Large gaps between primes

By James Maynard

Abstract

We show that there exist pairs of consecutive primes less than x whose difference is larger than

$$t(1 + o(1))(\log x)(\log\log x)(\log\log\log\log x)(\log\log\log x)^{-2}$$

for any fixed t. This answers a well-known question of Erdős.

1. Introduction

Let $G(X) = \sup_{p_n \leq X}(p_{n+1} - p_n)$ denote the maximal gap between primes of size at most X. Westzynthius [15] was the first to show that $G(X)$ could become arbitrarily large compared with the average gap $(1 + o(1))\log X$, and this was improved by Erdős [1] and then Rankin [12], who succeeded in showing that for X sufficiently large

$$(1.1) \qquad G(X) \geq (c + o(1))(\log X)(\log_2 X)(\log_4 X)(\log_3 X)^{-2},$$

with the constant $c = 1/3$, where $\log_v$ denotes the v-fold logarithm. Since Rankin's 1938 result, however, the only improvements have been in the constant c. Such improvements were obtained in the work of Schönhage [14], Rankin [13], Maier and Pomerance [7], with the best constant $c = 2e^\gamma$ due to Pintz [10]. In this paper we show that one can take the constant c to be arbitrarily large by incorporating sieve ideas based on the recent results on small gaps between primes [8], [9], [11] into the Erdős-Rankin method.

THEOREM 1. *We have*

$$\limsup_n \frac{p_{n+1} - p_n}{(\log p_n)(\log_2 p_n)(\log_4 p_n)(\log_3 p_n)^{-2}} = \infty.$$

We note that Ford, Green, Konyagin and Tao have independently obtained this result in the recent work [2] using a different method. Their work is based on incorporating results on linear equations in the primes [4], [6], [5] (rather than work on small gaps between primes) into the Erdős-Rankin construction.

Remark. The method presented here in fact allows one to obtain a quantitative improvement to Rankin's bound (1.1). We shall address this in forthcoming work.

2. **The Erdős-Rankin construction**

As with most approaches to the problem, we follow the Erdős-Rankin construction for large gaps, modifying only the final stage of the argument. We wish to choose residue classes $a_p \pmod{p}$ for each prime $p \le x$ such that every integer $n \in [1, U]$ satisfies $n \equiv a_p \pmod{p}$ for some prime $p \le x$.

We fix constants $C_U, \epsilon > 0$, (and we will assume ϵ is sufficiently small at various parts of the argument) and let y, z, U be defined in terms of x by

$$(2.1) \qquad y = \exp\left((1 - \epsilon)\frac{\log x \log_3 x}{\log_2 x}\right), \qquad z = \frac{x}{\log_2 x}, \qquad U = C_U \frac{x \log y}{\log_2 x}.$$

The only difference between these choices and those of [7] is that here U is determined in terms of an unspecified constant C_U (which we will show can be taken arbitrarily large) rather than a specific choice of C_U slightly less than $1.32 e^\gamma$.

If we can cover the interval $[1, U]$ by residue classes $a_p \pmod{p}$ for $p \le x$, then any U_0 satisfying $U_0 \equiv -a_p \pmod{p}$ for all $p \le x$ has the property that $U_0 + j$ has a prime factor of size at most x for each $j \in [1, U]$. By the Chinese remainder theorem, there is such a U_0 in the interval $[x, x + \exp((1 + o(1))x)]$, and this gives rise to an interval of length U which contains no primes since $U_0 + j$ is bigger than x but contains a prime factor less than x for all $j \in [1, U]$. Letting $x = (1 - \epsilon) \log X$, and recalling our choice of U and y above, we see that this would show there is an interval in $[1, X]$ of length $(1 - 2\epsilon + o(1))C_U(\log X)(\log_2 X)(\log_4 X)(\log_3 X)^{-2}$ containing no primes. Therefore we immediately obtain Theorem 1 if we can take C_U to be arbitrarily large whilst still covering $[1, U]$ by the residue classes $a_p \pmod{p}$ for $p \le x$.

We choose a_p for primes $p \le z$ by

$$(2.2) \qquad \begin{aligned} a_p &= 1 \quad \text{for every prime } 1 < p \le y, \\ a_p &= 0 \quad \text{for every prime } y < p \le z. \end{aligned}$$

After removing elements of $[1, U]$ in these residue classes we are left with the set

$$(2.3) \qquad\qquad\qquad \mathcal{R} \cup \mathcal{R}',$$

where

$$(2.4) \qquad \begin{aligned} \mathcal{R} &= \{mp \le U : p > z, m \text{ is } y\text{-smooth}, (mp - 1, P_y) = 1\}, \\ \mathcal{R}' &= \{m \le U : m \text{ is } y\text{-smooth}, (m - 1, P_y) = 1\}, \\ P_t &= \prod_{p \le t} p \qquad \text{for any } t \in \mathbb{R}. \end{aligned}$$

We first note that since $p > z$, the condition $(mp - 1, P_y) = 1$ requires that m be even. We split $\mathcal{R}$ according to this integer m. For even m, we let

$$(2.5) \qquad \mathcal{R}_m = \{z < p \leq U/m : (mp - 1, P_y) = 1\}.$$

LEMMA 2. *We have*

$$|\mathcal{R}'| \ll \frac{x}{(\log x)^{1+\epsilon}}.$$

Proof. This is [7, Th. 5.3]. Our slightly different choice of U does not affect the argument from [7]. $\qquad\square$

LEMMA 3. *We have uniformly for $z + z/\log x \leq V \leq x(\log x)^2$ and $m \leq x$,*

$$
\begin{aligned}
\#\{z < p \leq V : (mp - 1, P_y) = 1\} \\
= \frac{V - z}{\log x}\Big(\prod_{\substack{p \leq y \\ p \nmid m}} \frac{p-2}{p-1}\Big)\Big(1 + O(\exp(-(\log_2 x)^{1/2}))\Big).
\end{aligned}
$$

In particular, uniformly for even $m \leq U(1 - 1/\log x)/z$, we have

$$|\mathcal{R}_m| = \frac{2e^{-\gamma}U(1 + o(1))}{m(\log x)(\log y)}\Big(\prod_{p > 2} \frac{p(p-2)}{(p-1)^2}\Big)\Big(\prod_{p \mid m, p > 2} \frac{p-1}{p-2}\Big).$$

Proof. The first statement follows from a 'fundamental lemma' sieve and the Bombieri-Vinogradov theorem; see [3, Th. 6.12], for example. The second statement follows immediately from the first using Mertens' theorem. $\qquad\square$

LEMMA 4. *For any $M \geq 2$, we have*

$$\sum_{U/(zM) \leq m < U/z} |\mathcal{R}_m| \ll \frac{U \log M}{(\log x)(\log y)}.$$

In particular,

$$\sum_{U/(z(\log_2 x)^2) \leq m < U/z} |\mathcal{R}_m| = o\Big(\frac{C_U x}{\log x}\Big), \qquad \sum_{1 \leq m < U/(z(\log_2 x)^2)} |\mathcal{R}_m| = O\Big(\frac{C_U x}{\log x}\Big).$$

Proof. Let $w_1 = U/(zM)$, and let $w_2 = U(1 - 1/\log x)/z$. For $m \geq w_2$, we use the trivial bound $|\mathcal{R}_m| \ll U/(m \log x)$, and so we see that the contribution from $w_2 \leq m < U/z$ is at most $O(U/(\log x)^2)$. We now consider $w_1 \leq m < w_2$. Using the bound

$$\prod_{p \mid m,\, p > 2} \frac{p-1}{p-2} \ll \prod_{p \mid m} \frac{p+1}{p} = \sum_{d \mid m} \frac{1}{d},$$

letting $m = m'd$, and using Lemma 3, we have

$$\sum_{w_1 \leq m < w_2} |\mathcal{R}_m| \ll \frac{U}{(\log x)(\log y)} \sum_{d < w_2} \frac{1}{d^2} \sum_{w_1/d \leq m' < w_2/d} \frac{1}{m'} \ll \frac{U(\log M + O(1))}{(\log x)(\log y)}.$$

The final estimates follow immediately from recalling the definitions of U and y. $\qquad\square$

We now state our key proposition.

PROPOSITION 5. *Fix $\delta > 0$. Let $m < Uz^{-1}(\log_2 x)^{-2}$ be even, and let $\mathcal{I}_m \subseteq [x/2, x]$ be an interval of length at least $\delta|\mathcal{R}_m|\log x$.*

Then for $x > x_0(\delta, C_U)$, there exists a choice of residue classes $a_q \pmod{q}$ for each prime $q \in \mathcal{I}_m$ such that

$$p \in \mathcal{R}_m \Rightarrow p \equiv a_q \pmod{q} \text{ for some prime } q \in \mathcal{I}_m.$$

Theorem 1 now follows almost immediately from Proposition 5.

Proof of Theorem 1 assuming Propostion 5. By Lemma 4, we see that

$$(2.6) \qquad \sum_{m < Uz^{-1}(\log_2 x)^{-2}} \delta|\mathcal{R}_m|\log x \ll \delta C_U x.$$

Therefore, if δ is sufficiently small compared with C_U, we can choose intervals $\mathcal{I}_m$ of length $\delta|\mathcal{R}_m|\log x$ for each even $m < Uz^{-1}(\log_2 x)^{-2}$ such that all the $\mathcal{I}_m$ are disjoint and contained in $[x/2, x]$. By Proposition 5 we can cover $\mathcal{R}_m$ using a residue class for each prime in $\mathcal{I}_m$, for each such m. By Lemmas 2, 3 and 4, this means we can cover all but $o(x/\log x)$ elements of $\mathcal{R} \cup \mathcal{R}'$ using only residue classes of primes in $[x/2, x]$. By choosing the residue class of one remaining element for each prime in $[z, x/2]$, we can then cover all the remaining elements of $\mathcal{R} \cup \mathcal{R}'$. Therefore we can choose residue classes $a_p \pmod{p}$ for all $p \le x$ which cover all of $[1, U]$ for any fixed choice of C_U. This completes the proof. $\qquad\square$

We actually prove Proposition 5 in a slightly different (but equivalent) form: We show that for any fixed $\epsilon, \delta > 0$ and interval $\mathcal{I}_m \subseteq [x/2, x]$ of length at least $\delta|\mathcal{R}_m|\log x$, we can choose residue classes $a_q \pmod{q}$ for primes $q \in \mathcal{I}_m$ such that all but $\epsilon|\mathcal{R}_m|$ elements p of $\mathcal{R}_m$ satisfy $p \equiv a_q \pmod{q}$ for some prime $q \in \mathcal{I}_m$ (where m is as in Proposition 5).

By appending to $\mathcal{I}_m$ an interval of length $2\epsilon|\mathcal{R}_m|\log x$, we can choose the residue class of one of the remaining $\epsilon|\mathcal{R}_m|$ elements of $\mathcal{R}_m$ for each of the primes in this appended interval. Thus we can cover all of $\mathcal{R}_m$ by residue classes for primes in an interval of length $(2\epsilon + \delta)|\mathcal{R}_m|\log x$. Since ϵ and δ were arbitrary, we see these two forms of Proposition 5 are equivalent.

3. The probabilistic method

Given an even $m < U/(z(\log_2 x)^2)$ and a prime $q \in \mathcal{I}_m$, we will define a probability measure $\mu_{m,q}$ on the residue classes $a \pmod{q}$. We then consider the following situation: independently for each prime $q \in \mathcal{I}_m$, we randomly choose a residue class $a \pmod{q}$ with probability $\mu_{m,q}(a)$.

Given a prime $p \in \mathcal{R}_m$, we see that the probability that p is not in any of the chosen residue classes for any $q \in \mathcal{I}_m$ is

$$
\prod_{q \in \mathcal{I}_m \text{ prime}} (1 - \mu_{m,q}(p)) = \exp\Big(\sum_{q \in \mathcal{I}_m \text{ prime}} \log\left(1 - \mu_{m,q}(p)\right) \Big)
$$

(3.1)

$$
\leq \exp\Big(- \sum_{q \in \mathcal{I}_m \text{ prime}} \mu_{m,q}(p) \Big).
$$

Therefore, if for almost every $p \in \mathcal{R}_m$ we have that the expected number of $q \in \mathcal{I}_m$ for which the residue class $p \pmod{q}$ is chosen is at least t, then the probability that any such $p \in \mathcal{R}_m$ is not in any of the chosen residue classes is less than e^{-t}. Therefore the expected number of primes in $\mathcal{R}_m$ which are not in any of the chosen residue classes is at most $e^{-t}|\mathcal{R}_m|$. If t can be taken sufficiently large, we expect that all but at most $\epsilon|\mathcal{R}_m|$ elements of $\mathcal{R}_m$ are in at least one of the chosen residue classes. This means that there must be at least one configuration of residue classes $a \pmod{q}$ for $q \in \mathcal{I}_m$ which covers all but at most $\epsilon|\mathcal{R}_m|$ elements of $\mathcal{R}_m$, as required.

4. GPY probabilities

We have seen that to complete the argument we require a probability measure $\mu_{m,q}$ for each prime $q \in \mathcal{I}_m$, such that for almost every $p \in \mathcal{R}_m$, the expected number $\sum_{q \in \mathcal{I}_m} \mu_{m,q}(p)$ of times the residue class $p \pmod{q}$ is chosen is at least t, where t can be taken to be arbitrarily large.

We wish $\mu_{m,q}(a)$ to be large when the residue class $a \pmod{q}$ contains many primes in $\mathcal{R}_m$, and small otherwise. The key feature in this situation is that the modulus q is only slightly smaller than the size of the elements of $\mathcal{R}_m$, which makes it difficult to count the number of primes in a given residue class. To achieve such a measure, we adapt the weights used in [8], [9] to this situation, so that $\mu_{m,q}(a)$ is large when $a \pmod{q}$ contains many elements with no small prime factors.

Specifically, first we choose an admissible set $\mathcal{H} = \{h_1, \ldots, h_k\}$, with $h_j = p_{\pi(k)+j} P_w$ for each $j = 1, \ldots, k$ (i.e., h_j is the jth prime greater than k, multiplied by all primes less than w). Here w is a quantity which will go to infinity slowly with x, such that $P_w = o(\log_2 x)$ (we could take $w = \log_4 x$, for example), and k is a constant we will choose to be sufficiently large in terms of ϵ, δ, C_U. In particular, w will be large compared with k. We define

$$
(4.1) \quad \mu_{m,q}(a) = \alpha_{m,q} \sum_{\substack{n \leq U/m \\ n \equiv a \pmod{q} \\ (n(mn-1), P_w) = 1}} \Big(\sum_{\substack{d_1, \ldots, d_k \\ d_i | n + h_i q}} \sum_{\substack{e_1, \ldots, e_k \\ e_i | m(n+h_i q) - 1}} \lambda_{d_1, \ldots, d_k, e_1, \ldots, e_k} \Big)^2.
$$

Here $\lambda_{d_1, \ldots, d_k, e_1, \ldots, e_k}$ are real constants (which we will choose later), and $\alpha_{m,q}$ is a normalizing constant so that $\sum_{a \pmod{q}} \mu_{m,q}(a) = 1$.

The coefficients $\lambda_{d_1,\ldots,d_k,e_1\ldots,e_k}$ will factorize as $\lambda^{(1)}_{d_1,\ldots,d_k} \prod_{i=1}^{k} \lambda^{(2)}_{e_i}$. The $\lambda^{(1)}_{d_1,\ldots,d_k}$ correspond to a 'GPY' sieve and ensure that $\mu_{m,q}(a)$ can only be large if there exists an $n \equiv a \pmod q$ such that all of $\{n + h_1 q, \ldots, n + h_k q\}$ have no small prime factors (and so we expect many of them to be prime). The $\lambda^{(2)}_{e_j}$ correspond to a standard Selberg sieve[1] and ensure that the contribution from such an n is small unless $m(n + h_j q) - 1$ has no prime factors less than y^ϵ.

If we choose a residue class $a \pmod q$ randomly with probability $\mu_{m,q}(a)$, then for a suitable choice of λ coefficients, we would find from following the work [8], [9], [11] that the expected number of primes in $\mathcal{R}_m$ in the chosen residue class would be a constant multiple of $\log k$. One might hope that the primes found this way would be approximately independent for different $q \in \mathcal{I}_m$. If this were the case, then we would guess that the expected number of times a given prime in $\mathcal{R}_m$ would be in a picked residue class would be roughly the same for all primes in $\mathcal{R}_m$, in which case this would be approximately $(|\mathcal{I}_m| \log k)/(|\mathcal{R}_m| \log x)$, since there are roughly $|\mathcal{I}_m|/\log x$ primes in $\mathcal{I}_m$. Recalling that we choose $|\mathcal{I}_m| = \delta |\mathcal{R}_m| \log x$, we might therefore guess that the expected number of times $p \in \mathcal{R}_m$ is chosen is roughly a constant multiple of $\delta \log k$. (Normally this would actually depend on the arithmetic structure of $\mathcal{H}, m, p_0$, but by choosing all elements of $\mathcal{H}$ to be a multiple of all small primes this effect is negligible.) Therefore, if k is chosen sufficiently large, we expect to be able to make this quantity larger than any fixed constant. We now proceed to make these heuristic ideas rigorous.

In order for it to be feasible to estimate $\sum_q \mu_{m,q}(p)$, we exploit the linearity (in n and q) of the expressions $n + h_i q$ and $m(n + h_i q) - 1$ and make a choice of $\lambda_{d_1,\ldots,d_k,e_1,\ldots,e_k}$ which is independent of q. This allows us to estimate the resulting sums for fixed n and varying prime q and also for fixed q and varying n. In particular, this makes it more convenient to adopt the 'analytic' method for estimating the sums which appear, as in [11].

5. Setup

We let

$$
(5.1) \qquad
\begin{aligned}
\omega_{m,q}(p) = \# \{ 1 \leq n \leq p : n + h_i q \equiv 0 \pmod p \\
\text{or } m(n + h_i q) \equiv 1 \pmod p \text{ for some } 1 \leq i \leq k \},
\end{aligned}
$$

and we extend $\omega_{m,q}$ to a totally multiplicative function defined on $\mathbb{N}$. Similarly, we define the totally multiplicative function $\varphi_{m,q}$ by $\varphi_{m,q}(p) = p - \omega_{m,q}(p)$.

---

[1] We could apply a fundamental lemma type sieve here since $y = x^{o(1)}$, but we find it more convenient to apply a Selberg sieve, which is weaker only by an unimportant constant factor.

We put

$$(5.2) \qquad \mathfrak{S}_{m,q} = \prod_{p \le y} \left(1 - \frac{\omega_{m,q}(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-2k},$$

noting that this product is nonzero since $\omega_{m,q}(p) \le 2$ for $p \le w$ and $\omega_{m,q}(2) = 1$ since we are only considering m even.

We define $\lambda_{d_1,\dots,d_k,e_1\dots,e_k}$ by

$$(5.3) \qquad \lambda_{d_1,\dots,d_k,e_1,\dots,e_k} = \left(\prod_{i=1}^{k} \mu(d_i)\mu(e_i)\right)\sum_{j=1}^{J}\left(\prod_{\ell=1}^{k} F_{\ell,j}\left(\frac{\log d_\ell}{\log x}\right)G\left(\frac{\log e_\ell}{\log y}\right)\right)$$

for some smooth nonnegative functions $F_{i,j}, G : [0,\infty) \to \mathbb{R}$ which are not identically zero (which we declare later). The functions $F_{i,j}, G$ and the quantity J will be allowed to depend on k, but will be independent of x, q. Thus, in particular, $|\lambda_{d_1,\dots,d_k,e_1,\dots,e_k}| \ll_k 1$. We further require that for each $j \in \{1,\dots,J\}$, we have

$$(5.4) \qquad \sup\left\{\sum_{i=1}^{k} u_i : F_{i,j}(u_i) \ne 0\right\} \le 1/10,$$

and we restrict G to be supported on $[0,1]$. Finally, we put

$$(5.5) \qquad F(t_1,\dots,t_k) = \sum_{j=1}^{J}\prod_{\ell=1}^{k} F'_{\ell,j}(t_\ell),$$

and we assume that the $F_{\ell,j}$ are chosen such that F is symmetric. We emphasize that this choice of λ does not depend on q.

6. Sieve estimates

We first asymptotically evaluate the normalizing constant $\alpha_{m,q}$ and then estimate $\sum_{q \in \mathcal{I}_m} \mu_{m,q}(p)$.

LEMMA 6. *We have*

$$\alpha_{m,q}^{-1} = (1 + o_k(1))\frac{U\mathfrak{S}_{m,q}}{m(\log x)^k(\log y)^k}I_k^{(1)}(F)I_k^{(2)}(G),$$

where $\mathfrak{S}_{m,q}$ is given by (5.2), *and*

$$I_k^{(1)}(F) = \int\cdots\int_{t_1,\dots,t_k \ge 0} F(t_1,\dots,t_k)^2 dt_1\dots dt_k, \qquad I_k^{(2)}(G) = \left(\int_0^\infty G'(t)^2 dt\right)^k.$$

Proof. The quantity $\alpha_{m,q}^{-1}$ is somewhat analogous to that of [9, Prop. 9.1], although here we do not concern ourselves with uniformity in k. From the fact

that we have defined $\alpha_{m,q}$ to be such that $\sum_{a \pmod{q}} \mu_{m,q}(a) = 1$, we have that

$$(6.1) \qquad \alpha_{m,q}^{-1} = \sum_{\substack{n \leq U/m \\ (n(mn-1),P_w)=1}} \Big(\sum_{\substack{d_1,\ldots,d_k \\ d_i|n+h_iq}} \sum_{\substack{e_1,\ldots,e_k \\ e_i|m(n+h_iq)-1}} \lambda_{d_1,\ldots,d_k,e_1,\ldots,e_k} \Big)^2.$$

Expanding the squares and swapping the order of summation shows this sum is equal to

$$(6.2) \qquad \sum_{\substack{d_1,\ldots,d_k \, e_1,\ldots,e_k \\ d'_1,\ldots,d'_k \, e'_1,\ldots,e'_k}} \lambda_{\mathbf{d,e}} \lambda_{\mathbf{d',e'}} \sum_{\substack{n \leq U/m \\ (n(mn-1),P_w)=1 \\ [d_i,d'_i]|n+h_iq \, \forall i \\ [e_i,e'_i]|m(n+h_iq)-1 \, \forall i}} 1.$$

Here to ease notation we have written $\lambda_{\mathbf{d,e}}$ for $\lambda_{d_1,\ldots,d_k,e_1\ldots,e_k}$, and similarly for $\lambda_{\mathbf{d',e'}}$.

We concentrate on the inner sum. There is no contribution unless all of $d_1d'_1, \ldots, d_kd'_k, e_1e'_1, \ldots, e_ke'_k$ are coprime to P_w. Moreover, there is no contribution unless all of $d_1d'_1,\ldots,d_kd'_k$ are pairwise coprime, and all of $e_1e'_1,\ldots,e_ke'_k$ are pairwise coprime (since any $p|(e_ie'_i, e_je'_j)$, say, would have to divide $(h_i - h_j)q$, which is in contradiction to $p \leq x^{1/10} < q$ being prime and $(e_ie'_i, P_w) = 1$). Finally, we see that we must have $(d_id'_i, e_je'_j)|mq(h_j - h_i) - 1$ for all $1 \leq i,j \leq k$. If all of these conditions are satisfied, then, by the Chinese remainder theorem, we can combine the divisibility conditions in the inner sum to restrict n to lie in any of $\varphi_{m,q}(P_w)$ residue classes $\pmod{P_w[d_1,d'_1,e_1,e'_1,\ldots,d_k,d'_k,e_k,e'_k]}$.

Thus we see that (6.2) is given by

$$(6.3) \qquad {\sum_{\substack{d_1,\ldots,d_k \, e_1,\ldots,e_k \\ d'_1,\ldots,d'_k \, e'_1,\ldots,e'_k}}}' \lambda_{\mathbf{d,e}} \lambda_{\mathbf{d',e'}} \Big(\frac{U\varphi_{m,q}(P_w)}{m[\mathbf{d,d',e,e'}]P_w} + O(\varphi_{m,q}(P_w)) \Big),$$

where we have written $[\mathbf{d,d',e,e'}]$ for $[d_1,d'_1,e_1,e'_1,\ldots,d_k,d'_k,e_k,e'_k]$ and $\sum'$ for the conditions that $d_1d'_1,\ldots,d_kd'_k,P_w$ are pairwise coprime, $e_1e'_1,\ldots,e_ke'_k,P_w$ are pairwise coprime, and that $(e_ie'_i,d_jd'_j)|(mq(h_j - h_i) - 1, P_y)$ for all $1 \leq i,j \leq k$. The fact we have forced this common factor to divide P_y is redundant since $\lambda_{\mathbf{d,e}}$ is supported only on $e_i \leq y$, but is slightly convenient later.

We first estimate the error term trivially. We have $\sup_{\mathbf{d,e}}(|\lambda_{\mathbf{d,e}}|) \ll_k 1$. We write $d = \prod_{i=1}^k d_i$ (and similarly for d', e, e') and see that the support conditions of $F_{i,j}, G$ mean we only need to consider $d, d' \leq x^{1/10}$ and $e, e' \leq y^k \ll_k x^\epsilon$. Therefore the error term contributes

$$(6.4) \qquad \ll_k \varphi_{m,q}(P_w) \sum_{d,d' \leq x^{1/10}, e,e' \ll_k x^\epsilon} \tau_k(d)\tau_k(d')\tau_k(e)\tau_k(e') \ll_k x^{1/2}.$$

We now estimate the main term. The argument we use is standard and is a minor adaption of [11, Lemma 4.1].

We expand $\lambda_{\mathbf{d},\mathbf{e}}$, $\lambda_{\mathbf{d}',\mathbf{e}'}$ using (5.3). Thus we are left to evaluate

$$(6.5) \qquad \sum_{j=1}^{J}\sum_{j'=1}^{J}\sideset{}{'}\sum_{\substack{d_1,\dots,d_k\\d_1',\dots,d_k'}}\sideset{}{'}\sum_{\substack{e_1,\dots,e_k\\e_1',\dots,e_k'}} \frac{\prod_{\ell=1}^{k}\mu(d_\ell)\mu(d_\ell')\mu(e_\ell)\mu(e_\ell')H_{\ell,j,j'}(d_\ell,d_\ell',e_\ell,e_\ell')}{[\mathbf{d},\mathbf{d}',\mathbf{e},\mathbf{e}']},$$

where

$$H_{\ell,j,j'}(d_\ell,d_\ell',e_\ell,e_\ell') = F_{\ell,j}\Big(\frac{\log d_\ell}{\log x}\Big)F_{\ell,j'}\Big(\frac{\log d_\ell'}{\log x}\Big)G\Big(\frac{\log e_\ell}{\log y}\Big)G\Big(\frac{\log e_\ell'}{\log y}\Big).$$

The function $e^t F_{\ell,j}(t)$ can be extended to a smooth compactly supported function on $\mathbb{R}$, and so has a Fourier expansion $e^t F_{\ell,j}(t) = \int_{\mathbb{R}} e^{-it\xi} f_{\ell,j}(\xi)d\xi$, for a function $f_{\ell,j}$ which (from the smoothness of $e^t F_{\ell,j}(t)$ and integration by parts) satisfies $f_{\ell,j}(\xi) \ll_{k,A} (1+|\xi|)^{-A}$ for any $A > 0$, and so is rapidly decreasing. In particular, we have

$$(6.6) \qquad\qquad F_{\ell,j}\Big(\frac{\log d_\ell}{\log x}\Big) = \int_{\mathbb{R}} \frac{f_{\ell,j}(\xi_\ell)}{d_\ell^{(1+i\xi_\ell)/\log x}}d\xi_\ell.$$

We obtain an analogous expression for G. Thus the sum over the d, d', e, and e' variables in (6.5) can then be rewritten as

$$(6.7) \quad \begin{aligned} \int_{\mathbb{R}}\cdots\int_{\mathbb{R}} &\bigg(\sideset{}{'}\sum_{\substack{d_1,\dots,d_k\\d_1',\dots,d_k'}}\sideset{}{'}\sum_{\substack{e_1,\dots,e_k\\e_1',\dots,e_k'}} \frac{1}{[\mathbf{d},\mathbf{d}',\mathbf{e},\mathbf{e}']}\prod_{\ell=1}^{k}\frac{\mu(d_\ell)\mu(d_\ell')\mu(e_\ell)\mu(e_\ell')}{d_\ell^{\frac{1+i\xi_\ell}{\log x}}(d_\ell')^{\frac{1+i\xi_\ell'}{\log x}}e_\ell^{\frac{1+i\tau_\ell}{\log y}}(e_\ell')^{\frac{1+i\tau_\ell'}{\log y}}}\bigg)\\ &\times\bigg(\prod_{\ell=1}^{k} f_{\ell,j}(\xi_\ell)f_{\ell,j'}(\xi_\ell')g(\tau_\ell)g(\tau_\ell')d\xi_\ell d\xi_\ell' d\tau_\ell d\tau_\ell'\bigg). \end{aligned}$$

Here we have swapped the order of summation and integration (noting the expression is absolutely convergent).

We concentrate on the first term in parentheses in the integral. Since the restrictions imposed on the summation are multiplicative and the summand is also multiplicative, we can rewrite the sum as a product $\prod_p K_p$ for functions $K_p(\xi_1\dots,\xi_k,\xi_1',\dots,\xi_k',\tau_1,\dots,\tau_k,\tau_1',\dots,\tau_k')$. We first notice that

$$(6.8) \quad \begin{aligned} K_p &= \sideset{}{'}\sum_{\substack{d_1,\dots,d_k\\d_1',\dots,d_k'}}\sideset{}{'}\sum_{\substack{e_1,\dots,e_k\\e_1',\dots,e_k'\\[\mathbf{d},\mathbf{d}',\mathbf{e},\mathbf{e}']|p}} \frac{1}{[\mathbf{d},\mathbf{d}',\mathbf{e},\mathbf{e}']}\prod_{\ell=1}^{k}\frac{\mu(d_\ell)\mu(d_\ell')\mu(e_\ell)\mu(e_\ell')}{d_\ell^{\frac{1+i\xi_\ell}{\log x}}(d_\ell')^{\frac{1+i\xi_\ell'}{\log x}}e_\ell^{\frac{1+i\tau_\ell}{\log y}}(e_\ell')^{\frac{1+i\tau_\ell'}{\log y}}}\\ &= 1 + O_k(p^{-1-1/\log x}). \end{aligned}$$

Thus $\prod_p K_p \ll (\log x)^{O_k(1)}$. Since all the f,g functions are rapidly decreasing, this means that we can restrict the integral to $|\xi_\ell|$, $|\xi_\ell'|$, $|\tau_\ell|$, $|\tau_\ell'| \le \sqrt{\log x}$ for all ℓ at the cost of an error $O_k((\log x)^{-2k})$. To ease notation we let $s_j = (1+i\xi_j)/\log x$, $r_\ell = (1+i\tau_\ell)/\log y$ and similarly for s_j', r_ℓ'.

For $w < p \leq y$ with $p \nmid \prod_{h,h' \in \mathcal{H}}(mq(h - h') - 1)$, or for $p > y$, we have

(6.9)

$$K_p = \left(1 + O_k\left(\frac{1}{p^2}\right)\right) \prod_{\ell=1}^{k} \frac{\left(1 - p^{-1-s_\ell}\right)\left(1 - p^{-1-s'_\ell}\right)\left(1 - p^{-1-r_\ell}\right)\left(1 - p^{-1-r'_\ell}\right)}{\left(1 - p^{-1-s_\ell-s'_\ell}\right)\left(1 - p^{-1-r_\ell-r'_\ell}\right)}.$$

For $w < p \leq y$ with $p | \prod_{h,h' \in \mathcal{H}}(mq(h - h') - 1)$, we have terms involving the product of d_j and e_ℓ if $p|mq(h_\ell - h_j) - 1$. This means that we have an additional factor compared with (6.9) of

$$\left(1 + O_k(p^{-2})\right) \prod_{j,\ell: \, p|mq(h_\ell - h_j) - 1} \left(1 + \sum_{\substack{\mathcal{T} \subseteq \{s_j, s'_j, r_\ell, r'_\ell\} \\ \mathcal{T} \cap \{s_j, s'_j\} \neq \emptyset \\ \mathcal{T} \cap \{r_j, r'_j\} \neq \emptyset}} (-1)^{\#\mathcal{T}} p^{-\sum_{t \in \mathcal{T}} t}\right)$$

(6.10)

$$= \left(1 + \frac{\#\{j, \ell : \, p|mq(h_\ell - h_j) - 1\}}{p}\right)\left(1 + O_k\left(\frac{1}{p^2} + \frac{\log p \sqrt{\log x}}{p \log y}\right)\right).$$

Here we used the fact that we have restricted to $|s_j|, |r_\ell| < (\log x)^{1/2}/\log y$ (so, for example, $p^{-s_j} = 1 + O((\log p)(\log x)^{1/2}/(\log y))$ by Taylor expansion) to obtain the final error term.

For such p, we see that the $h_i q \pmod{p}$ are all distinct; $p > w$ implies $p \nmid \prod_{j \neq \ell}(h_j - h_\ell)$ and $p \leq y < q$ implies $p \nmid q$. Therefore, recalling the definition (5.1) of $\omega_{m,q}(p)$, we see that $\omega_{m,q}(p) = 2k - \#\{j, \ell : \, p|mq(h_\ell - h_j) - 1\}$, and so the first factor in (6.10) simplifies to $1 - (\omega_{m,q}(p) - 2k)/p$.

We see that since $\prod_{h,h' \in \mathcal{H}}(mq(h - h') - 1) \ll x^{O_k(1)}$, and $\log y > (\log x)^{1-\epsilon}$, we have

(6.11)

$$\prod_{\substack{p | \prod_{h,h'}(mq(h-h')-1) \\ p > w}} \left(1 + O_k\left(\frac{1}{p^2} + \frac{\log p \sqrt{\log x}}{p \log y}\right)\right)$$

$$= \exp\left(O_k\left(w^{-1} + \frac{\log \log x}{(\log x)^{1/2-\epsilon}}\right)\right)$$

$$= 1 + o_k(1),$$

and so the second factor in (6.10) has a negligible effect.

Finally, we see that since

$$\log w = o((\log x)^\epsilon) \text{ and } s_\ell, s'_\ell, r_\ell, r'_\ell = o(\log x)^{-1/2+\epsilon},$$

we have

(6.12)

$$\prod_{p \leq w} \prod_{\ell=1}^{k} \frac{\left(1 - p^{-1-s_\ell}\right)\left(1 - p^{-1-s'_\ell}\right)\left(1 - p^{-1-r_\ell}\right)\left(1 - p^{-1-r'_\ell}\right)}{\left(1 - p^{-1-s_\ell-s'_\ell}\right)\left(1 - p^{-1-r_\ell-r'_\ell}\right)}$$

$$= (1 + o_k(1)) \prod_{p < w} \left(1 - \frac{1}{p}\right)^{2k}.$$

Putting this all together gives

$$
\prod_{p>w} K_p = (1 + o_k(1)) \prod_{p<w} \left(1 - \frac{1}{p}\right)^{-2k} \prod_{w<p\le y} \left(1 - \frac{\omega_{m,q}(p) - 2k}{p}\right)
$$

(6.13)

$$
\times \prod_{\ell=1}^{k} \frac{\zeta\left(1 + \frac{2+i\xi_\ell+i\xi'_\ell}{\log x}\right)\zeta\left(1 + \frac{2+i\tau_\ell+i\tau'_\ell}{\log y}\right)}{\zeta\left(1 + \frac{1+i\xi_\ell}{\log x}\right)\zeta\left(1 + \frac{1+i\xi'_\ell}{\log x}\right)\zeta\left(1 + \frac{1+i\tau_\ell}{\log y}\right)\zeta\left(1 + \frac{1+i\tau'_\ell}{\log y}\right)}.
$$

Here we have extended the product of $1 - (\omega_{m,q}(p) - 2k)/p$ to all primes $w < p \le y$, which is valid since $\omega_{m,q}(p) = 2k$ if $p \nmid \prod_{h,h'}(mq(h - h') - 1)$ for such p.

In the region $|z| = o(1)$, we have the estimate $\zeta(1 + z) = (1 + o(1))/z$. Thus, recalling that $\log y \ge (\log x)^{1-\epsilon}$, we are left to estimate

(6.14)

$$
\int \cdots \int (1 + o_k(1)) \prod_{\ell=1}^{k} \left(\frac{(1 + i\xi_\ell)(1 + i\xi'_\ell)(1 + i\tau_\ell)(1 + i\tau'_\ell)}{(2 + i\xi_\ell + i\xi'_\ell)(2 + i\tau_\ell + i\tau'_\ell)} \right.
$$

$$
\left. \times f_{\ell,j}(\xi_\ell) f_{\ell,j'}(\xi'_\ell) g(\tau_\ell) g(\tau'_\ell)) d\xi_\ell d\xi'_\ell d\tau_\ell d\tau'_\ell \right),
$$

where the integral is over $|\xi_\ell|, |\xi'_\ell|, |\tau_\ell|, |\tau'_\ell| \le \sqrt{\log x}$. From the rapid decay of the f and g functions, we see that the $o_k(1)$ term contributes $o_k(1)$ in total, and we can extend the integrals to being over $\mathbb{R}$ at a cost of $o_k(1)$. Thus it suffices (since the integrals are absolutely convergent) to show that for any f_1, f_2 amongst the $f_{\ell,j}, g$, we have

(6.15)
$$
\int_{\mathbb{R}} \int_{\mathbb{R}} \frac{(1 + i\xi)(1 + i\xi')}{2 + i\xi + i\xi'} f_1(\xi) f_2(\xi') d\xi d\xi' = \int_0^\infty F'_1(t) F'_2(t) dt.
$$

This follows immediately from our definition of the Fourier transform. (Differentiating under the integral sign is valid due to absolute convergence.)

Putting everything together, we have that

$$
\alpha_{m,q}^{-1} = \frac{(1 + o_k(1))U\varphi_{m,q}(P_w)}{m(\log x)^k(\log y)^k P_w} \prod_{p\le w} \left(1 - \frac{1}{p}\right)^{-2k} \prod_{w\le p\le y} \left(1 - \frac{\omega_{m,q}(p) - 2k}{p}\right)
$$

(6.16)

$$
\times \left(\int_0^\infty G'(t)^2 dt\right)^k \sum_{j=1}^{J} \sum_{j'=1}^{J} \prod_{\ell=1}^{k} \left(\int_0^\infty F'_{\ell,j}(t) F'_{\ell,j'}(t) dt\right)
$$

$$
= (1 + o_k(1)) \frac{U\mathfrak{S}_{m,q} I_k(F)}{m(\log x)^k(\log y)^k} I_k^{(1)}(F) I_k^{(2)}(G).
$$

Here we have used the fact that G and the $F_{\ell,j}$ are all nonnegative (and not the zero function) to take the $o(1)$ errors as a factor at the front of the expression. $\qquad\square$

LEMMA 7. *Let $m < Uz^{-1}(\log x)^{-2}$ be even, and let $p_0 \in \mathcal{R}_m$ with $h_k x < p_0 < U/m - h_k x$. Then*

$$\sum_{q \in \mathcal{I}_m} \mu_{m,q}(p_0) \gg (1 + o_k(1)) \frac{k|\mathcal{I}_m|J_k^{(1)}(F)J_k^{(2)}(G)}{(\log x)|\mathcal{R}_m|I_k^{(1)}(F)I_k^{(2)}(G)},$$

where

$$J_k^{(1)}(F) = \int \cdots \int_{t_1,\ldots,t_{k-1} \geq 0} \left(\int_{t_k \geq 0} F(t_1, \ldots, t_k) dt_k \right)^2 dt_1 \ldots dt_{k-1},$$

$$J_k^{(2)}(G) = G(0)^2 \left(\int_0^\infty G'(t)^2 dt \right)^{k-1}.$$

Proof. We substitute the definition of $\mu_{m,q}$ to give

(6.17)

$$\sum_{\substack{q \in \mathcal{I}_m \text{ prime}}} \mu_{m,q}(p_0)$$

$$= \sum_{\substack{q \in \mathcal{I}_m \text{ prime}}} \alpha_{m,q} \sum_{\substack{n \leq U/m \\ n \equiv p_0 \pmod{q} \\ (n(mn-1),P_w)=1}} \left(\sum_{\substack{d_1,\ldots,d_k \\ d_i|n+h_iq}} \sum_{\substack{e_1,\ldots,e_k \\ e_i|m(n+h_iq)-1}} \lambda_{d_1,\ldots,d_k,e_1,\ldots,e_k} \right)^2.$$

Since all terms are nonnegative, we obtain a lower bound by dropping all terms in the sum over n except for when $n = p_0 - hq$ for some $h \in \mathcal{H}$. We see that $(p_0 + (h_i - h)q, P_w) = (p_0, P_w) = 1$ and $(mp_0 + m(h_i - h)q - 1, P_w) = (mp_0 - 1, P_w) = 1$ so all the terms $n = p_0 - hq$ appear in the sum (since, by assumption, $h_k x < p_0 < U/m - h_k x$). This gives

(6.18)

$$\sum_{\substack{q \in \mathcal{I}_m \text{ prime}}} \mu_{m,q}(p_0)$$

$$\geq \sum_{h \in \mathcal{H}} \sum_{\substack{q \in \mathcal{I}_m \text{ prime}}} \alpha_{m,q} \left(\sum_{\substack{d_1,\ldots,d_k \\ d_i|p_0+(h_i-h)q}} \sum_{\substack{e_1,\ldots,e_k \\ e_i|m(p_0+(h_i-h)q)-1}} \lambda_{d_1,\ldots,d_k,e_1,\ldots,e_k} \right)^2.$$

We split the sum over q into residue classes modulo P_w. This gives

(6.19)

$$\sum_{h \in \mathcal{H}} \sum_{\substack{w_0 \pmod{P_w} \\ (w_0,P_w)=1}}$$

$$\times \sum_{\substack{q \in \mathcal{I}_m \text{ prime} \\ q \equiv w_0 \pmod{P_w}}} \alpha_{m,q} \left(\sum_{\substack{d_1,\ldots,d_k \\ d_i|p_0+(h_i-h)q}} \sum_{\substack{e_1,\ldots,e_k \\ e_i|m(p_0+(h_i-h)q)-1}} \lambda_{d_1,\ldots,d_k,e_1,\ldots,e_k} \right)^2.$$

We now replace $\alpha_{m,q}$ with a slightly more manageable expression with less dependence on q. We first note that since all the h_i are a multiple of P_w, for

every prime $p \leq w$, we have that $\omega_{m,q}(p) = 2$ or 1 depending on whether or not $(m, p) = 1$. For $w < p \leq y$, we have $\omega_{m,q}(p) = 2k$ if

$$p \nmid \prod_{h',h'' \in \mathcal{H}} (mq(h' - h'') - 1).$$

Thus, recalling that we only consider even m, we have

$$\mathfrak{S}_{m,q}^{-1} = \prod_{p \leq y} \left(1 - \frac{\omega_{m,q}(p)}{p}\right)^{-1} \left(1 - \frac{1}{p}\right)^{2k}$$

(6.20)
$$\geq (1 + o_k(1))\mathfrak{S}_m \prod_{\substack{w < p \leq y \\ p \mid \prod_{h',h''} (mq(h'-h'')-1)}} \left(1 - \frac{2k}{p}\right),$$

where we have put

(6.21)
$$\mathfrak{S}_m = 2^{-(2k-1)} \left(\prod_{p \mid m,\, p > 2} \frac{p-2}{p-1} \right) \prod_{2 < p \leq w} \left(1 - \frac{2}{p}\right)^{-1} \left(1 - \frac{1}{p}\right)^{2k}.$$

Since $\prod_{h',h''}(mq(h' - h'') - 1) \ll x^{O_k(1)}$, we can restrict the primes occurring in the final product on the right hand side of (6.20) to be less than $z_0 = \log x / \log_2 x$ at a cost of a factor of $1 + o_k(1)$. Expanding out this product then gives

(6.22)
$$\mathfrak{S}_{m,q}^{-1} \geq (1 + o_k(1))\mathfrak{S}_m \sum_{\substack{a_{1,2},\ldots,a_{k,k-1} \mid P_{z_0}/P_w \\ a_{i,j} \mid mq(h_i-h_j)-1}} \frac{(-2k)^{\omega([\mathbf{a}])}}{[\mathbf{a}]}.$$

Here we have put $[\mathbf{a}] = [a_{1,2}, \ldots, a_{k,k-1}]$. Substituting this bound for $\mathfrak{S}_{m,q}$ and our bound for $\alpha_{m,q}$ from Lemma 6 into (6.18), we have

(6.23)

$$\sum_{q \in \mathcal{I}_m \text{ prime}} \mu_{m,q}(p_0) \geq \frac{(1 + o_k(1))m(\log x)^k (\log y)^k \mathfrak{S}_m}{U I_k^{(1)}(F) I_k^{(2)}(G)}$$

$$\times \sum_{\substack{h \in \mathcal{H}}} \sum_{\substack{w_0 \pmod{P_w} \\ (w_0, P_w)=1}} \sum_{\substack{a_{1,2},\ldots,a_{k,k-1} \mid P_{z_0}/P_w}} \frac{(-2k)^{\omega([\mathbf{a}])}}{[\mathbf{a}]}$$

$$\times \sum_{\substack{q \in \mathcal{I}_m \text{ prime} \\ q \equiv w_0 \pmod{P_w} \\ a_{i,j} \mid mq(h_i-h_j)-1 \, \forall \, i \neq j}} \left(\sum_{\substack{d_1,\ldots,d_k \\ d_i \mid p_0+(h_i-h)q}} \sum_{\substack{e_1,\ldots,e_k \\ e_i \mid m(p_0+(h_i-h)q)-1}} \lambda_{d_1,\ldots,d_k,e_1,\ldots,e_k} \right)^2.$$

We concentrate on the sum over q. For convenience we will consider the case when h in the outer sum is h_k; the other cases are entirely analogous.

Since p_0 is a prime larger than x, and $(mp_0 - 1, P_y) = 1$, we may restrict to $d_k = e_k = 1$ since no other divisors of p_0 or $mp_0 - 1$ occur. Inserting this

condition, expanding the square and swapping the order of summation then gives that the sum over q is equal to

$$(6.24) \qquad \sum_{\substack{d_1,\ldots,d_k \\ d_1',\ldots,d_k' \\ d_k=d_k'=1}} \sum_{\substack{e_1,\ldots,e_k \\ e_1',\ldots,e_k' \\ e_k=e_k'=1}} \lambda_{\mathbf{d},\mathbf{e}}\lambda_{\mathbf{d}',\mathbf{e}'} \sum_{\substack{q\in\mathcal{I}_m \text{ prime} \\ [d_i,d_i']\mid p_0+(h_i-h_k)q\,\forall\, i \\ [e_i,e_i']\mid mp_0+m(h_i-h_k)q-1\,\forall\, i \\ q\equiv w_0 \pmod{P_w} \\ a_{i,j}\mid mq(h_i-h_j)-1\,\forall\, i\neq j}} 1.$$

This is now an expression which is similar to that considered in the proof of [9, Prop. 9.2]. Let us be given $a_{1,2},\ldots,a_{k,k-1}\mid P_{z_0}/P_w$ and $\mathbf{d},\mathbf{d}',\mathbf{e},\mathbf{e}'$ with $\lambda_{\mathbf{d},\mathbf{e}}\lambda_{\mathbf{d}',\mathbf{e}'}\neq 0$ and with $d_k=d_k'=e_k=e_k'=1$. We see that the inner sum over q is empty unless $d_1d_1',\ldots,d_kd_k',P_w$ are pairwise coprime, $e_1e_1',\ldots,e_ke_k',P_w$ are pairwise coprime and $a_{1,2},\ldots,a_{k,k-1},m$ are pairwise coprime. Moreover, we must also have that

$$(6.25) \qquad \begin{aligned} (d_id_i',e_je_j')&\mid mp_0(h_i-h_j)+h_k-h_i & &\forall\, i,j, \\ (d_id_i',a_{j,\ell})&\mid (h_j-h_\ell)mp_0+h_i-h_k & &\forall\, i,j,\ell, \\ (e_ie_i',a_{j,\ell})&\mid (h_j-h_\ell)(1-p_0m)-h_i+h_k & &\forall\, i,j,\ell. \end{aligned}$$

If all of these conditions are satisfied, then the inner sum can be rewritten as a sum over primes in $\mathcal{I}_m$ in a single residue class modulo the least common multiple of $d_1,\, d_1',\, e_1,\, e_1',\ldots,\, d_k,\, d_k',\, e_k,\, e_k',\, a_{1,2}\ldots,\, a_{k,k-1},\, P_w$. Moreover, this residue class will be coprime to the modulus. Thus in this case the inner sum is

$$(6.26) \qquad \frac{\sum_{q\in\mathcal{I}_m \text{ prime}} 1}{\varphi(P_w)\varphi([\mathbf{d},\mathbf{d}',\mathbf{e},\mathbf{e}',\mathbf{a}])} + O(E(x;P_w[\mathbf{d},\mathbf{d}',\mathbf{e},\mathbf{e}',\mathbf{a}])),$$

where we have written $[\mathbf{d},\mathbf{d}',\mathbf{e},\mathbf{e}',\mathbf{a}]$ to represent the least common multiple of all the $d,\, d',\, e,\, e'$ and a variables, and where

$$(6.27) \qquad E(x;q) = \sup_{t\leq x} \sup_{(a,q)=1} \left| \pi(t;q,a) - \frac{\pi(t)}{\varphi(q)} \right|.$$

We put $r = P_w[\mathbf{d},\mathbf{d}',\mathbf{e},\mathbf{e}',\mathbf{a}]$ and note that $r \ll x^{1/5+o_k(1)}$, from the support conditions on $F_{\ell,j}$ and G and the fact that $a_{j,l}\mid P_{z_0} = x^{o(1)}$. Therefore, using the trivial bound $E(x;q) \ll x/q$ and the Bombieri-Vinogradov theorem, we see the contribution of the error to (6.24) is

$$(6.28) \qquad \begin{aligned} &\ll_k \sum_{r\ll x^{1/5+\epsilon}} \tau_{k^2+4k}(r)E(x;r) \\ &\ll \Big(x \sum_{r\ll x^{1/5+\epsilon}} \frac{\tau_{k^2+4k}(r)^2}{r} \Big)^{1/2} \Big(\sum_{r\ll x^{1/5+\epsilon}} E(x;r) \Big)^{1/2} \ll_k \frac{x}{(\log x)^{2k}}, \end{aligned}$$

which will be negligible. (Here we used $|\lambda_{\mathbf{d},\mathbf{e}}| \ll_k 1$ for all $\mathbf{d},\mathbf{e}$.)

$\mathcal{I}_m$ is an interval of length $\delta|\mathcal{R}_m|\log x \gg x(\log x)^{-2}$ by Lemma 3 and our bound on m. Since $\mathcal{I}_m$ is contained in $[x/2, x]$, the number of primes in $\mathcal{I}_m$ is $(1 + o(1))|\mathcal{I}_m|/\log x$ by the prime number theorem. Therefore the main term of (6.24) then simplifies to

(6.29)
$$\frac{(1 + o(1))|\mathcal{I}_m|}{\varphi(P_w)\log x} \sideset{}{^*}\sum_{\substack{d_1,\ldots,d_k \\ d_1',\ldots,d_k'}} \sideset{}{^*}\sum_{\substack{e_1,\ldots,e_k \\ e_1',\ldots,e_k'}} \frac{\lambda_{\mathbf{d},\mathbf{e}}\lambda_{\mathbf{d}',\mathbf{e}'}}{\varphi([\mathbf{d},\mathbf{d}',\mathbf{e},\mathbf{e}',\mathbf{a}])}$$

$$= \frac{(1 + o(1))|\mathcal{I}_m|G(0)^2}{\varphi(P_w)\log x} \sum_{j=1}^{J}\sum_{j'=1}^{J} F_{k,j}(0)F_{k,j'}(0) \sideset{}{^{**}}\sum_{\substack{d_1,\ldots,d_{k-1} \\ d_1',\ldots,d_{k-1}'}} \sideset{}{^{**}}\sum_{\substack{e_1,\ldots,e_{k-1} \\ e_1',\ldots,e_{k-1}'}}$$

$$\times \frac{\prod_{\ell=1}^{k-1}\mu(d_\ell)\mu(d_\ell')\mu(e_\ell)\mu(e_\ell')F_{\ell,j}\left(\frac{\log d_\ell}{\log x}\right)F_{\ell,j'}\left(\frac{\log d_\ell'}{\log x}\right)G\left(\frac{\log e_\ell}{\log y}\right)G\left(\frac{\log e_\ell'}{\log y}\right)}{\varphi([\mathbf{d},\mathbf{d}',\mathbf{e},\mathbf{e}',\mathbf{a}])}.$$

Here we have written $\sum^*$ to indicate that the summation is restricted to the conditions that $d_k = d_k' = e_k = e_k' = 1$, that $d_1d_1',\ldots,d_kd_k', P_w$ are pairwise coprime, that $e_1e_1',\ldots,e_ke_k', P_w$ are pairwise coprime, and that the divisibility constraints (6.25) are satisfied. Similarly, we have written $\sum^{**}$ for these constraints with the conditions on d_k, d_k', e_k, e_k' dropped (because we have separated their contribution).

We can evaluate this by an essentially identical argument to that used in Lemma 6. The presence of the Euler φ function in the denominator affects K_p by a factor $1 + O_k(p^{-2})$, and so has a negligible effect. The presence of the $\mathbf{a}$ factor in the denominator means that for any $p|a_{i,j}$, we have $K_p \ll_k p^{-1}$. This means that the contribution to (6.24) when $a_{1,2},\ldots,a_{k,k-1} \neq 1,\ldots,1$ is

(6.30)
$$\ll_k \frac{\mathfrak{S}_{m,p_0,h_k}^{(2)}|\mathcal{I}_m|}{\varphi(P_w)(\log x)^k(\log y)^{k-1}} \prod_{p|a_{i,j} \text{ for some } i,j} \frac{O_k(1)}{p},$$

where

(6.31) $\mathfrak{S}_{m,p_0,h}^{(2)} = \prod_{p\leq w}\left(1 - \frac{1}{p}\right)^{-(2k-2)} \prod_{w<p\leq y}\left(1 - \frac{\omega_{m,p_0,h}^{(2)}(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-(2k-2)},$

(6.32) $\begin{aligned}\omega_{m,p_0,h}^{(2)}(p) = \#\{1 \leq n \leq p &: p_0 + (h_i - h)n \equiv 0 \pmod{p} \\ &\text{or } m(p_0 + (h_i - h)n) \equiv 1 \pmod{p} \text{ for some } i\}.\end{aligned}$

Thus (6.30) shows that the total contribution from all $a_{1,2}, \ldots, a_{k,k-1} \neq 1, \ldots, 1$ is

(6.33)

$$
\sum_{\substack{a_{1,2},\ldots,a_{k,k-1} | P_{z_0}/P_w \\ a_{1,2},\ldots,a_{k,k-1} \neq 1,\ldots,1}} \frac{(-2k)^{\omega([\mathbf{a}])}}{[\mathbf{a}]}
$$

$$
\times \sum_{\substack{q \in \mathcal{I}_m \text{ prime} \\ q \equiv w_0 \ (\mathrm{mod}\ P_w) \\ a_{i,j} | mq(h_i - h_j) - 1\, \forall\, i \neq j}} \Big(\sum_{\substack{d_1,\ldots,d_k \\ d_i | p_0 + (h_i - h_k)q}} \sum_{\substack{e_1,\ldots,e_k \\ e_i | m(p_0 + (h_i - h_k)q) - 1}} \lambda_{d_1,\ldots,d_k,e_1,\ldots,e_k} \Big)^2
$$

$$
\ll_k \frac{\mathfrak{S}^{(2)}_{m,p_0,h_k} |\mathcal{I}_m|}{\varphi(P_w)(\log x)^k(\log y)^{k-1}} \Big(\prod_{w < p \leq z_0} \Big(1 + \frac{O_k(1)}{p^2} \Big) - 1 \Big)
$$

$$
= o_k\Big(\frac{\mathfrak{S}^{(2)}_{m,p_0,h_k} |\mathcal{I}_m|}{\varphi(P_w)(\log x)^k(\log y)^{k-1}} \Big).
$$

Hence the main contribution comes from when $a_{1,2} = \cdots = a_{k,k-1} = 1$. In this case, the contribution is

(6.34)

$$
\frac{(1 + o_k(1))\mathfrak{S}^{(2)}_{m,p_0,h_k} |\mathcal{I}_m| G(0)^2}{\varphi(P_w)(\log x)^k(\log y)^{k-1}} \sum_{j=1}^{J} \sum_{j'=1}^{J} F_{k,j}(0) F_{k,j'}(0) \Big(\int_0^\infty G'(t)^2 dt \Big)^{k-1}
$$

$$
\times \prod_{\ell=1}^{k-1} \int_0^\infty F'_{\ell,j}(t) F'_{\ell,j'}(t) dt
$$

$$
= \frac{(1 + o_k(1))\mathfrak{S}^{(2)}_{m,p_0,h_k} |\mathcal{I}_m|}{\varphi(P_w)(\log x)^k(\log y)^{k-1}} J_k^{(1)}(F) J_k^{(2)}(G).
$$

We obtain the same estimate (6.34) with $\mathfrak{S}^{(2)}_{m,p_0,h_k}$ replaced with $\mathfrak{S}^{(2)}_{m,p_0,h}$ for a different $h \in \mathcal{H}$. (Since F is symmetric, $J_k^{(1)}(F)$ does not depend on the choice of $h \in \mathcal{H}$.) Substituting (6.33) and (6.34) into our main term (6.23), we obtain

(6.35)

$$
\sum_{q \in \mathcal{I}_m \text{ prime}} \mu_{m,q}(p_0) \geq \frac{(1 + o_k(1)) m(\log y) J_k^{(1)}(F) J_k^{(2)}(G) |\mathcal{I}_m|}{U I_k^{(1)}(F) I_k^{(2)}(G) \varphi(P_w)}
$$

$$
\times \sum_{h \in \mathcal{H}} \mathfrak{S}_m \mathfrak{S}^{(2)}_{m,p_0,h} \sum_{\substack{w_0 \ (\mathrm{mod}\ P_w) \\ (w_0, P_w) = 1}} 1.
$$

The inner sum is clearly $\varphi(P_w)$. We note that for any $h \in \mathcal{H}$, we have $\omega^{(2)}_{m,p_0,h}(p) = 0$ for $p \leq w$ and $\omega^{(2)}_{m,p_0,h}(p) \leq 2k - 2$ for all p. Therefore, recalling the

definitions (6.21) and (6.31) of $\mathfrak{S}_m$ and $\mathfrak{S}_{m,p_0,h}$, we have

$$\mathfrak{S}_m\mathfrak{S}_{m,p_0,h}^{(2)} = 2^{-1} \prod_{p|m,\,p>2} \frac{p-2}{p-1} \prod_{p<w} \frac{\left(1-\frac{1}{p}\right)^2}{\left(1-\frac{2}{p}\right)} \prod_{w<p\leq y} \frac{\left(1-\frac{\omega_{m,p_0,h}^{(2)}(p)}{p}\right)}{\left(1-\frac{1}{p}\right)^{2k-2}}$$

(6.36)
$$\gg \prod_{p|m,\,p>2} \frac{p-2}{p-1} \prod_{w<p\leq y} \left(1-\frac{2k-2}{p}\right)\left(1-\frac{1}{p}\right)^{-(2k-2)}$$

$$\gg (1+o_k(1)) \prod_{p|m,\,p>2} \frac{p-2}{p-1}.$$

Recalling the estimate on the size of $\mathcal{R}_m$ from Lemma 3 then gives the result.
$\square$

Finally, we recall the key integral estimate from [8].

LEMMA 8. *There exists a suitable choice of smooth functions F,G such that*

$$\frac{kJ_k^{(1)}(F)J_k^{(2)}(G)}{I_k^{(1)}(F)I_k^{(2)}(G)} \gg \log k.$$

Proof. We choose $G(t)$ to be a smooth approximation to $1-t$ supported on $t \in [0,1]$ such that $J_k^{(2)}(G)/I_k^{(2)}(G) \geq 1-\epsilon$. We can choose such a G since the set of nonnegative continuous functions supported on $[0,1]$ is L^2- and L^1-dense in the set of continuous nonnegative functions supported on $[0,1]$ (by the Stone-Weierstrasss theorem).

We recall the choice of function giving [8, Prop. 4.3(iii)]. Let $F_k : [0,\infty)^k \to \mathbb{R}$ and $g : [0,\infty) \to \mathbb{R}$ be defined by

(6.37)
$$F_k(t_1,\ldots,t_k) = \begin{cases} \prod_{i=1}^k g(kt_i) & \text{if } \sum_{i=1}^k t_i \leq 1, \\ 0 & \text{otherwise,} \end{cases}$$

(6.38)
$$g(t) = \begin{cases} 1/(1+At) & t \in [0,T], \\ 0 & \text{otherwise,} \end{cases}$$

and where $A = \log k - 2\log_2 k$ and $T = (e^A-1)/A$. This is clearly a nonnegative symmetric function defined on $\sum_{i=1}^k t_i \leq 1$.

By [8, §7], we have that $J_k^{(1)}(F_k)/I_k^{(1)}(F_k) \gg (\log k)/k$. We choose $F_{\ell,j}$ such that F is a smooth approximation to $F_k(10t_1,\ldots,10t_k)$ supported on $\{(t_1,\ldots,t_k) \in [0,\infty)^k : \sum_{i=1}^k t_i \leq 1/10\}$ with

$$J_k^{(1)}(F)/I_k^{(1)}(F) \geq (1/10-\epsilon)J_k^{(1)}(F_k)/I_k^{(1)}(F_k),$$

which gives the result.

We can choose such an approximation F since the set of symmetric, non-negative linear combinations of direct products of smooth compactly supported functions on $[0,1]^k$ is L^2- and L^1-dense in the set of nonnegative symmetric L^2-integrable functions on $[0,1]^k$. $\qquad\square$

Completion of the proof of Proposition 5. Since

$$|\mathcal{I}_m| = (1 + o(1))\delta|\mathcal{R}_m| \log x,$$

by Lemmas 7 and 8 we have that any prime $p_0 \in \mathcal{R}_m$ with $h_k x < p_0 < U/m - h_k x$ has the expected number $\sum_q \mu_{m,q}(p_0)$ of times p_0 is chosen is $\gg \delta \log k$. By Lemma 3, the number of primes $p_0 \in \mathcal{R}_m$ which do not satisfy $h_k x < p_0 < U/m - h_k x$ is $o_k(|\mathcal{R}_m|)$ for $m < U/z(\log_2 x)^2$.

Thus we have shown that if we choose residue classes randomly according to $\mu_{m,q}$, for all but $o_k(|\mathcal{R}_m|)$ primes $p_0 \in \mathcal{R}_m$, the expected number $\sum_q \mu_{m,q}(p_0)$ of times $p_0 \in \mathcal{R}_m$ is chosen is $\gg \delta \log k$. By choosing k sufficiently large in terms of δ, ϵ, we can ensure that this expectation is larger than $\log \epsilon^{-1}$. But, by the argument of Section 3, this means that the expected number of $p_0 \in \mathcal{R}_m$ which are not in any of the chosen residue classes must be less than $\epsilon|\mathcal{R}_m|$. Therefore there must be at least one assignment of residue classes for which at most $\epsilon|\mathcal{R}_m|$ of the primes $p_0 \in \mathcal{R}_m$ are not chosen. By the argument at the end of Section 2, this implies Proposition 5, and hence Theorem 1. $\qquad\square$

7. **Acknowledgments**

References

[1] P. ERDÖS, The difference of consecutive primes, *Duke Math. J.* **6** (1940), 438–441. MR 0001759. Zbl 0023.29801. http://dx.doi.org/10.1215/S0012-7094-40-00635-4.

[2] K. FORD, B. GREEN, S. KONYAGIN, and T. TAO, Large gaps between consecutive prime numbers, *Ann. of Math.* **183** (2016), 935–974. http://dx.doi.org/10.4007/annals.2016.183.3.4.

[3] J. FRIEDLANDER and H. IWANIEC, *Opera de Cribro*, Amer. Math. Soc. Colloq. Publ. **57**, Amer. Math. Soc., Providence, RI, 2010. MR 2647984. Zbl 1226.11099.

[4] B. GREEN and T. TAO, Linear equations in primes, *Ann. of Math.* **171** (2010), 1753–1850. MR 2680398. Zbl 1242.11071. http://dx.doi.org/10.4007/annals.2010.171.1753.

[5] B. Green and T. Tao, The Möbius function is strongly orthogonal to nilsequences, *Ann. of Math.* **175** (2012), 541–566. MR 2877066. Zbl 06024998. http://dx.doi.org/10.4007/annals.2012.175.2.3.

[6] B. Green, T. Tao, and T. Ziegler, An inverse theorem for the Gowers $U^{s+1}[N]$-norm, *Ann. of Math.* **176** (2012), 1231–1372. MR 2950773. Zbl 1282. 11007. http://dx.doi.org/10.4007/annals.2012.176.2.11.

[7] H. Maier and C. Pomerance, Unusually large gaps between consecutive primes, *Trans. Amer. Math. Soc.* **322** (1990), 201–237. MR 0972703. Zbl 0706. 11052. http://dx.doi.org/10.2307/2001529.

[8] J. Maynard, Small gaps between primes, *Ann. of Math.* **181** (2015), 383–413. MR 3272929. Zbl 1306.11073. http://dx.doi.org/10.4007/annals.2015.181.1.7.

[9] J. Maynard, Dense clusters of primes in subsets, 2016, *Composito Math.*, to appear. arXiv 1405.2593.

[10] J. Pintz, Very large gaps between consecutive primes, *J. Number Theory* **63** (1997), 286–301. MR 1443763. Zbl 0870.11056. http://dx.doi.org/10.1006/jnth. 1997.2081.

[11] D. H. J. Polymath, Variants of the Selberg sieve, and bounded intervals containing many primes, *Res. Math. Sci.* **1** (2014), Art. 12, 83. MR 3373710. http://dx.doi.org/10.1186/s40687-014-0012-7.

[12] R. A. Rankin, The difference between consecutive prime numbers, *J. London Math. Soc.* **S1-13** (1938), 242–247. MR 1574971. Zbl 0019.39403. http://dx.doi. org/10.1112/jlms/s1-13.4.242.

[13] R. A. Rankin, The difference between consecutive prime numbers. V, *Proc. Edinburgh Math. Soc.* **13** (1962/1963), 331–332. MR 0160767. Zbl 0121.04705. http://dx.doi.org/10.1017/S0013091500025633.

[14] A. Schönhage, Eine Bemerkung zur Konstruktion grosser Primzahllücken, *Arch. Math.* (*Basel*) **14** (1963), 29–30. MR 0146154. Zbl 0108.04504. http: //dx.doi.org/10.1007/BF01234916.

[15] E. Westzynthius, Uber die Verteilung der Zahlen, die zu den n ersten Primzahlen teilerfremd sind., *Comm. Phys. Math. Soc. Sci. Fenn.* **25** (1931), 1–37. Zbl 0003.24601.

Magdalen College, Oxford, England
E-mail: james.alexander.maynard@gmail.com