**Research in
the Mathematical Sciences**
a SpringerOpen Journal

RESEARCH ARTICLE

Open Access

Variants of the Selberg sieve, and bounded intervals containing many primes

DHJ Polymath

Correspondence:
tao@math.ucla.edu
http://michaelnielsen.org/polymath1/
index.php

Abstract

For any $m \geq 1$, let H_m denote the quantity $\liminf_{n\to\infty}(p_{n+m} - p_n)$. A celebrated recent result of Zhang showed the finiteness of H_1, with the explicit bound $H_1 \leq 70,000,000$. This was then improved by us (the Polymath8 project) to $H_1 \leq 4680$, and then by Maynard to $H_1 \leq 600$, who also established for the first time a finiteness result for H_m for $m \geq 2$, and specifically that $H_m \ll m^3 e^{4m}$. If one also assumes the Elliott-Halberstam conjecture, Maynard obtained the bound $H_1 \leq 12$, improving upon the previous bound $H_1 \leq 16$ of Goldston, Pintz, and Yıldırım, as well as the bound $H_m \ll m^3 e^{2m}$. In this paper, we extend the methods of Maynard by generalizing the Selberg sieve further and by performing more extensive numerical calculations. As a consequence, we can obtain the bound $H_1 \leq 246$ unconditionally and $H_1 \leq 6$ under the assumption of the generalized Elliott-Halberstam conjecture. Indeed, under the latter conjecture, we show the stronger statement that for any admissible triple (h_1, h_2, h_3), there are infinitely many n for which at least two of $n + h_1, n + h_2, n + h_3$ are prime, and also obtain a related disjunction asserting that either the twin prime conjecture holds or the even Goldbach conjecture is asymptotically true if one allows an additive error of at most 2, or both. We also modify the 'parity problem' argument of Selberg to show that the $H_1 \leq 6$ bound is the best possible that one can obtain from purely sieve-theoretic considerations. For larger m, we use the distributional results obtained previously by our project to obtain the unconditional asymptotic bound $H_m \ll me^{(4 - \frac{28}{157})m}$ or $H_m \ll me^{2m}$ under the assumption of the Elliott-Halberstam conjecture. We also obtain explicit upper bounds for H_m when $m = 2, 3, 4, 5$.

Keywords: Selberg sieve; Elliott-Halberstam conjecture; Prime gaps

Background

For any natural number m, let H_m denote the quantity

$$H_m := \liminf_{n\to\infty}(p_{n+m} - p_n),$$

where p_n denotes the nth prime. The twin prime conjecture asserts that $H_1 = 2$; more generally, the Hardy-Littlewood prime tuples conjecture [1] implies that $H_m = H(m+1)$ for all $m \geq 1$, where $H(k)$ is the diameter of the narrowest admissible k-tuple (see the 'Outline of the key ingredients' section for a definition of this term). Asymptotically, one has the bounds

$$\left(\frac{1}{2} + o(1)\right)k \log k \leq H(k) \leq (1 + o(1))k \log k$$

Springer

as $k \to \infty$ (see Theorem 17 below); thus, the prime tuples conjecture implies that H_m is comparable to $m \log m$ as $m \to \infty$.

Until very recently, it was not known if any of the H_m were finite, even in the easiest case $m = 1$. In the breakthrough work of Goldston et al. [2], several results in this direction were established, including the following conditional result assuming the Elliott-Halberstam conjecture $\text{EH}[\vartheta]$ (see Claim 8 below) concerning the distribution of the prime numbers in arithmetic progressions:

Theorem 1 (GPY theorem). *Assume the Elliott-Halberstam conjecture* $\text{EH}[\vartheta]$ *for all* $0 < \vartheta < 1$. *Then,* $H_1 \leq 16$.

Furthermore, it was shown in [2] that any result of the form $\text{EH}\left[\frac{1}{2} + 2\varpi\right]$ for some fixed $0 < \varpi < 1/4$ would imply an explicit finite upper bound on H_1 (with this bound equal to 16 for $\varpi > 0.229855$). Unfortunately, the only results of the type $\text{EH}[\vartheta]$ that are known come from the Bombieri-Vinogradov theorem (Theorem 9), which only establishes $\text{EH}[\vartheta]$ for $0 < \vartheta < 1/2$.

The first unconditional bound on H_1 was established in a breakthrough work of Zhang [3]:

Theorem 2 (Zhang's theorem). $H_1 \leq 70,000,000$.

Zhang's argument followed the general strategy from [2] on finding small gaps between primes, with the major new ingredient being a proof of a weaker version of $\text{EH}\left[\frac{1}{2} + 2\varpi\right]$, which we call $\text{MPZ}[\varpi, \delta]$ (see Claim 10) below. It was quickly realized that Zhang's numerical bound on H_1 could be improved. By optimizing many of the components in Zhang's argument, we were able (Polymath, DHJ: New equidistribution estimates of Zhang type, submitted), [4] to improve Zhang's bound to

$$H_1 \leq 4,680.$$

Very shortly afterwards, a further breakthrough was obtained by Maynard [5] (with related work obtained independently in an unpublished work of Tao), who developed a more flexible 'multidimensional' version of the Selberg sieve to obtain stronger bounds on H_m. This argument worked without using any equidistribution results on primes beyond the Bombieri-Vinogradov theorem, and among other things was able to establish finiteness of H_m for all m, not just for $m = 1$. More precisely, Maynard established the following results.

Theorem 3 (Maynard's theorem). *Unconditionally, we have the following bounds:*

(i) $H_1 \leq 600$
(ii) $H_m \leq Cm^3 e^{4m}$ *for all* $m \geq 1$ *and an absolute (and effective) constant* C

Assuming the Elliott-Halberstam conjecture $\text{EH}[\vartheta]$ *for all* $0 < \vartheta < 1$, *we have the following improvements:*

(iii) $H_1 \leq 12$
(iv) $H_2 \leq 600$
(v) $H_m \leq Cm^3 e^{2m}$ *for all* $m \geq 1$ *and an absolute (and effective) constant* C

For a survey of these recent developments, see [6].

In this paper, we refine Maynard's methods to obtain the following further improvements.

Theorem 4. *Unconditionally, we have the following bounds:*

(i) $H_1 \leq 246$
(ii) $H_2 \leq 398,130$
(iii) $H_3 \leq 24,797,814$
(iv) $H_4 \leq 1,431,556,072$
(v) $H_5 \leq 80,550,202,480$
(vi) $H_m \leq Cm \exp\left(\left(4 - \frac{28}{157}\right)m\right)$ *for all* $m \geq 1$ *and an absolute (and effective) constant* C

Assume the Elliott-Halberstam conjecture EH[ϑ] *for all* $0 < \vartheta < 1$. *Then, we have the following improvements:*

(vii) $H_2 \leq 270$
(viii) $H_3 \leq 52,116$
(ix) $H_4 \leq 474,266$.
(x) $H_5 \leq 4,137,854$.
(xi) $H_m \leq Cme^{2m}$ *for all* $m \geq 1$ *and an absolute (and effective) constant* C

Finally, assume the generalized Elliott-Halberstam conjecture GEH[ϑ] *(see Claim 12 below) for all* $0 < \vartheta < 1$. *Then,*

(xii) $H_1 \leq 6$
(xiii) $H_2 \leq 252$

In the 'Outline of the key ingredients' section, we will describe the key propositions that will be combined together to prove the various components of Theorem 4. As with Theorem 1, the results in (vii)-(xiii) do not require EH[ϑ] or GEH[ϑ] for all $0 < \vartheta < 1$, but only for a single explicitly computable ϑ that is sufficiently close to 1.

Of these results, the bound in (xii) is perhaps the most interesting, as the parity problem [7] prohibits one from achieving any better bound on H_1 than 6 from purely sieve-theoretic methods; we review this obstruction in the 'The parity problem' section. If one only assumes the Elliott-Halberstam conjecture EH[ϑ] instead of its generalization GEH[ϑ], we were unable to improve upon Maynard's bound $H_1 \leq 12$; however, the parity obstruction does not exclude the possibility that one could achieve (xii) just assuming EH[ϑ] rather than GEH[ϑ], by some further refinement of the sieve-theoretic arguments (e.g. by finding a way to establish Theorem 20(ii) below using only EH[ϑ] instead of GEH[ϑ]).

The bounds (ii)-(vi) rely on the equidistribution results on primes established in our previous paper. However, the bound (i) uses only the Bombieri-Vinogradov theorem, and the remaining bounds (vii)-(xiii) of course use either the Elliott-Halberstam conjecture or a generalization thereof.

A variant of the proof of Theorem 4(xii), which we give in 'Additional remarks' section, also gives the following conditional 'near miss' to (a disjunction of) the twin prime conjecture and the even Goldbach conjecture:

Theorem 5 (Disjunction). *Assume the generalized Elliott-Halberstam conjecture* GEH[ϑ] *for all* $0 < \vartheta < 1$. *Then, at least one of the following statements is true:*

(a) *(Twin prime conjecture)* $H_1 = 2$.

(b) *(near-miss to even Goldbach conjecture) If n is a sufficiently large multiple of 6, then at least one of n and n − 2 is expressible as the sum of two primes, similarly with n − 2 replaced by n + 2. (In particular, every sufficiently large even number lies within 2 of the sum of two primes.)*

We remark that a disjunction in a similar spirit was obtained in [8], which established (prior to the appearance of Theorem 2) that either H_1 was finite or that every interval $[x, x + x^\varepsilon]$ contained the sum of two primes if x was sufficiently large depending on $\varepsilon > 0$.

There are two main technical innovations in this paper. The first is a further generalization of the multidimensional Selberg sieve introduced by Maynard and Tao, in which the support of a certain cutoff function F is permitted to extend into a larger domain than was previously permitted (particularly under the assumption of the generalized Elliott-Halberstam conjecture). As in [5], this largely reduces the task of bounding H_m to that of efficiently solving a certain multidimensional variational problem involving the cutoff function F. Our second main technical innovation is to obtain efficient numerical methods for solving this variational problem for small values of the dimension k, as well as sharpened asymptotics in the case of large values of k.

The methods of Maynard and Tao have been used in a number of subsequent applications [9-21]. The techniques in this paper should be able to be used to obtain slight numerical improvements to such results, although we did not pursue these matters here.

Organization of the paper

The paper is organized as follows. After some notational preliminaries, we recall in the 'Distribution estimates on arithmetic functions' section the known (or conjectured) distributional estimates on primes in arithmetic progressions that we will need to prove Theorem 4. Then, in the section 'Outline of the key ingredients', we give the key propositions that will be combined together to establish this theorem. One of these propositions, Lemma 18, is an easy application of the pigeonhole principle. Two further propositions, Theorem 19 and Theorem 20, use the prime distribution results from the 'Distribution estimates on arithmetic functions' section to give asymptotics for certain sums involving sieve weights and the von Mangoldt function; they are established in the 'Multidimensional Selberg sieves' section. Theorems 22, 24, 26, and 28 use the asymptotics established in Theorems 19 and 20, in combination with Lemma 18, to give various criteria for bounding H_m, which all involve finding sufficiently strong candidates for a variety of multidimensional variational problems; these theorems are proven in the 'Reduction to a variational problem' section. These variational problems are analysed in the asymptotic regime of large k in the 'Asymptotic analysis' section, and for small and medium k in the 'The case of small and medium dimension' section, with the results collected in Theorems 23, 25, 27, and 29. Combining these results with the previous propositions gives Theorem 16, which, when combined with the bounds on narrow admissible tuples in Theorem 17 that are established in the 'Narrow admissible tuples'

section, will give Theorem 4. (See also Table 1 for more details of the logical dependencies between the key propositions.)

Finally, in the 'The parity problem' section, we modify an argument of Selberg to show that the bound $H_1 \leq 6$ may not be improved using purely sieve-theoretic methods, and in the 'Additional remarks' section, we establish Theorem 5 and make some miscellaneous remarks.

Notation

The notation used here closely follows the notation in our previous paper.

We use $|E|$ to denote the cardinality of a finite set E, and $\mathbf{1}_E$ to denote the indicator function of a set E; thus, $\mathbf{1}_E(n) = 1$ when $n \in E$ and $\mathbf{1}_E(n) = 0$ otherwise.

All sums and products will be over the natural numbers $\mathbb{N} := \{1, 2, 3, \ldots\}$ unless otherwise specified, with the exceptions of sums and products over the variable p, which will be understood to be over primes.

The following important asymptotic notation will be in use throughout the paper.

Definition 6 (Asymptotic notation)**.** We use x to denote a large real parameter, which one should think of as going off to infinity; in particular, we will implicitly assume that it is larger than any specified fixed constant. Some mathematical objects will be independent of x and referred to as *fixed*; but unless otherwise specified, we allow all mathematical objects under consideration to depend on x (or to vary within a range that depends on x, e.g. the summation parameter n in the sum $\sum_{x \leq n \leq 2x} f(n)$). If X and Y are two quantities depending on x, we say that $X = O(Y)$ or $X \ll Y$ if one has $|X| \leq CY$ for some fixed C (which we refer to as the *implied constant*), and $X = o(Y)$ if one has $|X| \leq c(x)Y$ for some function $c(x)$ of x (and of any fixed parameters present) that goes to zero as $x \to \infty$ (for each choice of fixed parameters). We use $X \lll Y$ to denote the estimate $X \leq x^{o(1)} Y$, $X \sim Y$ to denote the estimate $Y \ll X \ll Y$, and $X \approx Y$ to denote the estimate $Y \lll X \lll Y$. Finally, we say that a quantity n is of *polynomial size* if one has $n = O(x^{O(1)})$.

If asymptotic notation such as $O()$ or $\lll$ appears on the left-hand side of a statement, this means that the assertion holds true for any specific interpretation of that notation. For instance, the assertion $\sum_{n=O(N)} |\alpha(n)| \lll N$ means that for each fixed constant $C > 0$, one has $\sum_{|n| \leq CN} |\alpha(n)| \lll N$.

If q and a are integers, we write $a|q$ if a divides q. If q is a natural number and $a \in \mathbb{Z}$, we use $a\ (q)$ to denote the residue class

$$a\ (q) := \{a + nq : n \in \mathbb{Z}\}$$

Table 1 Results used to prove various components of Theorem 16

Theorem 16	Results used
(i)	Theorems 9, 26, and 27
(ii)-(vi)	Theorems 11, 24, and 25
(vii)-(xi)	Theorems 22 and 23
(xii)	Theorems 28 and 29
(xiii)	Theorems 26 and 27

Note that Theorems 22, 24, 26, and 28 are in turn proven using Theorems 19 and 20 and Lemma 18.

and let $\mathbb{Z}/q\mathbb{Z}$ denote the ring of all such residue classes $a(q)$. The notation $b = a\ (q)$ is synonymous to $b \in a\ (q)$. We use (a, q) to denote the greatest common divisor of a and q, and $[a, q]$ to denote the least common multiple[a]. We also let

$$(\mathbb{Z}/q\mathbb{Z})^{\times} := \big\{ a\ (q) : (a, q) = 1 \big\}$$

denote the primitive residue classes of $\mathbb{Z}/q\mathbb{Z}$.

We use the following standard arithmetic functions:

(i) $\varphi(q) := |(\mathbb{Z}/q\mathbb{Z})^{\times}|$ denotes the Euler totient function of q.
(ii) $\tau(q) := \sum_{d|q} 1$ denotes the divisor function of q.
(iii) $\Lambda(q)$ denotes the von Mangoldt function of q; thus, $\Lambda(q) = \log p$ if q is a power of a prime p, and $\Lambda(q) = 0$ otherwise.
(iv) $\theta(q)$ is defined to equal $\log q$ when q is a prime, and $\theta(q) = 0$ otherwise.
(v) $\mu(q)$ denotes the Möbius function of q; thus, $\mu(q) = (-1)^k$ if q is the product of k distinct primes for some $k \geq 0$, and $\mu(q) = 0$ otherwise.
(vi) $\Omega(q)$ denotes the number of prime factors of q (counting multiplicity).

We recall the elementary *divisor bound*

$$\tau(n) \ll 1 \tag{1}$$

whenever $n \ll x^{O(1)}$, as well as the related estimate

$$\sum_{n \ll x} \frac{\tau(n)^C}{n} \ll \log^{O(1)} x \tag{2}$$

for any fixed $C > 0$ (see, e.g. [Lemma 1.5]).

The *Dirichlet convolution* $\alpha \star \beta \colon \mathbb{N} \to \mathbb{C}$ of two arithmetic functions $\alpha, \beta \colon \mathbb{N} \to \mathbb{C}$ is defined in the usual fashion as

$$\alpha \star \beta(n) := \sum_{d|n} \alpha(d)\beta\left(\frac{n}{d}\right) = \sum_{ab=n} \alpha(a)\beta(b).$$

Distribution estimates on arithmetic functions

As mentioned in the introduction, a key ingredient in the Goldston-Pintz-Yıldırım approach to small gaps between primes comes from distributional estimates on the primes, or more precisely on the von Mangoldt function Λ, which serves as a proxy for the primes. In this work, we will also need to consider distributional estimates on more general arithmetic functions, although we will not prove any new such estimates in this paper, relying instead on estimates that are already in the literature.

More precisely, we will need averaged information on the following quantity:

Definition 7 (Discrepancy)**.** For any function $\alpha \colon \mathbb{N} \to \mathbb{C}$ with finite support (that is, α is non-zero only on a finite set) and any primitive residue class $a\ (q)$, we define the (signed) *discrepancy* $\Delta(\alpha; a\ (q))$ to be the quantity

$$\Delta(\alpha; a\ (q)) := \sum_{n = a\ (q)} \alpha(n) - \frac{1}{\varphi(q)} \sum_{(n,q)=1} \alpha(n). \tag{3}$$

For any fixed $0 < \vartheta < 1$, let EH$[\vartheta]$ denote the following claim:

Claim 8 (Elliott-Halberstam conjecture, EH[ϑ]). *If $Q \lll x^{\vartheta}$ and $A \geq 1$ is fixed, then*

$$\sum_{q \leq Q} \sup_{a \in (\mathbb{Z}/q\mathbb{Z})^{\times}} \left| \Delta \left(\Lambda \mathbf{1}_{[x,2x]}; a\,(q) \right) \right| \ll x \log^{-A} x. \tag{4}$$

In [22], it was conjectured that EH[ϑ] held for all $0 < \vartheta < 1$. (The conjecture fails at the endpoint case $\vartheta = 1$; see [23,24] for a more precise statement.) The following classical result of Bombieri [25] and Vinogradov [26] remains the best partial result of the form EH[ϑ]:

Theorem 9 (Bombieri-Vinogradov theorem). *[25,26] EH[ϑ] holds for every fixed $0 < \vartheta < 1/2$.*

In [2], it was shown that any estimate of the form EH[ϑ] with some fixed $\vartheta > 1/2$ would imply the finiteness of H_1. While such an estimate remains unproven, it was observed by Motohashi and Pintz [27] and by Zhang [3] that a certain weakened version of EH[ϑ] would still suffice for this purpose. More precisely (and following the notation of our previous paper), let $\varpi, \delta > 0$ be fixed, and let MPZ[ϖ, δ] be the following claim:

Claim 10 (Motohashi-Pintz-Zhang estimate, MPZ[ϖ, δ]). *Let $I \subset \left[1, x^{\delta} \right]$ and $Q \lll x^{1/2+2\varpi}$. Let P_I denote the product of all the primes in I, and let $\mathcal{S}_I$ denote the square-free natural numbers whose prime factors lie in I. If the residue class $a\,(P_I)$ is primitive (and is allowed to depend on x), and $A \geq 1$ is fixed, then*

$$\sum_{\substack{q \leq Q \\ q \in \mathcal{S}_I}} \left| \Delta \left(\Lambda \mathbf{1}_{[x,2x]}; a\,(q) \right) \right| \ll x \log^{-A} x, \tag{5}$$

where the implied constant depends only on the fixed quantities (A, ϖ, δ), but not on a.

It is clear that EH$\left[\frac{1}{2} + 2\varpi \right]$ implies MPZ[ϖ, δ] whenever $\varpi, \delta \geq 0$. The first non-trivial estimate of the form MPZ[ϖ, δ] was established by Zhang [3], who (essentially) obtained MPZ[ϖ, δ] whenever $0 \leq \varpi, \delta < \frac{1}{1,168}$. In [Theorem 2.17], we improved this result to the following.

Theorem 11. *MPZ[ϖ, δ] holds for every fixed $\varpi, \delta \geq 0$ with $600\varpi + 180\delta < 7$.*

In fact, a stronger result was established, in which the moduli q were assumed to be *densely divisible* rather than smooth, but we will not exploit such improvements here. For our application, the most important thing is to get ϖ as large as possible; in particular, Theorem 11 allows one to get ϖ arbitrarily close to $\frac{7}{600} \approx 0.01167$.

In this paper, we will also study the following generalization of the Elliott-Halberstam conjecture:

Claim 12 (Generalized Elliott-Halberstam conjecture, GEH[ϑ]). *Let $\varepsilon > 0$ and $A \geq 1$ be fixed. Let N, M be quantities such that $x^{\varepsilon} \lll N, M \lll x^{1-\varepsilon}$ with $NM \asymp x$, and let*

$\alpha, \beta : \mathbb{N} \to \mathbb{R}$ *be sequences supported on* $[N, 2N]$ *and* $[M, 2M]$, *respectively, such that one has the pointwise bound*

$$|\alpha(n)| \ll \tau(n)^{O(1)} \log^{O(1)} x; \quad |\beta(m)| \ll \tau(m)^{O(1)} \log^{O(1)} x; \tag{6}$$

for all natural numbers n, m. *Suppose also that* β *obeys the Siegel-Walfisz type bound*

$$\left| \Delta \left(\beta \mathbf{1}_{(\cdot, r) = 1}; a\ (q) \right) \right| \ll \tau(qr)^{O(1)} M \log^{-A} x \tag{7}$$

for any $q, r \geq 1$, *any fixed* A, *and any primitive residue class* $a\ (q)$. *Then for any* $Q \lll x^{\vartheta}$, *we have*

$$\sum_{q \leq Q} \sup_{a \in (\mathbb{Z}/q\mathbb{Z})^{\times}} \left| \Delta \left(\alpha \star \beta; a\ (q) \right) \right| \ll x \log^{-A} x. \tag{8}$$

In [28, Conjecture 1], it was essentially conjectured[b] that GEH[ϑ] was true for all $0 < \vartheta < 1$. This is stronger than the Elliott-Halberstam conjecture:

Proposition 13. *For any fixed* $0 < \vartheta < 1$, GEH[ϑ] *implies* EH[ϑ].

Proof. (Sketch) As this argument is standard, we give only a brief sketch. Let $A > 0$ be fixed. For $n \in [x, 2x]$, we have Vaughan's identity[c] [29]

$$\Lambda(n) = \mu_< \star L(n) - \mu_< \star \Lambda_< \star 1(n) + \mu_{\geq} \star \Lambda_{\geq} \star 1(n),$$

where $L(n) := \log(n)$ and

$$\Lambda_{\geq}(n) := \Lambda(n)\mathbf{1}_{n \geq x^{1/3}}, \quad \Lambda_<(n) := \Lambda(n)\mathbf{1}_{n < x^{1/3}} \tag{9}$$

$$\mu_{\geq}(n) := \mu(n)\mathbf{1}_{n \geq x^{1/3}}, \quad \mu_<(n) := \mu(n)\mathbf{1}_{n < x^{1/3}}. \tag{10}$$

By decomposing each of the functions $\mu_<$, $\mu_{\geq}$, 1, $\Lambda_<$, $\Lambda_{\geq}$ into $O\left(\log^{A+1} x \right)$ functions supported on intervals of the form $[N, (1 + \log^{-A} x)N]$, and discarding those contributions which meet the boundary of $[x, 2x]$ (cf. [3,28,30,31]), and using GEH[ϑ] (with A replaced by a much larger fixed constant A') to control all remaining contributions, we obtain the claim (using the Siegel-Walfisz theorem; see, e.g. [32, Satz 4] or [33, Th. 5.29]). $\qquad \square$

By modifying the proof of the Bombieri-Vinogradov theorem, Motohashi [34] established the following generalization of that theorem:

Theorem 14 (Generalized Bombieri-Vinogradov theorem). [34] GEH[ϑ] *holds for every fixed* $0 < \vartheta < 1/2$.

One could similarly describe a generalization of the Motohashi-Pintz-Zhang estimate MPZ[ϖ, δ], but unfortunately, the arguments in [3] or Theorem 11 do not extend to this setting unless one is in the 'Type I/Type II' case in which N, M are constrained to be somewhat close to $x^{1/2}$, or if one has 'Type III' structure to the convolution $\alpha \star \beta$, in the sense that it can refactored as a convolution involving several 'smooth' sequences. In any event, our analysis would not be able to make much use of such incremental improvements to GEH[ϑ], as we only use this hypothesis effectively in the case when ϑ is very close to 1. In particular, we will not directly use Theorem 14 in this paper.

Outline of the key ingredients

In this section, we describe the key subtheorems used in the proof of Theorem 4, with the proofs of these subtheorems mostly being deferred to later sections.

We begin with a weak version of the Dickson-Hardy-Littlewood prime tuples conjecture [1], which (following Pintz [35]) we refer to as [k, j]. Recall that for any $k \in \mathbb{N}$, an *admissible k-tuple* is a tuple $\mathcal{H} = (h_1, \ldots, h_k)$ of k increasing integers $h_1 < \ldots < h_k$ which avoids at least one residue class $a_p \ (p) := \{a_p + np : n \in \mathbb{Z}\}$ for every p. For instance, $(0, 2, 6)$ is an admissible 3-tuple, but $(0, 2, 4)$ is not.

For any $k \geq j \geq 2$, we let DHL[k; j] denote the following claim:

Claim 15 (Weak Dickson-Hardy-Littlewood conjecture, DHL[k; j]). *For any admissible k-tuple $\mathcal{H} = (h_1, \ldots, h_k)$, there exist infinitely many translates $n + \mathcal{H} = (n + h_1, \ldots, n + h_k)$ of $\mathcal{H}$ which contain at least j primes.*

The full Dickson-Hardy-Littlewood conjecture is then the assertion that DHL[k; k] holds for all $k \geq 2$. In our analysis, we will focus on the case when j is much smaller than k; in fact, j will be of the order of $\log k$.

For any k, let $H(k)$ denote the minimal diameter $h_k - h_1$ of an admissible k-tuple; thus for instance, $H(3) = 6$. It is clear that for any natural numbers $m \geq 1$ and $k \geq m + 1$, the claim DHL[$k; m + 1$] implies that $H_m \leq H(k)$ (and the claim DHL[k; k] would imply that $H_{k-1} = H(k)$). We will therefore deduce Theorem 4 from a number of claims of the form DHL[k; j]. More precisely, we have

Theorem 16. *Unconditionally, we have the following claims:*

  (i)   DHL[50; 2].
 (ii)   DHL[35, 410; 3].
(iii)   DHL[1, 649, 821; 4].
(iv)   DHL[75, 845, 707; 5].
 (v)   DHL[3, 473, 955, 908; 6].
(vi)   DHL[k; m + 1] *whenever $m \geq 1$ and $k \geq C \exp\left(\left(4 - \frac{28}{157}\right)m\right)$ for some sufficiently large absolute (and effective) constant C.*

Assume the Elliott-Halberstam conjecture EH[θ] for all $0 < \theta < 1$. Then, we have the following improvements:

 (vii)   DHL[54; 3].
(viii)   DHL[5, 511; 4].
 (ix)   DHL[41, 588; 5].
  (x)   DHL[309, 661; 6].
 (xi)   DHL[k; m + 1] *whenever $m \geq 1$ and $k \geq C \exp(2m)$ for some sufficiently large absolute (and effective) constant C.*

Assume the generalized Elliott-Halberstam conjecture GEH[θ] for all $0 < \theta < 1$. Then

 (xii)   DHL[3; 2].
(xiii)   DHL[51; 3].

Theorem 4 then follows from Theorem 16 and the following bounds on $H(k)$ (ordered by increasing value of k):

Theorem 17 (Bounds on $H(k)$)**.**

- (xii) $H(3) = 6$.
- (i) $H(50) = 246$.
- (xiii) $H(51) = 252$.
- (vii) $H(54) = 270$.
- (viii) $H(5, 511) \le 52, 116$.
- (ii) $H(35, 410) \le 398, 130$.
- (ix) $H(41, 588) \le 474, 266$.
- (x) $H(309, 661) \le 4, 137, 854$.
- (iii) $H(1, 649, 821) \le 24, 797, 814$.
- (iv) $H(75, 845, 707) \le 1, 431, 556, 072$.
- (v) $H(3, 473, 955, 908) \le 80, 550, 202, 480$.
- (vi), (xi) In the asymptotic limit $k \to \infty$, one has $H(k) \le k \log k + k \log \log k - k + o(k)$, with the bounds on the decay rate $o(k)$ being effective.

We prove Theorem 17 in the 'Narrow admissible tuples' section. In the opposite direction, an application of the Brun-Titchmarsh theorem gives $H(k) \ge \left(\frac{1}{2} + o(1)\right) k \log k$ as $k \to \infty$ (see [4, §3.9] for this bound, as well as with some slight refinements).

The proof of Theorem 16 follows the Goldston-Pintz-Yıldırım strategy that was also used in all previous progress on this problem (e.g. [2,3,5,27]), namely that of constructing a sieve function adapted to an admissible k-tuple with good properties. More precisely, we set

$$w := \log \log \log x$$

and

$$W := \prod_{p \le w} p,$$

and observe the crude bound

$$W \ll \log \log^{O(1)} x. \tag{11}$$

We have the following simple 'pigeonhole principle' criterion for DHL$[k; m + 1]$ (cf. [Lemma 4.1], though the normalization here is slightly different):

Lemma 18 (Criterion for DHL)**.** *Let $k \ge 2$ and $m \ge 1$ be fixed integers and define the normalization constant*

$$B := \frac{\varphi(W)}{W} \log x. \tag{12}$$

Suppose that for each fixed admissible k-tuple $(h_1, \dots, h_k)$ and each residue class b (W) such that $b + h_i$ is coprime to W for all $i = 1, \dots, k$, one can find a non-negative weight function $v \colon \mathbb{N} \to \mathbb{R}^+$ and fixed quantities $\alpha > 0$ and $\beta_1, \dots, \beta_k \ge 0$, such that one has the asymptotic upper bound

$$\sum_{\substack{x \le n \le 2x \\ n = b \ (W)}} v(n) \le (\alpha + o(1)) B^{-k} \frac{x}{W}, \tag{13}$$

the asymptotic lower bound

$$\sum_{\substack{x \leq n \leq 2x \\ n \equiv b\ (W)}} \nu(n)\theta(n + h_i) \geq (\beta_i - o(1))B^{1-k}\frac{x}{\varphi(W)} \tag{14}$$

for all $i = 1, \ldots, k$, and the key inequality

$$\frac{\beta_1 + \cdots + \beta_k}{\alpha} > m. \tag{15}$$

Then, DHL$[k; m + 1]$ *holds.*

Proof. Let $(h_1, \ldots, h_k)$ be a fixed admissible k-tuple. Since it is admissible, there is at least one residue class $b\ (W)$ such that $(b + h_i, W) = 1$ for all $h_i \in \mathcal{H}$. For an arithmetic function ν as in the lemma, we consider the quantity

$$N := \sum_{\substack{x \leq n \leq 2x \\ n \equiv b\ (W)}} \nu(n)\left(\sum_{i=1}^{k} \theta(n + h_i) - m\log 3x\right).$$

Combining (13) and (14), we obtain the lower bound

$$N \geq (\beta_1 + \cdots + \beta_k - o(1))B^{1-k}\frac{x}{\varphi(W)} - (m\alpha + o(1))B^{-k}\frac{x}{W}\log 3x.$$

From (12) and the crucial condition (15), it follows that $N > 0$ if x is sufficiently large.

On the other hand, the sum

$$\sum_{i=1}^{k} \theta(n + h_i) - m\log 3x$$

can be positive only if $n + h_i$ is prime for *at least $m + 1$ indices* $i = 1, \ldots, k$. We conclude that, for all sufficiently large x, there exists some integer $n \in [x, 2x]$ such that $n + h_i$ is prime for at least $m + 1$ values of $i = 1, \ldots, k$.

Since $(h_1, \ldots, h_k)$ is an arbitrary admissible k-tuple, DHL$[k; m + 1]$ follows. □

The objective is then to construct non-negative weights ν whose associated ratio $\frac{\beta_1 + \cdots + \beta_k}{\alpha}$ has provable lower bounds that are as large as possible. Our sieve majorants will be a variant of the multidimensional Selberg sieves used in [5]. As with all Selberg sieves, the ν are constructed as the square of certain (signed) divisor sums. The divisor sums we will use will be finite linear combinations of products of 'one-dimensional' divisor sums. More precisely, for any fixed smooth compactly supported function $F : [0, +\infty) \to \mathbb{R}$, define the divisor sum $\lambda_F : \mathbb{Z} \to \mathbb{R}$ by the formula

$$\lambda_F(n) := \sum_{d | n} \mu(d)F(\log_x d) \tag{16}$$

where $\log_x$ denotes the base x logarithm

$$\log_x n := \frac{\log n}{\log x}. \tag{17}$$

One should think of λ_F as a smoothed out version of the indicator function to numbers n which are 'almost prime' in the sense that they have no prime factors less than x^ε for some small fixed $\varepsilon > 0$ (see Proposition 14 for a more rigorous version of this heuristic).

The functions ν we will use will take the form

$$\nu(n) = \left(\sum_{j=1}^{J} c_j \lambda_{F_{j,1}}(n+h_1) \dots \lambda_{F_{j,k}}(n+h_k) \right)^2 \tag{18}$$

for some fixed natural number J, fixed coefficients $c_1, \dots, c_J \in \mathbb{R}$ and fixed smooth compactly supported functions $F_{j,i} : [0, +\infty) \to \mathbb{R}$ with $j = 1, \dots, J$ and $i = 1, \dots, k$. (One can of course absorb the constant c_j into one of the $F_{j,i}$ if one wishes.) Informally, ν is a smooth restriction to those n for which $n + h_1, \dots, n + h_k$ are all almost prime.

Clearly, ν is a (positive-definite) linear combination of functions of the form

$$n \mapsto \prod_{i=1}^{k} \lambda_{F_i}(n+h_i) \lambda_{G_i}(n+h_i)$$

for various smooth functions $F_1, \dots, F_k, G_1, \dots, G_k : [0, +\infty) \to \mathbb{R}$. The sum appearing in (13) can thus be decomposed into linear combinations of sums of the form

$$\sum_{\substack{x \leq n \leq 2x \\ n = b \ (W)}} \prod_{i=1}^{k} \lambda_{F_i}(n+h_i) \lambda_{G_i}(n+h_i). \tag{19}$$

Also, since from (16) we clearly have

$$\lambda_F(n) = F(0) \tag{20}$$

when $n \geq x$ is prime and F is supported on $[0, 1]$, the sum appearing in (14) can be similarly decomposed into linear combinations of sums of the form

$$\sum_{\substack{x \leq n \leq 2x \\ n = b \ (W)}} \theta(n+h_i) \prod_{1 \leq i' \leq k; i' \neq i} \lambda_{F_{i'}}(n+h_{i'}) \lambda_{G_{i'}}(n+h_{i'}). \tag{21}$$

To estimate the sums (21), we use the following asymptotic, proven in the 'Multidimensional Selberg sieves' section. For each compactly supported $F : [0, +\infty) \to \mathbb{R}$, let

$$S(F) := \sup\{x \geq 0 : F(x) \neq 0\} \tag{22}$$

denote the upper range of the support of F (with the convention that $S(0) = 0$).

Theorem 19 (Asymptotic for prime sums). *Let $k \geq 2$ be fixed, let $(h_1, \dots, h_k)$ be a fixed admissible k-tuple, and let $b \ (W)$ be such that $b + h_i$ is coprime to W for each $i = 1, \dots, k$. Let $1 \leq i_0 \leq k$ be fixed, and for each $1 \leq i \leq k$ distinct from i_0, let $F_i, G_i : [0, +\infty) \to \mathbb{R}$ be fixed smooth compactly supported functions. Assume one of the following hypotheses:*

(i) *(Elliott-Halberstam) There exists a fixed $0 < \vartheta < 1$ such that EH$[\vartheta]$ holds and such that*

$$\sum_{1 \leq i \leq k; i \neq i_0} (S(F_i) + S(G_i)) < \vartheta. \tag{23}$$

(ii) *(Motohashi-Pintz-Zhang) There exists fixed $0 \leq \varpi < 1/4$ and $\delta > 0$ such that MPZ$[\varpi, \delta]$ holds and such that*

$$\sum_{1 \leq i \leq k; i \neq i_0} (S(F_i) + S(G_i)) < \frac{1}{2} + 2\varpi \tag{24}$$

and

$$\max_{1 \le i \le k; i \ne i_0} \{S(F_i), S(G_i)\} < \delta. \tag{25}$$

Then, we have

$$\sum_{\substack{x \le n \le 2x \\ n = b \, (W)}} \theta(n + h_{i_0}) \prod_{1 \le i \le k; i \ne i_0} \lambda_{F_i}(n + h_i)\lambda_{G_i}(n + h_i) = (c + o(1))B^{1-k}\frac{x}{\varphi(W)} \tag{26}$$

where

$$c := \prod_{1 \le i \le k; i \ne i_0} \left(\int_0^1 F_i'(t_i)G_i'(t_i) \, dt_i \right).$$

Here of course F' denotes the derivative of F.

To estimate the sums (19), we use the following asymptotic, also proven in the 'Multidimensional Selberg sieves' section.

Theorem 20 (Asymptotic for non-prime sums). *Let $k \ge 1$ be fixed, let $(h_1, \dots, h_k)$ be a fixed admissible k-tuple, and let b (W) be such that $b + h_i$ is coprime to W for each $i = 1, \dots, k$. For each fixed $1 \le i \le k$, let $F_i, G_i : [0, +\infty) \to \mathbb{R}$ be fixed smooth compactly supported functions. Assume one of the following hypotheses:*

(i) *(Trivial case) One has*

$$\sum_{i=1}^k (S(F_i) + S(G_i)) < 1. \tag{27}$$

(ii) *(Generalized Elliott-Halberstam) There exists a fixed $0 < \vartheta < 1$ and $i_0 \in \{1, \dots, k\}$ such that $\text{GEH}[\vartheta]$ holds, and*

$$\sum_{1 \le i \le k; i \ne i_0} (S(F_i) + S(G_i)) < \vartheta. \tag{28}$$

Then, we have

$$\sum_{\substack{x \le n \le 2x \\ n = b \, (W)}} \prod_{i=1}^k \lambda_{F_i}(n + h_i)\lambda_{G_i}(n + h_i) = (c + o(1))B^{-k}\frac{x}{W}, \tag{29}$$

where

$$c := \prod_{i=1}^k \left(\int_0^1 F_i'(t_i)G_i'(t_i) \, dt_i \right). \tag{30}$$

A key point in (ii) is that no upper bound on $S(F_{i_0})$ or $S(G_{i_0})$ is required (although, as we will see in the 'The generalized Elliott-Halberstam case' section, the result is a little easier to prove when one has $S(F_{i_0}) + S(G_{i_0}) < 1$). This flexibility in the F_{i_0}, G_{i_0} functions will be particularly crucial to obtain part (xii) of Theorem 16 and Theorem 4.

Remark 21. Theorems 19 and 20 can be viewed as probabilistic assertions of the following form: if n is chosen uniformly at random from the set $\{x \le n \le 2x : n = b \, (W)\}$, then the random variables $\theta(n + h_i)$ and $\lambda_{F_j}(n + h_j)\lambda_{G_j}(n + h_j)$ for $i, j = 1, \dots, k$ have mean $(1 + o(1))\frac{W}{\varphi(W)}$ and $\left(\int_0^1 F_j'(t)G_j'(t) \, dt + o(1) \right) B^{-1}$, respectively, and furthermore,

these random variables enjoy a limited amount of independence, except for the fact (as can be seen from (20)) that $\theta(n + h_i)$ and $\lambda_{F_i}(n + h_i)\lambda_{G_i}(n + h_i)$ are highly correlated. Note though that we do not have asymptotics for any sum which involves two or more factors of θ, as such estimates are of a difficulty at least as great as that of the twin prime conjecture (which is equivalent to the divergence of the sum $\sum_n \theta(n)\theta(n + 2)$).

Theorems 19 and 20 may be combined with Lemma 18 to reduce the task of establishing estimates of the form DHL$[k; m + 1]$ to that of establishing certain variational problems. For instance, in the 'Proof of Theorem 22' section, we reprove the following result of Maynard ([5, Proposition 4.2]):

Theorem 22 (Sieving on the standard simplex). *Let $k \geq 2$ and $m \geq 1$ be fixed integers. For any fixed compactly supported square-integrable function $F : [0, +\infty)^k \to \mathbb{R}$, define the functionals*

$$I(F) := \int_{[0,+\infty)^k} F(t_1, \ldots, t_k)^2 \, dt_1 \ldots t_k \tag{31}$$

and

$$J_i(F) := \int_{[0,+\infty)^{k-1}} \left(\int_0^\infty F(t_1, \ldots, t_k) \, dt_i \right)^2 dt_1 \ldots dt_{i-1} dt_{i+1} \ldots dt_k \tag{32}$$

for $i = 1, \ldots, k$, and let M_k be the supremum

$$M_k := \sup \frac{\sum_{i=1}^k J_i(F)}{I(F)} \tag{33}$$

over all square integrable functions F that are supported on the simplex

$$\mathcal{R}_k := \left\{ (t_1, \ldots, t_k) \in [0, +\infty)^k : t_1 + \cdots + t_k \leq 1 \right\}$$

and are not identically zero (up to almost everywhere equivalence, of course). Suppose that there is a fixed $0 < \vartheta < 1$ such that EH$[\vartheta]$ holds and such that

$$M_k > \frac{2m}{\vartheta}.$$

Then, DHL$[k; m + 1]$ holds.

Parts (vii)-(xi) of Theorem 16 (and hence Theorem 4) are then immediate from the following results, proven in the 'Asymptotic analysis' and 'The case of small and medium dimension' sections, and ordered by increasing value of k:

Theorem 23 (Lower bounds on M_k).

(vii) $M_{54} > 4.00238$.

(viii) $M_{5,511} > 6$.

(ix) $M_{41,588} > 8$.

(x) $M_{309,661} > 10$.

(xi) One has $M_k \geq \log k - C$ for all $k \geq C$, where C is an absolute (and effective) constant.

For the sake of comparison, in ([5, Proposition 4.3]), it was shown that $M_5 > 2$, $M_{105} > 4$, and $M_k \geq \log k - 2 \log \log k - 2$ for all sufficiently large k. As remarked in that paper, the sieves used on the bounded gap problem prior to the work in [5] would

essentially correspond, in this notation, to the choice of functions F of the special form $F(t_1, \ldots, t_k) := f(t_1 + \cdots + t_k)$, which severely limits the size of the ratio in (33) (in particular, the analogue of M_k in this special case cannot exceed 4, as shown in [36]).

In the converse direction, in Corollary 37, we will also show the upper bound $M_k \leq \frac{k}{k-1} \log k$ for all $k \geq 2$, which shows in particular that the bounds in (vii) and (xi) of the above theorem cannot be significantly improved. We remark that Theorem 23(vii) and the Bombieri-Vinogradov theorem also give a weaker version DHL[54; 2] of Theorem 16(i).

We also have a variant of Theorem 22 which can accept inputs of the form MPZ[ϖ, δ]:

Theorem 24 (Sieving on a truncated simplex). *Let $k \geq 2$ and $m \geq 1$ be fixed integers. Let $0 < \varpi < 1/4$ and $0 < \delta < 1/2$ be such that* MPZ[ϖ, δ] *holds. For any $\alpha > 0$, let $M_k^{[\alpha]}$ be defined as in* (33)*, but where the supremum now ranges over all square-integrable F supported in the truncated simplex*

$$\left\{ (t_1, \ldots, t_k) \in [0, \alpha]^k : t_1 + \cdots + t_k \leq 1 \right\} \tag{34}$$

and are not identically zero. If

$$M_k^{\left[\frac{\delta}{1/4+\varpi} \right]} > \frac{m}{1/4 + \varpi},$$

then DHL[$k; m + 1$] *holds.*

In the 'Asymptotic analysis' section, we will establish the following variant of Theorem 23, which when combined with Theorem 11, allows one to use Theorem 24 to establish parts (ii)-(vi) of Theorem 16 (and hence Theorem 4):

Theorem 25 (Lower bounds on $M_k^{[\alpha]}$).

(ii) *There exist $\delta, \varpi > 0$ with $600\varpi + 180\delta < 7$ and $M_{35\,410}^{\left[\frac{\delta}{1/4+\varpi} \right]} > \frac{2}{1/4+\varpi}$.*

(iii) *There exist $\delta, \varpi > 0$ with $600\varpi + 180\delta < 7$ and $M_{1\,649\,821}^{\left[\frac{\delta}{1/4+\varpi} \right]} > \frac{3}{1/4+\varpi}$.*

(iv) *There exist $\delta, \varpi > 0$ with $600\varpi + 180\delta < 7$ and $M_{75\,845\,707}^{\left[\frac{\delta}{1/4+\varpi} \right]} > \frac{4}{1/4+\varpi}$.*

(v) *There exist $\delta, \varpi > 0$ with $600\varpi + 180\delta < 7$ and $M_{3\,473\,955\,908}^{\left[\frac{\delta}{1/4+\varpi} \right]} > \frac{5}{1/4+\varpi}$.*

(vi) *For all $k \geq C$, there exist $\delta, \varpi > 0$ with $600\varpi + 180\delta < 7$, $\varpi \geq \frac{7}{600} - \frac{C}{\log k}$, and*

$$M_k^{\left[\frac{\delta}{1/4+\varpi} \right]} \geq \log k - C \text{ for some absolute (and effective) constant } C.$$

The implication is clear for (ii)-(v). For (vi), observe that from Theorem 25(vi), Theorem 11, and Theorem 24, we see that DHL[$k; m + 1$] holds whenever k is sufficiently large and

$$m \leq (\log k - C) \left(\frac{1}{4} + \frac{7}{600} - \frac{C}{\log k} \right)$$

which is in particular implied by

$$m \leq \frac{\log k}{4 - \frac{28}{157}} - C'$$

for some absolute constant C', giving Theorem 16(vi).

Now we give a more flexible variant of Theorem 22, in which the support of F is enlarged, at the cost of reducing the range of integration of the J_i.

Theorem 26 (Sieving on an epsilon-enlarged simplex). *Let $k \geq 2$ and $m \geq 1$ be fixed integers, and let $0 < \varepsilon < 1$ be fixed also. For any fixed compactly supported square-integrable function $F : [0, +\infty)^k \to \mathbb{R}$, define the functionals*

$$J_{i,1-\varepsilon}(F) := \int_{(1-\varepsilon) \cdot \mathcal{R}_{k-1}} \left(\int_0^\infty F(t_1, \ldots, t_k)\, dt_i \right)^2 dt_1 \ldots dt_{i-1} dt_{i+1} \ldots dt_k$$

for $i = 1, \ldots, k$, and let $M_{k,\varepsilon}$ be the supremum

$$M_{k,\varepsilon} := \sup \frac{\sum_{i=1}^k J_{i,1-\varepsilon}(F)}{I(F)}$$

over all square-integrable functions F that are supported on the simplex

$$(1 + \varepsilon) \cdot \mathcal{R}_k = \left\{ (t_1, \ldots, t_k) \in [0, +\infty)^k : t_1 + \cdots + t_k \leq 1 + \varepsilon \right\}$$

and are not identically zero. Suppose that there is a fixed $0 < \vartheta < 1$, such that one of the following two hypotheses hold:

(i) EH$[\vartheta]$ *holds, and $1 + \varepsilon < \frac{1}{\vartheta}$.*
(ii) GEH$[\vartheta]$ *holds, and $\varepsilon < \frac{1}{k-1}$.*

If

$$M_{k,\varepsilon} > \frac{2m}{\vartheta}$$

then DHL$[k; m + 1]$ *holds.*

We prove this theorem in the 'Proof of Theorem 26' section. We remark that due to the continuity of $M_{k,\varepsilon}$ in ε, the strict inequalities in (i) and (ii) of this theorem may be replaced by non-strict inequalities. Parts (i) and (xiii) of Theorem 16, and a weaker version DHL[4; 2] of part (xii), then follow from Theorem 9 and the following computations, proven in the 'Bounding $M_{k,\varepsilon}$ for medium k' and 'Bounding $M_{4,\varepsilon}$' sections:

Theorem 27 (Lower bounds on $M_{k,\varepsilon}$).

(i) $M_{50,1/25} > 4.0043$.
(xii') $M_{4,0.168} > 2.00558$.
(xiii) $M_{51,1/50} > 4.00156$.

We remark that computations in the proof of Theorem 27(xii') are simple enough that the bound may be checked by hand, without use of a computer. The computations used to establish the full strength of Theorem 16(xii) are however significantly more complicated.

In fact, we may enlarge the support of F further. We give a version corresponding to part (ii) of Theorem 26; there is also a version corresponding to part (i), but we will not give it here as we will not have any use for it.

Theorem 28 (Going beyond the epsilon enlargement). *Let $k \geq 2$ and $m \geq 1$ be fixed integers, let $0 < \vartheta < 1$ be a fixed quantity such that* GEH$[\vartheta]$ *holds, and let $0 < \varepsilon < \frac{1}{k-1}$ be fixed also. Suppose that there is a fixed non-zero square-integrable function*

$F : [0, +\infty)^k \to \mathbb{R}$ supported in $\frac{k}{k-1} \cdot \mathcal{R}_k$, such that for $i = 1, \ldots, k$, one has the vanishing marginal condition

$$\int_0^\infty F(t_1, \ldots, t_k) \, dt_i = 0 \tag{35}$$

whenever $t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_k \geq 0$ are such that

$$t_1 + \cdots + t_{i-1} + t_{i+1} + \cdots + t_k > 1 + \varepsilon.$$

Suppose that we also have the inequality

$$\frac{\sum_{i=1}^k J_{i,\varepsilon}(F)}{I(F)} > \frac{2m}{\vartheta}.$$

Then DHL[$k; m + 1$] holds.

This theorem is proven in the 'Proof of Theorem 28' section. Theorem 16(xii) is then an immediate consequence of Theorem 28 and the following numerical fact, established in the 'Three-dimensional cutoffs' section.

Theorem 29 (A piecewise polynomial cutoff). *Set* $\varepsilon := \frac{1}{4}$. *Then, there exists a piecewise polynomial function* $F : [0, +\infty)^3 \to \mathbb{R}$ *supported on the simplex*

$$\frac{3}{2} \cdot \mathcal{R}_3 = \left\{ (t_1, t_2, t_3) \in [0, +\infty)^3 : t_1 + t_2 + t_3 \leq \frac{3}{2} \right\}$$

and symmetric in the t_1, t_2, t_3 *variables, such that* F *is not identically zero and obeys the vanishing marginal condition*

$$\int_0^\infty F(t_1, t_2, t_3) \, dt_3 = 0$$

whenever $t_1, t_2 \geq 0$ *with* $t_1 + t_2 > 1 + \varepsilon$ *and such that*

$$\frac{3 \int_{t_1+t_2 \leq 1-\varepsilon} \left(\int_0^\infty F(t_1, t_2, t_3) \, dt_3 \right)^2 \, dt_1 dt_2}{\int_{[0,\infty)^3} F(t_1, t_2, t_3)^2 \, dt_1 dt_2 dt_3} > 2.$$

There are several other ways to combine Theorems 19 and 20 with equidistribution theorems on the primes to obtain results of the form DHL[$k; m+1$], but all of our attempts to do so either did not improve the numerology or else were numerically infeasible to implement.

Multidimensional Selberg sieves

In this section, we prove Theorems 19 and 20. A key asymptotic used in both theorems is the following:

Lemma 30 (Asymptotic). *Let* $k \geq 1$ *be a fixed integer, and let* N *be a natural number coprime to* W *with* $\log N = O\left(\log^{O(1)} x\right)$. *Let* $F_1, \ldots, F_k, G_1, \ldots, G_k : [0, +\infty) \to \mathbb{R}$ *be fixed smooth compactly supported functions. Then,*

$$\sum_{\substack{d_1, \ldots, d_k, d_1', \ldots, d_k' \\ [d_1, d_1'], \ldots, [d_k, d_k'], W, N \text{coprime}}} \prod_{j=1}^k \frac{\mu\left(d_j\right) \mu\left(d_j'\right) F_j\left(\log_x d_j\right) G_j\left(\log_x d_j'\right)}{\left[d_j, d_j'\right]} = (c + o(1)) B^{-k} \frac{N^k}{\varphi(N)^k}$$

$$\tag{36}$$

where B was defined in (12), and

$$c := \prod_{j=1}^{k} \int_0^{\infty} F_j'(t_j) G_j'(t_j) \, dt_j.$$

The same claim holds if the denominators $\left[d_j, d_j'\right]$ are replaced by $\varphi\left(\left[d_j, d_j'\right]\right)$.

Such asymptotics are standard in the literature (see, e.g. [37] for some similar computations). In older literature, it is common to establish these asymptotics via contour integration (e.g. via Perron's formula), but we will use the Fourier analytic approach here. Of course, both approaches ultimately use the same input, namely the simple pole of the Riemann zeta function at $s = 1$.

Proof. We begin with the first claim. For $j = 1, \dots, k$, the functions $t \mapsto e^t F_j(t)$, $t \mapsto e^t G_j(t)$ may be extended to smooth compactly supported functions on all of $\mathbb{R}$, and so we have Fourier expansions

$$e^t F_j(t) = \int_{\mathbb{R}} e^{-it\xi} f_j(\xi) \, d\xi \tag{37}$$

and

$$e^t G_j(t) = \int_{\mathbb{R}} e^{-it\xi} g_j(\xi) \, d\xi$$

for some fixed functions $f_j, g_j : \mathbb{R} \to \mathbb{C}$ that are smooth and rapidly decreasing in the sense that $f_j(\xi), g_j(\xi) = O\left((1 + |\xi|)^{-A}\right)$ for any fixed $A > 0$ and all $\xi \in \mathbb{R}$ (here the implied constant is independent of ξ and depends only on A).

We may thus write

$$F_j\left(\log_x d_j\right) = \int_{\mathbb{R}} \frac{f_j(\xi_j)}{d_j^{\frac{1+i\xi_j}{\log x}}} \, d\xi_j$$

and

$$G_j\left(\log_x d_j'\right) = \int_{\mathbb{R}} \frac{g_j\left(\xi_j'\right)}{\left(d_j'\right)^{\frac{1+i\xi_j'}{\log x}}} \, d\xi_j'$$

for all $d_j, d_j' \geq 1$. We note that

$$\sum_{d_j, d_j'} \frac{\left|\mu\left(d_j\right) \mu\left(d_j'\right)\right|}{\left[d_j, d_j'\right] d_j^{1/\log x} \left(d_j'\right)^{1/\log x}} = \prod_p \left(1 + \frac{2}{p^{1+1/\log x}} + \frac{1}{p^{1+2/\log x}}\right) \leq \exp(O(\log\log x)).$$

Therefore, if we substitute the Fourier expansions into the left-hand side of (36), the resulting expression is absolutely convergent. Thus, we can apply Fubini's theorem, and the left-hand side of (36) can thus be rewritten as

$$\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} K\left(\xi_1, \dots, \xi_k, \xi_1', \dots, \xi_k'\right) \prod_{j=1}^{k} f_j\left(\xi_j\right) g_j\left(\xi_j'\right) d\xi_j d\xi_j', \tag{38}$$

where

$$K(\xi_1, \dots, \xi_k, \xi_1', \dots, \xi_k') := \sum_{\substack{d_1, \dots, d_k, d_1', \dots, d_k' \\ \left[d_1, d_1'\right], \dots, \left[d_k, d_k'\right], W, N \text{ coprime}}} \prod_{j=1}^{k} \frac{\mu\left(d_j\right) \mu\left(d_j'\right)}{\left[d_j, d_j'\right] d_j^{\frac{1+i\xi_j}{\log x}} \left(d_j'\right)^{\frac{1+i\xi_j'}{\log x}}}.$$

This latter expression factorizes as an Euler product

$$K = \prod_{p \nmid WN} K_p,$$

where the local factors K_p are given by

$$K_p\left(\xi_1, \ldots, \xi_k, \xi_1', \ldots, \xi_k'\right) := 1 + \frac{1}{p} \sum_{\substack{d_1, \ldots, d_k, d_1', \ldots, d_k' \\ [d_1, \ldots, d_k, d_1', \ldots, d_k'] = p \\ [d_1, d_1'], \ldots, [d_k, d_k'] \text{coprime}}} \prod_{j=1}^{k} \frac{\mu\left(d_j\right)\mu\left(d_j'\right)}{d_j^{\frac{1+i\xi_j}{\log x}}\left(d_j'\right)^{\frac{1+i\xi_j'}{\log x}}}. \tag{39}$$

We can estimate each Euler factor as

$$K_p\left(\xi_1, \ldots, \xi_k, \xi_1', \ldots, \xi_k'\right) = \left(1 + O\left(\frac{1}{p^2}\right)\right) \prod_{j=1}^{k} \frac{\left(1 - p^{-1-\frac{1+i\xi_j}{\log x}}\right)\left(1 - p^{-1-\frac{1+i\xi_j'}{\log x}}\right)}{1 - p^{-1-\frac{2+i\xi_j+i\xi_j'}{\log x}}}. \tag{40}$$

Since

$$\prod_{p:p>w} \left(1 + O\left(\frac{1}{p^2}\right)\right) = 1 + o(1),$$

we have

$$K\left(\xi_1, \ldots, \xi_k, \xi_1', \ldots, \xi_k'\right) = (1 + o(1)) \prod_{j=1}^{k} \frac{\zeta_{WN}\left(1 + \frac{2+i\xi_j+i\xi_j'}{\log x}\right)}{\zeta_{WN}\left(1 + \frac{1+i\xi_j}{\log x}\right)\zeta_{WN}\left(1 + \frac{1+i\xi_j'}{\log x}\right)}$$

where the modified zeta function ζ_{WN} is defined by the formula

$$\zeta_{WN}(s) := \prod_{p \nmid WN} \left(1 - \frac{1}{p^s}\right)^{-1}$$

for $\Re(s) > 1$.

For $\Re(s) \geq 1 + \frac{1}{\log x}$, we have the crude bounds

$$|\zeta_{WN}(s)|, |\zeta_{WN}(s)|^{-1} \leq \prod_{p} \left(1 + \frac{1}{p^{1+1/\log x}} + O\left(\frac{1}{p^2}\right)\right)$$

$$\ll \exp\left(\sum_{p} \frac{1}{p^{1+1/\log x}}\right)$$

$$\leq \exp(\log\log x + O(1))$$

$$\ll \log x.$$

Thus,

$$K\left(\xi_1, \ldots, \xi_k, \xi_1', \ldots, \xi_k'\right) = O\left(\log^{3k} x\right).$$

Combining this with the rapid decrease of f_j, g_j, we see that the contribution to (38) outside of the cube $\left\{\max\left(\xi_1, \ldots, \xi_k, \xi_1', \ldots, \xi_k'\right) \leq \sqrt{\log x}\right\}$ (say) is negligible. Thus, it will suffice to show that

$$\int_{-\sqrt{\log x}}^{\sqrt{\log x}} \cdots \int_{-\sqrt{\log x}}^{\sqrt{\log x}} K\left(\xi_1, \ldots, \xi_k, \xi_1', \ldots, \xi_k'\right) \prod_{j=1}^{k} f_j\left(\xi_j\right) g_j\left(\xi_j'\right) d\xi_j d\xi_j' = (c + o(1))B^{-k}\frac{N^k}{\varphi(N)^k}.$$

When $|\xi_j| \le \sqrt{\log x}$, we see from the simple pole of the Riemann zeta function $\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}$ at $s = 1$ that

$$\zeta\left(1 + \frac{1 + i\xi_j}{\log x}\right) = (1 + o(1))\frac{\log x}{1 + i\xi_j}.$$

For $-\sqrt{\log x} \le \xi_j \le \sqrt{\log x}$, we see that

$$1 - \frac{1}{p^{1 + \frac{1 + i\xi_j}{\log x}}} = 1 - \frac{1}{p} + O\left(\frac{\log p}{p\sqrt{\log x}}\right).$$

Since $\log WN \ll \log^{O(1)} x$, this gives

$$\prod_{p|WN}\left(1 - \frac{1}{p^{1 + \frac{1 + i\xi_j}{\log x}}}\right) = \frac{\varphi(WN)}{WN}\exp\left(O\left(\sum_{p|WN}\frac{\log p}{p\sqrt{\log x}}\right)\right) = (1 + o(1))\frac{\varphi(WN)}{WN},$$

since the sum is maximized when WN is composed only of primes $p \ll \log^{O(1)} x$. Thus,

$$\zeta_{WN}\left(1 + \frac{1 + i\xi_j}{\log x}\right) = \frac{(1 + o(1))B\varphi(N)}{(1 + i\xi_j)N},$$

similarly with $1 + i\xi_j$ replaced by $1 + i\xi'_j$ or $2 + i\xi_j + i\xi'_j$. We conclude that

$$K\left(\xi_1, \ldots, \xi_k, \xi'_1, \ldots, \xi'_k\right) = (1 + o(1))B^{-k}\frac{N^k}{\varphi(N)^k}\prod_{j=1}^{k}\frac{\left(1 + i\xi_j\right)\left(1 + i\xi'_j\right)}{2 + i\xi_j + i\xi'_j}. \tag{41}$$

Therefore, it will suffice to show that

$$\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \prod_{j=1}^{k}\frac{\left(1 + i\xi_j\right)\left(1 + i\xi'_j\right)}{2 + i\xi_j + i\xi'_j}f_j(\xi_j)g_j\left(\xi'_j\right) d\xi_j d\xi'_j = c,$$

since the errors caused by the $1 + o(1)$ multiplicative factor in (41) or the truncation $|\xi_j|, |\xi'_j| \le \sqrt{\log x}$ can be seen to be negligible using the rapid decay of f_j, g_j. By Fubini's theorem, it suffices to show that

$$\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{(1 + i\xi)(1 + i\xi')}{2 + i\xi + i\xi'}f_j(\xi)g_j(\xi') \, d\xi \, d\xi' = \int_0^{+\infty} F'_j(t)G'_j(t) \, dt$$

for each $j = 1, \ldots, k$. But from dividing (37) by e^t and differentiating under the integral sign, we have

$$F'_j(t) = -\int_{\mathbb{R}}(1 + i\xi)e^{-t(1 + i\xi)}f_j(\xi) \, d\xi,$$

and the claim then follows from Fubini's theorem.

Finally, suppose that we replace $\left[d_j, d'_j\right]$ with $\varphi\left(\left[d_j, d'_j\right]\right)$. An inspection of the above argument shows that the only change that occurs is that the $\frac{1}{p}$ term in (39) is replaced by $\frac{1}{p-1}$; but this modification may be absorbed into the $1 + O\left(\frac{1}{p^2}\right)$ factor in (40), and the rest of the argument continues as before. $\qquad\square$

The trivial case

We can now prove the easiest case of the two theorems, namely case (i) of Theorem 20; a closely related estimate also appears in ([5, Lemma 6.2]). We may assume that x is suf-

ficiently large depending on all fixed quantities. By (16), the left-hand side of (29) may be expanded as

$$
\sum_{d_1,\ldots,d_k,d_1',\ldots,d_k'} \left(\prod_{i=1}^{k} \mu(d_i)\mu\left(d_i'\right) F_i\left(\log_x d_i\right) G_i\left(\log_x d_i'\right) \right) S\left(d_1,\ldots,d_k,d_1',\ldots,d_k'\right)
$$

(42)

where

$$
S\left(d_1,\ldots,d_k,d_1',\ldots,d_k'\right) := \sum_{\substack{x \leq n \leq 2x \\ n=b\ (W) \\ n+h_i=0\ ([d_i,d_i'])\ \forall i}} 1.
$$

By hypothesis, $b + h_i$ is coprime to W for all $i = 1,\ldots,k$, and $|h_i - h_j| < w$ for all distinct i, j. Thus, $S\left(d_1,\ldots,d_k,d_1',\ldots,d_k'\right)$ vanishes unless the $[d_i, d_i']$ are coprime to each other and to W. In this case, $S\left(d_1,\ldots,d_k,d_1',\ldots,d_k'\right)$ is summing the constant function 1 over an arithmetic progression in $[x, 2x]$ of spacing $W\left[d_1,d_1'\right]\ldots\left[d_k,d_k'\right]$, and so

$$
S\left(d_1,\ldots,d_k,d_1',\ldots,d_k'\right) = \frac{x}{W\left[d_1,d_1'\right]\ldots\left[d_k,d_k'\right]} + O(1).
$$

By Lemma 30, the contribution of the main term $\frac{x}{W[d_1,d_1']\ldots[d_k,d_k']}$ to (29) is $(c + o(1))B^{-k}\frac{x}{W}$; note that the restriction of the integrals in (30) to $[0, 1]$ instead of $[0, +\infty)$ is harmless since $S(F_i), S(G_i) < 1$ for all i. Meanwhile, the contribution of the $O(1)$ error is then bounded by

$$
O\left(\sum_{d_1,\ldots,d_k,d_1',\ldots,d_k'} \left(\prod_{i=1}^{k} |F_i(\log_x d_i)||G_i(\log_x d_i')| \right) \right).
$$

By the hypothesis in Theorem 20(i), we see that for $d_1,\ldots,d_k,d_1',\ldots,d_k'$ contributing a non-zero term here, one has

$$
\left[d_1, d_1'\right]\ldots\left[d_k, d_k'\right] \ll x^{1-\varepsilon}
$$

for some fixed $\varepsilon > 0$. From the divisor bound (1), we see that each choice of $\left[d_1, d_1'\right]\ldots\left[d_k, d_k'\right]$ arises from $\ll 1$ choices of $d_1,\ldots,d_k,d_1',\ldots,d_k'$. We conclude that the net contribution of the $O(1)$ error to (29) is $\ll x^{1-\varepsilon}$, and the claim follows.

The Elliott-Halberstam case

Now we show case (i) of Theorem 19. For the sake of notation, we take $i_0 = k$, as the other cases are similar. We use (16) to rewrite the left-hand side of (26) as

$$
\sum_{d_1,\ldots,d_{k-1},d_1',\ldots,d_{k-1}'} \left(\prod_{i=1}^{k-1} \mu(d_i)\mu\left(d_i'\right) F_i\left(\log_x d_i\right) G_i\left(\log_x d_i'\right) \right) \tilde{S}\left(d_1,\ldots,d_{k-1},d_1',\ldots,d_{k-1}'\right)
$$

(43)

where

$$
\tilde{S}\left(d_1,\ldots,d_{k-1},d_1',\ldots,d_{k-1}'\right) := \sum_{\substack{x \leq n \leq 2x \\ n=b\ (W) \\ n+h_i=0\ ([d_i,d_i'])\ \forall i=1,\ldots,k-1}} \theta(n + h_k).
$$

As in the previous case, $\tilde{S}\left(d_1, \ldots, d_{k-1}, d'_1, \ldots, d'_{k-1}\right)$ vanishes unless the $[d_i, d'_i]$ are coprime to each other and to W, and so the summand in (43) vanishes unless the modulus $q_{W, d_1, \ldots, d'_{k-1}}$ defined by

$$q_{W, d_1, \ldots, d'_{k-1}} := W\left[d_1, d'_1\right] \ldots \left[d_{k-1}, d'_{k-1}\right] \tag{44}$$

is square-free. In that case, we may use the Chinese remainder theorem to concatenate the congruence conditions on n into a single primitive congruence condition

$$n + h_k = a_{W, d_1, \ldots, d'_{k-1}} \left(q_{W, d_1, \ldots, d'_{k-1}}\right)$$

for some $a_{W, d_1, \ldots, d'_{k-1}}$ depending on $W, d_1, \ldots, d_{k-1}, d'_1, \ldots, d'_{k-1}$, and conclude using (3) that

$$\tilde{S}\left(d_1, \ldots, d_{k-1}, d'_1, \ldots, d'_{k-1}\right) = \frac{1}{\varphi\left(q_{W, d_1, \ldots, d'_{k-1}}\right)} \sum_{x + h_k \leq n \leq 2x + h_k} \theta(n)$$
$$+ \Delta\left(1_{[x+h_k, 2x+h_k]}\theta; a_{W, d_1, \ldots, d'_{k-1}} \left(q_{W, d_1, \ldots, d'_{k-1}}\right)\right). \tag{45}$$

From the prime number theorem, we have

$$\sum_{x + h_k \leq n \leq 2x + h_k} \theta(n) = (1 + o(1))x$$

and this expression is clearly independent of $d_1, \ldots, d'_{k-1}$. Thus, by Lemma 30, the contribution of the main term in (45) is $(c + o(1))B^{1-k} \frac{x}{\varphi(W)}$. By (11) and (12), it thus suffices to show that for any fixed A we have

$$\sum_{d_1, \ldots, d_{k-1}, d'_1, \ldots, d'_{k-1}} \left(\prod_{i=1}^{k-1} \left|F_i\left(\log_x d_i\right)\right| \left|G_i\left(\log_x d'_i\right)\right| \right) \left|\Delta\left(1_{[x+h_k, 2x+h_k]}\theta; a\,(q)\right)\right| \ll x \log^{-A} x, \tag{46}$$

where $a = a_{W, d_1, \ldots, d'_{k-1}}$ and $q = q_{W, d_1, \ldots, d'_{k-1}}$. For future reference, we note that we may restrict the summation here to those $d_1, \ldots, d'_{k-1}$ for which $q_{W, d_1, \ldots, d'_{k-1}}$ is square-free.

From the hypotheses of Theorem 19(i), we have

$$q_{W, d_1, \ldots, d'_{k-1}} \ll x^{\vartheta}$$

whenever the summand in (43) is non-zero, and each choice q of $q_{W, d_1, \ldots, d'_{k-1}}$ is associated to $O\left(\tau(q)^{O(1)}\right)$ choices of $d_1, \ldots, d_{k-1}, d'_1, \ldots, d'_{k-1}$. Thus, this contribution is

$$\ll \sum_{q \ll x^{\vartheta}} \tau(q)^{O(1)} \sup_{a \in (\mathbb{Z}/q\mathbb{Z})^{\times}} \left|\Delta\left(1_{[x+h_k, 2x+h_k]}\theta; a\,(q)\right)\right|.$$

Using the crude bound

$$\left|\Delta\left(1_{[x+h_k, 2x+h_k]}\theta; a\,(q)\right)\right| \ll \frac{x}{q} \log^{O(1)} x$$

and (2), we have

$$\sum_{q \ll x^{\vartheta}} \tau(q)^C \sup_{a \in (\mathbb{Z}/q\mathbb{Z})^{\times}} \left|\Delta\left(1_{[x+h_k, 2x+h_k]}\theta; a\,(q)\right)\right| \ll x \log^{O(1)} x$$

for any fixed $C > 0$. By the Cauchy-Schwarz inequality, it suffices to show that

$$\sum_{q \ll x^{\vartheta}} \sup_{a \in (\mathbb{Z}/q\mathbb{Z})^{\times}} \left|\Delta\left(1_{[x+h_k, 2x+h_k]}\theta; a\,(q)\right)\right| \ll x \log^{-A} x$$

for any fixed $A > 0$. However, since θ only differs from Λ on powers p^j of primes with $j > 1$, it is not difficult to show that

$$\left| \Delta \left(1_{[x+h_k, 2x+h_k]} \theta; a\,(q) \right) - \Delta \left(1_{[x+h_k, 2x+h_k]} \Lambda; a\,(q) \right) \right| \lll \sqrt{\frac{x}{q}},$$

so the net error in replacing θ here by Λ is $\lll x^{1-(1-\vartheta)/2}$, which is certainly acceptable. The claim now follows from the hypothesis EH$[\vartheta]$, thanks to Claim 8.

The Motohashi-Pintz-Zhang case

Now we show case (ii) of Theorem 19. We repeat the arguments from the 'The Elliott-Halberstam case' section, with the only difference being in the derivation of (46). As observed previously, we may restrict $q_{W,d_1,\dots,d'_{k-1}}$ to be square-free. From the hypotheses in Theorem 19(ii), we also see that

$$q_{W,d_1,\dots,d'_{k-1}} \lll x^\vartheta$$

and that all the prime factors of $q_{W,d_1,\dots,d'_{k-1}}$ are at most x^δ. Thus, if we set $I := [1, x^\delta]$, we see (using the notation from Claim 10) that $q_{W,d_1,\dots,d'_{k-1}}$ lies in $\mathcal{S}_I$ and is thus a factor of P_I. If we then let $\mathcal{A} \subset \mathbb{Z}/P_I\mathbb{Z}$ denote all the primitive residue classes $a\,(P_I)$ with the property that $a = b\,(W)$, and such that for each prime $w < p \le x^\delta$, one has $a + h_i = 0\,(p)$ for some $i = 1, \dots, k$, then we see that $a_{W,d_1,\dots,d'_{k-1}}$ lies in the projection of $\mathcal{A}$ to $\mathbb{Z}/q_{W,d_1,\dots,d'_{k-1}}\mathbb{Z}$. Each $q \in \mathcal{S}_I$ is equal to $q_{W,d_1,\dots,d'_{k-1}}$ for $O\left(\tau(q)^{O(1)}\right)$ choices of $d_1, \dots, d'_{k-1}$. Thus, the left-hand side of (46) is

$$\lll \sum_{q \in \mathcal{S}_I : q \lll x^\vartheta} \tau(q)^{O(1)} \sup_{a \in \mathcal{A}} \left| \Delta \left(1_{[x+h_k, 2x+h_k]} \theta; a\,(q) \right) \right|.$$

Note from the Chinese remainder theorem that for any given q, if one lets a range uniformly in $\mathcal{A}$, then $a\,(q)$ is uniformly distributed among $O\left(\tau(q)^{O(1)}\right)$ different moduli. Thus, we have

$$\sup_{a \in \mathcal{A}} \left| \Delta \left(1_{[x+h_k, 2x+h_k]} \theta; a\,(q) \right) \right| \lll \frac{\tau(q)^{O(1)}}{|\mathcal{A}|} \sum_{a \in \mathcal{A}} \left| \Delta \left(1_{[x+h_k, 2x+h_k]} \theta; a\,(q) \right) \right|,$$

and so it suffices to show that

$$\sum_{q \in \mathcal{S}_I : q \lll x^\vartheta} \frac{\tau(q)^{O(1)}}{|\mathcal{A}|} \sum_{a \in \mathcal{A}} \left| \Delta(1_{[x+h_k, 2x+h_k]} \theta; a\,(q)) \right| \lll x \log^{-A} x$$

for any fixed $A > 0$. We see it suffices to show that

$$\sum_{q \in \mathcal{S}_I : q \lll x^\vartheta} \tau(q)^{O(1)} \left| \Delta(1_{[x+h_k, 2x+h_k]} \theta; a\,(q)) \right| \lll x \log^{-A} x$$

for any given $a \in \mathcal{A}$. But this follows from the hypothesis MPZ$[\varpi, \delta]$ by repeating the arguments of the 'The Elliott-Halberstam case' section.

Crude estimates on divisor sums

To proceed further, we will need some additional information on the divisor sums λ_F (defined in (16)), namely that these sums are concentrated on 'almost primes'; results of this type have also appeared in [38].

Proposition 14 (Almost primality). *Let $k \ge 1$ be fixed, let $(h_1, \dots, h_k)$ be a fixed admissible k-tuple, and let $b\,(W)$ be such that $b + h_i$ is coprime to W for each $i = 1, \dots, k$.*

Let $F_1, \ldots, F_k : [0, +\infty) \to \mathbb{R}$ be fixed smooth compactly supported functions, and let $m_1, \ldots, m_k \geq 0$ and $a_1, \ldots, a_k \geq 1$ be fixed natural numbers. Then,

$$\sum_{x \leq n \leq 2x : n \equiv b \ (W)} \prod_{j=1}^{k} \left(|\lambda_{F_j}(n + h_j)|^{a_j} \tau(n + h_j)^{m_j} \right) \ll B^{-k} \frac{x}{W}. \tag{47}$$

Furthermore, if $1 \leq j_0 \leq k$ is fixed and p_0 is a prime with $p_0 \leq x^{\frac{1}{10k}}$, then we have the variant

$$\sum_{x \leq n \leq 2x : n \equiv b \ (W)} \prod_{j=1}^{k} \left(|\lambda_{F_j}(n + h_j)|^{a_j} \tau(n + h_j)^{m_j} \right) 1_{p_0 | n + h_{j_0}} \ll \frac{\log_x p_0}{p_0} B^{-k} \frac{x}{W}. \tag{48}$$

As a consequence, we have

$$\sum_{x \leq n \leq 2x : n \equiv b \ (W)} \prod_{j=1}^{k} \left(|\lambda_{F_j}(n + h_j)|^{a_j} \tau(n + h_j)^{m_j} \right) 1_{p(n + h_{j_0}) \leq x^\varepsilon} \ll \varepsilon B^{-k} \frac{x}{W}, \tag{49}$$

for any $\varepsilon > 0$, where $p(n)$ denotes the least prime factor of n.

The exponent $\frac{1}{10k}$ can certainly be improved here, but for our purposes, any fixed positive exponent depending only on k will suffice.

Proof. The strategy is to estimate the alternating divisor sums $\lambda_{F_j}(n + h_j)$ by non-negative expressions involving prime factors of $n + h_j$, which can then be bounded combinatorially using standard tools.

We first prove (47). As in the proof of Proposition 30, we can use Fourier expansion to write

$$F_j \left(\log_x d \right) = \int_{\mathbb{R}} \frac{f_j(\xi)}{d^{\frac{1+i\xi}{\log x}}} \, d\xi$$

for some rapidly decreasing $f_j : \mathbb{R} \to \mathbb{C}$ and all natural numbers d. Thus,

$$\lambda_{F_j}(n) = \int_{\mathbb{R}} \left(\sum_{d | n} \frac{\mu(d)}{d^{\frac{1+i\xi}{\log x}}} \right) f_j(\xi) \, d\xi,$$

which factorizes using Euler products as

$$\lambda_{F_j}(n) = \int_{\mathbb{R}} \prod_{p | n} \left(1 - \frac{1}{p^{\frac{1+i\xi}{\log x}}} \right) f_j(\xi) \, d\xi.$$

The function $s \mapsto p^{\frac{-s}{\log x}}$ has a magnitude of $O(1)$ and a derivative of $O\left(\log_x p\right)$ when $\Re(s) > 1$, and thus

$$1 - \frac{1}{p^{\frac{1+i\xi}{\log x}}} = O\left(\min((1 + |\xi|) \log_x p, 1) \right).$$

From the rapid decrease of f_j and the triangle inequality, we conclude that

$$|\lambda_{F_j}(n)| \ll \int_{\mathbb{R}} \left(\prod_{p | n} O\left(\min((1 + |\xi|) \log_x p, 1) \right) \right) \frac{d\xi}{(1 + |\xi|)^A}$$

for any fixed $A > 0$. Thus, noting that $\prod_{p|n} O(1) \ll \tau(n)^{O(1)}$, we have

$$|\lambda_{F_j}(n)|^{a_j} \ll \tau(n)^{O(1)} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \left(\prod_{p|n} \prod_{l=1}^{a_j} \min((1 + |\xi_l|) \log_x p, 1) \right)$$

$$\frac{d\xi_1 \dots d\xi_{a_j}}{(1 + |\xi_1|)^A \dots (1 + |\xi_{a_j}|)^A}$$

for any fixed a_j, A. However, we have

$$\prod_{i=1}^{a_j} \min \left((1 + |\xi_i|) \log_x p, 1 \right) \ll \min \left((1 + |\xi_1| + \cdots + |\xi_{a_j}|) \log_x p, 1 \right),$$

and so

$$|\lambda_{F_j}(n)|^{a_j} \ll \tau(n)^{O(1)} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \frac{\left(\prod_{p|n} \min \left((1 + |\xi_1| + \cdots + |\xi_{a_j}|) \log_x p, 1 \right) \right) d\xi_1 \dots d\xi_{a_j}}{\left(1 + |\xi_1| + \cdots + |\xi_{a_j}| \right)^A}.$$

Making the change of variables $\sigma := 1 + |\xi_1| + \cdots + |\xi_{a_j}|$, we obtain

$$|\lambda_{F_j}(n)|^{a_j} \ll \tau(n)^{O(1)} \int_1^{\infty} \left(\prod_{p|n} \min(\sigma \log_x p, 1) \right) \frac{d\sigma}{\sigma^A}$$

for any fixed $A > 0$. In view of this bound and the Fubini-Tonelli theorem, it suffices to show that

$$\sum_{x \leq n \leq 2x : n \equiv b \ (W)} \prod_{j=1}^{k} \left(\tau(n + h_j)^{O(1)} \prod_{p|n} \min(\sigma_j \log_x p, 1) \right) \ll B^{-k} \frac{x}{W} (\sigma_1 + \cdots + \sigma_k)^{O(1)}$$

for all $\sigma_1, \dots, \sigma_k \geq 1$. By setting $\sigma := \sigma_1 + \cdots + \sigma_k$, it suffices to show that

$$\sum_{x \leq n \leq 2x : n \equiv b \ (W)} \prod_{j=1}^{k} \left(\tau \left(n + h_j \right)^{O(1)} \prod_{p|n+h_j} \min \left(\sigma \log_x p, 1 \right) \right) \ll B^{-k} \frac{x}{W} \sigma^{O(1)} \qquad (50)$$

for any $\sigma \geq 1$.

To proceed further, we factorize $n + h_j$ as a product

$$n + h_j = p_1 \dots p_r$$

of primes $p_1 \leq \cdots \leq p_r$ in increasing order and then write

$$n + h_j = d_j m_j$$

where $d_j := p_1 \dots p_{i_j}$ and i_j is the largest index for which $p_1 \dots p_{i_j} < x^{\frac{1}{10k}}$, and $m_j := p_{i_j+1} \dots p_r$. By construction, we see that $0 \leq i_j < r$, $d_j \leq x^{\frac{1}{10k}}$. Also, we have

$$p_{i_j+1} \geq \left(p_1 \dots p_{i_j+1} \right)^{\frac{1}{i_j+1}} \geq x^{\frac{1}{10k(i_j+1)}}.$$

Since $n \leq 2x$, this implies that

$$r = O(i_j + 1)$$

and so

$$\tau(n + h_j) \leq 2^{O(1 + \Omega(d_j))},$$

where we recall that $\Omega(d_j) = i_j$ denotes the number of prime factors of d_j, counting multiplicity. We also see that

$$p(m_j) \geq x^{\frac{1}{10k(1+\Omega(d_j))}} \geq x^{\frac{1}{10k(1+\Omega(d_1\ldots d_k))}} =: R,$$

where $p(n)$ denotes the least prime factor of n. Finally, we have that

$$\prod_{p|n+h_j} \min\left(\sigma \log_x p, 1\right) \leq \prod_{p|d_j} \min\left(\sigma \log_x p, 1\right),$$

and we see that the $d_1, \ldots, d_k, W$ are coprime. We may thus estimate the left-hand side of (50) by

$$\ll \sum_* \left(\prod_{j=1}^k 2^{O(1+\Omega(d_j))} \prod_{p|d_j} \min(\sigma \log_x p, 1) \right) \sum_{**} 1$$

where the outer sum $\sum_*$ is over $d_1, \ldots, d_k \leq x^{\frac{1}{10k}}$ with $d_1, \ldots, d_k, W$ coprime, and the inner sum $\sum_{**}$ is over $x \leq n \leq 2x$ with $n = b\ (W)$ and $n + h_j = 0\ (d_j)$ for each j, with $p\left(\frac{n+h_j}{d_j}\right) \geq R$ for each j.

We bound the inner sum $\sum_{**} 1$ using a Selberg sieve upper bound. Let G be a smooth function supported on $[0, 1]$ with $G(0) = 1$, and let $d = d_1 \ldots d_k$. We see that

$$\sum_{**} 1 \leq \sum_{\substack{x \leq n \leq 2x \\ n+h_i \equiv 0\ (d_i) \\ n \equiv b\ (W)}} \prod_{i=1}^k \left(\sum_{\substack{e|n+h_i \\ (e, dW)=1}} \mu(e) G(\log_R e) \right)^2,$$

since the product is $G(0)^{2k} = 1$ if $p\left(\frac{n+h_j}{d_j}\right) \geq R$, and non-negative otherwise. The right-hand side may be expanded as

$$\sum_{\substack{e_1, \ldots, e_k, e_1', \ldots, e_k' \\ (e_i e_i', dW)=1 \forall i}} \left(\prod_{i=1}^k \mu(e_i)\mu(e_i') G\left(\log_R e_i\right) G\left(\log_R e_i'\right) \right) \sum_{\substack{x \leq n \leq 2x \\ n+h_i \equiv 0\ (d_i[e_i, e_i']) \\ n \equiv b\ (W)}} 1.$$

As in the 'The trivial case' section, the inner sum vanishes unless the $e_i e_i'$ are coprime to each other and dW, in which case it is

$$\frac{x}{dW[e_1, e_1'] \ldots [e_k, e_k']} + O(1).$$

The $O(1)$ term contributes $\ll R^k \ll x^{1/10}$, which is negligible. By Lemma 30, if $\Omega(d) \ll \log^{1/2} x$, then the main term contributes

$$\ll \left(\frac{d}{\varphi(d)} \right)^k \frac{x}{dW} (\log R)^{-k} \ll 2^{\Omega(d)} B^{-k} \frac{x}{dW}.$$

We see that this final bound applies trivially if $\Omega(d) \gg \log^{1/2} x$. The bound (50) thus reduces to

$$\sum_* \left(\prod_{j=1}^k \frac{2^{O(1+\Omega(d_j))}}{d_j} \prod_{p|d_j} \min(\sigma \log_x p, 1) \right) \ll \sigma^{O(1)}. \tag{51}$$

Ignoring the coprimality conditions on the d_j for an upper bound, we see this is bounded by

$$
\prod_{w < p \leq x^{\frac{1}{10k}}} \left(1 + \frac{O(\min(\sigma \log_x(p), 1))}{p} \sum_{j \geq 0} \frac{O(1)^j}{p^j} \right)^k
$$

$$
\ll \exp\left(O\left(\sum_{p \leq x} \frac{(\min(\sigma \log_x(p), 1))}{p} \right) \right).
$$

But from Mertens' theorem, we have

$$
\sum_{p \leq x} \frac{\min(\sigma \log_x p, 1)}{p} = O\left(\log \frac{1}{\sigma} \right),
$$

and the claim (47) follows.

The proof of (48) is a minor modification of the argument above used to prove (47). Namely, the variable d_{j_0} is now replaced by $[d_0, p_0] < x^{1/5k}$, which upon factoring out p_0 has the effect of multiplying the upper bound for (51) by $O\left(\frac{\sigma \log_x p_0}{p_0} \right)$ (at the negligible cost of deleting the prime p_0 from the sum $\sum_{p \leq x}$), giving the claim; we omit the details.

Finally, (49) follows immediately from (47) when $\varepsilon > \frac{1}{10k}$, and from (48) and Mertens' theorem when $\varepsilon \leq \frac{1}{10k}$. $\qquad \square$

Remark 32. As in [38], one can use Proposition 14, together with the observation that the quantity $\lambda_F(n)$ is bounded whenever $n = O(x)$ and $p(n) \geq x^\varepsilon$, to conclude that whenever the hypotheses of Lemma 18 are obeyed for some ν of the form (18), then there exists a fixed $\varepsilon > 0$ such that for all sufficiently large x, there are $\gg \frac{x}{\log^k x}$ elements n of $[x, 2x]$ such that $n + h_1, \ldots, n + h_k$ have no prime factor less than x^ε, and that at least m of the $n + h_1, \ldots, n + h_k$ are prime.

The generalized Elliott-Halberstam case

Now we show case (ii) of Theorem 20. For the sake of notation, we shall take $i_0 = k$, as the other cases are similar; thus, we have

$$
\sum_{i=1}^{k-1} (S(F_i) + S(G_i)) < \vartheta. \tag{52}
$$

The basic idea is to view the sum (29) as a variant of (26), with the role of the function θ now being played by the product divisor sum $\lambda_{F_k} \lambda_{G_k}$, and to repeat the arguments in the 'The Elliott-Halberstam case' section. To do this, we rely on Proposition 14 to restrict $n + h_i$ to the almost primes.

We turn to the details. Let $\varepsilon > 0$ be an arbitrary fixed quantity. From (49) and Cauchy-Schwarz, one has

$$
\sum_{\substack{x \leq n \leq 2x \\ n = b \, (W)}} \left(\prod_{i=1}^{k} \lambda_{F_i}(n + h_i) \lambda_{G_i}(n + h_i) \right) 1_{p(n + h_k) \leq x^\varepsilon} = O\left(\varepsilon B^{-k} \frac{x}{W} \right)
$$

with the implied constant uniform in ε, so by the triangle inequality and a limiting argument as $\varepsilon \to 0$, it suffices to show that

$$\sum_{\substack{x \leq n \leq 2x \\ n=b\ (W)}} \left(\prod_{i=1}^{k} \lambda_{F_i}(n+h_i) \lambda_{G_i}(n+h_i) \right) 1_{p(n+h_k)>x^\varepsilon} = (c_\varepsilon + o(1))B^{-k}\frac{x}{W} \tag{53}$$

where c_ε is a quantity depending on ε but not on x, such that

$$\lim_{\varepsilon \to 0} c_\varepsilon = \prod_{i=1}^{k} \int_0^1 F_i'(t) G_i'(t)\, dt.$$

We use (16) to expand out $\lambda_{F_i}, \lambda_{G_i}$ for $i = 1, \ldots, k-1$, but *not* for $i = k$, so that the left-hand side of (29) becomes

$$\sum_{d_1,\ldots,d_{k-1},d_1',\ldots,d_{k-1}'} \left(\prod_{i=1}^{k} \mu(d_i)\mu\left(d_i'\right) F_i\left(\log_x d_i\right) G_i\left(\log_x d_i'\right) \right) S'\left(d_1,\ldots,d_{k-1},d_1',\ldots,d_{k-1}'\right) \tag{54}$$

where

$$S'\left(d_1,\ldots,d_{k-1},d_1',\ldots,d_{k-1}'\right) := \sum_{\substack{x \leq n \leq 2x \\ n=b\ (W) \\ n+h_i=0\ ([d_i,d_i'])\ \forall i=1,\ldots,k-1}} \lambda_{F_k}(n+h_k)\lambda_{G_k}(n+h_k)1_{p(n+h_k)>x^\varepsilon}.$$

As before, the summand in (54) vanishes unless the modulus[d] $q_{W,d_1,\ldots,d_{k-1}'}$ defined in (44) is square-free, in which case we have the analogue

$$S'\left(d_1,\ldots,d_{k-1},d_1',\ldots,d_{k-1}'\right) = \frac{1}{\varphi(q)} \sum_{\substack{x+h_k \leq n \leq 2x+h_k \\ (n,q)=1}} \lambda_{F_k}(n)\lambda_{G_k}(n)1_{p(n)>x^\varepsilon}$$

$$+ \Delta\left(1_{[x+h_k,2x+h_k]}\lambda_{F_k}\lambda_{G_k}1_{p(\cdot)>x^\varepsilon}; a\ (q)\right) \tag{55}$$

of (45). Here we have put $q = q_{W,d_1,\ldots,d_{k-1}'}$ and $a = a_{W,d_1,\ldots,d_{k-1}'}$ for convenience. We thus split

$$S' = S_1' - S_2' + S_3',$$

where,

$$S_1'\left(d_1,\ldots,d_{k-1},d_1',\ldots,d_{k-1}'\right) = \frac{1}{\varphi(q)} \sum_{x+h_k \leq n \leq 2x+h_k} \lambda_{F_k}(n)\lambda_{G_k}(n)1_{p(n)>x^\varepsilon}, \tag{56}$$

$$S_2'\left(d_1,\ldots,d_{k-1},d_1',\ldots,d_{k-1}'\right) = \frac{1}{\varphi(q)} \sum_{x+h_k \leq n \leq 2x+h_k;(n,q)>1} \lambda_{F_k}(n)\lambda_{G_k}(n)1_{p(n)>x^\varepsilon}, \tag{57}$$

$$S_3'\left(d_1,\ldots,d_{k-1},d_1',\ldots,d_{k-1}'\right) = \Delta\left(1_{[x+h_k,2x+h_k]}\lambda_{F_k}\lambda_{G_k}1_{p(\cdot)>x^\varepsilon}; a\ (q)\right), \tag{58}$$

when $q = q_{W,d_1,\ldots,d_{k-1}'}$ is square-free, with $S_1' = S_2' = S_3' = 0$ otherwise.

For $j \in \{1, 2, 3\}$, let

$$\Sigma_j = \sum_{d_1,\ldots,d_{k-1},d_1',\ldots,d_{k-1}'} \left(\prod_{i=1}^{k} \mu(d_i)\mu\left(d_i'\right) F_i\left(\log_x d_i\right) G_i\left(\log_x d_i'\right) \right)$$

$$S_j'\left(d_1,\ldots,d_{k-1},d_1',\ldots,d_{k-1}'\right). \tag{59}$$

To show (53), it thus suffices to show the main term estimate

$$\Sigma_1 = (c_\varepsilon + o(1)) B^{-k} \frac{x}{W}, \tag{60}$$

the first error term estimate

$$\Sigma_2 \ll x^{1-\varepsilon}, \tag{61}$$

and the second error term estimate

$$\Sigma_3 \ll x \log^{-A} x \tag{62}$$

for any fixed $A > 0$.

We begin with (61). Observe that if $p(n) > x^\varepsilon$, then the only way that $\left(n, q_{W,d_1,\dots,d'_{k-1}}\right)$ can exceed 1 is if there is a prime $x^\varepsilon < p \ll x$ which divides both n and one of $d_1, \dots, d'_{k-1}$; in particular, this case can only occur when $k > 1$. For the sake of notation, we will just consider the contribution when there is a prime that divides p and d_1, as the other $2k - 3$ cases are similar. By (57), this contribution to Σ_2 can then be crudely bounded (using (1)) by

$$\Sigma_2 \ll \sum_{x^\varepsilon < p \ll x} \sum_{d_1,\dots,d_{k-1},d'_1,\dots,d'_{k-1} \le x; p|d_1} \frac{1}{[d_1, d'_1] \dots [d_{k-1}, d'_{k-1}]} \sum_{n \ll x; p|n} 1$$

$$\ll \sum_{x^\varepsilon < p \ll x} \frac{x}{p} \left(\sum_{e_1 \le x^2; p|e_1} \frac{\tau(e_1)}{e_1} \right) \prod_{i=2}^{k-1} \left(\sum_{e_i \le x^2} \frac{\tau(e_i)}{e_i} \right)$$

$$\ll \sum_{x^\varepsilon < p \ll x} \frac{x}{p^2}$$

$$\ll x^{1-\varepsilon}$$

as required, where we have made the change of variables $e_i := [d_i, d'_i]$, using the divisor bound to control the multiplicity.

Now we show (62). From the hypothesis (28), we have $q_{W,d_1,\dots,d'_{k-1}} \ll x^\theta$ whenever the summand in (62) is non-zero. From the divisor bound, for each $q \ll x^\theta$, there are $O\left(\tau(q)^{O(1)}\right)$ choices of $d_1, \dots, d'_{k-1}$ with $q_{W,d_1,\dots,d'_{k-1}} = q$. We see that the product in (59) is $O(1)$. Thus, by (58), we may bound Σ_3 by

$$\Sigma_3 \ll \sum_{q \ll x^\theta} \tau(q)^{O(1)} \sup_{a \in (\mathbb{Z}/q\mathbb{Z})^\times} \left| \Delta \left(1_{[x+h_k, 2x+h_k]} \lambda_{F_k} \lambda_{G_k} 1_{p(\cdot) > x^\varepsilon}; a \ (q) \right) \right|.$$

From (2), we easily obtain the bound

$$\Sigma_3 \ll \sum_{q \ll x^\theta} \tau(q)^{O(1)} \sup_{a \in (\mathbb{Z}/q\mathbb{Z})^\times} \left| \Delta \left(1_{[x+h_k, 2x+h_k]} \lambda_{F_k} \lambda_{G_k} 1_{p(\cdot) > x^\varepsilon}; a \ (q) \right) \right| \ll x \log^{O(1)} x,$$

so by Cauchy-Schwarz, it suffices to show that

$$\sum_{q \ll x^\theta} \sup_{a \in (\mathbb{Z}/q\mathbb{Z})^\times} \left| \Delta \left(1_{[x+h_k, 2x+h_k]} \lambda_{F_k} \lambda_{G_k} 1_{p(\cdot) > x^\varepsilon}; a \ (q) \right) \right| \ll x \log^{-A} x \tag{63}$$

for any fixed $A > 0$.

If we had the additional hypothesis $S(F_k) + S(G_k) < 1$, then this would follow easily from the hypothesis GEH[ϑ] thanks to Claim 12, since one can write $\lambda_{F_k} \lambda_{G_k} 1_{p(\cdot) > x^\varepsilon} = \alpha \star \beta$ with

$$\alpha(n) := 1_{p(n) > x^\varepsilon} \sum_{d,d':[d,d']=n} \mu(d) F_k \left(\log_x d \right) \mu\left(d'\right) G_k \left(\log_x d' \right)$$

and

$$\beta(n) := 1_{p(n) > x^\varepsilon}.$$

But even in the absence of the hypothesis $S(F_k) + S(G_k) < 1$, we can still invoke GEH$[\vartheta]$ after appealing to the fundamental theorem of arithmetic. Indeed, if $n \in [x + h_k, 2x + h_k]$ with $p(\cdot) > \varepsilon$, then we have

$$n = p_1 \dots p_r$$

for some primes $x^\varepsilon < p_1 \leq \cdots \leq p_r \leq 2x + h_k$, which forces $r \leq \frac{1}{\varepsilon} + 1$. If we then partition $[x^\varepsilon, 2x + h_k]$ by $O\left(\log^{A+1} x\right)$ intervals $I_1, \dots, I_m$, with each I_j contained in an interval of the form $\left[N, \left(1 + \log^{-A} x\right) N\right]$, then we have $p_i \in I_{j_i}$ for some $1 \leq j_1 \leq \cdots \leq j_r \leq m$, with the product interval $I_{j_1} \cdots I_{j_r}$ intersecting $[x + h_k, 2x + h_k]$. For fixed r, there are $O\left(\log^{Ar+r} x\right)$ such tuples $(j_1, \dots, j_r)$, and a simple application of the prime number theorem with classical error term (and crude estimates on the discrepancy Δ) shows that each tuple contributes $O\left(x \log^{-Ar+O(1)} x\right)$ to (63) (here, and for the rest of this section, implied constants will be independent of A unless stated otherwise). In particular, the $O\left(\log^{A(r-1)} x\right)$ tuples $(j_1, \dots, j_r)$ with one repeated j_i, or for which the interval $I_{j_1} \cdots I_{j_r}$ meets the boundary of $[x + h_k, 2x + h_k]$, contributes $O\left(\log^{-A+O(1)} x\right)$. This is an acceptable error to (63), and so these tuples may be removed. Thus, it suffices to show that

$$\sum_{q \ll x^\theta} \sup_{a \in (\mathbb{Z}/q\mathbb{Z})^\times} \left| \Delta\left(\lambda_{F_k} \lambda_{G_k} 1_{A_{j_1,\dots,j_r}}; a\,(q)\right) \right| \ll x \log^{-A(r+1)+O(1)} x$$

for any $1 \leq r \leq \frac{1}{\varepsilon} + 1$ and $1 \leq j_1 < \cdots < j_r \leq m$ with $I_{j_1} \cdots I_{j_r}$ contained in $[x + h_k, x + 2h_k]$, where $A_{j_1,\dots,j_r}$ is the set of all products $p_1 \dots p_r$ with $p_i \in I_{j_i}$ for $i = 1, \dots, r$, and where we allow implied constants in the $\ll$ notation to depend on ε. But for n in $A_{j_1,\dots,j_r}$, the 2^r factors of n are just the products of subsets of $\{p_1, \dots, p_r\}$, and from the smoothness of F_k, G_k, we see that $\lambda_{F_k}(n)$ is equal to some bounded constant (depending on $j_1, \dots, j_r$, but independent of $p_1, \dots, p_r$), plus an error of $O(\log^{-A} x)$. As before, the contribution of this error is $O\left(\log^{-A(r+1)+O(1)} x\right)$, so it suffices to show that

$$\sum_{q \ll x^\theta} \sup_{a \in (\mathbb{Z}/q\mathbb{Z})^\times} \left| \Delta\left(1_{A_{j_1,\dots,j_r}}; a\,(q)\right) \right| \ll x \log^{-A(r+1)+O(1)} x.$$

But one can write $1_{A_{j_1,\dots,j_r}}$ as a convolution $1_{A_{j_1}} \star \cdots \star 1_{A_{j_r}}$, where A_{j_i} denotes the primes in I_{j_i}; assigning A_{j_r} (for instance) to be β and the remaining portion of the convolution to be α, the claim now follows from the hypothesis GEH$[\vartheta]$, thanks to the Siegel-Walfisz theorem (see, e.g. [32, Satz 4] or [33, Th. 5.29]).

Finally, we show (60). By Lemma 30, we have

$$\sum_{\substack{d_1,\dots,d_{k-1},d'_1,\dots,d'_{k-1} \\ d_1 d'_1,\dots,d_{k-1} d'_{k-1}, W \,\text{coprime}}} \frac{\prod_{i=1}^{k-1} \mu(d_i) \mu(d'_i) F_i(\log_x d_i) G_i(\log_x d'_i)}{\varphi\left(q_{W,d_1,\dots,d'_{k-1}}\right)} = \frac{1}{\varphi(W)} \left(C' + o(1)\right) B^{-k+1},$$

where

$$C' := \prod_{i=1}^{k-1} \int_0^1 F'_i(t) G'_i(t)\, dt$$

(note that F_i, G_i are supported on $[0, 1]$ by hypothesis), so by (56) it suffices to show that

$$\sum_{x+h_k \leq n \leq 2x+h_k} \lambda_{F_k}(n)\lambda_{G_k}(n)1_{p(n)>x^\varepsilon} = \left(C_\varepsilon'' + o(1)\right)\frac{x}{\log x}, \tag{64}$$

where C_ε'' is a quantity depending on ε but not on x such that

$$\lim_{\varepsilon \to 0} C_\varepsilon'' = \int_0^1 F_k'(t)G_k'(t)\, dt.$$

In the case $S(F_k) + S(G_k) < 1$, this would follow easily from (the $k = 1$ case of) Theorem 20(i) and Proposition 14. In the general case, we may appeal once more to the fundamental theorem of arithmetic. As before, we may factor $n = p_1 \ldots p_r$ for some $x^\varepsilon \leq p_1 \leq \cdots \leq p_r \leq 2x + h_k$ and $r \leq \frac{1}{\varepsilon} + 1$. The contribution of those n with a repeated prime factor $p_i = p_{i+1}$ can easily be shown to be $\ll x^{1-\varepsilon}$ in the same manner we dealt with Σ_2, so we may restrict attention to the square-free n, for which the p_i are strictly increasing. In that case, one can write

$$\lambda_{F_k}(n) = (-1)^r \partial_{(\log_x p_1)} \ldots \partial_{(\log_x p_r)} F_k(0)$$

and

$$\lambda_{G_k}(n) = (-1)^r \partial_{(\log_x p_1)} \ldots \partial_{(\log_x p_r)} G_k(0)$$

where $\partial_{(h)}F(x) := F(x+h) - F(x)$. On the other hand, a standard application of Mertens' theorem and the prime number theorem (and an induction on r) shows that for any fixed $r \geq 1$ and any fixed continuous function $f : \mathbb{R}^r \to \mathbb{R}$, we have

$$\sum_{x^\varepsilon \leq p_1 < \cdots < p_r: x+h_k \leq p_1 \leq \ldots p_r \leq 2x+h_k} f\left(\log_x p_1, \ldots, \log_x p_r\right) = \left(c_f + o(1)\right)\frac{x}{\log x}$$

where c_f is the quantity

$$c_f := \int_{\varepsilon \leq t_1 < \cdots < t_r: t_1 + \cdots + t_r = 1} f(t_1, \ldots, t_r)\frac{n1 \ldots dt_{r-1}}{t_1 \ldots t_r}$$

where we lift Lebesgue measure $dt_1 \ldots dt_{r-1}$ up to the hyperplane $t_1 + \cdots + t_r = 1$, and thus

$$\int_{t_1 + \cdots + t_r = 1} F(t_1, \ldots, t_r)\, dt_1 \ldots dt_{r-1} := \int_{\mathbb{R}^{r-1}} F(t_1, \ldots, t_{r-1}, 1 - t_1 - \cdots - t_{r-1})dt_1 \ldots dt_{r-1}.$$

Putting all these together, we see that we obtain an asymptotic (64) with

$$C_\varepsilon'' := \sum_{1 \leq r \leq \frac{1}{\varepsilon}+1} \int_{\varepsilon \leq t_1 < \cdots < t_r: t_1 + \cdots + t_r = 1} \partial_{(t_1)} \ldots \partial_{(t_r)} F_k(0) \partial_{(t_1)} \ldots \partial_{(t_r)} G_k(0)\frac{dt_1 \ldots dt_{r-1}}{t_1 \ldots t_r}.$$

By Proposition 14, we have $C_\varepsilon'' + O(\varepsilon) = O(1)$. In the case $F_k = G_k$, we see that this implies $'_\varepsilon$ converges to a limit as $\varepsilon \to 0$, and the general case $F_k \neq G_k$ then follows from using the Cauchy-Schwarz inequality. Therefore, we have the absolute convergence

$$\sum_{r>0} \int_{0 < t_1 < \cdots < t_r: t_1 + \cdots + t_r = 1} |\partial_{t_1} \ldots \partial_{t_r} F_k(0)||\partial_{t_1} \ldots \partial_{t_r} G_k(0)|\frac{dt_1 \ldots dt_{r-1}}{t_1 \ldots t_r} < \infty, \tag{65}$$

and so, by the dominated convergence theorem, it suffices to establish the identity

$$\sum_{r>0} \int_{0 < t_1 < \cdots < t_r: t_1 + \cdots + t_r = 1} \partial_{t_1} \ldots \partial_{t_r} F_k(0) \partial_{t_1} \ldots \partial_{t_r} G_k(0)\frac{dt_1 \ldots dt_{r-1}}{t_1 \ldots t_r} = \int_0^1 F_k'(t)G_k'(t)\, dt.$$

$$\tag{66}$$

It will suffice to show the identity

$$\sum_{r>0} \int_{0<t_1<\cdots<t_r:t_1+\cdots+t_r=1} |\partial_{t_1}\ldots\partial_{t_r}F(0)|^2 \frac{dt_1\ldots dt_{r-1}}{t_1\ldots t_r} = \int_0^1 |F'(t)|^2\, dt \qquad (67)$$

for any smooth $F : [0,+\infty) \to \mathbb{R}$, since (66) follows by replacing F with $F_k + G_k$ and $F_k - G_k$ and then subtracting.

At this point, we use the following identity:

Lemma 33. *For any positive reals $t_1,\ldots,t_r$ with $r \geq 1$, we have*

$$\frac{1}{t_1\ldots t_r} = \sum_{\sigma\in S_r} \frac{1}{\prod_{i=1}^r \left(\sum_{j=i}^r t_{\sigma(j)}\right)}. \qquad (68)$$

Thus, for instance, when $r = 2$, we have

$$\frac{1}{t_1 t_2} = \frac{1}{(t_1+t_2)t_1} + \frac{1}{(t_1+t_2)t_2}.$$

Proof. If the right-hand side of (68) is denoted $f_r(t_1,\ldots,t_r)$, then one easily verifies the identity

$$f_r(t_1,\ldots,t_r) = \frac{1}{t_1+\cdots+t_r} \sum_{i=1}^r f_{r-1}(t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_r)$$

for any $r > 1$; but the left-hand side of (68) also obeys this identity, and the claim then follows from induction. □

From this lemma and symmetrisation, we may rewrite the left-hand side of (67) as

$$\sum_{r>0} \int_{\substack{t_1,\ldots,t_r\geq 0 \\ t_1+\cdots+t_r=1}} |\partial_{(t_1)}\ldots\partial_{(t_r)}F(0)|^2 \frac{dt_1\ldots dt_{r-1}}{\prod_{i=1}^r \left(\sum_{j=i}^r t_i\right)}.$$

Let

$$I_a(F) := \int_0^a F'(t)^2\, dt,$$

and

$$J_a(F) := (\partial_{(a)}F(0))^2.$$

One can then rewrite (67) as the identity

$$I_1(F) = \sum_{r=1}^\infty K_{1,r}(F), \qquad (69)$$

where

$$K_{a,r}(F) := \int_{\substack{t_1,\ldots,t_r\geq 0 \\ t_1+\cdots+t_r=a}} J_{t_r}\left(\partial_{(t_1)}\ldots\partial_{(t_{r-1})}F\right) \frac{dt_1\ldots dt_{r-1}}{a\,(a-t_1)\ldots(a-t_1-\cdots-t_{r-1})}.$$

To prove this, we first observe the identity

$$I_a(F) = \frac{1}{a}J_a(F) + \int_{0\leq t\leq a} I_{a-t}\left(\partial_{(t)}F\right) \frac{dt}{a}$$

for any $a > 0$; indeed, we have

$$
\begin{aligned}
\int_{0 \leq t \leq a} I_{a-t}(\partial_{(t)} F) \frac{dt}{a} &= \int_{0 \leq t \leq a; 0 \leq u \leq a-t} |F'(t+u) - F'(t)|^2 \frac{dudt}{a} \\
&= \int_{0 \leq t \leq s \leq a} |F'(s) - F'(t)|^2 \frac{dsdt}{a} \\
&= \frac{1}{2} \int_0^a \int_0^a |F'(s) - F'(t)|^2 \frac{dsdt}{a} \\
&= \int_0^a |F'(s)|^2 \, ds - \frac{1}{a} \left(\int_0^a F'(s) \, ds \right) \left(\int_0^a F'(t) \, dt \right) \\
&= I_a(F) - \frac{1}{a} J_a(F),
\end{aligned}
$$

and the claim follows. Iterating this identity k times, we see that

$$
I_a(F) = \sum_{r=1}^k K_{a,r}(F) + L_{a,k}(F) \tag{70}
$$

for any $k \geq 1$, where

$$
L_{a,k}(F) := \int_{\substack{t_1, \ldots, t_k \geq 0 \\ t_1 + \cdots + t_k \leq a}} I_{1-t_1-\cdots-t_k} \left(\partial_{(t_1)} \ldots \partial_{(t_k)} F \right) \frac{dt_1 \ldots dt_k}{a(a-t_1) \ldots (a-t_1-\cdots-t_{k-1})}.
$$

In particular, dropping the $L_{a,k}(F)$ term and sending $k \to \infty$ yields the lower bound

$$
\sum_{r=1}^{\infty} K_{a,r}(F) \leq I_a(F). \tag{71}
$$

On the other hand, we can expand $L_{a,k}(F)$ as

$$
\int_{\substack{t_1, \ldots, t_k, t \geq 0 \\ t_1 + \cdots + t_k + t \leq a}} |\partial_{(t_1)} \ldots \partial_{(t_k)} F'(t)|^2 \frac{dt_1 \ldots dt_k dt}{a(a-t_1) \ldots (a-t_1-\cdots-t_{k-1})}.
$$

Writing $s := t_1 + \cdots + t_k$, we obtain the upper bound

$$
L_{a,k}(F) \leq \int_{s,t \geq 0: s+t \leq a} K_{s,k}(F'_t) \, dt,
$$

where $F_t(x) := F(x+t)$. Summing this and using (71) and the monotone convergence theorem, we conclude that

$$
\sum_{k=1}^{\infty} L_{a,k}(F) \leq \int_{s,t \geq 0: s+t \leq a} I_s(F_t) \, dt < \infty,
$$

and in particular $L_{a,k}(F) \to 0$ as $k \to \infty$. Sending $k \to \infty$ in (70), we obtain (69) as desired.

Reduction to a variational problem

Now that we have proven Theorems 19 and 20, we can now establish Theorems 22, 24, 26 and 28. The main technical difficulty is to take the multidimensional measurable functions F appearing in these functions and approximate them by tensor products of smooth functions, for which Theorems 19 and 20 may be applied.

Proof of Theorem 22

We now prove Theorem 22. Let k, m, ϑ obey the hypotheses of that theorem, and thus we may find a fixed square-integrable function $F : [0, +\infty)^k \to \mathbb{R}$ supported on the simplex

$$\mathcal{R}_k := \left\{ (t_1, \ldots, t_k) \in [0, +\infty)^k : t_1 + \cdots + t_k \leq 1 \right\}$$

and not identically zero and with

$$\frac{\sum_{i=1}^{k} J_i(F)}{I(F)} > \frac{2m}{\vartheta}. \tag{72}$$

We now perform a number of technical steps to further improve the structure of *F*. Our arguments here will be somewhat convoluted and are not the most efficient way to prove Theorem 22 (which in any event was already established in [5]), but they will motivate the similar arguments given below to prove the more difficult results in Theorems 24, 26 and 28. In particular, we will use regularisation techniques which are compatible with the vanishing marginal condition (35) that is a key hypothesis in Theorem 28.

We first need to rescale and retreat a little bit from the slanted boundary of the simplex $\mathcal{R}_k$. Let $\delta_1 > 0$ be a sufficiently small fixed quantity, and write $F_1 : [0, +\infty)^k \to \mathbb{R}$ to be the rescaled function

$$F_1(t_1, \ldots, t_k) := F\left(\frac{t_1}{\vartheta/2 - \delta_1}, \ldots, \frac{t_k}{\vartheta/2 - \delta_1} \right).$$

Thus, F_1 is a fixed square-integrable measurable function supported on the rescaled simplex

$$(\vartheta/2 - \delta_1) \cdot \mathcal{R}_k = \left\{ (t_1, \ldots, t_k) \in [0, +\infty)^k : t_1 + \cdots + t_k \leq \vartheta/2 - \delta_1 \right\}.$$

From (72), we see that if δ_1 is small enough, then F_1 is not identically zero and

$$\frac{\sum_{i=1}^{k} J_i(F_1)}{I(F_1)} > m. \tag{73}$$

Let δ_1 and F_1 be as above. Next, let $\delta_2 > 0$ be a sufficiently small fixed quantity (smaller than δ_1), and write $F_2 : [0, +\infty)^k \to \mathbb{R}$ to be the shifted function, defined by setting

$$F_2(t_1, \ldots, t_k) := F_1(t_1 - \delta_2, \ldots, t_k - \delta_2)$$

when $t_1, \ldots, t_k \geq \delta_2$, and $F_2(t_1, \ldots, t_k) = 0$ otherwise. As F_1 was square-integrable, compactly supported, and not identically zero, and because spatial translation is continuous in the strong operator topology on L^2, it is easy to see that we will have F_2 not identically zero and that

$$\frac{\sum_{i=1}^{k} J_i(F_2)}{I(F_2)} > m \tag{74}$$

for δ_2 small enough (after restricting F_2 back to $[0, +\infty)^k$, of course). For δ_2 small enough, this function will be supported on the region

$$\left\{ (t_1, \ldots, t_k) \in \mathbb{R}^k : t_1 \cdots + t_k \leq \vartheta/2 - \delta_2; t_1, \ldots, t_k \geq \delta_2 \right\},$$

and thus F_2 stays away from all the boundary faces of $\mathcal{R}_k$.

By convolving F_2 with a smooth approximation to the identity that is supported sufficiently close to the origin, one may then find a *smooth* function $F_3 : [0, +\infty)^k \to \mathbb{R}$, supported on

$$\left\{ (t_1, \ldots, t_k) \in \mathbb{R}^k : t_1 \cdots + t_k \leq \vartheta/2 - \delta_2/2; t_1, \ldots, t_k \geq \delta_2/2 \right\},$$

which is not identically zero and such that

$$\frac{\sum_{i=1}^k J_i(F_3)}{I(F_3)} > m. \tag{75}$$

We extend F_3 by zero to all of $\mathbb{R}^k$ and then define the function $f_3 : \mathbb{R}^k \to \mathbb{R}$ by

$$f_3(t_1, \ldots, t_k) := \int_{s_1 \geq t_1, \ldots, s_k \geq t_k} F_3(s_1, \ldots, s_k) \, ds_1 \ldots ds_k,$$

and thus f_3 is smooth, not identically zero and supported on the region

$$\left\{ (t_1, \ldots, t_k) \in \mathbb{R}^k : \sum_{i=1}^k \max(t_i, \delta_2/2) \leq \vartheta/2 - \delta_2/2 \right\}. \tag{76}$$

From the fundamental theorem of calculus, we have

$$F_3(t_1, \ldots, t_k) := (-1)^k \frac{\partial^k}{\partial t_1 \ldots \partial t_k} f_3(t_1, \ldots, t_k), \tag{77}$$

and so $I(F_3) = \tilde{I}(f_3)$ and $J_i(F_3) = \tilde{J}_i(f_3)$ for $i = 1, \ldots, k$, where

$$\tilde{I}(f_3) := \int_{[0,+\infty)^k} \left| \frac{\partial^k}{\partial t_1 \ldots \partial t_k} f_3(t_1, \ldots, t_k) \right|^2 dt_1 \ldots dt_k \tag{78}$$

and

$$\tilde{J}_i(f_3) := \int_{[0,+\infty)^{k-1}} \left| \frac{\partial^{k-1}}{\partial t_1 \ldots \partial t_{i-1} \partial t_{i+1} \ldots \partial t_k} f_3(t_1, \ldots, t_{i-1}, 0, t_{i+1}, \ldots, t_k) \right|^2 \\ dt_1 \ldots dt_{i-1} dt_{i+1} \ldots dt_k. \tag{79}$$

In particular,

$$\frac{\sum_{i=1}^k \tilde{J}_i(f_3)}{\tilde{I}(f_3)} > m. \tag{80}$$

Now we approximate f_3 by linear combinations of tensor products. By the Stone-Weierstrass theorem, we may express f_3 as the uniform limit of functions of the form

$$(t_1, \ldots, t_k) \mapsto \sum_{j=1}^J c_j f_{1,j}(t_1) \ldots f_{k,j}(t_k) \tag{81}$$

where $c_1, \ldots, c_J$ are real scalars, and $f_{i,j} : \mathbb{R} \to \mathbb{R}$ are smooth compactly supported functions. Since f_3 is supported in (76), we can ensure that all the components $f_{1,j}(t_1) \ldots f_{k,j}(t_k)$ are supported in the slightly larger region

$$\left\{ (t_1, \ldots, t_k) \in \mathbb{R}^k : \sum_{i=1}^k \max(t_i, \delta_2/4) \leq \vartheta/2 - \delta_2/4 \right\}.$$

Observe that if one convolves a function of the form (81) by a smooth approximation to the identity which is of tensor product form $(t_1, \ldots, t_k) \mapsto \varphi_1(t_1) \ldots \varphi_1(t_k)$, one obtains another function of this form. Such a convolution converts a uniformly convergent sequence of functions to a *uniformly smoothly* convergent sequence of functions (that is to say, all derivatives of the functions converge uniformly). From this, we conclude that f_3 can be expressed as the *smooth* limit of functions of the form (81), with each component $f_{1,j}(t_1) \ldots f_{k,j}(t_k)$ supported in the region

$$\left\{ (t_1, \ldots, t_k) \in \mathbb{R}^k : \sum_{i=1}^{k} \max(t_i, \delta_2/8) \le \vartheta/2 - \delta_2/8 \right\}.$$

Thus, we may find such a linear combination

$$f_4(t_1, \ldots, t_k) = \sum_{j=1}^{J} c_j f_{1,j}(t_1) \ldots f_{k,j}(t_k) \tag{82}$$

with $J, c_j, f_{i,j}$ fixed and f_4 not identically zero, with

$$\frac{\sum_{i=1}^{k} \tilde{J}_i(f_4)}{\tilde{I}(f_4)} > m. \tag{83}$$

Furthermore, by construction we have

$$S(f_{1,j}) + \cdots + S(f_{k,j}) < \frac{\vartheta}{2} \le \frac{1}{2} \tag{84}$$

for all $j = 1, \ldots, J$, where $S()$ was defined in (22).

Now we construct the sieve weight $\nu : \mathbb{N} \to \mathbb{R}$ by the formula

$$\nu(n) := \left(\sum_{j=1}^{J} c_j \lambda_{f_{1,j}}(n + h_1) \ldots \lambda_{f_{k,j}}(n + h_k) \right)^2, \tag{85}$$

where the divisor sums λ_f were defined in (16).

Clearly ν is non-negative. Expanding out the square and using Theorem 20(i) and (84), we see that

$$\sum_{\substack{x \le n \le 2x \\ n = b \ (W)}} \nu(n) = (\alpha + o(1)) B^{-k} \frac{x}{\log x}$$

where

$$\alpha := \sum_{j=1}^{J} \sum_{j'=1}^{J} c_j c_{j'} \prod_{i=1}^{k} \int_{0}^{\infty} f'_{i,j}(t_i) f'_{i,j'}(t_i) \, dt_i$$

which factorizes using (82) and (78) as

$$\alpha = \int_{[0,+\infty)^k} \left| \frac{\partial^k}{\partial t_1 \ldots \partial t_k} f_4(t_1, \ldots, t_k) \right|^2 \, dt_1 \ldots dt_k$$

$$= \tilde{I}(f_4).$$

Now consider the sum

$$\sum_{\substack{x \leq n \leq 2x \\ n=b \ (W)}} \nu(n)\theta(n+h_k).$$

By (20), one has

$$\lambda_{f_{k,j}}(n+h_k) = f_{k,j}(0)$$

whenever n gives a non-zero contribution to the above sum. Expanding out the square in (85) again and using Theorem 19(i) and (84) (and the hypothesis EH[ϑ]), we thus see that

$$\sum_{\substack{x \leq n \leq 2x \\ n=b \ (W)}} \nu(n)\theta(n+h_k) = (\beta_k + o(1)) B^{1-k} \frac{x}{\varphi(W)}$$

where

$$\beta_k := \sum_{j=1}^{J} \sum_{j'=1}^{J} c_j c_{j'} f_{i,j}(0) f_{i,j'}(0) \prod_{i=1}^{k-1} \int_0^\infty f'_{i,j}(t_i) f'_{i,j'}(t_i) \, dt_i$$

which factorizes using (82) and (79) as

$$\beta_k = \int_{[0,+\infty)^k} \left| \frac{\partial^k}{\partial t_1 \dots \partial t_{k-1}} f_4(t_1, \dots, t_{k-1}, 0) \right|^2 dt_1 \dots dt_{k-1}$$
$$= \tilde{J}_k(f_4).$$

More generally, we see that

$$\sum_{\substack{x \leq n \leq 2x \\ n=b \ (W)}} \nu(n)\theta(n+h_k) = (\beta_i + o(1)) B^{1-k} \frac{x}{\varphi(W)}$$

for $i = 1, \dots, k$, with $\beta_i := \tilde{J}_i(f_4)$. Applying Lemma 18 and (75), we obtain DHL[$k; m+1$] as required.

Proof of Theorem 24

Now we prove Theorem 24, which uses a very similar argument to that of the previous section. Let k, m, ϖ, δ, F be as in Theorem 24. By performing the same rescaling as in the previous section (but with $1/2 + 2\varpi$ playing the role of ϑ), we see that we can find a fixed square-integrable measurable function F_1 supported on the rescaled truncated simplex

$$\left\{ (t_1, \dots, t_k) \in [0, +\infty)^k : t_1 + \dots + t_k \leq \frac{1}{4} + \varpi - \delta_1; t_1, \dots, t_k < \delta - \delta_1 \right\}$$

for some sufficiently small fixed $\delta_1 > 0$, such that (73) holds. By repeating the arguments of the previous section, we may eventually arrive at a smooth function $f_4 : \mathbb{R}^k \to \mathbb{R}$ of the form (82), which is not identically zero and obeys (83) and such that each component $f_{1,j}(t_1) \ldots f_{k,j}(t_k)$ is supported in the region

$$\left\{ (t_1, \ldots, t_k) \in \mathbb{R}^k : \sum_{i=1}^{k} \max(t_i, \delta_2/8) \leq \frac{1}{4} + \varpi - \delta_2/8; t_1, \ldots, t_k < \delta - \delta_2/8 \right\}$$

for some sufficiently small $\delta_2 > 0$. In particular, one has

$$S(f_{1,j}) + \cdots + S(f_{k,j}) < \frac{1}{4} + \varpi \leq \frac{1}{2}$$

and

$$S(f_{1,j}), \ldots, S(f_{k,j}) < \delta$$

for all $j = 1, \ldots, J$. If we then define ν by (85) as before, and repeat all of the above arguments (but use Theorem 19(ii) and MPZ$[\varpi, \delta]$ in place of Theorem 19(i) and EH$[\vartheta]$), we obtain the claim; we leave the details to the interested reader.

Proof of Theorem 26

Now we prove Theorem 26. Let $k, m, \varepsilon, \vartheta$ be as in that theorem. Then, one may find a square-integrable function $F : [0, +\infty)^k \to \mathbb{R}$ supported on $(1 + \varepsilon) \cdot \mathcal{R}_k$ which is not identically zero, and with

$$\frac{\sum_{i=1}^{k} J_{i,1-\varepsilon}(F)}{I(F)} > \frac{2m}{\vartheta}.$$

By truncating and rescaling as in the 'Proof of Theorem 22' section, we may find a fixed bounded measurable function $F_1 : [0, +\infty)^k \to \mathbb{R}$ on the simplex $(1 + \varepsilon) \left(\frac{\vartheta}{2} - \delta_1 \right) \cdot \mathcal{R}_k$ such that

$$\frac{\sum_{i=1}^{k} J_{i,(1-\varepsilon)\frac{\vartheta}{2}}(F_1)}{I(F_1)} > m.$$

By repeating the arguments in the 'Proof of Theorem 22' section, we may eventually arrive at a smooth function $f_4 : \mathbb{R}^k \to \mathbb{R}$ of the form (82), which is not identically zero and obeys

$$\frac{\sum_{i=1}^{k} \tilde{J}_{i,(1-\varepsilon)\frac{\vartheta}{2}}(f_4)}{\tilde{I}(f_4)} > m \tag{86}$$

with

$$\tilde{J}_{i,(1-\varepsilon)\frac{\vartheta}{2}}(f_4) := \int_{(1-\varepsilon)\frac{\vartheta}{2} \cdot \mathcal{R}_{k-1}} \left| \frac{\partial^{k-1}}{\partial t_1 \ldots \partial t_{i-1} \partial t_{i+1} \ldots \partial t_k} f_4(t_1, \ldots, t_{i-1}, 0, t_{i+1}, \ldots, t_k) \right|^2$$

$$dt_1 \ldots dt_{i-1} dt_{i+1} \ldots dt_k,$$

and such that each component $f_{1,j}(t_1)\dots f_{k,j}(t_k)$ is supported in the region

$$\left\{ (t_1,\dots,t_k) \in \mathbb{R}^k : \sum_{i=1}^{k} \max(t_i, \delta_2/8) \le (1+\varepsilon)\frac{\vartheta}{2} - \frac{\delta_2}{8} \right\}$$

for some sufficiently small $\delta_2 > 0$. In particular, we have

$$S(f_{1,j}) + \cdots + S(f_{k,j}) \le (1+\varepsilon)\frac{\vartheta}{2} - \frac{\delta_2}{8} \tag{87}$$

for all $1 \le j \le J$.

Let $\delta_3 > 0$ be a sufficiently small fixed quantity (smaller than δ_1 or δ_2). By a smooth partitioning, we may assume that all of the $f_{i,j}$ are supported in intervals of length at most δ_3, while keeping the sum

$$\sum_{j=1}^{J} |c_j| |f_{1,j}(t_1)| \dots |f_{k,j}(t_k)| \tag{88}$$

bounded uniformly in $t_1,\dots,t_k$ and in δ_3.

Now let ν be as in (85), and consider the expression

$$\sum_{\substack{x \le n \le 2x \\ n=b\ (W)}} \nu(n).$$

This expression expands as a linear combination of the expressions

$$\sum_{\substack{x \le n \le 2x \\ n=b\ (W)}} \prod_{i=1}^{k} \lambda_{f_{i,j}}(n+h_i)\lambda_{f_{i,j'}}(n+h_i)$$

for various $1 \le j, j' \le J$. We claim that this sum is equal to

$$\left(\prod_{i=1}^{k} \int_0^1 f'_{i,j}(t_i) f'_{i,j'}(t_i)\, dt_i + o(1) \right) B^{-k} \frac{x}{W}.$$

To see this, we divide into two cases. First, suppose that hypothesis (i) from Theorem 26 holds, then from (87) we have

$$\sum_{i=1}^{k} S(f_{i,j}) + S(f_{i,j'}) < (1+\varepsilon)\vartheta \ < 1$$

and the claim follows from Theorem 20(i). Now suppose instead that hypothesis (ii) from Theorem 26 holds, then from (87) one has

$$\sum_{i=1}^{k} S(f_{i,j}) + S(f_{i,j'}) < (1+\varepsilon)\vartheta \ < \frac{k}{k-1}\vartheta,$$

and so from the pigeonhole principle, we have

$$\sum_{1 \le i \le k: i \ne i_0} S(f_{i,j}) + S(f_{i,j'}) < \vartheta$$

for some $i_0 = 1,\dots,k$. The claim now follows from Theorem 20(ii).

Putting these together as in the 'Proof of Theorem 22' section, we conclude that

$$\sum_{\substack{x \le n \le 2x \\ n=b \ (W)}} v(n) = (\alpha + o(1))B^{-k}\frac{x}{W}$$

where

$$\alpha := \tilde{I}(f_4).$$

Now we consider the sum

$$\sum_{\substack{x \le n \le 2x \\ n=b \ (W)}} v(n)\theta(n + h_k). \tag{89}$$

From Proposition 13, we see that we have EH[ϑ] as a consequence of the hypotheses of Theorem 26. However, this and Theorem 19 are not strong enough to obtain an asymptotic for the sum (89), as there is an epsilon loss in (87). But observe that Lemma 18 only requires a *lower* bound on the sum (89), rather than an asymptotic.

To obtain this lower bound, we partition $\{1, \ldots, J\}$ into $\mathcal{J}_1 \cup \mathcal{J}_2$, where $\mathcal{J}_1$ consists of those indices $j \in \{1, \ldots, J\}$ with

$$S(f_{1,j}) + \cdots + S(f_{k-1,j}) < (1 - \varepsilon)\frac{\vartheta}{2} \tag{90}$$

and $\mathcal{J}_2$ is the complement. From the elementary inequality

$$(x_1 + x_2)^2 = x_1^2 + 2x_1 x_2 + x_2^2 \ge (x_1 + 2x_2)x_1,$$

we obtain the pointwise lower bound

$$v(n) \ge \left(\left(\sum_{j \in \mathcal{J}_1} + 2\sum_{j \in \mathcal{J}_2}\right) c_j \lambda_{f_{1,j}}(n + h_1) \ldots \lambda_{f_{k,j}}(n + h_k)\right)$$
$$\times \left(\sum_{j' \in \mathcal{J}_1} c_{j'} \lambda_{f_{1,j'}}(n + h_1) \ldots \lambda_{f_{k,j'}}(n + h_k)\right).$$

The point of performing this lower bound is that if $j \in \mathcal{J}_1 \cup \mathcal{J}_2$ and $j' \in \mathcal{J}_1$, then from (87) and (90) one has

$$\sum_{i=1}^{k-1} S\left(f_{i,j}\right) + S\left(f_{i,j'}\right) < \vartheta$$

which makes Theorem 19(i) available for use. Indeed, for any $j \in \{1, \ldots, J\}$ and $i = 1, \ldots, k$, we have from (87) that

$$S\left(f_{i,j}\right) \le (1 + \varepsilon)\frac{\vartheta}{2} < \vartheta < 1,$$

and so by (20), we have

$$v(n)\theta(n + h_k) \ge \left(\left(\sum_{j \in \mathcal{J}_1} + 2\sum_{j \in \mathcal{J}_2}\right) c_j \lambda_{f_{1,j}}(n + h_1) \ldots \lambda_{f_{k-1,j}}(n + h_{k-1})f_{k,j}(0)\right)$$
$$\times \left(\sum_{j' \in \mathcal{J}_1} c_{j'} \lambda_{f_{1,j'}}(n + h_1) \ldots \lambda_{f_{k-1,j'}}(n + h_{k-1})f_{k,j'}(0)\right)\theta(n + h_k)$$

$$\tag{91}$$

for $x \le n \le 2x$. If we then apply Theorem 19(i) and the hypothesis EH[ϑ], we obtain the lower bound

$$\sum_{\substack{x \le n \le 2x \\ n=b \ (W)}} \nu(n)\theta(n+h_k) \ge (\beta_k - o(1))B^{1-k}\frac{x}{\varphi(W)}$$

with

$$\beta_k := \left(\sum_{j \in \mathcal{J}_1} + 2\sum_{j \in \mathcal{J}_2}\right)\sum_{j' \in \mathcal{J}_1} c_j c_{j'} f_{k,j}(0) f_{k,j'}(0) \prod_{i=1}^{k-1} \int_0^\infty f'_{i,j}(t_i) f'_{i,j'}(t_i) \, dt_i$$

which we can rearrange as

$$\beta_k = \int_{[0,+\infty)^{k-1}} \left(\frac{\partial^{k-1}}{\partial t_1 \ldots \partial t_{k-1}} f_{4,1}(t_1,\ldots,t_{k-1},0) + 2\frac{\partial^{k-1}}{\partial t_1 \ldots \partial t_{k-1}} f_{4,2}(t_1,\ldots,t_{k-1},0)\right)$$
$$\frac{\partial^{k-1}}{\partial t_1 \ldots \partial t_{k-1}} f_{4,1}(t_1,\ldots,t_{k-1},0) \, dt_1 \ldots dt_{k-1}$$

where

$$f_{4,l}(t_1,\ldots,t_k) := \sum_{j \in \mathcal{J}_l} c_j f_{1,j}(t_1) \ldots f_{k,j}(t_k)$$

for $l = 1, 2$. Note that $f_{4,1}, f_{4,2}$ are both bounded pointwise by (88), and their supports only overlap on a set of measure $O(\delta_3)$. We conclude that

$$\beta_k = \tilde{J}_k(f_{4,1}) + O(\delta_3)$$

with the implied constant independent of δ_3, and thus

$$\beta_k = \tilde{J}_{k,(1-\varepsilon)\frac{\vartheta}{2}}(f_4) + O(\delta_3).$$

A similar argument gives

$$\sum_{\substack{x \le n \le 2x \\ n=b \ (W)}} \nu(n)\theta(n+h_i) \ge (\beta_i - o(1))B^{1-k}\frac{x}{\varphi(W)}$$

for $i = 1, \ldots, k$ with

$$\beta_i = \tilde{J}_{i,(1-\varepsilon)\frac{\vartheta}{2}}(f_4) + O(\delta_3).$$

If we choose δ_3 small enough, then the claim DHL[$k; m+1$] now follows from Lemma 18 and (86).

Proof of Theorem 28

Finally, we prove Theorem 28. Let k, m, ε, F be as in that theorem. By rescaling as in previous sections, we may find a square-integrable function $F_1 : [0, +\infty)^k \to \mathbb{R}$ supported on $\left(\frac{k}{k-1}\frac{\vartheta}{2} - \delta_1\right) \cdot \mathcal{R}_k$ for some sufficiently small fixed $\delta_1 > 0$, which is not identically zero, which obeys the bound

$$\frac{\sum_{i=1}^k J_{i,(1-\varepsilon)\frac{\vartheta}{2}}(F_1)}{I(F_1)} > m$$

and also obeys the vanishing marginal condition (35) whenever $t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_k \geq 0$ are such that

$$t_1 + \cdots + t_{i-1} + t_{i+1} + \cdots + t_k > (1 + \varepsilon)\frac{\vartheta}{2} - \delta_1.$$

As before, we pass from F_1 to F_2 by a spatial translation, and from F_2 to F_3 by a regularisation; crucially, we note that both of these operations interact well with the vanishing marginal condition (35), with the end product being that we obtain a smooth function $F_3 : [0, +\infty)^k \to \mathbb{R}$, supported on the region

$$\left\{ (t_1, \ldots, t_k) \in \mathbb{R}^k : t_1 \cdots + t_k \leq \frac{k}{k-1}\frac{\vartheta}{2} - \frac{\delta_2}{2}; t_1, \ldots, t_k \geq \frac{\delta_2}{2} \right\}$$

for some sufficiently small $\delta_2 > 0$, which is not identically zero, obeying the bound

$$\frac{\sum_{i=1}^k J_{i,(1-\varepsilon)\frac{\vartheta}{2}}(F_3)}{I(F_3)} > m$$

and also obeying the vanishing marginal condition (35) whenever $t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_k \geq 0$ are such that

$$t_1 + \cdots + t_{i-1} + t_{i+1} + \cdots + t_k > (1 + \varepsilon)\frac{\vartheta}{2} - \frac{\delta_2}{2}.$$

As before, we now define the function $f_3 : \mathbb{R}^k \to \mathbb{R}$ by

$$f_3(t_1, \ldots, t_k) := \int_{s_1 \geq t_1, \ldots, s_k \geq t_k} F_3(s_1, \ldots, s_k) \, ds_1 \ldots ds_k,$$

and thus, f_3 is smooth, not identically zero and supported on the region

$$\left\{ (t_1, \ldots, t_k) \in \mathbb{R}^k : \sum_{i=1}^k \max(t_i, \delta_2/2) \leq \frac{k}{k-1}\frac{\vartheta}{2} - \frac{\delta_2}{2} \right\}.$$

Furthermore, from the vanishing marginal condition, we see that we also have

$$f_3(t_1, \ldots, t_k) = 0$$

whenever we have some $1 \leq i \leq k$ for which $t_i \leq \delta_2/2$ and

$$t_1 + \cdots + t_{i-1} + t_{i+1} + \cdots + t_k \geq (1 + \varepsilon)\frac{\vartheta}{2} - \frac{\delta_2}{2}.$$

From the fundamental theorem of calculus as before, we have

$$\frac{\sum_{i=1}^k \tilde{J}_{i,(1-\varepsilon)\frac{\vartheta}{2}}(f_3)}{\tilde{I}(f_3)} > m.$$

Using the Stone-Weierstrass theorem as before, we can then find a function f_4 of the form

$$(t_1, \ldots, t_k) \mapsto \sum_{j=1}^J c_j f_{1,j}(t_1) \ldots f_{k,j}(t_k) \tag{92}$$

where $c_1, \ldots, c_J$ are real scalars, and $f_{i,j} : \mathbb{R} \to \mathbb{R}$ are smooth functions supported of intervals of length at most $\delta_3 > 0$ for some sufficiently small $\delta_3 > 0$, with the support of each component $f_{1,j}(t_1) \ldots f_{k,j}(t_k)$ supported in the region

$$\left\{ (t_1, \ldots, t_k) \in \mathbb{R}^k : \sum_{i=1}^{k} \max(t_i, \delta_2/8) \le \frac{k}{k-1} \frac{\vartheta}{2} - \delta_2/8 \right\}$$

and avoiding the regions

$$\left\{ (t_1, \ldots, t_k) \in \mathbb{R}^k : t_i \le \delta_2/8; \quad t_1 + \cdots + t_{i-1} + t_{i+1} + \cdots + t_k \ge (1+\varepsilon)\frac{\vartheta}{2} - \delta_2/8 \right\}$$

for each $i = 1, \ldots, k$, and such that

$$\frac{\sum_{i=1}^{k} \tilde{J}_{i,(1-\varepsilon)\frac{\vartheta}{2}}(f_4)}{\tilde{I}(f_4)} > m.$$

In particular, for any $j = 1, \ldots, J$ we have

$$S\left(f_{1,j}\right) + \cdots + S\left(f_{k,j}\right) < \frac{k}{k-1}\frac{\vartheta}{2} < \frac{1}{2}\frac{k}{k-1} \le 1 \tag{93}$$

and for any $i = 1, \ldots, k$ with $f_{k,i}$ not vanishing at zero, we have

$$S\left(f_{1,j}\right) + \cdots + S\left(f_{k,i-1}\right) + S\left(f_{k,i+1}\right) + \cdots + S\left(f_{k,j}\right) < (1+\varepsilon)\frac{\vartheta}{2}. \tag{94}$$

Let ν be defined by (85). From (93), the hypothesis GEH$[\vartheta]$, and the argument from the previous section used to prove Theorem 26(ii), we have

$$\sum_{\substack{x \le n \le 2x \\ n = b \ (W)}} \nu(n) = (\alpha + o(1)) B^{-k} \frac{x}{W}$$

where

$$\alpha := \tilde{I}\left(f_4\right).$$

Similarly, from (94) (and the upper bound $S\left(f_{i,j}\right) < 1$ from (93)), the hypothesis EH$[\vartheta]$ (which is available by Proposition 13), and the argument from the previous section, we have

$$\sum_{\substack{x \le n \le 2x \\ n = b \ (W)}} \nu(n)\theta(n + h_i) \ge (\beta_i - o(1)) B^{1-k} \frac{x}{\varphi(W)}$$

for $i = 1, \ldots, k$ with

$$\beta_i = \tilde{J}_{i,(1-\varepsilon)\frac{\vartheta}{2}}(f_4) + O(\delta_3).$$

Setting δ_3 small enough, the claim DHL$[k; m+1]$ now follows from Lemma 18.

Asymptotic analysis

We now establish upper and lower bounds on the quantity M_k defined in (33), as well as for the related quantities appearing in Theorem 24.

To obtain an upper bound on M_k, we use the following consequence of the Cauchy-Schwarz inequality.

Lemma 34 (Cauchy-Schwarz)**.** *Let* $k \geq 2$, *and suppose that there exist positive measurable functions* $G_i : \mathcal{R}_k \rightarrow (0, +\infty)$ *for* $i = 1, \ldots, k$ *such that*

$$\int_0^\infty G_i(t_1, \ldots, t_k) \, dt_i \leq 1 \tag{95}$$

for all $t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_k \geq 0$, *where we extend* G_i *by zero to all of* $[0, +\infty)^k$. *Then, we have*

$$M_k \leq \mathrm{ess} \sup_{(t_1, \ldots, t_k) \in \mathcal{R}_k} \sum_{i=1}^k \frac{1}{G_i(t_1, \ldots, t_k)}. \tag{96}$$

Here ess sup *refers to essential supremum (thus, we may ignore a subset of* $\mathcal{R}_k$ *of measure zero in the supremum).*

Proof. Let $F : [0, +\infty)^k \rightarrow \mathbb{R}$ be a square-integrable function supported on $\mathcal{R}_k$. From the Cauchy-Schwarz inequality and (95), we have

$$\left(\int_0^\infty F(t_1, \ldots, t_k) \, dt_i \right)^2 \leq \int_0^\infty \frac{F(t_1, \ldots, t_k)^2}{G_i(t_1, \ldots, t_k)} \, dt_i$$

for any $t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_k \geq 0$. Inserting this into (32) and integrating, we conclude that

$$J_i(F) \leq \int_{\mathcal{R}_k} \frac{F(t_1, \ldots, t_k)^2}{G_i(t_1, \ldots, t_k)} \, dt_1 \ldots dt_k.$$

Summing in i and using (31), (33), and (96), we obtain the claim. $\square$

As a corollary, we can compute M_k exactly if we can locate a positive eigenfunction:

Corollary 35. *Let* $k \geq 2$, *and suppose that there exists a positive function* $F : \mathcal{R}_k \rightarrow (0, +\infty)$ *obeying the eigenfunction equation*

$$\lambda F(t_1, \ldots, t_k) = \sum_{i=1}^k \int_0^\infty F(t_1, \ldots, t_{i-1}, t_i', t_{i+1}, \ldots, t_k) \, dt_i' \tag{97}$$

for some $\lambda > 0$ *and all* $(t_1, \ldots, t_k) \in \mathcal{R}_k$, *where we extend* F *by zero to all of* $[0, +\infty)^k$. *Then,* $\lambda = M_k$.

Proof. On the one hand, if we integrate (97) against F and use (31) and (32), we see that

$$\lambda I(F) = \sum_{i=1}^k J_i(F),$$

and thus by (33), we see that $M_k \geq \lambda$. On the other hand, if we apply Lemma 34 with

$$G_i(t_1, \ldots, t_k) := \frac{F(t_1, \ldots, t_k)}{\int_0^\infty F(t_1, \ldots, t_{i-1}, t_i', t_{i+1}, \ldots, t_k) \, dt_i'},$$

we see that $M_k \leq \lambda$, and the claim follows. $\square$

This allows for an exact calculation of M_2:

Corollary 36 (Computation of M_2). *We have*

$$M_2 = \frac{1}{1 - W(1/e)} = 1.38593\ldots$$

where the Lambert W-function $W(x)$ is defined for positive x as the unique positive solution to $x = W(x)e^{W(x)}$.

Proof. If we set $\lambda := \frac{1}{1-W(1/e)} = 1.38593\ldots$, then a brief calculation shows that

$$2\lambda - 1 = \lambda \log \lambda - \lambda \log(\lambda - 1). \tag{98}$$

Now if we define the function $f : [0, 1] \to [0, +\infty)$ by the formula

$$f(x) := \frac{1}{\lambda - 1 + x} + \frac{1}{2\lambda - 1} \log \frac{\lambda - x}{\lambda - 1 + x},$$

then a further brief calculation shows that

$$\int_0^{1-x} f(y)\, dy = \frac{\lambda - 1 + x}{2\lambda - 1} \log \frac{\lambda - x}{\lambda - 1 + x} + \frac{\lambda \log \lambda - \lambda \log(\lambda - 1)}{2\lambda - 1}$$

for any $0 \le x \le 1$, and hence by (98) that

$$\int_0^{1-x} f(y)\, dy = (\lambda - 1 + x)f(x).$$

If we then define the function $F : \mathcal{R}_2 \to (0, +\infty)$ by $F(x, y) := f(x) + f(y)$, we conclude that

$$\int_0^{1-x} F(x', y)\, dx' + \int_0^{1-y} F(x, y')\, dy' = \lambda F(x, y)$$

for all $(x, y) \in \mathcal{R}_2$, and the claim now follows from Corollary 35. □

We conjecture that a positive eigenfunction for M_k exists for all $k \ge 2$, not just for $k = 2$; however, we were unable to produce any such eigenfunctions for $k > 2$. Nevertheless, Lemma 34 still gives us a general upper bound:

Corollary 37. *We have $M_k \le \frac{k}{k-1} \log k$ for any $k \ge 2$.*

Thus, for instance, one has $M_2 \le 2 \log 2 = 1.38629\ldots$, which compares well with Corollary 36. On the other hand, Corollary 37 also gives

$$M_4 \le \frac{4}{3} \log 4 = 1.8454\ldots,$$

so that one cannot hope to establish DHL[4; 2] (or DHL[3; 2]) solely through Theorem 22 even when assuming GEH, and must rely instead on more sophisticated criteria for DHL[k; m] such as Theorem 26 or Theorem 28.

Proof. If we set $G_i : \mathcal{R}_k \to (0, +\infty)$ for $i = 1, \ldots, k$ to be the functions

$$G_i(t_1, \ldots, t_k) := \frac{k-1}{\log k} \frac{1}{1 - t_1 - \cdots - t_k + kt_i}$$

then direct calculation shows that

$$\int_0^\infty G_i(t_1,\ldots,t_k)\,dt_i \le 1$$

for all $t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_k \ge 0$, where we extend G_i by zero to all of $[0,+\infty)^k$. On the other hand, we have

$$\sum_{i=1}^k \frac{1}{G_i(t_1,\ldots,t_k)} = \frac{k}{k-1}\log k$$

for all $(t_1,\ldots,t_k) \in \mathcal{R}_k$. The claim now follows from Lemma 34. $\qquad\square$

The upper bound arguments for M_k can be extended to other quantities such as $M_{k,\varepsilon}$, although the bounds do not appear to be as sharp in that case. For instance, we have the following variant of Lemma 37, which shows that the improvement in constants when moving from M_k to $M_{k,\varepsilon}$ is asymptotically modest:

Proposition 38. *For any $k \ge 2$ and $\varepsilon \ge 0$, we have*

$$M_{k,\varepsilon} \le \frac{k}{k-1}\log(2k-1).$$

Proof. Let $F : [0,+\infty)^k \to \mathbb{R}$ be a square-integrable function supported on $(1+\varepsilon)\cdot\mathcal{R}_k$. If $i = 1,\ldots,k$ and $(t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_k) \in (1-\varepsilon)\cdot\mathcal{R}_k$, then if we write $s := 1 - t_1 - \cdots - t_{i-1} - t_{i+1} - \cdots - t_k$, we have $s \ge \varepsilon$ and hence

$$\int_0^{1-t_1-\cdots-t_{i-1}-t_{i+1}-\cdots-t_k+\varepsilon} \frac{1}{1-t_1-\cdots-t_k+kt_i}\,dt_i = \int_0^{s+\varepsilon} \frac{1}{s+(k-1)t_i}\,dt_i$$

$$= \frac{1}{k-1}\log\frac{ks+(k-1)\varepsilon}{s}$$

$$\le \frac{1}{k-1}\log(2k-1).$$

By Cauchy-Schwarz, we conclude that

$$\left(\int_0^\infty F(t_1,\ldots,t_k)\,dt_i\right)^2 \le \frac{1}{k-1}\log(2k-1)\int_0^\infty (1-t_1-\cdots-t_k+kt_i)F(t_1,\ldots,t_k)^2\,dt_i.$$

Integrating in $t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_k$ and summing in i, we obtain the claim. $\qquad\square$

Remark 39. The same argument, using the weight $1 + a(-t_1 - \cdots - t_k + kt_i)$, gives the more general inequality

$$M_{k,\varepsilon} \le \frac{k}{a(k-1)}\log\left(k + \frac{(a(1+\varepsilon)-1)(k-1)}{1-a(1-\varepsilon)}\right)$$

whenever $\frac{1}{1+\varepsilon} < a < \frac{1}{1-\varepsilon}$; the case $a = 1$ is Proposition 38, and the limiting case $a = \frac{1}{1+\varepsilon}$ recovers Lemma 37 when one sends ε to zero.

One can also adapt the computations in Corollary 36 to obtain exact expressions for $M_{2,\varepsilon}$, although the calculations are rather lengthy and will only be summarized here. For fixed $0 < \varepsilon < 1$, the eigenfunctions F one seeks should take the form

$$F(x,y) := f(x) + f(y)$$

for $x, y \geq 0$ and $x + y \leq 1 + \varepsilon$, where

$$f(x) := 1_{x \leq 1-\varepsilon} \int_0^{1+\varepsilon-x} F(x,t)\, dt.$$

In the regime $0 < \varepsilon < 1/3$, one can calculate that f will (up to scalar multiples) take the form

$$f(x) = 1_{x \leq 2\varepsilon} \frac{C_1}{\lambda - 1 - \varepsilon + x}$$
$$+ 1_{2\varepsilon \leq x \leq 1-\varepsilon} \left(\frac{\log(\lambda - x) - \log(\lambda - 1 - \varepsilon + x)}{2\lambda - 1 - \varepsilon} + \frac{1}{\lambda - 1 - \varepsilon + x} \right)$$

where

$$C_1 := \frac{\log(\lambda - 2\varepsilon) - \log(\lambda - 1 + \varepsilon)}{1 - \log(\lambda - 1 + \varepsilon) + \log(\lambda - 1 - \varepsilon)}$$

and λ is the largest root of the equation

$$1 = C_1(\log(\lambda - 1 + \varepsilon) - \log(\lambda - 1 - \varepsilon)) - \log(\lambda - 1 + \varepsilon)$$
$$+ \frac{(\lambda - 1 + \varepsilon)\log(\lambda - 1 + \varepsilon) - (\lambda - 2\varepsilon)\log(\lambda - 2\varepsilon)}{2\lambda - 1 - \varepsilon}.$$

In the regime $1/3 \leq \varepsilon < 1$, the situation is significantly simpler, and one has the exact expressions

$$f(x) = \frac{1_{x \leq 1-\varepsilon}}{\lambda - 1 - \varepsilon + x}$$

and

$$\lambda = \frac{e(1 + \varepsilon) - 2\varepsilon}{e - 1}.$$

In both cases, a variant of Corollary 35 can be used to show that $M_{2,\varepsilon}$ will be equal to λ; thus, for instance,

$$M_{2,\varepsilon} = \frac{e(1 + \varepsilon) - 2\varepsilon}{e - 1}$$

for $1/3 \leq \varepsilon < 1$. In particular, $M_{2,\varepsilon}$ increases to 2 in the limit $\varepsilon \to 1$; the lower bound $\liminf_{\varepsilon \to 1} M_{2,\varepsilon} \geq 2$ can also be established by testing with the function $F(x,y) := 1_{x \leq \delta, y \leq 1+\varepsilon-\delta} + 1_{y \leq \delta, x \leq 1+\varepsilon-\delta}$ for some sufficiently small $\delta > 0$.

Now we turn to lower bounds on M_k, which are of more relevance for the purpose of establishing results such as Theorem 23. If one restricts attention to those functions $F : \mathcal{R}_k \to \mathbb{R}$ of the special form $F(t_1, \ldots, t_k) = f(t_1 + \cdots + t_k)$ for some function $f : [0,1] \to \mathbb{R}$, then the resulting variational problem has been optimized in previous works [39] (and originally in an unpublished work of Conrey), giving rise to the lower bound

$$M_k \geq \frac{4k(k-1)}{j_{k-2}^2}$$

where j_{k-2} is the first positive zero of the Bessel function J_{k-2}. This lower bound is reasonably strong for small k; for instance, when $k = 2$ it shows that

$$M_2 \geq 1.383\ldots$$

which compares well with Corollary 36, and also shows that $M_6 > 2$, recovering the result of Goldston, Pintz, and Yıldırım that DHL[6; 2] (and hence $H_1 \leq 16$) was true on the Elliott-Halberstam conjecture. However, one can show that $\frac{4k(k-1)}{j_{k-2}^2} < 4$ for all k (see [36]), so this lower bound cannot be used to force M_k to be larger than 4.

In [5], the lower bound

$$M_k \geq \log k - 2 \log \log k - 2 \tag{99}$$

was established for all sufficiently large k. In fact, the arguments in [5] can be used to show this bound for all $k \geq 200$ (for $k < 200$, the right-hand side of (99) is either negative or undefined). Indeed, if we use the bound ([5], (7.19)) with A chosen so that $A^2 e^A = k$, then $3 < A < \log k$ when $k \geq 200$, hence $e^A = k/A^2 > k/\log^2 k$ and so $A \geq \log k - 2 \log \log k$. By using the bounds $\frac{A}{e^A - 1} < \frac{1}{6}$ (since $A > 3$) and $e^A/k = 1/A^2 < 1/9$, we see that the right-hand side of ([5], (8.17)) exceeds $A - \frac{1}{(1 - 1/6 - 1/9)^2} \geq A - 2$, which gives (99).

We will remove the $\log \log k$ term in (99) via the following explicit estimate.

Theorem 40. *Let $k \geq 2$, and let $c, T, \tau > 0$ be parameters. Define the function g : $[0, T] \to \mathbb{R}$ by*

$$g(t) := \frac{1}{c + (k-1)t} \tag{100}$$

and the quantities

$$m_2 := \int_0^T g(t)^2 \, dt \tag{101}$$

$$\mu := \frac{1}{m_2} \int_0^T t g(t)^2 \, dt \tag{102}$$

$$\sigma^2 := \frac{1}{m_2} \int_0^T t^2 g(t)^2 \, dt - \mu^2. \tag{103}$$

Assume the inequalities

$$k\mu \leq 1 - \tau \tag{104}$$

$$k\mu < 1 - T \tag{105}$$

$$k\sigma^2 < (1 + \tau - k\mu)^2. \tag{106}$$

Then, one has

$$\frac{k}{k-1} \log k - M_k^{[T]} \leq \frac{k}{k-1} \frac{Z + Z_3 + WX + VU}{(1 + \tau/2)\left(1 - \frac{k\sigma^2}{(1+\tau-k\mu)^2}\right)} \tag{107}$$

where Z, Z_3, W, X, V, U are the explicitly computable quantities

$$Z := \frac{1}{\tau} \int_1^{1+\tau} \left(r \left(\log \frac{r - k\mu}{T} + \frac{k\sigma^2}{4(r - k\mu)^2 \log \frac{r-k\mu}{T}} \right) + \frac{r^2}{4kT} \right) dr \tag{108}$$

$$Z_3 := \frac{1}{m_2} \int_0^T kt \log \left(1 + \frac{t}{T} \right) g(t)^2 \, dt \tag{109}$$

$$W := \frac{1}{m_2} \int_0^T \log \left(1 + \frac{\tau}{kt} \right) g(t)^2 \, dt \tag{110}$$

$$X := \frac{\log k}{\tau} c^2 \tag{111}$$

$$V := \frac{c}{m_2} \int_0^T \frac{1}{2c + (k-1)t} g(t)^2 \, dt \tag{112}$$

$$U := \frac{\log k}{c} \int_0^1 \left((1 + u\tau - (k-1)\mu - c)^2 + (k-1)\sigma^2 \right) \, du. \tag{113}$$

Of course, since $M_k^{[T]} \leq M_k$, the bound (107) also holds with $M_k^{[T]}$ replaced by M_k.

Proof. From (33), we have

$$\sum_{i=1}^k J_i(F) \leq M_k^{[T]} I(F)$$

whenever $F : [0, +\infty)^k \to \mathbb{R}$ is square-integrable and supported on $[0, T]^k \cap \mathcal{R}_k$. By rescaling, we conclude that

$$\sum_{i=1}^k J_i(F) \leq r M_k^{[T]} I(F)$$

whenever $r > 0$ and $F : [0, +\infty)^k \to \mathbb{R}$ is square-integrable and supported on $[0, rT]^k \cap r \cdot \mathcal{R}_k$. We apply this inequality with the function

$$F(t_1, \ldots, t_k) := 1_{t_1 + \cdots + t_k \leq r} g(t_1) \ldots g(t_k)$$

where $r > 1$ is a parameter which we will eventually average over, and g is extended by zero to $[0, +\infty)$. We thus have

$$I(F) = m_2^k \int_0^\infty \ldots \int_0^\infty 1_{t_1 + \cdots + t_k \leq r} \prod_{i=1}^k \frac{g(t_i)^2 \, dt_i}{m_2}.$$

We can interpret this probabilistically as

$$I(F) = m_2^k \mathbb{P}(X_1 + \cdots + X_k \leq r)$$

where $X_1, \ldots, X_k$ are independent random variables taking values in $[0, T]$ with probability distribution $\frac{1}{m_2} g(t)^2 \, dt$. In a similar fashion, we have

$$J_k(F) = m_2^{k-1} \int_0^\infty \ldots \int_0^\infty \left(\int_{[0, r - t_1 - \cdots - t_{k-1}]} g(t) \, dt \right)^2 \prod_{i=1}^{k-1} \frac{g(t_i)^2 \, dt_i}{m_2},$$

where we adopt the convention that $\int_{[a,b]}$ vanishes when $b < a$. In probabilistic language, we thus have

$$J_k(F) = m_2^{k-1} \mathbb{E} \left(\int_{[0, r - X_1 - \cdots - X_{k-1}]} g(t) \, dt \right)^2.$$

Also by symmetry, we see that $J_i(F) = J_k(F)$ for all $i = 1, \ldots, k$. Putting all these together, we conclude that

$$\mathbb{E} \left(\int_0^{r - X_1 - \cdots - X_{k-1}} g(t) \, dt \right)^2 \leq \frac{m_2 M_k^{[T]} r}{k} \mathbb{P}(X_1 + \cdots + X_k \geq r)$$

for all $r > 1$. Writing $S_i := X_1 + \cdots + X_i$, we abbreviate this as

$$\mathbb{E} \left(\int_{[0, r - S_{k-1}]} g(t) \, dt \right)^2 \leq \frac{m_2 M_k^{[T]} r}{k} \mathbb{P}(S_k \geq r). \tag{114}$$

Now we run a variant of the Cauchy-Schwarz argument used to prove Corollary 37. If, for fixed $r > 0$, we introduce the random function $h : (0, +\infty) \to \mathbb{R}$ by the formula

$$h(t) := \frac{1}{r - S_{k-1} + (k-1)t} 1_{S_{k-1} < r} \tag{115}$$

and observe that whenever $S_{k-1} < r$, we have

$$\int_{[0, r - S_{k-1}]} h(t) \, dt = \frac{\log k}{k - 1} \tag{116}$$

and thus by the Legendre identity, we have

$$\left(\int_{[0, r - S_{k-1}]} g(t) \, dt \right)^2 = \frac{\log k}{k - 1} \int_{[0, r - S_{k-1}]} \frac{g(t)^2}{h(t)} \, dt - \frac{1}{2} \int_{[0, r - S_{k-1}]} \int_{[0, r - S_{k-1}]} \frac{(g(s)h(t) - g(t)h(s))^2}{h(s)h(t)} \, ds \, dt$$

for $S_{k-1} < r$; but the claim also holds when $r \leq S_{k-1}$ since all integrals vanish in that case. On the other hand, we have

$$\mathbb{E} \int_{[0, r - S_{k-1}]} \frac{g(t)^2}{h(t)} \, dt = m_2 \mathbb{E} (r - S_{k-1} + (k-1)X_k) 1_{X_k \leq r - S_{k-1}}$$

$$= m_2 \mathbb{E} (r - S_k + k X_k) 1_{S_k \leq r}$$

$$= m_2 \mathbb{E} r 1_{S_k \leq r}$$

$$= m_2 r \mathbb{P}(S_k \leq r)$$

where we have used symmetry to get the third equality. We conclude that

$$\mathbb{E} \left(\int_{[0, r - S_{k-1}]} g(t) \, dt \right)^2 = \frac{\log k}{k - 1} m_2 r \mathbb{P}(S_k \leq r) - \frac{1}{2} \mathbb{E} \int_{[0, r - S_{k-1}]} \int_{[0, r - S_{k-1}]} \frac{(g(s)h(t) - g(t)h(s))^2}{h(s)h(t)} \, ds \, dt.$$

Combining this with (114), we conclude that

$$\Delta r \mathbb{P}(S_k \le r) \le \frac{k}{2m_2} \mathbb{E} \int_{[0,r-S_{k-1}]} \int_{[0,r-S_{k-1}]} \frac{\big(g(s)h(t) - g(t)h(s)\big)^2}{h(s)h(t)} \, ds dt$$

where

$$\Delta := \frac{k}{k-1} \log k - M_k^{[T]}.$$

Splitting into regions where s, t are less than T or greater than T, and noting that $g(s)$ vanishes for $s > T$, we conclude that

$$\Delta r \mathbb{P}(S_k \le r) \le Y_1(r) + Y_2(r)$$

where

$$Y_1(r) := \frac{k}{m_2} \mathbb{E} \int_{[0,T]} \int_{[T,r-S_{k-1}]} \frac{g(t)^2}{h(t)} h(s) \, ds dt$$

and

$$Y_2(r) := \frac{k}{2m_2} \mathbb{E} \int_{[0,\min(T,r-S_{k-1})]} \int_{[0,\min(T,r-S_{k-1})]} \frac{\big(g(s)h(t) - g(t)h(s)\big)^2}{h(s)h(t)} \, ds dt.$$

We average this from $r = 1$ to $r = 1 + \tau$, to conclude that

$$\Delta \left(\frac{1}{\tau} \int_1^{1+\tau} r \mathbb{P}(S_k \le r) \, dr \right) \le \frac{1}{\tau} \int_1^{1+\tau} Y_1(r) \, dr + \frac{1}{\tau} \int_1^{1+\tau} Y_2(r) \, dr.$$

Thus, to prove (107), it suffices (by (106)) to establish the bounds

$$\frac{1}{\tau} \int_1^{1+\tau} r \mathbb{P}(S_k \le r) \, dr \ge (1 + \tau/2) \left(1 - \frac{k\sigma^2}{(1 + \tau - k\mu)^2} \right), \tag{117}$$

$$\frac{k}{k-1} Y_1(r) \le Z + Z_3 \tag{118}$$

for all $1 < r \le 1 + \tau$, and

$$\frac{1}{\tau} \int_1^{1+\tau} Y_2(r) \, dr \le \frac{k}{k-1} (WX + VU). \tag{119}$$

We begin with (117). Since

$$\frac{1}{\tau} \int_1^{1+\tau} r \, dr = 1 + \frac{\tau}{2},$$

it suffices to show that

$$\mathbb{P}(S_k > 1 + \tau) \le \frac{k\sigma^2}{(1 + \tau - k\mu)^2}.$$

But from (102) and (103), we see that each X_i has mean μ and variance σ^2, so S_k has mean $k\mu$ and variance $k\sigma^2$. The claim now follows from Chebyshev's inequality and (104).

Now we show (118). The quantity $Y_1(r)$ is vanishing unless $r - S_{k-1} \geq T$. Using the crude bound $h(s) \leq \frac{1}{(k-1)s}$ from (115), we see that

$$\int_{[T, r - S_{k-1}]} h(s)\, ds \leq \frac{1}{k-1} \log_+ \frac{r - S_{k-1}}{T}$$

where $\log_+(x) := \max(\log x, 0)$. We conclude that

$$Y_1(r) \leq \frac{k}{k-1} \frac{1}{m_2} \mathbb{E} \int_{[0,T]} \frac{g(t)^2}{h(t)}\, dt \log_+ \frac{r - S_{k-1}}{T}.$$

We can rewrite this as

$$Y_1(r) \leq \frac{k}{k-1} \mathbb{E} \frac{1_{S_k \leq r}}{h(X_k)} \log_+ \frac{r - S_{k-1}}{T}.$$

By (115), we have

$$\frac{1_{S_k \leq r}}{h(X_k)} = (r - S_k + kX_k)\, 1_{S_k \leq r}.$$

Also, from the elementary bound $\log_+(x + y) \leq \log_+ x + \log(1 + y)$ for any $x, y \geq 0$, we see that

$$\log_+ \frac{r - S_{k-1}}{T} \leq \log_+ \frac{r - S_k}{T} + \log\left(1 + \frac{X_k}{T}\right).$$

We conclude that

$$Y_1(r) \leq \frac{k}{k-1} \mathbb{E}\, (r - S_k + kX_k)\left(\log_+ \frac{r - S_k}{T} + \log\left(1 + \frac{X_k}{T}\right)\right) 1_{S_k \leq r}$$

$$\leq \frac{k}{k-1}\left(\mathbb{E}(r - S_k + kX_k) \log_+ \frac{r - S_k}{T} + \max(r - S_k, 0)\frac{X_k}{T} + kX_k \log\left(1 + \frac{X_k}{T}\right)\right)$$

using the elementary bound $\log(1 + y) \leq y$. Symmetrizing in the $X_1, \ldots, X_k$, we conclude that

$$Y_1(r) \leq \frac{k}{k-1}(Z_1(r) + Z_2(r) + Z_3) \tag{120}$$

where

$$Z_1(r) := \mathbb{E} r \log_+ \frac{r - S_k}{T}$$

$$Z_2(r) := \mathbb{E}(r - S_k) 1_{S_k \leq r} \frac{S_k}{kT}$$

and Z_3 was defined in (109).

For the minor error term Z_2, we use the crude bound $(r - S_k) 1_{S_k \leq r} S_k \leq \frac{r^2}{4}$, so

$$Z_2(r) \leq \frac{r^2}{4kT}. \tag{121}$$

For Z_1, we upper bound $\log_+ x$ by a quadratic expression in x. More precisely, we observe the inequality

$$\log_+ x \leq \frac{(x - 2a \log a - a)^2}{4a^2 \log a}$$

for any $a > 1$ and $x \in \mathbb{R}$, since the left-hand side is concave in x for $x \geq 1$, while the right-hand side is convex in x, non-negative, and tangent to the left-hand side at $x = a$. We conclude that

$$\log_+ \frac{r - S_k}{T} \leq \frac{(r - S_k - 2aT \log a - aT)^2}{4a^2 T^2 \log a}.$$

On the other hand, from (102) and (103), we see that each X_i has mean μ and variance σ^2, so S_k has mean $k\mu$ and variance $k\sigma^2$. We conclude that

$$Z_1(r) \leq r \frac{(r - k\mu - 2aT \log a - aT)^2 + k\sigma^2}{4a^2 T^2 \log a}$$

for any $a > 1$.

From (105) and the assumption $r > 1$, we may choose $a := \frac{r - k\mu}{T}$ here, leading to the simplified formula

$$Z_1(r) \leq r \left(\log \frac{r - k\mu}{T} + \frac{k\sigma^2}{4(r - k\mu)^2 \log \frac{r - k\mu}{T}} \right). \tag{122}$$

From (120), (121), (122), and (108) we conclude (118).

Finally, we prove (119). Here, we finally use the specific form (100) of the function g. Indeed, from (100) and (115), we observe the identity

$$g(t) - h(t) = (r - S_{k-1} - c)g(t)h(t)$$

for $t \in [0, \min(r - S_{k-1}, T)]$. Thus,

$$Y_2(r) = \frac{k}{2m_2} \mathbb{E} \int_{[0,\min(r-S_{k-1},T)]} \int_{[0,\min(r-S_{k-1},T)]} \frac{\left((g-h)(s)h(t) - (g-h)(t)h(s)\right)^2}{h(s)h(t)} \, dsdt$$

$$= \frac{k}{2m_2} \mathbb{E}(r - S_{k-1} - c)^2 \int_{[0,\min(r-S_{k-1},T)]} \int_{[0,\min(r-S_{k-1},T)]} \left(g(s) - g(t)\right)^2 h(s)h(t) \, dsdt.$$

Using the crude bound $(g(s) - g(t))^2 \leq g(s)^2 + g(t)^2$ and using symmetry, we conclude

$$Y_2(r) \leq \frac{k}{m_2} \mathbb{E}\left(r - S_{k-1} - c\right)^2 \int_{[0,\min(r-S_{k-1},T)]} \int_{[0,\min(r-S_{k-1},T)]} g(s)^2 h(s)h(t) \, dsdt.$$

From (116) and (115), we conclude that

$$Y_2(r) \leq \frac{k}{k-1} Z_4(r)$$

where

$$Z_4(r) := \frac{\log k}{m_2} \mathbb{E}\left((r - S_{k-1} - c)^2 \int_{[0,\min(r-S_{k-1},T)]} \frac{g(s)^2}{r - S_{k-1} + (k-1)s} \, ds \right).$$

To prove (119), it thus suffices (after making the change of variables $r = 1 + u\tau$) to show that

$$\int_0^1 Z_4(1 + u\tau) \, du \leq WX + VU. \tag{123}$$

We will exploit the averaging in u to deal with the singular nature of the factor $\frac{1}{r - S_{k-1} + (k-1)s}$. By Fubini's theorem, the left-hand side of (123) may be written as

$$\frac{\log k}{m_2} \mathbb{E} \int_0^1 Q(u) \, du$$

where $Q(u)$ is the random variable

$$Q(u) := \left(1 + u\tau - S_{k-1} - c\right)^2 \int_{[0,\min(1+u\tau-S_{k-1},T)]} \frac{g(s)^2}{1 + u\tau - S_{k-1} + (k-1)s} \, ds.$$

Note that $Q(u)$ vanishes unless $1 + u\tau - S_{k-1} > 0$. Consider first the contribution of those $Q(u)$ for which

$$0 < 1 + u\tau - S_{k-1} \le 2c.$$

In this regime, we may bound

$$\left(1 + u\tau - S_{k-1} - c\right)^2 \le c^2,$$

so this contribution to (123) may be bounded by

$$\frac{\log k}{m_2} c^2 \mathbb{E} \int_{[0,T]} g(s)^2 \left(\int_0^1 \frac{1_{1+u\tau-S_{k-1} \ge s}}{1 + u\tau - S_{k-1} + (k-1)s} \, du \right) ds.$$

Observe on making the change of variables $v := 1 + u\tau - S_{k-1} + (k-1)s$ that

$$\int_0^1 \frac{1_{1+u\tau-S_{k-1} \ge s}}{1 + u\tau - S_{k-1} + (k-1)s} \, du = \frac{1}{\tau} \int_{[\max(ks, 1-S_{k-1}+(k-1)s), 1-S_{k-1}+\tau+(k-1)s]} \frac{dv}{v}$$

$$\le \frac{1}{\tau} \log \frac{ks + \tau}{ks}$$

and so this contribution to (123) is bounded by WX, where W, X are defined in (110) and (111).

Now we consider the contribution to (123) when[e]

$$1 + u\tau - S_{k-1} > 2c.$$

In this regime, we bound

$$\frac{1}{1 + u\tau - S_{k-1} + (k-1)s} \le \frac{1}{2c + (k-1)t},$$

and so this portion of $\int_0^1 Z_4[1 + u\tau] \, du$ may be bounded by

$$\int_0^1 \frac{\log k}{c} \mathbb{E} \left(1 + u\tau - S_{k-1} - c\right)^2 V \, du = VU$$

where V, U are defined in (112) and (113). The proof of the theorem is now complete. $\square$

We can now perform an asymptotic analysis in the limit $k \to \infty$ to establish Theorem 23(xi) and Theorem 25(vi). For k sufficiently large, we select the parameters

$$c := \frac{1}{\log k} + \frac{\alpha}{\log^2 k}$$

$$T := \frac{\beta}{\log k}$$

$$\tau := \frac{\gamma}{\log k}$$

for some real parameters $\alpha \in \mathbb{R}$ and $\beta, \gamma > 0$ independent of k to be optimized in later. From (100) and (101), we have

$$
\begin{aligned}
m_2 &= \frac{1}{k-1} \left(\frac{1}{c} - \frac{1}{c + (k-1)T} \right) \\
&= \frac{\log k}{k} \left(1 - \frac{\alpha}{\log k} + o\left(\frac{1}{\log k} \right) \right)
\end{aligned}
$$

where we use $o\left(f(k) \right)$ to denote a function $g(k)$ of k with $g(k)/f(k) \to 0$ as $k \to \infty$. On the other hand, we have from (100) and (102) that

$$
\begin{aligned}
m_2(c + (k-1)\mu) &= \int_0^T (c + (k-1)t)\, g(t)^2 \, dt \\
&= \frac{1}{k-1} \log \frac{c + (k-1)T}{c} \\
&= \frac{\log k}{k} \left(1 + \frac{\log \beta}{\log k} + o\left(\frac{1}{\log k} \right) \right)
\end{aligned}
$$

and thus

$$
\begin{aligned}
k\mu &= \frac{k}{k-1} \left(1 + \frac{\log \beta + \alpha}{\log k} + o\left(\frac{1}{\log k} \right) \right) - \frac{kc}{k-1} \\
&= 1 + \frac{\log \beta + \alpha}{\log k} + o\left(\frac{1}{\log k} \right) - \left(\frac{1}{\log k} + o\left(\frac{1}{\log k} \right) \right) \\
&= 1 + \frac{\log \beta + \alpha - 1}{\log k} + o\left(\frac{1}{\log k} \right).
\end{aligned}
$$

Similarly, from (100), (102), and (103), we have

$$
m_2 \left(c^2 + 2c(k-1)\mu + (k-1)^2 \left(\mu^2 + \sigma^2 \right) \right) = \int_0^T (c + (k-1)t)^2\, g(t)^2 \, dt
$$
$$
= T
$$

and thus

$$
\begin{aligned}
k\sigma^2 &= \frac{k}{(k-1)^2} \left(\frac{T}{m_2} - c^2 - 2c(k-1)\mu \right) - k\mu^2 \\
&= \frac{\beta}{\log^2 k} + o\left(\frac{1}{\log^2 k} \right).
\end{aligned}
$$

We conclude that the hypotheses (104), (105), and (106) will be obeyed for sufficiently large k if we have

$$
\log \beta + \alpha + \gamma < 1
$$
$$
\log \beta + \alpha + \beta < 1
$$
$$
\beta < \left(1 + \gamma - \alpha - \log \beta \right)^2.
$$

These conditions can be simultaneously obeyed, for instance by setting $\beta = \gamma = 1$ and $\alpha = -1$.

Now we crudely estimate the quantities Z, Z_3, W, X, V, U in (108)-(113). For $1 \le r \le 1 + \tau$, we have $r - k\mu \sim 1/\log k$, and so

$$
\frac{r - k\mu}{T} \asymp 1; \quad \frac{k\sigma^2}{(r - k\mu)^2} \asymp 1; \quad \frac{r^2}{4kT} = o(1)
$$

and so by (108) $Z = O(1)$. Using the crude bound $\log\left(1 + \frac{t}{T}\right) = O(1)$ for $0 \le t \le T$, we see from (109) and (102) that $Z_3 = O(k\mu) = O(1)$. It is clear that $X = O(1)$, and using the crude bound $\frac{1}{2c+(k-1)t} \le \frac{1}{c}$ we see from (112) and (101) that $V = O(1)$. For $0 \le u \le 1$ we have $1 + u\tau - (k-1)\mu - c = O(1/\log k)$, so from (113) we have $U = O(1)$. Finally, from (110) and the change of variables $t = \frac{s}{k\log k}$, we have

$$W = \frac{\log k}{km_2} \int_0^{kT\log k} \log\left(1 + \frac{\gamma}{s}\right) \frac{ds}{\left(1 + \frac{\alpha}{\log k} + \frac{k-1}{k}s\right)^2}$$

$$= O\left(\int_0^\infty \log\left(1 + \frac{\gamma}{s}\right) \frac{ds}{(1 + o(1))(1 + s)^2}\right)$$

$$= O(1).$$

Finally, we have

$$1 - \frac{k\sigma^2}{(1 + \tau - k\mu)^2} \sim 1.$$

Putting all these together, we see from (107) that

$$M_k \ge M_k^{[T]} \ge \frac{k}{k-1}\log k - O(1)$$

giving Theorem 23(xi). Furthermore, if we set

$$\varpi := \frac{7}{600} - \frac{C}{\log k}$$

and

$$\delta := \left(\frac{1}{4} + \frac{7}{600}\right)\frac{\beta}{\log k},$$

then we will have $600\varpi + 180\delta < 7$ for C large enough, and Theorem 25(vi) also follows (as one can verify from inspection that all implied constants here are effective).

Finally, Theorem 23(viii), (ix), and (x) follow by setting

$$c := \frac{\theta}{\log k}$$

$$T := \frac{\beta}{\log k}$$

$$\tau = 1 - k\mu$$

with θ, β given by Table 2, with (107) then giving the bound $M_k^{[T]} > M$ with M as given by the table, after verifying of course that the conditions (104), (105), and (106) are obeyed. Similarly, Theorem 25 (ii), (iii), (iv), and (v) follows with θ, β given by the same table, with ϖ chosen so that

$$M = \frac{m}{\frac{1}{4} + \varpi}$$

with $m = 2, 3, 4, 5$ for (ii), (iii), (iv), (v), respectively, and δ chosen by the formula

$$\delta := T\left(\frac{1}{4} + \varpi\right).$$

Table 2 Parameter choices for Theorems 23 and 25

k	θ	β	M
5,511	0.965	0.973	6.000048609
35,410	0.99479	0.85213	7.829849259
41,588	0.97878	0.94319	8.000001401
309,661	0.98627	0.92091	10.00000032
1,649,821	1.00422	0.80148	11.65752556
75,845,707	1.00712	0.77003	15.48125090
3,473,955,908	1.0079318	0.7490925	19.30374872

The case of small and medium dimension

In this section, we establish lower bounds for M_k (and related quantities, such as $M_{k,\varepsilon}$) both for small values of k (in particular, $k = 3$ and $k = 4$) and medium values of k (in particular, $k = 50$ and $k = 54$). Specifically, we will establish Theorem 23(vii), Theorem 27, and Theorem 29.

Bounding M_k for medium k

We begin with the problem of lower bounding M_k. We first formalize an observation[f] of Maynard [5] that one may restrict without loss of generality to symmetric functions:

Lemma 41. *For any $k \geq 2$, one has*

$$M_k := \sup \frac{kJ_1(F)}{I(F)}$$

where F ranges over symmetric square-integrable functions on $\mathcal{R}_k$ that are not identically zero.

Proof. Firstly, observe that if one replaces a square-integrable function $F : [0, +\infty)^k \to \mathbb{R}$ with its absolute value $|F|$, then $I(|F|) = I(F)$ and $J_i(|F|) \geq J_i(F)$. Thus, one may restrict the supremum in (33) to non-negative functions without loss of generality. We may thus find a sequence F_n of square-integrable non-negative functions on $\mathcal{R}_k$, normalized so that $I(F_n) = 1$, and such that $\sum_{i=1}^k J_i(F_n) \to M_k$ as $n \to \infty$.

Now let

$$\overline{F_n}(t_1, \ldots, t_k) := \frac{1}{k!} \sum_{\sigma \in S_k} F_n(t_{\sigma(1)}, \ldots, t_{\sigma(k)})$$

be the symmetrisation of F_n. Since the F_n are non-negative with $I(F_n) = 1$, we see that

$$I(\overline{F_n}) \geq I\left(\frac{1}{k!} F_n\right) = \frac{1}{(k!)^2}$$

and so $I(\overline{F_n})$ is bounded away from zero. Also, from (33), we know that the quadratic form

$$Q(F) := M_k I(F) - \sum_{i=1}^k J_i(F)$$

is positive semi-definite and is also invariant with respect to symmetries, and so from the triangle inequality for inner product spaces, we conclude that

$$Q\left(\overline{F_n}\right) \le Q\left(F_n\right).$$

By construction, $Q(F_n)$ goes to zero as $n \to \infty$, and thus $Q(\overline{F_n})$ also goes to zero. We conclude that

$$\frac{kJ_1\left(\overline{F_n}\right)}{I\left(\overline{F_n}\right)} = \frac{\sum_{i=1}^{k} J_i\left(\overline{F_n}\right)}{I\left(\overline{F_n}\right)} \to M_k$$

as $n \to \infty$, and so

$$M_k \ge \sup \frac{kJ_1(F)}{I(F)}.$$

The reverse inequality is immediate from (33), and the claim follows. $\qquad\square$

To establish a lower bound of the form $M_k > C$ for some $C > 0$, one thus seeks to locate a symmetric function $F : [0, +\infty)^k \to \mathbb{R}$ supported on $\mathcal{R}_k$ such that

$$kJ_1(F) > CI(F). \tag{124}$$

To do this numerically, we follow [5] (see also [2] for some related ideas) and can restrict attention to functions F that are linear combinations

$$F = \sum_{i=1}^{n} a_i b_i$$

of some explicit finite set $b_1, \ldots, b_n : [0, +\infty)^k \to \mathbb{R}$ supported on $\mathcal{R}_k$ and some real scalars $a_1, \ldots, a_n$ that we may optimize in. The condition (124) then may be rewritten as

$$\mathbf{a}^T \mathbf{M}_2 \mathbf{a} - C \mathbf{a}^T \mathbf{M}_1 \mathbf{a} > 0 \tag{125}$$

where $\mathbf{a}$ is the vector

$$\mathbf{a} := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

and $\mathbf{M}_1, \mathbf{M}_2$ are the real symmetric and positive semi-definite $n \times n$ matrices

$$\mathbf{M}_1 = \left(\int_{\mathbb{R}^k} b_i(t_1, \ldots, t_k) b_j(t_1, \ldots, t_k) \, dt_1 \ldots dt_k \right)_{1 \le i,j \le n} \tag{126}$$

$$\mathbf{M}_2 = \left(k \int_{\mathbb{R}^{k+1}} b_i(t_1, \ldots, t_k) b_j(t_1, \ldots, t_{k-1}, t_k') \, dt_1 \ldots dt_k dt_k' \right)_{1 \le i,j \le n}. \tag{127}$$

If the $b_1, \ldots, b_n$ are linearly independent in $L^2(\mathcal{R}_k)$, then $\mathbf{M}_1$ is strictly positive definite, and (as observed in [5, Lemma 8.3]), one can find $\mathbf{a}$ obeying (125) if and only if the largest eigenvalue of $\mathbf{M}_2 \mathbf{M}_1^{-1}$ exceeds C. This is a criterion that can be numerically verified for medium-sized values of n, if the $b_1, \ldots, b_n$ are chosen so that the matrix coefficients of $\mathbf{M}_1, \mathbf{M}_2$ are explicitly computable.

In order to facilitate computations, it is natural to work with bases $b_1, \ldots, b_n$ of symmetric polynomials. We have the following basic integration identity:

Lemma 42 (Beta function identity). *For any non-negative $a, a_1, \ldots, a_k$, we have*

$$\int_{\mathcal{R}_k} (1 - t_1 - \cdots - t_k)^a t_1^{a_1} \ldots t_k^{a_k} \, dt_1 \ldots dt_k = \frac{\Gamma(a+1)\Gamma(a_1+1)\ldots\Gamma(a_k+1)}{\Gamma(a_1 + \cdots + a_k + k + a + 1)}$$

where $\Gamma(s) := \int_0^\infty t^{s-1} e^{-t} \, dt$ is the Gamma function. In particular, if $a_1, \ldots, a_k$ are natural numbers, then

$$\int_{\mathcal{R}_k} (1 - t_1 - \cdots - t_k)^a t_1^{a_1} \ldots t_k^{a_k} \, dt_1 \ldots dt_k = \frac{a! \, a_1! \ldots a_k!}{(a_1 + \cdots + a_k + k + a)!}.$$

Proof. Since

$$\int_{\mathcal{R}_k} (1 - t_1 - \cdots - t_k)^a t_1^{a_1} \ldots t_k^{a_k} \, dt_1 \ldots dt_k = a \int_{\mathcal{R}_{k+1}} t_1^{a_1} \ldots t_k^{a_k} t_{k+1}^{a-1} \, dt_1 \ldots dt_{k+1},$$

we see that to establish the lemma, it suffices to do so in the case $a = 0$.

If we write

$$X := \int_{t_1 + \cdots + t_k = 1} t_1^{a_1} \ldots t_k^{a_k} \, dt_1 \ldots dt_{k-1},$$

then by homogeneity we have

$$r^{a_1 + \cdots + a_k + k - 1} X = \int_{t_1 + \cdots + t_k = r} t_1^{a_1} \ldots t_k^{a_k} \, dt_1 \ldots dt_{k-1}$$

for any $r > 0$, and hence on integrating r from 0 to 1, we conclude that

$$\frac{X}{a_1 + \cdots + a_k + k} = \int_{\mathcal{R}_k} t_1^{a_1} \ldots t_k^{a_k} \, dt_1 \ldots dt_k.$$

On the other hand, if we multiply by e^{-r} and integrate r from 0 to ∞, we obtain instead

$$\int_0^\infty r^{a_1 + \cdots + a_k + k - 1} X e^{-r} \, dr = \int_{[0,+\infty)^k} t_1^{a_1} \ldots t_k^{a_k} e^{-t_1 - \cdots - t_k} \, dt_1 \ldots dt_k.$$

Using the definition of the Gamma function, this becomes

$$\Gamma(a_1 + \cdots + a_k + k) X = \Gamma(a_1 + 1) \ldots \Gamma(a_k + 1)$$

and the claim follows. □

Define a *signature* to be a non-increasing sequence $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ of natural numbers; for brevity, we omit zeroes; thus, for instance if $k = 6$, then $(2, 2, 1, 1, 0, 0)$ will be

abbreviated as $(2, 2, 1, 1)$. The number of non-zero elements of α will be called the *length* of the signature α, and as usual the *degree* of α with be $\alpha_1 + \cdots + \alpha_k$. For each signature α, we then define the symmetric polynomials $P_\alpha = P_\alpha^{(k)}$ by the formula

$$P_\alpha(t_1, \ldots, t_k) = \sum_{a:s(a)=\alpha} t_1^{a_1} \ldots t_k^{a_k}$$

where the summation is over all tuples $a = (a_1, \ldots, a_k)$ whose non-increasing rearrangement $s(a)$ is equal to α. Thus, for instance

$$P_{(1)}(t_1, \ldots, t_k) = t_1 + \cdots + t_k$$
$$P_{(2)}(t_1, \ldots, t_k) = t_1^2 + \cdots + t_k^2$$
$$P_{(1,1)}(t_1, \ldots, t_k) = \sum_{1 \le i < j \le k} t_i t_j$$
$$P_{(2,1)}(t_1, \ldots, t_k) = \sum_{1 \le i < j \le k} t_i^2 t_j + t_i t_j^2$$

and so forth. Clearly, the P_α form a linear basis for the symmetric polynomials of $t_1, \ldots, t_k$. Observe that if $\alpha = (\alpha', 1)$ is a signature containing 1, then one can express P_α as $P_{(1)} P_{\alpha'}$ minus a linear combination of polynomials P_β with the length of β less than that of α. This implies that the functions $P_{(1)}^a P_\alpha$, with $a \ge 0$ and α avoiding 1, are also a basis for the symmetric polynomials. Equivalently, the functions $(1 - P_{(1)})^a P_\alpha$ with $a \ge 0$ and α avoiding 1 form a basis.

After extensive experimentation, we have discovered that a good basis $b_1, \ldots, b_n$ to use for the above problem comes by setting the b_i to be all the symmetric polynomials of the form $(1 - P_{(1)})^a P_\alpha$, where $a \ge 0$ and α consists entirely of even numbers, whose total degree $a + \alpha_1 + \cdots + \alpha_k$ is less than or equal to some chosen threshold d. For such functions, the coefficients of $\mathbf{M}_1, \mathbf{M}_2$ can be computed exactly using Lemma 42.

More explicitly, first we quickly compute a look-up table for the structure constants $c_{\alpha,\beta,\gamma} \in \mathbb{Z}$ derived from simple products of the form

$$P_\alpha P_\beta = \sum_\gamma c_{\alpha,\beta,\gamma} P_\gamma$$

where $\deg(\alpha) + \deg(\beta) \le d$. Using this look-up table, we rewrite the integrands of the entries of the matrices in (126) and (127) as integer linear combinations of nearly 'pure' monomials of the form $(1 - P_{(1)})^a t_1^{a_1} \ldots t_k^{a_k}$. We then calculate the entries of $\mathbf{M}_1$ and $\mathbf{M}_2$, as exact rational numbers, using Lemma 42.

We next run a generalized eigenvector routine on (real approximations to) $\mathbf{M}_1$ and $\mathbf{M}_2$ to find a vector $\mathbf{a}'$ which nearly maximize the quantity C in (125). Taking a rational approximation $\mathbf{a}$ to $\mathbf{a}'$, we then do the quick (and exact) arithmetic to verify that (125) holds for some constant $C > 4$. This generalized eigenvector routine is time-intensive when the sizes of $\mathbf{M}_1$ and $\mathbf{M}_2$ are large (say, bigger than $1,500 \times 1,500$) and in practice is the most computationally intensive step of our calculation. When one does not care about an exact arithmetic proof that $C > 4$, instead one can run a test for positive-definiteness for the matrix $C\mathbf{M}_1 - \mathbf{M}_2$, which is usually much faster and less RAM intensive.

Using this method, we were able to demonstrate $M_{54} > 4.00238$, thus establishing Theorem 23(vii). We took $d = 23$ and imposing the restriction on signatures α that they be composed only of even numbers. It is likely that $d = 22$ would suffice in the absence of

this restriction on signatures, but we found that the gain in M_{54} from lifting this restriction is typically only in the region of 0.005, whereas the execution time is increased by a large factor. We do not have a good understanding of why this particular restriction on signatures is so inexpensive in terms of the trade-off between the accuracy of M-values and computational complexity. The total run-time for this computation was under 1 h.

We now describe a second choice for the basis elements $b_1, \ldots, b_n$, which uses the Krylov subspace method; it gives faster and more efficient numerical results than the previous basis, but does not seem to extend as well to more complicated variational problems such as $M_{k,\varepsilon}$. We introduce the linear operator $\mathcal{L} : L^2(\mathcal{R}_k) \to L^2(\mathcal{R}_k)$ defined by

$$\mathcal{L}f(t_1, \ldots, t_k) := \sum_{i=1}^{k} \int_0^{1 - t_1 - \cdots - t_{i-1} - t_{i+1} - \cdots - t_k} f(t_1, \ldots, t_{i-1}, t_i', t_{i+1}, \ldots, t_k) \, dt_i'.$$

This is a self-adjoint and positive semi-definite operator on $L^2(\mathcal{R}_k)$. For symmetric $b_1, \ldots, b_n \in L^2(\mathcal{R}_k)$, one can then write

$$\mathbf{M}_1 = \left(\langle b_i, b_j \rangle \right)_{1 \le i,j \le n}$$
$$\mathbf{M}_2 = \left(\langle \mathcal{L}b_i, b_j \rangle \right)_{1 \le i,j \le n}.$$

If we then choose

$$b_i := \mathcal{L}^{i-1} 1$$

where 1 is the unit constant function on $\mathcal{R}_k$, then the matrices $\mathbf{M}_1, \mathbf{M}_2$ take the Hankel form

$$\mathbf{M}_1 = \left(\langle \mathcal{L}^{i+j-2} 1, 1 \rangle \right)_{1 \le i,j \le n}$$
$$\mathbf{M}_2 = \left(\langle \mathcal{L}^{i+j-1} 1, 1 \rangle \right)_{1 \le i,j \le n},$$

and so can be computed entirely in terms of the $2n$ numbers $\langle \mathcal{L}^i 1, 1 \rangle$ for $i = 0, \ldots, 2n - 1$.

The operator $\mathcal{L}$ maps symmetric polynomials to symmetric polynomials; for instance, one has

$$\mathcal{L}1 = k - (k-1)P_{(1)}$$
$$\mathcal{L}P_{(1)} = \frac{k}{2} - \frac{k-1}{2}P_{(2)} - (k-2)P_{(1,1)}$$

and so forth. From this and Lemma 42, the quantities $\langle \mathcal{L}^i 1, 1 \rangle$ are explicitly computable rational numbers; for instance, one can calculate

$$\langle 1, 1 \rangle = \frac{1}{k!}$$
$$\langle \mathcal{L}1, 1 \rangle = \frac{2k}{(k+1)!}$$
$$\langle \mathcal{L}^2 1, 1 \rangle = \frac{k(5k+1)}{(k+2)!}$$
$$\langle \mathcal{L}^3 1, 1 \rangle = \frac{2k^2(7k+5)}{(k+3)!}$$

and so forth.

With Maple, we were able to compute $\langle \mathcal{L}^i 1, 1 \rangle$ for $i \le 50$ and $k \le 100$, leading to lower bounds on M_k for these values of k, a selection of which is given in Table 3.

Table 3 Selected lower bounds on M_k obtained from the Krylov subspace method, with $\frac{k}{k-1} \log k$ upper bound displayed for comparison

k	Lower bound on M_k	$\frac{k}{k-1} \log k$
2	1.38593	1.38629
3	1.64644	1.64791
4	1.84540	1.84839
5	2.00714	2.01179
10	2.54547	2.55842
20	3.12756	3.15340
30	3.48313	3.51848
40	3.73919	3.78346
50	3.93586	3.99186
53	3.98621	4.04664
54	4.00223	4.06424
60	4.09101	4.16374
100	4.46424	4.65168

Bounding $M_{k,\varepsilon}$ for medium k

When bounding $M_{k,\varepsilon}$, we have not been able to implement the Krylov method because the analogue of $\mathcal{L}^i 1$ in this context is piecewise polynomial instead of polynomial, and we were only able to compute it explicitly for very small values of i, such as $i = 1, 2, 3$, which are insufficient for good numerics. Thus, we rely on the previously discussed approach, in which symmetric polynomials are used for the basis functions. Instead of computing integrals over the region $\mathcal{R}_k$, we pass to the regions $(1 \pm \varepsilon)\mathcal{R}_k$. In order to apply Lemma 42 over these regions, this necessitates working with a slightly different basis of polynomials. We chose to work with those polynomials of the form $(1 + \varepsilon - P_{(1)})^a P_\alpha$, where α is a signature with no 1's. Over the region $(1 + \varepsilon)\mathcal{R}_k$, a single change of variables converts the needed integrals into those of the form in Lemma 42, and we can then compute the entries of $\mathbf{M}_1$.

On the other hand, over the region $(1 - \varepsilon)\mathcal{R}_k$, we instead want to work with polynomials of the form $(1 - \varepsilon - P_{(1)})^a P_\alpha$. Since $(1 + \varepsilon - P_{(1)})^a = (2\varepsilon + (1 - \varepsilon - P_{(1)}))^a$, an expansion using the binomial theorem allows us to convert from our given basis to polynomials of the needed form.

With these modifications, and calculating as in the previous section, we find that $M_{50,1/25} > 4.00124$ if $d = 25$ and $M_{50,1/25} > 4.0043$ if $d = 27$, thus establishing Theorem 27(i). As before, we found it optimal to restrict signatures to contain only even entries, which greatly reduced execution time while only reducing M by a few thousandths.

One surprising additional computational difficulty introduced by allowing $\varepsilon > 0$ is that the 'complexity' of ε as a rational number affects the run-time of the calculations. We found that choosing $\varepsilon = 1/m$ (where $m \in \mathbb{Z}$ has only small prime factors) reduces this effect.

A similar argument gives $M_{51,1/50} > 4.00156$, thus establishing Theorem 27(xiii). In this case, our polynomials were of maximum degree $d = 22$.

Code and data for these calculations may be found at http://www.dropbox.com/sh/0xb4xrsx4qmua7u/WOhuo2Gx7f/Polymath8b.

Bounding $M_{4,\varepsilon}$

We now prove Theorem 27(xii'), which can be established by a direct numerical calculation. We introduce the explicit function $F : [0, +\infty)^4 \to \mathbb{R}$ defined by

$$F(t_1, t_2, t_3, t_4) := (1 - \alpha(t_1 + t_2 + t_3 + t_4))1_{t_1+t_2+t_3+t_4 \leq 1+\varepsilon}$$

with $\varepsilon := 0.168$ and $\alpha := 0.784$. As F is symmetric in t_1, t_2, t_3, t_4, we have $J_{i,1-\varepsilon}(F) = J_{1,1-\varepsilon}(F)$, so to show Theorem 27(xii') it will suffice to show that

$$\frac{4J_{1,1-\varepsilon}(F)}{I(F)} > 2.00558. \tag{128}$$

By making the change of variables $s = t_1 + t_2 + t_3 + t_4$, we see that

$$
\begin{aligned}
I(F) &= \int_{t_1+t_2+t_3+t_4 \leq 1+\varepsilon} (1 - \alpha(t_1 + t_2 + t_3 + t_4))^2 \ dt_1 dt_2 dt_3 dt_4 \\
&= \int_0^{1+\varepsilon} (1 - \alpha s)^2 \frac{s^3}{3!} \ ds \\
&= \alpha^2 \frac{(1+\varepsilon)^6}{36} - \alpha \frac{(1+\varepsilon)^5}{15} + \frac{(1+\varepsilon)^4}{24} \\
&= 0.00728001347\ldots
\end{aligned}
$$

and similarly by making the change of variables $u = t_1 + t_2 + t_3$

$$
\begin{aligned}
J_{1,1-\varepsilon}(F) &= \int_{t_1+t_2+t_3 \leq 1-\varepsilon} \left(\int_0^{1+\varepsilon-t_1-t_2-t_3} (1 - \alpha(t_1 + t_2 + t_3 + t_4)) \ dt_4 \right)^2 dt_1 dt_2 dt_3 \\
&= \int_0^{1-\varepsilon} \left(\int_0^{1+\varepsilon-u} (1 - \alpha(u + t_4)) \ dt_4 \right)^2 \frac{u^2}{2!} du \\
&= \int_0^{1-\varepsilon} (1 + \varepsilon - u)^2 \left(1 - \alpha \frac{1+\varepsilon+u}{2} \right)^2 \frac{u^2}{2} du \\
&= 0.003650160667\ldots
\end{aligned}
$$

and so (128) follows.

Remark 43. If we use the truncated function

$$\tilde{F}(t_1, t_2, t_3, t_4) := F(t_1, t_2, t_3, t_4)1_{t_1,t_2,t_3,t_4 \leq 1}$$

in place of F and set ε to 0.18 instead of 0.168, one can compute that

$$\frac{4J_{1,1-\varepsilon}(\tilde{F})}{I(\tilde{F})} > 2.00235.$$

Thus, it is possible to establish Theorem 27(xii') using a cutoff function F' that is also supported in the unit cube $[0, 1]^4$. This allows for a slight simplification to the proof of DHL[4; 2] assuming GEH, as one can add the additional hypothesis $S(F_{i_0}) + S(G_{i_0}) < 1$ to Theorem 20(ii) in that case.

Remark 44. By optimizing in ε and taking F to be a symmetric polynomial of degree higher than 1, one can get slightly better lower bounds for $M_{4,\varepsilon}$; for instance, setting $\varepsilon = 5/21$ and choosing F to be a cubic polynomial, we were able to obtain the bound $M_{4,\varepsilon} \geq 2.05411$. On the other hand, the best lower bound for $M_{3,\varepsilon}$ that we were able to obtain was 1.91726 (taking $\varepsilon = 56/113$ and optimizing over cubic polynomials).

Again, see www.dropbox.com/sh/0xb4xrsx4qmua7u/WOhuo2Gx7f/Polymath8b for the relevant code and data.

Three-dimensional cutoffs

In this section, we establish Theorem 29. We relabel the variables (t_1, t_2, t_3) as (x, y, z); thus, our task is to locate a piecewise polynomial function $F : [0, +\infty)^3 \to \mathbb{R}$ supported on the simplex

$$R := \left\{ (x, y, z) \in [0, +\infty)^3 : x + y + z \leq \frac{3}{2} \right\}$$

and symmetric in the x, y, z variables, obeying the vanishing marginal condition

$$\int_0^\infty F(x, y, z) \, dz = 0 \tag{129}$$

whenever $x, y \geq 0$ with $x + y > 1 + \varepsilon$, and such that

$$J(F) > 2I(F) \tag{130}$$

where

$$J(F) := 3 \int_{x+y \leq 1-\varepsilon} \left(\int_0^\infty F(x, y, z) \, dz \right)^2 \, dxdy \tag{131}$$

and

$$I(F) := \int_R F(x, y, z)^2 \, dxdydz \tag{132}$$

and

$$\varepsilon := 1/4.$$

Our strategy will be as follows. We will decompose the simplex R (up to null sets) into a carefully selected set of disjoint open polyhedra $P_1, \ldots, P_m$ (in fact m will be 60), and on each P_i we will take $F(x, y, z)$ to be a low-degree polynomial $F_i(x, y, z)$ (indeed, the degree will never exceed 3). The left-hand and right-hand sides of (130) become quadratic functions in the coefficients of the F_i. Meanwhile, the requirement of symmetry, as well as the marginal requirement (129), imposes some linear constraints on these coefficients. In principle, this creates a finite-dimensional quadratic program, which one can try to solve numerically. However, to make this strategy practical, one needs to keep the number of linear constraints imposed on the coefficients to be fairly small, as compared with the total number of coefficients. To achieve this, the following properties on the polynomials P_i are desirable:

- (Symmetry) If P_i is a polytope in the partition, then every reflection of P_i formed by permuting the x, y, z coordinates should also lie in the partition.
- (Graph structure) Each polytope P_i should be of the form

$$\left\{ (x, y, z) : z \in Q_i; a_i(x, y) < z < b_i(x, y) \right\}, \tag{133}$$

  where $a_i(x, y), b_i(x, y)$ are linear forms and Q_i is a polygon.
- (Epsilon splitting) Each Q_i is contained in one of the regions $\left\{ (x, y) : x + y < 1 - \varepsilon \right\}$, $\left\{ (x, y) : 1 - \varepsilon < x + y < 1 + \varepsilon \right\}$, or $\left\{ (x, y) : 1 + \varepsilon < x + y < 3/2 \right\}$.

Observe that the vanishing marginal condition (129) now takes the form

$$\sum_{i:(x,y)\in Q_i} \int_{a_i(x,y)}^{b_i(x,y)} F_i(x,y,z)\, dz = 0 \tag{134}$$

for every $x, y > 0$ with $x + y > 1 + \varepsilon$. If the set $\{i : (x, y) \in Q_i\}$ is fixed, then the left-hand side of (134) is a polynomial in x, y whose coefficients depend linearly on the coefficients on the F_i, and thus (134) imposes a set of linear conditions on these coefficients for each possible set $\{i : (x, y) \in Q_i\}$ with $x + y > 1 + \varepsilon$.

Now we describe the partition we will use. This partition can in fact be used for all ε in the interval $[1/4, 1/3]$, but the endpoint $\varepsilon = 1/4$ has some simplifications which allowed for reasonably good numerical results. To obtain the symmetry property, it is natural to split R (modulo null sets) into six polyhedra $R_{xyz}, R_{xzy}, R_{yxz}, R_{yzx}, R_{zxy}, R_{zyx}$, where

$$\begin{aligned} R_{xyz} &:= \left\{ (x, y, z) \in R \ : \ x + y < y + z < z + x \right\} \\ &= \left\{ (x, y, z) : 0 < y < x < z; x + y + z \leq 3/2 \right\} \end{aligned}$$

and the other polyhedra are obtained by permuting the indices x, y, z, thus for instance

$$\begin{aligned} R_{yxz} &:= \left\{ (x, y, z) \in R \ : \ y + x < x + z < z + y \right\} \\ &= \left\{ (x, y, z) : 0 < x < y < z; x + y + z \leq 3/2 \right\}. \end{aligned}$$

To obtain the epsilon splitting property, we decompose R_{xyz} (modulo null sets) into eight sub-polytopes

$$\begin{aligned} A_{xyz} &= \left\{ (x, y, z) \in R \ : \ x + y < y + z < z + x < 1 - \varepsilon \right\}, \\ B_{xyz} &= \left\{ (x, y, z) \in R \ : \ x + y < y + z < 1 - \varepsilon < z + x < 1 + \varepsilon \right\}, \\ C_{xyz} &= \left\{ (x, y, z) \in R \ : \ x + y < 1 - \varepsilon < y + z < z + x < 1 + \varepsilon \right\}, \\ D_{xyz} &= \left\{ (x, y, z) \in R \ : \ 1 - \varepsilon < x + y < y + z < z + x < 1 + \varepsilon \right\}, \\ E_{xyz} &= \left\{ (x, y, z) \in R \ : \ x + y < y + z < 1 - \varepsilon < 1 + \varepsilon < z + x \right\}, \\ F_{xyz} &= \left\{ (x, y, z) \in R \ : \ x + y < 1 - \varepsilon < y + z < 1 + \varepsilon < z + x \right\}, \\ G_{xyz} &= \left\{ (x, y, z) \in R \ : \ x + y < 1 - \varepsilon < 1 + \varepsilon < y + z < z + x \right\}, \\ H_{xyz} &= \left\{ (x, y, z) \in R \ : \ 1 - \varepsilon < x + y < y + z < 1 + \varepsilon < z + x \right\}; \end{aligned}$$

the other five polytopes $R_{xzy}, R_{yxz}, R_{yzx}, R_{zxy}, R_{zyx}$ are decomposed similarly, leading to a partition of R into $6 \times 8 = 48$ polytopes. This is almost the partition we will use; however, there is a technical difficulty arising from the fact that some of the permutations of F_{xyz} do not obey the graph structure property. So we will split F_{xyz} further into the three pieces

$$\begin{aligned} S_{xyz} &= \left\{ (x, y, z) \in F_{xyz} \ : \ z < 1/2 + \varepsilon \right\}, \\ T_{xyz} &= \left\{ (x, y, z) \in F_{xyz} \ : \ z > 1/2 + \varepsilon; x > 1/2 - \varepsilon \right\}, \\ U_{xyz} &= \left\{ (x, y, z) \in F_{xyz} \ : \ x < 1/2 - \varepsilon \right\}. \end{aligned}$$

Thus, R_{xyz} is now partitioned into ten polytopes $A_{xyz}, B_{xyz}, C_{xyz}, D_{xyz}, E_{xyz}, S_{xyz}, T_{xyz}, U_{xyz}, G_{xyz}, H_{xyz}$, and similarly for permutations of R_{xyz}, leading to a decomposition of R into $6 \times 10 = 60$ polytopes.

A symmetric piecewise polynomial function F supported on R can now be described (almost everywhere) by specifying a polynomial function $F \!\restriction_P : P \to \mathbb{R}$ for the ten

polytopes $P = A_{xyz}, B_{xyz}, C_{xyz}, D_{xyz}, E_{xyz}, S_{xyz}, T_{xyz}, U_{xyz}, G_{xyz}, H_{xyz}$, and then extending by symmetry, thus for instance

$$F \mid_{A_{yzx}} (x, y, z) = F \mid_{A_{xyz}} (z, x, y).$$

As discussed earlier, the expressions $I(F), J(F)$ can now be written as quadratic forms in the coefficients of the $F \mid_P$, and the vanishing marginal condition (129) imposes some linear constraints on these coefficients.

Observe that the polytope D_{xyz} and all of its permutations make no contribution to either the functional $J(F)$ or to the marginal condition (129), and give a non-negative contribution to $I(F)$. Thus, without loss of generality we may assume that

$$F \mid_{D_{xyz}} = 0.$$

However, the other nine polytopes $A_{xyz}, B_{xyz}, C_{xyz}, E_{xyz}, S_{xyz}, T_{xyz}, U_{xyz}, G_{xyz}, H_{xyz}$ have at least one permutation which gives a non-trivial contribution to either $J(F)$ or to (129), and cannot be easily eliminated.

Now we compute $I(F)$. By symmetry, we have

$$I(F) = 3! \, I(F \mid_{R_{xyz}}) = 6 \sum_P I(F \mid_P)$$

where P ranges over the nine polytopes $A_{xyz}, B_{xyz}, C_{xyz}, E_{xyz}, S_{xyz}, T_{xyz}, U_{xyz}, G_{xyz}, H_{xyz}$. A tedious but straightforward computation shows that

$$I(F \mid_{A_{xyz}}) = \int_{x=0}^{1/2-\varepsilon/2} \int_{y=0}^{x} \int_{z=x}^{1-\varepsilon-x} F \mid_{A_{xyz}}^2 \, dz \, dy \, dx$$

$$I(F \mid_{B_{xyz}}) = \left(\int_{z=1/2-\varepsilon/2}^{1/2+\varepsilon/2} \int_{x=1-\varepsilon-z}^{z} + \int_{z=1/2+\varepsilon/2}^{1-\varepsilon} \int_{x=1-\varepsilon-z}^{1+\varepsilon-z} \right) \int_{y=0}^{1-\varepsilon-z} F \mid_{B_{xyz}}^2 \, dy \, dx \, dz$$

$$I(F \mid_{C_{xyz}}) = \left(\int_{y=0}^{1/2-3\varepsilon/2} \int_{x=y}^{y+2\varepsilon} + \int_{y=1/2-3\varepsilon/2}^{1/2-\varepsilon} \int_{x=y}^{1-\varepsilon-y} \right) \int_{z=1-\varepsilon-y}^{1+\varepsilon-x}$$
$$+ \int_{y=1/2-\varepsilon}^{1/2-\varepsilon/2} \int_{x=y}^{1-\varepsilon-y} \int_{z=1-\varepsilon-y}^{3/2-x-y} F \mid_{C_{xyz}}^2 \, dz \, dx \, dy$$

$$I(F \mid_{E_{xyz}}) = \int_{z=1/2+\varepsilon/2}^{1-\varepsilon} \int_{x=1+\varepsilon-z}^{z} \int_{y=0}^{1-\varepsilon-z} F \mid_{E_{xyz}}^2 \, dy \, dx \, dz$$

$$I(F \mid_{S_{xyz}}) = \left(\int_{y=0}^{1/2-3\varepsilon/2} \int_{z=1-\varepsilon-y}^{1/2+\varepsilon} + \int_{y=1/2-3\varepsilon/2}^{1/2-\varepsilon} \int_{z=y+2\varepsilon}^{1/2+\varepsilon} \right) \int_{x=1+\varepsilon-z}^{1-\varepsilon-y} F \mid_{S_{xyz}} \, dx \, dz \, dy$$

$$I(F \mid_{T_{xyz}}) = \left(\int_{z=1/2+\varepsilon}^{1/2+2\varepsilon} \int_{x=1+\varepsilon-z}^{3/2-z} + \int_{z=1/2+2\varepsilon}^{1+\varepsilon} \int_{x=1/2-\varepsilon}^{3/2-z} \right) \int_{y=0}^{3/2-x-z} F \mid_{T_{xyz}}^2 \, dy \, dz \, dx$$

$$I(F \mid_{U_{xyz}}) = \int_{x=0}^{1/2-\varepsilon} \int_{y=0}^{x} \int_{z=1+\varepsilon-x}^{1+\varepsilon-y} F \mid_{U_{xyz}} \, dz \, dy \, dx$$

$$I(F \mid_{G_{xyz}}) = \int_{x=0}^{1/2-\varepsilon} \int_{y=0}^{x} \int_{z=1+\varepsilon-y}^{3/2-x-y} F \mid_{G_{xyz}}^2 \, dx \, dz \, dy$$

and

$$I(F \mid_{H_{xyz}}) = \left(\int_{x=1/2+\varepsilon/2}^{1-\varepsilon} \int_{y=1-\varepsilon-x}^{3/2-2x} + \int_{x=1-\varepsilon}^{3/4} \int_{y=0}^{3/2-2x} \right) \int_{z=x}^{3/2-x-y}$$
$$+ \int_{x=1/2}^{1/2+\varepsilon/2} \int_{y=1-\varepsilon-x}^{1/2-\varepsilon} \int_{z=1+\varepsilon-x}^{3/2-x-y} F \mid_{H_{xyz}}^2 \, dz \, dy \, dx.$$

Now we consider the quantity $J(F)$. Here we only have the symmetry of swapping x and y, so that

$$J(F) = 6 \int_{0<y<x; x+y<1-\varepsilon} \left(\int_0^{3/2-x-y} F(x,y,z)\, dz \right)^2 dx dy.$$

The region of integration meets the polytopes $A_{xyz}, A_{yzx}, A_{zyx}, B_{xyz}, B_{zyx}, C_{xyz}, E_{xyz}, E_{zyx}, S_{xyz}, T_{xyz}, U_{xyz}$, and G_{xyz}.

Projecting these regions to the (x,y)-plane, we have the diagram:
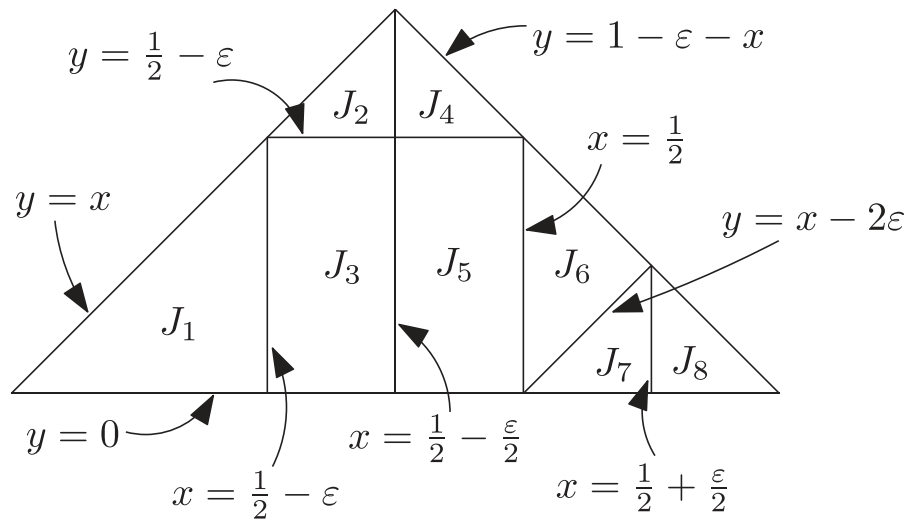
This diagram is drawn to scale in the case when $\varepsilon = 1/4$; otherwise, there is a separation between the J_5 and J_7 regions. For each of these eight regions, there are eight corresponding integrals $J_1, J_2, \ldots, J_8$, and thus

$$J = 2\,(J_1 + \cdots + J_8).$$

We have

$$J_1 = \int_{x=0}^{1/2-\varepsilon} \int_{y=0}^{x} \left(\int_{z=0}^{y} F \mid_{A_{yzx}} + \int_{z=y}^{x} F \mid_{A_{zyx}} + \int_{z=x}^{1-\varepsilon-x} F \mid_{A_{xyz}} + \int_{z=1-\varepsilon-x}^{1-\varepsilon-y} F \mid_{B_{xyz}} \right.$$
$$\left. + \int_{z=1-\varepsilon-y}^{1+\varepsilon-x} F \mid_{C_{xyz}} + \int_{z=1+\varepsilon-x}^{1+\varepsilon-y} F \mid_{U_{xyz}} + \int_{z=1+\varepsilon-y}^{3/2-x-y} F \mid_{G_{xyz}} dz \right)^2 dy\, dx.$$

Next comes

$$J_2 = \int_{x=1/2-\varepsilon}^{1/2-\varepsilon/2} \int_{y=1/2-\varepsilon}^{x} \left(\int_{z=0}^{y} F \mid_{A_{yzx}} + \int_{z=y}^{x} F \mid_{A_{zyx}} + \int_{z=x}^{1-\varepsilon-x} F \mid_{A_{xyz}} + \int_{z=1-\varepsilon-x}^{1-\varepsilon-y} F \mid_{B_{xyz}} \right.$$
$$\left. + \int_{z=1-\varepsilon-y}^{3/2-x-y} F \mid_{C_{xyz}} dz \right)^2 dy\, dx.$$

Third is the piece

$$J_3 = \int_{x=1/2-\varepsilon}^{1/2-\varepsilon/2} \int_{y=0}^{1/2-\varepsilon} \left(\int_{z=0}^{y} F \mid_{A_{yzx}} + \int_{z=y}^{x} F \mid_{A_{zyx}} + \int_{z=x}^{1-\varepsilon-x} F \mid_{A_{xyz}} + \int_{z=1-\varepsilon-x}^{1-\varepsilon-y} F \mid_{B_{xyz}} \right.$$
$$\left. + \int_{z=1-\varepsilon-y}^{1+\varepsilon-x} F \mid_{C_{xyz}} + \int_{z=1+\varepsilon-x}^{3/2-x-y} F \mid_{T_{xyz}} dz \right)^2 dy\, dx.$$

We now have dealt with all integrals involving A_{xyz}, and all remaining integrals pass through B_{zyx}. Continuing, we have

$$J_4 = \int_{x=1/2-\varepsilon/2}^{1/2} \int_{y=1/2-\varepsilon}^{1-\varepsilon-x} \left(\int_{z=0}^{y} F \lfloor_{A_{yzx}} + \int_{z=y}^{1-\varepsilon-x} F \lfloor_{A_{zyx}} + \int_{z=1-\varepsilon-x}^{x} F \lfloor_{B_{zyx}} \right.$$
$$\left. + \int_{z=x}^{1-\varepsilon-y} F \lfloor_{B_{xyz}} + \int_{z=1-\varepsilon-y}^{3/2-x-y} F \lfloor_{C_{xyz}} dz \right)^2 dy\, dx.$$

Another component is

$$J_5 = \int_{x=1/2-\varepsilon/2}^{1/2} \int_{y=0}^{1/2-\varepsilon} \left(\int_{z=0}^{y} F \lfloor_{A_{yzx}} + \int_{z=y}^{1-\varepsilon-x} F \lfloor_{A_{zyx}} \right.$$
$$+ \int_{z=1-\varepsilon-x}^{x} F \lfloor_{B_{zyx}} + \int_{z=x}^{1-\varepsilon-y} F \lfloor_{B_{xyz}} + \int_{z=1-\varepsilon-y}^{1+\varepsilon-x} F \lfloor_{C_{xyz}}$$
$$\left. + \int_{z=1+\varepsilon-x}^{3/2-x-y} F \lfloor_{T_{xyz}} dz \right)^2 dy\, dx.$$

The most complicated piece is

$$J_6 = \left(\int_{x=1/2}^{2\varepsilon} \int_{y=0}^{1-\varepsilon-x} + \int_{x=2\varepsilon}^{1/2+\varepsilon/2} \int_{y=x-2\varepsilon}^{1-\varepsilon-x} \right) \left(\int_{z=0}^{y} F \lfloor_{A_{yzx}} + \int_{z=y}^{1-\varepsilon-x} F \lfloor_{A_{zyx}} \right.$$
$$+ \int_{z=1-\varepsilon-x}^{x} F \lfloor_{B_{zyx}} + \int_{z=x}^{1-\varepsilon-y} F \lfloor_{B_{xyz}} + \int_{z=1-\varepsilon-y}^{1+\varepsilon-x} F \lfloor_{C_{xyz}} + \int_{z=1+\varepsilon-x}^{1/2+\varepsilon} F \lfloor_{S_{xyz}}$$
$$\left. + \int_{z=1/2+\varepsilon}^{3/2-x-y} F \lfloor_{T_{xyz}} dz \right)^2 dy\, dx.$$

Here we use $\left(\int_{x=1/2}^{2\varepsilon} \int_{y=0}^{1-\varepsilon-x} + \int_{x=2\varepsilon}^{1/2+\varepsilon/2} \int_{y=x-2\varepsilon}^{1-\varepsilon-x} \right) f(x,y)\, dy\,dx$ as an abbreviation for

$$\int_{x=1/2}^{2\varepsilon} \int_{y=0}^{1-\varepsilon-x} f(x,y)\, dy\,dx + \int_{x=2\varepsilon}^{1/2+\varepsilon/2} \int_{y=x-2\varepsilon}^{1-\varepsilon-x} f(x,y)\, dy\,dx.$$

We have now exhausted C_{xyz}. The seventh piece is

$$J_7 = \int_{x=2\varepsilon}^{1/2+\varepsilon/2} \int_{y=0}^{x-2\varepsilon} \left(\int_{z=0}^{y} F \lfloor_{A_{yzx}} + \int_{z=y}^{1-\varepsilon-x} F \lfloor_{A_{zyx}} + \int_{z=1-\varepsilon-x}^{x} F \lfloor_{B_{zyx}} \right.$$
$$+ \int_{z=x}^{1+\varepsilon-x} F \lfloor_{B_{xyz}} + \int_{z=1+\varepsilon-x}^{1-\varepsilon-y} F \lfloor_{E_{xyz}} + \int_{1-\varepsilon-y}^{1/2+\varepsilon} F \lfloor_{S_{xyz}}$$
$$\left. + \int_{1/2+\varepsilon}^{3/2-x-y} F \lfloor_{T_{xyz}} dz \right)^2 dy\, dx.$$

Finally, we have

$$J_8 = \int_{x=1/2+\varepsilon/2}^{1-\varepsilon} \int_{y=0}^{1-\varepsilon-x} \left(\int_{z=0}^{y} F \lfloor_{A_{yzx}} + \int_{z=y}^{1-\varepsilon-x} F \lfloor_{A_{zyx}} + \int_{z=1-\varepsilon-x}^{1+\varepsilon-x} F \lfloor_{B_{zyx}} \right.$$
$$+ \int_{z=1+\varepsilon-x}^{x} F \lfloor_{E_{zyx}} + \int_{z=x}^{1-\varepsilon-y} F \lfloor_{E_{xyz}} + \int_{1-\varepsilon-y}^{1/2+\varepsilon} F \lfloor_{S_{xyz}}$$
$$\left. + \int_{1/2+\varepsilon}^{3/2-x-y} F \lfloor_{T_{xyz}} dz \right)^2 dy\, dx.$$

In the case $\varepsilon = 1/4$, the marginal conditions (129) reduce to requiring

$$\int_{z=0}^{3/2-x-y} F \mid_{G_{yzx}} dz = 0 \tag{135}$$

$$\int_{z=0}^{y} F \mid_{G_{yzx}} + \int_{z=y}^{3/2-x-y} F \mid_{G_{zyx}} dz = 0 \tag{136}$$

$$\int_{z=0}^{1+\varepsilon-x} F \mid_{U_{yzx}} + \int_{z=1+\varepsilon-x}^{y} F \mid_{G_{yzx}} + \int_{z=y}^{3/2-x-y} F \mid_{G_{zyx}} dz = 0 \tag{137}$$

$$\int_{z=0}^{1+\varepsilon-x} F \mid_{U_{yzx}} + \int_{z=1+\varepsilon-x}^{3/2-x-y} F \mid_{G_{yzx}} dz = 0 \tag{138}$$

$$\int_{z=0}^{3/2-x-y} F \mid_{T_{yzx}} dz = 0 \tag{139}$$

$$\int_{z=0}^{1-\varepsilon-x} F \mid_{E_{yzx}} + \int_{z=1-\varepsilon-x}^{1-\varepsilon-y} F \mid_{S_{yzx}} + \int_{z=1-\varepsilon-y}^{3/2-x-y} F \mid_{H_{yzx}} dz = 0. \tag{140}$$

Each of these constraints is only required to hold for some portion of the parameter space $\{(x, y) : 1 + \varepsilon \leq x + y \leq 3/2\}$, but as the left-hand sides are all polynomial functions in x, y (using the signed definite integral $\int_b^a = -\int_a^b$), it is equivalent to require that all coefficients of these polynomial functions vanish.

Now we specify F. After some numerical experimentation, we have found that the simplest choice of F which still achieves the desired goal comes by taking $F(x, y, z)$ to be a polynomial of degree 1 on each of E_{xyz}, S_{xyz}, and H_{xyz}; degree 2 on T_{xyz}, vanishing on D_{xyz}; and degree 3 on the remaining five relevant components of R_{xyz}. After solving the quadratic program, rounding, and clearing denominators, we arrive at the choice

$$
\begin{aligned}
F \mid_{A_{xyz}} :=\ & -66 + 96x - 147x^2 + 125x^3 + 128y - 122xy + 104x^2y - 275y^2 + 394y^3 \\
& + 99z - 58xz + 63x^2z - 98yz + 51xyz + 41y^2z - 112z^2 + 24xz^2 + 72yz^2 \\
& + 50z^3
\end{aligned}
$$

$$
\begin{aligned}
F \mid_{B_{xyz}} :=\ & -41 + 52x - 73x^2 + 25x^3 + 108y - 66xy + 71x^2y - 294y^2 + 56xy^2 \\
& + 363y^3 + 33z + 15xz + 22x^2z - 40yz - 42xyz + 75y^2z - 36z^2 - 24xz^2 \\
& + 26yz^2 + 20z^3
\end{aligned}
$$

$$F \mid_{C_{xyz}} := -22 + 45x - 35x^2 + 63y - 99xy + 82x^2y - 140y^2 + 54xy^2 + 179y^3$$

$$F \mid_{E_{xyz}} := -12 + 8x + 32y$$

$$F \mid_{S_{xyz}} := -6 + 8x + 16y$$

$$F \mid_{T_{xyz}} := 18 - 30x + 12x^2 + 42y - 20xy - 66y^2 - 45z + 34xz + 22z^2$$

$$
\begin{aligned}
F \mid_{U_{xyz}} :=\ & 94 - 1{,}823x + 5{,}760x^2 - 5{,}128x^3 + 54y - 168x^2y + 105y^2 + 1{,}422xz \\
& - 2{,}340x^2z - 192y^2z - 128z^2 - 268xz^2 + 64z^3
\end{aligned}
$$

$$
\begin{aligned}
F \mid_{G_{xyz}} :=\ & 5{,}274 - 19{,}833x + 18{,}570x^2 - 5{,}128x^3 - 18{,}024y + 44{,}696xy \\
& - 20{,}664x^2y + 16{,}158y^2 - 19{,}056xy^2 - 4{,}592y^3 - 10{,}704z \\
& + 26{,}860xz - 12{,}588x^2z + 24{,}448yz - 30{,}352xyz - 10{,}980y^2z + 7{,}240z^2 \\
& - 9{,}092xz^2 - 8{,}288yz^2 - 1{,}632z^3
\end{aligned}
$$

$$F \mid_{H_{xyz}} := 8z.$$

One may compute that

$$I(F) = \frac{62,082,439,864,241}{507,343,011,840}$$

and

$$J(F) = \frac{9,933,190,664,926,733}{40,587,440,947,200}$$

with all the marginal conditions (135)-(140) obeyed, and thus

$$\frac{J(F)}{I(F)} = 2 + \frac{286,648,173}{4,966,595,189,139,280}$$

and (130) follows.

The parity problem

In this section, we argue why the 'parity barrier' of Selberg [7] prohibits sieve-theoretic methods, such as the ones in this paper, from obtaining any bound on H_1 that is stronger than $H_1 \leq 6$, even on the assumption of strong distributional conjectures such as the generalized Elliott-Halberstam conjecture GEH[ϑ] and even if one uses sieves other than the Selberg sieve. Our discussion will be somewhat informal and heuristic in nature.

We begin by briefly recalling how the bound $H_1 \leq 6$ on GEH (i.e., Theorem 4(xii)) was proven. This was deduced from the claim DHL[3; 2], or more specifically from the claim that the set

$$A := \{n \in \mathbb{N} : \text{at least two of} n, n+2, n+6 \text{ are prime}\} \tag{141}$$

was infinite.

To do this, we (implicitly) established a lower bound

$$\sum_n \nu(n)1_A(n) > 0$$

for some non-negative weight $\nu : \mathbb{N} \to \mathbb{R}^+$ supported on $[x, 2x]$ for a sufficiently large x. This bound was in turn established (after a lengthy sieve-theoretic analysis, and with a carefully chosen weight ν) from upper bounds on various discrepancies. More precisely, one required good upper bounds (on average) for the expressions

$$\left| \sum_{x \leq n \leq 2x : x = a \ (q)} f(n+h) - \frac{1}{\varphi(q)} \sum_{x \leq n \leq 2x : (n+h,q)=1} f(n+h) \right| \tag{142}$$

for all $h \in \{0, 2, 6\}$ and various residue classes $a \ (q)$ with $q \leq x^{1-\varepsilon}$ and arithmetic functions f, such as the constant function $f = 1$, the von Mangoldt function $f = \Lambda$, or Dirichlet convolutions $f = \alpha \star \beta$ of the type considered in Claim 12. (In the presentation of this argument in previous sections, the shift by h was eliminated using the change of variables $n' = n + h$, but for the current discussion, it is important that we do not use this shift.) One also required good asymptotic control on the main terms

$$\sum_{x \leq n \leq 2x : (n+h,q)=1} f(n+h). \tag{143}$$

An inspection of these arguments (which no longer exploit change of variables such as $n' = n + h$ in the n variable) shows that they would be equally valid if one inserted a

further non-negative weight $\omega : \mathbb{N} \to \mathbb{R}^+$ in the summation over n. More precisely, the above sieve-theoretic argument would also deduce the lower bound

$$\sum_n \nu(n) 1_A(n) \omega(n) > 0$$

if one had control on the weighted discrepancies

$$\left| \sum_{x \leq n \leq 2x : x \equiv a \ (q)} f(n+h)\omega(n) - \frac{1}{\varphi(q)} \sum_{x \leq n \leq 2x : (n+h,q)=1} f(n+h)\omega(n) \right| \qquad (144)$$

and on the weighted main terms

$$\sum_{x \leq n \leq 2x : (n+h,q)=1} f(n+h)\omega(n) \qquad (145)$$

that were of the same form as in the unweighted case $\omega = 1$.

Now suppose for instance that one was trying to prove the bound $H_1 \leq 4$. A natural way to proceed here would be to replace the set A in (141) with the smaller set

$$A' := \{n \in \mathbb{N} : n, n+2 \ \text{are both prime}\} \cup \{n \in \mathbb{N} : n+2, n+6 \ \text{are both prime}\} \quad (146)$$

and hope to establish a bound of the form

$$\sum_n \nu(n) 1_{A'}(n) > 0$$

for a well-chosen function $\nu : \mathbb{N} \to \mathbb{R}^+$ supported on $[x, 2x]$, by deriving this bound from suitable (averaged) upper bounds on the discrepancies (142) and control on the main terms (143). If the arguments were sieve-theoretic in nature, then (as in the $H_1 \leq 6$ case) one could then also deduce the lower bound

$$\sum_n \nu(n) 1_{A'}(n) \omega(n) > 0 \qquad (147)$$

for any non-negative weight $\omega : \mathbb{N} \to \mathbb{R}^+$, provided that one had the same control on the weighted discrepancies (144) and weighted main terms (145) that one did on (142) and (143).

We apply this observation to the weight

$$\omega(n) := (1 - \lambda(n)\lambda(n+2))(1 - \lambda(n+2)\lambda(n+6))$$
$$= 1 - \lambda(n)\lambda(n+2) - \lambda(n+2)\lambda(n+6) + \lambda(n)\lambda(n+6)$$

where $\lambda(n) := (-1)^{\Omega(n)}$ is the Liouville function. Observe that ω vanishes for any $n \in A'$, and hence

$$\sum_n \nu(n) 1_{A'}(n) \omega(n) = 0 \qquad (148)$$

for any ν. On the other hand, the 'Möbius randomness law' (see, e.g. [33]) predicts a significant amount of cancellation for any non-trivial sum involving the Möbius function μ or the closely related Liouville function λ. For instance, the expression

$$\sum_{x \leq n \leq 2x : n \equiv a \ (q)} \lambda(n+h)$$

is expected to be very small (of size[g] $O\left(\frac{x}{q}\log^{-A} x\right)$ for any fixed A) for any residue class a
(q) with $q \le x^{1-\varepsilon}$, and any $h \in \{0, 2, 6\}$; similarly for more complicated expressions such
as

$$\sum_{x \le n \le 2x: n=a \ (q)} \lambda(n+2)\lambda(n+6)$$

or

$$\sum_{x \le n \le 2x: n=a \ (q)} \Lambda(n)\lambda(n+2)\lambda(n+6)$$

or more generally

$$\sum_{x \le n \le 2x: n=a \ (q)} f(n)\lambda(n+2)\lambda(n+6)$$

where f is a Dirichlet convolution $\alpha \star \beta$ of the form considered in Claim 12. Similarly for
expressions such as

$$\sum_{x \le n \le 2x: n=a \ (q)} f(n)\lambda(n)\lambda(n+2);$$

note from the complete multiplicativity of λ that $(\alpha \star \beta)\lambda = (\alpha\lambda) \star (\beta\lambda)$, so if f is of the
form in Claim 12, then $f\lambda$ is also. In view of these observations (and similar observations
arising from permutations of $\{0, 2, 6\}$), we conclude (heuristically, at least) that all the
bounds that are believed to hold for (142) and (143) should also hold (up to minor changes
in the implied constants) for (144) and (145). Thus, if the bound $H_1 \le 4$ could be proven
in a sieve-theoretic fashion, one should be able to conclude the bound (147), which is in
direct contradiction to (148).

Remark 45. Similar arguments work for any set of the form

$$A_H := \{n \in \mathbb{N} : \exists n \le p_1 < p_2 \le n + H; p_1, p_2 \text{ both prime}, p_2 - p_1 \le 4\}$$

and any fixed $H > 0$, to prohibit any non-trivial lower bound on $\sum_n \nu(n)1_{A_H}(n)$ from
sieve-theoretic methods. Indeed, one uses the weight

$$\omega(n) := \prod_{0 \le i \le i' \le H; (n+i,3)=(n+i',3)=1; i'-i \le 4} (1 - \lambda(n+i)\lambda(n+i'));$$

we leave the details to the interested reader. This seems to block any attempt to use any
argument based only on the distribution of the prime numbers and related expressions in
arithmetic progressions to prove $H_1 \le 4$.

The same arguments of course also prohibit a sieve-theoretic proof of the twin prime
conjecture $H_1 = 2$. In this case, one can use the simpler weight $\omega(n) = 1 - \lambda(n)\lambda(n+2)$
to rule out such a proof, and the argument is essentially due to Selberg [7].

Of course, the parity barrier could be circumvented if one were able to introduce
stronger sieve-theoretic axioms than the 'linear' axioms currently available (which only
control sums of the form (142) or (143)). For instance, if one were able to obtain
non-trivial bounds for 'bilinear' expressions such as

$$\sum_{x \le n \le 2x} f(n)\Lambda(n+2) = \sum_d \sum_m \alpha(d)\beta(m)1_{[x,2x]}(dm)\Lambda(dm+2)$$

for functions $f = \alpha \star \beta$ of the form in Claim 12, then (by a modification of the proof of Proposition 13) one would very likely obtain non-trivial bounds on

$$\sum_{x \leq n \leq 2x} \Lambda(n)\Lambda(n+2)$$

which would soon lead to a proof of the twin prime conjecture. Unfortunately, we do not know of any plausible way to control such bilinear expressions. (Note however that there are some other situations in which bilinear sieve axioms may be established, for instance in the argument of Friedlander and Iwaniec [40] establishing an infinitude of primes of the form $a^2 + b^4$.)

Additional remarks

The proof of Theorem 16(xii) may be modified to establish the following variant:

Proposition 46. *Assume the generalized Elliott-Halberstam conjecture* GEH$[\theta]$ *for all* $0 < \theta < 1$. *Let* $0 < \varepsilon < 1/2$ *be fixed. Then, if x is a sufficiently large multiple of 6, there exists a natural number n with* $\varepsilon x \leq n \leq (1 - \varepsilon)x$ *such that at least two of* $n, n-2, x-n$ *are prime, and similarly if* $n - 2$ *is replaced by* $n + 2$.

Note that if at least two of $n, n-2, x-n$ are prime, then either $n, n+2$ are twin primes or else at least one of $x, x-2$ is expressible as the sum of two primes, and Theorem 5 easily follows.

Proof. (Sketch) We just discuss the case of $n - 2$, as the $n + 2$ case is similar. Observe from the Chinese remainder theorem (and the hypothesis that x is divisible by 6) that one can find a residue class $b\ (W)$ such that $b, b-2, x-b$ are all coprime to W (in particular, one has $b = 1\ (6)$). By a routine modification of the proof of Lemma 18, it suffices to find a non-negative weight function $\nu \colon \mathbb{N} \to \mathbb{R}^+$ and fixed quantities $\alpha > 0$ and $\beta_1, \beta_2, \beta_3 \geq 0$, such that one has the asymptotic upper bound

$$\sum_{\substack{\varepsilon x \leq n \leq (1-\varepsilon)x \\ n = b\ (W)}} \nu(n) \leq \mathfrak{S}(\alpha + o(1))B^{-k}\frac{(1-2\varepsilon)x}{W},$$

the asymptotic lower bounds

$$\sum_{\substack{\varepsilon x \leq n \leq (1-\varepsilon)x \\ n = b\ (W)}} \nu(n)\theta(n) \geq \mathfrak{S}(\beta_1 - o(1))B^{1-k}\frac{(1-2\varepsilon)x}{\varphi(W)}$$

$$\sum_{\substack{\varepsilon x \leq n \leq (1-\varepsilon)x \\ n = b\ (W)}} \nu(n)\theta(n+2) \geq \mathfrak{S}(\beta_2 - o(1))B^{1-k}\frac{(1-2\varepsilon)x}{\varphi(W)}$$

$$\sum_{\substack{\varepsilon x \leq n \leq (1-\varepsilon)x \\ n = b\ (W)}} \nu(n)\theta(x-n) \geq \mathfrak{S}(\beta_3 - o(1))B^{1-k}\frac{(1-2\varepsilon)x}{\varphi(W)}$$

and the inequality

$$\beta_1 + \beta_2 + \beta_3 > 2\alpha,$$

where $\mathfrak{S}$ is the singular series

$$\mathfrak{S} := \prod_{p|x(x-2);p>w} \frac{p}{p-1}.$$

We select ν to be of the form

$$\nu(n) = \left(\sum_{j=1}^{J} c_j \lambda_{F_{j,1}}(n) \lambda_{F_{j,2}}(n+2) \lambda_{F_{j,3}}(x-n) \right)^2$$

for various fixed coefficients $c_1, \ldots, c_J \in \mathbb{R}$ and fixed smooth compactly supported functions $F_{j,i} : [0, +\infty) \to \mathbb{R}$ with $j = 1, \ldots, J$ and $i = 1, \ldots, 3$. It is then routine[h] to verify that analogues of Theorem 19 and Theorem 20 hold for the various components of ν, with the role of x in the right-hand side replaced by $(1 - 2\varepsilon)x$, and the claim then follows by a suitable modification of Theorem 28, taking advantage of the function F constructed in Theorem 29. $\square$

It is likely that the bounds in Theorem 4 can be improved further by refining the sieve-theoretic methods employed in this paper, with the exception of part (xii) for which the parity problem prevents further improvement, as discussed in the 'The parity problem' section. We list some possible avenues to such improvements as follows:

1. In Theorem 27, the bound $M_{k,\varepsilon} > 4$ was obtained for some $\varepsilon > 0$ and $k = 50$. It is possible that k could be lowered slightly, for instance to $k = 49$, by further numerical computations, but we were only barely able to establish the $k = 50$ bound after 2 weeks of computation. However, there may be a more efficient way to solve the required variational problem (e.g. by selecting a more efficient basis than the symmetric monomial basis) that would allow one to advance in this direction; this would improve the bound $H_1 \le 246$ slightly. Extrapolation of existing numerics also raises the possibility that M_{53} exceeds 4, in which case the bound of 270 in Theorem 4(vii) could be lowered to 264.

2. To reduce k (and thus H_1) further, one could try to solve another variational problem, such as the one arising in Theorem 24 or in Theorem 28, rather than trying to lower bound M_k or $M_{k,\varepsilon}$. It is also possible to use the more complicated versions of MPZ$[\varpi, \delta]$ established (in which the modulus q is assumed to be densely divisible rather than smooth) to replace the truncated simplex appearing in Theorem 24 with a more complicated region (such regions also appear implicitly in [§4.5]). However, in the medium-dimensional setting $k \approx 50$, we were not able to accurately and rapidly evaluate the various integrals associated to these variational problems when applied to a suitable basis of functions. One key difficulty here is that whereas polynomials appear to be an adequate choice of basis for the M_k, an analysis of the Euler-Lagrange equation reveals that one should use piecewise polynomial basis functions instead for more complicated variational problems such as the $M_{k,\varepsilon}$ problem (as was done in the three-dimensional case in the 'Three-dimensional cutoffs' section), and these are difficult to work with in medium dimensions. From our experience with the low k problems, it looks like one should allow these piecewise polynomials to have relatively high degree on

some polytopes and low degree on other polytopes, and vanish completely on yet further polytopes[i], but we do not have a systematic understanding of what the optimal placement of degrees should be.

3. In Theorem 28, the function F was required to be supported in the simplex $\frac{k}{k-1} \cdot \mathcal{R}_k$. However, one can consider functions F supported in other regions R, subject to the constraint that all elements of the sumset $R + R$ lie in a region treatable by one of the cases of Theorem 20. This could potentially lead to other optimization problems that lead to superior numerology, although again it appears difficult to perform efficient numerics for such problems in the medium k regime $k \approx 50$. One possibility would be to adopt a 'free boundary' perspective, in which the support of F is not fixed in advance, but is allowed to evolve by some iterative numerical scheme.

4. To improve the bounds on H_m for $m = 2, 3, 4, 5$, one could seek a better lower bound on M_k than the one provided by Theorem 40; one could also try to lower bound more complicated quantities such as $M_{k,\varepsilon}$.

5. One could attempt to improve the range of ϖ, δ for which estimates of the form MPZ$[\varpi, \delta]$ are known to hold, which would improve the results of Theorem 4(ii)-(vi). For instance, we believe that the condition $600\varpi + 180\delta < 7$ in Theorem 11 could be improved slightly to $1,080\varpi + 330\delta < 13$ by refining the arguments, but this requires a hypothesis of square root cancellation in a certain four-dimensional exponential sum over finite fields, which we have thus far been unable to establish rigorously. Another direction to pursue would be to improve the δ parameter, or to otherwise relax the requirement of smoothness in the moduli, in order to reduce the need to pass to a truncation of the simplex $\mathcal{R}_k$, which is the primary reason why the $m = 1$ results are currently unable to use the existing estimates of the form MPZ$[\varpi, \delta]$. Another speculative possibility is to seek MPZ$[\varpi, \delta]$ type estimates which only control distribution for a positive proportion of smooth moduli, rather than for all moduli, and then to design a sieve ν adapted to just that proportion of moduli (cf. [41]). Finally, there may be a way to combine the arguments currently used to prove MPZ$[\varpi, \delta]$ with the automorphic forms (or 'Kloostermania') methods used to prove nontrivial equidistribution results with respect to a fixed modulus, although we do not have any ideas on how to actually achieve such a combination.

6. It is also possible that one could tighten the argument in Lemma 18, for instance by establishing a non-trivial lower bound on the portion of the sum $\sum_n \nu(n)$ when $n + h_1, \ldots, n + h_k$ are all composite, or a sufficiently strong upper bound on the pair correlations $\sum_n \theta(n + h_i)\theta(n + h_j)$ (see [9, §6] for a recent implementation of this latter idea). However, our preliminary attempts to exploit these adjustments suggested that the gain from the former idea would be exponentially small in k, whereas the gain from the latter would also be very slight (perhaps reducing k by $O(1)$ in large k regimes, e.g. $k \geq 5,000$).

7. All of our sieves used are essentially of Selberg type, being the square of a divisor sum. We have experimented with a number of non-Selberg type sieves (for instance trying to exploit the obvious positivity of $1 - \sum_{p \leq x : p | n} \frac{\log p}{\log x}$ when $n \leq x$); however, none of these variants offered a numerical improvement over the Selberg sieve. Indeed it appears that after optimizing the cutoff function F, the Selberg sieve is in

some sense a 'local maximum' in the space of non-negative sieve functions, and one would need a radically different sieve to obtain numerically superior results.

8. Our numerical bounds for the diameter $H(k)$ of the narrowest admissible k-tuple are known to be exact for $k \leq 342$, but there is scope for some slight improvement for larger values of k, which would lead to some improvements in the bounds on H_m for $m = 2, 3, 4, 5$. However, we believe that our bounds on H_m are already fairly close (e.g. within 10%) of optimal, so there is only a limited amount of gain to be obtained solely from this component of the argument.

Narrow admissible tuples

In this section, we outline the methods used to obtain the numerical bounds on $H(k)$ given by Theorem 17, which are reproduced below:

1. $H(3) = 6$,
2. $H(50) = 246$,
3. $H(51) = 252$,
4. $H(54) = 270$,
5. $H(5, 511) \leq 52, 116$,
6. $H(35, 410) \leq 398, 130$,
7. $H(41, 588) \leq 474, 266$,
8. $H(309, 661) \leq 4, 137, 854$,
9. $H(1, 649, 821) \leq 24, 797, 814$,
10. $H(75, 845, 707) \leq 1, 431, 556, 072$,
11. $H(3, 473, 955, 908) \leq 80, 550, 202, 480$.

H(k) values for small k

The equalities in the first four bounds (1)-(4) were previously known. The case $H(3) = 6$ is obvious: the admissible 3-tuples $(0, 2, 6)$ and $(0, 4, 6)$ have diameter 6 and no 3-tuple of smaller diameter is admissible. The cases $H(50) = 246$, $H(51) = 252$, and $H(54) = 270$ follow from results of Clark and Jarvis [42]. They define $\varrho^*(x)$ to be the largest integer k for which there exists an admissible k-tuple that lies in a half-open interval $(y, y + x]$ of length x. For each integer $k > 1$, the largest x for which $\varrho^*(x) = k$ is precisely $H(k + 1)$. Table 1 of [42] lists these largest x values for $2 \leq k \leq 170$, and we find that $H(50) = 246$, $H(51) = 252$, and $H(54) = 270$. Admissible tuples that realize these bounds are shown in Subsubsections "Admissible 50-tuple realizing H(50) = 246", "Admissible 51-tuple realizing H(51) = 252" and "Admissible 54-tuple realizing H(54) = 270".

Admissible 50-tuple realizing H(50) = 246

$0, 4, 6, 16, 30, 34, 36, 46, 48, 58, 60, 64, 70, 78, 84, 88, 90, 94, 100, 106,$

$108, 114, 118, 126, 130, 136, 144, 148, 150, 156, 160, 168, 174, 178, 184,$

$190, 196, 198, 204, 210, 214, 216, 220, 226, 228, 234, 238, 240, 244, 246.$

Admissible 51-tuple realizing H(51) = 252

$0, 6, 10, 12, 22, 36, 40, 42, 52, 54, 64, 66, 70, 76, 84, 90, 94, 96, 100, 106,$

$112, 114, 120, 124, 132, 136, 142, 150, 154, 156, 162, 166, 174, 180, 184,$

$190, 196, 202, 204, 210, 216, 220, 222, 226, 232, 234, 240, 244, 246, 250, 252.$

Admissible 54-tuple realizing H(54) = 270

0, 4, 10, 18, 24, 28, 30, 40, 54, 58, 60, 70, 72, 82, 84, 88, 94, 102, 108, 112, 114,

118, 124, 130, 132, 138, 142, 150, 154, 160, 168, 172, 174, 180, 184, 192, 198, 202,

208, 214, 220, 222, 228, 234, 238, 240, 244, 250, 252, 258, 262, 264, 268, 270.

H(k) bounds for mid-range k

As previously noted, exact values for $H(k)$ are known only for $k \leq 342$. The upper bounds on $H(k)$ for the five cases (5)-(9) were obtained by constructing admissible k-tuples using techniques developed during the first part of the Polymath8 project. These are described in detail in section 3 of [4], but for the sake of completeness, we summarize the most relevant methods here.

Fast admissibility testing

A key component of all our constructions is the ability to efficiently determine whether a given k-tuple $\mathcal{H} = (h_1, \ldots, h_k)$ is admissible. We say that $\mathcal{H}$ is *admissible modulo p* if its elements do not form a complete set of residues modulo p. Any k-tuple $\mathcal{H}$ is automatically admissible modulo all primes $p > k$, since a k-tuple cannot occupy more than k residue classes; thus, we only need to test admissibility modulo primes $p < k$.

A simple way to test admissibility modulo p is to enumerate the elements of $\mathcal{H}$ modulo p and keep track of which residue classes have been encountered in a table with p boolean-valued entries. Assuming the elements of $\mathcal{H}$ have absolute value bounded by $O(k \log k)$ (true of all the tuples we consider), this approach yields a total bit-complexity of $O(k^2 / \log k \ \mathsf{M}(\log k))$, where $\mathsf{M}(n)$ denotes the complexity of multiplying two n-bit integers, which, up to a constant factor, also bounds the complexity of division with remainder. Applying the Schönhage-Strassen bound $\mathsf{M}(n) = O(n \log n \log \log n)$ from [43], this is $O(k^2 \log \log k \log \log \log k)$, essentially quadratic in k.

This approach can be improved by observing that for most of the primes $p < k$, there are likely to be many unoccupied residue classes modulo p. In order to verify admissibility at p, it is enough to find one of them, and we typically do not need to check them all in order to do so. Using a heuristic model that assumes the elements of $\mathcal{H}$ are approximately equidistributed modulo p, one can determine a bound $m < p$ such that k random elements of $\mathbb{Z}/p\mathbb{Z}$ are unlikely to occupy all of the residue classes in $[0, m]$. By representing the k-tuple $\mathcal{H}$ as a boolean vector $\mathcal{B} = (b_0, \ldots, b_{h_k - h_1})$ in which $b_i = 1$ if and only if $i = h_j - h_1$ for some $h_j \in \mathcal{H}$, we can efficiently test whether $\mathcal{H}$ occupies every residue class in $[0, m]$ by examining the the entries

$$b_0, \ldots, b_m, b_p, \ldots, b_{p+m}, b_{2p}, \ldots, b_{2p+m}, \ldots$$

of $\mathcal{B}$. The key point is that when $p < k$ is large, say $p > (1+\epsilon)k/\log k$, we can choose m so that we only need to examine a small subset of the entries in $\mathcal{B}$. Indeed, for primes $p > k/c$ (for any constant c), we can take $m = O(1)$ and only need to examine $O(\log k)$ elements of $\mathcal{B}$ (assuming its total size is $O(k \log k)$, which applies to all the tuples we consider here).

Of course it may happen that $\mathcal{H}$ occupies every residue class in $[0, m]$ modulo p. In this case, we revert to our original approach of enumerating the elements of $\mathcal{H}$ modulo p, but we expect this to happen for only a small proportion of the primes $p < k$. Heuristically, this reduces the complexity of admissibility testing by a factor of $O(\log k)$, making

it sub-quadratic. In practice, we find this approach to be much more efficient than the straightforward method when k is large (see [§3.1] for further details.

Sieving methods

Our techniques for constructing admissible k-tuples all involve sieving an integer interval $[s, t]$ of residue classes modulo primes $p < k$ and then selecting an admissible k-tuple from the survivors. There are various approaches one can take, depending on the choice of interval and the residue classes to sieve. We list four of these below, starting with the classical sieve of Eratosthenes and proceeding to more modern variations.

- *Sieve of Eratosthenes.* We sieve an interval $[2, x]$ to obtain admissible k-tuples

  $$p_{m+1}, \ldots, p_{m+k}.$$

  with m as small as possible. If we sieve the the residue class $0(p)$ for all primes $p \leq k$, we have $m = \pi(k)$ and $p_{m+1} > k$. In this case, no admissibility testing is required, since the residue class $0(p)$ is unoccupied for all $p \leq k$. Applying the Prime Number Theorem in the forms

  $$p_k = k \log k + k \log \log k - k + O\left(k \frac{\log \log k}{\log k}\right),$$

  $$\pi(x) = \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right),$$

  this construction yields the upper bound

  $$H(k) \leq k \log k + k \log \log k - k + o(k). \tag{149}$$

  As an optimization, rather than sieving modulo every prime $p \leq k$, we instead sieve modulo increasing primes p and as soon as the first k survivors form an admissible tuple. This will typically happen for for some $p_m < k$.

- *Hensley-Richards sieve.* The bound in (149) was improved by Hensley and Richards [44-46], who observed that rather than sieving $[2, x]$ it is better to sieve the interval $[-x/2, x/2]$ to obtain admissible k-tuples of the form

  $$-p_{m+\lfloor k/2 \rfloor - 1}, \ldots, p_{m+1}, \ldots, -1, 1, \ldots, p_{m+1}, \ldots, p_{m+\lfloor (k+1)/2 \rfloor - 1},$$

  where we again wish to make m as small as possible. It follows from Lemma 5 of [45] that one can take $m = o(k/\log k)$, leading to the improved upper bound

  $$H(k) \leq k \log k + k \log \log k - (1 + \log 2)k + o(k). \tag{150}$$

- *Shifted Schinzel sieve.* As noted by Schinzel in [47], in the Hensley-Richards sieve, it is slightly better to sieve $1(2)$ rather than $0(2)$; this leaves unsieved powers of 2 near the center of the interval $[-x/2, x/2]$ that would otherwise be removed (more generally, one can sieve $1(p)$ for many small primes p, but we did not). Additionally, we find that shifting the interval $[-x/2, x/2]$ can yield significant improvements (one can also view this as changing the choices of residue classes).

  This leads to the following approach: we sieve an interval $[s, s + x]$ of odd integers and multiples of odd primes $p \leq p_m$, where x is large enough to ensure at least k survivors, and m is large enough to ensure that the survivors form an admissible tuple, with x and m minimal subject to these constraints. A tuple of exactly k survivors is then chosen to minimize the diameter. By varying s and comparing the results, we can choose a starting point $s \in [-x/2, x/2]$ that yields the smallest final

diameter. For large k, we typically find $s \approx k$ is optimal, as opposed to $s \approx -(k/2) \log k$ in the Hensley-Richards sieve.

- *Shifted greedy sieve.* As a further optimization, we can allow greater freedom in the choice of residue class to sieve. We begin as in the shifted Schinzel sieve, but for primes $p \le p_m$ that exceed $2\sqrt{k \log k}$, rather than sieving $0(p)$, we choose a minimally occupied residue class $a(p)$. As above, we sieve the interval $[s, s + x]$ for varying values of $s \in [-x/2, x/2]$ and select the best result, but unlike the shifted Schinzel sieve, for large k, we typically choose $s \approx -(k/\log k - k)/2$.

  We remark that while one might suppose that it would be better to choose a minimally occupied residue class at all primes, not just the larger ones, we find that this is generally not the case. Fixing a structured choice of residue classes for the small primes avoids the erratic behavior that can result from making greedy choices to soon (see [48, Fig. 1] for an illustration of this).

Table 4 lists the bounds obtained by applying each of these techniques (in the online version of this paper, each table entry includes a link to the constructed tuple). To the admissible tuples obtained using the shifted greedy sieve, we additionally applied various local optimizations that are detailed in ([§3.6]). As can be seen in the table, the additional improvement due to these local optimizations is quite small compared to that gained by using better sieving algorithms, especially when k is large.

Table 4 also lists the value $\lfloor k \log k + k \rfloor$ that we conjecture as an upper bound on $H(k)$ for all sufficiently large k.

$H(k)$ bounds for large k

The upper bounds on $H(k)$ for the last two cases (10) and (11) were obtained using modified versions of the techniques described above that are better suited for handling very large values of k. These entail three types of optimizations that are summarized in the subsections below.

Improved time complexity

As noted above, the complexity of admissibility testing is quasi-quadratic in k. Each of the techniques listed in the '$H(k)$ bounds for mid-range k' section involves optimizing over a parameter space whose size is at least quasi-linear in k, leading to an overall quasi-cubic time complexity for constructing a narrow admissible k-tuple; this makes it impractical to handle $k > 10^9$. We can reduce this complexity in a number of ways.

First, we can combine parameter optimization and admissibility testing. In both the sieve of Eratosthenes and Hensley-Richards sieves, taking $m = k$ guarantees an

Table 4 Upper bounds on $H(k)$ for selected values of k

k	5,511	35,410	41,588	309,661	1,649,821
k primes past k	56,538	433,992	516,586	4,505,700	26,916,060
Eratosthenes	55,160	424,636	505,734	4,430,212	26,540,720
Hensley-Richards	54,480	415,642	494,866	4,312,612	25,841,884
Shifted Schinzel	53,774	411,060	489,056	4,261,858	25,541,910
Shifted greedy	52,296	399,936	476,028	4,142,780	24,798,306
Best known	52,116	398,130	474,266	4,137,854	24,797,814
$\lfloor k \log k + k \rfloor$	52,985	406,320	483,899	4,224,777	25,268,951

admissible k-tuple. For $m < k$, if the corresponding k-tuple is inadmissible, it is typically because it is inadmissible modulo the smallest prime p_{m+1} that appears in the tuple. This suggests a heuristic approach in which we start with $m = k$, and then iteratively reduce m, testing the admissibility of each k-tuple modulo p_{m+1} as we go, until we can proceed no further. We then verify that the last k-tuple that was admissible modulo p_{m+1} is also admissible modulo all primes $p > p_{m+1}$ (we know it is admissible at all primes $p \leq p_m$ because we have sieved a residue class for each of these primes). We expect this to be the case, but if not we can increase m as required. Heuristically, this yields a quasi-quadratic running time, and in practice, it takes less time to find the minimal m than it does to verify the admissibility of the resulting k-tuple.

Second, we can avoid a complete search of the parameter space. In the case of the shifted Schinzel sieve, for example, we find empirically that taking $s = k$ typically yields an admissible k-tuple whose diameter is not much larger than that achieved by an optimal choice of s; we can then simply focus on optimizing m using the strategy described above. Similar comments apply to the shifted greedy sieve.

Improved space complexity

We expect a narrow admissible k-tuple to have diameter $d = (1 + o(1))k \log k$. Whether we encode this tuple as a sequence of k integers, or as a bitmap of $d + 1$ bits, as in the fast admissibility testing algorithm, we will need approximately $k \log k$ bits. For $k > 10^9$, this may be too large to conveniently fit in memory. We can reduce the space to $O(k \log \log k)$ bits by encoding the k-tuple as a sequence of $k - 1$ gaps; the average gap between consecutive entries has size $\log k$ and can be encoded in $O(\log \log k)$ bits. In practical terms, for the sequences we constructed, almost all gaps can be encoded using a single 8-bit byte for each gap.

One can further reduce space by partitioning the sieving interval into windows. For the construction of our largest tuples, we used windows of size $O(\sqrt{d})$ and converted to a gap-sequence representation only after sieving at all primes up to an $O(\sqrt{d})$ bound.

Parallelization

With the exception of the greedy sieve, all the techniques described above are easily parallelized. The greedy sieve is more difficult to parallelize because the choice of a minimally occupied residue class modulo p depends on the set of survivors obtained after sieving modulo primes less than p. To address this issue, we modified the greedy approach to work with batches of consecutive primes of size n, where n is a multiple of the number of parallel threads of execution. After sieving fixed residue classes modulo all small primes $p < 2\sqrt{k \log k}$, we determine minimally occupied residue classes for the next n primes in parallel, sieve these residue classes, and then proceed to the next batch of n primes.

In addition to the techniques described above, we also considered a modified Schinzel sieve in which we check admissibility modulo each successive prime p before sieving multiples of p, in order to verify that sieving modulo p is actually necessary. For values of p close to but slightly less than p_m, it will often be the case that the set of survivors is already admissibility modulo p, even though it does contain multiples of p (because some other residue class is unoccupied). As with the greedy sieve, when using this approach, we sieve residue classes in batches of size n to facilitate parallelization.

Table 5 Upper bounds on $H(k)$ for selected values of k

k	75,845,707	3,473,955,908
k primes past *k*	1,541,858,666	84,449,123,072
Eratosthenes	1,526,698,470	83,833,839,848
Hensley-Richards	1,488,227,220	81,912,638,914
Shifted Schinzel	1,467,584,468	80,761,835,464
Shifted greedy	1,431,556,072	Not available
Best known	1,431,556,072	80,550,202,480
$\lfloor k \log k + k \rfloor$	1,452,006,268	79,791,764,059

Results for large k

Table 5 lists the bounds obtained for the two largest values of *k*. For $k = 75,845,707$, the best results were obtained with a shifted greedy sieve that was modified for parallel execution as described above, using the fixed shift parameter $s = -(k \log k - k)/2$. A list of the sieved residue classes is available at math.mit.edu/~drew/n75845707_1431556072.txt. This file contains values of *k*, *s*, *d*, and *m*, along with a list of prime indices $n_i > m$ and residue classes r_i such that sieving the interval $[s, s + d]$ of odd integers, multiples of p_n for $1 < n \le m$, and at r_i modulo p_{n_i} yields an admissible *k*-tuple.

For $k = 3,473,955,908$, we did not attempt any form of greedy sieving due to practical limits on the time and computational resources available. The best results were obtained using a modified Schinzel sieve that avoids unnecessary sieving, as described above, using the fixed shift parameter $s = k0$. A list of the sieved residue classes is available at math.mit.edu/~drew/n75845707_1431556072.txt.

This file contains values of *k*, *s*, *d*, and *m*, along with a list of prime indices $n_i > m$ such that sieving the interval $[s, s + d]$ of odd integers, multiples of p_n for $1 < n \le m$, and multiples of p_{n_i} yields an admissible *k*-tuple.

Source code for our implementation is available at http://math.mit.edu/~drew/ompadm_v0.5.tar; this code can be used to verify the admissibility of both the tuples listed above.

Endnotes

[a]When a, b are real numbers, we will also need to use (a, b) and $[a, b]$ to denote the open and closed intervals, respectively, with endpoints a, b. Unfortunately, this notation conflicts with the notation given above, but it should be clear from the context which notation is in use.

[b]Actually, there are some differences between Conjecture 1 of [28] and the claim here. Firstly, we need an estimate that is uniform for all a, whereas in [28] only the case of a fixed modulus a was asserted. On the other hand, α, β were assumed to be controlled in ℓ^2 instead of via the pointwise bounds (6), and Q was allowed to be as large as $x \log^{-C} x$ for some fixed C (although, in view of the negative results in [23,24], this latter strengthening may be too ambitious).

[c]One could also use the Heath-Brown identity [49] here if desired.

[d]In the $k = 1$ case, we of course just have $q_{W,d_1,\dots,d'_{k-1}} = W$.

[e]One could obtain a small improvement to the bounds here by replacing the threshold $2c$ with a parameter to be optimized over.

[f]The arguments in [5] are rigorous under the assumption of a positive eigenfunction as in Corollary 35, but the existence of such an eigenfunction remains open for $k \ge 3$.

[g]Indeed, one might be even more ambitious and conjecture a square-root cancellation $\lll \sqrt{x/q}$ for such sums (see [50] for some similar conjectures), although such stronger cancellations generally do not play an essential role in sieve-theoretic computations.

[h]One new technical difficulty here is that some of the various moduli $[d_j, d_j']$ arising in these arguments are not required to be coprime at primes $p > w$ dividing x or $x - 2$; this requires some modification to Lemma 30 that ultimately leads to the appearance of the singular series $\mathfrak{S}$. However, these modifications are quite standard, and we do not give the details here.

[i]In particular, the optimal choice F for $M_{k,\varepsilon}$ should vanish on the polytope
$\{(t_1, \dots, t_k) \in (1 + \varepsilon) \cdot \mathcal{R}_k : \sum_{i \neq i_0} t_i \geq 1 - \varepsilon \text{ for all } i_0 = 1, \dots, k\}.$

References

1. Hardy, GH, Littlewood, JE: Some problems of "Partitio Numerorum", III: on the expression of a number as a sum of primes. Acta Math. **44**, 1–70 (1923)
2. Goldston, D, Pintz, J, Yıldırım, C: Primes in tuples. I. Ann. Math. **170**(2), 819–862 (2009)
3. Zhang, Y: Bounded gaps between primes. Annals Math. **179**, 1121–1174 (2014)
4. Polymath, DHJ: New equidistribution estimates of Zhang type, and bounded gaps between primes. arxiv.org/abs/1402.0811v2
5. Maynard, J: Small gaps between primes. Annals Math. to appear.
6. Granville, A: Bounded gaps between primes, preprint.
7. Selberg, A: On elementary methods in prime number-theory and their limitations. In: Proc. 11th Scand. Math. Cong. Trondheim (1949), Collected Works, vol. I, pp. 388–397. Springer-Verlag, Berlin-Göttingen-Heidelberg, (1989)
8. Pintz, J: The bounded gap conjecture and bounds between consecutive Goldbach numbers. Acta Arith. **155**(4), 397–405 (2012)
9. Banks, WD, Freiberg, T, Maynard, J: On limit points of the sequence of normalized prime gaps, preprint.
10. Banks, WD, Freiberg, T, Turnage-Butterbaugh, CL: Consecutive primes in tuples, preprint.
11. Benatar, J: The existence of small prime gaps in subsets of the integers, preprint.
12. Castillo, A, Hall, C, Lemke Oliver, RJ, Pollack, P, Thompson, L: Bounded gaps between primes in number fields and function fields, preprint.
13. Chua, L, Park, S, Smith, G. D: Bounded gaps between primes in special sequences, preprint.
14. Freiberg, T: A note on the theorem of Maynard and Tao, preprint.
15. Li, H, Pan, H: Bounded gaps between primes of the special form, preprint.
16. Maynard, J: Dense clusters of primes in subsets, preprint.
17. Pintz, J: On the ratio of consecutive gaps between primes, preprint.
18. Pintz, J: On the distribution of gaps between consecutive primes, preprint.
19. Pollack, P: Bounded gaps between primes with a given primitive root, preprint.
20. Pollack, P, Thompson, L: Arithmetic functions at consecutive shifted primes, preprint.
21. Thorner, J: Bounded Gaps Between Primes in Chebotarev Sets, preprint.
22. Elliott, PDTA, Halberstam, H: A conjecture in prime number theory. Symp. Math. **4**, 59–72 (1968)
23. Friedlander, J, Granville, A: Relevance of the residue class to the abundance of primes. In: Proceedings of the Amalfi Conference on Analytic Number Theory (Maiori, 1989), pp. 95–103. Univ. Salerno, Salerno, (1992)
24. Friedlander, J, Granville, A, Hildebrand, A, Maier, H: Oscillation theorems for primes in arithmetic progressions and for sifting functions. J. Amer. Math. Soc. **4**(1), 25–86 (1991)
25. Bombieri, E: Le Grand Crible dans la Théorie Analytique des Nombres (Seconde ed.) Astérisque. **18** (1987)
26. Vinogradov, AI: The density hypothesis for Dirichlet L-series. Izv. Akad. Nauk SSSR Ser. Mat. (in Russian). **29**, 903–934 (1956)
27. Motohashi, Y, Pintz, J: A smoothed GPY sieve. Bull. Lond. Math. Soc. **40**(2), 298–310 (2008)
28. Bombieri, E, Friedlander, J, Iwaniec, H: Primes in arithmetic progressions to large moduli. Acta Math. **156**(3–4), 203–251 (1986)
29. Vaughan, RC: Sommes trigonométriques sur les nombres premiers. C. R. Acad. Sci. Paris Sér. A. **285**, 981–983 (1977)
30. Fouvry É: Autour du théorème de Bombieri-Vinogradov. Acta Math. **152**(3–4), 219–244 (1984)
31. Fouvry, É, Iwaniec, H: Primes in arithmetic progressions. Acta Arith. **42**(2), 197–218 (1983)

32. Siebert, H: Einige Analoga zum Satz von Siegel-Walfisz. In: Zahlentheorie (Tagung, Math. Forschungsinst., Oberwolfach, 1970), pp. 173–184. Bibliographisches Inst., Mannheim (1971)

33. Iwaniec, H, Kowalski, E: Analytic number theory, Vol. 53. American Mathematical Society Colloquium Publications (2004)

34. Motohashi, Y: An induction principle for the generalization of Bombieri's prime number theorem. Proc. Japan Acad. **52**, 273–275 (1976)

35. Pintz, J: Polignac Numbers, Conjectures of Erdős on Gaps between Primes, Arithmetic Progressions in Primes, and the Bounded Gap Conjecture, preprint.

36. Soundararajan, K: Small gaps between prime numbers: the work of Goldston-Pintz-Yıldırım. Bull. Amer. Math. Soc. (N.S.) **44**(1), 1–18 (2007)

37. Goldston, D, Yıldırım, C: Higher correlations of divisor sums related to primes. I. Triple correlations. Integers. **3**(A5), 66 (2003)

38. Pintz, J: Are there arbitrarily long arithmetic progressions in the sequence of twin primes? An irregular mind. Szemerédi is 70. Bolyai Soc. Math. Stud. Springer. **21**, 525–559 (2010)

39. Farkas, B, Pintz, J, Révész, S: On the optimal weight function in the Goldston-Pintz-Yıldırım method for finding small gaps between consecutive primes. In: Paul Turán Memorial Volume: Number Theory, Analysis and Combinatorics. de Gruyter, Berlin, (2013)

40. Friedlander, J, Iwaniec, H: The polynomial $X^2 + Y^4$ captures its primes. Ann. Math. **148**(3), 945–1040 (1998)

41. Fouvry, E: Théorḿe de Brun-Titchmarsh: application au théorème de Fermat. Invent. Math. **79**(2), 383–407 (1985)

42. Clark, D, Jarvis, N: Dense admissible sequences. Math. Comp. **70**(236), 1713–1718 (2001)

43. Schönhage, A, Strassen, V: Schnelle Multiplikation großer Zahlen. Comput. Arch. Elektron. Rechnen. **7**, 281–292 (1971)

44. Hensley, D, Richards, I: On the incompatibility of two conjectures concerning primes. In: Analytic Number Theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo. 1972), pp. 123–127. Amer. Math. Soc., Providence, R.I. (1973)

45. Hensley, D, Richards, I: Primes in intervals. Acta Arith. **25**(1973/74), 375–391

46. Richards, I: On the incompatibility of two conjectures concerning primes; a discussion of the use of computers in attacking a theoretical problem. Bull. Amer. Math. Soc. **80**, 419–438 (1974)

47. Schinzel, A: Remarks on the paper "Sur certaines hypothèses concernant les nombres premiers". Acta Arith. **7**, 1–8 (1961/1962)

48. Gordon, D, Rodemich, G: Dense admissible sets. In: Algorithmic Number Theory (Portland, OR, 1998), Lecture Notes in Comput. Sci, pp. 216–225. Springer, Berlin, (1998)

49. Heath-Brown, DR: Prime numbers in short intervals and a generalized Vaughan identity. Canad. J. Math. **34**(6), 1365–1377 (1982)

50. Montgomery, HL: Topics in Multiplicative Number Theory, volume 227. Lecture Notes in Math. Springer, New York (1971)