Prime Running Functions

Jaeyoon Kim

EXPERIMENTAL MATHEMATICS
https://doi.org/10.1080/10586458.2020.1786863

Prime Running Functions

Jaeyoon Kim

Department of Mathematics, University of Michigan, Ann Arbor, MI, USA

ABSTRACT

We study arithmetic functions $\Phi(x; d, a)$, called prime running functions, whose value at x sums the gaps between primes $p_k \equiv a \pmod{d}$ below x and the next following prime p_{k+1}, up to x. (The following prime p_{k+1} may be in any residue class $\pmod{d}$.) We empirically observe systematic biases of order $x/\log x$ in $\Phi(x; d, a) - \Phi(x; d, b)$ for different a, b. We formulate modified Cramér models for primes and show that the corresponding sum of prime gap statistics exhibits systematic biases of this order of magnitude. The predictions of such modified Cramér models are compared with the experimental data.

KEYWORDS

Number theory; prime number statistics; probability

1. Introduction

This article studies a new class of prime counting statistics based on the size of gaps between primes, where the smaller prime in the gap is restricted to a fixed arithmetic progression. The *prime running function* $\Phi(x; d, a)$ counts the number of integers $n \le x$ having the property that the largest prime $p \le n$ satisfies $p \equiv a \pmod{d}$. Alternatively, these statistics may be thought of as counting the primes in a fixed arithmetic progression, each weighted by the length of the gap from that prime to the next larger prime. We present experimental evidence that

$$\Phi(x; d, a) = \frac{1}{\varphi(d)}x + R(d; a)\frac{x}{\log x} + o\left(\frac{x}{\log x}\right), \tag{1.1}$$

may hold as $x \to \infty$ (Conjecture 2.3). In this formula, even the main term $\Phi(x, d, a) \sim \frac{1}{\varphi(d)}x$ is conjectural for $d \ge 3$ (Conjecture 2.2). The main term is what one would expect from the mean of gap sizes not depending on the modulus $a \pmod{d}$, while the term $R(d; a)\frac{x}{\log x}$ quantifies a "bias term" which is the main focus of this article. We rigorously analyze a probabilistic model (modified Cramér model having a preliminary sieving on a modulus Q) which predicts a functional form of shape (1.1), with a bias term present. For small moduli d, we compare the model prediction for $R(d, a)$, taking Q to be a large primorial, against empirical estimates for $R(d, a)$.

The bias phenomenon was discovered in study of "prime running races" $\Phi(x; d, a) - \Phi(x; d, b)$, between two different residue classes a, b (with $(ab, d) = 1$). Such races are analogous to "prime number races" $\pi(x; d, a) - \pi(x; d, b)$, on which there has been a large amount of work (see Section 1.3). We present evidence that prime running races have biases asymptotically equivalent to $Cx/\log x$ for some constant $C = C(d; a, b)$. The conjectured formula (1.1) would give $C(d; a, b) = R(d; a) - R(d; b)$. This bias phenomenon was discovered experimentally for these statistics by plotting the simultaneous movements of two prime running races as n increases on a single figure (Figure 2). We plotted a walk on the square lattice $\mathbb{Z}^2$ with X component of the walk given by one prime running race and Y component of the walk given by a different prime running race. One can make similar plots for prime number races $\pi(x; d, a) - \pi(x; d, b)$. One sees a great difference in the appearance of the plots in the two cases. The plots for prime number races resemble 2-dimensional simple random walks, while the plots for prime running races do not resemble random walks at all, and exhibit systematic biases increasing with x. We illustrate this phenomenon with an example.

1.1. Prime walk

The following "prime walk" on the integer lattice $\mathbb{Z}^2$ takes steps according to the location of the two different prime number races (mod 5) as the variable n increments. We begin the walk from the origin $(0, 0)$ at time $n = 1$. From there, we repeatedly increment n by 1. Whenever $n = p_k$ is a prime, we do the following:

- if $p_k \equiv 1 \pmod{5}$, move down; add $(0, -1)$
- if $p_k \equiv 2 \pmod{5}$, move left; add $(-1, 0)$
- if $p_k \equiv 3 \pmod{5}$, move up; add $(0, 1)$
- if $p_k \equiv 4 \pmod{5}$, move right; add $(1, 0)$

If n is not prime (or if $n = 5$), we do not move.

CONTACT Jaeyoon Kim jaeykim@umich.edu Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1043, USA.

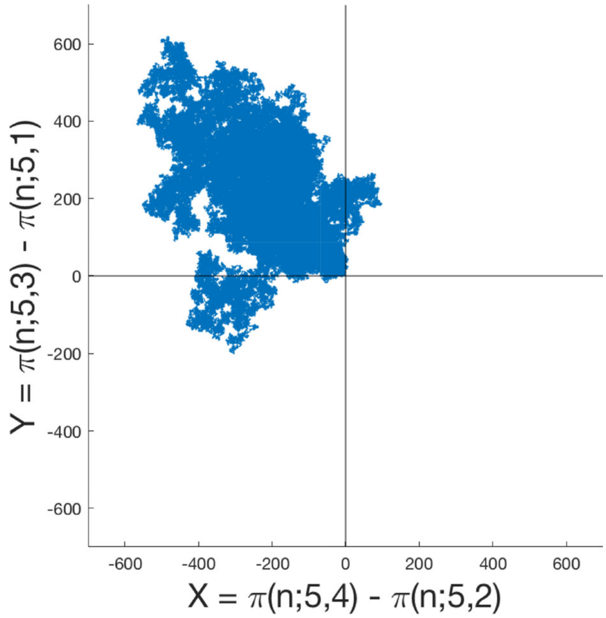

Figure 1. Plot of prime walk for $1 \leq n \leq 10^8$.

Figure 1 presents the plot of points of the "prime walk" for $n \leq 10^8$. The nth point of the walk is located at position

$$(\pi(n;5,4) - \pi(n;5,2),\ \pi(n;5,3) - \pi(n;5,1)) \quad 1 \leq n \leq 10^8.$$

Using the terminology of Granville and Martin [10], Figure 1 exhibits the motion of two "prime number races" (mod 5); the Y-component demonstrates the race between Team 3 and Team 1, while the X-component encodes the race between Team 4 and Team 2. The resulting walk exhibits a slight Northwest bias with a maximum magnitude of order 10^3. The Northwest bias is explained by Chebyshev's bias (mod 5) (see Section 1.3). Qualitatively, Figure 1 resembles a sample path of a simple random walk, in that its maximum distance from the origin is approximately proportional to the square root of the number of steps.

1.2. Prime run

We change the rules of the "prime walk" (mod 5) above to obtain "*prime run*." Whenever $n = p_k$ is prime, we move in the same direction as the prime walk. However, the prime run does not stop when n is composite, it continues taking steps in the same

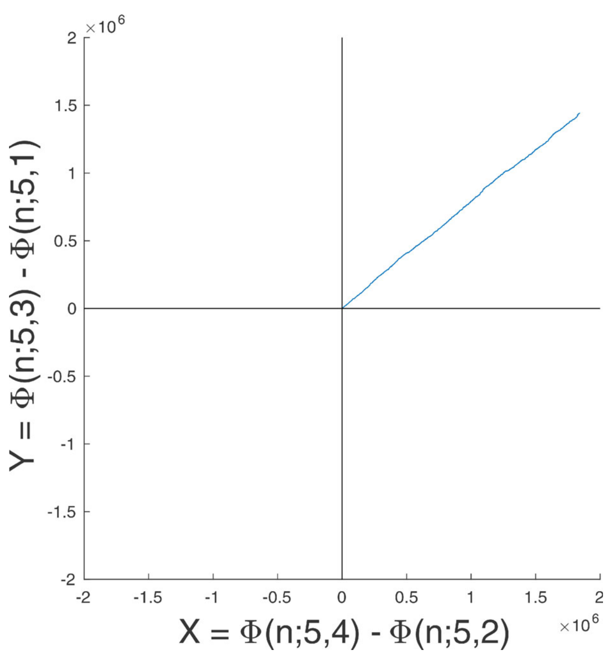

Figure 2. Plot of prime run for $1 \leq n \leq 10^8$.

direction that we were moving at time $n-1$. Each time $n = p_k$ is prime, we have an opportunity for changing directions. For the composite values of n in between, we move in a straight line at unit speed, following the previous direction.

To obtain the position when $n = p_{k+1} - 1$, we can apply the following algorithm to the position when $n = p_k - 1$.

- if $p_k \equiv 1 \pmod 5$, move down until the next prime; add $(0, -(p_{k+1} - p_k))$
- if $p_k \equiv 2 \pmod 5$, move left until the next prime; add $(-(p_{k+1} - p_k), 0)$
- if $p_k \equiv 3 \pmod 5$, move up until the next prime; add $(0, p_{k+1} - p_k)$
- if $p_k \equiv 4 \pmod 5$, move right until the next prime; add $(p_{k+1} - p_k, 0)$

If $n = p_3 = 5$, we stop the walk until the next prime $n = p_4 = 7$ is reached. Instead of moving one step, the prime run increments by the magnitude of the gap between primes. Since the average gap size between the primes is $x/\pi(x) \sim \log(x)$, one might expect that the prime running plot will look approximately like the prime walk scaled up by a factor of $\log(x)$.

Figure 2 presents the plot of points of the prime run for $n \le 10^8$.

It looks like a line! Also, we observe that the maximum distance reached away from the origin is of order 10^6, which is much larger than the 10^3 spread for the prime walk. We observe that the distance of order 10^6 from the origin reached is considerably smaller than the 10^8 steps taken, indicating that the line in the plot has some thickness. Another observation is that the direction of drift in Figure 2 is different from the direction of the "Chebyshev bias" in the prime walk shown in Figure 1. Experimentally, this plot of the prime run exhibits a much larger and more sharply focused drift than the drift in the prime walk.

1.3. Related work

The study of differences between the number of primes in different residue classes below a threshold x has a long history. In the paper "Comparative Prime Number Theory" by Knapowski and Turan [15, Problem 8], the study of $\pi(x; d, a) - \pi(x; d, b)$ was termed the (Shanks–Renyi) "prime number race." Let $\mathcal{P} = \{p_1 < p_2 < \cdots\}$ denote the set of primes, with $p_1 = 2, p_2 = 3$, etc. We recall that the counting function for primes in arithmetic progression $a \pmod d$ is

$$\pi(x; d, a) = \sum_{\substack{p_k \le x \\ p_k \equiv a \pmod d}} 1. \tag{1.2}$$

We assume $(a, d) = 1$, so that there are infinitely many primes in the class by Dirichlet's theorem.

The subject of prime number races trace back to an assertion of Chebyshev [3] in 1853 (without proof) that

$$\lim_{c \to 0^+} \sum_{k=1}^{\infty} (-1)^{\frac{p_k+1}{2}} e^{-p_k c} = +\infty, \tag{1.3}$$

which gave a sense in which there are more primes of the form $4n+3$ than of the form $4n+1$. In 1916, Hardy and Littlewood [11, pp. 141–148] proved Chebyshev's assertion under the assumption that the Riemann hypothesis holds for $L(s, \chi_{-4})$.

However, already in 1914, Littlewood [18] proved that $\pi(x; 4, 3) - \pi(x; 4, 1)$ has infinitely many sign changes. In 1995, by assuming the generalized Riemann hypothesis, Kaczorowski [13] extended Littlewood's result to races between all pairs of distinct nonzero residue classes (mod 5). It is now known that the lead of many prime races $\pi(x; d, a) - \pi(x; d, b)$ changes infinitely many times for many particular pairs of distinct reduced residue classes a, b for many moduli d. For a survey on the case of prime moduli d, see Granville and Martin [10]. For a general discussion of the distribution of the primes over different arithmetic progressions, see Kaczorowoski [14].

In 1994, Rubinstein and Sarnak [21] introduced another variant of prime number races which quantifies the degree to which one race is ahead of another. Their framework is to measure the set of values of x in which one member of a prime number race is ahead of another using *logarithmic density*. A set S of positive integers has a well-defined logarithmic density $d(S)$ if the following limit exists:

$$d(S) := \lim_{x \to \infty} \frac{1}{\log x} \left(\sum_{\{n \in S : n \le x\}} \frac{1}{n} \right).$$

Rubinstein and Sarnak showed, assuming strong conjectures on the distribution of zeros of L-functions, that a logarithmic density exists for the set of x such that $\pi(x; d, a) > \pi(x; d, b)$, where a and b are residues (mod d) having $(ab, d) = 1$. Their analysis predicted that the logarithmic density of x for which $\pi(x; 4, 3) > \pi(x; 4, 1)$ is approximately 0.9959. Rubinstein and Sarnak termed this phenomenon "Chebyshev's bias." See Feuerverger and Martin [7] and Fiorilli [8] for other examples of large biases in this sense.

The quantitative sizes of how far one member of a prime number race can be ahead of another (of such "Chebyshev biases") is always small compared to the average value of these functions separately, which is about $\frac{1}{\varphi(d)}\frac{x}{\log x}$. The prime number theorem for arithmetic progressions ([19, Corollary 11.21] and [6]) with $(a, d) = 1$ states

$$\pi(x; d, a) = \frac{1}{\varphi(d)}\mathrm{Li}(x) + \mathcal{O}\left(xe^{-c_d\sqrt{\log x}}\right), \tag{1.4}$$

where $\mathrm{Li}(x)$ denotes the logarithmic integral $\mathrm{Li}(x) = \int_2^x \frac{dt}{\log t}$ and c_d is some positive constant depending on d. Then each prime number race (mod d) with $\gcd(ab, d) = 1$ satisfies

$$|\pi(x; d, a) - \pi(x; d, b)| = \mathcal{O}\left(xe^{-c_d\sqrt{\log x}}\right).$$

Assuming the generalized Riemann hypothesis, this bound can be improved to

$$|\pi(x; d, a) - \pi(x; d, b)| = \mathcal{O}\left(x^{\frac{1}{2}+\epsilon}\right) \quad \text{for any } \epsilon > 0.$$

In 2016, Lemke Oliver and Soundararajan [17] introduced new prime statistics having "unexpected biases" which are quantitatively very large as a function of x. These statistics concerned the counts up to x for r-tuples of r consecutive primes whose residue classes (mod d) are specified. Restricting to $r = 2$, let $\pi(x; d, (a, b))$ count the number of primes $p_k \le x$ such that $p_k \equiv a \pmod d$ and $p_{k+1} \equiv b \mod d$. Here, we follow the standard notation that p_k denotes the k^{th} smallest prime. We call such functions "consecutive prime counting functions in arithmetic progressions." Here, one expects equidistribution of these counts as $x \to \infty$ in the sense that

$$\pi(x; d, (a, b)) \sim \frac{1}{\varphi(d)^2}\frac{x}{\log x} \quad \text{as } x \to \infty,$$

although such results remain conjectural. Lemke Oliver and Soundararajan formulated precise conjectures on the asymptotic growth of $\pi(x; d, (a, b))$ which predicts that the size of the bias terms can be as large as $x\frac{\log\log x}{(\log x)^2}$. Their main conjecture implies that differences of such functions

$$\pi(x; d, (a_1, b_1)) - \pi(x; d, (a_2, b_2)),$$

which we may call "consecutive prime number races," sometimes observe biases of order $x\frac{\log\log x}{(\log x)^2}$. Such a large systematic bias of the consecutive prime number races lead to a fixed sign for all sufficiently large x, which implies that one function wins the race for all sufficiently large x.

As an example, their main conjecture predicts[1]

$$\pi(x; 5, (1, 2)) - \pi(x; 5, (1, 1)) = \frac{1}{8}x\frac{\log\log x}{(\log x)^2} + \mathcal{O}\left(\frac{x}{(\log x)^2}\right),$$

an assertion implying that this bias will be positive for all large x. This bias term is smaller than the growth rate of $\pi(x; d, a)$ by a multiplicative factor $\frac{\log\log x}{\log x}$.

Unlike the functions studied by Lemke Oliver and Soundararajan which require two or more arithmetic progression conditions to exhibit bias, the prime running functions can exhibit a large bias even if we only restrict to a single arithmetic progression, as in (1.1).

1.4. Contents

Section 2 defines prime running functions and formulates conjectures regarding the asymptotic behavior of the prime running function. In Section 3, we present empirical evidence for $d = 3, 4, 5, 7$, and 25 which provided the original basis for some of the conjectures formulated in Section 2. In Section 4, we formulate probabilistic models for the primes which may explain the large bias terms. These probabilistic models are versions of the Cramér model of random primes, modified by first making initial sieving to remove any integers not co-prime to sieve modulus Q. These models predict that the prime running functions observe a bias of order $x/\log x$ (Theorems 4.3 and 4.5) and other behaviors (Theorems 4.6 and 4.8). These models provide heuristic justification for the conjectures made in Section 2. The proof of Theorem 4.3 is found in Section 4.2. Section 5 provides an efficient method for computing the predicted bias computation by the model. The predictions of the Cramér model is compared with empirical data. Section 6 makes concluding remarks on analyzing probabilistic models for prime running functions.

[1] We take $r = 2$ and $\frac{1}{8} = \frac{1}{2\varphi(5)}$ in their main conjecture, p. E4447.

2. Prime running functions: definitions and conjectures

2.1. Prime running functions

Now we introduce the prime running function.

Definition 2.1. *For* $a \pmod d$, *we define the* prime running function *as*

$$\Phi(x; d, a) = \sum_{\substack{1 \le n \le x \\ \lfloor n \rfloor_{\mathcal{P}} \equiv a \pmod d}} 1.$$

Here the $\mathcal{P}$*-floor function* $\lfloor n \rfloor_{\mathcal{P}}$ *gives the largest prime less than or equal to n. We define* $\lfloor 1 \rfloor_{\mathcal{P}} = 0$.

The prime running function is similar to the prime counting function "weighted" by the magnitude of the prime gaps.

$$\Phi(x; d, a) = \sum_{\substack{p_{k+1} \le x \\ p_k \equiv a \pmod d}} \left(p_{k+1} - p_k\right) + e(x; d, a), \tag{2.1}$$

where

$$e(x; d, a) = \begin{cases} -p_{k+1} + \lfloor x \rfloor + 1 & \text{if } \lfloor x \rfloor_{\mathcal{P}} = p_k,\ p_k \equiv a \pmod d, \\ 0 & \text{otherwise.} \end{cases}$$

The additional error term $e(x; d, a)$ is bounded by

$$|e(x; d, a)| = \mathcal{O}\left(x^{7/12+\epsilon}\right)$$

(see Huxley [12, Chap. 28]).

The plot of the prime run given in Figure 2 is a plot of two differences of prime running functions

$$\left(x_n, y_n\right) = (\Phi(n; 5, 4) - \Phi(n; 5, 2), \Phi(n; 5, 3) - \Phi(n; 5, 1))$$

for $1 \le n \le 10^8$.

2.2. Conjectures for prime running functions

It is natural to expect that the values of the prime running function are equidistributed among residue classes with $\gcd(a, d) = 1$.

Conjecture 2.2 (Prime running function main term). *For any integer* $d \ge 2$ *and any reduced residue* $a \pmod d$,

$$\Phi(x; d, a) \sim \frac{1}{\varphi(d)} x \quad \text{as } x \to \infty.$$

Aside from the trivial exception $d = 2$, there seem to be no results known to give unconditional asymptotic formulas for functions of this type. Furthermore, there does not even seem to be any lower bounds of the form $\Phi(x; d, a) > cx$ with $c > 0$.

Since the average spacing between primes is of order $\log x$, Conjecture 2.2 is equivalent to the statement that the average value of $p_{k+1} - p_k$ is independent of the congruent class of $p_k \pmod d$ to an error $o(\log x)$ as $x \to \infty$.

The main empirical observation of this article is the (apparent) existence of large biases in the prime running function away from the expected main term. We formulate a conjecture characterizing the bias of the prime running function between different residues.

Conjecture 2.3 (Prime running bias conjecture). *For any integer* $d \ge 2$ *and integer a with* $\gcd(a, d) = 1$, *there exists a constant* $R(d; a)$ *such that*

$$\Phi(x; d, a) = \frac{1}{\varphi(d)} x + R(d; a) \frac{x}{\log x} + o\left(\frac{x}{\log x}\right).$$

The order of magnitude $x/\log x$ for the bias term in Conjecture 2.3 is predicted by a probabilistic model in Section 4.

Assuming Conjecture 2.3, by taking the differences of two prime running functions, we can directly observe the bias term:

$$\Phi(x; d, a_1) - \Phi(x; d, a_2) = (R(d; a_1) - R(d; a_2)) \frac{x}{\log x} + o\left(\frac{x}{\log x}\right).$$

In Section 3, we present empirical estimates of the constants $R(d; a)$ for $d = 3, 5$, and 7. We call the constants $R(d; a)$, *bias constants*.

The empirical data and a probability model (see Theorem 4.6) suggest that the following anti-symmetry property of the bias constants may hold.

Conjecture 2.4 (Bias constant anti-symmetry conjecture). *The bias constants for prime running function for modulus d satisfy*

$$R(d;-a) = -R(d;a),$$

when $(a,d)=1$.

In addition, Conjecture 2.3 for $d=3$ implies anti-symmetry $R(3;1) = -R(3;2)$ since $\Phi(x;3,1)+\Phi(x;3,2) = x + \mathcal{O}(1)$.

Limited empirical data and a probabilistic model (see Theorem 4.8) support the conjecture that the bias constants (mod d) depend only on the square-free part $d_{\rm sf}$ of d, also called the radical of d, see [1].

$$d_{\rm sf} = \mathrm{rad}(d) := \prod_{p|d} p \tag{2.2}$$

Conjecture 2.5 (Radical equivalence conjecture). *For all* $d \geq 2$, *with* $(a,d)=1$,

$$R(d;a) = \frac{\varphi(d_{\rm sf})}{\varphi(d)} R(d_{\rm sf};a), \tag{2.3}$$

where $d_{\rm sf} = \mathrm{rad}(d)$ *is the square-free part of d.*

In particular, $R(d;a) = R(d;a')$ if $a \equiv a' \pmod{d_{\rm sf}}$. For special case $d=2$, we know unconditionally that $R(2;1)=0$. Thus Conjecture 2.5 predicts that

$$R(2^j;a) = 0 \tag{2.4}$$

for all $j \geq 1$ and $a \equiv 1 \pmod 2$.

3. Experimental results

In this section, we present numerical data on the prime running function for a few small modulus d over their residue classes. In Section 3.1, we provide data for $d = 3, 5$, and 7. In Section 3.2, we provide data for $d=4$ and 25.

3.1. Prime running function data for prime modulus

We first present data on the prime running functions for prime values of d and compare them to the predicted values from the main term Conjecture 2.2. Tables 1 and 2 give numerical data for $d=3$ and $d=5$ at $x = 10^8, 10^{10}, 10^{12}$.

This numerical data suggests that the main term is $\frac{1}{\varphi(d)}x$ and that systematic bias error terms are present.

The size of the bias appears to be growing more slowly than the main term $\frac{1}{\varphi(d)}x$ as x increases in powers of 10.

To fit the data to Conjecture 2.3, we introduce a new function.

Table 1. Value of the prime running function $\Phi(x;3,a)$ at different values of x and a (mod 3). Last row gives the main term from Conjecture 2.2.

$\Phi(x;3,a)$			
a \ x	$x=10^8$	$x=10^{10}$	$x=10^{12}$
$a=1$	51209542	5091131912	507317304782
$a=2$	48790455	4908868085	492682695215
2.2	50000000	5000000000	500000000000

Table 2. Value of the prime running function $\Phi(x;5,a)$ at different values of x and a (mod 5). Last row gives the main term from Conjecture 2.2.

$\Phi(x;5,a)$			
a \ x	$x=10^8$	$x=10^{10}$	$x=10^{12}$
$a=1$	24644198	2470292440	247456175258
$a=2$	23714857	2401583475	241999191675
$a=3$	26085716	2588759228	257451209200
$a=4$	25555226	2539364854	253093423864
2.2	25000000	2500000000	250000000000

Definition 3.1. For integer $d \geq 2$ and reduced residue $a \pmod d$, we define the *rescaled bias function* $R(x; d, a)$ by

$$R(x; d, a) := \left(\Phi(x; d, a) - \frac{1}{\varphi(d)}x\right)\frac{\log x}{x}. \tag{3.1}$$

Conjecture 2.3 can now be rewritten in the following form.

Conjecture 3.2. *For all $d \geq 2$, with $\gcd(a, d) = 1$ the following limit exists.*

$$R(d; a) = \lim_{x\to\infty} R(x; d, a).$$

Figures 3 and 4 plot the rescaled bias functions for $d = 3$ and $d = 5$ for $x \leq 10^{10}$. The resulting curves appear approximately flat, which supports the conjecture that the prime running functions approach $\frac{1}{\varphi(d)}x + R(d; a)\frac{x}{\log x}$, where $R(d; a)$ is the bias constant.

Tables 3–5 numerically computes the values of $R(x; d, a)$ for moduli $d = 3$, $d = 5$, and $d = 7$ at $x = 10^8, 10^{10}, 10^{12}$.

In Tables 3–5, slow trends are visible, but their directions (increase of decrease in magnitude) seems to vary with a. Furthermore, the data are consistent with the anti-symmetry Conjecture 2.4.

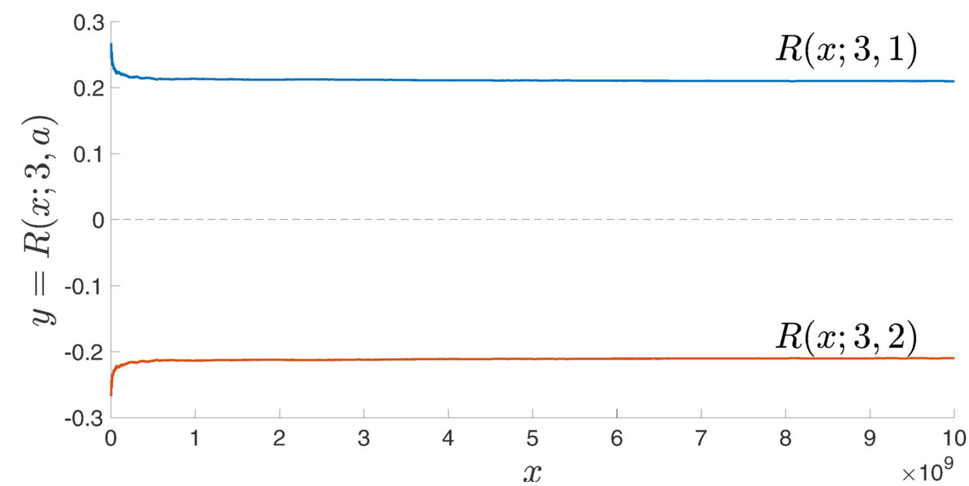

Figure 3. Plot of $R(x; 3, a)$ for all reduced residues $a \pmod 3$ and $x \leq 10^{10}$. The line $y = 0$ is marked with a dashed line.

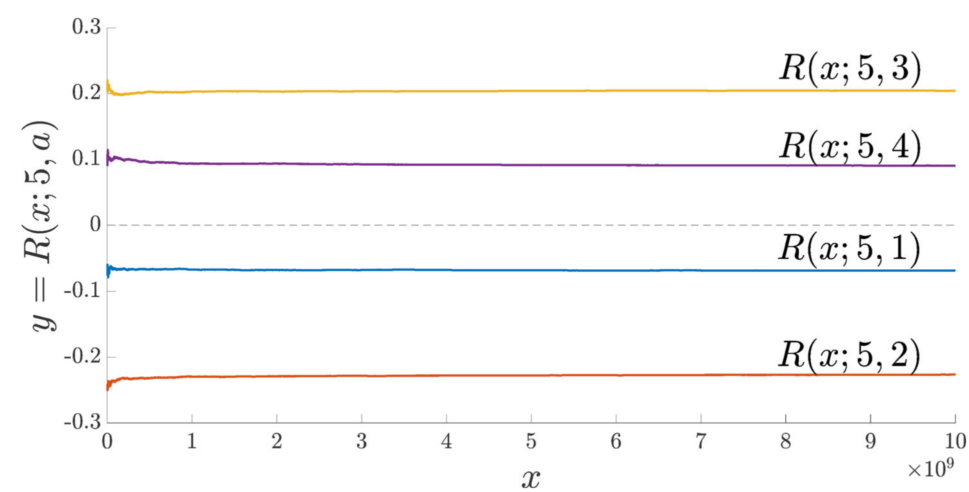

Figure 4. Plot of $R(x; 5, a)$ for all reduced residues $a \pmod 5$ and $x \leq 10^{10}$. The line $y = 0$ is marked with a dashed line.

Table 3. Values of $R(x; 3, a)$ for various values of x and a.

$R(x;3,a)$			
a \ x	$x = 10^8$	$x = 10^{10}$	$x = 10^{12}$
$a = 1$	0.2228	0.2098	0.2022
$a = 2$	−0.2228	−0.2098	−0.2022

Table 4. Values of $R(x; 5, a)$ for various values of x and a.

a \ x	$x = 10^8$	$x = 10^{10}$	$x = 10^{12}$
$a = 1$	−0.0655	−0.0684	−0.0703
$a = 2$	−0.2367	−0.2266	−0.2211
$a = 3$	0.2000	0.2044	0.2059
$a = 4$	0.1023	0.0906	0.0855

Table 5. Values of $R(x; 7, a)$ for various values of x and a.

a \ x	$x = 10^8$	$x = 10^{10}$	$x = 10^{12}$
$a = 1$	0.1530	0.1501	0.1461
$a = 2$	−0.0780	−0.0709	−0.0680
$a = 3$	0.0588	0.0527	0.0506
$a = 4$	−0.0681	−0.0601	−0.0571
$a = 5$	0.0583	0.0590	0.0626
$a = 6$	−0.1240	−0.1308	−0.1343

3.2. *Prime running function data for prime power modulus*

Figure 5 plots $R(x; 4, a)$ for $x \leq 10^{10}$. There appears to be a smaller bias for the prime running functions for 1 (mod 4) and 3 (mod 4). This is consistent with Conjecture 2.5 which would imply that $R(4; 1) = R(4; 3) = 0$.

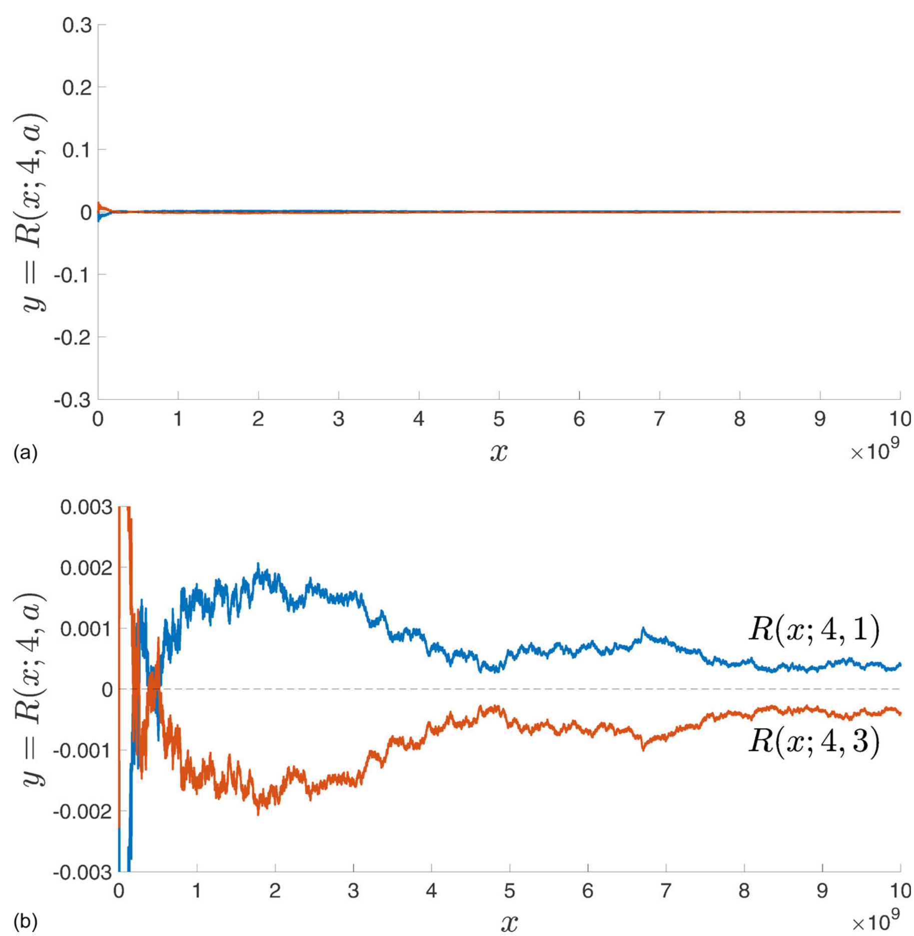

Figure 5. (a) Plot of $R(x; 4, a)$ for all reduced residues a (mod 4) and $x \leq 10^{10}$. The line $y = 0$ is marked with a dashed line. The axis is set to the same scale as Figures 3 and 4. (b) Y-axis is zoomed in by a scale of 100.

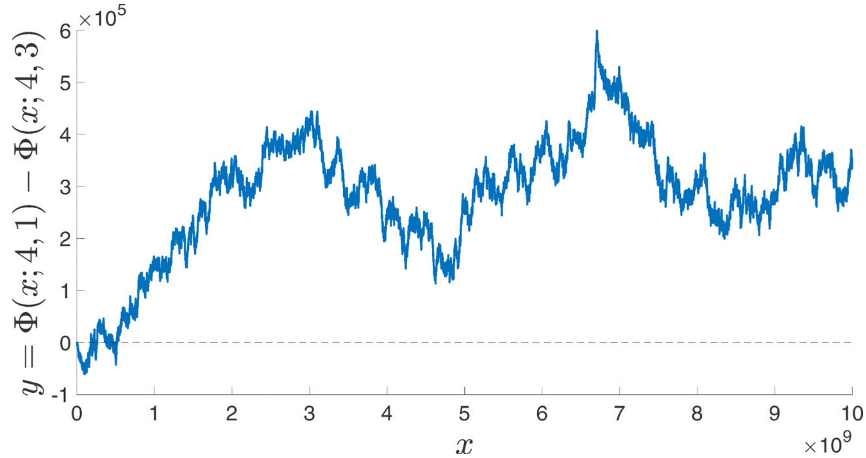

Figure 6. Plot of $\Phi(x; 4, 1) - \Phi(x; 4, 3)$ against x for $x \leq 10^{10}$. The line $y = 0$ is marked with a dashed line.

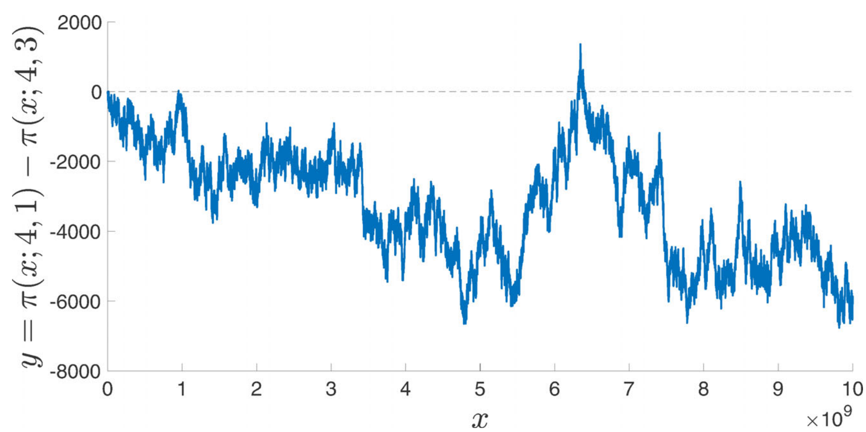

Figure 7. Plot of $\pi(x; 4, 1) - \pi(x; 4, 3)$ against x for $x \leq 10^{10}$. The line $y = 0$ is marked with a dashed line.

Table 6. Values of $R(x; 4, a)$ for various values of x and a.

	$R(x; 4, a)$		
a \ x	$x = 10^8$	$x = 10^{10}$	$x = 10^{12}$
1	−0.0041	0.0004	0.0002
3	0.0041	−0.0004	−0.0002

Table 7. Values of $R(x; 25, a)$ for various values of x and $a \equiv 1 \pmod 5$.

	$R(x; 25, a)$		
a \ x	$x = 10^8$	$x = 10^{10}$	$x = 10^{12}$
$a = 1$	−0.0129	−0.0139	−0.0140
$a = 6$	−0.0131	−0.0136	−0.0141
$a = 11$	−0.0144	−0.0139	−0.0141
$a = 16$	−0.0127	−0.0137	−0.0141
$a = 21$	−0.0125	−0.0134	−0.0140

Figure 6 presents the unscaled prime running race between 1 (mod 4) and 3 (mod 4). Chebyshev's bias for (mod 4) is illustrated in Figure 7. In the depicted domain, the sign of $\Phi(x; 4, 1) - \Phi(x; 4, 1)$ is predominantly positive, which is the opposite sign from Chebyshev's bias $\pi(x; 4, 1) - \pi(x; 4, 3)$.

We see that unlike prime running races between prime moduli, the bias for $d = 4$ is of much smaller order (roughly of order $\sqrt{x}$). We observe that for large values of x in the plot, $\Phi(x; 4, 1) - \Phi(x; 4, 3) > 0$.

Figure 8 plots $R(x; 25, a)$ for $x \leq 10^{10}$. Figure 8 follows a strong numerical agreement with Conjecture 2.5.

Table 6 numerically computes the values of $R(x; 4, a)$. Table 6 suggests that $R(4, 1) = R(4, 3) = 0$ as predicted by Conjecture 2.4.

From Table 7, it seems that values of $R(x; 25; 1 + 5k)$ for $k = 0, 1, 2, 3, 4$ become closer as the value of x increases. This behavior is consistent with Conjecture 2.5.

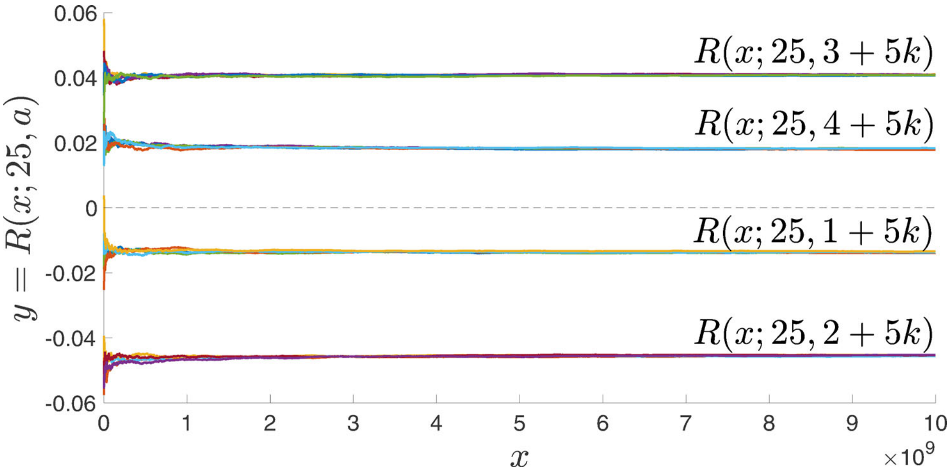

Figure 8. Plot of $R(x; 25, a)$ for all reduced residues a (mod 25) and $x \le 10^{10}$. The line $y = 0$ is marked with a dashed line. In the figure, one can see four "solid lines." However, each "line" is overlap of plots of $R(x; 25, a)$ for five different values of a. The top "line" is composed of plots of $R(x; 25, 3)$, $R(x; 25, 8)$, $R(x; 25, 13)$, $R(x; 25, 18)$, and $R(x; 25, 23)$. The figure is scaled down by a factor of 5 compared to Figures 3 and 4.

4. Probabilistic models for bias terms in prime running functions

We study probabilistic models for "random primes" which can model prime gaps and prime running functions (mod d). We show that modified Cramér models (defined in Section 4.1) produce bias terms of order $x/\log x$ associated with the prime running functions.

4.1. Modified Cramér models

The original probabilistic model of Cramér [4, 5] picks independently for each integer $n \ge 3$ to be "$\mathfrak{C}$-prime" with probability $\frac{1}{\log n}$. The Cramér model seems to accurately predict many statistics on primes. For example, the Cramér model predicts that $|\pi(x) - \mathrm{Li}(x)|$ lies within the predicted range by the Riemann hypothesis. However, it does not account for arithmetic restrictions on prime gaps and primes in arithmetic progressions. For example, almost all sample sequences of $\mathfrak{C}$-primes contain infinitely many gaps of size 1 between consecutive $\mathfrak{C}$-primes and contain infinitely many even numbers as $\mathfrak{C}$-primes.

We study a modified version of the Cramér model for the distribution of primes, that imposes initial sieving by an integers $Q \ge 2$ called the *sieve modulus*, followed by a probability model imposed on the unsieved elements. The initial sieving builds in arithmetic restrictions. In this model, we let integer n with $\gcd(n, Q) = 1$ be a "$\mathfrak{C}_Q$-prime" with probability $\frac{c_Q}{\log n}$ where c_Q is the prefactor

$$c_Q := \frac{Q}{\varphi(Q)} = \prod_{p|Q}\left(1 + \frac{1}{p-1}\right). \tag{4.1}$$

The prefactor c_Q quantifies the increased chance to be prime after the initial sieving. Modifications of Cramér models that make such an initial sieving were suggested in 1995 by Granville [9]. They were later studied by Pintz [20].

Formally, for fixed integer $Q \ge 2$, we define a sequence of independent Bernoulli random variables $Z_{n,Q}$ by

$$\Pr[Z_{n,Q} = 1] = \begin{cases} \frac{c_Q}{\log n} & \gcd(n, Q) = 1, \\ 0 & \gcd(n, Q) \ne 1. \end{cases} \tag{4.2}$$

If $\frac{c_Q}{\log n}$ from (4.2) exceeds 1, then we replace it by 1, a change that affects only finitely many values of n. If $Z_{n,Q} = 1$ then we say that n is a $\mathfrak{C}_Q$- *prime*.

In the modified Cramér model, we can define a random variable version of the prime running functions for these moduli d that divide the sieve modulus Q.

Definition 4.1. The *conditional gap* $W_{n,Q}$ is a random variable defined as a function of random variables $Z_{n,Q}, Z_{n+1,Q}, \dots$

$$W_{n,Q} := \begin{cases} m - n & \text{if } Z_{n,Q} = 1 \text{ and } Z_{n+1,Q} = Z_{n+2,Q} = \cdots = Z_{m-1,Q} = 0 \text{ and } Z_{m,Q} = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4.3}$$

We call $W_{n,Q}$ the *conditional gap* because if n is a $\mathfrak{C}_Q$-prime, then the value of $W_{n,Q}$ will equal the difference between n and the next $\mathfrak{C}_Q$-prime.

Definition 4.2 (Random prime running function). Let $Q \geq 2$ be an integer divisible by d. For fixed $x > 0$, we define the *random prime running function* $\tilde{\Phi}_Q(x; d, a)$ with Q as a sieve modulus to be a random variable

$$\tilde{\Phi}_Q(x; d, a) := \sum_{\substack{1 \leq n \leq x \\ n \equiv a \pmod{d}}} W_{n,Q}. \tag{4.4}$$

Definition 4.2 parallels the definition of prime running function in (2.1) in that they both sum over prime gaps (resp. $\mathfrak{C}_Q$ prime gaps) with smaller prime restricted to an arithmetic progression.

The function $\tilde{\Phi}_Q(x; d, a)$ is of interest when d divides Q and $\gcd(a, d) = 1$.

4.2. Modified Cramér model: expected value of the random prime running function

We demonstrate that the modified Cramér model, on average, predicts that prime running functions have a bias of the order $\frac{x}{\log x}$.

In what follows,

$$[n]_Q \equiv n \pmod{Q},\ 1 \leq [n]_Q \leq Q. \tag{4.5}$$

So $[n]_Q$ is least positive residue $\pmod{Q}$.

Theorem 4.3. *Fix an integer $d \geq 2$ and integer a such that $(a, d) = 1$. For the modified Cramér model with a fixed sieve modulus Q divisible by d, one has*

$$\mathbb{E}[\tilde{\Phi}_Q(x; d, a)] = \frac{x}{\varphi(d)} + R_Q(d; a)\frac{x}{\log x} + \mathcal{O}\left(\frac{x}{(\log x)^2}\right) \quad \text{as } x \to \infty$$

The bias constant $R_Q(d; a)$ is given by

$$R_Q(d; a) = R_Q^*(d; a) - \bar{R}_Q(d), \tag{4.6}$$

where

$$R_Q^*(d; a) := \frac{1}{\varphi(Q)^2} \sum_{\substack{s,t=1 \\ (st,Q)=1 \\ s \equiv a \pmod{d}}}^{Q} [t - s]_Q, \tag{4.7}$$

and

$$\bar{R}_Q(d) = \frac{1}{\varphi(d)} \frac{Q}{\varphi(Q)} \frac{\varphi(Q) + 1}{2}. \tag{4.8}$$

Proof. First, we recall the definition of the prime running function for a sample of the modified Cramér model, as a function of its random variables $Z_{i,Q}$. It is

$$\tilde{\Phi}_Q(x; d, a) = \sum_{\substack{n \leq x \\ n \equiv a \pmod{d}}} W_{n,Q}. \tag{4.9}$$

By linearity of expected values, it is sufficient to analyze the behavior of expected value of the conditional gaps $W_{n,Q}$ (Definition 4.1). By definition of expected value over a discrete space,

$$\mathbb{E}[W_{n,Q}] = \sum_{v} v \Pr[W_{n,Q} = v]. \tag{4.10}$$

The values v in (4.10) range over values of $W_{n,Q}$, which are the differences between two consecutive $\mathfrak{C}_Q$-primes. Since only positive integers co-prime to Q have a non-zero probability of being $\mathfrak{C}_Q$-prime, it is helpful to introduce a notation for the unsieved integers. Let U_Q be the set of the unsieved positive integers, i.e.

$$U_Q = \{1 = u_1 < u_2 < u_3, \ldots\} := \{u \in \mathbb{N} \mid \gcd(u, Q) = 1\}.$$

Since the random variables $\{Z_{k,Q}\}_{k=1}^{\infty}$ are independent, for $u_{i+l} > u_i$ we recover that

$$\Pr[W_{u_i,Q} = u_{i+l} - u_i] = \frac{c_Q}{\log u_i} \frac{c_Q}{\log u_{i+l}} \prod_{0<j<l} \left(1 - \frac{c_Q}{\log u_{i+j}}\right), \tag{4.11}$$

where $c_Q = \frac{Q}{\varphi(Q)}$ as defined in (4.1).

By substituting (4.11) into right hand side of (4.10), we conclude that

$$\mathbb{E}[W_{u_i,Q}] = \frac{c_Q{}^2}{\log u_i} \sum_{l>0} \left[\frac{u_{i+l} - u_i}{\log u_{i+l}} \prod_{0<j<l} \left(1 - \frac{c_Q}{\log u_{i+j}}\right) \right]. \tag{4.12}$$

While (4.12) gives us the exact value, it is difficult to work with. We proceed to approximating the expected value of $W_{u_i,Q}$ to a more convenient form.

Lemma 4.4. *Fix an integer $Q \geq 2$ and constant $c > 0$. Let m be a non-negative integer. Let u_i denote the i^{th} smallest positive integer co-prime to Q. Define*

$$T_1^m(n) = \sum_{k>0} \frac{(u_{n+k} - u_n)^m}{\log u_n} \left(1 - \frac{c}{\log u_n}\right)^{k-1},$$

$$T_2^m(n) = \sum_{k>0} \left(\frac{(u_{n+k} - u_n)^m}{\log u_{n+k}} \prod_{0<j<k} \left(1 - \frac{c}{\log u_{n+j}}\right) \right).$$

Then

$$T_2^m(n) = T_1^m(n) + \mathcal{O}\left(\frac{\log(n)^{m+\varepsilon}}{n}\right)$$

for any fixed $\varepsilon > 0$ as n tends to infinity.

The proof of Lemma 4.4 is postponed to Appendix A.

We continue the proof of Theorem 4.3. By substituting $c = c_Q$ and $m = 1$ into Lemma 4.4, we obtain

$$\mathbb{E}[W_{u_i,Q}] = \frac{c_Q{}^2}{\log u_i} \sum_{l>0} \left[\frac{u_{i+l} - u_i}{\log u_i} \left(1 - \frac{c_Q}{\log u_i}\right)^{l-1} \right] + \mathcal{O}\left(\frac{(\log u_i)^{\varepsilon}}{u_i}\right)$$

for any fixed $\varepsilon > 0$ as u_i tends to infinity.

To further simplify Lemma 4.4, we separate u_{i+l} into individual residue classes (mod Q).

$$\mathbb{E}[W_{u_i,Q}] = \frac{c_Q{}^2}{(\log u_i)^2} \sum_{h=1}^{\varphi(Q)} \sum_{l\geq 0} (u_{i+\varphi(Q)l+h} - u_i) \left(1 - \frac{c_Q}{\log u_i}\right)^{\varphi(Q)l+h-1} + \mathcal{O}\left(\frac{(\log u_i)^{\varepsilon}}{u_i}\right).$$

Now let $\alpha_i = 1 - \frac{c_Q}{\log u_i}$ and obtain

$$\mathbb{E}[W_{u_i,Q}] = \frac{c_Q{}^2}{(\log u_i)^2} \sum_{h=1}^{\varphi(Q)} \alpha_i^{h-1} \sum_{l\geq 0} (u_{i+h} - u_i + Ql)\alpha_i^{\varphi(Q)l} + \mathcal{O}\left(\frac{(\log u_i)^{\varepsilon}}{u_i}\right). \tag{4.13}$$

We utilize moments of a geometric distributed random variable Y_p with parameter $p \in (0, 1]$.

$$\mathbb{E}[Y_p^0] = \sum_{h=1}^{\infty} p(1-p)^{h-1} = 1, \tag{4.14}$$

$$\mathbb{E}[Y_p] = \sum_{h=1}^{\infty} hp(1-p)^{h-1} = \frac{1}{p}, \tag{4.15}$$

$$\mathbb{E}[Y_p^2] = \sum_{h=1}^{\infty} h^2 p(1-p)^{h-1} = \frac{2-p}{p^2}. \tag{4.16}$$

More specifically consider $Y_{1-\alpha_i^{\varphi(Q)}}$. Substituting the definition of moments to (4.13), we obtain

$$\mathbb{E}[W_{u_i,Q}] = \frac{c_Q{}^2}{(\log u_i)^2} \sum_{h=1}^{\varphi(Q)} \alpha_i^{h-1} \left(\frac{u_{i+h} - u_i}{1 - \alpha_i^{\varphi(Q)}} \mathbb{E}\left[Y^0_{1-\alpha_i^{\varphi(Q)}}\right] + \frac{Q\alpha_i^{\varphi(Q)}}{1 - \alpha_i^{\varphi(Q)}} \mathbb{E}\left[Y_{1-\alpha_i^{\varphi(Q)}}\right] \right) + \mathcal{O}\left(\frac{(\log u_i)^{\varepsilon}}{u_i}\right). \tag{4.17}$$

By substituting (4.14) and (4.15) into right hand side of (4.17), we obtain that

$$\mathbb{E}[W_{u_i,Q}] = \frac{{c_Q}^2}{(\log u_i)^2} \sum_{h=1}^{\varphi(Q)} \alpha_i^{h-1} \left[\frac{u_{i+h} - u_i}{1 - \alpha_i^{\varphi(Q)}} + \frac{Q\alpha_i^{\varphi(Q)}}{\left(1 - \alpha_i^{\varphi(Q)}\right)^2} \right] + \mathcal{O}\left(\frac{(\log u_i)^{\varepsilon}}{u_i} \right). \tag{4.18}$$

We further simplify (4.18) using the following series expansions.

$$\alpha_i^k = \left(\frac{\log u_i - c_Q}{\log u_i} \right)^k = 1 - \frac{kc_Q}{\log u_i} + \frac{k(k-1){c_Q}^2}{2(\log u_i)^2} + \mathcal{O}\left((\log u_i)^{-3}\right), \tag{4.19}$$

$$\frac{1}{1 - \alpha_i^{\varphi(Q)}} = \frac{\log u_i}{Q} \left(1 - \frac{(\varphi(Q) - 1)c_Q}{2\log u_i} + \mathcal{O}\left(\frac{1}{(\log u_i)^2} \right) \right)^{-1} = \frac{\log u_i}{Q} + \frac{\varphi(Q) - 1}{2\varphi(Q)} + \mathcal{O}\left((\log u_i)^{-1}\right), \tag{4.20}$$

$$\frac{\alpha_i^{\varphi(Q)}}{\left(1 - \alpha_i^{\varphi(Q)}\right)^2} = \frac{1}{\left(1 - \alpha_i^{\varphi(Q)}\right)^2} - \frac{1}{1 - \alpha_i^{\varphi(Q)}} = \frac{(\log u_i)^2}{Q^2} - \frac{\log u_i}{Q\varphi(Q)} + \mathcal{O}(1). \tag{4.21}$$

By substituting the series expansions (4.19)–(4.21) into the right-hand side of (4.18), we obtain the following equation.[2]

$$\begin{aligned}\mathbb{E}[W_{u_i,Q}] = \frac{{c_Q}^2}{(\log u_i)^2} \sum_{h=1}^{\varphi(Q)} \Bigg[&\left(1 - \frac{(h-1)c_Q}{\log u_i} + \mathcal{O}\left(\frac{1}{(\log u_i)^2} \right) \right) \\ &\times \left(\frac{(\log u_i)^2}{Q} + \left(\frac{u_{i+l} - u_i}{Q} - \frac{1}{\varphi(Q)} \right) \log u_i + \mathcal{O}(1) \right) \Bigg] + \mathcal{O}\left(\frac{(\log u_i)^{\varepsilon}}{u_i} \right).\end{aligned} \tag{4.22}$$

(4.22) simplifies to the following.

$$\mathbb{E}[W_{u_i,Q}] = c_Q + \frac{{c_Q}^2}{\log u_i} \sum_{h=1}^{\varphi(Q)} \left(\frac{u_{i+h} - u_i}{Q} - \frac{h}{\varphi(Q)} \right) + \mathcal{O}\left(\frac{1}{(\log u_i)^2} \right). \tag{4.23}$$

Note that the projection of $\{u_{i+h} : h = 1, 2, \ldots, \varphi(Q)\}$ to $(\mathbb{Z}/Q\mathbb{Z})^{\times}$ is a bijection. Also note that $1 \leq u_{i+h} - u_i \leq Q$ for $1 \leq h \leq \varphi(Q)$. Thus if $u_i \equiv s \pmod{Q}$, then

$$\sum_{h=1}^{\varphi(Q)} \left(\frac{u_{i+h} - u_i}{Q} - \frac{h}{\varphi(Q)} \right) = -\frac{\varphi(Q) + 1}{2} + \frac{1}{Q} \sum_{1 \leq t \leq Q, (t,Q)=1} [t - s]_Q. \tag{4.24}$$

Summing these contributions in (4.23) yields

$$\sum_{\substack{u_i \leq x \\ u_i \equiv s \ (\mathrm{mod}\ Q)}} \mathbb{E}[W_{u_i,Q}] = \frac{x}{\varphi(Q)} + \frac{Q}{\varphi(Q)^2} \left(-\frac{\varphi(Q) + 1}{2} + \frac{1}{Q} \sum_{\substack{1 \leq t \leq Q \\ (t,Q)=1}} [t - s]_Q \right) \frac{x}{\log x} + \mathcal{O}\left(\frac{x}{(\log x)^2} \right). \tag{4.25}$$

Finally, summing (4.25) over all $s \equiv a \pmod{d}$ for $s = 1, 2, \ldots Q$ that are co-prime to Q, we get

$$\mathbb{E}[\Phi(x; q, a)] = \sum_{\substack{u_i \leq x \\ u_i \equiv a \ (\mathrm{mod}\ d)}} \mathbb{E}[W_{u_i,Q}] = \frac{x}{\varphi(d)} + \left(R_Q^*(d; a) - \frac{1}{\varphi(d)} \frac{Q}{\varphi(Q)} \frac{\varphi(Q) + 1}{2} \right) \frac{x}{\log x} + \mathcal{O}\left(\frac{x}{(\log x)^2} \right). \qquad \square$$

4.3. *Modified Cramér model: variance of the random prime running function*

The next theorem shows that the probability distribution is centered around the mean value with a standard deviation of scale at most $\sqrt{x \log x}$. Note that the standard deviation is significantly smaller than the order of bias $\frac{x}{\log x}$.

Theorem 4.5. *Fix an integer $d \geq 2$ and an integer a such that $(a, d) = 1$. Then*

$$\mathrm{Var}(\tilde{\Phi}_Q(x; d, a)) = \mathcal{O}\left(x \log x\right). \tag{4.26}$$

[2] For fixed Q, we only sum over finite number of terms in (4.18). Thus the constants for the Big-O type bounds are bounded.

Proof. As with Theorem 4.3, let U_Q be the set of the unsieved integers and let c_Q be the prefactor, i.e.

$$U_Q = \{1 = u_1 < u_2 < u_3, \ldots\} := \{u \in \mathbb{N} \mid \gcd(u, Q) = 1\}$$

and $c_Q = \frac{Q}{\phi(Q)}$.

We first utilize the variance of sum of random variables formula.

$$\mathrm{Var}(\tilde{\Phi}_Q(x; q, a)) = \sum_{\substack{u_i \le x \\ u_i \equiv a \pmod d}} \mathrm{Var}(W_{u_i,Q}) + 2 \sum_{\substack{u_i < u_j \le x \\ u_i \equiv u_j \equiv a \pmod d}} \mathrm{Cov}(W_{u_i,Q}, W_{u_j,Q}).$$

We will bound the variance and the co-variance of $W_{u_i,Q}$ separately.

By definition of variance,

$$\mathrm{Var}(W_{u_i,Q}) = \mathbb{E}[W_{u_i,Q}^2] - \mathbb{E}[W_{u_i,Q}]^2 \le \mathbb{E}[W_{u_i,Q}^2].$$

It follows from (4.11) that

$$\mathbb{E}[W_{u_i,Q}^2] = \frac{{c_Q}^2}{\log u_i} \sum_{l>0} \left[\frac{(u_{i+l} - u_i)^2}{\log u_{i+l}} \prod_{0<j<l} \left(1 - \frac{c_Q}{\log u_{i+j}}\right) \right]. \tag{4.27}$$

By Lemma 4.4 and the inequality $u_{i+l} - u_i \le Ql$, (4.27) simplifies to

$$\mathrm{Var}(W_{u_i,Q}) \le \frac{{c_Q}^2 Q^2}{\log u_i} \sum_{l>0} \left[\frac{l^2}{\log u_i} \left(1 - \frac{c_Q}{\log u_i}\right)^{l-1} \right] + \mathcal{O}\left(\frac{(\log u_i)^{1+\epsilon}}{u_i} \right). \tag{4.28}$$

Letting Y_p be a geometrically distributed random variable with parameter $p = \frac{c_Q}{\log u_i}$. By substituting equation for $\mathbb{E}[Y_p^2]$ into (4.28), we obtain that

$$\mathrm{Var}(W_{u_i,Q}) \le \frac{c_Q Q^2}{\log u_i} \mathbb{E}[Y_p^2] + \mathcal{O}\left(\frac{(\log u_i)^{1+\epsilon}}{u_i} \right). \tag{4.29}$$

By second moment of geometric distribution (4.16), we obtain that

$$\mathrm{Var}(W_{u_i,Q}) = \mathcal{O}\left(\log u_i\right). \tag{4.30}$$

Thus there exists a constant $C > 0$ such that $\mathrm{Var}(W_{u_i,Q}) \le C \log(u_i)$ for sufficiently large u_i. We obtain that

$$\sum_{\substack{u_i \le x \\ u_i \equiv a \pmod d}} \mathrm{Var}(W_{u_i,Q}) \le Cx \log x + \mathcal{O}(1). \tag{4.31}$$

We now bound the covariance terms. Suppose that $i < j$. We will split the covariance into parts by conditioning on different events.

$$\mathrm{Cov}(W_{u_i,Q}, W_{u_j,Q}) = \sum_{k=1}^{3} \mathbb{E}[W_{u_i,Q} W_{u_j,Q} | E_k] \mathbb{P}(E_k) - \mathbb{E}[W_{u_i,Q}] \mathbb{E}[W_{u_j,Q}], \tag{4.32}$$

where E_1, E_2, E_3 are events $W_{u_i,Q} = u_j - u_i$, $W_{u_i,Q} < u_j - u_i$, $W_{u_i,Q} > u_j - u_i$, respectively. Suppose $W_{u_i,Q} > u_j - u_i$ (E_3). Then u_j cannot be $\mathfrak{C}_Q$-prime. Such event implies that $W_{u_j,Q} = 0$. Thus $\mathbb{E}[W_{u_i,Q} W_{u_j,Q} | E_3] = 0$. Now suppose that $W_{u_i,Q} < u_j - u_i$ (E_2). Such event implies that $W_{u_i,Q}$ and $W_{u_j,Q}$ are (conditionally) independent. Thus $\mathbb{E}[W_{u_i,Q} W_{u_j,Q} | E_2] = \mathbb{E}[W_{u_i,Q} | E_2] \mathbb{E}[W_{u_j,Q}] \le \mathbb{E}[W_{u_i,Q}] \mathbb{E}[W_{u_j,Q}]$. By combining these two observations, we conclude that

$$\mathrm{Cov}(W_{u_i,Q}, W_{u_j,Q}) \le \mathbb{P}(W_{u_i,Q} = u_j - u_i) \mathbb{E}(W_{u_i,Q} W_{u_j,Q} | W_{u_i,Q} = u_j - u_i),$$

which simplifies to

$$\mathrm{Cov}(W_{u_i,Q}, W_{u_j,Q}) \le \mathbb{P}(W_{u_i,Q} = u_j - u_i)](u_j - u_i) \frac{\log u_j}{c_Q} \mathbb{E}[W_{u_j,Q}].$$

By (4.23), $\mathbb{E}[W_{u_j,Q}] = \mathcal{O}(1)$ and by (4.11), $\mathbb{P}(W_{u_i,Q} = u_j - u_i) = \mathcal{O}\left(\frac{1}{(\log u_i)^2} \left(1 - \frac{c_Q}{\log u_j}\right)^{j-i} \right)$. Thus there exists a constant $A > 0$ such that if $\frac{Q}{\log u_j} \le 1$, then

$$\mathrm{Cov}(W_{u_i,Q}, W_{u_j,Q}) \le A \frac{\log u_j}{(1 + \log u_i)^2} (u_j - u_i) \left(1 - \frac{c_Q}{\log u_j}\right)^{j-i}. \tag{4.33}$$

Note that instead of $(\log u_i)^2$ as the denominator in equation (4.33), we have $(1+\log u_i)^2$. This allows us to avoid dividing by 0 when $u_i = 1$. This inconvenience occurs because the probability that u_i is $\mathfrak{C}_Q$ prime is equal to $\frac{c_Q}{\log u_i}$ only if u_i is sufficiently large. By summing equation (4.33) over different values of u_i, u_j, we obtain

$$\sum_{\substack{u_i<u_j\le x\\ u_i\equiv u_j\equiv a \pmod d}} \operatorname{Cov}(W_{u_i,Q}, W_{u_j,Q}) \le A \sum_{\substack{u_i<u_j\le x\\ u_i\equiv u_j\equiv a \pmod d}} \frac{\log u_j}{(1+\log u_i)^2}(u_j-u_i)\left(1-\frac{c_Q}{\log u_j}\right)^{j-i} + \mathcal{O}(1). \tag{4.34}$$

By utilizing (A.8), $\log u_j \le \log x$, and adding additional non-negative terms, (4.34) simplifies to

$$\sum_{\substack{u_i<u_j\le x\\ u_i\equiv u_j\equiv a \pmod d}} \operatorname{Cov}(W_{u_i,Q}, W_{u_j,Q}) \le AQ \sum_{u_i\le x} \frac{\log x}{(1+\log u_i)^2} \sum_{h=1}^{\infty} h\left(1-\frac{c_Q}{\log x}\right)^{h} + \mathcal{O}(1). \tag{4.35}$$

By setting $Y_{\frac{c_Q}{\log x}}$ to be a geometric random variable with parameter $p = \frac{c_Q}{\log x}$ and substituting the definition of $\mathbb{E}\left[Y_{\frac{c_Q}{\log x}}\right]$ into (4.35), we obtain that

$$\sum_{\substack{u_i<u_j\le x\\ u_i\equiv u_j\equiv a \pmod d}} \operatorname{Cov}(W_{u_i,Q}, W_{u_j,Q}) \le AQ \sum_{u_i\le x} \frac{\log x}{(1+\log u_i)^2} \frac{\log x}{c_Q}\mathbb{E}\left[Y_{\frac{c_Q}{\log x}}\right] + \mathcal{O}(1). \tag{4.36}$$

By (4.15) and the fact $\sum_{n\le x} \frac{1}{(1+\log n)^2} = \mathcal{O}\left(\frac{x}{(\log x)^2}\right)$, we obtain

$$\sum_{\substack{u_i<u_j\le x\\ u_i\equiv u_j\equiv a \pmod d}} \operatorname{Cov}(W_{u_i,Q}, W_{u_j,Q}) \le \mathcal{O}\left(x\log x\right). \tag{4.37}$$

□

4.4. Modified Cramér model: anti-symmetry properties

Theorem 4.6 suggests that the bias constant anti-symmetry Conjecture 2.4 should be true.

Theorem 4.6. *For any integer $d \ge 2$ and integer Q divisible by d, the following anti-symmetry holds.*

$$R_Q(d;-a) = -R_Q(d;a).$$

Proof. Fix an integer a co-prime to Q. By definition of bias constants,

$$R^*_Q(d;-a) = \frac{1}{\varphi(Q)^2} \sum_{\substack{s,t=1\\ s\equiv -a \pmod d\\ (st,Q)=1}}^{Q} [t-s]_Q.$$

Because $t \to -t$ is a permutation of $(\mathbb{Z}/Q\mathbb{Z})^\times$, we can sum over $-t$. Furthermore, $s \equiv a \pmod d$ implies $-s \equiv -a \pmod d$. Thus

$$R^*_Q(d;-a) = \frac{1}{\varphi(Q)^2} \sum_{\substack{s,t=1\\ s\equiv a \pmod d\\ (st,Q)=1}}^{Q} [s-t]_Q.$$

It follows that

$$R^*_Q(d;a) + R^*_Q(d;-a) = \frac{1}{\varphi(Q)^2} \sum_{\substack{s,t=1\\ s\equiv a \pmod d\\ (st,Q)=1}}^{Q} \left([t-s]_Q + [s-t]_Q\right). \tag{4.38}$$

Note that $[t-s]_Q + [s-t]_Q \equiv 0 \pmod{Q}$ and $[t-s]_Q + [s-t]_Q \in [2, 2Q]$. Thus

$$[t-s]_Q + [s-t]_Q = \begin{cases} Q & t \not\equiv s \pmod{Q}, \\ 2Q & t \equiv s \pmod{Q}. \end{cases} \tag{4.39}$$

By counting the number of times $s = t$ in (4.38) and utilizing (4.39), we obtain that

$$R_Q^*(d;a) + R_Q^*(d;-a) = \frac{1}{\varphi(Q)^2}\left(Q\frac{\varphi(Q)^2}{\varphi(d)} + Q\frac{\varphi(Q)}{\varphi(d)}\right) = 2\bar{R}_Q(d). \tag{4.40}$$

□

As an immediate corollary, we obtain that the bias constants add up to 0.

Corollary 4.7. *For $d \geq 2$ and Q divisible by d,*

$$\sum_{\substack{a=1 \\ (a,d)=1}}^{d} R_Q(d;a) = 0 \tag{4.41}$$

and

$$\bar{R}_Q(d) = \frac{1}{\varphi(d)} \sum_{\substack{a=1 \\ (a,d)=1}}^{d} R_Q^*(d;a). \tag{4.42}$$

4.5. Modified Cramér model: radical equivalence property

Theorem 4.8. *For all $d \geq 2$ with $(a,d) = 1$ and $d|Q$,*

$$R_Q(d;a) = \frac{\varphi(d_{\mathrm{sf}})}{\varphi(d)} R_Q(d_{\mathrm{sf}}, a), \tag{4.43}$$

where $d_{\mathrm{sf}} = \mathrm{rad}(d)$ is the maximal square-free divisor of d. Equivalently,

$$R_Q(d;a) = R_Q(d;a') \tag{4.44}$$

if $a \equiv a' \pmod{d_{\mathrm{sf}}}$.

Proof. Note that for any fixed sample sequence of $\mathfrak{C}_Q$ primes,

$$\tilde{\Phi}_Q(d_{\mathrm{sf}};a) = \sum_{\substack{a'=1 \\ a' \equiv a \pmod{d_{\mathrm{sf}}}}}^{d} \tilde{\Phi}_Q(d;a'). \tag{4.45}$$

By linearity of expected value and inspecting the $\frac{x}{\log x}$ order term from Theorem 4.3, we obtain that

$$R_Q(d_{\mathrm{sf}};a) = \sum_{\substack{a'=1 \\ a' \equiv a \pmod{d_{\mathrm{sf}}}}}^{d} R_Q(d;a'). \tag{4.46}$$

Thus (4.43) and (4.44) are equivalent. Fix a and a' such that $a \equiv a' \pmod{d_{\mathrm{sf}}}$. Let $Q_{\mathrm{sf}} = \mathrm{rad}(Q)$ denote the square-free part of Q. Because d_{sf} divides Q_{sf} and $Q_{\mathrm{sf}}/d_{\mathrm{sf}}$ is co-prime to d, there exists some integer k such that $a + kQ_{\mathrm{sf}} = a' \pmod{d}$. Thus it suffices to show that $R_Q(d;a) = R_Q(d;a+Q_{\mathrm{sf}})$ for any fixed a. By definition of bias constants given in Theorem 4.3,

$$R_Q(d;a+Q_{\mathrm{sf}}) = -\bar{R}_Q(d) + \frac{1}{\varphi(Q)^2} \sum_{\substack{s,t=1 \\ (st,Q)=1 \\ s \equiv a+Q_{\mathrm{sf}} \pmod{d}}}^{Q} [t-s]_Q.$$

Since $[t-s]_Q$ only depends on value of $t - s \pmod{Q}$, we can sum over $s \equiv a \pmod{d}$ and then add Q_{sf} to s.

$$R_Q(d;a+Q_{\mathrm{sf}}) = -\bar{R}_Q(d) + \frac{1}{\varphi(Q)^2} \sum_{\substack{s,t=1 \\ (st,Q)=1 \\ s \equiv a \pmod{d}}}^{Q} [t-(s+Q_{\mathrm{sf}})]_Q,$$

$$R_Q(d;a+Q_{\mathrm{sf}}) = -\bar{R}_Q(d) + \frac{1}{\varphi(Q)^2}\sum_{\substack{s,t=1\\(st,Q)=1\\ s\equiv a \pmod d}}^{Q} [(t-Q_{\mathrm{sf}})-s]_Q,$$

$$R_Q(d;a+Q_{\mathrm{sf}}) = -\bar{R}_Q(d) + \frac{1}{\varphi(Q)^2}\sum_{\substack{s,t=1\\(s(t+Q_{\mathrm{sf}}),Q)=1\\ s\equiv a \pmod d}}^{Q} [t-s]_Q.$$

Well, for any integer t, and any prime factor p of Q, $t \equiv t+Q_{\mathrm{sf}} \pmod p$. Thus t is co-prime to Q if and only if $t+Q_{\mathrm{sf}}$ is co-prime to Q. It follows that

$$R_Q(d;a+Q_{\mathrm{sf}}) = -\bar{R}_Q(d) + \frac{1}{\varphi(Q)^2}\sum_{\substack{s,t=1\\(st,Q)=1\\ s\equiv a \pmod d}}^{Q} [t-s]_Q. \tag{4.47}$$

We are done because (4.47) is the definition of $R_Q(d;a)$. □

5. Computation for modified Cramér model

In this section, we compute the bias constants $R_Q(d;a)$ for the modified Cramér model for various values of Q and d.

5.1. Recursive formula for bias constants

Brute force computation of bias constant $R_Q(d;a)$ has runtime complexity that is polynomial in Q, which is exponential in input bit size $O(\log Q)$. The following result gives a recursive formula yielding an improved method for computing the bias constants $R_Q(d;a)$ for fixed d and all $a \pmod d$ with $(a,d)=1$.

Theorem 5.1. *Suppose $d, p, Q_0 \geq 2$ are pairwise co-prime and p is a prime. Let $Q = dQ_0$. Then*

$$R_{pQ}(d;pa) = \frac{\varphi(p)^2-1}{\varphi(p)^2} R_Q(d;pa) + \frac{p}{\varphi(p)^2} R_Q(d;a). \tag{5.1}$$

Definition 5.2. Given $Q_1,\dots,Q_k$ pairwise co-prime, we define $[n_1,\dots,n_k]_{Q_1,\dots,Q_k}$ to be the unique element in $[1, Q_1Q_2\dots Q_k]$ such that

$$[n_1,\dots,n_k]_{Q_1,\dots,Q_k} \equiv n_i \pmod{Q_i}, \quad i=1,\dots,k.$$

Note that the definition is consistent with the definition of least positive residue $[n]_Q$. Because $[n]_{Q_1Q_2\dots Q_k}$ is congruent to $n \pmod{Q_i}$ for $i=1,\dots,k$, we obtain

$$[n]_{Q_1Q_2\dots Q_k} = [n,\dots,n]_{Q_1,\dots,Q_k}. \tag{5.2}$$

Proof of Theorem 5.1. By Corollary 4.7,

$$R_{pQ}(d;pa) := \frac{1}{\varphi(pQ)^2}\sum_{\substack{s,t=1\\(st,pQ)=1\\ s\equiv a \pmod d}}^{pQ} [t-s]_{pQ} - \frac{1}{\varphi(pQ)^2\varphi(d)}\sum_{\substack{s,t=1\\(st,pQ)=1}}^{pQ} [t-s]_{pQ}. \tag{5.3}$$

By substituting (5.2) into (5.3), we obtain that

$$R_{pQ}(d;pa) = \frac{1}{\varphi(pQ)^2}\left(\sum_{\substack{s=1\\ s\equiv pa \pmod d\\(s,pQ)=1}}^{pQ}\ \sum_{\substack{t=1\\(t,pQ)=1}}^{pQ} [t-s,t-s]_{pQ_0,d} - \frac{1}{\varphi(d)}\sum_{\substack{s,t=1\\(st,pQ)=1}}^{pQ} [t-s,t-s]_{pQ_0,d}\right). \tag{5.4}$$

For each integer s co-prime to pQ, there are exactly $\phi(d)$ many integers $1 \leq s' \leq pQ$ that are co-prime to pQ and congruent to $s \pmod{pQ_0}$. Furthermore, if $s \equiv pa \pmod d$, then for each s' that satisfies the conditions above, $[t-s,t-s,]_{pQ_0,d} = [t-s',t-pa]_{pQ_0,d}$.

Thus we can remove the restriction $s \equiv pa \pmod d$ from first sum in (5.4) by replacing $[t-s, t-s]_{pQ_0,d}$ with $[t-s, t-pa]pQ_0, d$ and accounting for multiplicity.

By substituting (5.5) into (5.4),

$$R_{pQ}(d;pa) = \frac{1}{\varphi(pQ)^2\varphi(d)} \sum_{\substack{s,t=1\\(st,pQ)=1}}^{pQ} \Big([t-s,t-pa]_{pQ_0,d} - [t-s,t-s]_{pQ_0,d}\Big) \tag{5.5}$$

and similarly,

$$R_Q(d;a) = \frac{1}{\varphi(Q)^2\varphi(d)} \sum_{\substack{s,t=1\\(st,Q)=1}}^{Q} \Big([t-s,t-a]_{Q_0,d} - [t-s,t-s]_{Q_0,d}\Big). \tag{5.6}$$

We now decompose (5.5) by the decomposition $\big(\mathbb{Z}/pQ_0\mathbb{Z}\big)^\times \cong \big(\mathbb{Z}/p\mathbb{Z}\big)^\times \times (\mathbb{Z}/Q_0\mathbb{Z})^\times$.

$$R_{pQ}(d;pa) = \frac{1}{\varphi(pQ)^2\varphi(d)} \sum_{\substack{s,t=1\\(st,Q)=1}}^{Q} \sum_{s',t'=1}^{p-1} \Big([t-s,t-pa,t'-s']_{Q_0,d,p} - [t-s,t-s,t'-s']_{Q_0,d,p}\Big). \tag{5.7}$$

Note that for any $r_1, r_2, r_3 \in \mathbb{Z}$, $[r_1, r_2, r_3]_{Q_0,d,p} - [r_1, r_2]_{Q_0,d}$ is congruent to $0 \pmod Q$ and $r_3 - [r_1, r_2]_{Q_0,d} \pmod p$. By Chinese remainder theorem, for any fixed $r_1, r_2 \in \mathbb{Z}$,

$$r_3 \mapsto \frac{1}{Q}([r_1, r_2, r_3]_{Q_0,d,p} - [r_1, r_2]_{Q_0,d})$$

is a permutation on $\{0, 1, 2, \dots, p-1\}$. By further fixing $r_2' \in \mathbb{Z}$ and summing over the set $\{0, 1, \dots, p-1\}$, we conclude that

$$\sum_{r_3=0}^{p-1} \big([r_1, r_2, r_3]_{Q_0,d,p} - [r_1, r_2', r_3]_{Q_0,d,p}\big) = p([r_1, r_2]_{Q_0,d} - [r_1, r_2']_{Q_0,d}). \tag{5.8}$$

We apply (5.8) to (5.7) as we sum over s'.

$$\begin{aligned} R_{pQ}(d;pa) = {} & \frac{1}{\varphi(pQ)^2\varphi(d)} \sum_{\substack{s,t=1\\(st,Q)=1}}^{Q} \sum_{t'=1}^{p-1} p([t-s,t-pa]_{Q_0,d} - [t-s,t-s]_{Q_0,d}) \\ & - \frac{1}{\varphi(pQ)^2\varphi(d)} \sum_{\substack{s,t=1\\(st,Q)=1}}^{Q} \sum_{t'=1}^{p-1} ([t-s,t-pa,t']_{Q_0,d,p} - [t-s,t-s,t']_{Q_0,d,p}). \end{aligned}$$

We apply (5.8) once more by summing over t'.

$$\begin{aligned} R_{pQ}(d;pa) = {} & \frac{p(p-1)}{\varphi(pQ)^2\varphi(d)} \sum_{\substack{s,t=1\\(st,Q)=1}}^{Q} ([t-s,t-pa]_{Q_0,d} - [t-s,t-s]_{Q_0,d}) \\ & - \frac{1}{\varphi(pQ)^2\varphi(d)} \sum_{\substack{s,t=1\\(st,Q)=1}}^{Q} p([t-s,t-pa]_{Q_0,d} - [t-s,t-s]_{Q_0,d}) \\ & + \frac{1}{\varphi(pQ)^2\varphi(d)} \sum_{\substack{s,t=1\\(st,Q)=1}}^{Q} ([t-s,t-pa,0]_{Q_0,d,p} - [t-s,t-s,0]_{Q_0,d,p}). \end{aligned}$$

This simplifies to

$$\begin{aligned} R_{pQ}(d;pa) = {} & \frac{p(p-2)}{\varphi(p)^2} \frac{1}{\varphi(Q)^2\varphi(d)} \sum_{\substack{s,t=1\\(st,Q)=1}}^{Q} ([t-s,t-pa]_{Q_0,d} - [t-s,t-s]_{Q_0,d}) \\ & + \frac{1}{\varphi(pQ)^2\varphi(d)} \sum_{\substack{s,t=1\\(st,Q)=1}}^{Q} ([t-s,t-pa,0]_{Q_0,d,p} - [t-s,t-s,0]_{Q_0,d,p}). \end{aligned} \tag{5.9}$$

By (5.6) the first term of (5.9) is $\frac{\varphi(p)^2-1}{\varphi(p)^2}R_Q(d;pa)$.

$$R_{pQ}(d;pa) = \frac{\varphi(p)^2-1}{\varphi(p)^2}R_Q(d;pa) + \frac{1}{\varphi(pQ)^2\varphi(d)}\sum_{\substack{s,t=1\\(st,Q)=1}}^{Q}([t-s,t-pa,0]_{Q_0,d,p} - [t-s,t-s,0]_{Q_0,d,p}). \tag{5.10}$$

Note that multiplication by p is a permutation of $(\mathbb{Z}/Q_0\mathbb{Z})^\times$ and $(\mathbb{Z}/d\mathbb{Z})^\times$. Thus one could sum over ps and pt instead of s and t.

$$R_{pQ}(d;pa) = \frac{\varphi(p)^2-1}{\varphi(p)^2}R_Q(d;pa) + \frac{1}{\varphi(pQ)^2\varphi(d)}\sum_{\substack{s,t=1\\(st,Q)=1}}^{Q}([pt-ps,pt-pa,0]_{Q_0,d,p} - [pt-ps,pt-ps,0]_{Q_0,d,p}). \tag{5.11}$$

By the Chinese remainder theorem,

$$[pr_1,pr_2,0]_{Q_0,d,p} = p[r_1,r_2]_{Q_0,d}, \quad r_1,r_2\in\mathbb{Z}. \tag{5.12}$$

$$R_{pQ}(d;pa) = \frac{\varphi(p)^2-1}{\varphi(p)^2}R_Q(d;pa) + p\frac{1}{\varphi(pQ)^2\varphi(d)}\sum_{\substack{s,t=1\\(st,Q)=1}}^{Q}([t-s,t-a]_{Q_0,d} - [t-s,t-s]_{Q_0,d}). \tag{5.13}$$

On substituting (5.6), we conclude that

$$R_{pQ}(d;pa) = \frac{\varphi(p)^2-1}{\varphi(p)^2}R_Q(d;pa) + \frac{p}{\varphi(p)^2}R_Q(d;a). \tag{5.14}$$

□

5.2. Computation of modified Cramér bias constants

We compute bias constants $R_Q(d;a)$ utilizing the recursive algorithm in Theorem 5.1. For modulus $d=p$ a prime, the simplest case is $Q=d$, and the bias constant is given by

$$R_d(d;a) = \frac{d}{\varphi(d)^2}\left(\frac{a}{d}-\frac{1}{2}\right), \quad 1\le a\le d-1. \tag{5.15}$$

These constants $R_d(d;a)$ are increasing as a function of a for $1\le a\le d-1$. For $d=3,5,7$ $R_d(d;a)$ significantly differ from the empirical data on bias constants $R(x;d,a)$ given in Tables 3–5 in Section 3. The empirical data also disagrees in sign for $d=3$ and the constants oscillate in a for $d=5$ and $d=7$.

We now study the effect of larger sieve modulus Q on the modified Cramér bias constants, which seems to improve our numerical result. In particular, we consider the case of a modified Cramér model with an initial sieve over all the prime numbers less than or equal to T. We let our sieve modulus $Q=T\#$, where the *primorial at* T, is defined by

$$T\# := \prod_{p\le T} p. \tag{5.16}$$

The notation $T\#$ for primorials follows Caldwell and Gallot [2]. Thus $\tilde{\Phi}_{T\#}(x;d,a)$ is a random prime running function corresponding to the modified Cramér model with initial sieving by all primes less than or equal to T.

Tables 8–10 give values of Cramér bias constants at various primorials.

The bias constant for the expected values in these modified Cramér models with sieve modulus of $Q=1000\#$ exhibit numerical resemblance with the empirical data for $d=5$ and 7. However, for the case $d=3$, there are significant deviations from the empirical data.

Note that as T varies in these tables, the values of the constants $R_{T\#}(d;a)$ may be showing oscillations as T increases.

Table 8. The bias constant $R_Q(3;a)$ for various sieve moduli Q.

a \ Q	Cramér model bias constants $Q=3$	$Q=3\#$	$Q=10\#$	$Q=100\#$	$Q=1000\#$	rescaled bias function $R(10^{12};3,a)$
$a=1$	−0.125	0.25	0.1823	0.1599	0.1569	0.2022
$a=2$	0.125	−0.25	−0.1823	−0.1599	−0.1569	−0.2022

The right most column is the empirical data $R(10^{12};3,a)$.

Table 9. The bias constant $R_Q(5;a)$ for various sieve moduli Q.

a \ Q	Cramér model bias constants $Q=5$	$Q=5\#$	$Q=10\#$	$Q=100\#$	$Q=1000\#$	rescaled bias function $R(10^{12};5,a)$
$a=1$	−0.09375	−0.0938	−0.0547	−0.0699	−0.0685	−0.0703
$a=2$	−0.03125	−0.1875	−0.2005	−0.2027	−0.2043	−0.2211
$a=3$	0.03125	0.1875	0.2005	0.2027	0.2043	0.2059
$a=4$	0.09375	0.0938	0.0547	0.0699	0.0685	0.0855

The right most column is the empirical data $R(10^{12};5,a)$.

Table 10. The bias constant $R_Q(7;a)$ for various sieve moduli Q.

a \ Q	Cramér model bias constants $Q=7$	$Q=10\#$	$Q=100\#$	$Q=1000\#$	rescaled bias function $R(10^{12};7;a)$
$a=1$	−0.0964	0.1432	0.1303	0.1310	0.1461
$a=2$	−0.0417	−0.0781	−0.0749	−0.0753	−0.0680
$a=3$	−0.0139	0.0651	0.0554	0.0557	0.0506
$a=4$	0.0139	−0.0651	−0.0554	−0.0557	−0.0571
$a=5$	0.0417	0.0781	0.0749	0.0753	0.0626
$a=6$	0.0964	−0.1432	−0.1303	−0.1310	−0.1343

The right most column is the empirical data $R(10^{12};7,a)$.

6. Concluding remarks

Section 4 presents a modified Cramér model which exhibits a mechanism that can lead to biases of order $x/\log x$. Our data in Section 5 computes bias constants for this model for primorials $T\#$ that roughly agree with the empirical data in Section 3 for $d=5$ and $d=7$.

The choice of taking the sieve modulus Q to run through primorials $T\#$ in the modified Cramér model is significant. Based on the choice of the sequence of integers $\{S_i\}_{i=1}^{\infty}$ with $S_i|S_{i+1}$, $R_{S_i}(d;a)$ could diverge or converge to a value that depends on the choice of $\{S_i\}_{i=1}^{\infty}$. For example, fix $d\geq 2$ prime and choose a with $(a,d)=1$. Define

$$Q_T = d\prod_{\substack{p\leq T\\ p\equiv 1 \pmod d}} p.$$

By Theorem 5.1,

$$R_{Q_T}(d;a) = \left(\prod_{\substack{p\leq T\\ p\equiv 1 \pmod d}} \frac{\phi(p)^2+p-1}{\phi(p)^2}\right) R_d(d;a) = \left(\prod_{\substack{p\leq T\\ p\equiv 1 \pmod d}} \frac{p}{p-1}\right) R_d(d;a). \tag{6.1}$$

It is known that

$$\left(\prod_{\substack{p\leq T\\ p\equiv 1 \pmod d}} \frac{p}{p-1}\right) \sim c(\log(x))^{1/\phi(d)}$$

for some constant $c>0$ (see [16], [22]). In particular, by taking Q_T as the sieve factor, the constants $R_{Q_T}(d;a)$ diverge as T grows to infinity.

We do not address the question of whether the bias constants $R_Q(d;a)$ produced by this model (letting $Q\to\infty$ through the primorials) will necessarily agree with the bias constants $R(d;a)$ asserted to exist in Conjecture 2.3.

We defined the prime running functions $\Phi(x;d,a)$ as summing gaps between primes $p_k \equiv a \pmod d$ below x and the next following prime p_{k+1}, up to x. However, one also consider the *reversed prime running functions* $\Phi^R(x;d,a)$ which puts instead a congruence condition on the upper endpoint of the interval $p_{k+1} \equiv a \pmod{d}$ and putting no congruence condition on p_k. By an analysis similar to that made in Section 4, the modified Cramér model predicts

$$\Phi^R(x; d, a) = \frac{1}{\phi(d)}x - R(d; a)\frac{x}{\log x} + o\left(\frac{x}{\log x}\right),$$

with the bias term having the opposite sign as for the prime running function.

A more refined analysis of the biases of prime running function and its generalizations can be done based on the Hardy–Littlewood k-tuple conjecture, following ideas in the paper of Lemke-Oliver and Soundararajan [17]. We leave this topic for future work.

Acknowledgments

The author made an initial observation similar to Figure 2 with Upamanyu Sharma, whom he thanks for the help in the initial computation of prime running functions. The author thanks J. C. Lagarias for mentoring this project, for advice on writing, and supplying references. The author thanks Corey Everlove and Djordje Milićević for helpful comments.

Declaration of interest

No potential conflict of interest was reported by the author(s).

Funding

This work was partially supported by NSF grant DMS-1701576.

Appendix A: Proof of Lemma 4.4

Proof. We begin by decomposing $|T_1(n) - T_2(n)|$ into two parts using the triangle inequality.

$$\left|T_1^m(n) - T_2^m(n)\right| \le H_1^m(n) + H_2^m(n),$$

where

$$H_1^m(n) := \sum_{k>0} (u_{n+k} - u_n)^m \left(\frac{1}{\log u_n} - \frac{1}{\log u_{n+k}}\right)\left(1 - \frac{c}{\log u_n}\right)^{k-1}, \tag{A.1}$$

$$H_2^m(n) := \sum_{k>1} \frac{(u_{n+k} - u_n)^m}{\log u_{n+k}} \left[\prod_{j=1}^{k-1}\left(1 - \frac{c}{\log u_{n+j}}\right) - \left(1 - \frac{c}{(\log u_n)}\right)^{k-1}\right]. \tag{A.2}$$

Thus it suffices to show the following inequalities.

$$H_1^m(n) = \mathcal{O}\left(\frac{(\log n)^m}{n}\right), \tag{A.3}$$

$$H_2^m(n) = \mathcal{O}\left(\frac{(\log n)^{m+\epsilon}}{n}\right). \tag{A.4}$$

We first prove (A.3). Note that $\frac{d}{dx}\frac{1}{\log x} = -\frac{1}{x(\log x)^2}$ is decreasing in magnitude. Thus by mean value theorem,

$$\frac{1}{\log x} - \frac{1}{\log(x+t)} \le \frac{t}{x(\log x)^2}, \quad t \ge 0. \tag{A.5}$$

By substituting (A.5) into (A.1), we establish that

$$H_1^m(n) \le \sum_{k>0} \frac{(u_{n+k} - u_n)^{m+1}}{u_n(\log u_n)^2}\left(1 - \frac{c}{\log u_n}\right)^{k-1}. \tag{A.6}$$

Note that for any positive integer a, there exists some integer $u \in [a, a+Q)$ co-prime to Q. Thus the following holds

$$n \le u_n \le Qn, \tag{A.7}$$

$$k \le u_{n+k} - u_n \le Qk, \quad k = 0, 1, 2, \ldots. \tag{A.8}$$

By substituting (A.7) and (A.8) into (A.6), we obtain that

$$H_1^m(n) \le \frac{Q^{m+1}}{u_n \log u_n} \sum_{k>0} \frac{k^{m+1}}{\log u_n}\left(1 - \frac{c}{\log u_n}\right)^{k-1}. \tag{A.9}$$

Let Y_p be a geometric random variable with parameter $p = \frac{c}{\log u_n}$. By substituting the definition for the $m+1$th moment, we obtain that

$$H_1^m(n) \leq \frac{Q^{m+1}}{cu_n \log u_n} \mathbb{E}\left[Y_p^{m+1}\right]. \tag{A.10}$$

We will use the moment generating function $M(t) = \mathbb{E}[\exp(tY_p)]$ to bound the growth of $\mathbb{E}\left[Y^{m+1}\right]$. By direct computation, $M(t) = \frac{pe^t}{1-e^t(1-p)}$. By utilizing the fact that $\frac{\partial^{m+1}}{\partial t^{m+1}} M(t)|_{t=0} = \mathbb{E}[Y_p^{m+1}]$, we conclude that

$$\mathbb{E}[Y_p^{m+1}] = \mathcal{O}\left(\frac{1}{p^{m+1}}\right)_{p\in(0,1]}. \tag{A.11}$$

By substituting (A.11) into (A.10) and $p = \frac{c}{\log u_n}$, we obtain that

$$H_1^m(n) = \mathcal{O}\left(\frac{\log(u_n)^m}{u_n}\right). \tag{A.12}$$

Now all we have left is to prove (A.4). Note that for any $k > 0$,

$$\left(1 - \frac{c}{\log u_n}\right)^{k-1} \leq \prod_{j=1}^{k-1}\left(1 - \frac{c}{\log u_{n+j}}\right) \leq \left(1 - \frac{c}{\log u_{n+k}}\right)^{k-1}.$$

Thus

$$H_2^m(n) \leq \sum_{k>0} \frac{(u_{n+k} - u_n)^m}{\log u_{n+k}} \left[\left(1 - \frac{c}{\log u_{n+k}}\right)^{k-1} - \left(1 - \frac{c}{\log u_n}\right)^{k-1}\right]. \tag{A.13}$$

Note that $\frac{d}{dx}\left(1 - \frac{c}{\log x}\right)^k = \frac{ck}{x(\log x)^2}\left(1 - \frac{c}{\log x}\right)^{k-1}$. Thus the derivative of the function $\left(1 - \frac{c}{\log x}\right)^k$ is bounded above by $\frac{ck}{a(\log a)^2}\left(1 - \frac{c}{\log b}\right)^{k-1}$ over the interval $x \in [a, b]$. By mean value theorem, we establish that for $e^c < a < b$,

$$\left(1 - \frac{c}{\log b}\right)^k - \left(1 - \frac{c}{\log a}\right)^k \leq \frac{ck(b-a)}{a(\log a)^2}\left(1 - \frac{c}{\log b}\right)^{k-1}. \tag{A.14}$$

By substituting (A.14) into (A.13), we establish that for sufficiently large n,

$$H_2^m(n) \leq c \sum_{k>0} \frac{(u_{n+k} - u_n)^{m+1}(k-1)}{\log(u_{n+k}) u_n (\log u_n)^2}\left(1 - \frac{c}{\log u_{n+k}}\right)^{k-2}. \tag{A.15}$$

By substituting (A.7) and (A.8) into (A.15), we obtain that

$$H_2^m(n) \leq \frac{cQ^{m+1}}{n(\log n)^3} \sum_{k>0} k^{m+1}(k-1)\left(1 - \frac{c}{\log(Q(n+k))}\right)^{k-2}.$$

By noting that $\left(1 - \frac{c}{\log(Q(n+k))}\right)^{-2} \leq 2$ for sufficiently large n, we obtain that

$$H_2^m(n) \leq 2\frac{cQ^{m+1}}{n(\log n)^3} \sum_{k>0} k^{m+2}\left(1 - \frac{c}{\log(Q(n+k))}\right)^k,$$

for sufficiently large n.

Because $1 - x \leq e^{-x}$ for all $x \in \mathbb{R}$, we know that

$$H_2^m(n) \leq 2\frac{cQ^{m+1}}{n(\log n)^3} \sum_{k>0} k^{m+2} e^{-ck/\log(Q(n+k))} \tag{A.16}$$

for sufficiently large n. Let $P = \frac{m+3}{m+3+\epsilon}$. Since $P < 1$, there exists a constant $C_P > 0$ such that for all sufficiently large n,

$$\frac{k}{\log(Q(n+k))} \geq \frac{k^P}{C_P \log n}, \quad k = 1, 2, \ldots. \tag{A.17}$$

Thus for sufficiently large n,

$$H_2^m(n) \leq 2\frac{cQ^{m+1}}{n(\log n)^3} \sum_{k>0} k^{m+2} e^{-c\frac{k^P}{C_a \log n}}, \tag{A.18}$$

$$H_2^m(n) \leq 2\frac{cQ^{m+1}}{n(\log n)^3} \int_0^\infty (t+1)^{m+2} e^{-c\frac{t^P}{C_P \log n}}\, dt. \tag{A.19}$$

By substituting $u = t^a$, we obtain that

$$\int_0^\infty (t+1)^{m+2} e^{-c\frac{t^P}{C_P \log n}}\, dt = \mathcal{O}\left((\log n)^{m+3+\epsilon}\right). \tag{A.20}$$

By substituting (A.20) into (A.19), we conclude (A.4). □

References

[1] Bombieri, E., Gubler, W. (2007). *Heights in Diophantine Geometry*. New Mathematical Monographs. Cambridge: Cambridge University Press, p. 402. DOI: 10.1017/CBO9780511542879
[2] Caldwell, C. K., Gallot, Y. (2002). On the primality of $n! \pm 1$ and $2 \cdot 3 \cdot 5 \cdots p \pm 1$. *Math. Comput.* 71(237): 441–448.
[3] Chebyshev, P. L. (1853). Lettre de M. le Professeur Tchébychev á M. Fuss sur un nouveaux théorème relatif aux nombres premiers contenus dans les formes $4n + 1$ et $4n + 3$. *Bull. Classe Phys. Acad. Imp. Sci. St. Petersb*. 11:208.
[4] Cramér, H. (1935). Prime numbers and probability. *Skand. Math. Kongr.* 8:107–115.
[5] Cramér, H. (1936). On the order of magnitude of the differences between consecutive prime numbers. *Acta Arith*. 2:23–46.
[6] Davenport, H. 2000. *Multiplicative Number Theory*. Graduate Texts in Mathematics, Vol. 74, 3rd ed. (Revised and with a preface by Hugh L. Montgomery). New York, NY: Springer-Verlag, +177 pp.
[7] Feuerverger, A., Martin, G. (2000). Biases in the Shanks-Renyi prime number race. *Exp. Math.* 9(4): 535–570.
[8] Fiorilli, D. (2014). Highly biased prime number races. *Algebra Number Theory*. 8: 1733–1767.
[9] Granville, A. (1995). Harald Cramér and the distribution of prime numbers. *Scand. Actuar. J.* 1995(1): 12–28.
[10] Granville, A., Martin, G. (2006). Prime number races. *Amer. Math. Monthly*. 113(1): 1–33.
[11] Hardy, G. H., Littlewood, J. E. (1916). Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes. *Acta Math.* 41(1): 141–148.
[12] Huxley, M. (1972). *The Distribution of Prime Numbers. Large Sieves and Zero Density Theorems*. Oxford: Clarendon Press.
[13] Kaczorowski, J. (1995). On the Shanks-Rényi race problem mod 5. *J. Number Theory*. 50: 106–118.
[14] Kaczorowoski, J. (1996). Boundary values of Dirichlet series and the distribution of primes. In: Balog, A., Katona, G. O. H., Recski, A., Szasz, D., eds. *European Congress of Mathematics*, Budapest, July 22–26, Vol. 1. Progress in Mathematics. Boston: Birkhäuser Verlag, pp. 237–254.
[15] Knapowski, S., Turán, P. (1962). Comparative prime number theory I. *Acta Math. Acad. Sci. Hung.* 13: 315–342.
[16] Languasco, A., Zaccagnini, A. (2007). A note on Mertens' formula for arithmetic progressions. *J. Number Theory*. 127(2007): 37–46.
[17] Lemke Oliver, R. J., Soundararajan, K. (2016). Unexpected biases in the distribution of consecutive primes. *Proc. Natl. Acad. Sci. U. S. A.* 113(31): E4446–E4454.
[18] Littlewood, J. E. (1914). Sur la distribution des nombres premiers. *C. R. Acad. Sci. Paris*. 158: 1869–1875.
[19] Montgomery, H. L., Vaughan, R. C. (2007). *Multiplicative Number Theory I. Classical Theory*. Cambridge: Cambridge University Press.
[20] Pintz, J. (2007). Cramér v. Cramér. On Cramér's probabilistic model for primes. *Funct. Approx.* 37(2): 361–376.
[21] Rubinstein, M., Sarnak, P. (1994). Chebyshev's bias. *Exp. Math.* 3(3): 173–197.
[22] Williams, K. S. (1974). Merten's theorem for arithmetic progressions. *J. Number Theory*. 6(5): 353–359.