GAPS BETWEEN PRIMES

James Maynard

Abstract

We discuss recent advances on weak forms of the Prime k-tuple Conjecture, and its role in proving new estimates for the existence of small gaps between primes and the existence of large gaps between primes.

1 Introduction

It follows from the Prime Number Theorem that the *average* gap between primes less than X is of size roughly $\log X$ when X is large. We expect, however, that occasionally these gaps are rather smaller than $\log X$, and occasionally rather larger. Specifically, based on random models and numerical evidence, we believe that the largest and smallest gaps are as described in the following two famous conjectures[1].

Conjecture 1 (Twin Prime Conjecture). *There are infinitely many pairs of primes which differ by exactly 2.*

Conjecture 2 (Cramér's Conjecture, weak form). *Let p_n denote the nth prime. Then*

$$\sup_{p_n \leq X} (p_{n+1} - p_n) = (\log X)^{2+o(1)}.$$

Moreover, the Twin Prime Conjecture can be thought of as a special case of the far-reaching Prime k-tuple Conjecture describing more general patterns of many primes.

Conjecture 3 (Prime k-tuple Conjecture). *Let $L_1, \ldots, L_k$ be integral linear functions $L_i(n) = a_i n + b_i$ such that for every prime p there is an integer n_p with $\prod_{i=1}^{k} L_i(n_p)$ coprime to p. Then there are infinitely many integers n such that all of $L_1(n), \ldots, L_k(n)$ are primes.*

[1]The original conjecture of Cramér [1936] was a stronger statement which is no longer fully believed to be true. We expect the weaker version given here to hold.

Unfortunately all these conjectures seem well beyond the current techniques, but we are able to make partial progress by showing that we do have gaps which are smaller or larger than the average gap. A key part of recent progress on results about gaps between primes has been weak versions of Conjecture 3, where one shows there are infinitely many integers n such that *several* (rather *all*) of the linear functions are prime at n.

Theorem 1 (Maynard [2015])**.** *Let $L_1, \ldots, L_k$ be integral linear functions $L_i(n) = a_i n + b_i$ such that for every prime p there is an integer n_p with $\prod_{i=1}^{k} L_i(n_p)$ coprime to p. Then there is a constant $c > 0$ such that there are infinitely many integers n where at least $c \log k$ of $L_1(n), \ldots, L_k(n)$ are primes.*

Variants of Theorem 1 have been important in recent results on small and large gaps between primes, and have proven useful because the method of proof is quite flexible and can generalize to other situations. Specifically, we now know the following approximations to Conjecture 1 and Conjecture 2.

Theorem 2 (Polymath [2014])**.** *There are infinitely many pairs of primes which differ by at most* 246.

Theorem 3 (Ford, Green, Konyagin, Maynard, and Tao [2018])**.** *There exists a constant $c > 0$ such that*

$$\sup_{p_n \leq X} (p_{n+1} - p_n) \geq c \frac{\log X \cdot \log \log X \cdot \log \log \log \log X}{\log \log \log X}.$$

Moreover, in the case of small gaps between primes, we can show the existence of *many* primes in bounded length intervals.

Theorem 4 (Maynard [2016a])**.** *There exists a constant $c > 0$ such that, for all $X \geq 2$ there are at least $c X \exp(-\sqrt{\log X})$ integers $x \in [X, 2X]$ such that*

$$\#\{primes \ in \ [x, x + y]\} \geq c \log y.$$

For example, this shows that there are infinitely many intervals of length $e^{m/c}$ containing m primes, and, for fixed $\epsilon > 0$, infinitely many x such that the interval $[x, x + (\log x)^\epsilon]$ contains $\epsilon c \log \log x$ primes.

2   The GPY sieve method

We aim to prove Theorem 1 by the 'GPY method', which can be interpreted as a first moment method which was introduced to study small gaps between primes by Goldston, Pintz, and Yıldırım [2009]. This takes the following basic steps, for some given set $\{L_1, \ldots, L_k\}$ of distinct functions satisfying the hypotheses of Conjecture 3:

1. We choose a probability measure w supported on integers in $[X, 2X]$.

2. We calculate the expected number of the functions $L_i(n)$ which are prime at n, if n is chosen randomly with probability $w(n)$.

3. If this expectation is at least m, then there must be some $n \in [X, 2X]$ such that at least m of $L_1(n), \ldots, L_k(n)$ are primes.

4. If this holds for all large X, then there are infinitely many such integers n.

This procedure only works if we can find a probability measure w which is suitably concentrated on integers n where many of the $L_i(n)$ are prime, but at the same time is simple enough that we can calculate this expectation unconditionally. We note that by linearity of expectation, it suffices to be able to calculate the probability that any one of the linear functions is prime. However, any slowly changing smooth function w is insufficient since the primes have density 0 in the integers, whereas any choice of w explicitly depending on the joint distribution of prime values of the L_i is likely to be too complicated to handle unconditionally.

Sieve methods are a flexible set of tools, developed over the past century, which provide natural choices for the probability measure w. The situation of simultaneous prime values of $L_1, \ldots, L_k$ is a 'k-dimensional' sieve problem. For such problems when k is large but fixed, the Selberg sieve tends to be the type of sieve which performs best. The standard choice of Selberg sieve weights (which are essentially optimal in closely related situations) are

$$(2\text{-}1) \qquad w(n) \propto \Big(\sum_{\substack{d \mid \prod_{i=1}^{k} L_i(n) \\ d < R}} \mu(d) \Big(\log \frac{R}{d} \Big)^k \Big)^2,$$

where R is a parameter which controls the complexity of the sieve weights, and w is normalized to sum to 1 on $[X, 2X]$.

To calculate the probability that $L_j(n)$ is prime with this choice of $w(n)$, we wish to estimate the sum of $w(n)$ over $n \in [X, 2X]$ such that $L_j(n)$ is prime. To do this we typically expand the divisor sum in the definition (2-1) and swap the order of summation. This reduces the problem to estimating the number of prime values of $L_j(n)$ for $n \in [X, 2X]$ in many different arithmetic progressions with moduli of size about R^2. The Elliott-Halberstam conjecture Elliott and Halberstam [1970] asserts that we should be able to do this when $R^2 < X^{1-\epsilon}$, but unconditionally we only know how to do this

when $R^2 < X^{1/2-\epsilon}$, using the Bombieri-Vinogradov Theorem Bombieri [1965] and Vinogradov [1965]. After some computation one finds that, provided we do have suitable estimates for primes in arithmetic progressions, the choice (2-1) gives

$$(2\text{-}2) \qquad \mathbb{E} \, \#\{1 \le i \le k : L_i(n) \text{ prime}\} = \left(2 - \frac{2}{k+1} + o(1)\right) \frac{\log R}{\log X}.$$

In particular, this is less than 1 for all large X, even if we optimistically assume the Elliott-Halberstam conjecture and take $R \approx X^{1/2-\epsilon}$, and so it appears that we will not be able to conclude anything about primes in this manner.

The groundbreaking work of Goldston, Pintz, and Yıldırım [2009] showed that a variant of this choice of weight actually performs much better. They considered

$$(2\text{-}3) \qquad w(n) \propto \left(\sum_{\substack{d | \prod_{i=1}^k L_i(n) \\ d < R}} \mu(d) \left(\log \frac{R}{d}\right)^{k+\ell} \right)^2,$$

where ℓ is an additional parameter to be optimized over. With the choice (2-3), one finds that provided we have suitable estimates for primes in arithmetic progressions, we obtain

$$\mathbb{E} \, \#\{1 \le i \le k : L_i(n) \text{ prime}\} = \left(4 + O\left(\frac{1}{\ell}\right) + O\left(\frac{\ell}{k}\right)\right) \frac{\log R}{\log X}.$$

This improves upon (2-2) by a factor of about 2 when k is large and $\ell \approx k^{1/2}$. This falls just short of proving that two of the linear functions are simultaneously prime when $R = X^{1/4-\epsilon}$ as allowed by the Bombieri-Vinogradov theorem, but any small improvement allowing $R = X^{1/4+\epsilon}$ would give bounded gaps between primes! By considering additional possible primes, Goldston Pintz and Yıldırım were able to show

$$\liminf_n \frac{p_{n+1} - p_n}{\log p_n} = 0,$$

finally extending a long sequence of improvements to upper bounds for the left hand side Erdős [1940], Rankin [1950], Ricci [1954], Bombieri and Davenport [1966], Pilt'ja [1972], Uchiyama [1975], M. N. Huxley [1973, 1977], M. Huxley [1984], and Maier [1988].

Building on earlier work of Fouvry and Iwaniec [1980] and Bombieri, Friedlander, and Iwaniec [1986, 1987, 1989], Zhang [2014] succeeded in establishing an extended version of the Bombieri-Vinogradov Theorem for moduli with no large prime factors allowing $R = X^{1/4+\epsilon}$, and ultimately this allowed him to show

$$\mathbb{E} \, \#\{1 \le i \le k : L_i(n) \text{ prime}\} > 1,$$

if $k > 3\,500\,000$ and X is sufficiently large. The key breakthrough in Zhang's work was this result on primes in arithmetic progressions of modulus slightly larger than $X^{1/2}$. By choosing suitable linear functions, this then showed that

$$\liminf_n (p_{n+1} - p_n) \le 7 \cdot 10^7.$$

3   A modified GPY sieve method

An alternative approach to extending the work of Goldston, Pintz and Yıldırım was developed independently by the author Maynard [2015] and Tao (unpublished). The key difference was to consider the multidimensional generalization

$$(3\text{-}1) \qquad w(n) \propto \Big(\sum_{\substack{d_1,\ldots,d_k \\ d_i | L_i(n) \\ \prod_{i=1}^k d_i < R}} \mu(d) F\Big(\frac{\log d_1}{\log R}, \ldots, \frac{\log d_k}{\log R}\Big) \Big)^2,$$

for suitable smooth functions $F : \mathbb{R}^k \to \mathbb{R}$ supported on $[0,\infty)^k$. The flexibility of allowing the function F to depend on each divisor $d_1,\ldots,d_k$ of $L_1(n),\ldots,L_k(n)$ allows us to make $w(n)$ more concentrated on integers n when many of the $L_i(n)$ are prime.

With this choice, after some computation, one finds that

$$(3\text{-}2) \qquad \mathbb{E} \,\#\{1 \le i \le k : L_i(n) \text{ prime}\} = \Big(\frac{\sum_{i=1}^k J_i(\tilde{F})}{I(\tilde{F})} + o(1) \Big) \frac{\log R}{\log X},$$

provided, as before, we are able to count primes in arithmetic progressions to modulus R^2 on average. Here the $J_i(\tilde{F})$ and $I(\tilde{F})$ are k-dimensional integrals depending on a transform[2] $\tilde{F}$ of F, given by

$$J_\ell(\tilde{F}) = \int \cdots \int_{\sum_{i \ne \ell} t_i \le 1} \Big(\int_0^{1-\sum_{i \ne \ell} t_i} \tilde{F}(t_1,\ldots,t_k) dt_\ell \Big)^2 dt_1 \ldots dt_{\ell-1} dt_{\ell+1} \ldots dt_k,$$

$$I(\tilde{F}) = \int \cdots \int_{\sum_{i=1}^k t_i \le 1} \tilde{F}(t_1,\ldots,t_k)^2 dt_1 \ldots dt_k.$$

For any piecewise smooth choice of $\tilde{F}$ supported on $\sum_{i=1}^k t_i \le 1$ there is a corresponding choice of F. In particular, we can show that many of the L_i are simultaneously prime infinitely often, if we can show that $\sup_{\tilde{F}} (\sum_{i=1}^k J_i(\tilde{F})/I(\tilde{F})) \to \infty$ as $k \to \infty$.

---

[2] $\tilde{F}$ is F differentiated with respect to each coordinate.

A key advantage of this generalization is that we can make use of high dimensional phenomena such as concentration of measure. We concentrate on functions $\tilde{F}$ the form

$$\tilde{F}(t_1,\ldots,t_k) = \begin{cases} \prod_{i=1}^{k} G(t_i), & \sum_{i=1}^{k} t_i < 1, \\ 0, & \text{otherwise,} \end{cases}$$

and make a probabilistic interpretation of the integrals $J_i(\tilde{F})$ and $I(\tilde{F})$. Let $Z_1,\ldots,Z_k$ be i.i.d. random variables on $[0,\infty]$ with probability density function G^2, expectation $\mu = \int_0^\infty t G(t)^2 dt$ and variance $\sigma^2 = \int_0^\infty (t-\mu)^2 G(t)^2 dt$. Then

$$I(\tilde{F}) = \mathbb{P}\Big(\sum_{i=1}^{k} Z_i < 1\Big), \qquad J_\ell(\tilde{F}) \ge \Big(\int_0^{1/2} G(t)dt\Big)^2 \mathbb{P}\Big(\sum_{i=1}^{k} Z_i < \frac{1}{2}\Big).$$

The random variable $\sum_{i=1}^{k} Z_i$ has mean $k\mu$ and variance $k\sigma^2$, and so it becomes concentrated on $k\mu$ when k is large provided $\sigma^2/\mu \to 0$ as $k \to \infty$. In particular, if $\mu < 1/3k$ and $\sigma^2 k \to 0$, then $\mathbb{P}(\sum_{i=1}^{k} Z_i < 1/2)$ approaches 1 as $k \to \infty$. Therefore, to show that $\sum_{\ell=1}^{k} J_\ell(\tilde{F})/I(\tilde{F}) \to \infty$ as $k \to \infty$, it suffices to find a function G satisfying

$$\int_0^\infty t G(t)^2 < \frac{1}{3k}, \qquad \int_0^\infty G(t)^2 dt = 1,$$

$$k \int_0^\infty t^2 G(t)^2 dt \to 0 \quad \text{and} \quad k\Big(\int_0^\infty G(t)dt\Big)^2 \to \infty \quad \text{as} \quad k \to \infty.$$

We find that choosing

(3-3)
$$G(t) \approx \begin{cases} \frac{\sqrt{k \log k}}{1+tk \log k}, & t < k^{-3/4}, \\ 0, & \text{otherwise,} \end{cases}$$

gives a function G satisfying these constraints. (This choice can be found via the calculus of variations, and further calculations show that this choice is essentially optimal.) Putting this all together, we find that for some constant $c > 0$

$$\mathbb{E} \,\#\{i \,:\, L_i(n) \text{ prime}\} \ge (c \log k + o(1)) \frac{\log R}{\log X}.$$

In particular, taking $R = X^{1/4-o(1)}$ (as allowed by the Bombieri-Vinogradov theorem) and letting k be sufficiently large, we find that there are infinitely many integers n such that $(c \log k)/4 + o(1)$ of the $L_i(n)$ are simultaneously prime. Performing these calculations carefully allows one to take $c \approx 1$ when k is large.

Morally, the effect of such a choice of function $\tilde{F}$ is to make 'typical' divisors $(d_1, \ldots, d_k)$ occurring in (3-1) to have $\prod_{i=1}^{k} d_i$ smaller, but for it to be more common that some of the components of $(d_1, \ldots, d_k)$ are unusually large when compared with (2-1) or (2-3). This correspondingly causes the random integer n chosen with probability $w(n)$ to be such that the $L_i(n)$ are more likely to have slightly smaller prime factors, but it is also more likely that some of the $L_i(n)$ are prime.

4   Consequences for small gaps between primes

The above argument shows that there is a constant $c > 0$ such that for any set $\{L_1, \ldots, L_k\}$ of integral linear functions satisfying the hypotheses of Conjecture 3, at least $c \log k$ of the linear functions L_i are simultaneously prime infinitely often. To show that there are primes close together, we simply take the linear functions to be of the form $L_i(n) = n + h_i$ for some integers $h_1 \le \cdots \le h_k$ chosen to make $\prod_{i=1}^{k} L_i(n)$ have no fixed prime divisor, and so that $h_k - h_1$ is small.

In general, a good choice of the h_i is to take h_i to be the ith prime after the integer k. With this choice, $\prod_{i=1}^{k} L_i(n)$ has no fixed prime divisor and $h_k - h_i \approx k \log k$, so we can find intervals of length $k \log k$ containing $c \log k$ primes infinitely often. By working out the best possible implied constants, the argument we have sketched allows us to show that there are m primes in an interval of length $O(m^3 e^{4m})$ infinitely often. By incorporating refinements of the work of Zhang on primes in arithmetic progressions by the Polymath 8a project Castryck et al. [2014], and using some bounds from Harman's sieve Harman [2007], this can be improved slightly to $O(e^{3.815m})$ by work of R. C. Baker and Irving [2017].

If we are only interested in just how small a *single* gap can be, then we can improve the analysis and get an explicit bound on the size of the gap by adopting a numerical analysis perspective. The key issue is to find the smallest value of k such that for all large X

$$\mathbb{E} \, \# \{1 \le i \le k : \, L_i(n) \text{ prime} \} > 1,$$

since this immediately implies that two of the linear functions are simultaneously prime. Recalling that we can choose $R = x^{1/4-\epsilon}$ for any $\epsilon > 0$, using (3-2) we reduce the problem to finding a value of k as small as possible, such that we can find a function $\tilde{F} : [0, \infty)^k \to \mathbb{R}$ with

$$\frac{\sum_{i=1}^{k} J_i(\tilde{F})}{I(\tilde{F})} > 4.$$

We fix some basis functions $g_1, \ldots, g_r : [0, \infty)^k \to \mathbb{R}$ which are supported on $(t_1, \ldots, t_k)$ satisfying $\sum_{i=1}^{k} t_i \le 1$, and restrict our attention to functions $\tilde{F}$ in the linear span of

the g_i; that is functions $\tilde{F}$ of the form

$$\tilde{F}(t_1, \ldots, t_k) = \sum_{i=1}^{r} f_i g_i(t_1, \ldots, t_r),$$

for some coefficients $\mathbf{f} = (f_1, \ldots f_r) \in \mathbb{R}^r$ which we think of as variables we will optimize over. For such a choice, we find that $\sum_{i=1}^{k} J_i(\tilde{F})$ and $I(\tilde{F})$ are both quadratic forms in the variables $f_1, \ldots, f_r$, and the coefficients of these quadratic forms are explicit integrals in terms of $g_1, \ldots, g_r$. If we choose a suitably nice basis $g_1, \ldots, g_r$, then these integrals are explicitly computable, and so we obtain explicit $r \times r$ real symmetric matrices M_1 and M_2 such that

$$\frac{\sum_{i=1}^{k} J_i(\tilde{F})}{I(\tilde{F})} = \frac{\mathbf{f}^T M_1 \mathbf{f}}{\mathbf{f}^T M_2 \mathbf{f}}.$$

We then find that the choice of coefficients $\mathbf{f}$ which maximizes this ratio is the eigenvector of $M_2^{-1} M_1$ corresponding to the largest eigenvalue, and the value of the ratio is given by this largest eigenvalue. Thus the existence of a good function $\tilde{F}$ can be reduced to checking whether the largest eigenvalue of a finite matrix is larger than 4, which can be performed numerically by a computer.

If we choose the g_i to be symmetric polynomials of low degree, then this provides a nice basis since the corresponding integrals have a closed form solution, and allows one to make arbitrarily accurate numerical approximations to the optimal function $\tilde{F}$ with enough computation. For the problem at hand, these numerical calculations are large but feasible. This approach ultimately allows one to show that if $k = 54$ there is a function $\tilde{F}$ such that $\sum_{i=1}^{54} J_i(\tilde{F})/I(\tilde{F}) > 4$.

To turn this into small gaps between primes we need to choose the shifts h_i in our linear functions $L_i(n) = n + h_i$ so that $\prod_{i=1}^{54} L_i(n)$ has no fixed prime divisor, and the h_i are in as short an interval as possible. This is a feasible exhaustive numerical optimization problem, with an optimal choice of the $\{h_1, \ldots, h_{54}\}$ given by [3]

$$\{0, 2, 6, 12, 20, 26, 30, 32, 42, 56, 60, 62, 72, 74, 84, 86, 90, 96,$$
$$104, 110, 114, 116, 120, 126, 132, 134, 140, 144, 152, 156,$$
$$162, 170, 174, 176, 182, 186, 194, 200, 204, 210, 216, 222,$$
$$224, 230, 236, 240, 242, 246, 252, 254, 260, 264, 266, 270\}.$$

Putting this together, we find

$$\liminf_{n} (p_{n+1} - p_n) \leq 270.$$

---

[3] Such computations were first performed by Engelsma - see http://www.opertech.com/primes/k-tuples.html

This result is not quite the current record - in Polymath [2014] we make some further technical refinements to the sieve, which corresponds to modifying the expressions $J_i(\tilde{F})$ and $I(\tilde{F})$ slightly. This ultimately allows us to improve $k = 54$ to $k = 50$ and correspondingly improve from gaps of size at most 270 to gaps of at most 246, giving Theorem 2. The main ideas are the same as above.

5   Large gaps between primes

It turns out that because Theorem 1 gives strong partial information about the joint distribution of prime solutions to linear equations, which lie at the heart of many basic estimates about primes, it can be used to make progress on the existence of *large* gaps between primes, although the connection is less direct than with small gaps.

The first major breakthrough on large gaps between primes was due to Westzynthius [1931], who showed that there were gaps which could be arbitrarily large compared with the average gap. During the 1930s this was refined with ideas due to Erdős [1940] and Rankin [1936], giving

$$(5\text{-}1) \qquad \sup_{p_n \leq X} (p_{n+1} - p_n) \geq c' \frac{\log X \cdot \log\log X \cdot \log\log\log\log X}{(\log\log\log X)^2},$$

for some constant $c' > 0$. Subsequent improvements over the next 75 years Schönhage [1963], Rankin [1962/1963], and Maier and Pomerance [1990] were only in improving the value of the constant c', the strongest being $c' = 2e^\gamma + o(1)$, due to Pintz [2014]. All of these approaches are based on a variant of the following lemma, which reduces the problem of constructing large gaps between primes to a combinatorial covering problem.

Lemma 5. *If one can choose residue classes a_p (mod p) for $p \leq x$ such that every element of $\{1, \ldots, y\}$ is congruent to a_p (mod p) for some $p \leq x$, then there is a prime $p_n \ll e^x$ such that $p_{n+1} - p_n \geq y$.*

This lemma (which is a simple consequence of the Chinese Remainder Theorem), can be thought of as a natural generalization of the argument that $n! + 2, \ldots, n! + n$ are n consecutive composite integers, and so explicitly demonstrates a prime gap of size at least n. (This roughly corresponds to choosing $a_p = 1$ for all primes p and taking $y = x$.)

The key idea in the Erdős–Rankin construction was to choose $a_p = 0$ for 'medium sized' primes, and choose a_p differently for small and large primes. Specifically, a version of their argument follows the following strategy to choose the a_p in turn, for some parameter $2 < z < x^{1/2}$:

1. Choose $a_p = 0$ for 'medium primes' $p \in [z, x/3]$.

2. Choose $a_p = 1$ for 'small primes' $p < z$.

3. Choose a_p greedily for 'large primes' $p \in (x/3, x]$.

By 'choosing greedily' we mean that we pick a residue class $a_q \pmod{q}$ which contains the largest number of integers in $\{1, \ldots, y\}$ which are not congruent to $a_p \pmod{p}$ for some previously chosen a_p. There must be a residue class containing at least one uncovered element if we have not already covered all elements.

By choosing z appropriately[4], this allows one to cover $\{1, \ldots, y\}$ by residue classes $a_p \pmod{p}$ for $p \le 7y(\log \log y)^2/(\log y \cdot \log \log \log y)$, which ultimately results in the bound (5-1). The key feature making this construction work is that choosing $a_p = 0$ for a very large number of 'medium primes' is much more efficient that a typical choice. This is because integers in $\{1, \ldots, y\}$ with no prime factors bigger than z_1 are much less common than integers avoiding a random residue class for each prime bigger than z_1.

It was a well-known challenge of Erdős as to whether one could improve upon (5-1) by an arbitrarily large constant, and this was verified independently by Ford, Green, Konyagin, and Tao [2016] and the author Maynard [2016b]. Ultimately the approach of Ford-Green-Konyagin-Tao relied on the work of Green-Tao and Green-Tao-Ziegler on linear equations in primes Green and Tao [2010], Green and Tao [2012], and Green, Tao, and Ziegler [2012], whereas the work of the author relied on versions of Theorem 1. So far this second approach has proved more flexible for gaining quantitative improvements over (5-1), the strongest known results being due to a collaboration between all these authors Ford, Green, Konyagin, Maynard, and Tao [2018].

We focus here on the approach based around Theorem 1, first thinking about obtaining an arbitrarily large constant improvement over (5-1). We follow the same overall strategy as Erdős–Rankin, but improve the analysis for the large primes. By choosing the residue classes in a more sophisticated manner, we are able to remove *many* uncovered elements on average, rather than just 1 element, and this is the key feature which allows us to improve on (5-1).

Using the same choice of a_p as Erdős–Rankin for $p \le x/3$, we see that the elements of $\{1, \ldots, y\}$ which are not covered by $a_p \pmod{p}$ for $p \le x/3$ are integers $n < y$ where n has no prime factors in $[z, x/3]$ and $n - 1$ has no prime factors less than z. Let us call the set of such integers $\mathcal{S}$. This is a set which is very similar to the primes, since most elements have a very large prime factor, and it is not clear that there are *any* possible a_p which might cover more than one additional element. Indeed, a typical residue class $a_p \pmod{p}$ for $p > x/3$ will contain *no* elements of $\mathcal{S}$. The problem of showing the existence of unusual residue classes containing many elements of $\mathcal{S}$ leads us to the following toy problem.

---

[4]$z = \exp(\log x \cdot \log \log \log x \, / \, 2 \log \log x)$ gives a suitable choice.

Problem 1. *Given a prime modulus q, can we find a residue class a_q (mod q) which contains many primes all less than $q(\log q)^{1/2}$?*

This toy problem can be answered in the positive by a variant of Theorem 1. If we choose our linear functions to be of the form $L_i(n) = n + h_i q$ for suitable constants h_i, then the existence of a residue class containing many small primes is implied by many of the L_i being simultaneously prime at n for some $n \leq q$. The underlying sieve methodology is flexible enough to handle the fact that now the linear functions depend on q, and so can solve the toy problem. This correspondingly shows that for each large prime p, there are some residue classes a_p (mod p) containing many elements of $\mathcal{S}$.[5]

To turn this into an actual covering, we need the residue classes for *different* large primes to be approximately 'independent' from one another. To do this we use the probabilistic method, by choosing a residue class at random for each large prime $p \in (x/3, x/2]$, and showing that with high probability this results in approximately independent behavior. Specifically, we choose a_p randomly with

$$\mathbb{P}(a_p = a \text{ (mod } p)) \propto \sum_{\substack{n \equiv a \pmod{p} \\ L_1(n),...,L_k(n) \in \mathcal{S}}} w_{\mathcal{L}_p}(n),$$

where $w_{L_p}(n)$ is the normalized sieve weight introduced in Section 3 for the functions $L_1, \ldots, L_k$ with $L_i(n) = n + h_i p$. We make these choices independently for all $p \in (x/3, x/2]$, so

$$\mathbb{P}(n \text{ not covered by large primes}) = \prod_{p \in [x/3, x/2]} \left(1 - \sum_{m \equiv n \pmod{p}} w_{\mathcal{L}_p}(m)\right)$$

$$\leq \exp\left(-\mathbb{E} \, \#\{p \in [x/3, x/2] : n \equiv a_p \text{ (mod } p)\}\right).$$

In particular, if the expected number of times any $n \in \mathcal{S}$ is congruent to a_p (mod p) for some $p \in (x/3, x/2]$ is at least t, then the expected number of elements of $\mathcal{S}$ which are not covered by a_p (mod p) for $p \leq x/2$ is at most $e^{-t} \#\mathcal{S}$. If t is large, we can then greedily choose residue classes a_p (mod p) for $p \in (x/2, x]$ to cover these few remaining elements.

Thus our new strategy for choosing the a_p is:

1. Choose $a_p = 0$ for 'medium primes' $p \in [z, x/3]$.

2. Choose $a_p = 1$ for 'small primes' $p < z$.

---

[5]The work of Green-Tao on linear equations in primes allows one to show for *most* large primes q there are k primes less than $q(\log q)^{1/2}$ for any fixed k, which is sufficient to improve the estimates for large primes.

3. Choose a_p randomly according to the sieve weights $w_{\mathfrak{L}_p}$ independently for 'large primes' $p \in (x/3, x/2]$.

4. Choose a_p greedily for 'very large primes' $p \in (x/2, x]$.

This gives an arbitrarily large constant improvement over (5-1) provided we show that we can take t arbitrarily large. Calculations similar to those of Section 3 allow us to take t to be of size $\log k$ when we consider k linear functions, and so letting k be large enough we succeed in showing that the constant in (5-1) can indeed be taken to be arbitrarily large.

To get a quantitative improvement over (5-1) we run the essentially the same argument, but we need a version of Theorem 1 which has uniformity with respect to the number k of linear functions we consider as well as uniformity with respect to the coefficients of the linear functions. Such a version of Theorem 1 was established in Maynard [2016a]. With some technical modifications, this would ultimately yield a bound

$$\sup_{p_n \leq X} (p_{n+1} - p_n) \geq c'' \frac{\log X \cdot \log \log X}{\log \log \log X}$$

for some constant $c'' > 0$. This is not quite as good as Theorem 3, because one can improve the quantitative argument further by being more careful about the manner in which we choose the a_p for different large primes. If instead of choosing residue classes independently at random we use ideas based on the 'semi-random' or 'Rödl nibble' method from combinatorics, we are able to establish a hypergraph covering lemma which allows our covering by residue classes to have almost no overlaps. After working through the technical details, this ultimately gives an additional improvement of a factor $\log \log \log \log X$, and hence gives Theorem 3.

6   Limitations

Both Theorem 1 and Theorem 3 appear to be the qualitative limit of what these methods can give, and require an entirely new approach to do better. Theorem 2 depends on the quantitative aspects of Theorem 1, and can potentially be improved slightly. New ideas are likely needed to significantly improve upon Theorem 2, however.

Optimal weights in high-dimensional sieves are poorly understood, and we do not know of general barriers beyond the parity phenomenon. In the context of Theorem 1 the parity phenomenon means that we cannot hope to prove $k/2$ of our linear functions are simultaneously prime based on a sieve argument. In particular, significant new ideas are required to attack the Twin Prime Conjecture.

Although in principle this leaves open the possibility of having rather more than $\log k$ of the linear functions being simultaneously prime, since the Selberg sieve weights seem

to perform best in high dimensional sieving situations, we expect that it is unlikely to be possible to much do better than (3-2) based on weights formed by short divisor sums, even if we cannot prove a direct obstruction. Heuristic arguments show that over fairly general classes of potential Selberg sieve weights, we do not expect to do better than the choice given by (3-1). Given the choice (3-1), it is possible to show that a choice of smooth function similar to (3-3) is essentially best possible.

For the question of explicit small gaps between primes, there is some potential for further progress. Stronger results about primes in arithmetic progressions allow us to take R in (3-2) larger, which should reduce the critical value of k, and hence the size of the gap. The current record of 246 in Theorem 2 does not make use of the new equidistribution estimates of Zhang or its refinements (but these would likely only lead to small improvements). If we assume optimistic conjectures on the distribution of primes then we can do significantly better. Under the Elliott-Halberstam conjecture one can show gaps of size 12 Maynard [2015], and a generalization of the Elliott-Halberstam conjecture to numbers with several prime factors allows us to reach the absolute limit of these methods. Specifically, we have the following.

Theorem 6 (Polymath [2014]). *Assume the 'Generalized Elliott-Halberstam Conjecture' (see Polymath [ibid.]). Then we have*

$$\liminf_n (p_{n+1} - p_n) \leq 6.$$

The parity phenomenon makes it impossible for a first moment method of this type to prove a result less than 6, and so Theorem 6 is the strongest result of this type we can hope to prove along these lines.

All proofs showing the existence of large gaps between primes rely on some variant of Lemma 5, which allows one to construct a sequence of consecutive composite integers $n_1, \ldots, n_1 + y$ all with a prime factor of size $O(\log n_1)$. This places a severe limitation on how large the gaps we produce can be, since we expect that a large gap between primes will involve many composites n whose smallest prime factor is much larger than $\log n$. Specifically, Maier and Pomerance [1990] conjectured that the largest string of consecutive integers less than X all containing such a small prime factor should be of length

$$\log X \cdot (\log \log X)^{2+o(1)}.$$

Therefore we do not expect to be able to produce gaps larger than this without a new approach. The 'semi-random' method used in Ford, Green, Konyagin, Maynard, and Tao [2018] to show the existence of a good choice residue classes for large primes is essentially as good as one can hope for, so any quantitative progress based on the same overall method would likely require an improvement to Problem 1, showing the existence of more

than $\log\log q$ primes less than $q(\log q)^{1/2}$ in some residue class modulo q. A uniform version of Conjecture 3 suggests that there should be residue classes containing roughly $(\log q)^{1/2}/\log\log q$ such primes, but we do not know how to prove this.

7 Other applications and further reading

For a more thorough survey of the details of these ideas on small gaps between primes, as well as the ideas behind the breakthrough of Zhang, we refer the reader to the excellent survey articles of Granville [2015] and Kowalski [2015]. For a more details of the original work of Goldston, Pintz and Yıldırım we refer the reader to survey by Soundararajan [2007].

One useful feature of the argument of Section 3 is that the full strength of the Bombieri-Vinogradov theorem was not required to prove bounded gaps between primes. Provided we can estimate primes in arithmetic progressions with modulus of size x^ϵ on average, we would obtain a version of Theorem 1 with $c\epsilon\log k$ of the linear functions are prime. This allows one to show the existence of bounded gaps between primes in many subsets of the primes where one has this type of weaker arithmetic information. A general statement of this type was established in Maynard [2016a]. The fact that one can restrict the entire argument to an arithmetic progression also allows one to get some control on the joint distribution of various arithmetic functions. There have been many recent works making use of these flexibilities in the setup of the sieve method, including Thorner [2014], Castillo, Hall, Lemke Oliver, Pollack, and Thompson [2015], Banks, Freiberg, and Turnage-Butterbaugh [2015], Freiberg [2016], Pollack [2014], H. Li and Pan [2015], R. C. Baker and Pollack [2016], Matomäki and Shao [2017], R. C. Baker and Zhao [2016a], Chua, Park, and Smith [2015], Vatwani [2017], Troupe [2016], Pintz [2015, 2017], Huang and Wu [2017], R. C. Baker and Zhao [2016b], Banks, Freiberg, and Maynard [2016], R. Baker and Freiberg [2016], Kaptan [2016], Parshall [2016], and Pollack and Thompson [2015].

New results on long gaps between primes have also found further applications to other situations R. Baker and Freiberg [2016], Maier and Rassias [2017], and J. Li, Pratt, and Shakan [2017]. It is hopeful that the ideas behind Theorem 1thrm:ManyGaps can find further applications in the future.

References

R. C. Baker and A. J. Irving (2017). "Bounded intervals containing many primes". *Math. Z.* 286.3-4, pp. 821–841. MR: 3671562 (cit. on p. 369).

Roger C. Baker and Paul Pollack (2016). "Bounded gaps between primes with a given primitive root, II". *Forum Math.* 28.4, pp. 675–687. MR: 3518382 (cit. on p. 376).

Roger C. Baker and Liangyi Zhao (2016a). "Gaps between primes in Beatty sequences". *Acta Arith.* 172.3, pp. 207–242. MR: 3460812 (cit. on p. 376).

– (2016b). "Gaps of smallest possible order between primes in an arithmetic progression". *Int. Math. Res. Not. IMRN* 23, pp. 7341–7368. MR: 3632085 (cit. on p. 376).

Roger Baker and Tristan Freiberg (2016). "Limit points and long gaps between primes". *Q. J. Math.* 67.2, pp. 233–260. MR: 3509992 (cit. on p. 376).

William D. Banks, Tristan Freiberg, and James Maynard (2016). "On limit points of the sequence of normalized prime gaps". *Proc. Lond. Math. Soc. (3)* 113.4, pp. 515–539. MR: 3556490 (cit. on p. 376).

William D. Banks, Tristan Freiberg, and Caroline L. Turnage-Butterbaugh (2015). "Consecutive primes in tuples". *Acta Arith.* 167.3, pp. 261–266. MR: 3316460 (cit. on p. 376).

E. Bombieri (1965). "On the large sieve". *Mathematika* 12, pp. 201–225. MR: 0197425 (cit. on p. 366).

E. Bombieri and H. Davenport (1966). "Small differences between prime numbers". *Proc. Roy. Soc. Ser. A* 293, pp. 1–18. MR: 0199165 (cit. on p. 366).

E. Bombieri, J. B. Friedlander, and H. Iwaniec (1986). "Primes in arithmetic progressions to large moduli". *Acta Math.* 156.3-4, pp. 203–251. MR: 834613 (cit. on p. 366).

– (1987). "Primes in arithmetic progressions to large moduli. II". *Math. Ann.* 277.3, pp. 361–393. MR: 891581 (cit. on p. 366).

– (1989). "Primes in arithmetic progressions to large moduli. III". *J. Amer. Math. Soc.* 2.2, pp. 215–224. MR: 976723 (cit. on p. 366).

Abel Castillo, Chris Hall, Robert J. Lemke Oliver, Paul Pollack, and Lola Thompson (2015). "Bounded gaps between primes in number fields and function fields". *Proc. Amer. Math. Soc.* 143.7, pp. 2841–2856. MR: 3336609 (cit. on p. 376).

Wouter Castryck et al. (2014). "New equidistribution estimates of Zhang type". *Algebra Number Theory* 8.9, pp. 2067–2199. MR: 3294387 (cit. on p. 369).

Lynn Chua, Soohyun Park, and Geoffrey D. Smith (2015). "Bounded gaps between primes in special sequences". *Proc. Amer. Math. Soc.* 143.11, pp. 4597–4611. MR: 3391020 (cit. on p. 376).

Harald Cramér (1936). "On the order of magnitude of the difference between consecutive prime numbers". *Acta Arithmetica* 2.1, pp. 23–46 (cit. on p. 363).

P. D. T. A. Elliott and H. Halberstam (1970). "A conjecture in prime number theory". In: *Symposia Mathematica, Vol. IV (INDAM, Rome, 1968/69)*. Academic Press, London, pp. 59–72. MR: 0276195 (cit. on p. 365).

P. Erdős (1940). "The difference of consecutive primes". *Duke Math. J.* 6, pp. 438–441. MR: 0001759 (cit. on pp. 366, 371).

Kevin Ford, Ben Green, Sergei Konyagin, James Maynard, and Terence Tao (2018). "Long gaps between primes". *J. Amer. Math. Soc.* 31.1, pp. 65–105. MR: 3718451 (cit. on pp. 364, 372, 375).

Kevin Ford, Ben Green, Sergei Konyagin, and Terence Tao (2016). "Large gaps between consecutive prime numbers". *Ann. of Math. (2)* 183.3, pp. 935–974. MR: 3488740 (cit. on p. 372).

E. Fouvry and H. Iwaniec (1980). "On a theorem of Bombieri-Vinogradov type". *Mathematika* 27.2, 135–152 (1981). MR: 610700 (cit. on p. 366).

Tristan Freiberg (2016). "Short intervals with a given number of primes". *J. Number Theory* 163, pp. 159–171. MR: 3459566 (cit. on p. 376).

Daniel A. Goldston, János Pintz, and Cem Y. Yıldırım (2009). "Primes in tuples. I". *Ann. of Math. (2)* 170.2, pp. 819–862. MR: 2552109 (cit. on pp. 364, 366).

Andrew Granville (2015). "Primes in intervals of bounded length". *Bull. Amer. Math. Soc. (N.S.)* 52.2, pp. 171–222. MR: 3312631 (cit. on p. 376).

Ben Green and Terence Tao (2012). "The Möbius function is strongly orthogonal to nilsequences". *Ann. of Math. (2)* 175.2, pp. 541–566. MR: 2877066 (cit. on p. 372).

Ben Green, Terence Tao, and Tamar Ziegler (2012). "An inverse theorem for the Gowers $U^{s+1}[N]$-norm". *Ann. of Math. (2)* 176.2, pp. 1231–1372. MR: 2950773 (cit. on p. 372).

Benjamin Green and Terence Tao (2010). "Linear equations in primes". *Ann. of Math. (2)* 171.3, pp. 1753–1850. MR: 2680398 (cit. on p. 372).

Glyn Harman (2007). *Prime-detecting sieves*. Vol. 33. London Mathematical Society Monographs Series. Princeton University Press, Princeton, NJ, pp. xvi+362. MR: 2331072 (cit. on p. 369).

Wen Huang and XiaoSheng Wu (2017). "On the set of the difference of primes". *Proc. Amer. Math. Soc.* 145.9, pp. 3787–3793. MR: 3665033 (cit. on p. 376).

M. N. Huxley (1973). "Small differences between consecutive primes". *Mathematika* 20, pp. 229–232. MR: 0352021 (cit. on p. 366).

– (1977). "Small differences between consecutive primes. II". *Mathematika* 24.2, pp. 142–152. MR: 0466042 (cit. on p. 366).

Martin Huxley (1984). "An application of the Fouvry-Iwaniec theorem". *Acta Arith.* 43.4, pp. 441–443. MR: 756293 (cit. on p. 366).

Deniz A. Kaptan (2016). "A note on small gaps between primes in arithmetic progressions". *Acta Arith.* 172.4, pp. 351–375. MR: 3471048 (cit. on p. 376).

Emmanuel Kowalski (2015). "Gaps between prime numbers and primes in arithmetic progressions [after Y. Zhang and J. Maynard]". *Astérisque* 367-368, Exp. No. 1084, ix, 327–366. MR: 3363595 (cit. on p. 376).

Hongze Li and Hao Pan (2015). "Bounded gaps between primes of a special form". *Int. Math. Res. Not. IMRN* 23, pp. 12345–12365. MR: 3431623 (cit. on p. 376).

Junxian Li, Kyle Pratt, and George Shakan (2017). "A lower bound for the least prime in an arithmetic progression". *Q. J. Math.* 68.3, pp. 729–758. MR: 3698292 (cit. on p. 376).

Helmut Maier (1988). "Small differences between prime numbers". *Michigan Math. J.* 35.3, pp. 323–344. MR: 978303 (cit. on p. 366).

Helmut Maier and Carl Pomerance (1990). "Unusually large gaps between consecutive primes". *Trans. Amer. Math. Soc.* 322.1, pp. 201–237. MR: 972703 (cit. on pp. 371, 375).

Helmut Maier and Michael Th. Rassias (2017). "Large gaps between consecutive prime numbers containing perfect k-th powers of prime numbers". *J. Funct. Anal.* 272.6, pp. 2659–2696. MR: 3603310 (cit. on p. 376).

Kaisa Matomäki and Xuancheng Shao (2017). "Vinogradov's three primes theorem with almost twin primes". *Compos. Math.* 153.6, pp. 1220–1256. MR: 3647436 (cit. on p. 376).

James Maynard (2015). "Small gaps between primes". *Ann. of Math. (2)* 181.1, pp. 383–413. MR: 3272929 (cit. on pp. 364, 367, 375).

– (2016a). "Dense clusters of primes in subsets". *Compos. Math.* 152.7, pp. 1517–1554. MR: 3530450 (cit. on pp. 364, 374, 376).

– (2016b). "Large gaps between primes". *Ann. of Math. (2)* 183.3, pp. 915–933. MR: 3488739 (cit. on p. 372).

Hans Parshall (2016). "Small gaps between configurations of prime polynomials". *J. Number Theory* 162, pp. 35–53. MR: 3448260 (cit. on p. 376).

JZ Pilt'ja (1972). "The magnitude of the difference between consecutive primes". *Studies in number theory*. Izdat. Saratov. Univ., Saratov (Russian), pp. 73–79 (cit. on p. 366).

János Pintz (2014). "Some new results on gaps between consecutive primes". In: *Number theory, analysis, and combinatorics*. De Gruyter Proc. Math. De Gruyter, Berlin, pp. 261–278. MR: 3220124 (cit. on p. 371).

– (2015). "On the ratio of consecutive gaps between primes". In: *Analytic number theory*. Springer, Cham, pp. 285–304. MR: 3467403 (cit. on p. 376).

– (2017). "Patterns of primes in arithmetic progressions". In: *Number theory – Diophantine problems, uniform distribution and applications*. Springer, Cham, pp. 369–379. MR: 3676411 (cit. on p. 376).

Paul Pollack (2014). "Bounded gaps between primes with a given primitive root". *Algebra Number Theory* 8.7, pp. 1769–1786. MR: 3272281 (cit. on p. 376).

Paul Pollack and Lola Thompson (2015). "Arithmetic functions at consecutive shifted primes". *Int. J. Number Theory* 11.5, pp. 1477–1498. MR: 3376222 (cit. on p. 376).

D. H. J. Polymath (2014). "Variants of the Selberg sieve, and bounded intervals containing many primes". *Res. Math. Sci.* 1, Art. 12, 83. MR: 3373710 (cit. on pp. 364, 371, 375).

R. A. Rankin (1936). "The Difference between Consecutive Prime Numbers". *J. London Math. Soc.* 11.4, pp. 242–245. MR: 1574971 (cit. on p. 371).

– (1950). "The difference between consecutive prime numbers. IV". *Proc. Amer. Math. Soc.* 1, pp. 143–150. MR: 0034800 (cit. on p. 366).

– (1962/1963). "The difference between consecutive prime numbers. V". *Proc. Edinburgh Math. Soc. (2)* 13, pp. 331–332. MR: 0160767 (cit. on p. 371).

Giovanni Ricci (1954). "Sull'andamento della differenza di numeri primi consecutivi". *Riv. Mat. Univ. Parma* 5, pp. 3–54. MR: 0067139 (cit. on p. 366).

Arnold Schönhage (1963). "Eine Bemerkung zur Konstruktion grosser Primzahllücken". *Arch. Math. (Basel)* 14, pp. 29–30. MR: 0146154 (cit. on p. 371).

K. Soundararajan (2007). "Small gaps between prime numbers: the work of Goldston-Pintz-Yıldırım". *Bull. Amer. Math. Soc. (N.S.)* 44.1, pp. 1–18. MR: 2265008 (cit. on p. 376).

Jesse Thorner (2014). "Bounded gaps between primes in Chebotarev sets". *Res. Math. Sci.* 1, Art. 4, 16. MR: 3344173 (cit. on p. 376).

Lee Troupe (2016). "Bounded gaps between prime polynomials with a given primitive root". *Finite Fields Appl.* 37, pp. 295–310. MR: 3426591 (cit. on p. 376).

S. Uchiyama (1975). "On the difference between consecutive prime numbers". *Acta Arith.* 27. Collection of articles in memory of JuriĭVladimirovič Linnik, pp. 153–157. MR: 0366839 (cit. on p. 366).

Akshaa Vatwani (2017). "Bounded gaps between Gaussian primes". *J. Number Theory* 171, pp. 449–473. MR: 3556693 (cit. on p. 376).

A. I. Vinogradov (1965). "The density hypothesis for Dirichet *L*-series". *Izv. Akad. Nauk SSSR Ser. Mat.* 29, pp. 903–934. MR: 0197414 (cit. on p. 366).

Erik Johan Westzynthius (1931). "Über die Verteilung der Zahlen: die zu den n ersten Primzahlen teilerfremd sind." *Commentat. Phys.-Math.* 5.25, pp. 1–37 (cit. on p. 371).

Yitang Zhang (2014). "Bounded gaps between primes". *Ann. of Math. (2)* 179.3, pp. 1121–1174. MR: 3171761 (cit. on p. 366).

JAMES MAYNARD
MATHEMATICAL INSTITUTE
OXFORD
ENGLAND OX1 4AU
james.alexander.maynard@gmail.com