Annals of Mathematics **181** (2015), 383–413
http://dx.doi.org/10.4007/annals.2015.181.1.7

Small gaps between primes

By James Maynard

Abstract

We introduce a refinement of the GPY sieve method for studying prime k-tuples and small gaps between primes. This refinement avoids previous limitations of the method and allows us to show that for each k, the prime k-tuples conjecture holds for a positive proportion of admissible k-tuples. In particular, $\liminf_n(p_{n+m}-p_n)<\infty$ for every integer m. We also show that $\liminf(p_{n+1}-p_n)\le 600$ and, if we assume the Elliott-Halberstam conjecture, that $\liminf_n(p_{n+1}-p_n)\le 12$ and $\liminf_n(p_{n+2}-p_n)\le 600$.

1. **Introduction**

We say that a set $\mathcal{H}=\{h_1,\dots,h_k\}$ of distinct nonnegative integers is 'admissible' if, for every prime p, there is an integer a_p such that $a_p\not\equiv h \pmod p$ for all $h\in\mathcal{H}$. We are interested in the following conjecture.

Conjecture (Prime k-tuples conjecture). *Let $\mathcal{H}=\{h_1,\dots,h_k\}$ be admissible. Then there are infinitely many integers n such that all of $n+h_1,\dots,n+h_k$ are prime.*

When $k>1$, no case of the prime k-tuples conjecture is currently known. Work on approximations to the prime k-tuples conjecture has been very successful in showing the existence of small gaps between primes, however. In their celebrated paper [5], Goldston, Pintz and Yıldırım introduced a new method for counting tuples of primes, and this allowed them to show that

$$\liminf_n \frac{p_{n+1}-p_n}{\log p_n}=0. \tag{1.1}$$

The recent breakthrough of Zhang [9] managed to extend this work to prove

$$\liminf_n (p_{n+1}-p_n)\le 70\,000\,000, \tag{1.2}$$

thereby establishing for the first time the existence of infinitely many bounded gaps between primes. Moreover, it follows from Zhang's theorem that the number of admissible sets of size 2 contained in $[1,x]^2$ which satisfy the prime 2-tuples conjecture is $\gg x^2$ for large x. Thus, in this sense, a positive proportion of admissible sets of size 2 satisfy the prime 2-tuples conjecture. The recent polymath project [7] has succeeded in reducing the bound (1.2) to 4680, by optimizing Zhang's arguments and introducing several new refinements.

The above results have used the 'GPY method' to study prime tuples and small gaps between primes, and this method relies heavily on the distribution of primes in arithmetic progressions. Given $\theta > 0$, we say the primes have 'level of distribution θ'[1] if, for every $A > 0$, we have

$$\sum_{q\le x^\theta}\max_{(a,q)=1}\left|\pi(x;q,a)-\frac{\pi(x)}{\varphi(q)}\right| \ll_A \frac{x}{(\log x)^A}. \tag{1.3}$$

The Bombieri-Vinogradov theorem establishes that the primes have level of distribution θ for every $\theta < 1/2$, and Elliott and Halberstam [1] conjectured that this could be extended to every $\theta < 1$. Friedlander and Granville [2] have shown that (1.3) cannot hold with x^θ replaced with $x/(\log x)^B$ for any fixed B, and so the Elliott-Halberstam conjecture is essentially the strongest possible result of this type.

The original work of Goldston, Pintz and Yıldırım showed the existence of bounded gaps between primes if (1.3) holds for some $\theta > 1/2$. Moreover, under the Elliott-Halberstam conjecture one had $\liminf_n(p_{n+1}-p_n) \le 16$. The key breakthrough of Zhang's work was in establishing that a slightly weakened form of (1.3) holds for some $\theta > 1/2$.

If one looks for bounded length intervals containing two or more primes, then the GPY method fails to prove such strong results. Unconditionally we are only able to improve upon the trivial bound from the prime number theorem by a constant factor [6], and even assuming the Elliott-Halberstam conjecture, the best available result [5] is

$$\liminf_n \frac{p_{n+2}-p_n}{\log p_n} = 0. \tag{1.4}$$

The aim of this paper is to introduce a refinement of the GPY method which removes the barrier of $\theta = 1/2$ to establishing bounded gaps between primes and allows us to show the existence of arbitrarily many primes in bounded length intervals. This answers the second and third questions posed in [5] on extensions of the GPY method (the first having been answered by Zhang's result). Our new method also has the benefit that it produces numerically superior results to previous approaches.

Theorem 1.1. *Let $m \in \mathbb{N}$. We have*

$$\liminf_n (p_{n+m}-p_n) \ll m^3 e^{4m}.$$

[1]We note that different authors have given slightly different names or definitions to this concept. For the purposes of this paper, (1.3) will be our definition of the primes having level of distribution θ.

Terence Tao (private communication) has independently proven Theorem 1.1 (with a slightly weaker bound) at much the same time. He uses a similar method; the steps are more-or-less the same but the calculations are done differently. We will indicate some of the differences in our proofs as we go along.

We see that the bound in Theorem 1.1 is quite far from the conjectural bound of approximately $m \log m$ predicted by the prime m-tuples conjecture.

Our proof naturally generalizes (but with a weaker upper bound) to many subsequences of the primes which have a level of distribution $\theta > 0$. For example, we can show corresponding results where the primes are contained in short intervals $[N, N+N^{7/12+\varepsilon}]$ for any $\varepsilon > 0$ or in an arithmetic progression modulo $q \ll (\log N)^A$. In particular, our method gives results for simultaneously prime values of linear functions, which might have specific interest. Given k distinct linear functions $L_i(n) = a_i n + b_i$ $(1 \le i \le k)$ with positive integer coefficients such that the product function $\Pi(n) = \prod_{i=1}^k L_i(n)$ has no fixed prime divisor, the method presented here shows that there are infinitely many integers n such that at least $(1/4 + o_{k\to\infty}(1)) \log k$ of the $L_i(n)$ are prime.

Theorem 1.2. *Let $m \in \mathbb{N}$. Let $r \in \mathbb{N}$ be sufficiently large depending on m, and let $\mathcal{A} = \{a_1, a_2, \dots, a_r\}$ be a set of r distinct integers. Then we have*

$$\frac{\begin{array}{l}\#\{\{h_1,\dots,h_m\} \subseteq \mathcal{A} : \textit{for infinitely many } n, \textit{ all} \\ \qquad\qquad\qquad\qquad \textit{of } n+h_1,\ \dots,\ n+h_m \textit{ are prime}\}\end{array}}{\#\{\{h_1,\dots,h_m\} \subseteq \mathcal{A}\}} \gg_m 1.$$

Therefore a positive proportion of admissible m-tuples satisfy the prime m-tuples conjecture for every m in an appropriate sense.

Theorem 1.3. *We have*

$$\liminf_n (p_{n+1} - p_n) \le 600.$$

We emphasize that the above result does not incorporate any of the technology used by Zhang to establish the existence of bounded gaps between primes. The proof is essentially elementary, relying only on the Bombieri-Vinogradov theorem. Naturally, if we assume that the primes have a higher level of distribution, then we can obtain stronger results.

Theorem 1.4. *Assume that the primes have level of distribution θ for every $\theta < 1$. Then*

$$\liminf_n (p_{n+1} - p_n) \le 12,$$

$$\liminf_n (p_{n+2} - p_n) \le 600.$$

Although the constant 12 of Theorem 1.4 appears to be optimal with our method in its current form, the constant 600 appearing in Theorem 1.3 and Theorem 1.4 is certainly not optimal. By performing further numerical calculations our method could produce a better bound, and also most of the ideas of Zhang's work (and the refinements produced by the polymath project) should be able to be combined with this method to reduce the constant further. We comment that the assumption of the Elliott-Halberstam conjecture allows us to improve the bound on Theorem 1.1 to $O(m^3e^{2m})$.

2. **An improved GPY sieve method**

We first give an explanation of the key idea behind our new approach. The basic idea of the GPY method is, for a fixed admissible set $\mathcal{H}=\{h_1,\dots,h_k\}$, to consider the sum

$$S(N,\rho)=\sum_{N\le n<2N}\Big(\sum_{i=1}^{k}\chi_{\mathbb{P}}(n+h_i)-\rho\Big)w_n. \tag{2.1}$$

Here $\chi_{\mathbb{P}}$ is the characteristic function of the primes, $\rho>0$ and w_n are nonnegative weights. If we can show that $S(N,\rho)>0$, then at least one term in the sum over n must have a positive contribution. By the nonnegativity of w_n, this means that there must be some integer $n\in[N,2N]$ such that at least $\lfloor\rho+1\rfloor$ of the $n+h_i$ are prime. (Here $\lfloor x\rfloor$ denotes the largest integer less than or equal to x.) Thus if $S(N,\rho)>0$ for all large N, there are infinitely many integers n for which at least $\lfloor\rho+1\rfloor$ of the $n+h_i$ are prime. (And so there are infinitely many bounded length intervals containing $\lfloor\rho+1\rfloor$ primes.)

The weights w_n are typically chosen to mimic Selberg sieve weights. Estimating (2.1) can be interpreted as a 'k-dimensional' sieve problem. The standard Selberg k-dimensional weights (which can be shown to be essentially optimal in other contexts) are

$$w_n=\Big(\sum_{\substack{d\mid\prod_{i=1}^{k}(n+h_i)\\ d<R}}\lambda_d\Big)^2,\qquad \lambda_d=\mu(d)(\log R/d)^k. \tag{2.2}$$

With this choice we find that we just fail to prove the existence of bounded gaps between primes if we assume the Elliott-Halberstam conjecture. The key new idea in the paper of Goldston, Pintz and Yıldırım [5] was to consider more general sieve weights of the form

$$\lambda_d=\mu(d)F(\log R/d) \tag{2.3}$$

for a suitable smooth function F. Goldston, Pintz and Yıldırım chose $F(x)=x^{k+l}$ for suitable $l\in\mathbb{N}$, which has been shown to be essentially optimal when k is large. This allows us to gain a factor of approximately 2 for large k over the previous choice of sieve weights. As a result we just fail to prove bounded gaps using the fact that the primes have exponent of distribution θ for any $\theta < 1/2$, but succeed in doing so if we assume they have level of distribution $\theta > 1/2$.

The new ingredient in our method is to consider a more general form of the sieve weights

$$w_n = \Big(\sum_{d_i|n+h_i \forall i} \lambda_{d_1,\dots,d_k}\Big)^2. \tag{2.4}$$

Using such weights with $\lambda_{d_1,\dots,d_k}$ is the key feature of our method. It allows us to improve on the previous choice of sieve weights by an arbitrarily large factor, provided that k is sufficiently large. It is the extra flexibility gained by allowing the weights to depend on the divisors of each factor individually which gives this improvement.

The idea to use such weights is not entirely new. Selberg [8, p. 245] suggested the possible use of similar weights in his work on approximations to the twin prime problem, and Goldston and Yıldırım [4] considered similar weights in earlier work on the GPY method, but with the support restricted to $d_i < R^{1/k}$ for all i.

We comment that our choice of $\lambda_{d_1,\dots,d_k}$ will look like

$$\lambda_{d_1,\dots,d_k} \approx \Big(\prod_{i=1}^{k} \mu(d_i)\Big) f(d_1,\dots,d_k) \tag{2.5}$$

for a suitable smooth function f. For our precise choice of $\lambda_{d_1,\dots,d_k}$ (given in Proposition 4.1), we find it convenient to give a slightly different form of $\lambda_{d_1,\dots,d_k}$, but weights of the form (2.5) should produce essentially the same results.

3. **Notation**

We shall view k as a fixed integer and $\mathcal{H} = \{h_1,\dots,h_k\}$ as a fixed admissible set. In particular, any constants implied by the asymptotic notation o, O or $\ll$ may depend on k and $\mathcal{H}$. We will let N denote a large integer, and all asymptotic notation should be interpreted as referring to the limit $N \to \infty$.

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, 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 Moebius function. We will let ε be a fixed positive real number, and we may assume without further comment that ε is sufficiently small at various stages of our argument. We let p_n denote the n^{th} prime and $\#\mathcal{A}$ denote the number of elements of a finite set $\mathcal{A}$. We use $\lfloor x \rfloor$ to denote the largest integer $n \le x$ and $\lceil x \rceil$ the smallest integer $n \ge x$. We let (a,b) be the greatest common divisor of integers a and b. Finally, $[a,b]$ will denote the closed interval on the real line with endpoints a and b, except for in Section 5, where it will denote the least common multiple of integers a and b instead.

4. **Outline of the proof**

We will find it convenient to choose our weights w_n to be zero unless n lies in a fixed residue class $v_0 \pmod W$, where $W = \prod_{p \le D_0} p$. This is a technical modification which removes some minor complications in dealing with the effect of small prime factors. The precise choice of D_0 is unimportant, but it will suffice to choose

$$D_0 = \log\log\log N, \tag{4.1}$$

so certainly $W \ll (\log\log N)^2$ by the prime number theorem. By the Chinese remainder theorem, we can choose v_0 such that $v_0 + h_i$ is coprime to W for each i since $\mathcal{H}$ is admissible. When $n \equiv v_0 \pmod W$, we choose our weights w_n of the form (2.4). We now wish to estimate the sums

$$S_1 = \sum_{\substack{N \le n < 2N \\ n \equiv v_0 \pmod W}} \Bigg(\sum_{d_i | n+h_i \forall i} \lambda_{d_1,\dots,d_k} \Bigg)^2, \tag{4.2}$$

$$S_2 = \sum_{\substack{N \le n < 2N \\ n \equiv v_0 \pmod W}} \Big(\sum_{i=1}^{k} \chi_{\mathbb{P}}(n+h_i) \Big) \Bigg(\sum_{d_i | n+h_i \forall i} \lambda_{d_1,\dots,d_k} \Bigg)^2. \tag{4.3}$$

We evaluate these sums using the following proposition.

Proposition 4.1. *Let the primes have exponent of distribution $\theta > 0$, and let $R = N^{\theta/2-\delta}$ for some small fixed $\delta > 0$. Let $\lambda_{d_1,\dots,d_k}$ be defined in terms of a fixed smooth function F by*

$$\lambda_{d_1,\dots,d_k} = \Big(\prod_{i=1}^{k} \mu(d_i) d_i \Big) \sum_{\substack{r_1,\dots,r_k \\ d_i | r_i \forall i \\ (r_i,W)=1 \forall i}} \frac{\mu(\prod_{i=1}^{k} r_i)^2}{\prod_{i=1}^{k} \varphi(r_i)} F\Big(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R} \Big),$$

whenever $(\prod_{i=1}^{k} d_i, W) = 1$, and let $\lambda_{d_1,\dots,d_k} = 0$ otherwise. Moreover, let F be supported on $\mathcal{R}_k = \{(x_1,\dots,x_k) \in [0,1]^k : \sum_{i=1}^{k} x_i \le 1\}$. Then we have

$$S_1 = \frac{(1+o(1))\varphi(W)^k N (\log R)^k}{W^{k+1}} I_k(F),$$

$$S_2 = \frac{(1+o(1))\varphi(W)^k N (\log R)^{k+1}}{W^{k+1} \log N} \sum_{m=1}^{k} J_k^{(m)}(F),$$

provided $I_k(F)\neq 0$ *and* $J_k^{(m)}(F)\neq 0$ *for each* m, *where*

$$I_k(F)=\int_0^1\cdots\int_0^1 F(t_1,\dots,t_k)^2dt_1\cdots dt_k,$$
$$J_k^{(m)}(F)=\int_0^1\cdots\int_0^1\left(\int_0^1 F(t_1,\dots,t_k)dt_m\right)^2 dt_1\cdots dt_{m-1}dt_{m+1}\cdots dt_k.$$

We recall that if S_2 is large compared to S_1, then using the GPY method we can show that there are infinitely many integers n such that several of the $n+h_i$ are prime. The following proposition makes this precise.

Proposition 4.2. *Let the primes have level of distribution* $\theta>0$. *Let* $\delta>0$ *and* $\mathcal{H}=\{h_1,\dots,h_k\}$ *be an admissible set. Let* $I_k(F)$ *and* $J_k^{(m)}(F)$ *be given as in Proposition* 4.1, *and let* $\mathcal{S}_k$ *denote the set of Riemann-integrable functions* $F:[0,1]^k\to\mathbb{R}$ *supported on* $\mathcal{R}_k=\{(x_1,\dots,x_k)\in[0,1]^k:\sum_{i=1}^k x_i\leq 1\}$ *with* $I_k(F)\neq 0$ *and* $J_k^{(m)}(F)\neq 0$ *for each* m. *Let*

$$M_k=\sup_{F\in\mathcal{S}_k}\frac{\sum_{m=1}^k J_k^{(m)}(F)}{I_k(F)},\qquad\qquad r_k=\left\lceil\frac{\theta M_k}{2}\right\rceil.$$

Then there are infinitely many integers n *such that at least* r_k *of the* $n+h_i$ $(1\leq i\leq k)$ *are prime. In particular,* $\liminf_n(p_{n+r_k-1}-p_n)\leq\max_{1\leq i,j\leq k}(h_i-h_j)$.

Proof of Proposition 4.2. We let $S=S_2-\rho S_1$, and we recall from Section 2 that if we can show $S>0$ for all large N, then there are infinitely many integers n such that at least $\lfloor\rho+1\rfloor$ of the $n+h_i$ are prime.

We put $R=N^{\theta/2-\delta}$ for a small $\delta>0$. By the definition of M_k, we can choose $F_0\in\mathcal{S}_k$ such that $\sum_{m=1}^k J_k^{(m)}(F_0)>(M_k-\delta)I_k(F_0)>0$. Since F_0 is Riemann-integrable, there is a smooth function F_1 such that $\sum_{m=1}^k J_k^{(m)}(F_1)>(M_k-2\delta)I_k(F_1)>0$. Using Proposition 4.1, we can then choose $\lambda_{d_1,\dots,d_k}$ such that

$$\begin{aligned}(4.4)\qquad S&=\frac{\varphi(W)^kN(\log R)^k}{W^{k+1}}\Big(\frac{\log R}{\log N}\sum_{j=1}^k J_k^{(m)}(F_1)-\rho I_k(F_1)+o(1)\Big)\\ &\geq\frac{\varphi(W)^kN(\log R)^kI_k(F_1)}{W^{k+1}}\Big(\Big(\frac{\theta}{2}-\delta\Big)\Big(M_k-2\delta\Big)-\rho+o(1)\Big).\end{aligned}$$

If $\rho=\theta M_k/2-\varepsilon$ then, by choosing δ suitably small (depending on ε), we see that $S>0$ for all large N. Thus there are infinitely many integers n for which at least $\lfloor\rho+1\rfloor$ of the $n+h_i$ are prime. Since $\lfloor\rho+1\rfloor=\lceil\theta M_k/2\rceil$ if ε is suitably small, we obtain Proposition 4.2. □

Thus, if the primes have a fixed level of distribution θ, to show the existence of many of the $n+h_i$ being prime for infinitely many $n\in\mathbb{N}$ we only require a suitable lower bound for M_k. The following proposition establishes such a bound for different values of k.

Proposition 4.3. *Let $k \in \mathbb{N}$, and let M_k be as given by Proposition* 4.2. *Then*

(1) *We have $M_5 > 2$.*
(2) *We have $M_{105} > 4$.*
(3) *If k is sufficiently large, we have $M_k > \log k - 2 \log\log k - 2$.*

We now prove Theorems 1.1, 1.2, 1.3 and 1.4 from Propositions 4.2 and 4.3.

First we consider Theorem 1.3. We take $k = 105$. By Proposition 4.3, we have $M_{105} > 4$. By the Bombieri-Vinogradov theorem, the primes have level of distribution $\theta = 1/2 - \varepsilon$ for every $\varepsilon > 0$. Thus, if we take ε sufficiently small, we have $\theta M_{105}/2 > 1$. Therefore, by Proposition 4.2, we have

$$\liminf (p_{n+1} - p_n) \le \max_{1 \le i,j \le 105} (h_i - h_j)$$

for any admissible set $\mathcal{H} = \{h_1, \dots, h_{105}\}$. By computations performed by Thomas Engelsma (unpublished), we can choose[2] $\mathcal{H}$ such that $0 \le h_1 < \cdots < h_{105}$ and $h_{105} - h_1 = 600$. This gives Theorem 1.3.

If we assume the Elliott-Halberstam conjecture then the primes have level of distribution $\theta = 1 - \varepsilon$. First we take $k = 105$ and see that $\theta M_{105}/2 > 2$ for ε sufficiently small (since $M_{105} > 4$). Therefore, by Proposition 4.2, $\liminf_n (p_{n+2} - p_n) \le \max_{1 \le i,j \le 105} (h_i - h_j)$. Thus, choosing the same admissible set $\mathcal{H}$ as above, we see $\liminf_n (p_{n+2} - p_n) \le 600$ under the Elliott-Halberstam conjecture.

Next we take $k = 5$ and $\mathcal{H} = \{0, 2, 6, 8, 12\}$, with $\theta = 1 - \varepsilon$ again. By Proposition 4.3 we have $M_5 > 2$, and so $\theta M_5/2 > 1$ for ε sufficiently small. Thus, by Proposition 4.2, $\liminf_n (p_{n+1} - p_n) \le 12$ under the Elliott-Halberstam conjecture. This completes the proof of Theorem 1.4.

Finally, we consider the case when k is large. For the rest of this section, any constants implied by asymptotic notation will be independent of k. By the Bombieri-Vinogradov theorem, we can take $\theta = 1/2 - \varepsilon$. Thus, by

[2]Explicitly, we can take $\mathcal{H} = \{0, 10, 12, 24, 28, 30, 34, 42, 48, 52, 54, 64, 70, 72, 78, 82, 90, 94, 100, 112, 114, 118, 120, 124, 132, 138, 148, 154, 168, 174, 178, 180, 184, 190, 192, 202, 204, 208, 220, 222, 232, 234, 250, 252, 258, 262, 264, 268, 280, 288, 294, 300, 310, 322, 324, 328, 330, 334, 342, 352, 358, 360, 364, 372, 378, 384, 390, 394, 400, 402, 408, 412, 418, 420, 430, 432, 442, 444, 450, 454, 462, 468, 472, 478, 484, 490, 492, 498, 504, 510, 528, 532, 534, 538, 544, 558, 562, 570, 574, 580, 582, 588, 594, 598, 600\}$. This set was obtained from the website http://math.mit.edu/~primegaps/ maintained by Andrew Sutherland.

Proposition 4.3, we have for k sufficiently large

$$\frac{\theta M_k}{2} \geq \Big(\frac{1}{4} - \frac{\varepsilon}{2}\Big)\Big(\log k - 2\log\log k - 2\Big). \tag{4.5}$$

We choose $\varepsilon = 1/k$ and see that $\theta M_k/2 > m$ if $k \geq Cm^2e^{4m}$ for some absolute constant C (independent of m and k). Thus, for any admissible set $\mathcal{H} = \{h_1, \dots, h_k\}$ with $k \geq Cm^2e^{4m}$, at least $m+1$ of the $n+h_i$ must be prime for infinitely many integers n. We can choose our set $\mathcal{H}$ to be the set $\{p_{\pi(k)+1}, \dots, p_{\pi(k)+k}\}$ of the first k primes which are greater than k. This is admissible, since no element is a multiple of a prime less than k (and there are k elements, so it cannot cover all residue classes modulo any prime greater than k.) This set has diameter $p_{\pi(k)+k} - p_{\pi(k)+1} \ll k\log k$. Thus $\liminf_n(p_{n+m} - p_n) \ll k\log k \ll m^3e^{4m}$ if we take $k = \lceil Cm^2e^{4m}\rceil$. This gives Theorem 1.1.

We can now establish Theorem 1.2 by a simple counting argument. Given m, we let $k = \lceil Cm^2e^{4m}\rceil$ as above. Therefore if $\{h_1, \dots, h_k\}$ is admissible, then there exists a subset $\{h_1', \dots, h_m'\} \subseteq \{h_1, \dots, h_k\}$ with the property that there are infinitely many integers n for which all of the $n + h_i'$ are prime ($1 \leq i \leq m$).

We let $\mathcal{A}_2$ denote the set formed by starting with the given set $\mathcal{A} = \{a_1, \dots, a_r\}$, and for each prime $p \leq k$, in turn removing all elements of the residue class modulo p which contains the fewest integers. We see that $\#\mathcal{A}_2 \geq r\prod_{p\leq k}(1-1/p) \gg_m r$. Moreover, any subset of $\mathcal{A}_2$ of size k must be admissible, since it cannot cover all residue classes modulo p for any prime $p \leq k$. We let $s = \#\mathcal{A}_2$, and since r is taken sufficiently large in terms of m, we may assume that $s > k$.

We see there are $\binom{s}{k}$ sets $\mathcal{H} \subseteq \mathcal{A}_2$ of size k. Each of these is admissible, and so contains at least one subset $\{h_1', \dots, h_m'\} \subseteq \mathcal{A}_2$ which satisfies the prime m-tuples conjecture. Any admissible set $\mathcal{B} \subseteq \mathcal{A}_2$ of size m is contained in $\binom{s-m}{k-m}$ sets $\mathcal{H} \subseteq \mathcal{A}_2$ of size k. Thus there are at least $\binom{s}{k}\binom{s-m}{k-m}^{-1} \gg_m s^m \gg_m r^m$ admissible sets $\mathcal{B} \subseteq \mathcal{A}_2$ of size m which satisfy the prime m-tuples conjecture. Since there are $\binom{r}{m} \leq r^m$ sets $\{h_1, \dots, h_m\} \subseteq \mathcal{A}$, Theorem 1.2 holds.

We are left to establish Propositions 4.1 and 4.3.

5. **Selberg sieve manipulations**

In this section we perform initial manipulations towards establishing Proposition 4.1. These arguments are multidimensional generalizations of the sieve arguments of [3]. In particular, our approach is based on the elementary combinatorial ideas of Selberg. The aim is to introduce a change of variables to rewrite our sums S_1 and S_2 in a simpler form.

Throughout the rest of the paper we assume that the primes have a fixed level of distribution θ, and $R = N^{\theta/2-\delta}$. We restrict the support of $\lambda_{d_1,\dots,d_k}$

to tuples for which the product $d=\prod_{i=1}^k d_i$ is less than R and also satisfies $(d,W)=1$ and $\mu(d)^2=1$. We note that the condition $\mu(d)^2=1$ implies that $(d_i,d_j)=1$ for all $i\neq j$.

Lemma 5.1. *Let*

$$y_{r_1,\dots,r_k}=\Big(\prod_{i=1}^k\mu(r_i)\varphi(r_i)\Big)\sum_{\substack{d_1,\dots,d_k\\ r_i|d_i\forall i}}\frac{\lambda_{d_1,\dots,d_k}}{\prod_{i=1}^k d_i}.$$

Let $y_{\max}=\sup_{r_1,\dots,r_k}|y_{r_1,\dots,r_k}|$. *Then*

$$S_1=\frac{N}{W}\sum_{r_1,\dots,r_k}\frac{y^2_{r_1,\dots,r_k}}{\prod_{i=1}^k\varphi(r_i)}+O\Big(\frac{y^2_{\max}\varphi(W)^kN(\log R)^k}{W^{k+1}D_0}\Big).$$

Proof. We expand out the square and swap the order of summation to give

(5.1)

$$S_1=\sum_{\substack{N\le n<2N\\ n\equiv v_0\pmod W}}\Big(\sum_{d_i|n+h_i\forall i}\lambda_{d_1,\dots,d_k}\Big)^2=\sum_{\substack{d_1,\dots,d_k\\ e_1,\dots,e_k}}\lambda_{d_1,\dots,d_k}\lambda_{e_1,\dots,e_k}\sum_{\substack{N\le n<2N\\ n\equiv v_0\pmod W\\ [d_i,e_i]|n+h_i\forall i}}1.$$

We recall that here, and throughout this section, we are using $[a,b]$ to denote the least common multiple of a and b.

By the Chinese remainder theorem, the inner sum can be written as a sum over a single residue class modulo $q=W\prod_{i=1}^k[d_i,e_i]$, provided that the integers $W,[d_1,e_1],\dots,[d_k,e_k]$ are pairwise coprime. In this case the inner sum is $N/q+O(1)$. If the integers are not pairwise coprime, then the inner sum is empty. This gives

$$S_1=\frac{N}{W}\sideset{}{'}\sum_{\substack{d_1,\dots,d_k\\ e_1,\dots,e_k}}\frac{\lambda_{d_1,\dots,d_k}\lambda_{e_1,\dots,e_k}}{\prod_{i=1}^k[d_i,e_i]}+O\Big(\sideset{}{'}\sum_{\substack{d_1,\dots,d_k\\ e_1,\dots,e_k}}|\lambda_{d_1,\dots,d_k}\lambda_{e_1,\dots,e_k}|\Big),\tag{5.2}$$

where $\sum'$ is used to denote the restriction that we require $W,[d_1,e_1],\dots,[d_k,e_k]$ to be pairwise coprime. To ease notation we will put $\lambda_{\max}=\sup_{d_1,\dots,d_k}|\lambda_{d_1,\dots,d_k}|$. We now see that since $\lambda_{d_1,\dots,d_k}$ is nonzero only when $\prod_{i=1}^k d_i<R$, the error term contributes

$$\ll\lambda^2_{\max}\Big(\sum_{d<R}\tau_k(d)\Big)^2\ll\lambda^2_{\max}R^2(\log R)^{2k},\tag{5.3}$$

which will be negligible.

In the main sum we wish to remove the dependencies between the d_i and the e_j variables. We use the identity

$$\frac{1}{[d_i,e_i]}=\frac{1}{d_ie_i}\sum_{u_i|d_i,e_i}\varphi(u_i)\tag{5.4}$$

to rewrite the main term as

$$\frac{N}{W}\sum_{u_1,\dots,u_k}\Big(\prod_{i=1}^k\varphi(u_i)\Big)\sideset{}{'}\sum_{\substack{d_1,\dots,d_k\\ e_1,\dots,e_k\\ u_i|d_i,e_i\forall i}}\frac{\lambda_{d_1,\dots,d_k}\lambda_{e_1,\dots,e_k}}{(\prod_{i=1}^k d_i)(\prod_{i=1}^k e_i)}. \tag{5.5}$$

We recall that $\lambda_{d_1,\dots,d_k}$ is supported on integers $d_1,\dots,d_k$ with $(d_i,W)=1$ for each i and $(d_i,d_j)=1$ for all $i\neq j$. Thus we may drop the requirement that W is coprime to each of the $[d_i,e_i]$ from the summation, since these terms have no contribution. Similarly, we may drop the requirement that the d_i variables are all pairwise coprime and the requirement that the e_i variables are all pairwise coprime. Thus the only remaining restriction coming from the pairwise coprimality of $W,[d_1,e_1],\dots,[d_k,e_k]$ is that $(d_i,e_j)=1$ for all $i\neq j$.

We can remove the requirement that $(d_i,e_j)=1$ by multiplying our expression by $\sum_{s_{i,j}|d_i,e_j}\mu(s_{i,j})$. We do this for all i,j with $i\neq j$. This transforms the main term to

(5.6)

$$\frac{N}{W}\sum_{u_1,\dots,u_k}\Big(\prod_{i=1}^k\varphi(u_i)\Big)\sum_{s_{1,2},\dots,s_{k,k-1}}\Big(\prod_{\substack{1\le i,j\le k\\ i\neq j}}\mu(s_{i,j})\Big)\sum_{\substack{d_1,\dots,d_k\\ e_1,\dots,e_k\\ u_i|d_i,e_i\forall i\\ s_{i,j}|d_i,e_j\forall i\neq j}}\frac{\lambda_{d_1,\dots,d_k}\lambda_{e_1,\dots,e_k}}{(\prod_{i=1}^k d_i)(\prod_{i=1}^k e_i)}.$$

We can restrict the $s_{i,j}$ to be coprime to u_i and u_j, because terms with $s_{i,j}$ not coprime to u_i or u_j make no contribution to our sum. This is because $\lambda_{d_1,\dots,d_k}=0$ unless $(d_i,d_j)=1$. Similarly we can further restrict our sum so that $s_{i,j}$ is coprime to $s_{i,a}$ and $s_{b,j}$ for all $a\neq j$ and $b\neq i$. We denote the summation over $s_{1,2},\dots,s_{k,k-1}$ with these restrictions by $\sum^*$.

We now introduce a change of variables to make the estimation of the sum more straightforward. We let

$$y_{r_1,\dots,r_k}=\Big(\prod_{i=1}^k\mu(r_i)\varphi(r_i)\Big)\sum_{\substack{d_1,\dots,d_k\\ r_i|d_i\forall i}}\frac{\lambda_{d_1,\dots,d_k}}{\prod_{i=1}^k d_i}. \tag{5.7}$$

This change is invertible. For $d_1,\dots,d_k$ with $\prod_{i=1}^k d_i$ square-free, we find that

$$\begin{aligned}\sum_{\substack{r_1,\dots,r_k\\ d_i|r_i\forall i}}\frac{y_{r_1,\dots,r_k}}{\prod_{i=1}^k\varphi(r_i)}&=\sum_{\substack{r_1,\dots,r_k\\ d_i|r_i\forall i}}\Big(\prod_{i=1}^k\mu(r_i)\Big)\sum_{\substack{e_1,\dots,e_k\\ r_i|e_i\forall i}}\frac{\lambda_{e_1,\dots,e_k}}{\prod_{i=1}^k e_i}\\ &=\sum_{e_1,\dots,e_k}\frac{\lambda_{e_1,\dots,e_k}}{\prod_{i=1}^k e_i}\sum_{\substack{r_1,\dots,r_k\\ d_i|r_i\forall i\\ r_i|e_i\forall i}}\prod_{i=1}^k\mu(r_i)=\frac{\lambda_{d_1,\dots,d_k}}{\prod_{i=1}^k\mu_i(d_i)d_i}.\end{aligned} \tag{5.8}$$

Thus any choice of $y_{r_1,\dots,r_k}$ supported on $r_1,\dots,r_k$, with the product $r=\prod_{i=1}^k r_i$ square-free and satisfying $r<R$ and $(r,W)=1$, will give a suitable choice of $\lambda_{d_1,\dots,d_k}$. We let $y_{\max}=\sup_{r_1,\dots,r_k}|y_{r_1,\dots,r_k}|$. Now, since $d/\varphi(d)=\sum_{e|d}1/\varphi(e)$ for square-free d, we find by taking $r'=\prod_{i=1}^k r_i/d_i$ that

$$
\begin{aligned}
\text{(5.9)}\qquad \lambda_{\max} &\le \sup_{\substack{d_1,\dots,d_k\\ \prod_{i=1}^k d_i \text{ square-free}}} y_{\max}\Big(\prod_{i=1}^k d_i\Big) \sum_{\substack{r_1,\dots,r_k\\ d_i|r_i\forall i\\ \prod_{i=1}^k r_i<R\\ \prod_{i=1}^k r_i \text{ square-free}}} \Big(\prod_{i=1}^k \frac{\mu(r_i)^2}{\varphi(r_i)}\Big)\\
&\le y_{\max} \sup_{\substack{d_1,\dots,d_k\\ \prod_{i=1}^k d_i \text{ square-free}}} \Big(\prod_{i=1}^k \frac{d_i}{\varphi(d_i)}\Big) \sum_{\substack{r'<R/\prod_{i=1}^k d_i\\ (r',\prod_{i=1}^k d_i)=1}} \frac{\mu(r')^2\tau_k(r')}{\varphi(r')}\\
&\le y_{\max} \sup_{d_1,\dots,d_k} \sum_{d|\prod_{i=1}^k d_i} \frac{\mu(d)^2}{\varphi(d)} \sum_{\substack{r'<R/\prod_{i=1}^k d_i\\ (r',\prod_{i=1}^k d_i)=1}} \frac{\mu(r')^2\tau_k(r')}{\varphi(r')}\\
&\le y_{\max}\sum_{u<R}\frac{\mu(u)^2\tau_k(u)}{\varphi(u)} \ll y_{\max}(\log R)^k.
\end{aligned}
$$

In the last line we have taken $u=dr'$ and used the fact $\tau_k(dr')\ge\tau_k(r')$. Hence the error term $O(\lambda_{\max}^2R^2(\log N)^{2k})$ is of size $O(y_{\max}^2R^2(\log N)^{4k})$.

Substituting our change of variables (5.7) into the main term (5.6), and using the above estimate for the error term, we obtain

$$
\begin{aligned}
\text{(5.10)}\qquad S_1 = \frac{N}{W}&\sum_{u_1,\dots,u_k}\Big(\prod_{i=1}^k\varphi(u_i)\Big)\\
&\times \sideset{}{^*}\sum_{s_{1,2},\dots,s_{k,k-1}}\Big(\prod_{\substack{1\le i,j\le k\\ i\ne j}}\mu(s_{i,j})\Big)\Big(\prod_{i=1}^k\frac{\mu(a_i)\mu(b_i)}{\varphi(a_i)\varphi(b_i)}\Big)y_{a_1,\dots,a_k}y_{b_1,\dots,b_k}\\
&+O\Big(y_{\max}^2R^2(\log R)^{4k}\Big),
\end{aligned}
$$

where $a_j=u_j\prod_{i\ne j}s_{j,i}$ and $b_j=u_j\prod_{i\ne j}s_{i,j}$. In these expressions we have used the fact that we have restricted $s_{i,j}$ to be coprime to the other terms in the expression for a_i and b_j. For the same reason we may rewrite $\mu(a_j)$ as $\mu(u_j)\prod_{i\ne j}\mu(s_{i,j})$, and similarly for $\varphi(a_j)$, $\mu(b_j)$ and $\varphi(b_j)$. This gives us

$$
\begin{aligned}
\text{(5.11)}\quad S_1 = \frac{N}{W}\sum_{u_1,\dots,u_k}\Big(\prod_{i=1}^k\frac{\mu(u_i)^2}{\varphi(u_i)}\Big)\sideset{}{^*}\sum_{s_{1,2},\dots,s_{k,k-1}}\Big(\prod_{\substack{1\le i,j\le k\\ i\ne j}}\frac{\mu(s_{i,j})}{\varphi(s_{i,j})^2}\Big)y_{a_1,\dots,a_k}y_{b_1,\dots,b_k}&\\
+O\Big(y_{\max}^2R^2(\log R)^{4k}\Big).&
\end{aligned}
$$

We see that there is no contribution from $s_{i,j}$ with $(s_{i,j}, W) \neq 1$ because of the restricted support of y. Thus we only need to consider $s_{i,j} = 1$ or $s_{i,j} > D_0$. The contribution when $s_{i,j} > D_0$ is

$$\begin{aligned}(5.12)\qquad &\ll \frac{y_{\max}^2 N}{W}\Big(\sum_{\substack{u<R\\(u,W)=1}} \frac{\mu(u)^2}{\varphi(u)}\Big)^k\Big(\sum_{s_{i,j}>D_0}\frac{\mu(s_{i,j})^2}{\varphi(s_{i,j})^2}\Big)\Big(\sum_{s\geq 1}\frac{\mu(s)^2}{\varphi(s)^2}\Big)^{k^2-k-1}\\ &\ll \frac{y_{\max}^2\varphi(W)^k N(\log R)^k}{W^{k+1}D_0}.\end{aligned}$$

Thus we may restrict our attention to the case when $s_{i,j} = 1$ for all $i \neq j$. This gives

(5.13)
$$S_1 = \frac{N}{W}\sum_{u_1,\dots,u_k}\frac{y_{u_1,\dots,u_k}^2}{\prod_{i=1}^k \varphi(u_i)} + O\left(\frac{y_{\max}^2\varphi(W)^k N(\log R)^k}{W^{k+1}D_0} + y_{\max}^2 R^2(\log R)^{4k}\right).$$

We recall that $R^2 = N^{\theta-2\delta} \leq N^{1-2\delta}$ and $W \ll N^{\delta}$, and so the first error term dominates. This gives the result. $\square$

We now consider S_2. We write $S_2 = \sum_{m=1}^k S_2^{(m)}$, where

$$(5.14)\qquad S_2^{(m)} = \sum_{\substack{N\leq n<2N\\ n\equiv v_0 \pmod W}} \chi_{\mathbb{P}}(n+h_m)\Big(\sum_{\substack{d_1,\dots,d_k\\ d_i|n+h_i\forall i}}\lambda_{d_1,\dots,d_k}\Big)^2.$$

We now estimate $S_2^{(m)}$ in a similar way to our treatment of S_1.

Lemma 5.2. *Let*

$$y_{r_1,\dots,r_k}^{(m)} = \Big(\prod_{i=1}^k \mu(r_i)g(r_i)\Big)\sum_{\substack{d_1,\dots,d_k\\ r_i|d_i\forall i\\ d_m=1}}\frac{\lambda_{d_1,\dots,d_k}}{\prod_{i=1}^k\varphi(d_i)},$$

where g is the totally multiplicative function defined on primes by $g(p) = p-2$. Let $y_{\max}^{(m)} = \sup_{r_1,\dots,r_k}|y_{r_1,\dots,r_k}^{(m)}|$. Then for any fixed $A > 0$, we have

$$\begin{aligned}S_2^{(m)} = &\frac{N}{\varphi(W)\log N}\sum_{r_1,\dots,r_k}\frac{(y_{r_1,\dots,r_k}^{(m)})^2}{\prod_{i=1}^k g(r_i)} + O\Big(\frac{(y_{\max}^{(m)})^2\varphi(W)^{k-2}N(\log N)^{k-2}}{W^{k-1}D_0}\Big)\\ &+ O\Big(\frac{y_{\max}^2 N}{(\log N)^A}\Big).\end{aligned}$$

Proof. We first expand out the square and swap the order of summation to give

$$(5.15)\qquad S_2^{(m)} = \sum_{\substack{d_1,\dots,d_k\\ e_1,\dots,e_k}}\lambda_{d_1,\dots,d_k}\lambda_{e_1,\dots,e_k}\sum_{\substack{N\leq n<2N\\ n\equiv v_0 \pmod W\\ [d_i,e_i]|n+h_i\forall i}}\chi_{\mathbb{P}}(n+h_m).$$

As with S_1, the inner sum can be written as a sum over a single residue class modulo $q = W\prod_{i=1}^k [d_i, e_i]$, provided that $W, [d_1, e_1], \dots, [d_k, e_k]$ are pairwise coprime. The integer $n+h_m$ will lie in a residue class coprime to the modulus if and only if $d_m = e_m = 1$. In this case the inner sum will contribute $X_N/\varphi(q) + O(E(N, q))$, where

$$E(N, q) = 1 + \sup_{(a,q)=1} \Big| \sum_{\substack{N \le n < 2N \\ n \equiv a \pmod q}} \chi_{\mathbb{P}}(n) - \frac{1}{\varphi(q)} \sum_{N \le n < 2N} \chi_{\mathbb{P}}(n) \Big|, \tag{5.16}$$

$$X_N = \sum_{N \le n < 2N} \chi_{\mathbb{P}}(n). \tag{5.17}$$

If either one pair of $W, [d_1, e_1], \dots, [d_k, e_k]$ share a common factor, or if either d_m or e_m are not 1, then the contribution of the inner sum is zero. Thus we obtain

$$S_2^{(m)} = \frac{X_N}{\varphi(W)} \sideset{}{'}\sum_{\substack{d_1,\dots,d_k \\ e_1,\dots,e_k \\ e_m = d_m = 1}} \frac{\lambda_{d_1,\dots,d_k} \lambda_{e_1,\dots,e_k}}{\prod_{i=1}^k \varphi([d_i, e_i])} + O\Big(\sum_{\substack{d_1,\dots,d_k \\ e_1,\dots,e_k}} |\lambda_{d_1,\dots,d_k} \lambda_{e_1,\dots,e_k}| E(N, q) \Big), \tag{5.18}$$

where we have written $q = W\prod_{i=1}^k [d_i, e_i]$.

We first deal with the contribution from the error terms. From the support of $\lambda_{d_1,\dots,d_k}$, we see that we only need to consider square-free q with $q < R^2 W$. Given a square-free integer r, there are at most $\tau_{3k}(r)$ choices of $d_1, \dots, d_k, e_1, \dots, e_k$ for which $W\prod_{i=1}^k [d_i, e_i] = r$. We also recall from (5.9) that $\lambda_{\max} \ll y_{\max}(\log R)^k$. Thus the error term contributes

$$\ll y_{\max}^2 (\log R)^{2k} \sum_{r < R^2 W} \mu(r)^2 \tau_{3k}(r) E(N, r). \tag{5.19}$$

By Cauchy-Schwarz, the trivial bound $E(N, q) \ll N/\varphi(q)$, and our hypothesis that the primes have level of distribution θ, this contributes for any fixed $A > 0$

$$\begin{aligned} &\ll y_{\max}^2 (\log R)^{2k} \Big(\sum_{r < R^2 W} \mu(r)^2 \tau_{3k}^2(r) \frac{N}{\varphi(r)} \Big)^{1/2} \Big(\sum_{r < R^2 W} \mu(r)^2 E(N, r) \Big)^{1/2} \\ &\ll \frac{y_{\max}^2 N}{(\log N)^A}. \end{aligned} \tag{5.20}$$

We now concentrate on the main sum. As in the treatment of S_1 in the proof of Lemma 5.1, we rewrite the conditions $(d_i, e_j) = 1$ by multiplying our expression by $\sum_{s_{i,j} | d_i, e_j} \mu(s_{i,j})$. Again we may restrict $s_{i,j}$ to be coprime to u_i, u_j, $s_{i,a}$ and $s_{b,j}$ for all $a \neq j$ and $b \neq i$. We denote the summation subject to these restrictions by $\sum^*$. We also split the $\varphi([d_i, e_i])$ terms by using the equation

(valid for square-free d_i, e_i)

$$\frac{1}{\varphi([d_i,e_i])}=\frac{1}{\varphi(d_i)\varphi(e_i)}\sum_{u_i|d_i,e_i}g(u_i), \tag{5.21}$$

where g is the totally multiplicative function defined on primes by $g(p)=p-2$. This gives us a main term of

$$\frac{X_N}{\varphi(W)}\sum_{u_1,\dots,u_k}\Big(\prod_{i=1}^k g(u_i)\Big)\sideset{}{^*}\sum_{s_{1,2},\dots,s_{k,k-1}}\Big(\prod_{\substack{1\le i,j\le k\\ i\neq j}}\mu(s_{i,j})\Big)\sum_{\substack{d_1,\dots,d_k\\ e_1,\dots,e_k\\ u_i|d_i,e_i\forall i\\ s_{i,j}|d_i,e_j\forall i\neq j\\ d_m=e_m=1}}\frac{\lambda_{d_1,\dots,d_k}\lambda_{e_1,\dots,e_k}}{\prod_{i=1}^k\varphi(d_i)\varphi(e_i)}. \tag{5.22}$$

We have now separated the dependencies between the e and d variables, so again we make a substitution. We let

$$y^{(m)}_{r_1,\dots,r_k}=\Big(\prod_{i=1}^k\mu(r_i)g(r_i)\Big)\sum_{\substack{d_1,\dots,d_k\\ r_i|d_i\forall i\\ d_m=1}}\frac{\lambda_{d_1,\dots,d_k}}{\prod_{i=1}^k\varphi(d_i)}. \tag{5.23}$$

We note $y^{(m)}_{r_1,\dots,r_k}=0$ unless $r_m=1$. Substituting this into (5.22), we obtain a main term of

$$\frac{X_N}{\varphi(W)}\sum_{u_1,\dots,u_k}\Big(\prod_{i=1}^k\frac{\mu(u_i)^2}{g(u_i)}\Big)\sideset{}{^*}\sum_{s_{1,2},\dots,s_{k,k-1}}\Big(\prod_{\substack{1\le i,j\le k\\ i\neq j}}\frac{\mu(s_{i,j})}{g(s_{i,j})^2}\Big)y^{(m)}_{a_1,\dots,a_k}y^{(m)}_{b_1,\dots,b_k}, \tag{5.24}$$

where $a_j=u_j\prod_{i\neq j}s_{j,i}$ and $b_j=u_j\prod_{i\neq j}s_{i,j}$ for each $1\le j\le k$. As before, we have replaced $\mu(a_j)$ with $\mu(u_j)\prod_{i\neq j}\mu(s_{j,i})$ (and similarly for $g(a_j)$, $\mu(b_j)$ and $g(b_j)$). This is valid since terms with a_j or b_j not square-free make no contribution.

We see the contribution from $s_{i,j}\neq 1$ is of size

$$\begin{aligned}&\ll\frac{(y^{(m)}_{\max})^2N}{\varphi(W)\log N}\Big(\sum_{\substack{u<R\\ (u,W)=1}}\frac{\mu(u)^2}{g(u)}\Big)^{k-1}\Big(\sum_s\frac{\mu(s)^2}{g(s)^2}\Big)^{k(k-1)-1}\sum_{s_{i,j}>D_0}\frac{\mu(s_{i,j})^2}{g(s_{i,j})^2}\\ &\ll\frac{(y^{(m)}_{\max})^2\varphi(W)^{k-2}N(\log R)^{k-1}}{W^{k-1}D_0\log N}.\end{aligned} \tag{5.25}$$

Thus we find that

$$(5.26)\quad S_2^{(m)}=\frac{X_N}{\varphi(W)}\sum_{u_1,\dots,u_k}\frac{(y^{(m)}_{u_1,\dots,u_k})^2}{\prod_{i=1}^k g(u_i)}+O\Big(\frac{(y^{(m)}_{\max})^2\varphi(W)^{k-2}N(\log R)^{k-2}}{D_0W^{k-1}}\Big)+O\Big(\frac{y_{\max}^2N}{(\log N)^A}\Big).$$

Finally, by the prime number theorem, $X_N=N/\log N+O(N/(\log N)^2)$. This error term contributes

$$(5.27)\quad \ll\frac{(y^{(m)}_{\max})^2N}{\varphi(W)(\log N)^2}\Big(\sum_{\substack{u<R\\(u,W)=1}}\frac{\mu(u)^2}{g(u)}\Big)^{k-1}\ll\frac{(y^{(m)}_{\max})^2\varphi(W)^{k-2}N(\log R)^{k-3}}{W^{k-1}},$$

which can be absorbed into the first error term of (5.26). This completes the proof. □

Remark. In our proof of Lemma 5.2 we only really require $\lambda_{d_1,\dots,d_k}$ to be supported on $d_1,\dots,d_k$ satisfying $\prod_{i\ne j}d_i<R$ for all j instead of $\prod_{i=1}^k d_i<R$. For $k\ge 3$, the numerical benefit of this extension is small, and so we do not consider it further.

Remark. As our result relies on the Bombieri-Vinogradov theorem, the implied constant in the error term is not effectively computable. However, if we restrict the $\lambda_{d_1,\dots,d_k}$ to be supported on d_i which are coprime to the largest prime factor of a possible exceptional modulus of a primitive character, then we can make this error term (and all others in this paper) effective at the cost of a negligible error.

We now relate our new variables $y^{(m)}_{r_1,\dots,r_k}$ to the $y_{r_1,\dots,r_k}$ variables from S_1.

LEMMA 5.3. *If $r_m=1$, then*

$$y^{(m)}_{r_1,\dots,r_k}=\sum_{a_m}\frac{y_{r_1,\dots,r_{m-1},a_m,r_{m+1},\dots,r_k}}{\varphi(a_m)}+O\Big(\frac{y_{\max}\varphi(W)\log R}{WD_0}\Big).$$

Proof. We assume throughout the proof that $r_m=1$. We first substitute our expression (5.8) into the definition (5.23). This gives

$$(5.28)\qquad y^{(m)}_{r_1,\dots,r_k}=\Big(\prod_{i=1}^k\mu(r_i)g(r_i)\Big)\sum_{\substack{d_1,\dots,d_k\\ r_i|d_i\forall i\\ d_m=1}}\Big(\prod_{i=1}^k\frac{\mu(d_i)d_i}{\varphi(d_i)}\Big)\sum_{\substack{a_1,\dots,a_k\\ d_i|a_i\forall i}}\frac{y_{a_1,\dots,a_k}}{\prod_{i=1}^k\varphi(a_i)}.$$

We swap the summation of the d and a variables to give

$$y^{(m)}_{r_1,\dots,r_k}=\Big(\prod_{i=1}^k\mu(r_i)g(r_i)\Big)\sum_{\substack{a_1,\dots,a_k\\ r_i|a_i\forall i}}\frac{y_{a_1,\dots,a_k}}{\prod_{i=1}^k\varphi(a_i)}\sum_{\substack{d_1,\dots,d_k\\ d_i|a_i,r_i|d_i\forall i\\ d_m=1}}\prod_{i=1}^k\frac{\mu(d_i)d_i}{\varphi(d_i)}. \tag{5.29}$$

We can now evaluate the sum over $d_1,\dots,d_k$ explicitly. This gives

$$y^{(m)}_{r_1,\dots,r_k}=\Big(\prod_{i=1}^k\mu(r_i)g(r_i)\Big)\sum_{\substack{a_1,\dots,a_k\\ r_i|a_i\forall i}}\frac{y_{a_1,\dots,a_k}}{\prod_{i=1}^k\varphi(a_i)}\prod_{i\neq m}\frac{\mu(a_i)r_i}{\varphi(a_i)}. \tag{5.30}$$

We see that from the support of $y_{a_1,\dots,a_k}$ that we may restrict the summation over a_j to $(a_j,W)=1$. Thus either $a_j=r_j$ or $a_j>D_0r_j$. For $j\neq m$, the total contribution from $a_j\neq r_j$ is

(5.31)

$$\ll y_{\max}\Big(\prod_{i=1}^k g(r_i)r_i\Big)\Big(\sum_{\substack{a_j>D_0r_j\\ r_j|a_j}}\frac{\mu(a_j)^2}{\varphi(a_j)^2}\Big)\Big(\sum_{\substack{a_m<R\\ (a_m,W)=1}}\frac{\mu(a_m)^2}{\varphi(a_m)}\Big)\prod_{\substack{1\le i\le k\\ i\neq j,m}}\Big(\sum_{r_i|a_i}\frac{\mu(a_i)^2}{\varphi(a_i)^2}\Big)$$

$$\ll\Big(\prod_{i=1}^k\frac{g(r_i)r_i}{\varphi(r_i)^2}\Big)\frac{y_{\max}\varphi(W)\log R}{WD_0}\ll\frac{y_{\max}\varphi(W)\log R}{WD_0}.$$

Thus we find that the main contribution is when $a_j=r_j$ for all $j\neq m$. We have

(5.32)

$$y^{(m)}_{r_1,\dots,r_k}=\Big(\prod_{i=1}^k\frac{g(r_i)r_i}{\varphi(r_i)^2}\Big)\sum_{a_m}\frac{y_{r_1,\dots,r_{m-1},a_m,r_{m+1},\dots,r_k}}{\varphi(a_m)}+O\Big(\frac{y_{\max}\varphi(W)\log R}{WD_0}\Big).$$

We note that $g(p)p/\varphi(p)^2=1+O(p^{-2})$. Thus, since the contribution is zero unless $\prod_{i=1}^k r_i$ is coprime to W, we see that the product in the above expression may be replaced by $1+O(D_0^{-1})$. This gives the result. $\square$

6. Smooth choice of y

We now choose suitable values for our y variables and complete the proof of Proposition 4.1.

We first give some comments to motivate our choice of the y variables, which we believe should be close to optimal. We wish to choose y so as to maximize the ratio of the main terms of S_2 and S_1. If we use Lagrangian multipliers to maximize this ratio (treating all error terms as zero), we arrive at the condition that

$$\lambda y_{r_1,\dots,r_k}=\Big(\prod_{i=1}^k\frac{\varphi(r_i)}{g(r_i)}\Big)\sum_{m=1}^k\frac{g(r_m)}{\varphi(r_m)}y^{(m)}_{r_1,\dots,r_{m-1},1,r_{m+1},\dots,r_k} \tag{6.1}$$

for some fixed constant λ. The y terms are supported on integers free of small prime factors, and for most integers r free of small prime factors we have $g(r) \approx \varphi(r) \approx r$, and so the above condition reduces to

$$\lambda y_{r_1,\dots,r_k} \approx \sum_{m=1}^{k} y^{(m)}_{r_1,\dots,r_{m-1},1,r_{m+1},\dots,r_k}. \tag{6.2}$$

This condition looks smooth (it has no dependence on the prime factorization of the r_i) and should be able to be satisfied if $y_{r_1,\dots,r_k}$ is a smooth function of the r_i variables. Motivated by the above, when the product $r = \prod_{i=1}^k r_i$ satisfies $(r, W) = 1$ and $\mu(r)^2 = 1$, we choose

$$y_{r_1,\dots,r_k} = F\Big(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\Big), \tag{6.3}$$

for some smooth function $F : \mathbb{R}^k \to \mathbb{R}$, supported on $\mathcal{R}_k = \{(x_1, \dots, x_k) \in [0,1]^k : \sum_{i=1}^k x_i \le 1\}$. As previously required, we set $y_{r_1,\dots,r_k} = 0$ if the product r is either not coprime to W or is not square-free. With this choice of y, we can obtain suitable asymptotic estimates for S_1 and S_2.

We will use the following lemma to estimate our sums S_1 and S_2 with this choice of y.

Lemma 6.1. *Let $A_1, A_2, L > 0$. Let γ be a multiplicative function satisfying*

$$0 \le \frac{\gamma(p)}{p} \le 1 - A_1,$$

and

$$-L \le \sum_{w \le p \le z} \frac{\gamma(p) \log p}{p} - \log z/w \le A_2$$

for any $2 \le w \le z$. 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

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

where

$$\mathfrak{S} = \prod_p \Big(1 - \frac{\gamma(p)}{p}\Big)^{-1}\Big(1 - \frac{1}{p}\Big).$$

Here the constant implied by the 'O' term is independent of G and L.

Proof. This is [3, Lemma 4], with $\kappa = 1$ and slight changes to the notation. □

We now finish our estimations of S_1 and $S_2^{(m)}$, completing the proof of Proposition 4.1. We first estimate S_1.

LEMMA 6.2. *Let $y_{r_1,\dots,r_k}$ be given in terms of a smooth function F by* (6.3), *with F supported on $\mathcal{R}_k = \{(x_1,\dots,x_k) \in [0,1]^k : \sum_{i=1}^k x_i \le 1\}$. Let*

$$F_{\max} = \sup_{(t_1,\dots,t_k)\in[0,1]^k} |F(t_1,\dots,t_k)| + \sum_{i=1}^k |\frac{\partial F}{\partial t_i}(t_1,\dots,t_k)|.$$

Then we have

$$S_1 = \frac{\varphi(W)^k N(\log R)^k}{W^{k+1}} I_k(F) + O\Big(\frac{F_{\max}^2 \varphi(W)^k N(\log R)^k}{W^{k+1}D_0}\Big),$$

where

$$I_k(F) = \int_0^1 \cdots \int_0^1 F(t_1,\dots,t_k)^2 dt_1 \cdots dt_k.$$

Proof. We substitute our choice (6.3) of y into our expression of S_1 in terms of $y_{r_1,\dots,r_k}$ given by Lemma 5.1. This gives

$$S_1 = \frac{N}{W} \sum_{\substack{u_1,\dots,u_k \\ (u_i,u_j)=1 \forall i\ne j \\ (u_i,W)=1 \forall i}} \Big(\prod_{i=1}^k \frac{\mu(u_i)^2}{\varphi(u_i)}\Big) F\Big(\frac{\log u_1}{\log R}, \dots, \frac{\log u_k}{\log R}\Big)^2 + O\Big(\frac{F_{\max}^2 \varphi(W)^k N(\log R)^k}{W^{k+1}D_0}\Big). \tag{6.4}$$

We note that two integers a and b with $(a,W) = (b,W) = 1$ but $(a,b) \ne 1$ must have a common prime factor which is greater than D_0. Thus we can drop the requirement that $(u_i,u_j) = 1$, at the cost of an error of size

$$\begin{aligned} &\ll \frac{F_{\max}^2 N}{W} \sum_{p > D_0} \sum_{\substack{u_1,\dots,u_k < R \\ p | u_i, u_j \\ (u_i,W)=1 \forall i}} \prod_{i=1}^k \frac{\mu(u_i)^2}{\varphi(u_i)} \\ &\ll \frac{F_{\max}^2 N}{W} \sum_{p>D_0} \frac{1}{(p-1)^2} \Big(\sum_{\substack{u<R \\ (u,W)=1}} \frac{\mu(u)^2}{\varphi(u)}\Big)^k \ll \frac{F_{\max}^2 \varphi(W)^k N(\log R)^k}{W^{k+1}D_0}. \end{aligned} \tag{6.5}$$

Thus we are left to evaluate the sum

$$\sum_{\substack{u_1,\dots,u_k \\ (u_i,W)=1 \forall i}} \Big(\prod_{i=1}^k \frac{\mu(u_i)^2}{\varphi(u_i)}\Big) F\Big(\frac{\log u_1}{\log R}, \dots, \frac{\log u_k}{\log R}\Big)^2. \tag{6.6}$$

We can now estimate this sum by k applications of Lemma 6.1, dealing with the sum over each u_i in turn. For each application, we take

$$\gamma(p)=\begin{cases}1, & p\nmid W,\\ 0, & \text{otherwise},\end{cases} \tag{6.7}$$

$$L\ll 1+\sum_{p|W}\frac{\log p}{p}\ll \log D_0, \tag{6.8}$$

and A_1 and A_2 fixed constants of suitable size. This gives

$$\begin{aligned}\sum_{\substack{u_1,\dots,u_k\\(u_i,W)=1\forall i}}\Big(\prod_{i=1}^k\frac{\mu(u_i)^2}{\varphi(u_i)}\Big)F\Big(\frac{\log u_1}{\log R},\dots,\frac{\log u_k}{\log R}\Big)^2\\ =\frac{\varphi(W)^k(\log R)^k}{W^k}I_k(F)+O\Big(\frac{F_{\max}^2\varphi(W)^k(\log D_0)(\log R)^{k-1}}{W^k}\Big).\end{aligned} \tag{6.9}$$

We now combine (6.9) with (6.4) and (6.5) to obtain the result. □

LEMMA 6.3. *Let $y_{r_1,\dots,r_k}$, F and $F_{\max}$ be as described in Lemma 6.2. Then we have*

$$S_2^{(m)}=\frac{\varphi(W)^kN(\log R)^{k+1}}{W^{k+1}\log N}J_k^{(m)}(F)+O\Big(\frac{F_{\max}^2\varphi(W)^kN(\log R)^k}{W^{k+1}D_0}\Big),$$

where

$$J_k^{(m)}(F)=\int_0^1\cdots\int_0^1\Big(\int_0^1F(t_1,\dots,t_k)dt_m\Big)^2dt_1\cdots dt_{m-1}dt_{m+1}\cdots dt_k.$$

Proof. The estimation of $S_2^{(m)}$ is similar to the estimation of S_1. We first estimate $y^{(m)}_{r_1,\dots,r_k}$. We recall that $y^{(m)}_{r_1,\dots,r_k}=0$ unless $r_m=1$ and $r=\prod_{i=1}^k r_i$ satisfies $(r,W)=1$ and $\mu(r)^2=1$, in which case $y^{(m)}_{r_1,\dots,r_k}$ is given in terms of $y_{r_1,\dots,r_k}$ by Lemma 5.3. We first concentrate on this case when $y^{(m)}_{r_1,\dots,r_k}\neq 0$. We substitute our choice (6.3) of y into our expression from Lemma 5.3. This gives

$$\begin{aligned}y^{(m)}_{r_1,\dots,r_k}&=O\Big(\frac{F_{\max}\varphi(W)\log R}{WD_0}\Big)\\ &+\sum_{(u,W\prod_{i=1}^k r_i)=1}\frac{\mu(u)^2}{\varphi(u)}F\Big(\frac{\log r_1}{\log R},\dots,\frac{\log r_{m-1}}{\log R},\frac{\log u}{\log R},\frac{\log r_{m+1}}{\log R},\dots,\frac{\log r_k}{\log R}\Big).\end{aligned} \tag{6.10}$$

We can see from this that $y^{(m)}_{\max} \ll \varphi(W) F_{\max}(\log R)/W$. We now estimate the sum over u in (6.10). We apply Lemma 6.1 with

$$\gamma(p) = \begin{cases} 1, & p \nmid W \prod_{i=1}^k r_i, \\ 0, & \text{otherwise}, \end{cases} \tag{6.11}$$

$$L \ll 1 + \sum_{p | W \prod_{i=1}^k r_i} \frac{\log p}{p} \ll \sum_{p < \log R} \frac{\log p}{p} + \sum_{\substack{p | W \prod_{i=1}^k r_i \\ p > \log R}} \frac{\log \log R}{\log R} \ll \log \log N, \tag{6.12}$$

and with A_1, A_2 suitable fixed constants. This gives us

$$y^{(m)}_{r_1,\dots,r_k} = (\log R) \frac{\varphi(W)}{W} \Big(\prod_{i=1}^k \frac{\varphi(r_i)}{r_i} \Big) F^{(m)}_{r_1,\dots,r_k} + O\Big(\frac{F_{\max} \varphi(W) \log R}{W D_0} \Big), \tag{6.13}$$

where

$$F^{(m)}_{r_1,\dots,r_k} = \int_0^1 F\Big(\frac{\log r_1}{\log R}, \dots, \frac{\log r_{m-1}}{\log R}, t_m, \frac{\log r_{m+1}}{\log R}, \dots, \frac{\log r_k}{\log R} \Big) dt_m. \tag{6.14}$$

Thus we have shown that if $r_m = 1$ and $r = \prod_{i=1}^k r_i$ satisfies $(r, W) = 1$ and $\mu(r)^2 = 1$, then $y^{(m)}_{r_1,\dots,r_k}$ is given by (6.13), and otherwise $y^{(m)}_{r_1,\dots,r_k} = 0$. We now substitute this into our expression from Lemma 5.2; namely,

$$\begin{aligned} S_2^{(m)} = &\frac{N}{\varphi(W) \log N} \sum_{r_1,\dots,r_k} \frac{(y^{(m)}_{r_1,\dots,r_k})^2}{\prod_{i=1}^k g(r_i)} \\ &+ O\Big(\frac{(y^{(m)}_{\max})^2 \varphi(W)^{k-2} N (\log N)^{k-2}}{W^{k-1} D_0} \Big) + O\Big(\frac{y^2_{\max} N}{(\log N)^A} \Big). \end{aligned} \tag{6.15}$$

We obtain

$$\begin{aligned} S_2^{(m)} = &\frac{\varphi(W) N (\log R)^2}{W^2 \log N} \sum_{\substack{r_1,\dots,r_k \\ (r_i, W) = 1 \forall i \\ (r_i, r_j) = 1 \forall i \neq j \\ r_m = 1}} \Big(\prod_{i=1}^k \frac{\mu(r_i)^2 \varphi(r_i)^2}{g(r_i) r_i^2} \Big) (F^{(m)}_{r_1,\dots,r_k})^2 \\ &+ O\Big(\frac{F^2_{\max} \varphi(W)^k N (\log R)^k}{W^{k+1} D_0} \Big). \end{aligned} \tag{6.16}$$

We remove the condition that $(r_i, r_j) = 1$ in the same way we did when considering S_1. Instead of (6.5), this introduces an error which is of size

$$\begin{aligned} (6.17) \qquad &\ll \frac{\varphi(W)N(\log R)^2F_{\max}^2}{W^2\log N}\Big(\sum_{p>D_0}\frac{\varphi(p)^4}{g(p)^2p^4}\Big)\Big(\sum_{\substack{r<R\\(r,W)=1}}\frac{\mu(r)^2\varphi(r)^2}{g(r)r^2}\Big)^{k-1}\\ &\ll \frac{F_{\max}^2\varphi(W)^kN(\log N)^k}{W^{k+1}D_0}. \end{aligned}$$

Thus we are left to evaluate the sum

$$(6.18) \qquad \sum_{\substack{r_1,\dots,r_{m-1},r_{m+1},\dots,r_k\\(r_i,W)=1\forall i}}\Big(\prod_{\substack{1\le i\le k\\i\ne j}}\frac{\mu(r_i)^2\varphi(r_i)^2}{g(r_i)r_i^2}\Big)(F^{(m)}_{r_1,\dots,r_k})^2.$$

We estimate this by applying Lemma 6.1 to each summation variable in turn. In each case we take

$$(6.19) \qquad \gamma(p) = \begin{cases} 1-\frac{p^2-3p+1}{p^3-p^2-2p+1}, & p\nmid W,\\ 0, & \text{otherwise,}\end{cases}$$

$$(6.20) \qquad L \ll 1+\sum_{p|W}\frac{\log p}{p} \ll \log D_0,$$

and A_1, A_2 suitable fixed constants. This gives

$$(6.21) \qquad S_2^{(m)} = \frac{\varphi(W)^kN(\log R)^{k+1}}{W^{k+1}\log N}J_k^{(m)} + O\Big(\frac{F_{\max}^2\varphi(W)^kN(\log N)^k}{W^{k+1}D_0}\Big),$$

where

$$(6.22) \qquad J_k^{(m)} = \int_0^1\cdots\int_0^1\Big(\int_0^1 F(t_1,\dots,t_k)dt_m\Big)^2dt_1\cdots dt_{m-1}dt_{m+1}\cdots dt_k,$$

as required. □

Remark. If $F(t_1,\dots,t_k) = G(\sum_{i=1}^k t_i)$ for some function G, then $I_k(F)$ and $J_k^{(m)}(F)$ simplify to $I_k(F) = \int_0^1 G(t)^2t^{k-1}dt/(k-1)!$ and $J_k^{(m)}(F) = \int_0^1(\int_t^1 G(v)dv)^2t^{k-2}dt/(k-2)!$ for each m, which is equivalent to the results obtained using the original GPY method using weights given by (2.3).

Remark. Tao gives an alternative approach to arrive at his equivalent of Proposition 4.1. His approach is to define $\lambda_{d_1,\dots,d_k}$ in terms of a suitable smooth function $f(t_1,\dots,t_k)$ as in (2.5). He then estimates the corresponding sums directly using Fourier integrals. This is somewhat similar to the original paper of Goldston, Pintz and Yıldırım [5]. Our function F corresponds to $f(t_1,\dots,t_k)$ differentiated with respect to each coordinate.

7. **Choice of smooth weight for large** k

In this section we establish part (3) of Proposition 4.3. Our argument here is closely related to that of Tao, who uses a probability theory proof.

We let $\mathcal{S}_k$ denote the set of Riemann-integrable functions $F:[0,1]^k\rightarrow\mathbb{R}$ supported on $\mathcal{R}_k=\{(x_1,\dots,x_k)\in[0,1]^k:\sum_{i=1}^k x_i\le 1\}$ with $I_k(F)\ne 0$ and $J_k^{(m)}(F)\ne 0$ for each m. We would like to obtain a lower bound for

$$M_k=\sup_{F\in\mathcal{S}_k}\frac{\sum_{m=1}^k J_k^{(m)}(F)}{I_k(F)}. \tag{7.1}$$

Remark. Let $\mathcal{L}_k$ denote the linear operator defined by

(7.2)

$$\mathcal{L}_kF(u_1,\dots,u_k)=\sum_{m=1}^k\int_0^{1-\sum_{i\ne m}u_i}F(u_1,\dots,u_{m-1},t_m,u_{m+1},\dots,u_k)dt_m$$

whenever $(u_1,\dots,u_k)\in\mathcal{R}_k$, and zero otherwise. We expect that if F maximizes the ratio $\sum_{m=1}^k J_k^{(m)}(F)/I_k(F)$, then F is an eigenfunction for $\mathcal{L}_k$, and the corresponding eigenvalue is the value of ratio at F. Unfortunately the author has not been able to solve the eigenvalue equation for $\mathcal{L}_k$ when $k>2$.

We obtain a lower bound for M_k by constructing a function $F=F_k$ which makes the ratio $\sum_{m=1}^k J_k^{(m)}(F)/I_k(F)$ large provided k is large. We choose F to be of the form

$$F(t_1,\dots,t_k)=\begin{cases}\prod_{i=1}^k g(kt_i), & \text{if } \sum_{i=1}^k t_i\le 1,\\ 0, & \text{otherwise,}\end{cases} \tag{7.3}$$

for some smooth function $g:[0,\infty]\rightarrow\mathbb{R}$, supported on $[0,T]$. We see that with this choice F is symmetric, and so $J_k^{(m)}(F)$ is independent of m. Thus we only need to consider $J_k=J_k^{(1)}(F)$. Similarly we write $I_k=I_k(F)$.

The key observation is that if the center of mass $\int_0^\infty ug(u)^2du/\int_0^\infty g(u)^2du$ of g^2 is strictly less than 1, then for large k we expect that the constraints $\sum_{i=1}^k t_i\le 1$ to be able to be dropped at the cost of only a small error. This is because (by concentration of measure) the main contribution to the unrestricted integrals

$$I'_k=\int_0^\infty\cdots\int_0^\infty\prod_{i=1}^k g(kt_i)^2dt_1\cdots dt_k$$

and

$$J'_k=\int_0^\infty\cdots\int_0^\infty(\int_0^\infty\prod_{i=1}^k g(kt_i)dt_1)^2dt_2\cdots dt_k$$

should come primarily from when $\sum_{i=1}^k t_i$ is close to the center of mass. Therefore we would expect the contribution when $\sum_{i=1}^k t_i>1$ to be small if the center of mass is less than 1, and so I_k and J_k are well approximated by I_k' and J_k' in this case.

To ease notation we let $\gamma = \int_{u\geq 0} g(u)^2 du$, and we restrict our attention to g such that $\gamma > 0$. We have

$$I_k = \underset{\mathcal{R}_k}{\int \cdots \int} F(t_1, \dots, t_k)^2 dt_1 \cdots dt_k \leq \Big(\int_0^\infty g(kt)^2 dt\Big)^k = k^{-k}\gamma^k. \tag{7.4}$$

We now consider J_k. Since squares are nonnegative, we obtain a lower bound for J_k if we restrict the outer integral to $\sum_{i=2}^k t_i < 1 - T/k$. This has the advantage that, by the support of g, there are no further restrictions on the inner integral. Thus

$$J_k \geq \underset{\substack{t_2,\dots,t_k \geq 0\\ \sum_{i=2}^k t_i \leq 1 - T/k}}{\int \cdots \int} \Big(\int_0^{T/k} \Big(\prod_{i=1}^k g(kt_i)\Big) dt_1\Big)^2 dt_2 \cdots dt_k. \tag{7.5}$$

We write the right-hand side of (7.5) as $J_k' - E_k$, where

$$\begin{aligned} J_k' &= \underset{t_2,\dots,t_k \geq 0}{\int \cdots \int} \Big(\int_0^{T/k} \Big(\prod_{i=1}^k g(kt_i)\Big) dt_1\Big)^2 dt_2 \cdots dt_k \\ &= \Big(\int_0^\infty g(kt_1) dt_1\Big)^2 \Big(\int_0^\infty g(kt)^2 dt\Big)^{k-1} = k^{-k-1}\gamma^{k-1}\Big(\int_0^\infty g(u) du\Big)^2, \end{aligned} \tag{7.6}$$

$$\begin{aligned} E_k &= \underset{\substack{t_2,\dots,t_k \geq 0\\ \sum_{i=2}^k t_i > 1 - T/k}}{\int \cdots \int} \Big(\int_0^{T/k} \Big(\prod_{i=1}^k g(kt_i)\Big) dt_1\Big)^2 dt_2 \cdots dt_k \\ &= k^{-k-1}\Big(\int_0^\infty g(u) du\Big)^2 \underset{\substack{u_2,\dots,u_k \geq 0\\ \sum_{i=2}^k u_i > k - T}}{\int \cdots \int} \Big(\prod_{i=2}^k g(u_i)^2\Big) du_2 \cdots du_k. \end{aligned} \tag{7.7}$$

First we wish to show the error integral E_k is small. We do this by comparison with a second moment. We expect the bound (7.13) for E_k to be small if the center of mass of g^2 is strictly less than $(k-T)/(k-1)$. Therefore we introduce the restriction on g that

$$\mu = \frac{\int_0^\infty u g(u)^2 du}{\int_0^\infty g(u)^2 du} < 1 - \frac{T}{k}. \tag{7.8}$$

To simplify notation, we put $\eta = (k-T)/(k-1) - \mu > 0$. If $\sum_{i=2}^k u_i > k - T$, then $\sum_{i=2}^k u_i > (k-1)(\mu + \eta)$, and so we have

$$1 \leq \eta^{-2}\Big(\frac{1}{k-1}\sum_{i=2}^k u_i - \mu\Big)^2. \tag{7.9}$$

Since the right-hand side of (7.9) is nonnegative for all u_i, we obtain an upper bound for E_k if we multiply the integrand by $\eta^{-2}(\sum_{i=2}^k u_i/(k-1)-\mu)^2$ and then drop the requirement that $\sum_{i=1}^k u_i > k-T$. This gives us

(7.10)
$$E_k \le \eta^{-2}k^{-k-1}\Big(\int_0^\infty g(u)du\Big)^2\int_0^\infty\cdots\int_0^\infty\Big(\frac{\sum_{i=2}^k u_i}{k-1}-\mu\Big)^2\Big(\prod_{i=2}^k g(u_i)^2\Big)du_2\cdots du_k.$$

We expand out the inner square. All the terms which are not of the form u_j^2 we can calculate explicitly as an expression in μ and γ. We find

(7.11)
$$\int_0^\infty\cdots\int_0^\infty\Big(\frac{2\sum_{2\le i<j\le k}u_iu_j}{(k-1)^2}-\frac{2\mu\sum_{i=2}^k u_i}{k-1}+\mu^2\Big)\Big(\prod_{i=2}^k g(u_i)^2\Big)du_2\cdots du_k = \frac{-\mu^2\gamma^{k-1}}{k-1}.$$

For the u_j^2 terms, we see that $u_j^2g(u_j)^2 \le Tu_jg(u_j)^2$ from the support of g. Thus

(7.12)
$$\int_0^\infty\cdots\int_0^\infty u_j^2\Big(\prod_{i=2}^k g(u_i)^2\Big)du_2\cdots du_k \le T\gamma^{k-2}\int_0^\infty u_jg(u_j)^2du_j = \mu T\gamma^{k-1}.$$

This gives

$$\begin{aligned} E_k &\le \eta^{-2}k^{-k-1}\Big(\int_0^\infty g(u)du\Big)^2\Big(\frac{\mu T\gamma^{k-1}}{k-1}-\frac{\mu^2\gamma^{k-1}}{k-1}\Big)\\ &\le \frac{\eta^{-2}\mu Tk^{-k-1}\gamma^{k-1}}{k-1}\Big(\int_0^\infty g(u)du\Big)^2. \end{aligned} \tag{7.13}$$

Since $(k-1)\eta^2 \ge k(1-T/k-\mu)^2$ and $\mu\le 1$, we find that putting together (7.4), (7.5), (7.6) and (7.13), we obtain

$$\frac{kJ_k}{I_k} \ge \frac{(\int_0^\infty g(u)du)^2}{\int_0^\infty g(u)^2du}\Big(1-\frac{T}{k(1-T/k-\mu)^2}\Big). \tag{7.14}$$

To maximize our lower bound (7.14), we wish to maximize $\int_0^T g(u)du$ subject to the constraints that $\int_0^T g(u)^2du=\gamma$ and $\int_0^T ug(u)^2du=\mu\gamma$. Thus we wish to maximize the expression

$$\int_0^T g(u)du-\alpha\Big(\int_0^T g(u)^2du-\gamma\Big)-\beta\Big(\int_0^T ug(u)^2du-\mu\gamma\Big) \tag{7.15}$$

with respect to α,β and the function g. By the Euler-Lagrange equation, this occurs when $\frac{\partial}{\partial g}(g(t)-\alpha g(t)^2-\beta tg(t)^2)=0$ for all $t\in[0,T]$. Thus we see that

$$g(t)=\frac{1}{2\alpha+2\beta t}\qquad \text{for } 0\le t\le T. \tag{7.16}$$

Since the ratio we wish to maximize is unaffected if we multiply g by a positive constant, we restrict our attention to functions g of the form $1/(1+At)$ for $t \in [0,T]$ and for some constant $A>0$. With this choice of g we find that

$$\int_0^T g(u)du = \frac{\log(1+AT)}{A}, \qquad \int_0^T g(u)^2 du = \frac{1}{A}\Big(1-\frac{1}{1+AT}\Big), \tag{7.17}$$

$$\int_0^T ug(u)^2 du = \frac{1}{A^2}\Big(\log(1+AT)-1+\frac{1}{1+AT}\Big). \tag{7.18}$$

We choose T such that $1+AT=e^A$ (which is close to optimal). With this choice we find that $\mu = 1/(1-e^{-A})-A^{-1}$ and $T\le e^A/A$. Thus $1-T/k-\mu \ge A^{-1}(1-A/(e^A-1)-e^A/k)$. Substituting (7.17) into (7.14), and then using these expressions, we find that

$$\frac{kJ_k}{I_k} \ge \frac{A}{1-e^{-A}}\Big(1-\frac{T}{k(1-T/k-\mu)^2}\Big) \ge A\Big(1-\frac{Ae^A}{k(1-A/(e^A-1)-e^A/k)^2}\Big), \tag{7.19}$$

provided the right-hand side is positive. Finally, we choose $A=\log k - 2\log\log k>0$. For k sufficiently large, we have

$$1-\frac{T}{k}-\mu \ge A^{-1}\Big(1-\frac{(\log k)^3}{k}-\frac{1}{(\log k)^2}\Big)>0, \tag{7.20}$$

and so $\mu<1-T/k$, as required by our constraint (7.8). This choice of A gives

$$M_k \ge \frac{kJ_k}{I_k} \ge (\log k-2\log\log k)\Big(1-\frac{\log k}{(\log k)^2+O(1)}\Big) \ge \log k - 2\log\log k - 2 \tag{7.21}$$

when k is sufficiently large.

8. Choice of weight for small k

In this section we establish parts (1) and (2) of Proposition 4.3. In order to get a suitable lower bound for M_k when k is small, we will consider approximations to the optimal function F of the form

$$F(t_1,\dots,t_k) = \begin{cases} P(t_1,\dots,t_k), & \text{if } (t_1,\dots,t_k)\in\mathcal{R}_k,\\ 0, & \text{otherwise}\end{cases} \tag{8.1}$$

for polynomials P. By the symmetry of $\sum_{m=1}^k J_k^{(m)}(F)$ and $I_k(F)$, we restrict our attention to polynomials which are symmetric functions of $t_1,\dots,t_k$. (If F satisfies $\mathcal{L}_kF=\lambda F$, then $F_\sigma = F(\sigma(t_1),\dots,\sigma(t_k))$ also satisfies this for every permutation σ of $t_1,\dots,t_k$. Thus the symmetric function which is the average of F_σ over all such permutations would satisfy this eigenfunction equation, and so we expect there to be an optimal function which is symmetric.) Any such polynomial can be written as a polynomial expression in the power sum polynomials $P_j=\sum_{i=1}^k t_i^j$.

LEMMA 8.1. *Let $P_j=\sum_{i=1}^k t_i^j$ denote the j^{th} symmetric power sum polynomial. Then we have*

$$\underset{\mathcal{R}_k}{\int\cdots\int}(1-P_1)^a P_j^b dt_1\cdots dt_k=\frac{a!}{(k+jb+a)!}G_{b,j}(k),$$

where

$$G_{b,j}(x)=b!\sum_{r=1}^{b}\binom{x}{r}\sum_{\substack{b_1,\dots,b_r\geq 1\\ \sum_{i=1}^r b_i=b}}\prod_{i=1}^r\frac{(jb_i)!}{b_i!}$$

is a polynomial of degree b which depends only on b and j.

Proof. We first show by induction on k that

$$\underset{\mathcal{R}_k}{\int\cdots\int}\Big(1-\sum_{i=1}^k t_i\Big)^a\prod_{i=1}^k t_i^{a_i}dt_1\cdots dt_k=\frac{a!\prod_{i=1}^k a_i!}{(k+a+\sum_{i=1}^k a_i)!}. \tag{8.2}$$

We consider the integration with respect to t_1. The limits of integration are 0 and $1-\sum_{i=2}^k t_i$ for $(t_2,\dots,t_k)\in\mathcal{R}_{k-1}$. By substituting $v=t_1/(1-\sum_{i=2}^k t_i)$ we find

$$\begin{aligned}\int_0^{1-\sum_{i=2}^k t_i}&\Big(1-\sum_{i=1}^k t_i\Big)^a\Big(\prod_{i=1}^k t_i^{a_i}\Big)dt_1 \\ &=\Big(\prod_{i=2}^k t_i^{a_i}\Big)\Big(1-\sum_{i=2}^k t_i\Big)^{a+a_1+1}\int_0^1(1-v)^a v^{a_1}dv \\ &=\frac{a!a_1!}{(a+a_1+1)!}\Big(\prod_{i=2}^k t_i^{a_i}\Big)\Big(1-\sum_{i=2}^k t_i\Big)^{a+a_1+1}.\end{aligned} \tag{8.3}$$

Here we used the beta function identity $\int_0^1 t^a(1-t)^b dt=a!b!/(a+b+1)!$ in the last line. We now see (8.2) follows by induction.

By the binomial theorem,

$$P_j^b=\sum_{\substack{b_1,\dots,b_k\\ \sum_{i=1}^k b_i=b}}\frac{b!}{\prod_{i=1}^k b_i!}\prod_{i=1}^k t_i^{jb_i}. \tag{8.4}$$

Thus, applying (8.2), we obtain

$$\underset{\mathcal{R}_k}{\int\cdots\int}(1-P_1)^a P_j^b dt_1\cdots dt_k=\frac{b!a!}{(k+a+jb)!}\sum_{\substack{b_1,\dots,b_k\\ \sum_{i=1}^k b_i=b}}\prod_{i=1}^k\frac{(jb_i)!}{b_i!}. \tag{8.5}$$

For computations, b will be small, and so we find it convenient to split the summation depending on how many of the b_i are nonzero. Given an integer r, there are $\binom{k}{r}$ ways of choosing r of $b_1, \dots, b_k$ to be nonzero. Thus

$$\sum_{\substack{b_1,\dots,b_k\\ \sum_{i=1}^k b_i=b}} \prod_{i=1}^k \frac{(jb_i)!}{b_i!} = \sum_{r=1}^b \binom{k}{r} \sum_{\substack{b_1,\dots,b_r\geq 1\\ \sum_{i=1}^r b_i=b}} \prod_{i=1}^r \frac{(jb_i)!}{b_i!}. \tag{8.6}$$

This gives the result. □

It is straightforward to extend Lemma 8.1 to more general combinations of the symmetric power polynomials. In this paper we will concentrate on the case when P is a polynomial expression in only P_1 and P_2 for simplicity. We comment the polynomials $G_{b,j}$ are not problematic to calculate numerically for small values of b. We now use Lemma 8.1 to obtain a manageable expression for $I_k(F)$ and $J_k^{(m)}(F)$ with this choice of P.

Lemma 8.2. *Let F be given in terms of a polynomial P by* (8.1). *Let P be given in terms of a polynomial expression in the symmetric power polynomials $P_1 = \sum_{i=1}^k t_i$ and $P_2 = \sum_{i=1}^k t_i^2$ by $P = \sum_{i=1}^d a_i(1-P_1)^{b_i}P_2^{c_i}$ for constants $a_i \in \mathbb{R}$ and nonnegative integers b_i, c_i. Then for each $1 \leq m \leq k$, we have*

$$I_k(F) = \sum_{1\leq i,j\leq d} a_i a_j \frac{(b_i+b_j)!G_{c_i+c_j,2}(k)}{(k+b_i+b_j+2c_i+2c_j)!},$$

$$J_k^{(m)}(F) = \sum_{1\leq i,j\leq d} a_i a_j \sum_{c_1'=0}^{c_i}\sum_{c_2'=0}^{c_j} \binom{c_i}{c_1'}\binom{c_j}{c_2'} \frac{\gamma_{b_i,b_j,c_i,c_j,c_1',c_2'}G_{c_1'+c_2',2}(k-1)}{(k+b_i+b_j+2c_i+2c_j+1)!},$$

where

$$\gamma_{b_i,b_j,c_i,c_j,c_1',c_2'} = \frac{b_i!b_j!(2c_i-2c_1')!(2c_j-2c_2')!(b_i+b_j+2c_i+2c_j-2c_1'-2c_2'+2)!}{(b_i+2c_i-2c_1'+1)!(b_j+2c_j-2c_2'+1)!}$$

and where G is the polynomial given by Lemma 8.1.

Proof. We first consider $I_k(F)$. We have, using Lemma 8.1,

$$\begin{aligned} I_k(F) &= \int\cdots\int_{\mathcal{R}_k} P^2 dt_1\cdots dt_k \\ &= \sum_{1\leq i,j\leq d} a_i a_j \int\cdots\int_{\mathcal{R}_k} (1-P_1)^{b_i+b_j}P_2^{c_i+c_j} dt_1\cdots dt_k \\ &= \sum_{1\leq i,j\leq d} a_i a_j \frac{(b_i+b_j)!G_{c_i+c_j,2}(k)}{(k+b_i+b_j+2c_i+2c_j)!}. \end{aligned} \tag{8.7}$$

We now consider $J_k^{(m)}(F)$. Since F is symmetric in $t_1, \dots, t_k$, we see that $J_k^{(m)}(F)$ is independent of m, and so it suffices to only consider $J_k^{(1)}(F)$. We have

$$\int_0^{1-\sum_{i=2}^k t_i}(1-P_1)^b P_2^c dt_1 \tag{8.8}$$
$$= \sum_{c'=0}^{c}\binom{c}{c'}\Big(\sum_{i=2}^k t_i^2\Big)^{c'}\int_0^{1-\sum_{i=2}^k t_i}\Big(1-\sum_{i=1}^k t_i\Big)^b t_1^{2c-2c'}dt_1$$
$$= \sum_{c'=0}^{c}\binom{c}{c'}(P_2')^{c'}(1-P_1')^{b+2c-2c'+1}\int_0^1(1-u)^b u^{2c-2c'}du$$
$$= \sum_{c'=0}^{c}\binom{c}{c'}(P_2')^{c'}(1-P_1')^{b+2c-2c'+1}\frac{b!(2c-2c')!}{(b+2c-2c'+1)!},$$

where $P_1'=\sum_{i=2}^k t_i$ and $P_2'=\sum_{i=2}^k t_i^2$. Thus

(8.9)
$$\Big(\int_0^1 F dt_1\Big)^2 = \Big(\sum_{i=1}^d a_i\int_0^{1-\sum_{j=2}^k t_j}(1-P_1)^{b_i}P_2^{c_i}dt_1\Big)^2$$
$$= \sum_{1\le i,j\le d}a_ia_j\sum_{c_1'=0}^{c_i}\sum_{c_2'=0}^{c_j}\binom{c_i}{c_1'}\binom{c_j}{c_2'}(P_2')^{c_1'+c_2'}(1-P_1')^{b_i+b_j+2c_i+2c_j-2c_1'-2c_2'+2}$$
$$\times\frac{b_i!b_j!(2c_i-2c_1')!(2c_j-2c_2')!}{(b_i+2c_i-2c_1'+1)!(b_j+2c_j-2c_2'+1)!}.$$

Applying Lemma 8.1 again, we see that

$$\int\cdots\int_{\mathcal{R}_{k-1}}(1-P_1')^b(P_2')^{c'}dt_2\cdots dt_k = \frac{b!}{(k+b+c-1)!}G_{c,2}(k-1). \tag{8.10}$$

Combining (8.9) and (8.10) gives the result. □

We see from Lemma 8.2 that $I_k(F)$ and $\sum_{m=1}^k J_k^{(m)}(F)$ can both be expressed as quadratic forms in the coefficients $\mathbf{a}=(a_1,\dots,a_d)$ of P. Moreover, these will be positive definite real quadratic forms. Thus, in particular, we find that

$$\frac{\sum_{m=1}^k J_k^{(m)}(F)}{I_k(F)} = \frac{\mathbf{a}^TA_2\mathbf{a}}{\mathbf{a}^TA_1\mathbf{a}}, \tag{8.11}$$

for two rational symmetric positive definite matrices A_1,A_2, which can be calculated explicitly in terms of k for any choice of the exponents b_i,c_i. Maximizing expressions of this form has a known solution.

Lemma 8.3. *Let A_1, A_2 be real, symmetric positive definite matrices. Then*

$$\frac{\mathbf{a}^TA_2\mathbf{a}}{\mathbf{a}^TA_1\mathbf{a}}$$

is maximized when $\mathbf{a}$ *is an eigenvector of* $A_1^{-1}A_2$ *corresponding to the largest eigenvalue of* $A_1^{-1}A_2$. *The value of the ratio at its maximum is this largest eigenvalue.*

Proof. We see that multiplying $\mathbf{a}$ by a nonzero scalar does not change the ratio, so we may assume without loss of generality that $\mathbf{a}^T A_1 \mathbf{a} = 1$. By the theory of Lagrangian multipliers, $\mathbf{a}^T A_2 \mathbf{a}$ is maximized subject to $\mathbf{a}^T A_1 \mathbf{a} = 1$ when

$$L(\mathbf{a}, \lambda) = \mathbf{a}^T A_2 \mathbf{a} - \lambda(\mathbf{a}^T A_1 \mathbf{a} - 1) \tag{8.12}$$

is stationary. This occurs when (using the symmetricity of A_1, A_2)

$$0 = \frac{\partial L}{\partial a_i} = ((2A_2 - 2\lambda A_1)\mathbf{a})_i \tag{8.13}$$

for each i. This implies that (recalling that A_1 is positive definite so invertible)

$$A_1^{-1} A_2 \mathbf{a} = \lambda \mathbf{a}. \tag{8.14}$$

It then is clear that $\mathbf{a}^T A_1 \mathbf{a} = \lambda^{-1} \mathbf{a}^T A_2 \mathbf{a}$. □

Proof of parts (1) *and* (2) *of Proposition* 4.3. To establish Proposition 4.3 we rely on some computer calculation to calculate a lower bound for M_k. We let F be given in terms of a polynomial P by (8.1). We let P be given by a polynomial expression in $P_1 = \sum_{i=1}^k t_i$ and $P_2 = \sum_{i=1}^k t_i^2$ which is a linear combination of all monomials $(1-P_1)^b P_2^c$ with $b + 2c \le 11$. There are 42 such monomials, and with $k = 105$ we can calculate the 42×42 rational symmetric matrices A_1 and A_2 corresponding to the coefficients of the quadratic forms $I_k(F)$ and $\sum_{m=1}^k J_k^{(m)}(F)$. We then find[3] that the largest eigenvalue of $A_1^{-1}A_2$ is

$$\lambda \approx 4.0020697 \cdots > 4. \tag{8.15}$$

Thus $M_{105} > 4$. This verifies part (2) of Proposition 4.3. We comment that by taking a rational approximation to the corresponding eigenvector, we can verify this lower bound by calculating the ratio $\sum_{m=1}^k J_k^{(m)}(F)/I_k(F)$ using only exact arithmetic.

For part (1) of Proposition 4.3, we take $k = 5$ and

$$P = (1-P_1)P_2 + \frac{7}{10}(1-P_1)^2 + \frac{1}{14}P_2 - \frac{3}{14}(1-P_1). \tag{8.16}$$

With this choice we find that

$$M_5 \ge \frac{\sum_{m=1}^k J_k^{(m)}(F)}{I_k(F)} = \frac{1\,417\,255}{708\,216} > 2. \tag{8.17}$$

This completes the proof of Proposition 4.3. □

[3] An ancillary *Mathematica*® file detailing these computations is available alongside this paper at www.arxiv.org.

9. Acknowledgements

The author would like to thank Andrew Granville, Roger Heath-Brown, Dimitris Koukoulopoulos and Terence Tao for many useful conversations and suggestions.

The work leading to this paper was started whilst the author was a D.Phil student at Oxford and funded by the EPSRC (Doctoral Training Grant EP/P505216/1), and it was finished when the author was a CRM-ISM post-doctoral fellow at the Université de Montréal.

References

[1] P. D. T. A. Elliott and H. Halberstam, A conjecture in prime number theory, in *Symposia Mathematica, Vol. IV* (INDAM, Rome, 1968/69), Academic Press, London, 1970, pp. 59–72. MR 0276195. Zbl 0238.10030.

[2] J. Friedlander and A. Granville, Limitations to the equi-distribution of primes. I, *Ann. of Math.* **129** (1989), 363–382. MR 0986796. Zbl 0671.10041. http://dx.doi.org/10.2307/1971450.

[3] D. A. Goldston, S. W. Graham, J. Pintz, and C. Y. Yildirim, Small gaps between products of two primes, *Proc. Lond. Math. Soc.* **98** (2009), 741–774. MR 2500871. Zbl 1213.11171. http://dx.doi.org/10.1112/plms/pdn046.

[4] D. A. Goldston and C. Y. Yildirim, Higher correlations of divisor sums related to primes. III. Small gaps between primes, *Proc. Lond. Math. Soc.* **95** (2007), 653–686. MR 2368279. Zbl 1118.11040. http://dx.doi.org/10.1112/plms/pdm021.

[5] D. A. Goldston, J. Pintz, and C. Y. Yildirim, Primes in tuples. I, *Ann. of Math.* **170** (2009), 819–862. MR 2552109. Zbl 1207.11096. http://dx.doi.org/10.4007/annals.2009.170.819.

[6] D. A. Goldston, J. Pintz, and C. Y. Yildirim, Primes in tuples. III. On the difference $p_{n+\nu} - p_n$, *Funct. Approx. Comment. Math.* **35** (2006), 79–89. MR 2271608. Zbl 1196.11123. http://dx.doi.org/10.7169/facm/1229442618.

[7] D. H. J. Polymath, New equidistribution estimates of Zhang type, and bounded gaps between primes, preprint. arXiv 1402.0811.

[8] A. Selberg, *Collected Papers. Vol. II*, Springer-Verlag, New York, 1991, with a foreword by K. Chandrasekharan. MR 1295844. Zbl 0729.11001.

[9] Y. Zhang, Bounded gaps between primes, *Ann. of Math.* **179** (2014), 1121–1174. MR 3171761. Zbl 06302171. http://dx.doi.org/10.4007/annals.2014.179.3.7.

(Received: January 3, 2014)
(Revised: March 11, 2014)

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