HIGHLY BIASED PRIME NUMBER RACES

DANIEL FIORILLI

Abstract. 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 $\bmod q$. 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 $\mathrm{Li}(x)$ 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..., 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 exists prime number races whose density can be arbitrarily close to 1. An example of race whose density exceeds the above number is the race between quadratic residues and non-residues modulo 4849845, for which the density is 0.999999928... We also give fairly general criteria to decide whether a prime number race is highly biased or not. Our main result depends on the General Riemann Hypothesis and on 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 claims without proof that as $c \to 0$, we have

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

However, as Hardy and Littlewood [HL] and Landau [Lan1, Lan2] have shown, this statement is equivalent to the Riemann hypothesis for $L(s,\chi_{-4})$, where χ_{-4} denotes the primitive character modulo 4.

The modern way to study this question 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 [K]. To remedy to this problem we define the *logarithmic density* of a set $P \subset \mathbb{N}$ by

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

if the limit exists. In general we define $\underline{\delta}(P)$ and $\overline{\delta}(P)$ 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... (see [RS]).

The **General Riemann Hypothesis** states that for every primitive character χ mod q, all non-trivial 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 $\mathbb{Q}$.

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 [GM].

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

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

In other words, the bias dissolves as $q \to \infty$. For $r = 2$, this phenomenon can readily seen in [FiMa], where the authors exhibit the list of the 117 densities which are greater than or equal to 9/10. By the trivial inequality

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

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

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

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

$$\delta(1) := \delta(\{n : \mathrm{Li}(n) > \pi(n)\}) = 0.99999973...$$

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

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

and denote by $\delta(q; A, B)$ the logarithmic density of the set of x for which it is satisfied, if it 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 $q \to \infty$, $\delta(q; NR, R) \to \frac{1}{2}$ (but at a much slower rate than two-way races, see [FiMa]).

In general, one can see ([BFHR], [FiMa]) that low-lying zeros (excluding real zeros) have a significant effect on decreasing the bias. However, real zeros have the reverse effect, and increase the bias. Nonetheless, real zeros are very rare, in fact Chowla's conjecture asserts that Dirichlet L-functions never vanish in the interval $s \in (0,1]$.

Odlyzko [O] has shown that the Dirichlet L-function having the highest first zero in the critical strip is the Riemann zeta function, which is $\rho_0 = \frac{1}{2} + i \cdot 14.134725...$ Subsequently, Miller [Mi] generalized this result by showing that each member of a very large class of cuspidal GL_n L-functions 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 $1/2$ (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. By these considerations, one might conjecture that the highest density one will ever find by doing prime number races is $\delta(1) = 0.99999973...$

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 non-residue race as in [RS], but for a general modulus q. We take $A = NR := \{a \bmod q : a \equiv \square \bmod q\}$ and $B = R := \{b \bmod q : b \not\equiv \square \bmod q\}$ in (1). Note that $|B| = \phi(q)/\rho(q)$ and $|A| = \phi(q)\left(1 - \frac{1}{\rho(q)}\right)$, where

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

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*

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

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

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

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 hypothesis on the multiplicity of the elements of the multiset of all non-trivial 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.

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

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

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 $q = 2^e \prod_{\substack{p|q \\ p \neq 2}} p^{e_p}$ and $\ell := \prod_{\substack{p|q \\ p \neq 2}} p$. One can see that

$$\delta(q; NR, R) = \delta(2^{\min(3,e)}\ell; NR, R),$$

since 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 $e \geq 4$ (see Lemma 3.1). Therefore, when studying $\delta(q; NR, R)$ one can assume without loss of generality that q is of the form $2^m\ell$, where ℓ 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:

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

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

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 $\lambda = 1 - \frac{1 + \log\log 2}{\log 2} = 0.086071...$ Interestingly, Ford's work on integers having a divisor in a given interval (see [Fo1], [Fo2]) shows that these integers appear in the Erdős multiplication table.

In terms of random variables, this 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 $\mathbb{E}[X]/\sqrt{\mathrm{Var}[X]}$ is large. Instead, one should study the large deviations of $X - \mathbb{E}[X]$. The theory of large deviations of error terms arising from prime counting functions was initiated by Montgomery [Mo], and has since then been developed by Monach [Mn], Montgomery and Odlyzko [MoOd], Rubinstein and Sarnak [RS], and more recently Lamzouri [Lam]. Exploiting such ideas 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 absolute constants.

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

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	0.999999999954

Remark 1.7. As remarked by Rubinstein and Sarnak [RS], 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 following 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 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, races with normal moduli have a very moderate bias.

It is possible to analyse 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 $\overrightarrow{a} = (a_1, ..., a_k)$ a vector of invertible reduced residues modulo q and $\overrightarrow{\alpha} = (\alpha_1, ..., \alpha_k)$ a non-zero 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 $\overrightarrow{\alpha}$, that is we define

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

Moreover, we define

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

and we assume without loss of generality that

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

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

There are many choices of vectors $\overrightarrow{a}$ and $\overrightarrow{\alpha}$ 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* $k_R \le \frac{\rho(q)}{\phi(q)}$ *and* $k_N \le \left(1 - \frac{1}{\rho(q)}\right)\phi(q)$ *be two positive integers. Take* $a_1, ..., a_{k_N}$ *to be any distinct quadratic non-residues* mod$q$*, with coefficients* $\alpha_1 = ... = \alpha_{k_N} = k_R$*, and* $a_{k_N+1}, ..., a_{k_N+k_R}$ *to be any distinct quadratic residues* mod$q$*, with* $\alpha_{k_N+1} = ... = \alpha_{k_N+k_R} = -k_N$*. There exists an absolute constant* $c > 0$ *such that if for some* $0 < \epsilon < \frac{1}{2c}$ *we have*

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

then

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

Remark 1.11. Fix $0 < \epsilon < \frac{1}{2c}$ and define $N_\epsilon(q)$ to be to number of positive integers k_N, k_R for which $k_N \le (1 - \frac{1}{\rho(q)})\phi(q)$, $k_R \le \frac{\phi(q)}{\rho(q)}$, 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 $\rho(q) \ge \epsilon^{-2} \log q$, we have that $N_\epsilon(q)$ tends to infinity as $q \to \infty$. 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 $a_1, ..., a_{k_N+k_R}$ gives a race with $\delta(q; \overrightarrow{a}, \overrightarrow{\alpha}) = 1 - o(1)$.

Remark 1.13. In Theorem 1.1, we have $k_N = \left(1 - \frac{1}{\rho(q)}\right)\phi(q)$ and $k_R = \frac{\phi(q)}{\rho(q)}$, 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* $0 < \epsilon < \frac{1}{2c}$ *we have*

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

then

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

Remark 1.15. Trivially, one has

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

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 $k_R \leq \frac{\phi(q)}{\rho(q)}$, this imposes the following 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 $\delta(q; \overrightarrow{a}, \overrightarrow{\alpha})$. One can use the Montgomery-Odlyzko bounds [MoOd] 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 are very particular, in the sense that they must satisfy precise conditions.

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

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

then

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

(Hence this race cannot be too biased.)

Remark 1.18. Applying the Cauchy-Schwartz 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 $\overrightarrow{a}$ and $\overrightarrow{\alpha}$ 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

$$\pi(x_0) > Li(x_0).$$

This number has been extensively studied since Skewes' 1933 paper in which he showed under GRH that

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

The GRH assumption has since then be removed and the upper bound greatly reduced; we refer the reader to [BH] for the list of such improvements. The current record is due to Bays and Hudson [BH], 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 $\delta(q; a, b) \to \frac{1}{2}$ 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 we constructed, in fact we expect the Skewes number

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

to tend to infinity as $\rho(q)/\log q'$ tends to infinity (q' is the radical of q). One can then ask the following question: how fast does it tend to infinity? Similar arguments to those of Montgomery [Mo] and of Ng [N] allow us to make the speculation that the answer is double-exponentially.

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

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

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 2.3 and for suggesting the current proof of Lemma 2.4. 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. Finally, I thank Byungchul Cha, Ke Gong, Chen Meiri, Nathan Ng and Anders Södergren for helpful conversations. This work was supported by an NSERC Postdoctoral Fellowship, as well as NSF grant DMS-0635607.

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. *Assuming GRH, we have that*

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

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

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

The explicit formula gives

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

where ρ_χ runs over the non-trivial zeros of $L(s,\chi)$. The left hand side of (12) is equal to

$$\sum_{\substack{\chi \bmod q\\ \chi^2=\chi_0\\ \chi\neq\chi_0}} \sum_{p\leq x} \chi(p)\log p + \sum_{\substack{\chi \bmod q\\ \chi^2=\chi_0\\ \chi\neq\chi_0}} \sum_{p^2\leq x} \chi(p)^2\log p + O(x^{\frac{1}{3}})$$
$$= (\rho(q)-1) \sum_{\substack{p\leq x\\ p\equiv\square \bmod q}} \log p - \sum_{\substack{p\leq x\\ p\not\equiv\square \bmod q}} \log p + (\rho(q)-1)\sqrt{x} + o(\sqrt{x}),$$

by (11) and the Prime Number Theorem. Combining this with a standard summation by parts we get that

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

$\square$

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

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

Proof. This follows from Rubinstein and Sarnak's anaysis [RS], and from [ANS]. $\square$

Remark 2.3. As Schlage-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

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

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

$$\sup_{x<c} F_q(x) \leq \liminf_{Y\to\infty}\frac{1}{Y} meas\{y\leq Y : E_q(e^y)\leq c\}$$
$$\leq \limsup_{Y\to\infty}\frac{1}{Y} meas\{y\leq Y : E_q(e^y)\leq c\} \leq \inf_{x>c} F_q(x),$$

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 μ_q. We will show that X_q can be very biased, in the sense that $\mathrm{Prob}[X_q>0]$ 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 2.4. *We have that*

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

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

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

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

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

$$\limsup_{Y\to\infty}\frac{1}{Y}\int_{\substack{0\le y\le Y\\ |E_q(e^y)|>M}} |E_q(e^y)|^2dy \longrightarrow 0. \tag{13}$$

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

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

and so

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

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

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

We then have

$$\frac{1}{Y}\int_2^Y E_q(e^y)^2dy = \frac{1}{Y}\int_{2\le y\le Y} H_M(E_q(e^y))dy - \frac{1}{Y}\int_{\substack{2\le y\le Y\\ M<|E_q(e^y)|\le M+1}} H_M(E_q(e^y))dy + \frac{1}{Y}\int_{\substack{2\le y\le Y\\ |E_q(e^y)|>M}} E_q(e^y)^2dy,$$

therefore by (13) and by Lemma 2.2 we get that

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

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

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

(see Theorem 1.2 of [RS]) we get by taking $M\to\infty$ that

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

The same reasoning applies to the lim inf, and thus the proof is finished.

$\square$

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

Lemma 2.5. *Assume GRH. Then,*

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

where the last sum runs over the imaginary parts of the non-trivial zeros of

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

m_γ denotes the multiplicity of the zero $\frac{1}{2}+i\gamma$, the star meaning 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

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

by absolute convergence. Taking $Y \to \infty$ and applying Lemma 2.4 gives that

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

The calculation of the variance follows from Lemma 2.1 and from Parseval's identity for B^2 almost-periodic functions [B1]. (An alternative way to compute the variance is to argue as in [SP].) $\square$

Remark 2.6. It is a general fact that Besicovitch almost-periodic functions always have a mean value [B2]. Moreover, Parseval's identity [B1, B2] shows that Besicovitch B^2 (and thus also Stepanov S^2, Weil W^2 and Bohr) almost periodic functions $f(y)$ have a second moment given by

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

where the A_n are the Fourier coefficients of f.

Lemma 2.7. *Assume GRH. If*

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

is large enough, then

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

Proof. It is clear from Lemma 2.5 and the Riemann-von Mangoldt formula that $\mathrm{Var}[X_q] \gg \log q'$, and therefore our assumption that $B(q)$ is large enough implies that $\mathbb{E}[X_q]$ is also large enough, say at least 4. Define

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

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

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

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:

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

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

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

□

Proof of Theorem 1.2. By Lemma 2.5, 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 the Riemann von-Mangoldt formula. Lemma 2.5 also implies that $\mathbb{E}[X_q] \gg \rho(q)$, and hence Lemma 2.7 implies that

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

The last inequality to show, that is $\overline{\delta}(q; NR, R) < 1$, follows from an analysis using the functions $f(x)$ and $H(x)$ of Lemma 2.7, combined with a lower bound on $\mu_E(-\infty, -1]$ similar to that in Theorem 1.2 of [RS], which holds in greater generality [ANS].

□

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 hypothesis 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 equals*

$$Prob[X_q] > 0,$$

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

$$X_q \sim \rho(q) - 1 + \sum_{\substack{\chi \bmod q \\ \chi^2=\chi_0 \\ \chi\neq\chi_0}} \sum_{\gamma_\chi>0} \frac{2\Re(Z_{\gamma_\chi})}{\sqrt{\frac14+\gamma_\chi^2}}, \tag{14}$$

where the Z_{γ_χ} 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 $\mathrm{Prob}[X_q>0]$ (their analysis shows that the distribution function of X_q is continuous). Moreover, an argument similar to the proof of Proposition 2.3 of [FiMa] shows that (14) holds. □

One can show that the random variables appearing in (14) have variance $\mathrm{Var}[\Re(Z_{\gamma_\chi})] = \frac12$, and have mean $\mathbb{E}[Z_{\gamma_\chi}]=0$. Using this and the fact that they are mutually independent, we recover Lemma 2.5:

$$\mathbb{E}[X_q] = \rho(q)-1, \qquad \mathrm{Var}[X_q] = \sum_{\substack{\chi \bmod q \\ \chi^2=\chi_0 \\ \chi\neq\chi_0}} \sum_{\gamma_\chi} \frac{1}{\frac14+\gamma_\chi^2}, \tag{15}$$

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

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

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

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

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

Proof. By Remark 1.4, we have that

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

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

Lemma 3.5 of [FiMa] gives that

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

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

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

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

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

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

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

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

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

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

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

□

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

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

It is $B(q)$ which dictates the behaviour 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 the estimate

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

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 $e_c := \min\{e \geq 1 : 2^{-e}c < \frac{2}{\log 4}\}$ and $c_1 := 2^{-e_c}c < \frac{2}{\log 4}$. Define for $\ell = 1, 2, ...$ the intervals

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

Since $c_1 < \frac{2}{\log 4}$ 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_ℓ to be any prime in the interval I_ℓ, 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. *Let X_q be the random variable defined in* (14)*, and define*

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

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

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

Moreover, in the same range we have

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

Proof. The proof is very similar to that of Theorem 3.22 of [FiMa]. Using the additivity of the cumulant-generating function of X_q, one can show that

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

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

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

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

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

One can show using the Riemann-von Mangoldt formula that

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

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

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

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 $|\xi| \leq \frac{12}{5}$:

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

□

Lemma 3.5 (Berry-Essen inequality)**.** *Denote by F_q the distribution function of*

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

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 Martin-Feuerverger formula [FeMa]. 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 $\mathrm{Var}[X_q] \geq 1$ (by Lemma 3.2).

The Berry-Esseen inequality in the form given by Esseen (Theorem 2a of [E]) gives that for any $T > 0$,

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

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

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

We now bound the remaining part of the integral using an argument analogous to Proposition 2.14 of [FiMa]. Fix $0 \leq \lambda \leq \frac{5}{6}$. By the properties of the Bessel function $J_0(x)$, we have that if $|\xi| > \lambda$, 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 $|\xi| > \frac{5}{12}\mathrm{Var}[X_q]^{-\frac{1}{4}}$ we have $|\hat{X}_q(\xi)| \leq |\hat{X}_q(\frac{5}{12}\mathrm{Var}[X_q]^{-\frac{1}{4}})|$ (since $\mathrm{Var}[X_q] \geq 1$), and so

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

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

□

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 $\delta(q_n, NR, R) \to \eta$. 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 of Rubinstein and Sarnak [RS]. Therefore we can assume that $\frac{1}{2} < \eta < 1$.

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

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

Let moreover $\{q_n\}$ be the sequence of squarefree odd integers coming from Lemma 3.3 for which

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

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

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

Define

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

We will use the central limit theorem of Lemma 3.4, as well as the Berry-Essen 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

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

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 $\rho(q_n) \to \infty$, hence this last quantity tends to zero as $n \to \infty$, 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 Theorem 2 of [MoOd].

Theorem 4.1 (Montgomery and Odlyzko)**.** *For $n = 1, 2, ...$ 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

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

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

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

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 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

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

□

Lemma 4.3. *Let $\mathcal{F}(q)$ be a subset of the invertible residues* mod$q$ *such that $\chi \in \mathcal{F}(q) \Rightarrow \overline{\chi} \in \mathcal{F}(q)$. 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_{γ_χ} are i.i.d. uniformly distributed on the unit circle. Then, we have for q large enough that

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

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. It is a direct application of Theorem 4.1. Taking the sequence $\{r_i\}$ to be the $\frac{2}{\sqrt{\frac{1}{4}+\gamma_\chi^2}}$ ordered by size, we have for $0 < \alpha \leq 4$ that

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

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

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

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

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

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 χ and large enough such that $N(2T_1,\chi) - N(T_1,\chi) \gg \log q^*$ (this is possible by the von-Mangoldt formula), then we have

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

□

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

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

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.

□

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, $\overrightarrow{a} = (a_1, ..., a_k)$ will be a vector of invertible reduced residues $\bmod q$ and $\overrightarrow{\alpha} = (\alpha_1, ..., \alpha_k)$ will be a non-zero vector of real numbers such that $\sum_{i=1}^k \alpha_i = 0$. Recall that

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

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*

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

has the same distribution as that of the random variable

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

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

Remark 5.2. If we take $a_1, ..., a_{\phi(q)(1-\rho(q)^{-1})}$ to be the set of all quadratic non-residues $\bmod q$ with $\alpha_1 = ... = \alpha_{\phi(q)(1-\rho(q)^{-1})} = \frac{1}{\phi(q)}$, and we take $a_{\phi(q)(1-\rho(q)^{-1})+1}, ..., a_{\phi(q)}$ to be the set of all quadratic residues $\bmod q$ with $\alpha_{\phi(q)(1-\rho(q)^{-1})+1} = ... = \alpha_{\phi(q)} = \frac{1-\rho(q)}{\phi(q)}$, then we recover the 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

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

(the main terms cancel since $\sum_{i=1}^k \alpha_i = 0$). By the work of Rubinstein and Sarnak [RS], $F(y; q, \overrightarrow{a}, \overrightarrow{\alpha})$ has the same distribution as $X_{q;\overrightarrow{a},\overrightarrow{\alpha}} - \mathbb{E}[X_{q;\overrightarrow{a},\overrightarrow{\alpha}}]$, 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

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

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 mod$q$, we have

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

$\square$

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

$$\phi(q)\|\overrightarrow{\alpha}\|_2^2 \log\left(\frac{3\phi(q)}{k}\right) \ll V(q;\overrightarrow{a},\overrightarrow{\alpha}) \ll \phi(q)\|\overrightarrow{\alpha}\|_2^2 \log q, \tag{23}$$

where

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

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 bounded number of distinct prime factors and consider the race between residues and non-residues with the weights of Remark 5.2, we obtain by Lemma 3.2 that $V(q;\overrightarrow{a},\overrightarrow{\alpha}) = O(1)$, and this is of the same order of magnitude as the lower bound in (23).

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

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

(LI implies that there are no real zeros.) By (16), there exists q_0 such that whenever $q^* \geq q_0$ we have

$$\sum_{\gamma_\chi} \frac{1}{\frac{1}{4}+\gamma_\chi^2} \asymp \log q^*, \tag{24}$$

and the same estimate clearly holds for $q^* < q_0$, since the left hand side of (24) is positive. We conclude that

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

Now, $\alpha_1\chi(a_1)+...+\alpha_k\chi(a_k)=0$, so

$$\sum_{\chi\neq\chi_0}|\alpha_1\chi(a_1)+...+\alpha_k\chi(a_k)|^2=\sum_{\chi \bmod q}|\alpha_1\chi(a_1)+...+\alpha_k\chi(a_k)|^2$$
$$=\sum_{1\leq i,j\leq k}\alpha_i\alpha_j\sum_{\chi \bmod q}\chi(a_ia_j^{-1})$$
$$=\phi(q)\sum_{i=1}^{k}\alpha_i^2.$$

Using this and (25), the upper bound follows from the fact that $\log q^*\leq\log q$. This also gives the lower bound $V(q;\overrightarrow{a},\overrightarrow{\alpha})\geq\log 3\phi(q)\|\overrightarrow{\alpha}\|_2^2$, 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:

$$V(q;\overrightarrow{a},\overrightarrow{\alpha})\geq\log L\sum_{\substack{\chi \bmod q:\\ q^*>L}}|\alpha_1\chi(a_1)+...+\alpha_k\chi(a_k)|^2$$
$$=\log L\sum_{1\leq i,j\leq k}\alpha_i\alpha_j\sum_{\substack{\chi \bmod q:\\ q^*>L}}\chi(a_ia_j^{-1})$$
$$=\log L\left[\sum_{i=1}^{k}\alpha_i^2\sum_{\substack{\chi \bmod q\\ q^*>L}}1+\sum_{1\leq i\neq j\leq k}\alpha_i\alpha_j\sum_{\substack{\chi \bmod q:\\ q^*>L}}\chi(a_ia_j^{-1})\right],$$

which by Lemma 5.3 and the orthogonality relations is

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

by the Cauchy-Schwartz 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 Lemma 5.3 and the Cauchy-Schwartz inequality, their contribution is

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

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

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

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

∎

Proof of Theorem 1.10. It is a particular case of Theorem 1.14. ∎

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. *Let*

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

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

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

in the range $|\xi|\leq\frac{3}{5\|\overrightarrow{\alpha}\|_1}$, where $\|\overrightarrow{\alpha}\|_1 := \sum_{i=1}^k|\alpha_i|$.

Proof. As in Lemma 3.4, we compute

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

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

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

Using (24) and the analogous estimate $\sum_{\gamma_\chi}(\frac{1}{4}+\gamma_\chi^2)^{-2}\asymp\log q^*$, we get

$$\log\hat{Y}_{q;\overrightarrow{a},\overrightarrow{\alpha}}(\xi) = -\frac{\xi^2}{2} + O\left(\xi^4\frac{\sum_{\chi\neq\chi_0}|\alpha_1\chi(a_1)+...+\alpha_k\chi(a_k)|^4\log q^*}{\left(\sum_{\chi\neq\chi_0}|\alpha_1\chi(a_1)+...+\alpha_k\chi(a_k)|^2\log q^*\right)^2}\right). \tag{26}$$

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 $\Theta_\chi := |\alpha_1\chi(a_1)+...+\alpha_k\chi(a_k)|^2$, we have

$$\sum_{\chi\neq\chi_0}\Theta_\chi\log q^* \geq \sum_{\substack{\chi\neq\chi_0\\ q^*>L}}\Theta_\chi\log q^* \geq \sqrt{\log L}\sum_{\substack{\chi\neq\chi_0\\ q^*>L}}\Theta_\chi\sqrt{\log q^*}$$
$$\geq \sqrt{\log L}\left(\sum_{\chi\neq\chi_0}\Theta_\chi\sqrt{\log q^*} - kL^2\sqrt{\log L}\|\overrightarrow{\alpha}\|_2^2\right) \tag{27}$$

by Lemma 5.3 and the Cauchy-Schwartz inequality. Now, by our choice of L and by the fact that $\phi(q)/k \geq 576$ we have

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

hence (27) 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 (26) 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-Schwartz inequality:

$$\frac{\sum_{\chi\neq\chi_0}|\alpha_1\chi(a_1)+...+\alpha_k\chi(a_k)|^4\log q^*}{\left(\sum_{\chi\neq\chi_0}|\alpha_1\chi(a_1)+...+\alpha_k\chi(a_k)|^2\log q^*\right)^2} \ll \frac{\log q}{\log(3\phi(q)/k)^2}\frac{\sum_{\chi\neq\chi_0}|\alpha_1\chi(a_1)+...+\alpha_k\chi(a_k)|^4}{(\phi(q)\|\overrightarrow{\alpha}\|_2^2)^2}$$
$$= \frac{\log q}{\log(3\phi(q)/k)^2}\frac{\sum_{\substack{i,j,i',j'\\ a_ia_{i'}\equiv a_ja_{j'} \bmod q}}\alpha_i\alpha_{i'}\alpha_j\alpha_{j'}}{\phi(q)\|\overrightarrow{\alpha}\|_2^4}$$
$$\leq \frac{\log q}{\log(3\phi(q)/k)^2}\frac{\left(\sqrt{\sum_{i=1}^k\alpha_i}\sqrt{\sum_{j=1}^k 1}\right)^4}{\phi(q)\|\overrightarrow{\alpha}\|_2^4},$$

which gives the claimed bound. □

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 $K_2=K$. Define the vector $\overrightarrow{\beta} := \frac{e^{-K}}{\|\overrightarrow{\alpha}\|_1}\overrightarrow{\alpha}$, so that $\|\overrightarrow{\beta}\|_1 = e^{-K}$, which will allow us to apply Lemma 5.7. Clearly,

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

since multiplying $\overrightarrow{\alpha}$ by a positive constant does not affect the inequality $\alpha_1\pi(n;q,a_1)+...+\alpha_k\pi(n;q,a_k)>0$.

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

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

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

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

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

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

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

$$\begin{aligned}\mathrm{Prob}[Y_{q;\vec{a},\vec{\beta}} > -B(q;\vec{a},\vec{\beta})] &- \mathrm{Prob}[W > -B(q;\vec{a},\vec{\beta})] \\ &\ll \int_{-\frac{3}{5}e^K}^{\frac{3}{5}e^K} \frac{\hat{Y}_{q;\vec{a},\vec{\beta}}(\xi) - e^{-\frac{\xi^2}{2}}}{\xi} d\xi + \frac{5}{3}e^{-K} \\ &\ll \int_{-\frac{3}{5}e^K}^{\frac{3}{5}e^K} \frac{\xi^3 e^{-\frac{\xi^2}{2}}}{e^{4K}} d\xi + e^{-K} \\ &\ll e^{-K}.\end{aligned} \tag{28}$$

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

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

for some absolute constant c_1. Therefore, applying (28) gives

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

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

To end this section we give an exact expression 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. *We have that*

$$V(q;\vec{a},\vec{\alpha})=\phi(q)\|\vec{\alpha}\|_2\left(\log q-\sum_{p\mid q}\frac{\log p}{p-1}\right)-\phi(q)\sum_{i\neq j}\alpha_i\alpha_j\frac{\Lambda\left(\frac{q}{(q,a_ia_j^{-1}-1)}\right)}{\phi\left(\frac{q}{(q,a_ia_j^{-1}-1)}\right)}. \tag{29}$$

Proof. Using Proposition 3.3 of [FiMa], we get

$$\begin{aligned}V(q;\vec{a},\vec{\alpha})&=\sum_{\chi \bmod q}|\alpha_1\chi(a_1)+...+\alpha_k\chi(a_k)|^2\log q^*\\&=\sum_{1\leq i,j\leq k}\alpha_i\alpha_j\sum_{\chi \bmod q}\chi(a_ia_j^{-1})\log q^*\\&=\phi(q)\left(\log q-\sum_{p\mid q}\frac{\log p}{p-1}\right)\sum_{i=1}^k\alpha_i^2-\phi(q)\sum_{i\neq j}\alpha_i\alpha_j\frac{\Lambda\left(\frac{q}{(q,a_ia_j^{-1}-1)}\right)}{\phi\left(\frac{q}{(q,a_ia_j^{-1}-1)}\right)}.\end{aligned}$$

□

It might seem like the second term of (29) 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

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

References

[ANS] Amir Akbary, Nathan Ng and Majid Shahabi, *The distribution of the classical error terms of prime number theory*, Preprint.

[BFHR] Carter Bays, Kevin Ford, Richard H. Hudson, Michael Rubinstein, *Zeros of Dirichlet L-functions near the real axis and Chebyshev's bias.* J. Number Theory **87** (2001), no. 1, 54-76.

[BH] Carter Bays, Richard H. Hudson, *A new bound for the smallest x with $\pi(x)>li(x)$.* Math. Comp. **69** (2000), no. 231, 1285–1296 (electronic).

[B1] A. S. Besicovitch, *On Generalized Almost Periodic Functions.* Proc. London Math. Soc. S2-25 no. 1, 495.

[B2] A. S. Besicovitch, *Almost periodic functions.* Dover Publications, Inc., New York, 1955. xiii+180 pp.

[E] Carl-Gustav Esseen, *Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law.* Acta Math. **77** (1945), 1-125.

[FeMa] Andrey Feuerverger and Greg Martin, *Biases in the Shanks-Rényi prime number race.* Experiment. Math. **9** (2000), no. 4, 535-570.

[FiMa] Daniel Fiorilli and Greg Martin, *Inequities in the Shanks-Rényi Prime Number Race: An asymptotic formula for the densities.* To appear in J. Reine Angew. Math.

[Fo1] Kevin Ford, *The distribution of integers with a divisor in a given interval.* Ann. of Math. (2) **168** (2008), no. 2, 367-433.

[Fo2] Kevin Ford, *Integers with a divisor in $(y,2y]$.* Anatomy of integers, 65?80, CRM Proc. Lecture Notes, **46**, Amer. Math. Soc., Providence, RI, 2008.

[HL] Godfrey H. Hardy and John E. Littlewood, *Contributions to the theory of the Riemann Zeta-function and the theory of the distribution of primes.* Acta Math. **41** (1917), 119–196 .

[K] Jerzy Kaczorowski, *On the distribution of primes (mod 4).* Analysis **15** (1995), no. 2, 159-171.

[GM] Andrew Granville, Greg Martin, *Prime number races.* Amer. Math. Monthly **113** (2006), no. 1, 1–33.

[Lam] Youness Lamzouri, *Large deviations of the limiting distribution in the Shanks-Rényi prime number race.* To appear in Mathematical Proceedings of the Cambridge Philosophical Society.

[Lan1] Edmund Landau, *Über einige ältere Vermutungen und Behauptungen in der Primzahltheorie. I.* Math. Z., **1** (1918), 1–24.

[Lan2] Edmund Landau, *Über einige ältere Vermutungen und Behauptungen in der Primzahltheorie. II.* Math. Z. **1**, (1918), 213–219.

[Mi] Stephen D. Miller, *The highest lowest zero and other applications of positivity.* Duke Math. J. **112** (2002), no. 1, 83-116.

[Mn] William R. Monach, *Numerical investigation of several problems in number theory.* Ph.D. Thesis, University of Michigan, 1980, 180 pp.

[Mo] Hugh L. Montgomery, *The zeta function and prime numbers.* Proceedings of the Queen's Number Theory Conference, 1979, pp. 1-31, Queen's Papers in Pure and Appl. Math., 54, Queen's Univ., Kingston, Ont., 1980.

[MoOd] Hugh L. Montgomery, Andrew M. Odlyzko, *Large deviations of sums of independent random variables.* Acta Arith. **49** (1988), no. 4, 427–434.

[My] Colin Mysercough, *Application of an improved remainder term in the calculation of residue class distributions.* Preprint.

[N] Nathan Ng, *Limiting distributions and zeros of Artin L-functions.* Ph.D. thesis, University of British Columbia, 2000.

[O] Andrew Odlyzko, *Bounds for discriminants and related estimates for class numbers, regulators and zeroes of zeta functions: a survey of recent results.* Séminaire de Théorie des Nombres, Bordeaux **2** (1990), 119–141.

[RS] Michael Rubinstein and Peter Sarnak, *Chebyshev's bias.* Experiment. Math. **3** (1994), no. 3, 173-197.

[SP] J.-C. Schlage-Puchta, *On large oscillations of the remainder of the prime number theorems.* Acta Math. Hungar. **87** (2000), no. 3, 213–227.

School of Mathematics, Institute for Advanced Study, 1 Einstein Drive, Princeton NJ 08540 USA

Current address: Department of Mathematics, University of Michigan, 530 Church Street, Ann Arbor MI 48109 USA

E-mail address: `fiorilli@umich.edu`