EXPLICIT CALCULATIONS FOR SONO’S MULTIDIMENSIONAL SIEVE OF $E_2$-NUMBERS

DANIEL A. GOLDSTON, APOORVA PANIDAPU, AND JORDAN SCHETTLER

Abstract. We derive explicit formulas for integrals of certain symmetric polynomials used in Keiju Sono’s multidimensional sieve of $E_2$-numbers, i.e., integers which are products of two distinct primes. We use these computations to produce the currently best-known bounds for gaps between multiple $E_2$-numbers. For example, we show there are infinitely many occurrences of four $E_2$-numbers within a gap size of 94 unconditionally and within a gap size of 32 assuming the Elliott-Halberstam conjecture for primes and sifted $E_2$-numbers.

1. Introduction

Let $p_n$ denote the $n$th prime. The prime number theorem implies the average of the gap $p_{n+1} - p_n$ between successive primes is asymptotic to $\log p_n$. On the other hand, the still unproven twin prime conjecture asserts that the gap $p_{n+1} - p_n = 2$ occurs infinitely often. The groundbreaking development in [GPY09] by Goldston, Pintz, and Yıldırım of the GPY sieve, a variant of the Selberg sieve, established that there are always gaps much smaller than average:

$$\liminf_{n \to \infty} \frac{p_{n+1} - p_n}{\log p_n} = 0.$$ 

In fact, assuming the Elliott-Halberstam conjecture for primes (denoted BV[1, P] as below), they also showed that

$$\liminf_{n \to \infty} (p_{n+1} - p_n) \leq 16.$$ 

Yitang Zhang was the first to establish an unconditional result on finite gaps between pairs of primes. By using ideas from the GPY sieve along with proving a sufficiently weakened version of BV[1, P], he showed in [Zha14] that

$$\liminf_{n \to \infty} (p_{n+1} - p_n) \leq 70000000.$$ 

In general, for any positive integer $\nu$ we define

$$H_\nu := \liminf_{n \to \infty} (p_{n+\nu} - p_n).$$ 

Shortly after Zhang’s result, Maynard (in [May15]) and Tao (unpublished work) developed a multidimensional GPY sieve and showed that $H_\nu$ was finite for all $\nu$ with $H_1 \leq 600$ unconditionally. The Polymath8 Project further refined these techniques to obtain the currently best-known unconditional estimate $H_1 \leq 246$ in [Pol14].

The GPY sieve methods can establish even stronger results for gaps between $E_2$-numbers (products of two distinct primes), which are more plentiful than prime numbers. Let $q_n$ denote the $n$th $E_2$-number, i.e.,

$$q_1 = 6, q_2 = 10, q_3 = 14, q_4 = 15, \ldots,$$
and let $G_\nu$ denote the gap for $\nu + 1$ successive $E_2$-numbers: 

$$G_\nu := \liminf_{n \to \infty} (q_{n+\nu} - q_n).$$

We say that an $E_2$-number is “sifted” when it takes the form $p_{k_1} p_{k_2}$ where the prime factors satisfy the inequalities $N^\eta < p_{k_1} \leq N^{1/2} < p_{k_2}$ for arbitrary constants $\eta$ and $N$, with $0 < \eta \leq 1/4$ and $N$ a large positive integer. Now let $\tilde{q}_n$ denote the $n$th sifted $E_2$-number and let $\tilde{G}_\nu$ denote the gap for $\nu + 1$ such sifted $E_2$-numbers as $N$ becomes large:

$$\tilde{G}_\nu = \lim_{N \to \infty} \liminf_{n \to \infty} (\tilde{q}_{n+\nu} - \tilde{q}_n).$$

Sifted $E_2$-numbers have conjectured gaps comparable to prime numbers, while unsifted $E_2$-numbers should have smaller gaps. The Hardy-Littlewood $k$-tuple conjectures imply $\tilde{G}_\nu = H_\nu = H(\nu+1)$ and $G_\nu = G(\nu+1)$ where $H(k)$ (respectively $G(k)$) is the smallest diameter of an admissible (respectively $E_2$-admissible) $k$-tuple. The definitions for admissible and $E_2$-admissible are given in the section below. Goldston, Graham, Pintz, and Yildirim showed in [GGPY09] that we have finite gaps $G_\nu \leq \tilde{G}_\nu \leq \nu e^{\gamma - \gamma (1 + o(1))}$ for all $\nu$ where $\gamma$ is Euler’s constant, and they also established the explicit bound $\tilde{G}_1 \leq 6$.

Following Maynard’s ideas, Sono developed a multidimensional sieve for sifted $E_2$-numbers in [Son20] and showed that $\tilde{G}_2 \leq 12$ under the assumption of the Elliott-Halberstam conjecture for primes $\text{BV}[1, \mathcal{P}]$ and sifted $E_2$-numbers $\text{BV}[1, \tilde{E}_2]$. The present authors in [GPS22] were able to obtain this result of Sono’s using only the one-dimensional GGPY sieve and further proved that $\tilde{G}_2 \leq 32$ unconditionally. Sono’s multidimensional sieve, however, should be able to produce results that could not be obtained from the one-dimensional sieve. The purpose of the present paper is to prove the following theorem using Sono’s sieve, which gives Table 1 showing the currently best-known gaps for $E_2$-numbers. In particular, the last two columns from the third row and beyond are entirely new, while the first two theoretical columns are provided for comparison.

**Theorem 1.** We have the following table for gaps between $\nu + 1$ successive $E_2$-numbers where the Elliott-Halberstam column indicates that we are assuming both $\text{BV}[1, \mathcal{P}]$ and $\text{BV}[1, \tilde{E}_2]$ there.

| $\nu$ | Hardy-Littlewood | Hardy-Littlewood | Elliott-Halberstam | Unconditional |
|-------|------------------|------------------|-------------------|--------------|
| 1     | $G_1 = G(2) = 1$ | $\tilde{G}_1 = H(2) = 2$ | $\tilde{G}_1 \leq H(3) = 6$ | $\tilde{G}_1 \leq H(3) = 6$ |
| 2     | $G_2 = G(3) = 2$ | $\tilde{G}_2 = H(3) = 6$ | $\tilde{G}_2 \leq H(5) = 12$ | $\tilde{G}_2 \leq H(10) = 32$ |
| 3     | $G_3 = G(4) = 4$ | $\tilde{G}_3 = H(4) = 8$ | $\tilde{G}_3 \leq H(10) = 32$ | $\tilde{G}_3 \leq H(23) = 94$ |
| 4     | $G_4 = G(5) = 5$ | $\tilde{G}_4 = H(5) = 12$ | $\tilde{G}_4 \leq H(16) = 60$ | $\tilde{G}_4 \leq H(49) = 240$ |
| 5     | $G_5 = G(6) = 6$ | $\tilde{G}_5 = H(6) = 16$ | $\tilde{G}_5 \leq H(25) = 110$ | $\tilde{G}_5 \leq H(102) = 576$ |
| 6     | $G_6 = G(7) = 8$ | $\tilde{G}_6 = H(7) = 20$ | $\tilde{G}_6 \leq H(37) = 168$ | $\tilde{G}_6 \leq H(225) = 1440$ |

Table 1. Bounds for Gap Sizes of $\nu + 1$ Successive $E_2$-numbers

The proof of Theorem 1 will follow from Sono’s Theorem combined with an intricate analysis of certain integrals $I_k(F)$, $J_k(F)$, $L_k(F)$, $M_k(F)$ done via extensions of the explicit computations for $I_k(F)$, $J_k(F)$ from Section 8 in [May15] to $L_k(F)$, $M_k(F)$. See Theorems 4 and 5 and Lemma 6 below.
2. Theorems of Maynard and Sono

In order to state the results of the multidimensional sieves of Maynard and Sono, we first need to state some definitions and results about the distribution of primes and \( E_2 \)-numbers.

Let \( \pi(x; q) \) denote the number of primes which are relatively prime to \( q \) and less than or equal to \( x \), and let \( \pi(x; q, a) \) denote the number of primes which are congruent to a unit \( a \) modulo \( q \) and less than or equal to \( x \). Then \( \pi(x; q, a) \sim \pi(x; q)/\varphi(q) \) as \( x \to \infty \), and a level of distribution for primes describes the average error in this asymptotic over large intervals. More precisely, if we take \( x \) as \( E \), we first need to state some definitions and results about the distribution of primes and \( E_2 \)-numbers.

Assumption 1 \((BV[\theta, \mathcal{P}])\). For all \( \varepsilon > 0 \) and \( A > 0 \), we have

\[
\sum_{q \leq x^{1-\varepsilon}} \max_{(a, q) = 1} \left| \pi^b(x; q, a) - \frac{\pi^b(x; q)}{\varphi(q)} \right| \ll A \frac{x}{\log^A x}
\]
as \( x \to \infty \).

The Bombieri-Vinogradov theorem \([\text{Bom87, Vin65}]\) states that \( BV[1/2, \mathcal{P}] \) holds, while the unproven Elliott-Halberstam conjecture for primes \([\text{EH70}]\) states that \( BV[1, \mathcal{P}] \) holds. To state the analogs for \( E_2 \)-numbers, we define \( \tilde{\pi}_2(x; q) \) to be the number of sifted \( E_2 \)-numbers which are relatively prime to \( q \) and less than or to equal to \( x \), and take \( \tilde{\pi}_2(x; q, a) \) to be the number of sifted \( E_2 \)-numbers which are congruent to a unit \( a \) modulo \( q \) and less than or to equal to \( x \). As above, we take

\[
\tilde{\pi}_2^b(x; q) = \tilde{\pi}_2(2x; q) - \tilde{\pi}_2(x; q)
\]
and

\[
\tilde{\pi}_2^b(x; q, a) = \tilde{\pi}_2(2x; q, a) - \tilde{\pi}_2(x; q, a).
\]
Then we say the sifted \( E_2 \)-numbers have a level of distribution \( \theta \in (0, 1] \) when the following statement holds.

Assumption 2 \((BV[\theta, \tilde{E}_2])\). For all \( \varepsilon > 0 \) and \( A > 0 \), we have

\[
\sum_{q \leq x^{1-\varepsilon}} \max_{(a, q) = 1} \left| \tilde{\pi}_2^b(x; q, a) - \frac{\tilde{\pi}_2^b(x; q)}{\varphi(q)} \right| \ll A \frac{x}{\log^A x}
\]
as \( x \to \infty \).

Motohoshi in \([\text{Mot76}]\) proved that \( BV[1/2, \tilde{E}_2] \) holds, and we expect that the Elliott-Halberstam conjecture for sifted \( E_2 \)-numbers \( BV[1, \tilde{E}_2] \) holds here as well. Note that our definitions for level of distributions \( \theta \) follow Sono but vary from other authors. In particular, we can always take \( \theta = 1/2 \) unconditionally and \( \theta = 1 \) is equivalent to assuming the corresponding Elliott-Halberstam conjecture.

Let \( \mathcal{H} = \{h_1, h_2, \ldots, h_k\} \) denote a set of \( k \) distinct nonnegative integers, and let \( n \) denote a positive integer variable. In order for all \( n + h_i \) to be prime simultaneously infinitely often there cannot be a prime \( p \) such that the \( h_i \) cover all congruence classes modulo \( p \) since then we would always have at least one of the \( n + h_i \) as a multiple of \( p \). We say \( \mathcal{H} \) is admissible if for all primes \( p \) there is a positive integer \( n_0 \) (depending on \( p \)) such that \( p \) does not divide any \( n_0 + h_i \). Similarly, we say \( \mathcal{H} \) is \( E_2 \)-admissible if for all pairs of primes \( p_{k_1}, p_{k_2} \) there is a positive integer \( n_0 \) (depending on \( p_{k_1}, p_{k_2} \)) such that \( p_{k_1}p_{k_2} \) does not divide any \( n_0 + h_i \). The Hardy-Littlewood \( k \)-tuple conjectures imply that if \( \mathcal{H} \) is admissible (respectively
$E_2$-admissible), then there are infinitely many $n$ such that $n + h_1$, $n + h_2$, $n + h_k$ are all primes/sifted $E_2$-numbers (respectively unsifted $E_2$-numbers). We call $h_k - h_1$ the diameter of $H$, and the smallest diameter for an admissible (respectively $E_2$-admissible) set of size $k$ is denoted $H(k)$ (respectively $G(k)$). For example, $H = \{0, 2, 6, 8\}$ is admissible with minimal diameter $H(4) = 8$, so we expect there are infinitely many positive integers $n$ such that $n$, $n + 2$, $n + 6$, $n + 8$ are all prime. Thus we should have $H_3 = H(3 + 1) = 8$ according to Hardy-Littlewood, but the best-known unconditional result is that $H_3 \leq 24797814 [Pol14]$. In general, sieve methods allow us to conclude that for sufficiently large $k$ at least $\nu + 1$ of the $n + h_i$ are all primes/sifted $E_2$-numbers infinitely often for $H$ admissible. The parity problem inherent in most sieve methods, however, restricts how small $k$ can be for a given $\nu$. In particular, we must have $k \geq 2\nu + 1$ in these cases, and frequently the optimal $k = 2\nu + 1$ is not attainable. For example, when $\nu = 1$, we can get $\nu + 1 = 2$ of the $n + h_i$ as sifted $E_2$-numbers infinitely often for admissible sets $H$ of size $k = 2\nu + 1 = 3$, but for primes we can take $k = 3$ only by assuming a generalized Elliott-Halberstam conjecture. Likewise, when $\nu = 2$, we can take $k = 2\nu + 1 = 5$ for sifted $E_2$-numbers only by assuming the Elliott-Halberstam conjecture for primes and $E_2$-numbers, and we can take $k = 10$ with no such assumptions [Son20] [GPS22]. Hence we get $G_2 \leq H(5) = 12$ conditionally, and we can currently conclude $G_2 \leq H(10) = 32$ unconditionally.

Now we will outline the basic ideas in the sieve methods of Maynard-Tao and Sono. Let $\chi_S(n)$ denote the characteristic function of some set $S$ of interest like the set of primes $P$ or sifted $E_2$-numbers $\tilde{E}_2$. Fix a positive integer $\nu$ and an admissible set $H = \{h_1, h_2, \ldots, h_k\}$ as above. Consider the sum

$$S(x) = \sum_{x < n \leq 2x} \left( \sum_{i=1}^{k} \chi_S(n + h_i) - \nu \right) w_n$$

where $w_n$ are nonnegative weights (depending on $H$ and $n$). Note that if $S(x) > 0$, then we must have at least one $n$ with $x < n \leq 2x$ such that $\nu + 1$ or more of the $n + h_i$ are in $S$. The standard Selberg weights are of the form

$$w_n = \left( \sum_{d | \prod_{i=1}^{k} (n + h_i)} \lambda_d \right)^2,$$

while the Maynard-Tao sieve weights give more flexibility by allowing weights to depend on divisors of individual factors $n + h_i$:

$$w_n = \left( \sum_{d_i | n + h_i} \lambda_{d_1, d_2, \ldots, d_k} \right)^2.$$

Here the $\lambda_{d_1, d_2, \ldots, d_k}$ are given in terms of a smooth function $F: \mathbb{R}^k \to \mathbb{R}$ supported on the region $R_k$ of points $(x_1, x_2, \ldots, x_k)$ where $x_i \geq 0$ for all $i$ and $x_1 + x_2 + \cdots + x_k \leq 1$. The sum $S(x)$ is decomposed as $S = S_2 - \nu S_1$ where

$$S_1(x) = \sum_{x < n \leq 2x} w_n.$$
and

\[ S_2(x) = \sum_{x < n \leq 2x} \left( \sum_{i=1}^{k} \chi_S(n + h_i) \right) w_n. \]

Then one estimates \( S_1, S_2 \) by certain iterated integrals of \( F \) in order to conclude that \( S_2(x)/S_1(x) > \nu \) for all sufficiently large \( x \). Here we consider a symmetric polynomial \( P \) in \( k \) variables and take \( F \) to have the shape

\[ F(x_1, x_2, \ldots, x_k) = \begin{cases} \frac{P(x_1, x_2, \ldots, x_k)}{(x_1, x_2, \ldots, x_k) \in R_k} & \text{if } \theta > 0 \\ 0 & \text{otherwise.} \end{cases} \]

Now define quantities \( I_k(F), J_k(F) \):

\[ I_k(F) := \int_0^1 \cdots \int_0^1 F^2 \, dx_1 \cdots dx_k, \]
\[ J_k(F) := \int_0^1 \cdots \int_0^1 \left( \int_0^1 F \, dx_1 \right)^2 \, dx_2 \cdots dx_k. \]

**Theorem 2** (Maynard, [May15]). Let \( \mathcal{H} = \{h_1, h_2, \ldots, h_k\} \) be an admissible set, and suppose \( \theta \) is a level of distribution for the primes. If \( \nu > 0 \) is an integer such that

\[ \frac{\theta k J_k(F)}{I_k(F)} > \nu \]

for some \( F \) as above, then there are infinitely many integers \( n \) such that at least \( \nu + 1 \) of the \( n + h_i \) are prime.

To state the analogous result proved by Sono for sifted \( E_2 \)-numbers, we need to define two more integral quantities \( L_k(F), M_k(F) \) which depend on \( \theta \) and the parameter \( \eta \) mentioned above:

\[ L_k(F) := \int_0^{\theta/2} \frac{\theta/2 - \xi}{\xi(1-\xi)} \int_0^1 \cdots \int_0^1 \left( \int_0^1 F_{\xi} \, dx_1 \right) \left( \int_0^1 F \, dx_1 \right) \, dx_2 \cdots dx_k \, d\xi \]
\[ M_k(F) := \int_0^{\theta/2} \frac{(\theta/2 - \xi)^2}{\xi(1-\xi)} \int_0^1 \cdots \int_0^1 \left( \int_0^1 F_{\xi} \, dx_1 \right)^2 \, dx_2 \cdots dx_k \, d\xi \]

where

\[ F_{\xi}(x_1, x_2, \ldots, x_k) = F \left( \frac{2\xi}{\theta} + \left( 1 - \frac{2\xi}{\theta} \right) x_1, x_2, x_3, \ldots, x_k \right). \]

**Theorem 3** (Sono, [Son20]). Let \( \mathcal{H} = \{h_1, h_2, \ldots, h_k\} \) be an admissible set, and suppose \( \theta \) is a common level of distribution for the primes and sifted \( E_2 \)-numbers. If \( \nu > 0 \) is an integer such that

\[ \lim_{\eta \to 0^+} \frac{-\theta k L_k(F) + \frac{\theta^2}{2} \log \left( \frac{1-\eta}{\eta} \right) k J_k(F) + k M_k(F)}{\frac{\theta}{2} I_k(F)} > \nu \]

for some \( F \) as above, then there are infinitely many integers \( n \) such that at least \( \nu + 1 \) of the \( n + h_i \) are sifted \( E_2 \)-numbers.
3. Explicit Computations

Let \( P_i = x_i^1 + x_i^2 + \cdots + x_i^k \) denote the \( i \)th power sum symmetric polynomial in \( k \) variables. Maynard computed \( I_k(F) \) and \( J_k(F) \) explicitly when \( P \) has a special form.

**Assumption 3.** Suppose the symmetric polynomial \( P \) is of the form
\[
P = \sum_{i=0}^{n} a_i (1 - P_1)^b P_2^c i
\]
where the \( a_i \) are real constants and \( b_i, c_i \) are nonnegative integers.

Central to these calculations are the beta function identity \( \int_0^1 (1 - u)^b u^c \, du = b!/(b + c + 1)! \) and certain polynomials \( Q_c(x) \) of degree \( c \) with \( Q_0(x) = 1 \) and for \( c > 1 \) are given by
\[
Q_c(x) = c! \sum_{r=1}^{c} \binom{x}{r} \prod_{i=1}^{r} \frac{(2c_i)!}{c_i!}.
\]
For example, \( Q_1(x) = 2x \), \( Q_2(x) = 20x + 4x^2 \), \( Q_3(x) = 592x + 120x^2 + 8x^3 \), \( Q_4(x) = 33888x + 5936x^2 + 480x^3 + 16x^4 \).

**Theorem 4** (Maynard, [May15]). Under Assumption 3, we have
\[
I_k(F) = \sum_{0 \leq i,j \leq n} a_i a_j \frac{(b_i + b_j)! Q_{c_i+c_j}(k)}{(k + b_i + b_j + 2c_i + 2c_j)!}
\]
and
\[
J_k(F) = \sum_{0 \leq i,j \leq n} a_i a_j \sum_{c_1=0}^{c_i} \sum_{c_2=0}^{c_j} \frac{(c_i)}{c_i!} \frac{(c_j)}{c_j!} \frac{\gamma_{i,j,c_i,c_j} Q_{c_i+c_j}(k-1)}{(k + b_i + b_j + 2c_i + 2c_j + 1)!}
\]
where
\[
\gamma_{i,j,c_i,c_j} = \frac{b_i!b_j!(2c_i - 2c'_i)(2c_j - 2c'_j)(b_i + b_j + 2c_i + 2c_j - 2c'_i - 2c'_j + 2)!}{(b_i + 2c_i - 2c'_i + 1)!(b_j + 2c_j - 2c'_j + 1)!}.
\]

Sono did not derive explicit calculations for \( L_k(F) \) and \( M_k(F) \) as Maynard did for \( I_k(F) \) and \( J_k(F) \), but we will do so here. First, we need a lemma.

**Lemma 5.** We have
\[
\int_{\mathcal{R}_k} \cdots \int (1 - P_1)^b P_2^c \, dx_1 \cdots \, dx_k = \frac{b! Q_c(k)}{(k + 2c + b)!}
\]
and
\[
\int_{\mathcal{R}_k} \cdots \int P_1^b P_2^c \, dx_1 \cdots \, dx_k = \frac{Q_c(k)}{(k + 2c - 1)! (k + 2c + b)}.
\]

**Proof.** Equation (2) was shown in [May15], and we will use this to prove Equation (3). We start with the binomial theorem and then simplify by applying the identity \( \sum_{j=m}^{n} (-1)^j \binom{n}{j} = (-1)^m \binom{n-1}{n-m} \) for \( 0 < m \leq n \) to get
\[
\int_{\mathcal{R}_k} \cdots \int (1 - P_1)^b P_2^c \, dx_1 \cdots \, dx_k
\]
Theorem 6. Under Assumption 3, we have

\[ J(\tilde{\theta}) := \lim_{F_{\tilde{\theta}} \rightarrow \tilde{\theta}} \left( F_{\tilde{\theta}} - \frac{\theta}{2} \log \left( \frac{1 - \eta}{\eta(\frac{3}{2} - 1)} \right) \right) J_k(F) \]

\[ = \lim_{\eta \rightarrow 0^+} \left( -\theta k L_k(F) + \frac{\theta^2}{4} \log \left( \frac{1 - \eta}{\eta(\frac{3}{2} - 1)} \right) \right) J_k(F) + \frac{\theta^2}{4} \log \left( \frac{2}{\theta} - 1 \right) k J_k(F) + k \tilde{M}_k(F). \]

Definition 1. Now define limits

\[ \tilde{L}_k(F) = \lim_{\eta \rightarrow 0^+} \left( L_k(F) - \frac{\theta}{2} \log \left( \frac{1 - \eta}{\eta(\frac{3}{2} - 1)} \right) J_k(F) \right), \]

and

\[ \tilde{M}_k(F) = \lim_{\eta \rightarrow 0^+} \left( M_k(F) - \frac{\theta^2}{4} \log \left( \frac{1 - \eta}{\eta(\frac{3}{2} - 1)} \right) J_k(F) \right). \]

Then the limit of the numerator on the left-hand side of the inequality in Equation \( \Box \) becomes

\[ \tilde{J}_k(F) := \lim_{\eta \rightarrow 0^+} \left( -\theta k L_k(F) + \frac{\theta^2}{4} \log \left( \frac{1 - \eta}{\eta(\frac{3}{2} - 1)} \right) k J_k(F) + k \tilde{M}_k(F) \right) \]

\[ = -\theta k \tilde{L}_k(F) + \frac{\theta^2}{4} \log \left( \frac{2}{\theta} - 1 \right) k J_k(F) + k \tilde{M}_k(F). \]

Theorem 6. Under Assumption 3, we have

\[ \tilde{L}_k(F) = \sum_{0 \leq i, j \leq n} a_i a_j \sum_{c_i = 0}^{c_i} \sum_{c_j = 0}^{c_j} \left( c_i \choose c_i' \right) \left( c_j \choose c_j' \right) Q_{c_i + c_j'}(k - 1) \]

\[ \cdot \left( \sum_{d_i = 0}^{\tilde{c}_i} \frac{(-1)^{d_i} \tilde{c}_i!}{(d_i)!} \delta_{i, j, c_i', c_j'} \lambda k - 1 + d_i + 2c_i' + 2c_j' \right) \]

\[ - \sum_{b_i = 0}^{b} \sum_{c_i = 0}^{c} \left( -1 \right)^{c_i + k_1} \left( \tilde{c}_i \choose b_i \right) \frac{(c_i)_{b_i}}{(k + 2c_i' + 2c_j' - 2)!} \left( k - 1 + 2c_i' + 2c_j' + 2c_i + 2c_j \right) \]

and

\[ \tilde{M}_k(F) = \frac{\theta}{2} \sum_{0 \leq i, j \leq n} a_i a_j \sum_{c_i = 0}^{c_i} \sum_{c_j = 0}^{c_j} \left( c_i \choose c_i' \right) \left( c_j \choose c_j' \right) Q_{c_i + c_j'}(k - 1) \]

\[ \cdot \left( \sum_{d_i = 0}^{\tilde{c}_i} \frac{(-1)^{d_i} \tilde{c}_i!}{(d_i)!} \delta_{i, j, c_i', c_j'} \lambda k - 1 + d_i + 2c_i' + 2c_j' \right) \]

\[ - \sum_{b_i = 0}^{b} \sum_{c_i = 0}^{c} \left( -1 \right)^{c_i + k_1} \left( \tilde{c}_i \choose b_i \right) \frac{(c_i)_{b_i}}{(k + 2c_i' + 2c_j' - 2)!} \left( k - 1 + 2c_i' + 2c_j' + 2c_i + 2c_j \right). \]
\[ -2 \sum_{b_i=0}^{b_i} \sum_{e_i=0}^{e_i} \frac{(-1)^{e_1+b'_1} \left( \int \right) \left( \int \right)}{(k+2c'_1+2c'_2-2)(k-1+2c'_1+2c'_2+e_1)} \\
+ \sum_{b_1=0}^{b_1} \sum_{b_2=0}^{b_2} \sum_{f_1=0}^{f_1} \frac{(-1)^{f_1+b'_1+b'_2} \left( \int \right) \left( \int \right)}{(k+2c'_1+2c'_2-2)(k-1+2c'_1+2c'_2+f_1)} \]

where \( \tilde{d} = b_i + b_j + 2c_i + 2c_j - 2c'_1 - 2c'_2 + 2 \), \( \tilde{e} = b_i - b'_1 + b_j + 2c_j - 2c'_2 + 1 \), \( \tilde{f} = b_i + b_j - b'_1 - b'_2 \),

\[ \delta_{i,j,c'_1,c'_2} = \frac{b_j!(2c_i - 2c'_1)}{(b_1 + 2c_i - 2c'_1 + 1)(b_j + 2c_j - 2c'_2 + 1)!} \]

\[ \varepsilon_{i,j,c'_1,c'_2,b'_1,b'_2} = \frac{b_j!(2c_j - 2c'_2)}{(b_j + 2c_j - 2c'_2 + 1)(b'_1 + 2c_i - 2c'_1 + 1)!} \]

\[ \zeta_{i,j,c'_1,c'_2,b'_1,b'_2} = \frac{b'_1 + b'_2 + 2c_i - 2c'_1 + 1}{b'_1 + 2c_i - 2c'_1 + 1} \]

\[ \lambda_n = \frac{\theta}{2} \left( \frac{2F_1(1,1; n+2; \theta/2)}{n+1} - H_n + \log \left( 1 - \frac{\theta}{2} \right) \right) \]

and

\[ \mu_{m,n} = \frac{\theta}{2} \left( \frac{2F_1(1,m; m+n+1; \theta/2)}{m(m+n)} \right) \]

Here \( 2F_1(a,b;c;z) \) denotes the standard hypergeometric function and \( H_n \) denotes the \( n \)th harmonic number.

Proof. First, we make a change of variables \( x_0 = 2\xi/\theta \) and \( \tilde{x}_1 = x_0 + (1 - x_0)x_1 \) for the inner integral of \( F_k \). Then \( L_k(F) \) becomes

\[ \int_{2\eta/\theta}^{1} \frac{1}{x_0(x_0 + (1 - x_0)x_1)} \int \cdots \int \left( \int_{x_0}^{1-P'_1} P \, d\tilde{x}_1 \right) \left( \int_{x_0}^{1-P'_1} P \, dx_1 \right) dx_2 \cdots dx_k \]

where \( P'_1 = x_2 + x_3 + \cdots + x_k \) and \( R'_k \) is the region of points \( (x_2, x_3, \ldots, x_k) \) such that \( x_i \geq 0 \) for all \( i \) and \( x_0 + P'_1 \leq 1 \). Likewise \( M_k(F) \) becomes

\[ \frac{\theta}{2} \int_{2\eta/\theta}^{1} \frac{1}{x_0(x_0 + (1 - x_0)x_1)} \int \cdots \int \left( \int_{x_0}^{1-P'_1} P \, d\tilde{x}_1 \right)^2 dx_2 \cdots dx_k \]

To evaluate the inner integrals we use the binomial theorem on \( P \), another substitution \( u = \tilde{x}_1/(1 - P'_1) \), and the beta function identity to get

\[ \int_{x_0}^{1-P'_1} P \, d\tilde{x}_1 = \sum_{i=0}^{n} a_i \int_{x_0}^{1-P'_1} (1 - P_1)^{b_i} (P_2)^{c_i} d\tilde{x}_1 \]

\[ = \sum_{i=0}^{n} a_i \sum_{c'_i=0}^{c_i} \left( \frac{c_i}{c'_i} \right) (P_2)^{c'_i} \int_{x_0}^{1-P'_1} (1 - P_1)^{b_i} \tilde{x}_1^{2c_i - 2c'_i} d\tilde{x}_1 \]

\[ = \sum_{i=0}^{n} a_i \sum_{c'_i=0}^{c_i} \left( \frac{c_i}{c'_i} \right) (1 - P'_1)^{b_i} (P_2)^{c'_i} \int_{x_0/(1-P'_1)}^{1} (1 - u)^{b_i} u^{2c_i - 2c'_i} du \]

\[ = \sum_{i=0}^{n} a_i \sum_{c'_i=0}^{c_i} \left( \frac{c_i}{c'_i} \right) \left( \frac{b_i!(2c_i - 2c'_i)!}{(b_i + 2c_i - 2c'_1 + 1)!} \right) (1 - P'_1)^{b_i} (P_2)^{c'_i} \]
\[ - \sum_{b'_i = 0}^{b_i} \frac{(-1)^{b'_i} \binom{b'_i}{b'_i} x_0^{b'_i + 2 c_i - 2 c'_i + 1}}{b'_i + 2 c_i - 2 c'_i + 1} \left(1 - P'_1 \right)^{b'_i} \binom{P'_1}{P'_2}^{c_i} \]

This argument also implies

\[ \int_{0}^{1-P'_1} P \, dx_1 = \sum_{j=0}^{n} a_j \sum_{c'_j = 0}^{c_j} \binom{c_j}{c'_j} b_j! (2 c_j - 2 c'_j)! \left(1 - P'_1 \right)^{b_j} \binom{P'_1}{P'_2}^{c'_j} \]

so

\[ \left( \int_{x_0}^{1-P'_1} P \, dx \right) \left( \int_{0}^{1-P'_1} P \, dx_1 \right) = \sum_{0 \leq i, j \leq n} a_i a_j \sum_{c'_i = 0}^{c_i} \sum_{c'_j = 0}^{c_j} \binom{c_i}{c'_i} \binom{c_j}{c'_j} \left(1 - P'_1 \right)^{b_i} \binom{P'_1}{P'_2}^{c'_i} \binom{P'_1}{P'_2}^{c'_j} \]

and

\[ \left( \int_{x_0}^{1-P'_1} P \, dx \right)^2 = \sum_{0 \leq i, j \leq n} a_i a_j \sum_{c'_i = 0}^{c_i} \sum_{c'_j = 0}^{c_j} \binom{c_i}{c'_i} \binom{c_j}{c'_j} \left(1 - P'_1 \right)^{b_i} \binom{P'_1}{P'_2}^{c'_i} \binom{P'_1}{P'_2}^{c'_j} \]

Next we use homogeneity and then apply Lemma 5 to obtain the intermediate result

\[ \int \cdots \int (1 - P'_1)^{\tilde{b}} \binom{P'_1}{P'_2}^{c} \, dx_2 \cdots \, dx_k \]

\[ = \sum_{\tilde{b}' = 0}^{\tilde{b}} (\tilde{b}')! \int \cdots \int (P'_1)^{\tilde{b}'} \binom{P'_1}{P'_2}^{c} \, dx_2 \cdots \, dx_k \]

\[ = \sum_{\tilde{b}' = 0}^{\tilde{b}} (\tilde{b}')! (1 - x_0)^{k-1+\tilde{b}'+2c} \int \cdots \int (P'_1)^{\tilde{b}'} \binom{P'_1}{P'_2}^{c} \, dx_2 \cdots \, dx_k \]

\[ = \sum_{\tilde{b}' = 0}^{\tilde{b}} (\tilde{b}')! Q_c(k-1) \frac{x_0^{k-1+\tilde{b}'+2c}}{(k + 2c - 2)! (k - 1 + 2c + \tilde{b}')}. \]
Combining the previous three equations yields

\[
L_k(F) = \sum_{0 \leq i,j \leq n} a_i a_j \sum_{c_1=0}^{c_i} \sum_{c_2=0}^{c_j} \binom{c_i}{c_1} \binom{c_j}{c_2} Q_{c_1+c_2}(k-1) \\
\cdot \left( \sum_{d_1=0}^{ \tilde{d}} \frac{(-1)^{d_1} (\frac{\tilde{d}}{d_1}) \delta_{i,j,c_1,c_2}}{2^{n/\theta}} \int_{2^{n/\theta}}^1 \frac{(1 - x_0)^{k-1+d_1+2c_1+2c_2}}{x_0(\frac{\theta}{\theta} - x_0)} dx_0 \right) \\
- \sum_{b_i'=0}^{b_i} \sum_{c_1} \frac{(-1)^{c_1+b_i'} (\tilde{e}_{i,j}) (\tilde{e}_{b_i'}) \varepsilon_{b_i,c_1}}{(k + 2c_1 + 2c_2 - 2)! (k-1 + 2c_1 + 2c_2 + d_1)} \\
\int_{2^{n/\theta}}^1 \frac{(1 - x_0)^{k-1+c_1+2c_2}}{x_0(\frac{\theta}{\theta} - x_0)} dx_0 \\
\left( \sum_{d_1=0}^{ \tilde{d}} \frac{(-1)^{d_1} (\frac{\tilde{d}}{d_1}) \delta_{i,j,c_1,c_2}}{2^{n/\theta}} \int_{2^{n/\theta}}^1 \frac{(1 - x_0)^{k-1+d_1+2c_1+2c_2}}{x_0(\frac{\theta}{\theta} - x_0)} dx_0 \right) \\
- 2 \sum_{b_i'=0}^{b_i} \sum_{c_1} \frac{(-1)^{c_1+b_i'} (\tilde{e}_{i,j}) (\tilde{e}_{b_i'}) \varepsilon_{b_i,c_1}}{(k + 2c_1 + 2c_2 - 2)! (k-1 + 2c_1 + 2c_2 + d_1)} \\
\int_{2^{n/\theta}}^1 \frac{(1 - x_0)^{k-1+c_1+2c_2}}{x_0(\frac{\theta}{\theta} - x_0)} dx_0 \\
+ \sum_{b_i'=0}^{b_i} \sum_{b_j'=0}^{b_j} \sum_{f_1=0}^{f_1} \frac{(-1)^{f_1+b_i'+b_j'} (\tilde{f}_i) (\tilde{f}_j)}{(\tilde{f}_i) (\tilde{f}_j)} \varepsilon_{b_i,b_j} \\
\int_{2^{n/\theta}}^1 \frac{(1 - x_0)^{k-1+f_1+2c_1+2c_2}}{x_0(\frac{\theta}{\theta} - x_0)} dx_0 \right).
\]

and

\[
M_k(F) = \frac{\theta}{2} \sum_{0 \leq i,j \leq n} a_i a_j \sum_{c_1=0}^{c_i} \sum_{c_2=0}^{c_j} \binom{c_i}{c_1} \binom{c_j}{c_2} Q_{c_1+c_2}(k-1) \\
\cdot \left( \sum_{d_1=0}^{ \tilde{d}} \frac{(-1)^{d_1} (\frac{\tilde{d}}{d_1}) \delta_{i,j,c_1,c_2}}{2^{n/\theta}} \int_{2^{n/\theta}}^1 \frac{(1 - x_0)^{k-1+d_1+2c_1+2c_2}}{x_0(\frac{\theta}{\theta} - x_0)} dx_0 \right) \\
- \sum_{b_i'=0}^{b_i} \sum_{c_1} \frac{(-1)^{c_1+b_i'} (\tilde{e}_{i,j}) (\tilde{e}_{b_i'}) \varepsilon_{b_i,c_1}}{(k + 2c_1 + 2c_2 - 2)! (k-1 + 2c_1 + 2c_2 + d_1)} \\
\int_{2^{n/\theta}}^1 \frac{(1 - x_0)^{k-1+c_1+2c_2}}{x_0(\frac{\theta}{\theta} - x_0)} dx_0 \\
\right).
\]

To evaluate the integrals in \( x_0 \) we make use of the formula

\[
\frac{\text{2F1}(a, c; z)}{b(\frac{a}{b})} = \int_0^1 \frac{t^b(1-t)^{c-b-1}}{t(1-zt)^n} dt
\]

which is valid when \( z \) is non-real or strictly less than 1 and where \( b, c \) are integers with \( c > b > 0 \). Immediately, we get

\[
\lim_{\eta \to 0^+} \int_{2^{n/\theta}}^1 \frac{x_0^n (1 - x_0)^n}{x_0(\frac{\theta}{\theta} - x_0)} dx_0 = \frac{\theta}{2} \frac{\text{2F1}(1, m; m + n + 1; \theta/2)}{m(m+n)} = \mu_{m,n}
\]

when \( m > 0, n \geq 0 \) are integers. Next we evaluate

\[
\int_{2^{n/\theta}}^1 \frac{1}{x_0(\frac{\theta}{\theta} - x_0)} dx_0 = \frac{\theta}{2} \int_{2^{n/\theta}}^1 \frac{1}{x_0} dx_0 + \frac{\theta}{2} \int_{2^{n/\theta}}^1 \frac{1}{\frac{\theta}{\theta} - x_0} dx_0
\]

\[
= -\frac{\theta}{2} \log \frac{\theta}{\theta} + \frac{\theta}{2} \log \left( \frac{2/\theta - 2\eta/\theta}{2/\theta - 1} \right)
\]
A positive definite quadratic forms with assumption 3. As Maynard noted in [May15],

\[ \sim \]

Thus ratio \( a \) also take \( \sim \) variables. In particular, \( \sim \)

Lastly,

\[
\lim_{\eta \to 0^+} \frac{1}{\eta} \int_{2n/\eta}^1 \frac{(1-x_0)^n-1}{x_0(\frac{n}{\theta} - x_0)} \, dx_0 = \frac{\theta}{2} \int_0^1 \frac{(1-x_0)^n-1}{x_0(1-\frac{n}{\theta} x_0)} \, dx_0
\]

\[ = \frac{\theta}{2} \left( \int_0^1 \frac{\theta}{2} x_0(1-x_0)^n \, dx_0 + \int_0^1 \frac{(1-x_0)^n-1}{x_0} \, dx_0 - \int_0^1 \frac{\theta}{2} x_0 \, dx_0 \right) \]

\[ = \frac{\theta}{2} \left( \frac{\theta}{2} F(1,1;n+2;\theta/2) \right) - H_n + \log \left( 1 - \frac{\theta}{2} \right) = \lambda_n. \]

\[ \square \]

4. Proof of Theorem 1

In this last section, we use Theorem 6 to establish Theorem 1, which gives conditional and unconditional bounds on gaps for \( \nu+1 \) successive sifted \( E_2 \)-numbers as seen in Table 1.

Proof of Theorem 7. Assume the symmetric polynomial \( P \) defining \( F \) satisfies Assumption 3. As Maynard noted in [May15], \( I_k(F) \) and \( J_k(F) \) can be regarded as positive definite quadratic forms with \( I_k(F) = a^T A_1 a \) and \( J_k(F) = a^T A_2 a \) where \( A_1, A_2 \) are real symmetric matrices with coefficient vector \( a = (a_0, a_1, \ldots, a_n) \). We also take \( a_m = 0 \) when \( m > n \) for convenience. Note that

\[
0 \leq \left[ \frac{\theta}{2} \int_0^1 F \, dx_1 - \left( \frac{\theta}{2} - \xi \right) \int_0^1 F \, dx_1 \right]^2 \, dx_2 \cdots dx_k \, d\xi
\]

\[ = -\theta L_k(F) + \frac{\theta^2}{4} \log \left( 1 - \frac{\eta}{\eta} \right) J_k(F) + M_k(F) - \frac{\theta^2}{4} \log \left( \frac{2}{\theta} - 1 \right) J_k(F), \]

so

\[ \tilde{J}_k(F) \geq \frac{\theta^2}{4} \log \left( \frac{2}{\theta} - 1 \right) J_k(F) > 0. \]

Thus \( \tilde{J}_k(F) \) can also be regarded as positive definite quadratic forms in the \( a_i \) variables. In particular, \( \tilde{J}_k(F) = a^T A_2 a \) for some real symmetric matrix \( A_2 \). The ratio

\[ R_k(F) := \frac{\tilde{J}_k(F)}{\tilde{J}_k(\tilde{F})} = \frac{2}{\theta} \frac{a^T A_2 a}{a^T A_2 a} \]

is maximized when \( a \) is an eigenvector for the largest eigenvalue of \( A_1^{-1} A_2 \). We use Mathematica to find a numerical approximation to this eigenvector and then rationalize it. The rational vector \( a \) then gives an exact value for \( R_k(F) \). Recall that Sono’s Theorem 3 states that if \( \mathcal{H} \) is admissible and \( R_k(F) > \nu \), then there are infinitely many integers \( n \) such that at least \( \nu + 1 \) of the \( n + h_i \) are sifted \( E_2 \)-numbers, and thus \( G(\nu) \leq H(k) \). We fix the exponents in a symmetric polynomial

\[ P = \sum_{i=0}^n a_i (1 - P_i)^{b_i} P_i^{c_i} \]

as follows:

\[
(b_0, b_1, \ldots) = (0, 1, 2, 0, 3, 1, 4, 2, 0, 5, 3, 1, 6, 4, 2, 0, 7, 5, 3, 1, 8, 6, 4, 2, 0, 9, 7, 5, 3, 1, 10, 8, 6, 4, 2, 0, 11, 9, 7, 5, 3, 1, \ldots)
\]

\[
(c_0, c_1, \ldots) = (0, 0, 0, 1, 0, 1, 2, 0, 1, 2, 0, 1, 2, 0, 1, 2, 3, 0, 1, 2, 3, 0,
\]

\[
\]
We will only derive the entries in the last two columns of Table 1 from the third row and beyond since the first two columns follow directly from the Hardy-Littlewood conjectures and the first two rows in the last two columns were known previously [GGPY09, Son20, GPS22].

We first assume the Elliott-Halberstam conjecture for primes and sifted \( E_2 \)-numbers, i.e., we assume \( \text{BV}[\theta, \mathcal{P}] \) and \( \text{BV}[\theta, \tilde{E}_2] \) with \( \theta = 1 \). When \( k = 10 \) we take

\[
a = \left( \begin{array}{c}
230160446403391652 \\
124720775947120337501 \\
838332478853543802
\end{array} \right)
\]

which gives \( R_{10}(F) = 3.0353 \ldots > 3 \), so \( \tilde{G}_3 \leq H(10) = 32 \). When \( k = 16 \) we take

\[
a = \left( \begin{array}{c}
414930787723574773 \\
10942207624908766578 \\
5637296939984330619
\end{array} \right)
\]

which gives \( R_{16}(F) = 4.000399 \ldots > 4 \), so \( \tilde{G}_4 \leq H(16) = 60 \). When \( k = 25 \) we take

\[
a = \left( \begin{array}{c}
1119449503899613 \\
5344293655026775953 \\
6822903858709892
\end{array} \right)
\]

which gives \( R_{25}(F) = 5.0454 \ldots > 5 \), so \( \tilde{G}_5 \leq H(25) = 110 \). When \( k = 37 \) we take

\[
a = \left( \begin{array}{c}
21148272786657 \\
97959366205232913548 \\
668790979930374 \\
981953056213600467 \\
220482561042490655 \\
87523836340436194254 \\
629705249842009709 \\
107036774210681509668 \\
18172970799879376836 \\
6040191999214638261 \\
91902595669759901
\end{array} \right)
\]

which gives \( R_{37}(F) = 6.01020 \ldots > 6 \), so \( \tilde{G}_6 \leq H(37) = 168 \).

We now prove our unconditional results, so we have \( \text{BV}[\theta, \mathcal{P}] \) and \( \text{BV}[\theta, \tilde{E}_2] \) with \( \theta = 1/2 \). When \( k = 23 \), we take

\[
a = \left( \begin{array}{c}
16068472196488 \\
57228638216292482079 \\
3190274818823462
\end{array} \right)
\]

which gives \( R_{23}(F) = 3.13958 \ldots > 3 \), so \( \tilde{G}_3 \leq H(23) = 106 \). When \( k = 37 \), we take

\[
a = \left( \begin{array}{c}
32660633769801409 \\
132753576873497845795 \\
15378962969533792839
\end{array} \right)
\]

which gives \( R_{37}(F) = 6.01020 \ldots > 6 \), so \( \tilde{G}_6 \leq H(37) = 168 \).
which gives \( R_{23}(F) = 3.0000564254 \ldots > 3 \), so \( \tilde{G}_3 \leq H(23) = 94 \). When \( k = 49 \), we take \( a = \left( \begin{array}{c} 470254915649 \\ 150765533574864473017 \\ 50254705151197 \\ 13182673466835769000 \\ 18464862673265278 \\ 9934907194135507801 \\ 358924377910039906 \\ 26947695639277577771 \\ 3426771083907690033 \\ 194982267383830416128 \\ 6233483905322588837 \\ 113763328332070308845 \end{array} \right) \), which gives \( R_{49}(F) = 4.00096634233 \ldots > 4 \), so \( \tilde{G}_4 \leq H(49) = 240 \). When \( k = 102 \), we take \( a = \left( \begin{array}{c} 882233146 \\ 1139042176880719802317 \\ 4430088519001806748 \\ 226430143350338954387 \\ 38370825062791 \\ 25034766287281234113 \\ 1893082290928153 \\ 7410164071227538557 \\ 9384583779292 \\ 1214563427050811905 \\ 1422926881003116425 \\ 175664537100675106683 \\ 97003526743921763 \\ 8408291389318959293 \end{array} \right) \), which gives \( R_{49}(F) = 4.00096634233 \ldots > 4 \), so \( \tilde{G}_4 \leq H(49) = 240 \).
which gives \( R_{102}(F) = 5.01623513164 \ldots \geq 5 \), so \( \tilde{G}_5 \leq H(102) = 576 \).

When \( k = 225 \) we take \( a = (\ldots) \) which gives \( R_{225}(F) = 6.0098418048817 \ldots \geq 6 \), so \( \tilde{G}_6 \leq H(225) = 1440 \).

\[ \Box \]

**References**

[Bom87] Enrico Bombieri, *Le grand crible dans la théorie analytique des nombres*, Astérisque (1987), no. 18, 103.

[EH70] P. D. T. A. Elliott and H. Halberstam, *A conjecture in prime number theory*, Symposia Mathematica, Vol. IV (INDAM, Rome, 1968/69), Academic Press, London, 1970, pp. 59–72.

[GGPY09] D. A. Goldston, S. W. Graham, J. Pintz, and C. Y. Yıldırım, *Small gaps between products of two primes*, Proceedings of the London Mathematical Society 98 (2009), no. 3, 741–774.

[GPS22] Daniel A. Goldston, Apoorva Panidapu, and Jordan Schettler, *Small gaps between three almost primes and almost prime powers*, Acta Arith. 203 (2022), no. 1, 1–18. MR 4415993

[GPY09] D. A. Goldston, J. Pintz, and C. Y. Yıldırım, *Primes in tuples. I.*, Annals of Mathematics 179 (2009), no. 2, 819–862.

[May15] James Maynard, *Small gaps between primes*, Ann. of Math. (2) 181 (2015), no. 1, 383–413.

[Mot76] Yoichi Motohashi, *An induction principle for the generalization of Bombieri’s prime number theorem*, Proc. Japan Acad. 52 (1976), no. 6, 273–275.

[Pol14] D. H. J. Polymath, *Variants of the Selberg sieve, and bounded intervals containing many primes*, Res. Math. Sci. 1 (2014), Art. 12, 83. MR 3373710
[Son20] Keiju Sono, Small gaps between the set of products of at most two primes, J. Math. Soc. Japan 72 (2020), no. 1, 81–118.

[Vin65] A. I. Vinogradov, The density hypothesis for Dirichlet L-series, Izv. Akad. Nauk SSSR Ser. Mat. 29 (1965), 903–934.

[Zha14] Yitang Zhang, Bounded gaps between primes, Ann. of Math. (2) 179 (2014), no. 3, 1121–1174.