

# *Algebra & Number Theory*

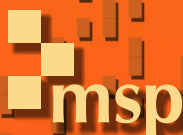
Volume 8

2014

No. 7

**Highly biased prime number races**

Daniel Fiorilli





# Highly biased prime number races

Daniel Fiorilli

Chebyshev observed in a letter to Fuss that there tends to be more primes of the form  $4n + 3$  than of the form  $4n + 1$ . The general phenomenon, which is referred to as Chebyshev's bias, is that primes tend to be biased in their distribution among the different residue classes mod  $q$ . It is known that this phenomenon has a strong relation with the low-lying zeros of the associated  $L$ -functions, that is, if these  $L$ -functions have zeros close to the real line, then it will result in a lower bias. According to this principle one might believe that the most biased prime number race we will ever find is the  $\text{Li}(x)$  versus  $\pi(x)$  race, since the Riemann zeta function is the  $L$ -function of rank one having the highest first zero. This race has density  $0.99999973 \dots$ , and we study the question of whether this is the highest possible density. We will show that it is not the case; in fact, there exist prime number races whose density can be arbitrarily close to 1. An example of a race whose density exceeds the above number is the race between quadratic residues and nonresidues modulo 4849845, for which the density is  $0.999999928 \dots$ . We also give fairly general criteria to decide whether a prime number race is highly biased or not. Our main result depends on the generalized Riemann hypothesis and a hypothesis on the multiplicity of the zeros of a certain Dedekind zeta function. We also derive more precise results under a linear independence hypothesis.

## 1. Introduction and statement of results

The study of prime number races started in 1853, when Chebyshev noted in a letter to Fuss that there seemed to be more primes of the form  $4n + 3$  than of the form  $4n + 1$ . More precisely, Chebyshev claimed without proof that as  $c \rightarrow 0$ , we have

$$-\sum_p \left( \frac{-4}{p} \right) e^{-pc} = e^{-3c} - e^{-5c} + e^{-7c} + e^{-11c} - e^{-13c} - \dots \rightarrow \infty.$$

However, as Hardy and Littlewood [1916] and Landau [1918a; 1918b] have shown, this statement is equivalent to the Riemann hypothesis for  $L(s, \chi_{-4})$ , where  $\chi_{-4}$  denotes the primitive character modulo 4.

---

*MSC2010:* primary 11N13; secondary 11M26.

*Keywords:* prime number races, primes in arithmetic progressions.

Chebyshev's observation created a new area of research which goes under the names of either comparative prime number theory, Chebyshev's bias or, more colloquially, prime number races. This rich research area has a long history, encompassing authors such as Chebyshev, Littlewood, Wintner, Shanks, Knapowski, Turan and Kaczorowski, to name a few, and more recently Rubinstein, Sarnak, Schlage-Puchta, Ng, Martin, Ford, Konyagin and Lamzouri. For a good account of the history of the subject as well as recent developments, the reader is encouraged to consult the great expository paper [Granville and Martin 2006].

The modern way to study Chebyshev's observation is to look at the set of integers  $n$  for which  $\pi(n; 4, 3) > \pi(n; 4, 1)$ , which we denote by  $P_{4;3,1}$ . One would like to understand the size of this set; however, it is known that its natural density does not exist [Kaczorowski 1995]. To remedy this problem we define the *logarithmic density* of a set  $P \subset \mathbb{N}$  by

$$\delta(P) := \lim_{N \rightarrow \infty} \frac{1}{\log N} \sum_{\substack{n \leq N \\ n \in P}} \frac{1}{n},$$

if the limit exists. In general we define  $\underline{\delta}(P)$  and  $\bar{\delta}(P)$  to be the lim inf and lim sup of this sequence. If  $P = P_{4;3,1}$ , then this last limit exists under the assumption of the generalized Riemann hypothesis (GRH) and the linear independence hypothesis (LI), and equals  $0.9959 \dots$  (see [Rubinstein and Sarnak 1994]).

The *generalized Riemann hypothesis* states that for every primitive character  $\chi \bmod q$ , all nontrivial zeros of  $L(s, \chi)$  lie on the line  $\Re(s) = \frac{1}{2}$ .

The *linear independence hypothesis* states that for every fixed modulus  $q$ , the set

$$\bigcup_{\substack{\chi \bmod q \\ \chi \text{ primitive}}} \{ \Im(\rho_\chi) : L(\rho_\chi, \chi) = 0, 0 < \Re(\rho_\chi) < 1, \Im(\rho_\chi) \geq 0 \}$$

is linearly independent over  $\mathbb{Q}$ .

Rubinstein and Sarnak developed a framework to study this question and more general prime number races. Assuming GRH and LI, they have shown that for any  $r$ -tuple  $(a_1, \dots, a_r)$  of admissible residue classes mod  $q$  (that is,  $(a_i, q) = 1$ ), the logarithmic density of the set

$$P_{q;a_1, \dots, a_r} := \{ n : \pi(n; q, a_1) > \pi(n; q, a_2) > \dots > \pi(n; q, a_r) \},$$

which we denote by  $\delta(q; a_1, \dots, a_r)$ , exists and is not equal to 0 or 1 (we call this an  $r$ -way prime number race). Moreover, they have shown that if  $r$  is fixed, then as  $q \rightarrow \infty$ ,

$$\max_{\substack{1 \leq a_1, \dots, a_r \leq q \\ (a_i, q) = 1}} \left| \delta(q; a_1, \dots, a_r) - \frac{1}{r!} \right| \rightarrow 0.$$

In other words, the bias dissolves as  $q \rightarrow \infty$ . For  $r = 2$ , this phenomenon can be readily seen in [Fiorilli and Martin 2013], where the authors exhibited the list of the 117 densities which are greater than or equal to  $\frac{9}{10}$ . By the trivial inclusion

$$P_{q;a_1,\dots,a_r} \subset P_{q;a_1,a_2},$$

we see that the most biased  $r$ -way prime number race is the two-way race appearing on top of the list in that article, that is,

$$\delta(24; 5, 1) = 0.999988 \dots$$

Only one race is known to be more biased: it is the race between  $\text{Li}(x)$  and  $\pi(x)$ , for which the density is

$$\delta(1) := \delta(\{n : \text{Li}(n) > \pi(n)\}) = 0.99999973 \dots$$

One can also combine different residue classes mod  $q$  to make prime number races. For two subsets  $A, B \subset (\mathbb{Z}/q\mathbb{Z})^\times$ , we consider the inequality

$$\frac{1}{|A|} \sum_{a \in A} \pi(n; q, a) > \frac{1}{|B|} \sum_{b \in B} \pi(n; q, b), \quad (1)$$

and denote by  $\delta(q; A, B)$  the logarithmic density of the set of  $n$  for which it is satisfied, if the density exists. An example of such race was given by Rubinstein and Sarnak, who studied the race between

$$\pi(x; q, NR) = \#\{p \leq x : p \text{ is not a quadratic residue mod } q\}$$

and

$$\pi(x; q, R) = \#\{p \leq x : p \text{ is a quadratic residue mod } q\},$$

for moduli  $q$  having a primitive root. This race appears naturally in their work, since, as they have shown, it is the property of the competitors being a quadratic residue or not which determines whether a two-way prime number race is biased or not. These are good candidates for biased races, however, it can be shown that as  $q \rightarrow \infty$ ,  $\delta(q; NR, R) \rightarrow \frac{1}{2}$  (but at a much slower rate than two-way races [Fiorilli and Martin 2013]).

It is known [Bays et al. 2001; Fiorilli and Martin 2013] that under GRH and LI, low-lying zeros of  $L(s, \chi)$  have a significant effect on decreasing the bias, which explains why races of high moduli are very moderately biased. Odlyzko [1990] has shown that the Dedekind zeta function  $\zeta_K(s)$  having the highest first zero in the critical strip is the Riemann zeta function, which is  $\rho_0 = \frac{1}{2} + i \cdot 14.134725 \dots$ . Subsequently, Miller [2002] generalized this result by showing that each member of a very large class of cuspidal  $\text{GL}_n$   $L$ -functions of real archimedean type has the property of either having a zero in the interval  $[\frac{1}{2} - 14.13472i, \frac{1}{2} + 14.13472i]$ , or

having a zero whose real part is strictly larger than  $\frac{1}{2}$  (violating GRH). In particular, this class contains all Dirichlet, rational elliptic curve and modular form  $L$ -functions, and possibly also contains all Artin and rational abelian variety  $L$ -functions.<sup>1</sup> By these considerations, one might conjecture that the highest density one will ever find by doing prime number races with  $L$ -functions of real archimedean type is  $\delta(1) = 0.99999973 \dots$ .

As it turns out, this is false, and we can find races which are arbitrarily biased. This is achieved by considering races between linear combinations of prime counting functions, and we will see in Section 5 that the key to finding such biased races is to take a very large number of residue classes.

The first (and most extreme) example we give is a quadratic residue versus quadratic nonresidue race as in [Rubinstein and Sarnak 1994], but for a general modulus  $q$ . We take  $A = NR := \{a \in (\mathbb{Z}/q\mathbb{Z})^\times : a \not\equiv \square \pmod{q}\}$  and  $B = R := \{b \in (\mathbb{Z}/q\mathbb{Z})^\times : b \equiv \square \pmod{q}\}$  in (1). An elementary argument using the Chinese remainder theorem shows that  $|B| = \phi(q)/\rho(q)$  and  $|A| = \phi(q)(1 - 1/\rho(q))$ , where for  $G = (\mathbb{Z}/q\mathbb{Z})^\times$ ,

$$\rho(q) := [G : G^2] = \begin{cases} 2^{\omega(q)} & \text{if } 2 \nmid q, \\ 2^{\omega(q)-1} & \text{if } 2 \mid q \text{ but } 4 \nmid q, \\ 2^{\omega(q)} & \text{if } 4 \mid q \text{ but } 8 \nmid q, \\ 2^{\omega(q)+1} & \text{if } 8 \mid q, \end{cases}$$

and  $\omega(q)$  denotes the number of distinct prime factors of  $q$ .

**Theorem 1.1.** *Assume GRH and LI. Then for any  $\epsilon > 0$  there exists  $q$  such that*

$$1 - \epsilon < \delta(q; NR, R) < 1. \quad (2)$$

*Moreover, for any fixed  $\frac{1}{2} \leq \eta \leq 1$  there exists a sequence of moduli  $\{q_n\}$  such that*

$$\lim_{n \rightarrow \infty} \delta(q_n; NR, R) = \eta. \quad (3)$$

*In concise form,*

$$\overline{\{\delta(q; NR, R)\}} = \left[\frac{1}{2}, 1\right].$$

To prove the existence of highly biased races we do not need the full strength of LI; in fact, we only need a much weaker hypothesis on the multiplicity of the elements of the multiset of all nontrivial zeros of quadratic Dirichlet  $L$ -functions modulo  $q$ , which we will denote by  $Z(q)$ . Note that LI implies that the elements of this set have multiplicity one.

<sup>1</sup>The restriction to  $L$ -functions of real archimedean type is crucial here, since Bober et al. [2014] have given an example of an  $L$ -function having a first zero whose imaginary part is  $t_0 \approx 14.496$ . They have also shown that under certain conditions, all  $L$ -functions have a zero in the interval  $(-t_2, t_2)$ , with  $t_2 \approx 22.661$ .

**Theorem 1.2.** *Assume GRH, and assume that there exists an increasing sequence of moduli  $q$  such that  $\log q = o(\rho(q))$  and such that each element of  $Z(q)$  has multiplicity  $o(\rho(q)/\log q)$ . Then for any  $\epsilon > 0$  there exists  $q$  such that*

$$1 - \epsilon < \underline{\delta}(q; NR, R) \leq \bar{\delta}(q; NR, R) < 1. \quad (4)$$

**Remark 1.3.** The difference between (2) and (4) is explained by the fact that it is not known whether  $\delta(q; NR, R)$  exists under GRH alone.

**Remark 1.4.** For a fixed modulus  $q \geq 2$ , write

$$q = 2^e \prod_{\substack{p|q \\ p \neq 2}} p^{e_p} \quad \text{and} \quad \ell := \prod_{\substack{p|q \\ p \neq 2}} p.$$

One can show that under GRH,<sup>2</sup>  $\bar{\delta}(q; NR, R) = \bar{\delta}(2^{\min(3,e)}\ell; NR, R)$ . Therefore, when studying  $\bar{\delta}(q; NR, R)$  one can assume without loss of generality that  $q$  is of the form  $2^m\ell$ , where  $\ell$  is an odd squarefree integer and  $m \leq 3$ .

**Remark 1.5.** We will see that what controls the bias in these races is the number of prime factors of  $q$  and the size of  $q$ . More precisely, under GRH and LI the two following statements are equivalent:

$$\sum_{p|q} \log p = o(2^{\omega(q)}), \quad (5)$$

$$\delta(q; NR, R) = 1 - o(1). \quad (6)$$

Using this, we can show that the set of moduli  $q \leq x$  such that  $\delta(q; NR, R) = 1 - o(1)$  has density  $(\log x)^{-\lambda + o(1)}$ , where  $\lambda = 1 - (1 + \log \log 2)/\log 2 = 0.086071 \dots$ . It is an interesting coincidence that the integers satisfying (5) also appear in the Erdős multiplication table (see Ford's work [2008a; 2008b] on integers having a divisor in a given interval).

In terms of random variables, Theorem 1.2 can be explained by saying that the extreme examples we are considering correspond to random variables whose mean is much larger than their standard deviation. The easy way to show that this implies a very large bias is to use Chebyshev's inequality; however this approach is quite imprecise when the ratio  $\mathbb{E}[X]/\sqrt{\text{Var}[X]}$  is large. Instead, one should study the large deviations of  $X - \mathbb{E}[X]$ . The theory of large deviations of remainder terms arising from prime counting functions was initiated by Montgomery [1980], and has since

<sup>2</sup>First note that there are no real primitive characters modulo  $p^e$  with  $p \neq 2$  and  $e \geq 2$ , and there are no real primitive characters modulo  $2^e$  for  $e \geq 4$ . That is, the conductor of any real character modulo  $q$  divides  $2^{\min(3,e)}\ell$ . The claimed equality follows from Lemma 2.1, since  $L(s, \chi^*)$  and  $L(s, \chi)$  have the same nontrivial zeros, and thus  $E_q(x) = E_{2^{\min(3,e)}\ell}(x) + o(1)$ .

$q$	$\omega(q)$	$\rho(q)/\log q'$	$\delta(q; NR, R)$
3	1	1.82	0.999063
15	2	1.47	0.999907
105	3	1.71	0.999928
1155	4	2.26	0.999877
15015	5	3.33	0.999950
255255	6	5.14	0.9999946
4849845	7	8.31	0.999999928
111546435	8	13.81	0.99999999954

**Table 1.** First few values of  $\delta(q; NR, R)$  for half-primorial moduli.

then been developed by Monach [1980], Montgomery and Odlyzko [1988], Rubinstein and Sarnak [1994], and more recently Ng [2004] and Lamzouri [2012]. Exploiting the results of Montgomery and Odlyzko we are able to be more precise in (2).

**Theorem 1.6.** Assume GRH and LI, and define  $q' := \prod_{p|q} p$ . If  $\rho(q)/\log q'$  is large enough, then we have

$$\exp\left(-a_1 \frac{\rho(q)}{\log q'}\right) \leq 1 - \delta(q; NR, R) \leq \exp\left(-a_2 \frac{\rho(q)}{\log q'}\right),$$

where  $a_1$  and  $a_2$  are positive absolute constants.

This last theorem shows that the convergence in (2) can be quite fast. It is actually possible to explicitly compute a density which exceeds  $\delta(1)$ , namely  $\delta(4849845; NR, R) = 0.999999928\dots$ . In Table 1 we list the first few values of  $\delta(q; NR, R)$  for half-primorial moduli (that is, for  $q$  the product of the first  $k$  primes excluding  $p = 2$ ). These values were computed using Myerscough's method [2013] and Rubinstein's `lcalc` package.

**Remark 1.7.** As remarked in [Rubinstein and Sarnak 1994], these densities can theoretically be computed to any given level of accuracy under GRH alone. Indeed, using the  $B^2$  almost-periodicity of these races, this amounts to computing a finite number of zeros of Dirichlet  $L$ -functions to a certain level of accuracy.

**Remark 1.8.** One can summarize Remark 1.5, Theorem 1.1 and Theorem 1.6 by the statement

$$\delta(q; NR, R) \approx \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{2^{\omega(q)-1}/\log q'}}^{\infty} e^{-\frac{x^2}{2}} dx.$$



**Remark 1.9.** Using our analysis, one can show under GRH and LI that for almost all squarefree integers  $q$ ,

$$\delta(q; NR, R) - \frac{1}{2} = (\log q)^{\frac{\log 2 - 1}{2}} + o(1).$$

That is to say, most such races have a very moderate bias.

It is possible to analyze highly biased races in a more general setting, and to determine which features are needed for this bias to appear. To do this we take  $\vec{a} = (a_1, \dots, a_k)$  a vector of invertible reduced residues modulo  $q$  and  $\vec{\alpha} = (\alpha_1, \dots, \alpha_k)$  a nonzero vector of real numbers such that  $\sum_{i=1}^k \alpha_i = 0$ . We will be interested in the race between positive and negative entries of  $\vec{\alpha}$ ; that is, we define

$$\delta(q; \vec{a}, \vec{\alpha}) := \delta(\{n : \alpha_1 \pi(n; q, a_1) + \dots + \alpha_k \pi(n; q, a_k) > 0\}).$$

Moreover, we define

$$\epsilon_i := \begin{cases} 1 & \text{if } a_i \equiv \square \pmod{q}, \\ 0 & \text{if } a_i \not\equiv \square \pmod{q}, \end{cases}$$

and we assume without loss of generality that

$$\sum_{i=1}^k \epsilon_i \alpha_i < 0.$$

(By Lemma 5.1, this will force  $\delta(q; \vec{a}, \vec{\alpha}) > \frac{1}{2}$ . If  $\sum_{i=1}^k \epsilon_i \alpha_i = 0$ , then  $\delta(q; \vec{a}, \vec{\alpha}) = \frac{1}{2}$ . If  $\sum_{i=1}^k \epsilon_i \alpha_i > 0$ , then we multiply  $\vec{\alpha}$  by  $-1$  and study the complementary probability  $\delta(q; \vec{a}, -\vec{\alpha}) = 1 - \delta(q; \vec{a}, \vec{\alpha})$ .)

There are many choices of vectors  $\vec{a}$  and  $\vec{\alpha}$  which yield highly biased races. We give some examples with constant coefficients, which we believe are the most natural.

**Theorem 1.10.** *Assume GRH and LI, and let*

$$k_R \leq \frac{\phi(q)}{\rho(q)} \quad \text{and} \quad k_N \leq \left(1 - \frac{1}{\rho(q)}\right) \phi(q)$$

*be two positive integers. Take  $a_1, \dots, a_{k_N}$  to be any distinct quadratic nonresidues mod  $q$ , with coefficients  $\alpha_1 = \dots = \alpha_{k_N} = k_R$ , and  $a_{k_N+1}, \dots, a_{k_N+k_R}$  to be any distinct quadratic residues mod  $q$ , with  $\alpha_{k_N+1} = \dots = \alpha_{k_N+k_R} = -k_N$ . There exists an absolute constant  $c > 0$  such that if for some  $0 < \epsilon < 1/(2c)$  we have*

$$\frac{1}{k_N} + \frac{1}{k_R} < \epsilon \frac{\rho(q)^2}{\phi(q) \log q}, \tag{7}$$

*then*

$$\delta(q; \vec{a}, \vec{\alpha}) > 1 - c\epsilon.$$

**Remark 1.11.** Fix  $0 < \epsilon < 1/(2c)$  and define  $N_\epsilon(q)$  to be the number of positive integers  $k_N, k_R$  for which  $k_N \leq (1 - 1/\rho(q))\phi(q)$ ,  $k_R \leq \phi(q)/\rho(q)$  and

$$\frac{1}{k_N} + \frac{1}{k_R} < \epsilon \frac{\rho(q)^2}{\phi(q) \log q}.$$

Then, for values of  $q$  for which  $\rho(q) \geq \epsilon^{-2} \log q$ , we have that  $N_\epsilon(q)$  tends to infinity as  $q \rightarrow \infty$ . Hence, for values of  $q$  for which  $\log q = o(\rho(q))$ , (7) has a large number of solutions.

**Remark 1.12.** Theorem 1.10 shows the existence of highly biased races with the same number of residue classes on each side of the inequality. Indeed, for moduli  $q$  with  $\log q = o(\rho(q))$ , taking  $k_R = k_N$  with  $\phi(q) \log q / \rho(q)^2 = o(k_R)$  and choosing any residue classes  $a_1, \dots, a_{k_N+k_R}$  gives a race with  $\delta(q; \vec{a}, \vec{\alpha}) = 1 - o(1)$ .

**Remark 1.13.** In Theorem 1.1, we have

$$k_N = \left(1 - \frac{1}{\rho(q)}\right)\phi(q) \quad \text{and} \quad k_R = \frac{\phi(q)}{\rho(q)},$$

which explains why we obtained a highly biased race when  $\rho(q)$  was large compared to  $\log q$ .

Here is our most general class of highly biased races.

**Theorem 1.14.** Assume GRH and LI. There exists an absolute constant  $c > 0$  such that if for some  $0 < \epsilon < 1/(2c)$  we have

$$\frac{\sum_{i=1}^k \alpha_i^2}{\left(\sum_{i=1}^k \epsilon_i \alpha_i\right)^2} < \epsilon \frac{\rho(q)^2}{\phi(q) \log q}, \quad (8)$$

then

$$\delta(q; \vec{a}, \vec{\alpha}) > 1 - c\epsilon.$$

**Remark 1.15.** Trivially, one has

$$\frac{\sum_{i=1}^k \alpha_i^2}{\left(\sum_{i=1}^k \epsilon_i \alpha_i\right)^2} \geq \frac{1}{k_R},$$

where  $k_R := \sum_{i=1}^k \epsilon_i$ . Hence, for (8) to be satisfied, one needs  $k_R$  to be larger than

$$\epsilon^{-1} \frac{\phi(q) \log q}{\rho(q)^2}.$$

Since  $k_R \leq \phi(q)/\rho(q)$ , this imposes the condition on  $q$

$$\rho(q) \geq \epsilon^{-1} \log q.$$

**Remark 1.16.** The goal of Theorem 1.14 is to give a large class of biased races, without necessarily being precise on the value of  $\delta(q; \vec{a}, \vec{\alpha})$ . One can use the Montgomery–Odlyzko bounds [1988] to obtain more precise estimates in some particular cases.

The previous examples of highly biased races all have the property that the number of residue classes involved is very large in terms of  $q$  (it is at least  $q^{1-o(1)}$ ). In the next theorem we show that this condition is necessary, and that moreover highly biased races are very particular, in the sense that they must satisfy precise conditions.

**Theorem 1.17.** *Assume GRH and LI. There exist absolute positive constants  $K_1, K_2$  and  $0 < \eta < \frac{1}{2}$  such that if  $k \leq K_1 \phi(q)$  and*

$$\frac{\left(\sum_{i=1}^k \epsilon_i \alpha_i\right)^2}{\sum_{i=1}^k \alpha_i^2} \leq K_2 \frac{\phi(q) \log \frac{3\phi(q)}{k}}{\rho(q)^2}, \quad (9)$$

then

$$\delta(q; \vec{a}, \vec{\alpha}) \leq 1 - \eta. \quad (10)$$

(Hence this race cannot be too biased.)

**Remark 1.18.** Applying the Cauchy–Schwarz inequality and using that  $k_R := \sum_{i=1}^k \epsilon_i \leq \phi(q)/\rho(q)$ , one sees that if  $\rho(q) \leq K_2 \log(3\phi(q)/k)$ , then whatever  $\vec{a}$  and  $\vec{\alpha}$  are, (9) holds. Moreover, in the range  $\rho(q) > K_2 \log(3\phi(q)/k)$  we have that if  $k_R \leq K_2 \phi(q)/\rho(q)^2$ , then (9) holds. We conclude that a necessary condition to obtain a highly biased race is that  $k_R \gg \phi(q)/\rho(q)^2$ .

An interesting feature of prime number races is Skewes' number. It is by definition the smallest  $x_0$  for which

$$\pi(x_0) > \text{Li}(x_0).$$

This number has been extensively studied since Skewes' paper [1933] in which he showed under the Riemann hypothesis that

$$x_0 < 10^{10^{10^{34}}}.$$

The Riemann hypothesis has since then been removed and the upper bound greatly reduced; we refer the reader to [Bays and Hudson 2000] for the list of such improvements. The current record is due to the authors of that work, who showed that

$x_0 < 1.3983 \times 10^{316}$ , and moreover this bound is believed to be close to the true size of  $x_0$ .

One could also study the generalized Skewes' number

$$x_{q;a,b} := \inf\{x : \pi(x; q, a) < \pi(x; q, b)\}.$$

However, two-way prime number races become less and less biased as  $q$  grows; that is,  $\delta(q; a, b) \rightarrow \frac{1}{2}$  uniformly in  $a$  and  $b$  coprime to  $q$ . Hence, for large  $q$  we expect this generalized Skewes number to be small and uninteresting.

The situation is quite different with the highly biased race we constructed; in fact, we expect the Skewes number

$$x_q := \inf\{x : \pi(x; q, NR) < (\rho(q) - 1)\pi(x; q, R)\}$$

to tend to infinity as  $\rho(q)/\log q'$  tends to infinity ( $q'$  is the radical of  $q$ ). Using similar arguments to those of [Montgomery 1980; Ng 2000], we can speculate on the exact growth rate of  $x_q$ .

**Conjecture 1.19.** *As  $\rho(q)/\log q'$  tends to infinity we have*

$$\log \log x_q \asymp \frac{\rho(q)}{\log q'}.$$

## 2. Results without the linear independence hypothesis

The goal of this section is to prove Theorem 1.2 (from which the first part of Theorem 1.1 clearly follows). We first note that if  $A = NR$  and  $B = R$ , then (1) is equivalent to

$$\pi(x; q, NR) > (\rho(q) - 1)\pi(x; q, R).$$

**Lemma 2.1.** *Fix  $q \geq 3$ . Assuming GRH, we have that*

$$\begin{aligned} E_q(x) &:= \frac{\pi(x; q, NR) - (\rho(q) - 1)\pi(x; q, R)}{\sqrt{x}/\log x} \\ &= \rho(q) - 1 + \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{\gamma_\chi} \frac{x^{i\gamma_\chi}}{\rho_\chi} + o_{x \rightarrow \infty}(1). \end{aligned}$$

*Proof.* Let  $b$  be an invertible reduced residue mod  $q$ . We will use the orthogonality relation

$$\sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \chi(b) = \begin{cases} \rho(q) - 1 & \text{if } b \equiv \square \bmod q, \\ -1 & \text{if } b \not\equiv \square \bmod q. \end{cases} \quad (11)$$

The explicit formula gives

$$\sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \psi(x, \chi) = - \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{\rho_\chi} \frac{x^{\rho_\chi}}{\rho_\chi} + O_q(\log x), \quad (12)$$

where  $\rho_\chi$  runs over the nontrivial zeros of  $L(s, \chi)$ . The left side of (12) is equal to

$$\begin{aligned} & \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{p \leq x} \chi(p) \log p + \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{p^2 \leq x} \chi(p)^2 \log p + O_q(x^{\frac{1}{3}}) \\ &= (\rho(q) - 1) \sum_{\substack{p \leq x \\ p \equiv \square \bmod q}} \log p - \sum_{\substack{p \leq x \\ p \not\equiv \square \bmod q}} \log p + (\rho(q) - 1)\sqrt{x} + o_q(\sqrt{x}), \end{aligned}$$

by (11) and the prime number theorem. Combining this with a standard summation by parts we get that

$$\frac{\pi(x; q, NR) - (\rho(q) - 1)\pi(x; q, R)}{\sqrt{x}/\log x} = \rho(q) - 1 + \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{\gamma_\chi} \frac{x^{i\gamma_\chi}}{\rho_\chi} + o_{x \rightarrow \infty}(1). \quad \square$$

**Lemma 2.2.** *Assuming GRH, the quantity  $E_q(x)$  defined in Lemma 2.1 has a limiting logarithmic distribution; that is, there exists a Borel measure  $\mu_q$  on  $\mathbb{R}$  such that for any bounded Lipschitz continuous function  $f : \mathbb{R} \rightarrow \mathbb{R}$  we have*

$$\lim_{Y \rightarrow \infty} \frac{1}{Y} \int_2^Y f(E_q(e^y)) dy = \int_{\mathbb{R}} f(t) d\mu_q(t).$$

*Proof.* This follows from analysis in [Rubinstein and Sarnak 1994; Akbary et al. 2013].  $\square$

**Remark 2.3.** By the Portmanteau theorem, the Lipschitz assumption in the last lemma can be removed.

**Remark 2.4.** As Schläge-Puchta has pointed out to me, it is possible to show under GRH that for all but a countable set of values of  $c$ , the density

$$F_q(c) := \lim_{Y \rightarrow \infty} \frac{1}{Y} \text{meas}\{y \leq Y : E_q(e^y) \leq c\}$$

exists. Moreover, one can show that in the domain where  $F$  is defined,

$$\begin{aligned} \sup_{x < c} F_q(x) &\leq \liminf_{Y \rightarrow \infty} \frac{1}{Y} \text{meas}\{y \leq Y : E_q(e^y) \leq c\} \\ &\leq \limsup_{Y \rightarrow \infty} \frac{1}{Y} \text{meas}\{y \leq Y : E_q(e^y) \leq c\} \leq \inf_{x > c} F_q(x), \end{aligned}$$

and so in particular if  $F_q(x)$  is continuous at  $x = c$ , then the set  $\{y \leq Y : E_q(e^y) \leq c\}$  has a density.

Let  $X_q$  be the random variable associated to  $\mu_q$ . We will show that  $X_q$  can be very biased, in the sense that  $\text{Prob}[X_q > 0]$  can be very close to 1. To do so we will compute the first two moments of  $E_q(e^y)$ , which we relate to the random variable  $X_q$ .

**Lemma 2.5.** *Under GRH, we have that*

$$\begin{aligned} \lim_{Y \rightarrow \infty} \frac{1}{Y} \int_2^Y E_q(e^y) dy &= \int_{\mathbb{R}} t d\mu_q(t), \\ \lim_{Y \rightarrow \infty} \frac{1}{Y} \int_2^Y E_q(e^y)^2 dy &= \int_{\mathbb{R}} t^2 d\mu_q(t). \end{aligned}$$

*Proof.* We will only prove the second statement, as the first follows along the same lines. Similarly as in [Schlage-Puchta 2000], we can compute that

$$\lim_{Y \rightarrow \infty} \frac{1}{Y} \int_0^Y |E_q(e^y)|^4 dy = \sum_{\rho_1 + \rho_2 + \rho_3 + \rho_4 = 0} \frac{1}{\rho_1 \rho_2 \rho_3 \rho_4} < \infty,$$

where the last sum runs over quadruples of nontrivial zeros of quadratic Dirichlet  $L$ -functions modulo  $q$ . This implies that as  $M \rightarrow \infty$ ,

$$\limsup_{Y \rightarrow \infty} \frac{1}{Y} \int_{\substack{0 \leq y \leq Y \\ |E_q(e^y)| > M}} |E_q(e^y)|^2 dy \rightarrow 0. \quad (13)$$

Indeed, if this was not the case then we would have that for all  $M > M_0$ ,

$$\limsup_{Y \rightarrow \infty} \frac{1}{Y} \int_{\substack{0 \leq y \leq Y \\ |E_q(e^y)| > M}} |E_q(e^y)|^2 dy \geq \eta > 0,$$

and so

$$\limsup_{Y \rightarrow \infty} \frac{1}{Y} \int_{\substack{0 \leq y \leq Y \\ |E_q(e^y)| > M}} |E_q(e^y)|^4 dy \geq \eta M^2,$$

which would contradict the fact that the fourth moment is finite. We now define the bounded Lipschitz function

$$H_M(t) := \begin{cases} t^2 & \text{if } |t| \leq M, \\ M^2(M+1-|t|) & \text{if } M < |t| \leq M+1, \\ 0 & \text{if } |t| \geq M+1. \end{cases}$$

We then have

$$\begin{aligned} \frac{1}{Y} \int_2^Y E_q(e^y)^2 dy &= \frac{1}{Y} \int_{2 \leq y \leq Y} H_M(E_q(e^y)) dy - \frac{1}{Y} \int_{\substack{2 \leq y \leq Y \\ M < |E_q(e^y)| \leq M+1}} H_M(E_q(e^y)) dy \\ &\quad + \frac{1}{Y} \int_{\substack{2 \leq y \leq Y \\ |E_q(e^y)| > M}} E_q(e^y)^2 dy; \end{aligned}$$

therefore by (13) and Lemma 2.2 we get that

$$\limsup_{Y \rightarrow \infty} \frac{1}{Y} \int_2^Y E_q(e^y)^2 dy = \int_{\mathbb{R}} H_M(t) d\mu_q(t) + \epsilon_M,$$

where  $\epsilon_M$  tends to zero as  $M \rightarrow \infty$ . Using the bound

$$\mu_q((-\infty, -M] \cup [M, \infty)) \ll \exp(-c_2 \sqrt{M})$$

(see [Rubinstein and Sarnak 1994, Theorem 1.2]), we get by taking  $M \rightarrow \infty$  that

$$\limsup_{Y \rightarrow \infty} \frac{1}{Y} \int_2^Y E_q(e^y)^2 dy = \int_{\mathbb{R}} t^2 d\mu_q(t).$$

The same reasoning applies to the  $\liminf$ , and thus the proof is finished.  $\square$

The following calculation is similar to that of Schlage-Puchta [2000], who computed the moments of  $e^{-t/2}\psi(e^t; \chi)$ .

**Lemma 2.6.** *Assume GRH. Then,*

$$\mathbb{E}[X_q] = \rho(q) - 1 + z(q) \quad \text{and} \quad \text{Var}[X_q] = \sum_{\gamma \neq 0}^* \frac{m_\gamma^2}{\frac{1}{4} + \gamma^2},$$

where the last sum runs over the imaginary parts of the nontrivial zeros of

$$Z_q(s) := \prod_{\substack{\chi^2 = \chi_0 \\ \chi \neq \chi_0}} L(s, \chi),$$

$m_\gamma$  denotes the multiplicity of the zero  $\frac{1}{2} + i\gamma$ , the star means that we count the zeros without multiplicity, and  $z(q)$  denotes the multiplicity of the (possible) real zero  $\gamma = 0$ .

*Proof.* By Lemma 2.1 we have that

$$\begin{aligned} \int_2^Y E_q(e^y) dy &= (\rho(q) - 1 + z(q))(Y - 2) \\ &\quad + \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{\gamma_\chi \neq 0} \frac{1}{\frac{1}{2} + i\gamma_\chi} \int_2^Y e^{i\gamma_\chi y} dy + o_{Y \rightarrow \infty}(Y) \\ &= (\rho(q) - 1 + z(q))(Y - 2) + O_q(1) + o_{Y \rightarrow \infty}(Y), \end{aligned}$$

by absolute convergence. Taking  $Y \rightarrow \infty$  and applying Lemma 2.5 gives that

$$\mathbb{E}[X_q] = \lim_{Y \rightarrow \infty} \frac{1}{Y} \int_2^Y E_q(e^y) dy = \rho(q) - 1 + z(q).$$

The calculation of the variance follows from Lemma 2.1 and from Parseval's identity for  $B^2$  almost-periodic functions [Besicovitch 1926]. (An alternative way to compute the variance is to argue as in [Schlage-Puchta 2000].)  $\square$

**Remark 2.7.** It is a general fact that Besicovitch almost-periodic functions [1955] always have a mean value. Moreover, Parseval's identity [Besicovitch 1955; 1926] shows that Besicovitch  $B^2$  almost-periodic functions  $f(y)$  have a second moment, given by

$$\lim_{Y \rightarrow \infty} \frac{1}{Y} \int_0^Y f(y)^2 dy = \sum_{n \geq 1} A_n^2,$$

where the  $A_n$  are the Fourier coefficients of  $f$ .

**Lemma 2.8.** Let  $\chi \bmod q$  be a Dirichlet character. We have for  $k \geq 1$  that

$$b_k(\chi) := \sum_{\gamma_\chi} \frac{1}{\left(\frac{1}{4} + \gamma_\chi^2\right)^k} \asymp_k \log q^*,$$

where the sum is counted with multiplicity.

**Remark 2.9.** One has an exact formula for  $b_k(\chi)$ , in terms of the values of the derivatives of  $\log L(s, \chi)$  evaluated at  $s = 1$  [Fiorilli and Martin 2013, Lemma 3.15].



*Proof.* This follows from applying the Riemann–von Mangoldt formula and summation by parts:

$$\begin{aligned} b_k(\chi) &= \sum_{|\gamma_\chi| < 1} \frac{1}{(\frac{1}{4} + \gamma_\chi^2)^k} + \int_1^\infty \frac{dN(t, \chi)}{(\frac{1}{4} + t^2)^k} \\ &\asymp_k N(1, \chi) + 2k \int_1^\infty \frac{tN(t, \chi)}{(\frac{1}{4} + t^2)^{k+1}} dt \asymp_k \log q^*. \end{aligned} \quad \square$$

**Lemma 2.10.** *Assume GRH. If*

$$B(q) := \frac{\mathbb{E}[X_q]}{\sqrt{\text{Var}[X_q]}}$$

*is large enough, then*

$$\underline{\delta}(q; NR, R) \geq 1 - 2 \frac{\text{Var}[X_q]}{\mathbb{E}[X_q]^2}.$$

*Proof.* It is clear from Lemmas 2.6 and 2.8 that  $\text{Var}[X_q] \gg \log q'$ , and therefore our assumption that  $B(q)$  is large enough implies that  $\mathbb{E}[X_q]$  is also large enough, say at least 4. Now let

$$H(x) := \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } x \geq 0, \end{cases} \quad f(x) := \begin{cases} 0 & \text{if } x < 0, \\ x & \text{if } 0 \leq x < 1, \\ 1 & \text{if } x \geq 1. \end{cases}$$

Clearly,  $f(x)$  is bounded Lipschitz continuous and  $f(x) \leq H(x)$ . Therefore,

$$\underline{\delta}(q; NR, R) = \liminf_{Y \rightarrow \infty} \frac{1}{Y} \int_2^Y H(E_q(e^y)) dy \geq \liminf_{Y \rightarrow \infty} \frac{1}{Y} \int_2^Y f(E_q(e^y)) dy,$$

which by Lemma 2.2 is equal to

$$\begin{aligned} \int_{\mathbb{R}} f(t) d\mu_q(t) &= 1 - \int_{\mathbb{R}} (1 - f(t)) d\mu_q(t) \\ &= 1 - \int_{-\infty}^1 (1 - f(t)) d\mu_q(t) \geq 1 - \mu_q(-\infty, 1]. \end{aligned}$$

We now apply Chebyshev's inequality:

$$\begin{aligned} \mu_q(-\infty, 1] &= \text{Prob}[X_q \leq 1] = \text{Prob}[X_q - \mathbb{E}[X_q] \leq 1 - \mathbb{E}[X_q]] \\ &\leq \text{Prob}[|X_q - \mathbb{E}[X_q]| \geq \mathbb{E}[X_q] - 1] \leq \frac{\text{Var}[X_q]}{(\mathbb{E}[X_q] - 1)^2} \leq 2 \frac{\text{Var}[X_q]}{\mathbb{E}[X_q]^2}, \end{aligned}$$

since  $\mathbb{E}[X_q] \geq 4$ , and therefore

$$\underline{\delta}(q; NR, R) \geq 1 - 2 \frac{\text{Var}[X_q]}{\mathbb{E}[X_q]^2}. \quad \square$$

*Proof of Theorem 1.2.* By Lemma 2.6, our hypothesis implies that for the sequence of moduli  $q$  under consideration,

$$\mathrm{Var}[X_q] \leq \max_{\gamma} (m_{\gamma}) \sum_{\gamma}^* \frac{m_{\gamma}}{\frac{1}{4} + \gamma_{\chi}^2} = o\left(\frac{\rho(q)}{\log q} \rho(q) \log q\right) = o(\rho(q)^2),$$

by Lemma 2.8. Lemma 2.6 also implies that  $\mathbb{E}[X_q] \gg \rho(q)$ , and hence Lemma 2.10 implies that

$$\underline{\delta}(q; NR, R) \geq 1 - o(1).$$

The last inequality to show, that is,  $\bar{\delta}(q; NR, R) < 1$ , follows from a lower bound on  $\mu_E(-\infty, -1]$  similar to that in [Rubinstein and Sarnak 1994, Theorem 1.2], which holds in greater generality [Akbari et al. 2013]. Using this lower bound, one does an analysis similar to that in the proof of Lemma 2.10, replacing the function  $f(x)$  with

$$g(x) := \begin{cases} 1 & \text{if } x < -1, \\ -x & \text{if } -1 \leq x < 0, \\ 0 & \text{if } x \geq 0, \end{cases}$$

in order to obtain a lower bound for

$$1 - \bar{\delta}(q; NR, R) = \limsup_{Y \rightarrow \infty} \frac{1}{Y} \int_2^{\infty} (1 - H(E_q(e^y))) dy. \quad \square$$

### 3. A central limit theorem

The goal of this section is to show a central limit theorem under GRH and LI, from which the second part of Theorem 1.1 will follow. We first translate our problem to questions on sums of independent random variables, which can be done thanks to LI. Recall that we are interested in the set of  $n$  such that

$$\pi(n; q, NR) > (\rho(q) - 1)\pi(n; q, R).$$

**Lemma 3.1.** *Assume GRH and LI. Then the logarithmic density of the set of  $n$  for which  $\pi(n; q, NR) > (\rho(q) - 1)\pi(n; q, R)$  exists and is equal to*

$$\mathrm{Prob}[X_q > 0],$$

where  $X_q$  is the random variable defined in Section 2. Moreover, we have

$$X_q \sim \rho(q) - 1 + \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{\gamma_{\chi} > 0} \frac{2\Re(Z_{\gamma_{\chi}})}{\sqrt{\frac{1}{4} + \gamma_{\chi}^2}}, \quad (14)$$

where the  $Z_{\gamma_{\chi}}$  are independent identically distributed random variables following a uniform distribution on the unit circle in  $\mathbb{C}$ .

*Proof.* By Lemma 2.1, we have that

$$\frac{\pi(x; q, NR) - (\rho(q) - 1)\pi(x; q, R)}{\sqrt{x}/\log x} = \rho(q) - 1 + \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{\gamma_\chi} \frac{x^{i\gamma_\chi}}{\rho_\chi} + o(1),$$

since LI implies that there are no real zeros. It follows by the work of Rubinstein and Sarnak that  $\delta(q; NR, R)$  exists and equals  $\text{Prob}[X_q > 0]$  (their analysis shows that the distribution function of  $X_q$  is continuous). Moreover, an argument similar to the proof of [Fiorilli and Martin 2013, Proposition 2.3] shows that (14) holds.  $\square$

One can show that the random variables in (14) have variance  $\text{Var}[\Re(Z_{\gamma_\chi})] = \frac{1}{2}$  and have mean  $\mathbb{E}[Z_{\gamma_\chi}] = 0$ . Using this and the fact that they are mutually independent, we recover Lemma 2.6:

$$\mathbb{E}[X_q] = \rho(q) - 1, \quad \text{Var}[X_q] = \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{\gamma_\chi} \frac{1}{\frac{1}{4} + \gamma_\chi^2}, \quad (15)$$

since the zeros come in conjugate pairs ( $\chi$  is real). We will see in the following lemma that  $\text{Var}[X_q] \asymp \rho(q) \log q'$  (recall that  $q' := \prod_{p|q} p$ ), and this is a crucial fact in our analysis.

**Lemma 3.2.** *Assume GRH and LI, and let  $X_q$  be the random variable defined in (14). We have that*

$$\text{Var}[X_q] = 2^{\omega(q)-1-\epsilon_q} \log q' \left[ 1 + O\left(\frac{\log \log q'}{\log q'}\right) \right],$$

where  $\epsilon_q = 1$  if  $2 \mid q$ , and  $\epsilon_q = 0$  otherwise. In particular,

$$\text{Var}[X_q] \asymp \rho(q) \log q'.$$

*Proof.* By Remark 1.4, we have that

$$\text{Var}[X_q] = \text{Var}[X_{2^e \ell}],$$

where  $e \leq 3$ ,  $2^e \parallel q$  and  $\ell := \prod_{p|q, p \neq 2} p$ . Therefore we assume from now on (without loss of generality) that  $q = 2^e \ell$ , with  $e \leq 3$  and  $\ell$  an odd squarefree integer.

Lemma 3.5 of [Fiorilli and Martin 2013] gives that

$$\begin{aligned} \sum_{\gamma_\chi} \frac{1}{\frac{1}{4} + \gamma_\chi^2} &= \log q^* - \log \pi - \gamma - (1 + \chi(-1)) \log 2 + 2\Re \frac{L'}{L}(1, \chi^*) \\ &= \log q^* + O(\log \log q^*), \end{aligned} \quad (16)$$

by Littlewood's GRH bound on  $(L'/L)(1, \chi)$ . Plugging this into (15), we get

$$\text{Var}[X_q] = \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0}} \log q^* + O(2^{\omega(q)} \log \log q).$$

If  $q$  is odd, then there is exactly one primitive real character mod  $d$  for every  $d \mid q$ , hence

$$\sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0}} \log q^* = \sum_{d \mid q} \log d = \sum_{d \mid q} \sum_{p \mid d} \log p = \sum_{p \mid q} (\log p) 2^{\omega(q)-1} = 2^{\omega(q)-1} \log q.$$

If  $2 \parallel q$ , then there are no primitive characters modulo even divisors of  $q$ , so

$$\sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0}} \log q^* = \sum_{d \mid \frac{q}{2}} \log d = 2^{\omega(q)-2} \log \frac{q}{2}.$$

If  $4 \parallel q$ , then there is exactly one primitive real character modulo divisors which are multiples of 4, so

$$\sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0}} \log q^* = \sum_{d \mid \frac{q}{4}} \log d + \sum_{4 \mid d \mid q} \log d = 2^{\omega(q)-2} \log(2q).$$

If  $8 \parallel q$ , then there are exactly two primitive real characters modulo divisors which are multiples of 8, so

$$\sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0}} \log q^* = \sum_{d \mid \frac{q}{8}} \log d + \sum_{\substack{4 \mid d \mid q \\ 8 \nmid d}} \log d + 2 \sum_{8 \mid d \mid q} \log d = 2^{\omega(q)-2} \log(8q). \quad \square$$

Let  $X_q$  be the random variable defined in (14), and define

$$B(q) := \frac{\mathbb{E}[X_q]}{\sqrt{\text{Var}[X_q]}}.$$

It is  $B(q)$  which dictates the behavior of the race we are considering: if  $B(q)$  is small, then the race will not be very biased, whereas if  $B(q)$  is large, then the race will have a significant bias. By Lemma 3.2, we have under GRH and LI the estimate

$$B(q) = \sqrt{\frac{2^{\omega(q)+1+\epsilon_q}}{\log q'}} \left[ 1 + O\left(2^{-\omega(q)} + \frac{\log \log q'}{\log q'}\right) \right]. \quad (17)$$

To prove the second part of Theorem 1.1 we will need a sequence of moduli for which  $B(q)$  is very regular.

**Lemma 3.3.** *For any fixed  $0 < c < \infty$ , there exists an increasing sequence of squarefree odd integers  $\{q_n\}$  such that*

$$2^{\omega(q_n)+1} = (c + o(1)) \log q_n.$$

*Proof.* Fix  $0 < c < \infty$ , and define  $e_c := \min\{e \geq 1 : 2^{-e}c < 2/\log 4\}$  and  $c_1 := 2^{-e_c}c < 2/\log 4$ . Define for  $\ell = 1, 2, \dots$  the intervals

$$I_\ell := (\exp(c_1^{-1}2^\ell), 2\exp(c_1^{-1}2^\ell)), \quad J_\ell := (2\exp(c_1^{-1}2^\ell), 4\exp(c_1^{-1}2^\ell)).$$

Since  $c_1 < 2/\log 4$ , we have that for all  $\ell \geq 1$ ,

$$4\exp(c_1^{-1}2^\ell) < \exp(c_1^{-1}2^{\ell+1});$$

hence our intervals are all disjoint. We define  $p_\ell$  to be any prime in the interval  $I_\ell$ , and similarly for  $p'_\ell \in J_\ell$ . The existence of such primes is granted by Bertrand's postulate (note that  $\exp(c_1^{-1}2^1) > 4$ ). Now, the sequence of moduli we are looking for is

$$q_n := \prod_{1 \leq \ell \leq e_c} p'_\ell \prod_{1 \leq \ell \leq n} p_\ell,$$

since

$$\begin{aligned} \frac{2^{\omega(q_n)+1}}{\log q_n} &= \frac{2^{n+e_c+1}}{O_c(1) + \sum_{1 \leq \ell \leq n} (c_1^{-1}2^\ell + O(1))} = \frac{2^{n+e_c+1}}{c_1^{-1}2^{n+1} + O_c(n)} \\ &= 2^{e_c}c_1 \left(1 + O_c\left(\frac{n}{2^n}\right)\right) = c(1 + o(1)), \end{aligned}$$

by definition of  $c_1$ . □

Before proving the second part of Theorem 1.1, we give some information about the characteristic function of the random variables we are interested in. The following lemma implies a central limit theorem.

**Lemma 3.4.** *Assume GRH and LI. Let  $X_q$  be the random variable defined in (14), and define*

$$Y_q := \frac{X_q - \mathbb{E}[X_q]}{\sqrt{\text{Var}[X_q]}} = \frac{1}{\sqrt{\text{Var}[X_q]}} \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{\gamma_\chi > 0} \frac{2\Re(Z_{\gamma_\chi})}{\sqrt{\frac{1}{4} + \gamma_\chi^2}}.$$

*The characteristic function of  $Y_q$  satisfies, for  $|\xi| \leq \frac{3}{5}\sqrt{\text{Var}[X_q]}$ ,*

$$\hat{Y}_q(\xi) = -\frac{\xi^2}{2} + O\left(\frac{\xi^4}{\rho(q) \log q'}\right).$$

Moreover, in the same range, we have

$$\widehat{Y}_q(\xi) \leq -\frac{\xi^2}{2}. \quad (18)$$

*Proof.* The proof is very similar to that of [Fiorilli and Martin 2013, Theorem 3.22]. Using the additivity of the cumulant-generating function of  $X_q$ , one can show that

$$\log \widehat{X}_q(\xi) = i\mathbb{E}[X_q]\xi + \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{\gamma_\chi > 0} \log \left( J_0 \left( \frac{2\xi}{\sqrt{\frac{1}{4} + \gamma_\chi^2}} \right) \right), \quad (19)$$

where  $J_0(x)$  is the Bessel function of the first kind:

$$J_0(x) = \sum_{n=0}^{\infty} \frac{(-1)^n (x/2)^{2n}}{n!^2}.$$

We will use the following Taylor expansion, which is valid for  $|\xi| \leq \frac{12}{5}$  (see [Fiorilli and Martin 2013, Section 2.2]):

$$\log J_0(\xi) = -\frac{\xi^2}{4} + O(\xi^4). \quad (20)$$

Plugging this estimate into (19), we get that for  $|\xi| \leq \frac{3}{5}$ ,

$$\log \widehat{X}_q(\xi) = i\mathbb{E}[X_q]\xi - \xi^2 \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{\gamma_\chi > 0} \frac{1}{\frac{1}{4} + \gamma_\chi^2} + O\left(\xi^4 \sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{\gamma_\chi > 0} \frac{1}{(\frac{1}{4} + \gamma_\chi^2)^2}\right).$$

Applying Lemma 2.8 gives

$$\sum_{\substack{\chi \bmod q \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{\gamma_\chi > 0} \frac{1}{(\frac{1}{4} + \gamma_\chi^2)^2} \ll \rho(q) \log q'.$$

Moreover, by Lemma 3.2 we have  $\text{Var}[X_q] \asymp \rho(q) \log q'$ . Putting these together and using (15), we get that

$$\log \widehat{Y}_q(\xi) = \log \widehat{X}_q\left(\frac{\xi}{\sqrt{\text{Var}[X_q]}}\right) - i\mathbb{E}[X_q] \frac{\xi}{\sqrt{\text{Var}[X_q]}} = -\frac{\xi^2}{2} + O\left(\frac{\xi^4}{\rho(q) \log q'}\right),$$

showing the first assertion. For the second we use the same argument, but we replace the estimate (20) with the following inequality, valid in the range  $|\xi| \leq \frac{12}{5}$ :

$$\log J_0(\xi) \leq -\frac{\xi^2}{4}. \quad \square$$

**Lemma 3.5** (Berry–Esseen inequality). *Assume GRH and LI. Denote by  $F_q$  the distribution function of*

$$Y_q := \frac{X_q - \mathbb{E}[X_q]}{\sqrt{\text{Var}[X_q]}},$$

*and by  $F$  that of the Gaussian distribution. We have that*

$$\sup_{x \in \mathbb{R}} |F_q(x) - F(x)| \ll \frac{1}{\rho(q) \log q'}.$$

**Remark 3.6.** One could get a more precise estimate using the Feuerverger–Martin formula [2000]. However, the estimate of Lemma 3.5 is sufficient for our purposes.

*Proof.* Since the statement is trivial if  $\rho(q) \log q'$  is bounded, we can assume without loss of generality that  $\text{Var}[X_q] \geq 1$  (by Lemma 3.2).

The Berry–Esseen inequality in the form of [Esseen 1945, Theorem 2a] gives that for any  $T > 0$ ,

$$\sup_{x \in \mathbb{R}} |F_q(x) - F(x)| \ll \int_{-T}^T \frac{\hat{Y}_q(\xi) - e^{-\frac{\xi^2}{2}}}{\xi} d\xi + \frac{1}{T}. \quad (21)$$

We take  $T := \text{Var}[X_q]$ . By Lemma 3.4, the part of the integral with  $|\xi| \leq \frac{3}{5} \text{Var}[X_q]^{\frac{1}{4}}$  is at most

$$\begin{aligned} \int_{-\frac{3}{5} \text{Var}[X_q]^{\frac{1}{4}}}^{\frac{3}{5} \text{Var}[X_q]^{\frac{1}{4}}} \frac{e^{-\frac{\xi^2}{2}} (e^{O(\frac{\xi^4}{\rho(q) \log q'})} - 1)}{\xi} d\xi \\ \ll \frac{1}{\rho(q) \log q'} \int_{\mathbb{R}} \xi^3 e^{-\frac{\xi^2}{2}} d\xi \ll \frac{1}{\rho(q) \log q'}. \end{aligned}$$

We now bound the remaining part of the integral using an argument analogous to [Fiorilli and Martin 2013, Proposition 2.14]. Fix  $0 \leq \lambda \leq \frac{5}{6}$ . By the properties of the Bessel function  $J_0(x)$ , we have that if  $|\xi| > \lambda/2$ , then whatever  $\gamma_\chi \in \mathbb{R}$  is,

$$\left| J_0 \left( \frac{2\xi}{\sqrt{\frac{1}{4} + \gamma_\chi^2}} \right) \right| \leq J_0 \left( \frac{\lambda}{\sqrt{\frac{1}{4} + \gamma_\chi^2}} \right).$$

By (19), this shows that in the range  $|\xi| > \frac{5}{12} \text{Var}[X_q]^{\frac{1}{4}}$  we have  $|\hat{X}_q(\xi)| \leq |\hat{X}_q(\frac{5}{12} \text{Var}[X_q]^{\frac{1}{4}})|$  (since  $\text{Var}[X_q] \geq 1$ ), and so

$$\begin{aligned} & \int_{\frac{3}{5} \text{Var}[X_q]^{\frac{1}{4}} < |\xi| \leq \text{Var}[X_q]} \frac{\hat{Y}_q(\xi) - e^{-\frac{\xi^2}{2}}}{\xi} d\xi \\ & \ll \hat{Y}_q\left(\frac{5}{12} \text{Var}[X_q]^{\frac{1}{4}}\right) \log \text{Var}[X_q] + \int_{|\xi| > \frac{3}{5} \text{Var}[X_q]^{\frac{1}{4}}} \frac{e^{-\frac{\xi^2}{2}}}{\xi} d\xi \\ & \ll \exp\left(-\frac{25}{577} \text{Var}[X_q]^{\frac{1}{2}}\right) + \exp\left(-\frac{9}{51} \text{Var}[X_q]^{\frac{1}{2}}\right), \end{aligned}$$

by (18). Applying Lemma 3.2, we conclude that the right-hand side of (21) is at most a constant times  $(\rho(q) \log q')^{-1}$ .  $\square$

*Proof of Theorem 1.1, second part.* Fix  $\eta \in [\frac{1}{2}, 1]$ . We wish to find a sequence of moduli  $\{q_n\}$  such that  $\delta(q_n, NR, R) \rightarrow \eta$ . The case  $\eta = 1$  was already covered in part (1), and the case  $\eta = \frac{1}{2}$  follows from taking prime values of  $q$ , by the central limit theorem [Rubinstein and Sarnak 1994]. Therefore we can assume that  $\frac{1}{2} < \eta < 1$ .

Let  $\kappa > 0$  be the unique real solution of the equation

$$\frac{1}{\sqrt{2\pi}} \int_{-\kappa}^{\infty} e^{-\frac{t^2}{2}} dt = \eta.$$

Moreover, let  $\{q_n\}$  be the sequence of squarefree odd integers from Lemma 3.3 for which

$$2^{\omega(q_n)+1} = \log q'_n(\kappa^2 + o(1)).$$

By (17), this gives that as  $n \rightarrow \infty$ ,

$$B(q_n) := \frac{\mathbb{E}[X_{q_n}]}{\sqrt{\text{Var}[X_{q_n}]}} \rightarrow \kappa.$$

Define

$$Y_{q_n} := \frac{X_{q_n} - \mathbb{E}[X_{q_n}]}{\sqrt{\text{Var}[X_{q_n}]}} = \frac{X_{q_n}}{\sqrt{\text{Var}[X_{q_n}]}} - B(q_n).$$

We will use the central limit theorem of Lemma 3.4, as well as the Berry–Esseen inequality (21). Denoting by  $F_{q_n}$  the distribution function of  $Y_{q_n}$  and by  $F$  that of the Gaussian distribution, we have that



$$\begin{aligned}
|\delta(q_n, NR, R) - \eta| &= |\text{Prob}[X_{q_n} > 0] - \eta| = |\text{Prob}[X_{q_n} \leq 0] - (1 - \eta)| \\
&= |F_{q_n}(-B(q_n)) - F(-\kappa)| \\
&\leq |F_{q_n}(-B(q_n)) - F(-B(q_n))| + |F(-B(q_n)) - F(-\kappa)| \\
&\ll \frac{1}{\rho(q_n) \log q'_n} + |\kappa - B(q_n)|,
\end{aligned}$$

by Lemma 3.5 and by the fact that the probability density function of the Gaussian is bounded on  $\mathbb{R}$ . Looking at the proof of Lemma 3.3, we see that  $\rho(q_n) \rightarrow \infty$ , hence this last quantity tends to zero as  $n \rightarrow \infty$ , concluding the proof.  $\square$

#### 4. A more precise estimation of the bias using the theory of large deviations

To give a more precise estimate for the bias we are interested in under LI, we use the theory of large deviations of independent random variables. The fundamental estimate of this section is given in the following theorem.

**Theorem 4.1** [Montgomery and Odlyzko 1988, Theorem 2]. *For  $n = 1, 2, \dots$ , let  $Y_n$  be independent real-valued random variables such that  $\mathbb{E}[Y_n] = 0$  and  $|Y_n| \leq 1$ . Suppose that there is a constant  $c > 0$  such that  $\mathbb{E}[Y_n^2] \geq c$  for all  $n$ . Put  $Y = \sum r_n Y_n$  where  $\sum r_n^2 < \infty$ .*

*If  $\sum_{|r_n| \geq \alpha} |r_n| \leq V/2$  then*

$$\text{Prob}[Y \geq V] \leq \exp\left(-\frac{1}{16} V^2 \left(\sum_{|r_n| < \alpha} r_n^2\right)^{-1}\right).$$

*If  $\sum_{|r_n| \geq \alpha} |r_n| \geq 2V$  then*

$$\text{Prob}[Y \geq V] \geq a_1 \exp\left(-a_2 V^2 \left(\sum_{|r_n| < \alpha} r_n^2\right)^{-1}\right).$$

Here  $a_1 > 0$  and  $a_2 > 0$  depend only on  $c$ .

To make use of these bounds we need to give estimates on sums over zeros.

**Lemma 4.2.** *For  $T \geq 1$ , we have*

$$\sum_{|\gamma_\chi| < T} \frac{1}{\sqrt{\frac{1}{4} + \gamma_\chi^2}} = \frac{1}{\pi} \log(q^* \sqrt{T}) \log T + O(\log(q^* T)).$$

*Proof.* We start from the Riemann–von Mangoldt formula,

$$N(T, \chi) = \frac{T}{\pi} \log \frac{q^* T}{2\pi e} + O(\log q^* T).$$

With a summation by parts we get

$$\begin{aligned}
 \sum_{|\gamma_\chi| < T} \frac{1}{\sqrt{\frac{1}{4} + \gamma_\chi^2}} &= O(\log q^*) + \int_1^T \frac{dN(t, \chi)}{\sqrt{\frac{1}{4} + t^2}} \\
 &= \frac{N(T, \chi)}{\sqrt{\frac{1}{4} + T^2}} + \int_1^T \frac{tN(t, \chi)}{(\frac{1}{4} + t^2)^{\frac{3}{2}}} dt + O(\log q^*) \\
 &= \int_1^T \frac{t^2 \log \frac{q^* t}{2\pi e}}{(\frac{1}{4} + t^2)^{\frac{3}{2}}} dt + O(\log(q^* T)) \\
 &= \frac{1}{\pi} \log(q^* \sqrt{T}) \log T + O(\log q^* T). \quad \square
 \end{aligned}$$

**Lemma 4.3.** Assume LI, and let  $\mathcal{F}(q)$  be a subset of the group of Dirichlet characters mod  $q$  such that  $\chi \in \mathcal{F}(q) \Rightarrow \bar{\chi} \in \mathcal{F}(q)$ . Define the random variable

$$Y := \sum_{\chi \in \mathcal{F}(q)} \sum_{\gamma_\chi > 0} \frac{2\Re(Z_{\gamma_\chi})}{\sqrt{\frac{1}{4} + \gamma_\chi^2}},$$

where the  $Z_{\gamma_\chi}$  are i.i.d. uniformly distributed on the unit circle. Then, we have for  $q$  large enough that

$$a_1 \exp\left(-a_2 \frac{|\mathcal{F}(q)|}{L(q)}\right) \leq \text{Prob}[Y \geq |\mathcal{F}(q)|] \leq \exp\left(-a_3 \frac{|\mathcal{F}(q)|}{L(q)}\right),$$

where the  $a_i$  are absolute constants and

$$L(q) := \frac{\sum_{\chi \in \mathcal{F}(q)} \log q^*}{|\mathcal{F}(q)|} \geq \frac{\log 2}{2}.$$

*Proof.* This is a direct application of Theorem 4.1. Taking the sequence  $\{r_i\}$  to be the  $2/\sqrt{\frac{1}{4} + \gamma_\chi^2}$  ordered by size, and denoting by  $C$  the constant  $\sqrt{4/\alpha^2 - 1/4}$ , we have for  $0 < \alpha \leq 4$  that

$$\begin{aligned}
 \sum_{|r_n| \geq \alpha} |r_n| &= \sum_{\chi \in \mathcal{F}(q)} \sum_{0 < \gamma_\chi \leq C} \frac{2}{\sqrt{\frac{1}{4} + \gamma_\chi^2}}, \\
 \sum_{|r_n| > \alpha} |r_n|^2 &= \sum_{\chi \in \mathcal{F}(q)} \sum_{\gamma_\chi > C} \frac{4}{\frac{1}{4} + \gamma_\chi^2}.
 \end{aligned}$$

For the upper bound we take  $\alpha = 4$ , then we trivially have  $\sum_{|r_n| \geq \alpha} |r_n| \leq |\mathcal{F}(q)|/2$ , so

$$\text{Prob}[Y \geq |\mathcal{F}(q)|] \leq \exp\left(-\frac{1}{16} |\mathcal{F}(q)|^2 \left(c_1 \sum_{\chi \in \mathcal{F}(q)} \log q^*\right)^{-1}\right)$$

for some absolute constant  $c_1$ . For the lower bound we take  $\alpha = 2/\sqrt{\frac{1}{4} + T_0^2}$ , where  $T_0 > 1$  is a fixed large real number (independent of  $q$  and  $\mathcal{F}(q)$ ) such that

$$\sum_{\chi \in \mathcal{F}(q)} \sum_{|\gamma_\chi| \leq T_0} \frac{1}{\sqrt{\frac{1}{4} + \gamma_\chi^2}} \geq \frac{4}{\log 2} L(q) |\mathcal{F}(q)| \geq 2 |\mathcal{F}(q)|,$$

whose existence is granted by Lemma 4.2 (we grouped together conjugate characters). Then Theorem 4.1 gives the bound

$$\begin{aligned} \text{Prob}[Y \geq |\mathcal{F}(q)|] &\geq c_2 \exp\left(-c_3 |\mathcal{F}(q)|^2 \left(\sum_{\chi \in \mathcal{F}(q)} \sum_{\gamma_\chi > T_0} \frac{4}{\frac{1}{4} + \gamma_\chi^2}\right)^{-1}\right) \\ &\geq c_2 \exp\left(-c_3 |\mathcal{F}(q)|^2 \left(c_4 \sum_{\chi \in \mathcal{F}(q)} \log q^*\right)^{-1}\right) \end{aligned}$$

for  $q$  large enough and some absolute constants  $c_2$ ,  $c_3$  and  $c_4$ , since if we choose  $T_1 > T_0$  independent of  $\chi$  and large enough such that  $N(2T_1, \chi) - N(T_1, \chi) \gg \log q^*$  (this is possible by the Riemann–von Mangoldt formula), then we have

$$\begin{aligned} \sum_{\chi \in \mathcal{F}(q)} \sum_{\gamma_\chi > T_0} \frac{4}{\frac{1}{4} + \gamma_\chi^2} &\geq \sum_{\chi \in \mathcal{F}(q)} \sum_{T_1 < \gamma_\chi < 2T_1} \frac{4}{\frac{1}{4} + \gamma_\chi^2} \\ &\geq \sum_{\chi \in \mathcal{F}(q)} \frac{4}{\frac{1}{4} + (2T_1)^2} (N(2T_1, \chi) - N(T_1, \chi)) \\ &\gg \sum_{\chi \in \mathcal{F}(q)} \log q^*. \end{aligned} \quad \square$$

*Proof of Theorem 1.6.* Let  $X_q$  be the random variable in (14) and define the symmetric random variable

$$Y_q := X_q - \mathbb{E}[X_q].$$

By Lemma 3.1,

$$\begin{aligned} \delta(q; NR, R) &= \text{Prob}[X_q > 0] = \text{Prob}[Y_q > -\mathbb{E}[X_q]] \\ &= \text{Prob}[Y_q < \mathbb{E}[X_q]] = 1 - \text{Prob}[Y_q \geq \mathbb{E}[X_q]]. \end{aligned}$$

The proof follows by taking  $\mathcal{F}(q) := \{\chi \bmod q : \chi^2 = \chi_0, \chi \neq \chi_0\}$  in Lemma 4.3 and by estimating  $L(q)$  as in the proof of Lemma 3.2.  $\square$

## 5. A more general analysis

In this section we do a more general analysis by studying arbitrary linear combinations of prime counting functions.

Throughout the section,  $\vec{a} = (a_1, \dots, a_k)$  will be a vector of invertible reduced residues mod  $q$  and  $\vec{\alpha} = (\alpha_1, \dots, \alpha_k)$  will be a nonzero vector of real numbers such that  $\sum_{i=1}^k \alpha_i = 0$ . Recall that

$$\epsilon_i = \begin{cases} 1 & \text{if } a_i \equiv \square \pmod{q}, \\ 0 & \text{if } a_i \not\equiv \square \pmod{q}, \end{cases}$$

and we assume without loss of generality that

$$\sum_{i=1}^k \epsilon_i \alpha_i < 0.$$

To prove Theorems 1.10, 1.14 and 1.17, we need a few lemmas.

**Lemma 5.1.** *Assume GRH and LI. Then the quantity*

$$E(y; q, \vec{a}; \vec{\alpha}) := \phi(q) \frac{\alpha_1 \pi(e^y; q, a_1) + \dots + \alpha_k \pi(e^y; q, a_k)}{e^{y/2}/y}$$

*has the same distribution as the random variable*

$$X_{q; \vec{a}, \vec{\alpha}} := -\rho(q) \sum_{i=1}^k \epsilon_i \alpha_i + \sum_{\chi \neq \chi_0} |\alpha_1 \chi(a_1) + \dots + \alpha_k \chi(a_k)| \sum_{\gamma_\chi > 0} \frac{2\Re(Z_{\gamma_\chi})}{\sqrt{\frac{1}{4} + \gamma_\chi^2}}, \quad (22)$$

*where the  $Z_{\gamma_\chi}$  are independent random variables following a uniform distribution on the unit circle in  $\mathbb{C}$ .*

**Remark 5.2.** If we take  $a_1, \dots, a_{\phi(q)(1-\rho(q)^{-1})}$  to be the set of all quadratic nonresidues mod  $q$  with  $\alpha_1 = \dots = \alpha_{\phi(q)(1-\rho(q)^{-1})} = 1/\phi(q)$ , and we take  $a_{\phi(q)(1-\rho(q)^{-1})+1}, \dots, a_{\phi(q)}$  to be the set of all quadratic residues mod  $q$  with  $\alpha_{\phi(q)(1-\rho(q)^{-1})+1} = \dots = \alpha_{\phi(q)} = (1 - \rho(q))/\phi(q)$ , then we recover formula (14).

*Proof.* In the same way as in the proof of Lemma 3.1, we get by the explicit formula and by applying GRH that

$$\begin{aligned} F(y; q, \vec{a}, \vec{\alpha}) &:= \phi(q) \frac{\alpha_1 \psi(e^y; q, a_1) + \dots + \alpha_k \psi(e^y; q, a_k)}{e^{y/2}} \\ &= - \sum_{\chi \neq \chi_0} (\alpha_1 \bar{\chi}(a_1) + \dots + \alpha_k \bar{\chi}(a_k)) \sum_{\gamma_\chi} \frac{e^{i\gamma_\chi y}}{\rho_\chi} + o_q(1) \end{aligned}$$

(the main terms are canceled since  $\sum_{i=1}^k \alpha_i = 0$ ). By [Rubinstein and Sarnak 1994],  $F(y; q, \vec{a}, \vec{\alpha})$  has the same distribution as  $X_{q; \vec{a}, \vec{\alpha}} - \mathbb{E}[X_{q; \vec{a}, \vec{\alpha}}]$ , since LI implies that there are no real zeros. The second step is to use summation by parts and to remove

squares and other prime powers; this gives that

$$E(y; q, \vec{a}, \vec{\alpha}) + \rho(q) \sum_{i=1}^k \epsilon_i \alpha_i + o(1) = F(y; q, \vec{a}, \vec{\alpha}),$$

completing the proof.  $\square$

Before we give a bound on the variance of this distribution, we prove a lemma about conductors.

**Lemma 5.3.** *Let  $1 \leq L \leq \phi(q)$ . Then,*

$$\#\{\chi \bmod q : q^* \leq L\} \leq \min\{L\tau(q), L^2\}.$$

*Proof.* Denoting by  $\phi^*(d)$  the number of primitive characters modulo  $d$ , we have

$$\sum_{\substack{d|q \\ d \leq L}} \phi^*(d) \leq \min\left\{\sum_{d \leq L} d, L \sum_{d|q} 1\right\}. \quad \square$$

**Lemma 5.4.** *Assume LI. Let  $V(q; \vec{a}, \vec{\alpha}) := \text{Var}[X_{q; \vec{a}, \vec{\alpha}}]$ , where  $X_{q; \vec{a}, \vec{\alpha}}$  is the random variable defined in (22). Then,*

$$\phi(q) \|\vec{a}\|_2^2 \log \frac{3\phi(q)}{k} \ll V(q; \vec{a}, \vec{\alpha}) \ll \phi(q) \|\vec{a}\|_2^2 \log q, \quad (23)$$

where

$$\|\vec{a}\|_2^2 := \sum_{i=1}^k \alpha_i^2.$$

**Remark 5.5.** The upper bound in (23) is attained when  $q$  is prime by Lemma 5.8. As for the lower bound, if we take moduli  $q$  with a fixed set of distinct prime factors (for instance powers of a fixed prime) and consider the race between residues and non-residues with the weights of Remark 5.2, we obtain by Lemma 3.2 that  $V(q; \vec{a}, \vec{\alpha}) = O(1)$ , and this is of the same order of magnitude as the lower bound in (23).

*Proof.* Since the  $Z_{\gamma_\chi}$  in (22) are independent and have variance  $\frac{1}{2}$ , we have that

$$\text{Var}[X_{q; \vec{a}, \vec{\alpha}}] = \sum_{\chi \neq \chi_0} |\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)|^2 \sum_{\gamma_\chi} \frac{1}{\frac{1}{4} + \gamma_\chi^2} \quad (24)$$

(LI implies there are no real zeros). Combining this with Lemma 2.8 gives

$$V(q; \vec{a}, \vec{\alpha}) \asymp \sum_{\chi \neq \chi_0} |\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)|^2 \log q^*. \quad (25)$$

Now,  $\alpha_1 \chi_0(a_1) + \cdots + \alpha_k \chi_0(a_k) = 0$ , so

$$\begin{aligned}
 \sum_{\chi \neq \chi_0} |\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)|^2 &= \sum_{\chi \bmod q} |\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)|^2 \\
 &= \sum_{1 \leq i, j \leq k} \alpha_i \alpha_j \sum_{\chi \bmod q} \chi(a_i a_j^{-1}) \\
 &= \phi(q) \sum_{i=1}^k \alpha_i^2.
 \end{aligned} \tag{26}$$

Using this and (25), the upper bound follows from the fact that  $\log q^* \leq \log q$ . This also gives the lower bound  $V(q; \vec{a}, \vec{\alpha}) \geq (\log 3) \phi(q) \|\vec{\alpha}\|_2^2$ , which proves the claim for bounded values of  $\phi(q)/k$ . Hence we assume from now on that  $\phi(q)/k \geq 576$ . We fix a parameter  $1 < L < \phi(q)$  and discard the characters of conductor at most  $L$ :

$$\begin{aligned}
 V(q; \vec{a}, \vec{\alpha}) &\gg \log L \sum_{\substack{\chi \bmod q: \\ q^* > L}} |\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)|^2 \\
 &= \log L \sum_{1 \leq i, j \leq k} \alpha_i \alpha_j \sum_{\substack{\chi \bmod q: \\ q^* > L}} \chi(a_i a_j^{-1}) \\
 &= \log L \left[ \sum_{i=1}^k \alpha_i^2 \sum_{\substack{\chi \bmod q \\ q^* > L}} 1 + \sum_{1 \leq i \neq j \leq k} \alpha_i \alpha_j \sum_{\substack{\chi \bmod q: \\ q^* > L}} \chi(a_i a_j^{-1}) \right],
 \end{aligned}$$

which by Lemma 5.3 and the orthogonality relations is

$$\begin{aligned}
 &\geq \log L \left[ \sum_{i=1}^k \alpha_i^2 (\phi(q) - \min\{L\tau(q), L^2\}) - \sum_{1 \leq i \neq j \leq k} |\alpha_i \alpha_j| \min\{L\tau(q), L^2\} \right] \\
 &\geq \log L \|\vec{\alpha}\|_2^2 [\phi(q) - (k+1) \min\{L\tau(q), L^2\}]
 \end{aligned}$$

by the Cauchy–Schwarz inequality. Taking  $L := (3\phi(q)/k)^{\frac{1}{3}}$  gives the result, since then  $\phi(q)/k \geq 576$  implies that  $(k+1)L^2 \leq \phi(q)/2$ .  $\square$

**Remark 5.6.** In the last proof, we did not lose a lot by discarding the characters of conductor at most  $(3\phi(q)/k)^{\frac{1}{3}}$ , since by (26) their contribution is

$$\ll \phi(q) \|\vec{\alpha}\|_2^2 \log \frac{3\phi(q)}{k}.$$

*Proof of Theorem 1.14.* We have by Lemma 5.4 that there exists an absolute constant  $c > 0$  such that

$$B(q; \vec{a}, \vec{\alpha}) := \frac{\mathbb{E}[X_{q; \vec{a}, \vec{a}}]}{\sqrt{\text{Var}[X_{q; \vec{a}, \vec{a}}]}} \geq \frac{\rho(q) \left| \sum_{i=1}^k \epsilon_i \alpha_i \right|}{\sqrt{c \phi(q) \log q \sum_{i=1}^k \alpha_i^2}},$$

a quantity which is greater or equal to  $(c\epsilon)^{-\frac{1}{2}}$  by the condition of the theorem. We conclude that  $1 - \delta(q; \vec{a}, \vec{\alpha}) \leq c\epsilon$  by using Chebyshev's bound in the same way as in the proof of Theorem 1.1.  $\square$

*Proof of Theorem 1.10.* It is a particular case of Theorem 1.14.  $\square$

We now prove our negative results. To do so, we need to provide a central limit theorem, analogous to Lemma 3.4.

**Lemma 5.7.** *Assume LI, and let*

$$Y_{q; \vec{a}, \vec{\alpha}} := \frac{X_{q; \vec{a}, \vec{\alpha}} - \mathbb{E}[X_{q; \vec{a}, \vec{\alpha}}]}{\sqrt{\text{Var}[X_{q; \vec{a}, \vec{\alpha}}]}}.$$

*The characteristic function of  $Y_{q; \vec{a}, \vec{\alpha}}$  satisfies*

$$\log \hat{Y}_{q; \vec{a}, \vec{\alpha}}(\xi) = -\frac{\xi^2}{2} + O\left(\frac{\xi^4}{\log(3\phi(q)/k)} \min\left\{1, \frac{k^2 \log q}{\phi(q) \log(3\phi(q)/k)}\right\}\right)$$

*in the range  $|\xi| \leq 3/(5\|\vec{\alpha}\|_1)$ , where  $\|\vec{\alpha}\|_1 := \sum_{i=1}^k |\alpha_i|$ .*

*Proof.* As in Lemma 3.4, we compute

$$\begin{aligned} \log \hat{X}_{q; \vec{a}, \vec{\alpha}}(\xi) &= i \mathbb{E}[X_{q; \vec{a}, \vec{\alpha}}] \xi + \sum_{\chi \neq \chi_0} \sum_{\gamma_\chi > 0} \log \left( J_0 \left( \frac{2|\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)| \xi}{\sqrt{\frac{1}{4} + \gamma_\chi^2}} \right) \right). \end{aligned}$$

We now use the Taylor expansion (20), which is valid as soon as  $|\xi| \leq 3/(5\|\vec{\alpha}\|_1)$ , since under this condition we have

$$\frac{2|\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)| |\xi|}{\sqrt{\frac{1}{4} + \gamma_\chi^2}} \leq \frac{2\|\vec{\alpha}\|_1}{\frac{1}{2}} \frac{3}{5\|\vec{\alpha}\|_1} = \frac{12}{5}.$$

Applying Lemma 2.8, we obtain that

$$\begin{aligned} \log \widehat{Y}_{q;\vec{a},\vec{\alpha}}(\xi) \\ = -\frac{\xi^2}{2} + O\left(\xi^4 \frac{\sum_{\chi \neq \chi_0} |\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)|^4 \log q^*}{(\sum_{\chi \neq \chi_0} |\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)|^2 \log q^*)^2}\right). \end{aligned} \quad (27)$$

If  $\phi(q)/k$  is bounded, then the statement trivially follows from the bound  $\sum_i a_i^4 \leq (\sum_i a_i^2)^2$ . Therefore we assume from now on that  $\phi(q)/k \geq 576$ .

We now use two different approaches to bound the error term. The first idea is to “factor out  $\sqrt{\log q^*}$ ” before applying the trivial inequality  $\sum_i a_i^4 \leq (\sum_i a_i^2)^2$ . We have seen in Remark 5.6 that the main contribution to the variance is that of the characters with  $q^* \geq L := (3\phi(q)/k)^{\frac{1}{3}}$ . We use the same idea here. Setting  $\Theta_\chi := |\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)|^2$ , we have

$$\begin{aligned} \sum_{\chi \neq \chi_0} \Theta_\chi \log q^* &\geq \sum_{\substack{\chi \neq \chi_0 \\ q^* > L}} \Theta_\chi \log q^* \\ &\geq \sqrt{\log L} \sum_{\substack{\chi \neq \chi_0 \\ q^* > L}} \Theta_\chi \sqrt{\log q^*} \\ &\geq \sqrt{\log L} \left( \sum_{\chi \neq \chi_0} \Theta_\chi \sqrt{\log q^*} - kL^2 \sqrt{\log L} \|\vec{\alpha}\|_2^2 \right) \end{aligned} \quad (28)$$

by Lemma 5.3 and the Cauchy–Schwarz inequality. Now, by our choice of  $L$ , the fact that  $\phi(q)/k \geq 576$  and the equality  $\sum_{\chi \neq \chi_0} \Theta_\chi = \phi(q) \|\vec{\alpha}\|_2^2$  (see (26)), we have

$$\begin{aligned} kL^2 \sqrt{\log L} \|\vec{\alpha}\|_2^2 &\leq \frac{1}{2} \sqrt{\log L} \left[ \phi(q) \|\vec{\alpha}\|_2^2 - kL^2 \|\vec{\alpha}\|_2^2 \right] \\ &\leq \frac{1}{2} \sum_{\substack{\chi \neq \chi_0 \\ q^* \geq L}} \Theta_\chi \sqrt{\log q^*} \leq \frac{1}{2} \sum_{\chi \neq \chi_0} \Theta_\chi \sqrt{\log q^*}, \end{aligned}$$

hence (28) gives that

$$\sum_{\chi \neq \chi_0} \Theta_\chi \log q^* \gg \sqrt{\log L} \sum_{\chi \neq \chi_0} \Theta_\chi \sqrt{\log q^*}.$$

Plugging this into (27) and using the trivial bound

$$\sum_{\chi \neq \chi_0} \Theta_\chi^2 \log q^* \leq \left( \sum_{\chi \neq \chi_0} \Theta_\chi \sqrt{\log q^*} \right)^2,$$

we get that the error term is  $\ll \xi^4 / \log L$ .



For the second upper bound we use Lemma 5.4 and the Cauchy–Schwarz inequality:

$$\begin{aligned}
 & \frac{\sum_{\chi \neq \chi_0} |\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)|^4 \log q^*}{\left(\sum_{\chi \neq \chi_0} |\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)|^2 \log q^*\right)^2} \\
 & \ll \frac{\log q}{\log(3\phi(q)/k)^2} \frac{\sum_{\chi \neq \chi_0} |\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)|^4}{(\phi(q) \|\vec{\alpha}\|_2^2)^2} \\
 & = \frac{\log q}{\log(3\phi(q)/k)^2} \frac{\sum_{\substack{i,j,i',j' \\ a_i a_{i'} \equiv a_j a_{j'} \pmod q}} \alpha_i \alpha_{i'} \alpha_j \alpha_{j'}}{\phi(q) \|\vec{\alpha}\|_2^4} \\
 & \leq \frac{\log q}{\log(3\phi(q)/k)^2} \frac{\left(\sqrt{\sum_{i=1}^k \alpha_i^2} \sqrt{\sum_{j=1}^k 1}\right)^4}{\phi(q) \|\vec{\alpha}\|_2^4},
 \end{aligned}$$

which gives the claimed bound.  $\square$

*Proof of Theorem 1.17.* Let  $K \geq 1$  and define  $c > 0$  to be the constant implied in the lower bound in Lemma 5.4. Assume that  $k \leq e^{-e^{4K}} \phi(q)$  and that (9) holds with  $K_2 = cK$ . Define the vector  $\vec{\beta} := (e^{-K} / \|\vec{\alpha}\|_1) \vec{\alpha}$ , so that  $\|\vec{\beta}\|_1 = e^{-K}$ , which will allow us to apply Lemma 5.7. Clearly,

$$\delta(q; \vec{a}, \vec{\alpha}) = \delta(q; \vec{a}, \vec{\beta}),$$

since multiplying  $\vec{\alpha}$  by a positive constant does not affect the inequality

$$\alpha_1 \pi(n; q, a_1) + \cdots + \alpha_k \pi(n; q, a_k) > 0.$$

We have by Lemma 5.4 and by the definition of  $c$  that

$$\begin{aligned}
 B(q; \vec{a}, \vec{\beta}) &:= \frac{\mathbb{E}[X_{q; \vec{a}, \vec{\beta}}]}{\sqrt{\text{Var}[X_{q; \vec{a}, \vec{\beta}}]}} \leq \frac{\rho(q) \left| \sum_{i=1}^k \epsilon_i \beta_i \right|}{\sqrt{c \phi(q) \log(3\phi(q)/k) \sum_{i=1}^k \beta_i^2}} \\
 &= c^{-\frac{1}{2}} \frac{\rho(q) \left| \sum_{i=1}^k \epsilon_i \alpha_i \right|}{\sqrt{\phi(q) \log(3\phi(q)/k) \sum_{i=1}^k \alpha_i^2}},
 \end{aligned}$$

a quantity which is at most  $\sqrt{K}$  by (9). Defining

$$Y_{q; \vec{a}, \vec{\beta}} := \frac{X_{q; \vec{a}, \vec{\beta}} - \mathbb{E}[X_{q; \vec{a}, \vec{\beta}}]}{\sqrt{\text{Var}[X_{q; \vec{a}, \vec{\beta}}]}},$$

we have by Lemma 5.7 and by our condition on  $k$  that in the range  $|\xi| \leq \frac{3}{5}e^K$ ,

$$\log \hat{Y}_{q;\vec{a},\vec{\beta}}(\xi) = -\frac{\xi^2}{2} + O\left(\frac{\xi^4}{e^{4K}}\right).$$

Combining this with the Berry–Esseen inequality (21) and taking  $W$  to be a standard Gaussian random variable with mean 0 and variance 1, we get

$$\begin{aligned} \text{Prob}[Y_{q;\vec{a},\vec{\beta}} > -B(q;\vec{a},\vec{\beta})] - \text{Prob}[W > -B(q;\vec{a},\vec{\beta})] \\ &\ll \int_{-\frac{3}{5}e^K}^{\frac{3}{5}e^K} \frac{\hat{Y}_{q;\vec{a},\vec{\beta}}(\xi) - e^{-\frac{\xi^2}{2}}}{\xi} d\xi + \frac{5}{3}e^{-K} \\ &\ll \int_{-\frac{3}{5}e^K}^{\frac{3}{5}e^K} \frac{\xi^3 e^{-\frac{\xi^2}{2}}}{e^{4K}} d\xi + e^{-K} \\ &\ll e^{-K}. \end{aligned} \quad (29)$$

However, since  $B(q;\vec{a},\vec{\beta}) \leq \sqrt{K}$ , we have that

$$\text{Prob}[W \leq -B(q;\vec{a},\vec{\beta})] \geq c_1 \frac{e^{-\frac{K}{2}}}{K}$$

for some absolute constant  $c_1$ . Therefore, applying (29) gives

$$\begin{aligned} \delta(q;\vec{a},\vec{\beta}) &= \text{Prob}[Y_{q;\vec{a},\vec{\beta}} > -B(q;\vec{a},\vec{\beta})] \\ &= \text{Prob}[W > -B(q;\vec{a},\vec{\beta})] + O(e^{-K}) \\ &\leq 1 - c_1 e^{-\frac{K}{2}} / K + c_2 e^{-K}, \end{aligned}$$

a quantity which is less than the right-hand side of (10) for  $K$  large enough. The proof is finished since  $\delta(q;\vec{a},\vec{\alpha}) = \delta(q;\vec{a},\vec{\beta})$ .  $\square$

To end this section we give an estimate for the variance  $V(q;\vec{a},\vec{\alpha})$ . While we have not explicitly made use of this expression, we include it for its intrinsic interest, and for its ability to give a precise evaluation of the variance  $V(q;\vec{a},\vec{\alpha})$  for values of  $q$  having prescribed prime factors.

**Lemma 5.8.** *Assuming GRH and LI, we have that*

$$\begin{aligned} V(q;\vec{a},\vec{\alpha}) \\ = \phi(q)\|\vec{\alpha}\|_2(\log q + O(\log \log q)) - \phi(q) \sum_{i \neq j} \alpha_i \alpha_j \frac{\Lambda\left(\frac{q}{(q, a_i a_j^{-1} - 1)}\right)}{\phi\left(\frac{q}{(q, a_i a_j^{-1} - 1)}\right)}. \end{aligned} \quad (30)$$

*Proof.* Using [Fiorilli and Martin 2013, Proposition 3.3], we obtain that

$$\begin{aligned} \sum_{\chi \bmod q} |\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)|^2 \log q^* \\ = \sum_{1 \leq i, j \leq k} \alpha_i \alpha_j \sum_{\chi \bmod q} \chi(a_i a_j^{-1}) \log q^* \\ = \phi(q) \left( \log q - \sum_{p|q} \frac{\log p}{p-1} \right) \sum_{i=1}^k \alpha_i^2 - \phi(q) \sum_{i \neq j} \alpha_i \alpha_j \frac{\Lambda\left(\frac{q}{(q, a_i a_j^{-1} - 1)}\right)}{\phi\left(\frac{q}{(q, a_i a_j^{-1} - 1)}\right)}. \end{aligned}$$

We have

$$\sum_{p|q} \frac{\log p}{p-1} \leq \sum_{i=1}^{\omega(q)} \frac{\log p_i}{p_i - 1} \ll \log \log q,$$

where  $p_i$  denotes the  $i$ -th prime. The claimed estimate then follows by combining this with the formula

$$V(q; \vec{a}, \vec{\alpha}) = \sum_{\chi \bmod q} |\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)|^2 \sum_{\gamma_\chi} \frac{1}{\frac{1}{4} + \gamma_\chi^2}$$

(see (24)) and with (16). Note that by the Littlewood bound  $(L'/L)(1, \chi) \ll \log \log q^*$ , the implied error term is

$$\ll \sum_{\chi \bmod q} |\alpha_1 \chi(a_1) + \cdots + \alpha_k \chi(a_k)|^2 \log \log q = \phi(q) \|\vec{\alpha}\|_2 \log \log q. \quad \square$$

It might seem like the second term of (30) is an error term; however, this is not necessarily true for large values of  $k$  (see Lemma 3.2). Nevertheless, we expect many cancellations to occur since

$$\sum_{i \neq j} \alpha_i \alpha_j = \left( \sum_{i=1}^k \alpha_i \right)^2 - \sum_{i=1}^k \alpha_i^2 = - \sum_{i=1}^k \alpha_i^2.$$

### Acknowledgements

I would like to thank my former advisor Andrew Granville for his very interesting question which motivated this work, and for his comments and encouragement. I would also like to thank Enrico Bombieri and Peter Sarnak for their advice and their encouragement, and the Institute for Advanced Study for providing excellent research conditions. I thank Kevin Ford, Youness Lamzouri and the people at the University of Illinois at Urbana–Champaign for their hospitality and for very fruitful conversations which led me to consider general linear combinations of prime counting functions. I thank Jan-Christoph Schlage-Puchta for Remark 2.4

and for suggesting the current proof of Lemma 2.5. I also thank Carl Pomerance for suggesting Remark 1.9, and Greg Martin for suggesting Remark 1.12. I thank Barry Mazur for motivating me to weaken the linear independence hypothesis. I thank Byungchul Cha, Ke Gong, Peter Humphries, Chen Meiri, Nathan Ng and Anders Södergren for helpful conversations and for their comments. Finally I thank the anonymous referees for their careful reading of the manuscript and their very useful comments. This work was supported by an NSERC Postdoctoral Fellowship, as well as NSF grant DMS-0635607.

## References

- [Akbari et al. 2013] A. Akbari, N. Ng, and M. Shahabi, “Limiting distributions of the classical error terms of prime number theory”, preprint, 2013. arXiv 1306.1657
- [Bays and Hudson 2000] C. Bays and R. H. Hudson, “A new bound for the smallest  $x$  with  $\pi(x) > \text{li}(x)$ ”, *Math. Comp.* **69**:231 (2000), 1285–1296. MR 2001c:11138 Zbl 1042.11001
- [Bays et al. 2001] C. Bays, K. Ford, R. H. Hudson, and M. Rubinstein, “Zeros of Dirichlet  $L$ -functions near the real axis and Chebyshev’s bias”, *J. Number Theory* **87**:1 (2001), 54–76. MR 2001m:11148 Zbl 1009.11057
- [Besicovitch 1926] A. S. Besicovitch, “On generalized almost periodic functions”, *Proc. London Math. Soc.* **S2-25**:1 (1926), 495–512. MR 1575297 JFM 52.0263.01
- [Besicovitch 1955] A. S. Besicovitch, *Almost periodic functions*, Dover, New York, 1955. MR 16,817a Zbl 0065.07102
- [Bober et al. 2014] J. Bober, J. B. Conrey, D. W. Farmer, A. Fujii, S. Koutsoliotas, S. Lemurell, M. Rubinstein, and H. Yoshida, “The highest lowest zero of general  $L$ -functions”, preprint, 2014. arXiv 1211.5996
- [Esseen 1945] C.-G. Esseen, “Fourier analysis of distribution functions: A mathematical study of the Laplace–Gaussian law”, *Acta Math.* **77** (1945), 1–125. MR 7,312a Zbl 0060.28705
- [Feuerverger and Martin 2000] A. Feuerverger and G. Martin, “Biases in the Shanks–Rényi prime number race”, *Experiment. Math.* **9**:4 (2000), 535–570. MR 2002d:11111 Zbl 0976.11041
- [Fiorilli and Martin 2013] D. Fiorilli and G. Martin, “Inequities in the Shanks–Rényi prime number race: An asymptotic formula for the densities”, *J. Reine Angew. Math.* **676** (2013), 121–212. MR 3028758 Zbl 1276.11150
- [Ford 2008a] K. Ford, “The distribution of integers with a divisor in a given interval”, *Ann. of Math.* (2) **168**:2 (2008), 367–433. MR 2009m:11152 Zbl 1181.11058
- [Ford 2008b] K. Ford, “Integers with a divisor in  $(y, 2y]$ ”, pp. 65–80 in *Anatomy of integers*, edited by J.-M. De Koninck et al., CRM Proc. Lecture Notes **46**, Amer. Math. Soc., Providence, RI, 2008. MR 2009i:11113 Zbl 1175.11053
- [Granville and Martin 2006] A. Granville and G. Martin, “Prime number races”, *Amer. Math. Monthly* **113**:1 (2006), 1–33. MR 2006h:11112 Zbl 1139.11037
- [Hardy and Littlewood 1916] G. H. Hardy and J. E. Littlewood, “Contributions to the theory of the Riemann Zeta-function and the theory of the distribution of primes”, *Acta Math.* **41**:1 (1916), 119–196. MR 1555148 JFM 46.0498.01
- [Kaczorowski 1995] J. Kaczorowski, “On the distribution of primes (mod 4)”, *Analysis* **15**:2 (1995), 159–171. MR 96h:11095 Zbl 0826.11042

- [Lamzouri 2012] Y. Lamzouri, “Large deviations of the limiting distribution in the Shanks–Rényi prime number race”, *Math. Proc. Cambridge Philos. Soc.* **153**:1 (2012), 147–166. MR 2943671 Zbl 1286.11143
- [Landau 1918a] E. Landau, “Über einige ältere Vermutungen und Behauptungen in der Primzahltheorie, I”, *Math. Z.* **1** (1918), 1–24. JFM 46.0263.02
- [Landau 1918b] E. Landau, “Über einige ältere Vermutungen und Behauptungen in der Primzahltheorie, II”, *Math. Z.* **1**:2-3 (1918), 213–219. MR 1544293 JFM 46.0263.02
- [Miller 2002] S. D. Miller, “The highest lowest zero and other applications of positivity”, *Duke Math. J.* **112**:1 (2002), 83–116. MR 2003b:11047 Zbl 1014.11036
- [Monach 1980] W. R. Monach, *Numerical investigation of several problems in number theory*, Ph.D. thesis, University of Michigan, 1980, Available at <http://search.proquest.com/docview/302978421>.
- [Montgomery 1980] H. L. Montgomery, “The zeta function and prime numbers”, pp. 1–31 in *Proceedings of the Queen’s Number Theory Conference, 1979* (Kingston, ON, 1979), edited by P. Ribenboim, Queen’s Papers in Pure and Appl. Math. **54**, Queen’s Univ., Kingston, ON, 1980. MR 82k:10047 Zbl 0451.10023
- [Montgomery and Odlyzko 1988] H. L. Montgomery and A. M. Odlyzko, “Large deviations of sums of independent random variables”, *Acta Arith.* **49**:4 (1988), 427–434. MR 89m:11075 Zbl 0641.60032
- [Myerscough 2013] C. Myerscough, “Application of an accurate remainder term in the calculation of residue class distributions”, preprint, 2013. arXiv 1301.1434
- [Ng 2000] N. Ng, *Limiting distributions and zeros of Artin L-functions*, Ph.D. thesis, University of British Columbia, 2000, Available at <http://www.cs.uleth.ca/~nathanng/RESEARCH/phd.thesis.pdf>.
- [Ng 2004] N. Ng, “The distribution of the summatory function of the Möbius function”, *Proc. London Math. Soc.* (3) **89**:2 (2004), 361–389. MR 2005f:11215 Zbl 1138.11341
- [Odlyzko 1990] A. M. Odlyzko, “Bounds for discriminants and related estimates for class numbers, regulators and zeros of zeta functions: A survey of recent results”, *Sém. Théor. Nombres Bordeaux* (2) **2**:1 (1990), 119–141. MR 91i:11154 Zbl 0722.11054
- [Rubinstein and Sarnak 1994] M. Rubinstein and P. Sarnak, “Chebyshev’s bias”, *Experiment. Math.* **3**:3 (1994), 173–197. MR 96d:11099 Zbl 0823.11050
- [Schlage-Puchta 2000] J.-C. Puchta, “On large oscillations of the remainder of the prime number theorems”, *Acta Math. Hungar.* **87**:3 (2000), 213–227. MR 2001g:11146 Zbl 0963.11051
- [Skewes 1933] S. Skewes, “On the difference  $\pi(x) - \text{li}(x)$ , (I)”, *J. Lond. Math. Soc.* **S8-4**:4 (1933), 277–283. Zbl 0007.34003

Communicated by Andrew Granville

Received 2014-03-26

Revised 2014-03-26

Accepted 2014-05-23

daniel.fiorilli@uottawa.ca

Department of Mathematics and Statistics,  
University of Ottawa, 585 King Edward Avenue,  
Ottawa, Ontario, K1N 6N5, Canada



# Algebra & Number Theory

msp.org/ant

## EDITORS

### MANAGING EDITOR

Bjorn Poonen  
Massachusetts Institute of Technology  
Cambridge, USA

### EDITORIAL BOARD CHAIR

David Eisenbud  
University of California  
Berkeley, USA

## BOARD OF EDITORS

Georgia Benkart	University of Wisconsin, Madison, USA	Shigefumi Mori	RIMS, Kyoto University, Japan
Dave Benson	University of Aberdeen, Scotland	Raman Parimala	Emory University, USA
Richard E. Borcherds	University of California, Berkeley, USA	Jonathan Pila	University of Oxford, UK
John H. Coates	University of Cambridge, UK	Anand Pillay	University of Notre Dame, USA
J-L. Colliot-Thélène	CNRS, Université Paris-Sud, France	Victor Reiner	University of Minnesota, USA
Brian D. Conrad	University of Michigan, USA	Peter Sarnak	Princeton University, USA
Hélène Esnault	Freie Universität Berlin, Germany	Joseph H. Silverman	Brown University, USA
Hubert Flenner	Ruhr-Universität, Germany	Michael Singer	North Carolina State University, USA
Edward Frenkel	University of California, Berkeley, USA	Vasudevan Srinivas	Tata Inst. of Fund. Research, India
Andrew Granville	Université de Montréal, Canada	J. Toby Stafford	University of Michigan, USA
Joseph Gubeladze	San Francisco State University, USA	Bernd Sturmfels	University of California, Berkeley, USA
Roger Heath-Brown	Oxford University, UK	Richard Taylor	Harvard University, USA
Craig Huneke	University of Virginia, USA	Ravi Vakil	Stanford University, USA
Yujiro Kawamata	University of Tokyo, Japan	Michel van den Bergh	Hasselt University, Belgium
János Kollár	Princeton University, USA	Marie-France Vignéras	Université Paris VII, France
Yuri Manin	Northwestern University, USA	Kei-Ichi Watanabe	Nihon University, Japan
Barry Mazur	Harvard University, USA	Efim Zelmanov	University of California, San Diego, USA
Philippe Michel	École Polytechnique Fédérale de Lausanne	Shou-Wu Zhang	Princeton University, USA
Susan Montgomery	University of Southern California, USA		

## PRODUCTION

production@msp.org  
Silvio Levy, Scientific Editor

---

See inside back cover or [msp.org/ant](http://msp.org/ant) for submission instructions.

---

The subscription price for 2014 is US \$225/year for the electronic version, and \$400/year (+\$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscribers address should be sent to MSP.


---

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

---

ANT peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

 **mathematical sciences publishers**  
nonprofit scientific publishing

<http://msp.org/>

© 2014 Mathematical Sciences Publishers

# Algebra & Number Theory

Volume 8      No. 7      2014

---

Double Dirichlet series and quantum unique ergodicity of weight one-half Eisenstein series	1539
YIANNIS N. PETRIDIS, NICOLE RAULF and MORTEN S. RISAGER	
Monodromy and local-global compatibility for $l = p$	1597
ANA CARAIANI	
Finite generation of the cohomology of some skew group algebras	1647
VAN C. NGUYEN and SARAH WITHERSPOON	
On the supersingular locus of the $\mathrm{GU}(2,2)$ Shimura variety	1659
BENJAMIN HOWARD and GEORGIOS PAPPAS	
Poincaré–Birkhoff–Witt deformations of smash product algebras from Hopf actions on Koszul algebras	1701
CHELSEA WALTON and SARAH WITHERSPOON	
Highly biased prime number races	1733
DANIEL FIORILLI	
Bounded gaps between primes with a given primitive root	1769
PAUL POLLACK	



1937-0652(2014)8:7;1-4