v_0),W)=1$ then gives the result. $\square$

PROPOSITION 9.2. *Let w_n be as described in § 7. Let $L\in\mathcal{L}$ satisfy $L(n)>R$ for $n\in[x,2x]$ and*

$$\sum_{\substack{q\leqslant x^\theta\\(q,B)=1}}\max_{(L(a),q)=1}\left|\#\mathcal{P}_{L,\mathcal{A}}(x;q,a)-\frac{\#\mathcal{P}_{L,\mathcal{A}}(x)}{\varphi_L(q)}\right|\ll\frac{\#\mathcal{P}_{L,\mathcal{A}}(x)}{(\log x)^{100k^2}}.$$

Then we have

$$\begin{aligned}\sum_{n\in\mathcal{A}(x)}\mathbf{1}_{\mathcal{P}}(L(n))w_n&=\left(1+O\left(\frac{1}{(\log x)^{1/10}}\right)\right)\frac{B^{k-1}}{\varphi(B)^{k-1}}\mathfrak{S}_B(\mathcal{L})\#\mathcal{P}_{L,\mathcal{A}}(x)(\log R)^{k+1}J_k(F)\prod_{\substack{p|a_m\\ p\nmid WB}}\frac{p-1}{p}\\&\quad+O\left(\frac{B^k}{\varphi(B)^k}\mathfrak{S}_B(\mathcal{L})\#\mathcal{A}(x)(\log R)^{k-1}I_k(F)\right).\end{aligned}$$

Proof. We let $L(n)=L_m(n)=a_mn+b_m$ be the mth element of $\mathcal{L}$.

Again we split the sum into residue classes $n\equiv v_0$ (mod W). If $(\prod_{i=1}^kL_i(v_0),W)\neq1$ then we have $w_n=0$, and so we restrict our attention to v_0 with $(\prod_{i=1}^kL_i(v_0),W)=1$. We substitute the definition (7.5) of w_n, expand the square and swap the order of summation. This gives

$$\sum_{\substack{n\in\mathcal{A}(x)\\ n\equiv v_0\ (\mathrm{mod}\ W)}}\mathbf{1}_{\mathcal{P}}(L(n))w_n=\sum_{\mathbf{d},\mathbf{e}\in\mathcal{D}_k}\lambda_{\mathbf{d}}\lambda_{\mathbf{e}}\sum_{\substack{n\in\mathcal{A}(x)\\ n\equiv v_0\ (\mathrm{mod}\ W)\\ [d_i,e_i]|L_i(n)\ \forall i}}\mathbf{1}_{\mathcal{P}}(L(n)). \tag{9.16}$$

We first show that there is no contribution to our sum from $\lambda_{\mathbf{d}}$ for which $(d_j,a_jb_m-a_mb_j)\neq1$ for some $j\neq m$. If $p|d_j$ then the inner sum requires that $p|a_jn+b_j$. However, if we also have $p|a_jb_m-b_ja_m$ then this means $p|a_mn+b_m$ (since $(a_j,b_j)=1$ by admissibility of $\mathcal{L}$). Since there is

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

no contribution to our sum unless $L(n)=L_m(n)=a_mn+b_m$ is a prime and since $d_j<R<L(n)$ by the support of $\lambda_{\mathbf{d}}$ and assumption of the lemma, we see that there is no contribution from $\lambda_{\mathbf{d}}$ with $(d_j,a_jb_m-a_mb_j)\neq 1$.

Thus we may restrict the support of $\lambda_{\mathbf{d}}$ to $\mathcal{D}_k'$, defined by

$$\mathcal{D}_k'=\{\mathbf{d}\in\mathbb{R}^k:\mu^2(d)=1,(d_j,W_j')=1\ \forall j\},\quad W_j'=\prod_{p|W_j(a_jb_m-a_mb_j)}p. \tag{9.17}$$

We write $\lambda_{\mathbf{d}}'$ for $\lambda_{\mathbf{d}}$ with this restricted support. We see from this that $p|W_j'/W_j$ if and only if $p\nmid a_m$ and j was the chosen index for the residue class $-b_m\overline{a_m}\pmod p$ (for our fixed set of choices of residue classes given in § 7).

We now observe that given $\mathbf{d},\mathbf{e}\in\mathcal{D}_k'$, the inner sum of (9.16) is empty unless $(d_ie_i,d_je_j)=1$ for all $i\neq j$ (since otherwise the divisibility conditions are incompatible). If $(d_ie_i,d_je_j)=1\ \forall i\neq j$, then we can combine the conditions by the Chinese remainder theorem. This shows the sum is $\#\mathcal{P}_{L,\mathcal{A}}(x;q,a)$ for $q=W[\mathbf{d},\mathbf{e}]$ and some a. We note $\#\mathcal{P}_{L,\mathcal{A}}(x;q,a)\neq 0$ if and only if $(L(a),q)=1$, which occurs if and only if $d_m=e_m=1$. For such a choice of $\mathbf{d},\mathbf{e}$, we write $\#\mathcal{P}_{L,\mathcal{A}}(x;q,a)=\#\mathcal{P}_{L,\mathcal{A}}(x)/\varphi_L(q)+O(E_q^{(2)})$, where $E_q^{(2)}=\max_{(a,q)=1}|\#\mathcal{P}_{L,\mathcal{A}}(x;q,a)-\#\mathcal{P}_{L,\mathcal{A}}(x)/\varphi_L(q)|$.

We treat the error term $E_q^{(2)}$ in the same manner as we treated $E_q^{(1)}$ in the proof of Proposition 9.1. We note that for all $\mathbf{d},\mathbf{e}\in\mathcal{D}_k'$ we have $(q,B)=1$, allowing us to use Proposition 6.1 for the average of $E_q^{(2)}$. We also note that trivially $\#\mathcal{P}_{L,\mathcal{A}}(x;q,a)\ll\#\mathcal{A}(x;q,a)$, which gives us the bound $E_q^{(2)}\ll\#\mathcal{A}(x)/\varphi_L(q)$. Thus the same argument shows that these error terms contribute $O(\#\mathcal{A}(x)W^{-1}(\log x)^{-2k^2})$.

We now consider the main term, given by

$$\frac{\#\mathcal{P}_{L,\mathcal{A}}(x)}{\varphi_L(W)}\sideset{}{'}\sum_{\substack{\mathbf{d},\mathbf{e}\in\mathcal{D}_k'\\ d_m=e_m=1}}\frac{\lambda_{\mathbf{d}}'\lambda_{\mathbf{e}}'}{\varphi_L([\mathbf{d},\mathbf{e}])}, \tag{9.18}$$

where we recall $\sum'$ indicates the sum is restricted to $(d_ie_i,d_je_j)=1$ for all $i\neq j$. We change variables to $y_{\mathbf{r}}^{(m)}$, satisfying

$$y_{\mathbf{r}}^{(m)}=\mu(r)\varphi_\omega(r)\sum_{\mathbf{r}|\mathbf{d},\,d_m=1}\frac{\lambda_{\mathbf{d}}'}{\varphi_L(d)},\quad \lambda_{\mathbf{d}}'=\mu(d)\varphi_L(d)\sum_{\mathbf{d}|\mathbf{r}}\frac{y_{\mathbf{r}}^{(m)}}{\varphi_\omega(r)}. \tag{9.19}$$

We see from (9.19) that the $y_{\mathbf{r}}^{(m)}$ are supported on $\mathbf{r}\in\mathcal{D}_k'$ with $r_m=1$. Substituting our expression (9.19) for $\lambda_{\mathbf{d}}'$ into our main term (9.18) gives

$$\frac{\#\mathcal{P}_{L,\mathcal{A}}(x)}{\varphi_L(W)}\sideset{}{'}\sum_{\substack{\mathbf{d},\mathbf{e}\in\mathcal{D}_k'\\ d_m=e_m=1}}\frac{\lambda_{\mathbf{d}}'\lambda_{\mathbf{e}}'}{\varphi_L([\mathbf{d},\mathbf{e}])}=\frac{\#\mathcal{P}_{L,\mathcal{A}}(x)}{\varphi_L(W)}\sum_{\substack{\mathbf{r},\mathbf{s}\in\mathcal{D}_k'\\ r_m=s_m=1}}\frac{y_{\mathbf{r}}^{(m)}y_{\mathbf{s}}^{(m)}}{\varphi_\omega(r)\varphi_\omega(s)}\prod_{p|rs}S_p^{(m)}(\mathbf{r},\mathbf{s}), \tag{9.20}$$

where now, if $\mathbf{r}$ and $\mathbf{s}$ are such that $y_{\mathbf{r}}^{(m)}y_{\mathbf{s}}^{(m)}\neq 0$ (so $r_m=s_m=1$, $\mathbf{r},\mathbf{s}\in\mathcal{D}_k'$) and $p|rs$, we have

$$S_p^{(m)}(\mathbf{r},\mathbf{s})=\sideset{}{'}\sum_{\substack{\mathbf{d}|\mathbf{r},\mathbf{e}|\mathbf{s}\\ d_i,e_i|p\ \forall i\\ d_m=e_m=1}}\frac{\mu(d)\mu(e)\varphi_L(d)\varphi_L(e)}{\varphi_L([\mathbf{d},\mathbf{e}])}=\begin{cases}p-2, & p|(\mathbf{r},\mathbf{s}),p\nmid a_m,\\ p-1, & p|(\mathbf{r},\mathbf{s}),p|a_m,\\ -1, & p|r,p|s,p\nmid(\mathbf{r},\mathbf{s}),\\ 0, & (p|r\text{ and }p\nmid s)\text{ or }(p|s\text{ and }p\nmid r),\end{cases} \tag{9.21}$$

so again we may restrict to $r=s$. We use the following lemma to relate $y_{\mathbf{r}}^{(m)}$ to $y_{\mathbf{r}}$.

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

Lemma 9.3. *Let $\mathbf{r}\in\mathcal{D}_k'$ with $r_m=1$, and let $t_i=\log r_i/\log R$ for $i\neq m$. Then we have*

$$y_{\mathbf{r}}^{(m)}=\log R\frac{\varphi(a_mWB)W^{k-1}B^{k-1}\mathfrak{S}_{WB}(\mathcal{L})}{a_m\varphi(WB)^k}\int_0^\infty H(t_1,\dots,t_k)\,dt_m,$$

where

$$H(t_1,\dots,t_k)=F(t_1,\dots,t_k)+O\biggl(\frac{T_k(\log\log R)^2}{\log R}F_2(t_1,\dots,t_k)\biggr).$$

We first complete the proof of Proposition 9.2, and then establish the lemma. Given $\mathbf{r},\mathbf{s}\in\mathcal{D}_k'$ with $r_m=s_m=1$ and $r=s$, let $A=A(\mathbf{r},\mathbf{s})$ be the product of primes dividing r but not $(\mathbf{r},\mathbf{s})$. We note that Lemma 9.3 implies the bound

$$y_{\mathbf{r}}^{(m)}\ll\frac{\varphi(a_mWB)\log R}{kWB}Y_{\mathbf{r}},\tag{9.22}$$

since we have the bound $\int_0^\infty F_2(t_1,\dots,t_k)\,dt_m\ll F_2(t_1,\dots,t_{m-1},0,t_{m+1},\dots,t_k)/k$ (this follows from the definition (7.12) of F_2). Analogously to Lemma 8.2, we have (for $A>1$)

$$\begin{aligned}y_{\mathbf{r}}^{(m)}&=y_{\mathbf{s}}^{(m)}+O\biggl(\frac{T_k}{k}(Y_{\mathbf{r}}+Y_{\mathbf{s}})\frac{\varphi(a_mWB)}{a_mWB}(\log A+(\log\log R)^2)\biggr)\\&=y_{\mathbf{s}}^{(m)}+O\biggl(\frac{T_k}{k}(Y_{\mathbf{r}}+Y_{\mathbf{s}})\frac{\varphi(a_mWB)}{a_mWB}(\log A)(\log\log R)^2\biggr).\end{aligned}\tag{9.23}$$

Substituting this into our main term (9.20), we are left to estimate

$$\sum_{\substack{\mathbf{r},\mathbf{s}\in\mathcal{D}_k'\\ r_m=s_m=1\\ r=s}}\frac{y_{\mathbf{r}}^{(m)}}{\varphi_\omega(r)^2}\biggl(\prod_{p|rs}S_p^{(m)}(\mathbf{r},\mathbf{s})\biggr)\biggl(y_{\mathbf{r}}^{(m)}+O\biggl(\frac{T_k}{k}(Y_{\mathbf{r}}+Y_{\mathbf{s}})\frac{\varphi(a_mWB)}{a_mWB}(\log A)(\log\log R)^2\biggr)\biggr).\tag{9.24}$$

We note that for $r=s$ the value of $\prod_{p|rs}S_p^{(m)}(\mathbf{r},\mathbf{s})$ depends only on r and A. Substituting this value for $S_p^{(m)}(\mathbf{r},\mathbf{s})$ gives a main term

$$\sum_{\substack{\mathbf{r}\in\mathcal{D}_k'\\ r_m=1}}\frac{(y_{\mathbf{r}}^{(m)})^2}{\varphi_\omega(r)^2}\biggl(\prod_{p|r}(\varphi_L(p)-1)\biggr)\sum_{A|r}\biggl(\prod_{p|A}\frac{-1}{\varphi_L(p)-1}\biggr)\sum_{\substack{\mathbf{s}\in\mathcal{D}_k'\\ A(\mathbf{r},\mathbf{s})=A}}1,\tag{9.25}$$

and (using (9.22) and $Y_{\mathbf{r}}Y_{\mathbf{s}}\leqslant Y_{\mathbf{r}}^2+Y_{\mathbf{s}}^2$) an error term of size

$$\ll\frac{T_k\varphi(a_mWB)^2\log R}{k^2a_m^2B^2W^2}\sum_{\substack{\mathbf{r}\in\mathcal{D}_k'\\ r_m=1}}\frac{Y_{\mathbf{r}}^2\prod_{p|r}(\varphi_L(p)-1)}{\varphi_\omega(r)^2}\sum_{A|r}\frac{\log A}{\prod_{p|A}(\varphi_L(p)-1)}\sum_{\substack{\mathbf{s}\in\mathcal{D}_k'\\ A(\mathbf{r},\mathbf{s})=A}}(\log\log R)^2.\tag{9.26}$$

We first estimate the inner sum over $\mathbf{s}$ which occurs in both terms. We fix a choice of $\mathbf{r}\in\mathcal{D}_k'$ and $A=A(\mathbf{r},\mathbf{s})$ with $A|r$. For each prime $p|A$, we count how many components of $\mathbf{s}$ can be a multiple of p, subject to the constraints that $p\nmid(s_i,r_i)$ and $p\nmid(s_i,W_i')$ for all i. If $p|A,p\nmid a_m$ then there are $\omega(p)-2$ possible choices of which component of $\mathbf{s}$ can be a multiple of p (there are $\omega(p)-1$ indices $j\neq m$ for which $p\nmid W_j'$, but for one of these indices $p|r_j$). If $p|A$ and $p|a_m$, then instead there are $\omega(p)-1$ choices (since there are $\omega(p)$ indices $j\neq m$ for which $p\nmid W'_j$, but for one of these indices $p|r_j$). Thus we have

$$\sum_{\substack{\mathbf{s}\in\mathcal{D}'_k\\ A(\mathbf{r},\mathbf{s})=A}} 1=\prod_{p|A,\, p\nmid a_m}(\omega(p)-2)\prod_{p|A,\, p|a_m}(\omega(p)-1). \tag{9.27}$$

We now consider the error term (9.26). We follow an analogous argument to that in the proof of Proposition 9.1. Substituting our expression (9.27) for the inner sum and crudely bounding the multiplicative functions gives a bound

$$\ll \frac{T_k\varphi(a_mWB)^2(\log R)(\log\log R)^2}{k^2a_m^2B^2W^2}\sum_{(A,WB)=1}\log A\frac{\omega(A)}{\varphi_\omega(A)^2}\sum_{\substack{\mathbf{r}\in\mathcal{D}'_k\\ r_m=1\\ A|r}}\frac{Y_{\mathbf{r}}^2\varphi(r/A)}{\varphi_\omega(r/A)^2}. \tag{9.28}$$

We let $\mathbf{r}'$ be given by $r'_i=r_i/(r_i,A)$ and see that $Y_{\mathbf{r}'}\geqslant Y_{\mathbf{r}}$. Moreover, we see that there are $O(\omega(A))$ choices of $\mathbf{r}$ given $\mathbf{r}'$. Therefore we obtain the bound

$$\sum_{(A,WB)=1}\log A\frac{\omega(A)}{\varphi_\omega(A)^2}\sum_{\substack{\mathbf{r}\in\mathcal{D}'_k\\ r_m=1\\ A|r}}\frac{Y_{\mathbf{r}}^2\varphi(r/A)}{\varphi_\omega(r/A)^2}\ll\Bigg(\sum_{(A,WB)=1}\log A\frac{\omega(A)^2}{\varphi_\omega(A)^2}\Bigg)\sum_{\substack{\mathbf{r}'\in\mathcal{D}_k\\ r'_m=1}}\frac{Y_{\mathbf{r}'}^2\varphi(r')}{\varphi_\omega(r')^2}. \tag{9.29}$$

Here we have dropped the requirement that $(r',A)=1$ for an upper bound. We substitute $\log A=\sum_{p|A}\log p$, $A=pA'$, and swap the order of summation. This shows the right-hand side of (9.29) is

$$\ll\Bigg(\sum_{p>2k^2}\frac{\omega(p)^2\log p}{\varphi_\omega(p)^2}\Bigg)\Bigg(\sum_{(A',WB)=1}\frac{\omega(A')^2}{\varphi_\omega(A')^2}\Bigg)\Bigg(\sum_{\substack{\mathbf{r}'\in\mathcal{D}'_k\\ r_m=1}}\frac{\varphi(r')Y_{\mathbf{r}'}^2}{\varphi_\omega(r')^2}\Bigg). \tag{9.30}$$

The first two sums are seen to be $O(1)$ since they only involve primes $p>2k^2$. The final sum we estimate using Lemma 8.4. This gives a bound for (9.30) of

$$\begin{aligned}&\ll\frac{W^{k+1}B^{k+1}\mathfrak{S}_{WB}(\mathcal{L})^2(\log R)^{k-1}}{\varphi(WB)^{k+1}}\prod_{p\nmid WB}\left(1+\frac{\omega(p)-1}{p+O(k)}\right)\left(1-\frac{1}{p}\right)^{k-1}\\ &\quad\times\int\cdots\int_{t_1,\dots,t_k\geqslant 0,\, t_m=0}F_2(t_1,\dots,t_k)^2\prod_{i\neq m}dt_i.\end{aligned} \tag{9.31}$$

We see that the product is $O(\mathfrak{S}_{WB}(\mathcal{L})^{-1})$ analogously to (9.11). We also have, from the definition (7.12) of F_2,

$$\begin{aligned}&\int\cdots\int_{t_1,\dots,t_k\geqslant 0,\, t_m=0}F_2(t_1,\dots,t_k)^2\prod_{i\neq m}dt_i\\ &\qquad\leqslant k^2\left(\int_0^\infty\frac{\psi(t/U_k)^2\,dt}{(1+T_kt)^2}\right)^{k-2}\left(\int_0^\infty\frac{\psi(t/2)^2\,dt}{(1+T_kt)^2}\right)\ll k^2T_k^2J_k(F).\end{aligned} \tag{9.32}$$

Putting this together, the contribution of the error term to (9.24) is

$$\ll\frac{T_k^3\varphi(a_mWB)^2W^{k-1}B^{k-1}\mathfrak{S}_{WB}(\mathcal{L})(\log R)^k(\log\log R)^2}{a_m^2\varphi(WB)^{k+1}}J_k(F), \tag{9.33}$$

which contributes only to the error term in the statement of the lemma.

We now consider the main term in (9.24), given by (9.25). Substituting our expression (9.27) for the inner sum and evaluating the sum over A gives

$$\sum_{\substack{\mathbf{r}\in\mathcal{D}_k\\ r_m=1}} \frac{(y_{\mathbf{r}}^{(m)})^2}{\varphi_\omega(r)^2}\bigg(\prod_{p\mid r}(\varphi_L(p)-1)\bigg)\sum_{A\mid r}\bigg(\prod_{p\mid A}\frac{-1}{\varphi_L(p)-1}\bigg)\sum_{\substack{\mathbf{s}\in\mathcal{D}_k'\\ A(\mathbf{r},\mathbf{s})=A}} 1 = \sum_{\substack{\mathbf{r}\in\mathcal{D}_k'\\ r_m=1}}\frac{(y_{\mathbf{r}}^{(m)})^2}{\varphi_\omega(r)}. \tag{9.34}$$

We evaluate this sum using Lemma 8.4. This gives

$$\begin{aligned}\sum_{\substack{\mathbf{r}\in\mathcal{D}_k'\\ r_m=1}}\frac{(y_{\mathbf{r}}^{(m)})^2}{\varphi_\omega(r)} &= \prod_{p\nmid a_m BW}\bigg(1+\frac{\omega(p)-1}{\varphi_\omega(p)}\bigg)\bigg(1-\frac{1}{p}\bigg)^{k-1}\prod_{p\mid a_m, p\nmid WB}\bigg(1+\frac{\omega(p)}{\varphi_\omega(p)}\bigg)\bigg(1-\frac{1}{p}\bigg)^{k-1}\\ &\quad\times(\log R)^{k+1}\frac{\varphi(a_m WB)^2 W^{k-1}B^{k-1}\mathfrak{S}_{WB}(\mathcal{L})^2}{a_m^2\varphi(WB)^{k+1}}\\ &\quad\times\bigg(J_k(H)+O\bigg(\frac{kT_k^2(\log\log R)^2}{\log R}J_k(F_1)\bigg)\bigg).\end{aligned} \tag{9.35}$$

By Lemma 8.6, we have $J_k(F_1)\gg J_k(F)$. From the definition of H, we have

$$\begin{aligned}J_k(H) = J_k\bigg(F+O\bigg(\frac{T_k\log\log R}{\log R}F_2\bigg)\bigg) &= J_k(F)+O\bigg(\frac{T_k\log\log R}{\log R}J_k(F_2)\bigg)\\ &= J_k(F)\bigg(1+O\bigg(\frac{k^2T_k\log\log R}{\log R}\bigg)\bigg).\end{aligned} \tag{9.36}$$

We recall $k\leqslant(\log x)^{1/5}$ and $T_k=k\log k$, so the errors appearing are $o((\log x)^{-1/10})$. Therefore, simplifying the products in (9.35) gives

$$\sum_{\substack{\mathbf{r}\in\mathcal{D}_k'\\ r_m=1}}\frac{(y_{\mathbf{r}}^{(m)})^2}{\varphi_\omega(r)} = \bigg(1+\bigg(\frac{1}{(\log x)^{1/10}}\bigg)\bigg)(\log R)^{k+1}\frac{W^{k-1}B^{k-1}\mathfrak{S}_{WB}(\mathcal{L})}{\varphi(WB)^{k-1}}J_k(F)\prod_{p\mid a_m, p\nmid WB}\frac{p-1}{p}. \tag{9.37}$$

Thus, putting everything together, we have

$$\begin{aligned}\sum_{\substack{n\in\mathcal{A}(x)\\ n\equiv v_0 \pmod W}}\mathbf{1}_{\mathcal{P}}(L(n))w_n &= \frac{\#\mathcal{P}_{L,\mathcal{A}}(x)}{\varphi_L(W)}\sideset{}{'}\sum_{\substack{\mathbf{d},\mathbf{e}\\ d_m=e_m=1}}\frac{\lambda_{\mathbf{d}}'\lambda_{\mathbf{e}}'}{\varphi_L([\mathbf{d},\mathbf{e}])}+O\bigg(\frac{\#\mathcal{A}(x)}{W(\log x)^{2k^2}}\bigg)\\ &= \bigg(1+O\bigg(\frac{1}{(\log x)^{1/10}}\bigg)\bigg)(\log R)^{k+1}\frac{W^{k-1}B^{k-1}\mathfrak{S}_{WB}(\mathcal{L})\#\mathcal{P}_{L,\mathcal{A}}(x)}{\varphi(WB)^{k-1}\varphi_L(W)}J_k(F)\prod_{p\mid a_m, p\nmid WB}\frac{p-1}{p}\\ &\quad+O\bigg(\frac{\#\mathcal{A}(x)}{W(\log x)^{2k^2}}\bigg).\end{aligned} \tag{9.38}$$

Summing over the $\varphi_\omega(W)$ residue classes $v_0 \pmod W$ then gives the result (recalling $(W,B)=1$). $\square$

We now prove Lemma 9.3.

Proof of Lemma 9.3. We recall that $\lambda_{\mathbf{d}}'=\lambda_{\mathbf{d}}$ for $\mathbf{d}\in\mathcal{D}_k'$, and $\lambda_{\mathbf{d}}'=0$ otherwise. Using this, we substitute our expression (7.8) for $\lambda_{\mathbf{d}}$ into the definition (9.19) of $y_{\mathbf{r}}^{(m)}$. For $r_m=1$ and $\mathbf{r}\in\mathcal{D}_k'$,

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

we obtain

$$
\begin{aligned}
y_{\mathbf{r}}^{(m)} &= \mu(r)\varphi_\omega(r) \sum_{\substack{\mathbf{d}\in\mathcal{D}_k' \\ \mathbf{r}|\mathbf{d}, d_m=1}} \frac{\lambda_{\mathbf{d}}}{\varphi_L(d)} = \mu(r)\varphi_\omega(r) \sum_{\substack{\mathbf{e}\in\mathcal{D}_k \\ \mathbf{r}|\mathbf{e}}} \frac{y_{\mathbf{e}}}{\varphi_\omega(e)} \sum_{\substack{\mathbf{d}\in\mathcal{D}_k' \\ \mathbf{r}|\mathbf{d},\, \mathbf{d}|\mathbf{e},\, d_m=1}} \frac{\mu(d)d}{\varphi_L(d)} \\
&= \frac{r\varphi_\omega(r)}{\varphi_L(r)} \sum_{\substack{\mathbf{e}\in\mathcal{D}_k \\ \mathbf{r}|\mathbf{e}}} \frac{y_{\mathbf{e}}}{\varphi_\omega(e)} \prod_{p|e/r} S_p^{\prime(m)}(\mathbf{e},\mathbf{r}),
\end{aligned}
\tag{9.39}
$$

where if $p|e_m$ then $S_p^{\prime(m)}(\mathbf{e},\mathbf{r})=1$ and if $p|e_j/r_j$ with $j\neq m$ then

$$
S_p^{\prime(m)}(\mathbf{e},\mathbf{r}) = \sum_{\substack{\mathbf{d}\in\mathcal{D}_k' \\ d_j|(e_j/r_j,p),\, d_m=1}} \frac{\mu(d)d}{\varphi_L(d)} = \begin{cases} -1/(p-1), & p\nmid a_m W_j', \\ 0, & p|a_m, p\nmid W_j', \\ 1, & p|W_j'/W_j. \end{cases}
\tag{9.40}
$$

(Since $\mathbf{e}\in\mathcal{D}_k$, we have $(e_j,W_j)=1$ and so if $p|e_j/r_j$ we only need consider $p\nmid W_j$.)

We let $e_j=r_ju_jv_j$ for each $j\neq m$, where u_j is the product of primes dividing e_j/r_j but not W_j', and v_j is the product of primes dividing both e_j/r_j and W_j'/W_j. We put $u_m=v_m=1$, and consider e_m separately.

We can restrict to the case when $(u_j,a_m)=1$ for all j, since otherwise the product of $S_p^{\prime(m)}(\mathbf{e},\mathbf{r})$ vanishes. For $\mathbf{e}\in\mathcal{D}_k$, the product in (9.39) is then $\mu(u)/\varphi(u)$ by (9.40). (We recall that in our notation $u=\prod_{i=1}^k u_i$, and similarly for v.) Since $(a_m,W_j'/W_j)=1$ for all j, we can also restrict to $(v_j,a_m)=1$ for all j. (If $p|W_j'/W_j$ then $p|a_mb_j-a_jb_m$, so if $p|a_m$ and $p|W_j'/W_j$, then $p|a_j$ and hence $p|W_j$, meaning $p\nmid W_j'/W_j$.)

We let $\mathbf{r}'=(r_1,\ldots,r_{m-1},e_m,r_{m+1},\ldots,r_k)$. By Lemma 8.2, we have

$$
y_{\mathbf{e}} = y_{\mathbf{r}'} + O\biggl(T_kY_{\mathbf{r}'}\frac{\log uv}{\log R}\biggr).
\tag{9.41}
$$

Substituting this into (9.39) gives

$$
y_{\mathbf{r}}^{(m)} = \frac{r}{\varphi_L(r)}\sum_{e_m}\frac{y_{\mathbf{r}'}}{\varphi_\omega(e_m)}\sum_{\mathbf{u},\mathbf{v}}\frac{\mu(u)}{\varphi(u)\varphi_\omega(uv)} + O\biggl(\frac{T_kr}{\varphi_L(r)\log R}\sum_{\substack{e_m \\ \mathbf{r}'\in\mathcal{D}_k}}\frac{Y_{\mathbf{r}'}}{\varphi_\omega(e_m)}\sum_{\mathbf{u},\mathbf{v}}\frac{\log uv}{\varphi(u)\varphi_\omega(uv)}\biggr),
\tag{9.42}
$$

where the sums over $\mathbf{u},\mathbf{v}$ are subject to the constraints $\mathbf{u}\in\mathcal{D}_k',\mathbf{v}\in\mathcal{D}_k$, $u_m=v_m=1$, $(u,v)=(uv,re_ma_m)=1$, and $v_j|W_j'/W_j$.

We first estimate the error term from (9.42). We have $\log uv\ll u^{1/2}(1+\log v)$, and we drop the requirement that $(u,v)=1$. The sum over u then factorizes as an Euler product, which can be seen to be $O(1)$ since there are $O(\omega(s))$ choices of $\mathbf{u}$ with $u=s$, and we only consider primes $p>2k^2$. Thus we are left to consider the sum over $\mathbf{v}$. For any choice of v there is at most one possible $\mathbf{v}$ with $\prod_{i=1}^k v_i=v$ (since for every prime $p|v$ with $p|W_m$ there is a unique index j such that $p|W_j'/W_j$, and if $p\nmid W_m$ there is no such index). Any such v must have $v|\Delta=\prod_{i\neq m}(a_mb_i-a_ib_m)$ since $v_j|W_j'/W_j$. Thus the sum over $\mathbf{v}$ contributes at most

$$
\begin{aligned}
\sum_{v\in\mathcal{D}_k:v|\Delta}\frac{1+\sum_{p|v}\log p}{\varphi_\omega(v)} &\ll \biggl(1+\sum_{p>2k^2:p|\Delta}\frac{\log p}{p}\biggr)\prod_{p>2k^2:p|\Delta}\biggl(1+\frac{1}{\varphi_\omega(p)}\biggr) \\
&\ll (\log\log\Delta)^2 \ll (\log\log R)^2,
\end{aligned}
\tag{9.43}
$$

since both sum and product are largest if Δ is composed of primes $\ll\log\Delta$, and $\Delta\ll x^{O(k)}$.

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

Thus, relaxing the constraint $(e_m, rW_m)=1$ to $(e_m, a_mWBr)=1$, and using Lemma 8.4 to estimate the sum over e_m, we see the error contributes a total

$$\begin{aligned}&\ll \frac{T_k(\log\log R)^2}{\log R}\frac{r}{\varphi_L(r)}\sum_{(e_m,a_mWBr)=1}\frac{Y_{\mathbf{r}'}}{\varphi_\omega(e_m)}\\&\ll \frac{T_k(\log\log R)^2\varphi(a_mWB)W^{k-1}B^{k-1}\mathfrak{S}_{WB}(\mathcal{L})}{a_m\varphi(WB)^k}\int_0^\infty F_2(t_1,\dots,t_k)\,dt_m. \qquad (9.44)\end{aligned}$$

Here, as in the statement of the lemma, $t_i=\log r_i/\log R$ for $i\neq m$.

We now return to the main term from (9.42). We first consider the inner sum, which by multiplicativity we can rewrite as a product

$$\sideset{}{^*}\sum_{\mathbf{u},\mathbf{v}}\frac{\mu(u)}{\varphi(u)\varphi_\omega(uv)}=\prod_p \sideset{}{^*}\sum_{\substack{\mathbf{u},\mathbf{v}\\ u_i|p,\, v_i|p\ \forall i}}\frac{\mu(u)}{\varphi(u)\varphi_\omega(uv)}, \qquad (9.45)$$

where the asterisk indicates that sums are subject to the additional constraints that $u_m=v_m=1$, $(u,v)=1$ and that $(u_i,W_i're_ma_m)=1$, $(v_i,W_ire_ma_m)=1$, and $v_i|W_i'/W_i$ for all $1\leqslant i\leqslant k$. Since the summand only depends on u and v we can evaluate it by counting how many pairs $\mathbf{u},\mathbf{v}$ correspond to a given choice of u,v.

If $p|WBre_ma_m$ then our coprimality restrictions mean no component of $\mathbf{u}$ or $\mathbf{v}$ can be a multiple of p. If $p\nmid WBre_ma_m$ then there are $\omega(p)-1$ components of $\mathbf{u}$ which can be a multiple of p (corresponding to the indices for all residue classes chosen mod p except for the index corresponding to $-b_m\overline{a_m}$). If $p\nmid W_m$ then no components of $\mathbf{v}$ can be a multiple of p ($p\nmid W_m$ means m was the chosen index for the residue class $-b_m\overline{a}_m \pmod p$, so $p\nmid W_j'/W_j$ for any j). If $p|W_m$ and $p\nmid WBre_ma_m$ then exactly one component of $\mathbf{v}$ can be a multiple of p (v_j can be a multiple of p if j was the chosen index for the residue class $-b_m\overline{a}_m \pmod p$). Finally, since $(u,v)=1$, no component of $\mathbf{u}$ can be a multiple of p if v is a multiple of p. Putting this together, we obtain (since $(e_m,rW_m)=1$)

$$\begin{aligned}\sideset{}{^*}\sum_{\mathbf{u},\mathbf{v}}\frac{\mu(u)}{\varphi(u)\varphi_\omega(uv)}&=\prod_{p|W_m,p\nmid WBra_m}\left(1-\frac{\omega(p)-1}{\varphi(p)\varphi_\omega(p)}+\frac{1}{\varphi_\omega(p)}\right)\prod_{p\nmid W_mre_m}\left(1-\frac{\omega(p)-1}{\varphi(p)\varphi_\omega(p)}\right)\\&=\prod_{p|W_m,p\nmid WBra_m}\frac{p}{p-1}\prod_{p\nmid W_mr}\left(\frac{p}{p-1}-\frac{1}{\varphi_\omega(p)}\right)\prod_{p|e_m}\left(\frac{p}{p-1}-\frac{1}{\varphi_\omega(p)}\right)^{-1}.\end{aligned} \qquad (9.46)$$

Now, using Lemma 8.3, we estimate the summation over e_m. This gives

$$\begin{aligned}&\sum_{(e_m,rW_m)=1}\frac{y_{\mathbf{r}'}}{\varphi_\omega(e_m)}\prod_{p|e_m}\left(\frac{p}{p-1}-\frac{1}{\varphi_\omega(p)}\right)^{-1}\\&=\log R\frac{\mathfrak{S}_{WB}(\mathcal{L})W^kB^k}{\varphi(WB)^k}\prod_{p|rW_m}\left(1-\frac{1}{p}\right)\prod_{p\nmid rW_m}\left(\frac{p}{p-1}-\frac{1}{\varphi_\omega(p)}\right)^{-1}\int_0^\infty H(t_1,\dots,t_k)\,dt_m,\end{aligned} \qquad (9.47)$$

where, as in the statement of the lemma, $t_i=\log r_i/\log R$ for $i\neq m$, and where

$$H(t_1,\dots,t_k)=F(t_1,\dots,t_k)+O\biggl(\frac{T_k(\log\log R)^2}{\log R}F_2(t_1,\dots,t_k)\biggr).$$

We have added an additional factor of $\log\log R$ into the error term for H so we can absorb (9.44) into the error term.

Thus, combining (9.46) and (9.47) gives

$$\begin{aligned}
&\frac{r}{\varphi_L(r)}\sum_{e_m}\frac{y_{\mathbf{r}'}}{\varphi_\omega(e_m)}\sum_{\mathbf{u},\mathbf{v}}^{*}\frac{\mu(u)}{\varphi(u)\varphi_\omega(uv)}\\
&\quad=\log R\frac{W^kB^k\mathfrak{S}_{WB}(\mathcal{L})}{\varphi(WB)^k}\frac{r}{\varphi_L(r)}\prod_{p|r}\Big(1-\frac{1}{p}\Big)\prod_{\substack{p|WBa_m\\ p\nmid r}}\Big(1-\frac{1}{p}\Big)\int_0^\infty H(t_1,\ldots,t_k)\,dt_m\\
&\quad=\log R\frac{\varphi(a_mWB)W^{k-1}B^{k-1}\mathfrak{S}_{WB}(\mathcal{L})}{a_m\varphi(WB)^k}\int_0^\infty H(t_1,\ldots,t_k)\,dt_m. \qquad (9.48)
\end{aligned}$$

Here we have used the fact that $\mathbf{r}\in\mathcal{D}_k'$, and so $(r,WB)=1$. Combining (9.44) and (9.48) gives the result. □

PROPOSITION 9.4. *Let w_n be as described in § 7. Given $D\ll x^{O(1)}$ and ξ satisfying $k(\log\log x)^2/(\log x)\ll\xi\leqslant\theta/10$, let*

$$\mathcal{S}(\xi;D)=\{n\in\mathbb{N}:p|n\implies(p>x^\xi\text{ or }p|D)\}.$$

For $L=a_0n+b_0\notin\mathcal{L}$, with $|a_0|,|b_0|\ll x^{O(1)}$ and $\Delta_L\neq 0$, we have

$$\sum_{n\in\mathcal{A}(x)}\mathbf{1}_{\mathcal{S}(\xi;D)}(L(n))w_n\ll\xi^{-1}\frac{\Delta_L}{\varphi(\Delta_L)}\frac{D}{\varphi(D)}\frac{B^k}{\varphi(B)^k}\mathfrak{S}_B(\mathcal{L})\#\mathcal{A}(x)(\log R)^{k-1}I_k(F),$$

where

$$\Delta_L=a_0\prod_{i=1}^k|a_jb_0-a_0b_j|.$$

Proof. We first split the sum into residue classes v_0 modulo $V=\prod_{p\leqslant 2k^2}p$ for which $L(v_0)$ is coprime to $\prod_{p\leqslant 2k^2,p\nmid D}p$ and each of the $L_i(v_0)$ are coprime to W (the other residue classes make no contribution because of the support of w_n and $\mathbf{1}_{\mathcal{S}(\xi;D)}$). We use the Selberg sieve upper bound

$$\mathbf{1}_{\mathcal{S}(\xi;D)}(L(n))\leqslant\tilde{\lambda}_1^{-2}\Bigg(\sum_{\substack{d_0|L(n)\\ d_0<x^\xi\\ (d_0,D)=1}}\tilde{\lambda}_{d_0}\Bigg)^2. \qquad (9.49)$$

(This holds for any choice of the values of $\tilde{\lambda}_d\in\mathbb{R}$ with $\tilde{\lambda}_1\neq 0$). For the residue class $v_0\pmod V$, this gives

$$\sum_{\substack{n\in\mathcal{A}(x)\\ n\equiv v_0\ (\mathrm{mod}\ V)}}\mathbf{1}_{\mathcal{S}(\xi;D)}(L(n))w_n\leqslant\frac{1}{\tilde{\lambda}_1^2}\sum_{\substack{n\in\mathcal{A}(x)\\ n\equiv v_0\ (\mathrm{mod}\ V)}}\Bigg(\sum_{\substack{d_0|L(n)\\ (d_0,D)=1,\,d_0<x^\xi}}\tilde{\lambda}_{d_0}\Bigg)^2\Bigg(\sum_{\substack{\mathbf{d}\in\mathcal{D}_k\\ d_i|L_i(n)\ \forall 1\leqslant i\leqslant k}}\lambda_{\mathbf{d}}\Bigg)^2. \qquad (9.50)$$

We restrict the support of $\tilde{\lambda}_{d_0}$ in a similar way to that of $\lambda_{\mathbf{d}}$. We force $\tilde{\lambda}_{d_0}=0$ if $p|d_0$ for any prime with $p|W_0$ where

$$W_0=DV\Delta_L. \qquad (9.51)$$

Similarly, we force $\tilde{\lambda}_{d_0}=0$ if $d_0>x^\xi$. Note that we allow $\tilde{\lambda}_{d_0}\neq 0$ if $(d_0,B)\neq 1$. These conditions mean we can drop the constraints $d_0<x^\xi$, $(d_0,D)=1$ since $\tilde{\lambda}_{d_0}=0$ if either of these do not hold.

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

We return to (9.50). Expanding the squares and swapping the order of summation gives

$$\begin{aligned}\frac{1}{\tilde{\lambda}_1^2}&\sum_{\substack{n\in\mathcal{A}(x)\\ n\equiv v_0 \pmod V}}\Bigg(\sum_{\substack{d_0|L(n)\\ (d_0,W_0)=1}}\tilde{\lambda}_{d_0}\Bigg)^2\Bigg(\sum_{\substack{\mathbf{d}\in\mathcal{D}_k\\ d_i|L_i(n)\ \forall 1\leqslant i\leqslant k}}\lambda_{\mathbf{d}}\Bigg)^2\\ &=\frac{1}{\tilde{\lambda}_1^2}\sum_{\substack{d_0,e_0\\ (d_0e_0,W_0)=1}}\tilde{\lambda}_{d_0}\tilde{\lambda}_{e_0}\sum_{\mathbf{d},\mathbf{e}\in\mathcal{D}_k}\lambda_{\mathbf{d}}\lambda_{\mathbf{e}}\sum_{\substack{n\in\mathcal{A}(x)\\ n\equiv v_0 \pmod V\\ [d_i,e_i]|L_i(n)\ \forall 0\leqslant i\leqslant k}}1. \qquad (9.52)\end{aligned}$$

We see that by our restrictions on the support of $\lambda_{\mathbf{d}},\tilde{\lambda}_{d_0}$, there is no contribution to (9.52) unless $d_0,e_0,\mathbf{d},\mathbf{e}$ are such that $(d_ie_i,d_je_j)=1$ for all $0\leqslant i\neq j\leqslant k$, and $d,e<R$ and $d_0,e_0<x^{\xi}$. (To avoid confusion, we recall that $d=\prod_{i=1}^k d_i$ and $e=\prod_{i=1}^k e_i$.) For such values, we can combine the congruence conditions using the Chinese remainder theorem, which shows the inner sum is $\#\mathcal{A}(x;q,a)$ for some a and $q=V\prod_{i=0}^k[d_i,e_i]$. We see that $q<VR^2x^{2\xi}<x^{\theta}$ since $\xi\leqslant\theta/10$. We substitute $\#\mathcal{A}(x;q,a)=\#\mathcal{A}(x)/q+O(E_q^{(1)})$, and the contribution from $E_q^{(1)}$ can be seen to be negligible by an identical argument to that in the proof of Proposition 9.1. We are therefore left to evaluate

$$\frac{\#\mathcal{A}(x)}{V\tilde{\lambda}_1^2}\sum_{\substack{d_0,e_0\\ (d_0,e_0,W_0)=1}}\frac{\tilde{\lambda}_{d_0}\tilde{\lambda}_{e_0}}{[d_0,e_0]}\sideset{}{'}\sum_{\substack{\mathbf{d},\mathbf{e}\in\mathcal{D}_k\\ (de,d_0e_0)=1}}\frac{\lambda_{\mathbf{d}}\lambda_{\mathbf{e}}}{[\mathbf{d},\mathbf{e}]}. \qquad (9.53)$$

We let ω^* be the totally multiplicative function defined by

$$\omega^*(p)=\begin{cases}\#\Big\{1\leqslant n\leqslant p: L(n)\prod_{i=1}^k L_i(n)\equiv 0 \pmod p\Big\}, & p\nmid B,\\ 1, & p|B.\end{cases} \qquad (9.54)$$

We note that with this choice, $\omega^*(p)=\omega(p)$ if $p|\Delta_L$ and $p\nmid B$, and $\omega^*(p)=\omega(p)+1$ otherwise. We also define

$$y_{\mathbf{r},r_0}=\mu(r_0r)\varphi_{w*}(r_0r)\sum_{\substack{\mathbf{d},d_0\\ \mathbf{r}|\mathbf{d},\,r_0|d_0\\ (d_0,d)=1}}\frac{\lambda_{\mathbf{d}}\tilde{\lambda}_{d_0}}{dd_0},\quad \tilde{y}_{r_0}=\mu(r_0)\varphi(r_0)\sum_{\substack{d_0\\ r_0|d_0}}\frac{\tilde{\lambda}_{d_0}}{d_0}. \qquad (9.55)$$

By Möbius inversion, we see that this definition of $\tilde{y}_{r_0}$ implies that

$$\tilde{\lambda}_{d_0}=\mu(d_0)d_0\sum_{d_0|r_0}\frac{\tilde{y}_{r_0}}{\varphi(r_0)}. \qquad (9.56)$$

For $(r_0,W_0)=1$ and $r_0<x^{\xi}$ we choose

$$\tilde{y}_{r_0}=\frac{W_0}{\varphi(W_0)}, \qquad (9.57)$$

and $\tilde{y}_{r_0}=0$ otherwise. This gives rise to a suitable choice of $\tilde{\lambda}_{d_0}$ supported on $d_0<x^{\xi}$ with $(d_0,W_0)=1$. Since $\xi\gg k(\log\log x)^2/(\log x)$, Lemma 8.3 shows that

$$\tilde{\lambda}_1=\sum_{\substack{r_0<x^{\xi}\\ (r_0,W_0)=1}}\frac{\tilde{y}_{r_0}\mu^2(r_0)}{\varphi(r_0)}=\xi\log x+O(\log\log x)\gg\xi\log x. \qquad (9.58)$$

As in the proof of Proposition 9.1 (this is exactly the same argument but for $(k+1)$-dimensional vectors instead of k-dimensional ones) changing variables using (9.55) shows that

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

$$\sum_{d_0,e_0} \frac{\tilde\lambda_{d_0}\tilde\lambda_{e_0}}{[d_0,e_0]} \sideset{}{'}\sum_{\substack{\mathbf{d},\mathbf{e}\in\mathcal{D}_k\\(de,d_0e_0)=1}} \frac{\lambda_{\mathbf{d}}\lambda_{\mathbf{e}}}{[\mathbf{d},\mathbf{e}]} = \sum_{\mathbf{r},\mathbf{s},r_0,s_0} \frac{y_{\mathbf{r},r_0}y_{\mathbf{s},s_0}}{\varphi_{\omega*}(rr_0)\varphi_{\omega*}(ss_0)} \prod_{p|rr_0ss_0} S_p(\mathbf{r},\mathbf{s},r_0,s_0), \tag{9.59}$$

where

$$S_p(\mathbf{r},\mathbf{s},r_0,s_0) = \begin{cases} p-1, & p|(\mathbf{r},\mathbf{s})(r_0,s_0),\\ -1, & p|rr_0 \text{ and } p|ss_0 \text{ but } p\nmid(\mathbf{r},\mathbf{s})(r_0,s_0),\\ 0, & (p|rr_0 \text{ and } p\nmid ss_0) \text{ or } (p|ss_0 \text{ and } p\nmid rr_0).\end{cases} \tag{9.60}$$

Thus we may restrict to $rr_0 = ss_0$. Using the bound $y_{\mathbf{r},r_0}y_{\mathbf{s},s_0} \ll y_{\mathbf{r},r_0}^2 + y_{\mathbf{s},s_0}^2$, we see that (by symmetry) the right-hand side of (9.59) may be bounded by

$$\begin{aligned}\sum_{\mathbf{r},r_0} \frac{y_{\mathbf{r},r_0}^2}{\varphi_{\omega*}(rr_0)^2} \sum_{\substack{\mathbf{s},s_0\\ ss_0=rr_0}} \prod_{p|rr_0} |S_p(\mathbf{r},\mathbf{s},r_0,s_0)| &\leqslant \sum_{\mathbf{r},r_0} y_{\mathbf{r},r_0}^2 \prod_{p|rr_0} \frac{p+\omega^*(p)-2}{(p-\omega^*(p))^2}\\ &= \sum_{\mathbf{r},r_0} \frac{y_{\mathbf{r},r_0}^2}{\prod_{p|rr_0}(p+O(k))}.\end{aligned} \tag{9.61}$$

To evaluate this, we express $y_{\mathbf{r},r_0}$ in terms of $y_{\mathbf{r}}$ and $\tilde y_{r_0}$. Substituting (7.8) into (9.55), we find that for $(r_0, rW_0) = 1$ and $\mathbf{r} \in \mathcal{D}_k$,

$$\begin{aligned} y_{\mathbf{r},r_0} &= \mu(r_0r)\varphi_{\omega*}(r_0r) \sum_{\substack{d_0\\ r_0|d_0}} \mu(d_0) \sum_{\substack{f_0\\ d_0|f_0}} \frac{\tilde y_{f_0}}{\varphi(f_0)} \sum_{\substack{\mathbf{d}\\ \mathbf{r}|\mathbf{d}\\ (d,d_0)=1}} \mu(d) \sum_{\substack{\mathbf{f}\\ \mathbf{d}|\mathbf{f}}} \frac{y_{\mathbf{f}}}{\varphi_\omega(f)}\\ &= \mu(r_0r)\varphi_{\omega*}(r_0r) \sum_{\substack{f_0,\mathbf{f}\\ r_0|f_0,\,\mathbf{r}|\mathbf{f}}} \frac{y_{\mathbf{f}}\tilde y_{f_0}}{\varphi_\omega(f)\varphi(f_0)} \sum_{\substack{d_0,\mathbf{d}\\ r_0|d_0,\,\mathbf{r}|\mathbf{d}\\ d_0|f_0,\,\mathbf{d}|\mathbf{f}\\ (d,d_0)=1}} \mu(d)\mu(d_0). \end{aligned} \tag{9.62}$$

The inner sum is 0 unless every prime dividing one of f, f_0 but not the other is a divisor of rr_0. In this case the sum is ±1. Thus, using the fact that $y_{\mathbf{r}} \geqslant y_{\mathbf{f}}$ and $\tilde y_{r_0} \geqslant \tilde y_{f_0}$ (since F is decreasing), we have the crude bound

$$y_{\mathbf{r},r_0} \leqslant \varphi_{\omega*}(r_0r) y_{\mathbf{r}} \tilde y_{r_0} \sum_{\substack{f_0\\ r_0|f_0\\ (f_0,W_0)=1}} \sum_{\substack{\mathbf{f}\in\mathcal{D}_k\\ \mathbf{r}|\mathbf{f}\\ f/(f,f_0)|rr_0}} \frac{\mu^2(f_0)}{\varphi(f_0)\varphi_\omega(f)}. \tag{9.63}$$

We let $f_0 = r_0 f_0' g_0$ and $f_i = r_i f_i' g_i$ for $1 \leqslant i \leqslant k$, where $f_i' = f_i/(f_i, rr_0)$ is f_i with any factors of rr_0 removed, $g_0|r$ and $g_i|r_0$ for $1 \leqslant i \leqslant k$. We see the constraint $f/(f,f_0)|rr_0$ means that $f_0' = \prod_{i=1}^k f_i'$. Therefore we can bound the double sum above by

$$\begin{aligned} &\frac{1}{\varphi(r_0)\varphi_\omega(r)} \sum_{\mathbf{f}'\in\mathcal{D}_k} \frac{1}{\varphi(f')\varphi_\omega(f')} \sum_{\substack{\mathbf{g}\in\mathcal{D}_k\\ g_i|r_0\ \forall 1\leqslant i\leqslant k}} \frac{1}{\varphi_\omega(g)} \sum_{\substack{g_0|r\\ (g_0,W_0)=1}} \frac{1}{\varphi(g_0)}\\ &= \frac{1}{\varphi(r_0)\varphi_\omega(r)} \prod_{p\nmid WB}\left(1+\frac{\omega(p)}{(p-1)(p-\omega(p))}\right) \prod_{p|r_0}\left(1+\frac{\omega(p)}{p-\omega(p)}\right) \prod_{p|r,p\nmid W_0}\left(1+\frac{1}{p-1}\right). \end{aligned} \tag{9.64}$$

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

The first product is $O(1)$ since it is over primes $p > 2k^2$. Thus, simplifying the remaining products, we obtain

$$y_{\mathbf{r},r_0} \ll y_{\mathbf{r}}\tilde{y}_{r_0}\bigg(\prod_{p|(r,W_0)}\frac{p-\omega^*(p)}{p-\omega(p)}\bigg)\bigg(\prod_{p|rr_0,\, p\nmid W_0}\frac{p(p-\omega^*(p))}{(p-1)(p-\omega(p))}\bigg) \leqslant y_{\mathbf{r}}\tilde{y}_{r_0}. \tag{9.65}$$

Here we have used the fact that $\omega^*(p) = \omega(p) + 1$ if $p \nmid W_0$.

Recalling the definitions (9.57) and (7.8) of $\tilde{y}_{r_0}$ and $y_{\mathbf{r}}$, and applying Lemma 8.4, we find that (since $\xi \gg k(\log\log x)^2/(\log x)$)

$$\begin{aligned}\sum_{\mathbf{r},r_0}\frac{(y_{\mathbf{r},r_0})^2}{\prod_{p|r_0r}(p+O(k))} &\ll \bigg(\sum_{\substack{r_0<x^\xi\\(r_0,W_0)=1}}\frac{\tilde{y}_{r_0}^2}{\prod_{p|r_0}(p+O(k))}\bigg)\bigg(\sum_{\mathbf{r}\in\mathcal{D}_k}\frac{y_{\mathbf{r}}^2}{\prod_{p|r}(p+O(k))}\bigg)\\ &\ll \xi(\log R)^{k+1}\frac{W^kB^kW_0\mathfrak{S}_{WB}(\mathcal{L})^2}{\varphi(WB)^k\varphi(W_0)}\\ &\quad\times\prod_{p\nmid W_0}\bigg(1+\frac{O(k)}{p^2}\bigg)\prod_{p\nmid WB}\bigg(1+\frac{\omega(p)}{p+O(k)}\bigg)\bigg(1-\frac{1}{p}\bigg)^k I_k(F). \end{aligned}\tag{9.66}$$

We note that the first product is $O(1)$ and the second product is $O(\mathfrak{S}_{WB}(\mathcal{L})^{-1})$, since all primes in the products are greater than $2k^2$ and $\omega(p) \leqslant k$. Thus, we obtain (recalling $\tilde{\lambda}_1 \gg \xi\log x$)

$$\frac{\#\mathcal{A}(x)}{V\tilde{\lambda}_1^2}\sum_{\mathbf{r},r_0}\frac{(y_{\mathbf{r},r_0})^2}{\prod_{p|rr_0}(p+O(k))} \ll \xi^{-1}(\log R)^{k-1}\#\mathcal{A}(x)\frac{W_0W^kB^k\mathfrak{S}_{WB}(\mathcal{L})}{V\varphi(W_0)\varphi(WB)^k}I_k(F). \tag{9.67}$$

We now sum over residue classes $v_0 \bmod V$, for which $L(v_0)$ is coprime to $\prod_{p\leqslant 2k^2, p\nmid D}p$ and each of the $L_i(v_0)$ are coprime to W. The number N of such residue classes is given by

$$N = \prod_{\substack{p|W\\p\nmid D\Delta_L}}(p-\omega(p)-1)\prod_{\substack{p|W\\p|D\Delta_L}}(p-\omega(p))\prod_{\substack{p|V/W\\p\nmid Da_0}}(p-1)\prod_{\substack{p|V/W\\p|Da_0}}p. \tag{9.68}$$

This then gives

$$\sum_{n\in\mathcal{A}(x)}\mathbf{1}_{\mathcal{S}(\xi;D)}(L(n))w_n \ll \xi^{-1}\frac{B^k}{\varphi(B)^k}(\log R)^{k-1}\#\mathcal{A}(x)\mathfrak{S}_{WB}(\mathcal{L})I_k(F)\frac{NW_0W^k}{V\varphi(W_0)\varphi(W)^k}. \tag{9.69}$$

Finally, by calculation we find that

$$\begin{aligned}\frac{NW_0W^k}{V\varphi(W_0)\varphi(W)^k} &= \frac{\mathfrak{S}_B(\mathcal{L})\Delta_LD}{\mathfrak{S}_{WB}(\mathcal{L})\varphi(\Delta_LD)}\prod_{\substack{p|(\Delta_L,V)\\p\nmid a_0WD}}\frac{p-1}{p}\prod_{\substack{p|W\\p\nmid\Delta_LD}}\frac{(p-\omega(p)-1)p}{(p-\omega(p))(p-1)}\\ &\leqslant \frac{\mathfrak{S}_B(\mathcal{L})\Delta_LD}{\mathfrak{S}_{WB}(\mathcal{L})\varphi(\Delta_L)\varphi(D)}.\end{aligned}\tag{9.70}$$

This gives the result. $\square$

PROPOSITION 9.5. *Let w_n be as described in § 7. For $L \in \mathcal{L}$ and $\rho \leqslant \theta/10$, we have*

$$\sum_{n\in\mathcal{A}(x)}\bigg(\sum_{\substack{p|L(n)\\p<x^\rho\\p\nmid B}}1\bigg)w_n \ll \rho^2k^4(\log k)^2\mathfrak{S}_B(\mathcal{L})\#\mathcal{A}(x)(\log R)^kI_k(F).$$

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

Proof. We let $L(n)=L_m(n)=a_mn+b_m$ be the mth function in $\mathcal{L}$. As with Propositions 9.1 and 9.2, we consider the sum restricted to $n\equiv v_0 \pmod W$ for some v_0 with $(\prod_{i=1}^k L_i(v_0),W)=1$, since the other choices of v_0 make no contribution. This means we can also restrict the sum to $p\nmid W$.

Expanding the square and swapping the order of summation gives

$$\sum_{\substack{n\in\mathcal{A}(x)\\ n\equiv v_0 \pmod W}}\Bigg(\sum_{\substack{p|L(n)\\ p<x^\rho\\ p\nmid WB}}1\Bigg)w_n=\sum_{\substack{p<x^\rho\\ p\nmid WB}}\sum_{\mathbf{d},\mathbf{e}\in\mathcal{D}_k}\lambda_{\mathbf{d}}\lambda_{\mathbf{e}}\sum_{\substack{n\in\mathcal{A}(x)\\ [d_i,e_i]|L_i(n)\\ n\equiv v_0 \pmod W\\ p|L_m(n)}}1. \tag{9.71}$$

The inner sum is empty unless $(d_ie_i,d_je_j)=1$ for all $i\ne j$ and $(d_ie_i,p)=1$ for all $i\ne m$. In this case, by the Chinese remainder theorem, we can combine the congruence conditions and see that the inner sum is $\#\mathcal{A}(x;q,a)$ for $q=[d_m,e_m,p]\prod_{i\ne m}[d_i,e_i]$ and some a. We write $\#\mathcal{A}(x;q,a)=\#\mathcal{A}(x)/q+O(E_q^{(1)})$ as in the proof of Proposition 9.1. We treat the error $E_q^{(1)}$ from making this change in the same manner as in Proposition 9.1, noting that all moduli q we need to consider are square-free and satisfy $q<WR^2x^\rho<x^\theta$, and for any q there are $O(\tau_{3k+4}(q))$ choices of $\mathbf{d},\mathbf{e},p$ which give rise to the modulus q. Thus these error terms make a negligible contribution.

We use (7.8) to change to our $y_{\mathbf{r}}$ variables, which gives us a main term of

$$\frac{\#\mathcal{A}(x)}{W}\sum_{\substack{p\leqslant x^\rho\\ p\nmid WB}}\ \sideset{}{'}\sum_{\substack{\mathbf{d},\mathbf{e}\in\mathcal{D}_k\\ (d_ie_i,p)=1}}\frac{\lambda_{\mathbf{d}}\lambda_{\mathbf{e}}}{[d_m,e_m,p]\prod_{i\ne m}[d_i,e_i]}=\frac{\#\mathcal{A}(x)}{W}\sum_{\substack{p<x^\rho\\ p\nmid WB}}\frac{1}{p}\sum_{\mathbf{r},\mathbf{s}\in\mathcal{D}_k}\frac{y_{\mathbf{r}}y_{\mathbf{s}}}{\varphi_\omega(r)\varphi_\omega(s)}\prod_{p'|rs}S_{p'}(\mathbf{r},\mathbf{s},p). \tag{9.72}$$

Here if $p'\ne p$ then $S_{p'}(\mathbf{r},\mathbf{s},p)=S_{p'}(\mathbf{r},\mathbf{s})$, given by (9.5), whereas if $p'=p$ then

$$S_p(\mathbf{r},\mathbf{s},p)=\sideset{}{'}\sum_{\substack{\mathbf{d}|\mathbf{r},\,\mathbf{e}|\mathbf{s}\\ d_i,e_i|p\ \forall i\\ (d_ie_i,p)=1\ \forall i\ne m}}\frac{p\mu(d)\mu(e)de}{[d_m,e_m,p]\prod_{i\ne m}[d_i,e_i]}=\begin{cases}(p-1)^2, & p|(r_m,s_m),\\ -(p-1), & p|r_ms_m, p\nmid(r_m,s_m),\\ 1, & p\nmid r_ms_m.\end{cases} \tag{9.73}$$

We let $\mathbf{u}=(r_1/(r_1,p),\dots,r_k/(r_k,p))$ be the vector formed by removing a possible factor of p from the components of $\mathbf{r}$. We note that for any $\mathbf{s}\in\mathcal{D}_k$ and $p\nmid W$ we have

$$\begin{aligned}\sum_{\substack{\mathbf{r}\in\mathcal{D}_k\\ r_i/(r_i,p)=u_i\ \forall i\\ (r_i,W_i)=1}}\frac{\prod_{p'|rs}S_{p'}(\mathbf{r},\mathbf{s},p)}{\varphi_\omega(r)}&=\Bigg(\prod_{p'|su}S_{p'}(\mathbf{s},\mathbf{u})\Bigg)\sum_{\substack{\mathbf{r}\in\mathcal{D}_k\\ r_i/(r_i,p)=u_i\ \forall i\\ (r_i,W_i)=1}}\frac{S_p(\mathbf{r},\mathbf{s},p)}{\varphi_\omega(r)}\\ &=\frac{\mu((s_m,p))\varphi((s_m,p))}{\varphi_\omega(u)}\bigg(1+\frac{\omega(p)-1}{p-\omega(p)}-\frac{p-1}{p-\omega(p)}\bigg)=0.\end{aligned} \tag{9.74}$$

Here the first term in parentheses represents the contribution when $(r,p)=1$, the second term represents the contribution when $p|r$ but $p\nmid r_m$ (and so there are $\omega(p)-1$ choices of which index can be a multiple of p) and the final term represents the contribution when $p|r_m$.

We substitute $y_{\mathbf{r}}=y_{\mathbf{u}}+(y_{\mathbf{r}}-y_{\mathbf{u}})$ into our main term. By (9.74) we find the $y_{\mathbf{u}}$ term makes a total contribution of 0, leaving only the contribution from $(y_{\mathbf{r}}-y_{\mathbf{u}})$. Similarly, we let $\mathbf{v}$ be the

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

vector obtained by removing a possible factor of p from $\mathbf{s}$. We make the equivalent substitution $y_{\mathbf{s}} = y_{\mathbf{v}} + (y_{\mathbf{s}} - y_{\mathbf{v}})$, with the $y_{\mathbf{v}}$ term making no contribution. By Lemma 8.2 we have

$$(y_{\mathbf{r}} - y_{\mathbf{u}})(y_{\mathbf{s}} - y_{\mathbf{v}}) \ll Y_{\mathbf{u}} Y_{\mathbf{v}} T_k^2 (\log p)^2/(\log R)^2. \tag{9.75}$$

Substituting this bound into our main term (9.72), we obtain the bound

$$\ll \frac{T_k^2 \#\mathcal{A}(x)}{W} \sum_{\substack{p<x^\rho \\ p\nmid WB}} \frac{1}{p}\left(\frac{\log p}{\log R}\right)^2 \sum_{\substack{\mathbf{u},\mathbf{v}\in\mathcal{D}_k \\ (u,p)=(v,p)=1}} Y_{\mathbf{u}} Y_{\mathbf{v}} \prod_{p'|uv} |S_{p'}(\mathbf{u},\mathbf{v})| \sum_{\substack{\mathbf{r},\mathbf{s}\in\mathcal{D}_k \\ r_i/(r_i,p)=u_i \ \forall i \\ s_i/(s_i,p)=v_i \ \forall i}} \frac{|S_p(\mathbf{r},\mathbf{s},p)|}{\varphi_\omega(r)\varphi_\omega(s)}. \tag{9.76}$$

A calculation reveals that the inner sum is $O(\varphi_\omega(u)^{-1}\varphi_\omega(v)^{-1})$ for all $p\nmid WB$. This gives the bound

$$\begin{aligned}
&\ll \frac{T_k^2 \#\mathcal{A}(x)}{W} \sum_{\substack{p<x^\rho \\ p\nmid WB}} \frac{1}{p}\left(\frac{\log p}{\log R}\right)^2 \sum_{\substack{\mathbf{u},\mathbf{v}\in\mathcal{D}_k \\ (u,p)=(v,p)=1}} \frac{Y_{\mathbf{u}} Y_{\mathbf{v}}}{\varphi_\omega(u)\varphi_\omega(v)} \prod_{p'|uv} |S_{p'}(\mathbf{u},\mathbf{v})| \\
&\ll \frac{T_k^2 \rho^2 \#\mathcal{A}(x)}{W} \sum_{\mathbf{u},\mathbf{v}\in\mathcal{D}_k} \frac{Y_{\mathbf{u}}^2 + Y_{\mathbf{v}}^2}{\varphi_\omega(u)\varphi_\omega(v)} \prod_{p'|uv} |S_{p'}(\mathbf{u},\mathbf{v})|.
\end{aligned} \tag{9.77}$$

Here we have dropped the requirement that $(u,p)=(v,p)=1$ and used $Y_{\mathbf{u}} Y_{\mathbf{v}} \leqslant Y_{\mathbf{u}}^2 + Y_{\mathbf{v}}^2$ for an upper bound.

We recall from (9.5) that $S_{p'}(\mathbf{u},\mathbf{v}) = 0$ unless $u = v$. By multiplicativity and from the definition (9.5) of $S_{p'}(\mathbf{u},\mathbf{v})$, we find that, given $\mathbf{u}\in\mathcal{D}_k$,

$$\sum_{\mathbf{v}\in\mathcal{D}_k} \frac{\prod_{p'|uv} |S_{p'}(\mathbf{u},\mathbf{v})|}{\varphi_\omega(v)} = \prod_{p'|u}\Bigg(\sum_{\substack{\mathbf{w}\in\mathcal{D}_k \\ w_i|p' \ \forall i}} \frac{|S_{p'}(\mathbf{u},\mathbf{w})|}{\varphi_\omega(w)}\Bigg) = \prod_{p'|u}\left(\frac{p-1}{p-\omega(p)} + \frac{\omega(p)-1}{p-\omega(p)}\right). \tag{9.78}$$

(Here the first term in parentheses in the final product corresponds to the $\mathbf{w}$ such that $p|(\mathbf{u},\mathbf{w})$ and the second term to the $\omega(p)-1$ choices of $\mathbf{w}$ such that $p\nmid(\mathbf{u},\mathbf{w})$.) Thus, we find

$$\sum_{\mathbf{u},\mathbf{v}\in\mathcal{D}_k} \frac{Y_{\mathbf{u}}^2 + Y_{\mathbf{v}}^2}{\varphi_\omega(u)\varphi_\omega(v)} \prod_{p'|uv} |S_{p'}(\mathbf{u},\mathbf{v})| \ll \sum_{\mathbf{r}\in\mathcal{D}_k} \frac{Y_{\mathbf{r}}^2}{g(r)}, \tag{9.79}$$

where g is the multiplicative function defined by $g(p) = (p-\omega(p))^2/(p+\omega(p)-2)$. Applying Lemma 8.4, we see that this is

$$\ll \frac{B^k W^k \mathfrak{S}_{WB}(\mathcal{L})^2}{\varphi(WB)^k} (\log R)^k \prod_{p\nmid WB}\left(1+\frac{\omega(p)}{g(p)}\right)\left(1-\frac{1}{p}\right)^k I_k(F_2). \tag{9.80}$$

By Lemma 8.6 we have $I_k(F_2) \ll k^2 I_k(F)$. Since any prime $p\nmid WB$ has $p > 2k^2$ and $g(p) = p + O(k)$, we see the product is $\ll \mathfrak{S}_{WB}(\mathcal{L})^{-1}$. Thus (9.80) is

$$\ll k^2 \frac{B^k W^k \mathfrak{S}_{WB}(\mathcal{L})}{\varphi(WB)^k} (\log R)^k I_k(F). \tag{9.81}$$

Putting this all together gives

$$\sum_{\substack{n\in\mathcal{A}(x) \\ n\equiv v_0 \pmod{W}}} \Bigg(\sum_{\substack{p|L(n) \\ p<x^\rho \\ p\nmid B}} 1\Bigg) w_n \ll k^2 T_k^2 \rho^2 \#\mathcal{A}(x) \frac{B^k W^{k-1} \mathfrak{S}_{WB}(\mathcal{L})}{\varphi(WB)^k} (\log R)^k I_k(F). \tag{9.82}$$

Summing over the $\varphi_\omega(W)$ residue classes mod W then gives the result. $\square$

https://doi.org/10.1112/S0010437X16007296 Published online by Cambridge University Press

Acknowledgements

I would like to thank Sary Drappeau, Andrew Granville, Dimitris Koukoulopoulos and Jesse Thorner for many useful comments and suggestions. This research was carried out while the author was funded by a CRM-ISM postdoctoral fellowship at the Université de Montréal.

References

BFT15 W. D. Banks, T. Freiberg and C. Turnage-Butterbaugh, *Consecutive primes in tuples*, Acta Arith. **167** (2015), 261–266.

Ben00 J. Benatar, *The existence of small prime gaps in subsets of the integers*, Preprint (2013), arXiv:1305.0348 [math.NT].

CHLPT15 A. Castillo, C. Hall, R. Lemke Oliver, P. Pollack and L. Thompson, *Bounded gaps between primes in number fields and function fields*, Proc. Amer. Math. Soc. **143** (2015), 2841–2856.

Dav00 H. Davenport, *Multiplicative number theory*, Graduate Texts in Mathematics, vol. 74, third edition (Springer, New York, 2000), revised and with a preface by Hugh L. Montgomery.

Fre11 T. Freiberg, *Strings of congruent primes in short intervals*, J. Lond. Math. Soc. (2) **84** (2011), 344–364.

GGPY09 D. A. Goldston, S. W. Graham, J. Pintz and C. Y. Yıldırım, *Small gaps between products of two primes*, Proc. Lond. Math. Soc. (3) **98** (2009), 741–774.

GPY09 D. A. Goldston, J. Pintz and C. Y. Yıldırım, *Primes in tuples. I*, Ann. of Math. (2) **170** (2009), 819–862.

GPY11 D. A. Goldston, J. Pintz and C. Y. Yıldırım, *Positive proportion of small gaps between consecutive primes*, Publ. Math. Debrecen **79** (2011), 433–444.

HWW04 G. Harman, N. Watt and K. Wong, *A new mean-value result for Dirichlet L-functions and polynomials*, Q. J. Math. **55** (2004), 307–324.

HI75 M. N. Huxley and H. Iwaniec, *Bombieri's theorem in short intervals*, Mathematika **22** (1975), 188–194.

KT77 S. Knapowski and P. Turán, *On prime numbers $\equiv 1$ resp. 3 mod 4*, in *Number theory and algebra* (Academic Press, New York, 1977), 157–165.

Kum02 A. Kumchev, *The difference between consecutive primes in an arithmetic progression*, Q. J. Math. **53** (2002), 479–501.

LP00 H. Li and H. Pan, *Bounded gaps between primes of the special form*, Preprint (2014), arXiv:1403.4527 [math.NT].

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

MM87 M. R. Murty and V. K. Murty, *A variant of the Bombieri–Vinogradov theorem*, in *Number theory (Montreal, Quebec, 1985)*, CMS Conference Proceedings, vol. 7 (American Mathematical Society, Providence, RI, 1987), 243–272.

PPS85 A. Perelli, J. Pintz and S. Salerno, *Bombieri's theorem in short intervals II*, Invent. Math. **79** (1985), 1–9.

Pol14 P. Pollack, *Bounded gaps between primes with a given primitive root*, Algebra Number Theory **8** (2014), 1769–1786.

Shi00 D. K. L. Shiu, *Strings of congruent primes*, J. Lond. Math. Soc. (2) **61** (2000), 359–373.

Tho14 J. Thorner, *Bounded gaps between primes in Chebotarev sets*, Res. Math. Sci. **1** (2014), doi:10.1186/2197-9847-1-4.

Tim87 N. M. Timofeev, *Distribution of arithmetic functions in short intervals in the mean with respect to progressions*, Izv. Akad. Nauk SSSR Ser. Mat. **51** (1987), 341–362, 447.

Zha14 Y. Zhang, *Bounded gaps between primes*, Ann. of Math. (2) **179** (2014), 1121–1174.

James Maynard james.alexander.maynard@gmail.com
Centre de recherches mathématiques, Université de Montréal,
Pavillon André-Aisenstadt, 2920 Chemin de la tour, Room 5357,
Montréal (Québec), Canada H3T 1J4

Current address: Mathematical Institute, Radcliffe Observatory Quarter,
Woodstock Road, Oxford, OX1 6GG, UK