Math. Res. Lett. **19** (2012), *no.* 03, 649–666

THE SHANKS–RÉNYI PRIME NUMBER RACE WITH MANY CONTESTANTS

Youness Lamzouri

Abstract. Under certain plausible assumptions, M. Rubinstein and P. Sarnak solved the Shanks–Rényi race problem by showing that the set of real numbers $x \geq 2$ such that $\pi(x;q,a_1) > \pi(x;q,a_2) > \cdots > \pi(x;q,a_r)$ has a positive logarithmic density $\delta_{q;a_1,\dots,a_r}$. Furthermore, they established that if r is fixed, $\delta_{q;a_1,\dots,a_r} \to 1/r!$ as $q \to \infty$. In this paper, we investigate the size of these densities when the number of contestants r tends to infinity with q. In particular, we deduce a strong form of a recent conjecture of Feuerverger and Martin which states that $\delta_{q;a_1,\dots,a_r} = o(1)$ in this case. Among our results, we prove that $\delta_{q;a_1,\dots,a_r} \sim 1/r!$ in the region $r = o(\sqrt{\log q})$ as $q \to \infty$. We also bound the order of magnitude of these densities beyond this range of r. For example, we show that when $\log q \leq r \leq \phi(q)$, $\delta_{q;a_1,\dots,a_r} \ll_\epsilon q^{-1+\epsilon}$.

1. Introduction

A classical problem in analytic number theory is the so-called "Shanks–Rényi prime number race" which concerns the distribution of prime numbers in arithmetic progressions. As colorfully described by Knapowski and Turán in [11], let $q \geq 3$ and $2 \leq r \leq \phi(q)$ be positive integers, and denote by $\mathcal{A}_r(q)$ the set of ordered r-tuples of distinct residue classes $(a_1, a_2, \dots, a_r)$ modulo q which are coprime to q. For $(a_1, a_2, \dots, a_r) \in \mathcal{A}_r(q)$ consider a game with r players called "1" through "r", where at time x, the player "j" has a score of $\pi(x;q,a_j)$ (where $\pi(x;q,a)$ denotes the number of primes $p \leq x$ with $p \equiv a \bmod q$). As $x \to \infty$, will all $r!$ orderings of the players occur for infinitely many integers x?

It is generally believed that the answer to this question is yes for all q and all $(a_1, a_2, \dots, a_r) \in \mathcal{A}_r(q)$. An old result of Littlewood [14] shows that this is indeed true in the special cases $(q, a_1, a_2) = (4, 1, 3)$ and $(q, a_1, a_2) = (3, 1, 2)$. Since then, this problem has been extensively studied by many authors, including Knapowski and Turán [11], Bays and Hudson [1, 2], Kaczorowski [8–10], Feuerverger and Martin [4], Martin [15], Ford and Konyagin [6, 7], Fiorilli and Martin [5], and Lamzouri [12] and [13].

A major breakthrough was made in 1994 by Rubinstein and Sarnak who completely solved this problem in [16], conditionally on the two following assumptions:

- The Generalized Riemann Hypothesis (GRH): all nontrivial zeros of Dirichlet L-functions have real part equal 1/2.

Received by the editors August 15, 2011.

2010 *Mathematics Subject Classification.* Primary 11N13; Secondary 11N69, 11M26.

Key words and phrases. The Shanks–Rényi race problem, primes in arithmetic progressions, zeros of Dirichlet L-functions.

- The Linear Independence Hypothesis (LI) (also known as the Grand Simplicity Hypothesis): the nonnegative imaginary parts of the nontrivial zeros of Dirichlet L-functions attached to primitive characters are linearly independent over $\mathbb{Q}$.

Rubinstein and Sarnak [16] proved, under these two hypotheses, the stronger result that for any $(a_1, \dots, a_r) \in \mathcal{A}_r(q)$ the set of real numbers $x \geq 2$ such that

$$\pi(x; q, a_1) > \pi(x; q, a_2) > \cdots > \pi(x; q, a_r),$$

has a positive logarithmic density, which shall be denoted throughout this paper by $\delta_{q;a_1,\dots,a_r}$. (Recall that the logarithmic density of a subset S of $\mathbb{R}$ is defined as

$$\delta_S := \lim_{x\to\infty} \frac{1}{\log x} \int_{t\in S\cap[2,x]} \frac{dt}{t},$$

provided that this limit exists.) To establish this result, they constructed an absolutely continuous measure $\mu_{q;a_1,\dots,a_r}$ for which

$$\delta_{q;a_1,\dots,a_r} = \int_{x_1>x_2>\cdots>x_r} d\mu_{q;a_1,\dots,a_r}(x_1, \dots, x_r). \tag{1.1}$$

Among the results they derived on these densities, Rubinstein and Sarnak [16] showed that in an r-way race with r fixed, all biases disappear when $q \to \infty$. More specifically they proved

$$\lim_{q\to\infty} \max_{(a_1,\dots,a_r)\in\mathcal{A}_r(q)} |r!\delta_{q;a_1,\dots,a_r} - 1| = 0. \tag{1.2}$$

Recently, Fiorilli and Martin [5] established an asymptotic formula for the density in a two-way race, which allows them to determine the exact rate at which $\delta_{q;a_1,a_2}$ converges to $1/2$ as q grows. Shortly after, the author [12] obtained an asymptotic formula for $\delta_{q;a_1,\dots,a_r}$ for any fixed $r \geq 3$ as $q \to \infty$, in which the rate of convergence to $1/r!$ is surprisingly different from the case $r = 2$.

However, as far as the author of the present paper knows, no results have been obtained on the size of the densities $\delta_{q;a_1,\dots,a_r}$ if $r \to \infty$ as $q \to \infty$. In [4], Feuerverger and Martin conjectured that in this case we should have $\delta_{q;a_1,\dots,a_r} = o(1)$. They also asked whether one can prove a uniform version of the result of Rubinstein and Sarnak (1.2), namely that this statement holds in a certain range $r \leq r_0(q)$ for some $r_0(q) \to \infty$ as $q \to \infty$.

Conjecture 1.1 (Feuerverger–Martin)**.** *We have*

$$\lim_{q\to\infty} \max_{(a_1,\dots,a_r)\in\mathcal{A}_r(q)} \delta_{q;a_1,\dots,a_r} = 0,$$

for any arbitrary function $r = r(q)$ tending to infinity with q.

In the present paper, we investigate the order of magnitude of $\delta_{q;a_1,\dots,a_r}$ when the number of contestants $r \to \infty$ as $q \to \infty$. In particular, answering the question of Feuerverger and Martin, we establish a uniform version of (1.2), and obtain a strong quantitative form of Conjecture 1.1.

Theorem 1.1. *Assume GRH and LI. Let q be a large positive integer. Then, for any integer r such that $2 \le r \le \sqrt{\log q}$ we have*

$$\delta_{q;a_1,\dots,a_r} = \frac{1}{r!}\left(1 + O\left(\frac{r^2}{\log q}\right)\right),$$

uniformly for all r-tuples $(a_1, \dots, a_r) \in \mathcal{A}_r(q)$.

As a consequence, Theorem 1.1 implies that (1.2) holds true in the range $r = o(\sqrt{\log q})$ as $q \to \infty$. Indeed in this region of r, all biases disappear when $q \to \infty$, namely

$$\delta_{q;a_1,\dots,a_r} \sim \frac{1}{r!}, \tag{1.3}$$

uniformly for all r-tuples $(a_1, \dots, a_r) \in \mathcal{A}_r(q)$. Moreover, one can also deduce that if $c_0 > 0$ is a suitably small constant and $r \le c_0\sqrt{\log q}$, then uniformly for all r-tuples $(a_1, \dots, a_r) \in \mathcal{A}_r(q)$ we have

$$\delta_{q;a_1,\dots,a_r} \asymp \frac{1}{r!}. \tag{1.4}$$

Note that $1/r! = \exp(-r\log r + r + O(\log r))$ by Stirling's formula. Our next result shows that the densities $\delta_{q;a_1,\dots,a_r}$ have roughly the same asymptotic decay in the range $\sqrt{\log q} \ll r \le (1-\epsilon)\log q/\log\log q$, for any $\epsilon > 0$.

Theorem 1.2. *Assume GRH and LI. For any $\epsilon > 0$, if q is large and $\sqrt{\log q} \ll r \le (1-\epsilon)\log q/\log\log q$ is an integer, then*

$$\delta_{q;a_1,\dots,a_r} = \exp\left(-r\log r + r + O\left(\log r + \frac{r^2}{\log q}\right)\right),$$

uniformly for all r-tuples $(a_1, \dots, a_r) \in \mathcal{A}_r(q)$.

It would be interesting to determine the order of magnitude of the densities $\delta_{q;a_1,\dots,a_r}$ beyond the region $r \le (1-\epsilon)\log q/\log\log q$. Unfortunately, this range seems to be the limit of what can be achieved using our method. Nevertheless, we can use Theorem 1.2 to obtain an upper bound for $\delta_{q;a_1,\dots,a_r}$ beyond this range of r.

Theorem 1.3. *Assume GRH and LI. For any $\epsilon > 0$, if q is large and r is a positive integer such that $(1-\epsilon/2)\log q/\log\log q \le r \le \phi(q)$, then*

$$\max_{(a_1,\dots,a_r)\in\mathcal{A}_r(q)} \delta_{q;a_1,\dots,a_r} \ll_\epsilon \frac{1}{q^{1-\epsilon}}.$$

The paper is organized as follows. In Section 2, following the work of Rubinstein and Sarnak, we shall construct the measure $\mu_{q;a_1,\dots,a_r}$ as a probability distribution corresponding to a certain random vector and study its covariance matrix and large deviations. In Section 3, we investigate the Fourier transform of $\mu_{q;a_1,\dots,a_r}$ and show that in a certain range $\hat{\mu}_{q;a_1,\dots,a_r}$ can be approximated by the Fourier transform of a multivariate normal distribution having the same covariance matrix. In Section 4 we study properties of multivariate normal distributions and prove Theorems 1.1, 1.2 and 1.3.

2. The measure $\mu_{q;a_1,\dots,a_r}$

We begin by developing the necessary notation to construct the measure $\mu_{q;a_1,\dots,a_r}$, following the work of Rubinstein and Sarnak [16]. For $(a_1, a_2, \dots, a_r) \in \mathcal{A}_r(q)$ we introduce the vector-valued function

$$E_{q;a_1,\dots,a_r}(x) := (E(x;q,a_1), \dots, E(x;q,a_r)),$$

where

$$E(x;q,a) := \frac{\log x}{\sqrt{x}} \left(\phi(q)\pi(x;q,a) - \pi(x)\right).$$

The normalization is such that, if we assume GRH, $E_{q;a_1,\dots,a_r}(x)$ varies roughly boundedly as x varies. Moreover, for a nontrivial character χ modulo q, we denote by $\{\gamma_\chi\}$ the sequence of imaginary parts of the nontrivial zeros of $L(s,\chi)$. Let χ_0 denote the principal character modulo q and define $S = \cup_{\chi \neq \chi_0 \bmod q}\{\gamma_\chi\}$. Furthermore, let $\{U(\gamma_\chi)\}_{\gamma_\chi \in S}$ be a sequence of independent random variables uniformly distributed on the unit circle.

Rubinstein and Sarnak established, under GRH and LI, that the vector-valued function $E_{q;a_1,\dots,a_r}$ has a limiting distribution $\mu_{q;a_1,\dots,a_r}$, where $\mu_{q;a_1,\dots,a_r}$ is the probability measure corresponding to the random vector

$$X_{q;a_1,\dots,a_r} = (X(q,a_1), \dots, X(q,a_r)),$$

where

$$X(q,a) = -C_q(a) + \sum_{\substack{\chi \neq \chi_0 \\ \chi \bmod q}} \sum_{\gamma_\chi > 0} \frac{2\mathrm{Re}(\chi(a)U(\gamma_\chi))}{\sqrt{\frac{1}{4} + \gamma_\chi^2}},$$

and

$$C_q(a) := -1 + \sum_{\substack{b^2 \equiv a \bmod q \\ 1 \leq b \leq q}} 1.$$

Note that for $(a,q) = 1$ the function $C_q(a)$ takes only two values: $C_q(a) = -1$ if a is a non-square modulo q, and $C_q(a) = C_q(1)$ if a is a square modulo q. Furthermore, an elementary argument shows that $C_q(a) < d(q) \ll_\epsilon q^\epsilon$ for any $\epsilon > 0$, where $d(q) = \sum_{m|q} 1$ is the usual divisor function.

To investigate the distribution of the random vector $X_{q;a_1,\dots,a_r}$ we shall first compute its covariance matrix $\mathrm{Cov}_{q;a_1,\dots,a_r}$ (the covariance matrix generalizes the notion of variance to multiple dimensions). Recall that the j,k entry of the covariance matrix corresponds to the covariance between the j-th and k-th entry of the random vector.

Lemma 2.1. *The entries of* $\mathrm{Cov}_{q;a_1,\dots,a_r}$ *are*

$$\mathrm{Cov}_{q;a_1,\dots,a_r}(j,k) = \begin{cases} \mathrm{Var}(q) & \text{if } j = k \\ B_q(a_j, a_k) & \text{if } j \neq k, \end{cases}$$

where

$$\mathrm{Var}(q) := 2 \sum_{\substack{\chi \neq \chi_0 \\ \chi \bmod q}} \sum_{\gamma_\chi > 0} \frac{1}{\frac{1}{4} + \gamma_\chi^2}, \text{ and } B_q(a,b) := \sum_{\substack{\chi \neq \chi_0 \\ \chi \bmod q}} \sum_{\gamma_\chi > 0} \frac{\chi\left(\frac{b}{a}\right) + \chi\left(\frac{a}{b}\right)}{\frac{1}{4} + \gamma_\chi^2}.$$

Proof. First, note that $\mathbb{E}(X(q,a)) = -C_q(a)$ since $\mathbb{E}(U(\gamma_\chi)) = 0$ for all γ_χ. Therefore, $\mathrm{Cov}_{q;a_1,\ldots,a_r}(j,k)$ equals

$$\mathbb{E}\Big(\big(X(q,a_j)+C_q(a_j)\big)\big(X(q,a_k)+C_q(a_k)\big)\Big)$$

$$= \mathbb{E}\left(\sum_{\substack{\chi\neq\chi_0\\ \chi \bmod q}}\sum_{\gamma_\chi>0}\sum_{\substack{\psi\neq\chi_0\\ \psi \bmod q}}\sum_{\widetilde{\gamma}_\psi>0}\frac{\begin{array}{c}\big(\chi(a_j)U(\gamma_\chi)+\overline{\chi(a_j)U(\gamma_\chi)}\big)\\ \big(\psi(a_k)U(\widetilde{\gamma}_\psi)+\overline{\psi(a_k)U(\widetilde{\gamma}_\psi)}\big)\end{array}}{\sqrt{\frac14+\gamma_\chi^2}\sqrt{\frac14+\widetilde{\gamma}_\psi^2}}\right).$$

Since $\mathbb{E}(U(\gamma_\chi)U(\widetilde{\gamma}_\psi)) = 0$ for all $\gamma_\chi, \widetilde{\gamma}_\psi$ and

$$\mathbb{E}\left(U(\gamma_\chi)\overline{U(\widetilde{\gamma}_\psi)}\right) = \begin{cases}1 & \text{if } \chi=\psi \text{ and } \gamma_\chi=\widetilde{\gamma}_\psi,\\ 0 & \text{otherwise,}\end{cases}$$

we deduce that

$$\mathrm{Cov}_{q;a_1,\ldots,a_r}(j,k) = \sum_{\substack{\chi\neq\chi_0\\ \chi \bmod q}}\sum_{\gamma_\chi>0}\frac{\chi(a_j/a_k)+\chi(a_k/a_j)}{\frac14+\gamma_\chi^2},$$

which implies the result. □

Our next lemma gives the asymptotic behavior of $\mathrm{Var}(q)$ along with the maximal order of $B_q(a_j,a_k)$. This was established in [12], and we should also note that it follows implicitly from the results of [5].

Lemma 2.2. *Assume GRH. Then*

$$\mathrm{Var}(q) = \phi(q)\log q + O(\phi(q)\log\log q), \tag{2.1}$$

and

$$\max_{(a,b)\in\mathcal{A}_2(q)} B_q(a,b) \asymp \phi(q). \tag{2.2}$$

Proof. First, the asymptotic formula (2.1) is proved in Lemma 3.1 of [12]. Now, the fact that $B_q(a_j,a_k) \ll \phi(q)$ is proved in Corollary 5.4 of [12], while Proposition 5.1 of [12] implies $B_q(a,-a) \gg \phi(q)$. □

Here and throughout we shall use the notations $\|\mathbf{t}\| = \sqrt{\sum_{j=1}^r t_j^2}$ and $|\mathbf{t}|_\infty = \max_{1\le j\le r}|t_j|$ for the Euclidean norm and the maximum norm of $\mathbf{t}\in\mathbb{R}^r$, respectively. Our next result is an upper bound for the tail of the distribution $\mu_{q;a_1,\ldots,a_r}$. This was established in Proposition 4.1 of [12] in the case where r is fixed.

Lemma 2.3. *Let q be large and $2\le r\le\phi(q)$ be a positive integer. Then for $R\ge\sqrt{\phi(q)\log q}$ we have*

$$\mu_{q;a_1,\ldots,a_r}(|\mathbf{x}|_\infty > R) \le 2r\exp\left(-\frac{R^2}{4\phi(q)\log q}\right),$$

uniformly for all $(a_1,\ldots,a_r)\in\mathcal{A}_r(q)$.

Proof. First, we have

$$\mu_{q;a_1,\dots,a_r}(|\mathbf{x}|_\infty > R) = \mathrm{P}(|X_{q;a_1,\dots,a_r}|_\infty > R)$$
$$\leq \sum_{j=1}^{r} \mathrm{P}(X(q,a_j) > R) + \sum_{j=1}^{r} \mathrm{P}(X(q,a_j) < -R).$$

We shall bound only $\mathrm{P}(X(q,a_j) > R)$, since the corresponding bound for $\mathrm{P}(X(q,a_j) < -R)$ can be obtained similarly. Let $s > 0$ and $(a,q) = 1$. Then we have

$$\mathbb{E}\left(e^{sX(q,a)}\right) = e^{-sC_q(a)} \prod_{\substack{\chi\neq\chi_0 \\ \chi \bmod q}} \prod_{\gamma_\chi>0} \mathbb{E}\left(\frac{2s\mathrm{Re}(\chi(a)U(\gamma_\chi))}{\sqrt{\frac14+\gamma_\chi^2}}\right)$$
$$= e^{-sC_q(a)} \prod_{\substack{\chi\neq\chi_0 \\ \chi \bmod q}} \prod_{\gamma_\chi>0} I_0\left(\frac{2s}{\sqrt{\frac14+\gamma_\chi^2}}\right),$$

where $I_0(t) := \sum_{n=0}^{\infty}(t/2)^{2n}/n!^2$ is the modified Bessel function of order 0. Hence, using the Chernoff bound along with the fact that $I_0(s) \leq \exp(s^2/4)$ for all $s \in \mathbb{R}$ we derive

$$\mathrm{P}(X(q,a) > R) \leq e^{-sR}\mathbb{E}\left(e^{sX(q,a)}\right) \leq \exp\left(-sR - sC_q(a) + \frac{s^2}{2}\mathrm{Var}(q)\right).$$

The lemma follows upon choosing $s = R/(\phi(q)\log q)$, since $C_q(a) = q^{o(1)}$ and $\mathrm{Var}(q) \sim \phi(q)\log q$ by Lemma 2.2. □

3. The Fourier transform $\hat{\mu}_{q;a_1,\dots,a_r}$

Throughout the remaining part of the paper we shall assume both GRH and LI. Moreover, we will use the following normalization for the Fourier transform of an integrable function $f : \mathbb{R}^n \to \mathbb{C}$

$$\hat{f}(t_1,\dots,t_n) = \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}(t_1x_1+\dots+t_nx_n)} f(x_1,\dots,x_n)dx_1\dots dx_n.$$

Then if $\hat{f}$ is integrable on $\mathbb{R}^n$ we have the Fourier inversion formula

$$f(x_1,\dots,x_n) = (2\pi)^{-n}\int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}(t_1x_1+\dots+t_nx_n)} \hat{f}(t_1,\dots,t_n)dt_1\dots dt_n.$$

Similarly we write

$$\hat{\nu}(t_1,\dots,t_n) = \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}(t_1x_1+\dots+t_nx_n)} d\nu(x_1,\dots,x_n)$$

for the Fourier transform of a finite measure ν on $\mathbb{R}^n$.

Rubinstein and Sarnak [16] established the following explicit formula for the Fourier transform of $\mu_{q;a_1,\dots,a_r}$:

(3.1)

$$\hat{\mu}_{q;a_1,\dots,a_r}(t_1,\dots,t_r) = \exp\left(\mathrm{i}\sum_{j=1}^{r} C_q(a_j)t_j\right) \prod_{\substack{\chi\neq\chi_0\\ \chi \bmod q}} \prod_{\gamma_\chi>0} J_0\left(\frac{2\left|\sum_{j=1}^{r}\chi(a_j)t_j\right|}{\sqrt{\frac14+\gamma_\chi^2}}\right),$$

where $J_0(z)=\sum_{m=0}^{\infty}(-1)^m(z/2)^{2m}/m!^2$ is the Bessel function of order 0.

Our first result shows that in the range $\|\mathbf{t}\|\le \mathrm{Var}(q)^{-1/2+o(1)}$, the Fourier transform $\hat{\mu}_{q;a_1,\dots,a_r}(t_1,\dots,t_r)$ is very close to the Fourier transform of a multivariate normal distribution whose covariance matrix equals $\mathrm{Cov}_{q;a_1,\dots,a_r}$.

Proposition 3.1. *Let q be large, $2\le r\le \log q$ be a positive integer, and $(a_1,\dots,a_r)\in\mathcal{A}_r(q)$. Then in the range $\|\mathbf{t}\|\le \mathrm{Var}(q)^{-1/2}\log^2 q$ we have*

$$\hat{\mu}_{q;a_1,\dots,a_r}(t_1,\dots,t_r) = \exp\left(-\frac12\mathbf{t}^T\mathrm{Cov}_{q;a_1,\dots,a_r}\mathbf{t}\right)\left(1+O\left(\frac{d(q)\log^3 q}{\sqrt{q}}\right)\right).$$

Proof. First, the explicit formula (3.1) yields

$$\log\hat{\mu}_{q;a_1,\dots,a_r}(t_1,\dots,t_r) = \sum_{\substack{\chi\neq\chi_0\\ \chi \bmod q}}\sum_{\gamma_\chi>0}\log J_0\left(\frac{2\left|\sum_{j=1}^{r}\chi(a_j)t_j\right|}{\sqrt{\frac14+\gamma_\chi^2}}\right) + O\left(\|\mathbf{t}\|\sum_{j=1}^{r}|C_q(a_j)|\right).$$

Using Lemma 2.2 along with the standard estimate $\phi(q)\gg q/\log\log q$, we deduce that the error term above is $\ll q^{-1/2}d(q)\log^3 q$. On the other hand note that

$$\frac{2\left|\sum_{j=1}^{r}\chi(a_j)t_j\right|}{\sqrt{\frac14+\gamma_\chi^2}} \ll r\|\mathbf{t}\|\le 1$$

if q is large enough. Hence, using that $\log J_0(z)=-z^2/4+O(z^4)$ for $|z|\le 1$ we obtain

(3.2)

$$\log\hat{\mu}_{q;a_1,\dots,a_r}(t_1,\dots,t_r) = -\sum_{\substack{\chi\neq\chi_0\\ \chi \bmod q}}\sum_{\gamma_\chi>0}\frac{\left|\sum_{j=1}^{r}\chi(a_j)t_j\right|^2}{\frac14+\gamma_\chi^2} + O\left(r^4\|\mathbf{t}\|^4\sum_{\substack{\chi\neq\chi_0\\ \chi \bmod q}}\sum_{\gamma_\chi>0}\frac{1}{\left(\frac14+\gamma_\chi^2\right)^2}+\frac{d(q)\log^3 q}{\sqrt{q}}\right).$$

Since $\sum_{\substack{\chi\neq\chi_0\\ \chi \bmod q}}\sum_{\gamma_\chi>0}1/\left(\frac14+\gamma_\chi^2\right)^2\ll \mathrm{Var}(q)$, it follows that the error term in the above estimate is $\ll q^{-1/2}d(q)\log^3 q$. On the other hand, the main term on the RHS of (3.2) equals

$$-\sum_{\substack{\chi\neq\chi_0\\ \chi \bmod q}}\sum_{\gamma_\chi>0}\frac{1}{\frac14+\gamma_\chi^2}\sum_{1\leq j,k\leq r}\chi(a_j)\overline{\chi(a_k)}t_jt_k = -\frac{1}{2}\sum_{1\leq j,k\leq r}\mathrm{Cov}_{q;a_1,\dots,a_r}(j,k)t_jt_k$$
$$= -\frac{1}{2}\mathbf{t}^T\mathrm{Cov}_{q;a_1,\dots,a_r}\mathbf{t},$$

by Lemma 2.1. □

Next, we show that $\hat{\mu}_{q;a_1,\dots,a_r}(t_1,\dots,t_r)$ is rapidly decreasing in the range $\|\mathbf{t}\|\geq \mathrm{Var}(q)^{-1/2}$. In particular, the following result is a refinement of Proposition 3.2 of [12], which takes into account the dependence of the upper bounds on r.

Proposition 3.2. *There exists a constant $c_1>0$ such that, if q is large and $2\leq r\leq c_1\log q$, then uniformly for all $(a_1,\dots,a_r)\in\mathcal{A}_r(q)$ we have*

$$|\hat{\mu}_{q;a_1,\dots,a_r}(t_1,\dots,t_r)|\leq\begin{cases}\exp\left(-\dfrac{\phi(q)}{8r}\|\mathbf{t}\|\right) & \text{if } \|\mathbf{t}\|\geq 400,\\ \exp\left(-\dfrac{\phi(q)}{(\log q)^8}\right) & \text{if } (\log q)^{-2}\leq\|\mathbf{t}\|\leq 400,\\ \exp\left(-\dfrac{\phi(q)\log q}{4}\|\mathbf{t}\|^2\right) & \text{if } \|\mathbf{t}\|\leq(\log q)^{-2}.\end{cases}$$

Before proving this result we first require the following lemma.

Lemma 3.3. *Let q be large and $2\leq r\leq\phi(q)/4$ be an integer. For $\mathbf{a}=(a_1,\dots,a_r)\in\mathcal{A}_r(q)$ and $\mathbf{t}\in\mathbb{R}^r$ we denote by $M_{q,\mathbf{a}}(\mathbf{t})$ the set of nontrivial characters $\chi \bmod q$ such that $\left|\sum_{j=1}^r\chi(a_j)t_j\right|\geq\|\mathbf{t}\|/2$. Then*

$$|M_{q,\mathbf{a}}(\mathbf{t})|\geq\frac{\phi(q)}{2r}.$$

Proof. Let

$$S(\mathbf{t})=\sum_{\substack{\chi\neq\chi_0\\ \chi \bmod q}}\left|\sum_{j=1}^r\chi(a_j)t_j\right|^2=\sum_{\chi \bmod q}\left|\sum_{j=1}^r\chi(a_j)t_j\right|^2-\left(\sum_{j=1}^r t_j\right)^2 \tag{3.3}$$
$$=\sum_{j=1}^r\sum_{k=1}^r t_jt_k\sum_{\chi \bmod q}\chi(a_j)\overline{\chi(a_k)}-\left(\sum_{j=1}^r t_j\right)^2=\phi(q)\sum_{j=1}^r t_j^2-\left(\sum_{j=1}^r t_j\right)^2$$
$$\geq(\phi(q)-r)\|\mathbf{t}\|^2,$$

by the Cauchy–Schwarz inequality. Therefore, using that $\left|\sum_{j=1}^r\chi(a_j)t_j\right|^2\leq\left(\sum_{j=1}^r|t_j|\right)^2\leq r\|\mathbf{t}\|^2$, we deduce

$$S(\mathbf{t})=\sum_{\chi\in M_{q,\mathbf{a}}(\mathbf{t})}\left|\sum_{j=1}^r\chi(a_j)t_j\right|^2+\sum_{\chi\notin M_{q,\mathbf{a}}(\mathbf{t})}\left|\sum_{j=1}^r\chi(a_j)t_j\right|^2\leq r|M_{q,\mathbf{a}}(\mathbf{t})|\|\mathbf{t}\|^2+\frac{\phi(q)}{4}\|\mathbf{t}\|^2.$$

Combining this estimate with (3.3) completes the proof. □

Proof of Proposition 3.2. First, assume that $\|\mathbf{t}\| \geq 400$. For any Dirichlet character χ we define

$$F(x,\chi) := \prod_{\gamma_\chi > 0} J_0\left(\frac{2x}{\sqrt{\frac{1}{4}+\gamma_\chi^2}}\right).$$

Then it follows from Lemma 2.16 of [5] that for any nontrivial character χ mod q we have

$$|F(x,\chi)F(x,\overline{\chi})| \leq e^{-x} \tag{3.4}$$

for $x \geq 200$. Moreover, the explicit formula (3.1) implies

$$|\hat{\mu}_{q;a_1,\dots,a_r}(t_1,\dots,t_r)| = \prod_{\substack{\chi \neq \chi_0 \\ \chi \bmod q}} \left| F\left(\left|\sum_{j=1}^r \chi(a_j)t_j\right|, \chi\right)\right|.$$

Note that $\chi \in M_{q,\mathbf{a}}(\mathbf{t})$ if and only if $\overline{\chi} \in M_{q,\mathbf{a}}(\mathbf{t})$. Hence

$$\prod_{\chi \in M_{q,\mathbf{a}}(\mathbf{t})} \left| F\left(\left|\sum_{j=1}^r \chi(a_j)t_j\right|, \chi\right)\right| = \prod_{\chi \in M_{q,\mathbf{a}}(\mathbf{t})} \left| F\left(\left|\sum_{j=1}^r \chi(a_j)t_j\right|, \overline{\chi}\right)\right|. \tag{3.5}$$

Furthermore, if $\chi \in M_{q,\mathbf{a}}(\mathbf{t})$ then $\left|\sum_{j=1}^r \chi(a_j)t_j\right| \geq \|\mathbf{t}\|/2 \geq 200$. Therefore combining (3.4) and (3.5) along with the trivial bound $|F(x,\chi)| \leq 1$ (since $|J_0(x)| \leq 1$) we derive

$$\begin{aligned}
|\hat{\mu}_{q;a_1,\dots,a_r}(t_1,\dots,t_r)|^2 &\leq \left(\prod_{\chi \in M_{q,\mathbf{a}}(\mathbf{t})} \left| F\left(\left|\sum_{j=1}^r \chi(a_j)t_j\right|, \chi\right)\right|\right)^2 \\
&= \prod_{\chi \in M_{q,\mathbf{a}}(\mathbf{t})} \left| F\left(\left|\sum_{j=1}^r \chi(a_j)t_j\right|, \chi\right) F\left(\left|\sum_{j=1}^r \chi(a_j)t_j\right|, \overline{\chi}\right)\right| \\
&\leq \exp\left(-\sum_{\chi \in M_{q,\mathbf{a}}(\mathbf{t})} \left|\sum_{j=1}^r \chi(a_j)t_j\right|\right) \leq \exp\left(-\frac{1}{2}|M_{q,\mathbf{a}}(\mathbf{t})|\|\mathbf{t}\|\right).
\end{aligned}$$

Thus, we infer from Lemma 3.3 that

$$|\hat{\mu}_{q;a_1,\dots,a_r}(t_1,\dots,t_r)| \leq \exp\left(-\frac{1}{4}|M_{q,\mathbf{a}}(\mathbf{t})|\|\mathbf{t}\|\right) \leq \exp\left(-\frac{\phi(q)}{8r}\|\mathbf{t}\|\right),$$

as desired.

Let $\epsilon = (\log q)^{-2}$ and suppose that $\epsilon \leq \|t\| \leq 400$. If $\chi \in M_{q,\mathbf{a}}(\mathbf{t})$ then

$$\frac{2\left|\sum_{j=1}^r \chi(a_j)t_j\right|}{\sqrt{\frac{1}{4}+\gamma_\chi^2}} \geq \frac{\epsilon}{\sqrt{\frac{1}{4}+\gamma_\chi^2}}.$$

We also note that if q is sufficiently large then $\epsilon\left(\frac{1}{4}+\gamma_\chi^2\right)^{-1/2} \leq 2\epsilon \leq 1$. Therefore, since J_0 is a positive decreasing function on $[0,1]$ and $|J_0(z)| \leq J_0(1)$ for all $z \geq 1$, we get

$$|\hat{\mu}_{q;a_1,\ldots,a_r}(t_1,\ldots,t_r)| \leq \prod_{\chi\in M_{q,\mathbf{a}}(\mathbf{t})}\prod_{\gamma_\chi>0}\left|J_0\left(\frac{\epsilon}{\sqrt{\frac14+\gamma_\chi^2}}\right)\right|.$$

Furthermore, using the standard bound $|J_0(x)| \leq \exp(-x^2/4)$ for $|x|\leq 1$, we deduce that

$$|\hat{\mu}_{q;a_1,\ldots,a_r}(t_1,\ldots,t_r)| \leq \exp\left(-\frac{\epsilon^2}{4}\sum_{\chi\in M_{q,\mathbf{a}}(\mathbf{t})}\sum_{\gamma_\chi>0}\frac{1}{\frac14+\gamma_\chi^2}\right). \tag{3.6}$$

Let $N(T,\chi)$ denote the number of γ_χ in the interval $[0,T]$. Then, we have the classical estimate (see Chapters 15 and 16 of [3])

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

where q^* is the conductor of χ. Hence, if $T=\log^2 q$ then $N(T,\chi)\gg \log^2 q$. This yields

$$\sum_{\gamma_\chi>0}\frac{1}{\frac14+\gamma_\chi^2}\geq \sum_{0<\gamma_\chi\leq \log^2 q}\frac{1}{\frac14+\gamma_\chi^2}\gg\frac{1}{\log^2 q}.$$

The upper bound on $|\hat{\mu}_{q;a_1,\ldots,a_r}(t_1,\ldots,t_r)|$ then follows upon inserting this estimate in (3.6) and using Lemma 3.3.

Finally assume that $\|\mathbf{t}\|\leq (\log q)^{-2}$. If q is large enough then

$$\frac{2\left|\sum_{j=1}^r\chi(a_j)t_j\right|}{\sqrt{\frac14+\gamma_\chi^2}}\ll r\|\mathbf{t}\|\leq 1.$$

Hence, using that $|J_0(x)|\leq\exp(-x^2/4)$ for $|x|\leq 1$ we obtain from the explicit formula (3.1)

$$|\hat{\mu}_{q;a_1,\ldots,a_r}(t_1,\ldots,t_r)| \leq \exp\left(-\sum_{\substack{\chi\neq\chi_0\\ \chi \bmod q}}\sum_{\gamma_\chi>0}\frac{\left|\sum_{j=1}^r\chi(a_j)t_j\right|^2}{\frac14+\gamma_\chi^2}\right). \tag{3.7}$$

Furthermore, Lemma 2.2 yields

$$\begin{aligned}\sum_{\substack{\chi\neq\chi_0\\ \chi \bmod q}}\sum_{\gamma_\chi>0}\frac{\left|\sum_{j=1}^r\chi(a_j)t_j\right|^2}{\frac14+\gamma_\chi^2} &= \sum_{\substack{\chi\neq\chi_0\\ \chi \bmod q}}\sum_{\gamma_\chi>0}\frac{1}{\frac14+\gamma_\chi^2}\sum_{1\leq j,k\leq r}\chi(a_j)\overline{\chi(a_k)}t_jt_k\\ &= \frac{\mathrm{Var}(q)}{2}\left(t_1^2+\cdots+t_r^2\right)+\sum_{1\leq j<k\leq r}B_q(a_j,a_k)t_jt_k\\ &= \frac{\phi(q)\log q}{2}\|\mathbf{t}\|^2\left(1+O\left(\frac{r+\log\log q}{\log q}\right)\right),\end{aligned}$$

since

$$\sum_{1\le j<k\le r} |t_j t_k| \le \left(\sum_{j=1}^r |t_j|\right)^2 \le r\|\mathbf{t}\|^2,$$

by the Cauchy–Schwarz inequality. Thus, if $r \le c_1 \log q$ where $c_1 > 0$ is suitably small, then

$$\sum_{\substack{\chi\ne\chi_0\\ \chi \bmod q}} \sum_{\gamma_\chi>0} \frac{\left|\sum_{j=1}^r \chi(a_j)t_j\right|^2}{\frac14+\gamma_\chi^2} \ge \frac{\phi(q)\log q}{4}\|\mathbf{t}\|^2.$$

Inserting this estimate in (3.7) completes the proof. □

4. The asymptotic behavior of the densities $\delta_{q;a_1,\dots,a_r}$: Proof of Theorems 1.1, 1.2 and 1.3

We showed in the previous section that in a small region around 0, the Fourier transform of $\mu_{q;a_1,\dots,a_r}$ can be approximated by the Fourier transform of a multivariate normal distribution whose covariance matrix equals $\mathrm{Cov}_{q;a_1,\dots,a_r}$. If we normalize by $\sqrt{\mathrm{Var}(q)}$ then Proposition 3.1 above implies that in the range $\|\mathbf{t}\| \le \log^2 q$ we have

(4.1)
$$\hat{\mu}_{q;a_1,\dots,a_r}\left(\frac{t_1}{\sqrt{\mathrm{Var}(q)}},\dots,\frac{t_r}{\sqrt{\mathrm{Var}(q)}}\right) = \exp\left(-\frac12\mathbf{t}^T\mathcal{C}\mathbf{t}\right)\left(1+O\left(\frac{d(q)\log^3 q}{\sqrt{q}}\right)\right),$$

where $\mathcal{C}$ is an $r\times r$ symmetric matrix whose entries are

$$\mathcal{C}_{jk} = \begin{cases} 1 & \text{if } j=k,\\ \dfrac{B_q(a_j,a_k)}{\mathrm{Var}(q)} \ll \dfrac{1}{\log q} & \text{if } j\ne k.\end{cases}$$

Let $\mathcal{M}_r(\epsilon)$ denote the set of $r\times r$ symmetric matrices $A=(a_{jk})$ such that $a_{jj}=1$ for all $1\le j\le r$ and $|a_{jk}|\le\epsilon$ for all $1\le j\ne k\le r$. In order to prove Theorems 1.1-1.3, we need to investigate multivariate normal distributions whose covariance matrices belong to $\mathcal{M}_r(\epsilon)$ where $\epsilon \ll 1/\log q$ is small. To this end we shall study the density function of a multivariate normal distribution, which is given by

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{r/2}\sqrt{\det(A)}}\exp\left(-\frac12\mathbf{x}^T A^{-1}\mathbf{x}\right), \tag{4.2}$$

if A is the covariance matrix of the distribution.

Our first lemma shows that the determinant of any matrix $A\in\mathcal{M}_r(\epsilon)$ is close to 1 if ϵ is small enough.

Lemma 4.1. *If $\epsilon\le 1/(2r)$ then for any $A\in\mathcal{M}_r(\epsilon)$ we have* $\det(A) = 1+O(\epsilon^2 r^2)$.

Proof. Let S_r be the set of all permutations σ of $\{1,\dots,r\}$. Then we have

$$\det(A) = \sum_{\sigma\in S_r}\mathrm{sgn}(\sigma)a_{1\sigma(1)}\cdots a_{r\sigma(r)} = 1+\sum_{\substack{\sigma\in S_r\\ \sigma\ne\mathbf{1}}}\mathrm{sgn}(\sigma)a_{1\sigma(1)}\cdots a_{r\sigma(r)}, \tag{4.3}$$

where $\mathbf{1}$ denotes the identity permutation. For $0\le k\le r$ let $S_r(k)$ be the set of permutations $\sigma\in S_r$ such that the equation $\sigma(j)=j$ has exactly $r-k$ solutions in

$\{1,\ldots,r\}$. Then $S_r(0)=\{\mathbf{1}\}$, $S_r(1)=\emptyset$ and more generally one has

$$|S_r(k)|\leq \binom{r}{r-k}(k-1)!\leq r^k, \text{ for } 2\leq k\leq r.$$

Moreover, note that $|a_{1\sigma(1)}\cdots a_{r\sigma(r)}|\leq \epsilon^k$, for all $\sigma\in S_r(k)$.

Hence, we deduce

$$\sum_{\substack{\sigma\in S_r\\ \sigma\neq \mathbf{1}}}\operatorname{sgn}(\sigma)a_{1\sigma(1)}\cdots a_{r\sigma(r)}=\sum_{k=2}^{r}\sum_{\sigma\in S_r(k)}\operatorname{sgn}(\sigma)a_{1\sigma(1)}\cdots a_{r\sigma(r)}\ll\sum_{k=2}^{r}(\epsilon r)^k\ll \epsilon^2r^2.$$

Inserting this estimate in (4.3) implies the result. □

In order to understand the behavior of the density function $f(\mathbf{x})$ we need to determine the size of the entries $\{\tilde{a}_{jk}\}$ of A^{-1}, if $A\in\mathcal{M}_r(\epsilon)$. The next lemma shows that if ϵ is small then the diagonal entries are close to 1 and the off-diagonal ones are small.

Lemma 4.2. *If $\epsilon\leq 1/(2r)$ then for any $A\in\mathcal{M}_r(\epsilon)$ we have*

$$\tilde{a}_{jk}=\begin{cases}1+O(\epsilon^2r^2) & \text{if } j=k,\\ O(\epsilon) & \text{if } j\neq k.\end{cases}$$

Proof. Recall that

$$\tilde{a}_{jk}=\frac{1}{\det(A)}(-1)^{j+k}M_{kj},$$

where M_{kj} is the minor of the entry a_{kj}, which is given by $M_{kj}=\det(A_{kj})$ and A_{kj} is the matrix obtained from A by deleting the kth row and the jth column.

First, we determine the size of the diagonal entries $\tilde{a}_{jj}$. In this case, remark that $A_{jj}\in\mathcal{M}_{r-1}(\epsilon)$. Hence, it follows from Lemma 4.1 that

$$\tilde{a}_{jj}=\frac{\det(A_{jj})}{\det(A)}=1+O(\epsilon^2r^2).$$

Now, we handle the off-diagonal entries. For $1\leq j\neq k\leq r$, let $\mathcal{B}_{j,k}$ denote the set of all bijections σ from $\{1,\ldots,r\}\setminus\{j\}$ to $\{1,\ldots,r\}\setminus\{k\}$. Then, we have

$$|M_{jk}|=|\det(A_{jk})|\leq\sum_{\sigma\in\mathcal{B}_{j,k}}\prod_{1\leq n\neq j\leq r}|a_{n\sigma(n)}|.$$

For $0\leq l\leq r-1$ we define $\mathcal{B}_{j,k}(l)$ to be the set of bijections $\sigma\in\mathcal{B}_{j,k}$ such that the equation $\sigma(m)=m$ has exactly $r-1-l$ solutions. Since $\sigma(k)\neq k$ then it follows that $\mathcal{B}_{j,k}(0)=\emptyset$, and more generally one has

$$|\mathcal{B}_{j,k}(l)|\leq\binom{r-2}{r-1-l}(l-1)!\leq r^{l-1}, \text{ for } 1\leq l\leq r-1.$$

Hence we obtain

$$|M_{jk}|\leq\sum_{l=1}^{r-1}\sum_{\sigma\in\mathcal{B}_{j,k}(l)}\prod_{1\leq n\neq j\leq r}|a_{n\sigma(n)}|\ll\sum_{l=1}^{r-1}r^{l-1}\epsilon^l\ll\epsilon.$$

Combining this bound with Lemma 4.1 yield the desired bound $\tilde{a}_{jk}\ll\epsilon$. □

We know that the Fourier transform of a multivariate Gaussian of covariance matrix A is (up to normalization) a multivariate Gaussian of covariance A^{-1}. The last ingredient we need to prove Theorems 1.1–1.3 is an approximate version of this statement when $A \in \mathcal{M}_r(\epsilon)$.

Lemma 4.3. *Let $r \geq 2$ be a positive integer, $R \geq 10\sqrt{r}$ be a real number and $\mathbf{x} \in \mathbb{R}^r$. If $\epsilon \leq 1/(2r)$ then for any $A \in \mathcal{M}_r(\epsilon)$ we have*

$$\begin{aligned}(2\pi)^{-r} \int_{\|\mathbf{t}\| \leq R} & \mathrm{e}^{\mathrm{i}(t_1x_1+\cdots+t_rx_r)} \exp\left(-\frac{1}{2}\mathbf{t}^T A\mathbf{t}\right) d\mathbf{t} \\ &= \frac{1}{(2\pi)^{r/2}\sqrt{\det(A)}} \exp\left(-\frac{1}{2}\mathbf{x}^T A^{-1}\mathbf{x}\right) \\ &\quad + O\left(\exp\left(-\frac{R^2}{5}\right)\right).\end{aligned}$$

Proof. Since $\exp\left(-\frac{1}{2}\mathbf{t}^T A\mathbf{t}\right)$ is the Fourier transform of the multivariate normal distribution whose density equals

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{r/2}\sqrt{\det(A)}} \exp\left(-\frac{1}{2}\mathbf{x}^T A^{-1}\mathbf{x}\right),$$

then the Fourier inversion formula yields

$$\begin{aligned}(2\pi)^{-r} \int_{\mathbf{t}\in\mathbb{R}^r} & \mathrm{e}^{\mathrm{i}(t_1x_1+\cdots+t_rx_r)} \exp\left(-\frac{1}{2}\mathbf{t}^T A\mathbf{t}\right) d\mathbf{t} \\ &= \frac{1}{(2\pi)^{r/2}\sqrt{\det(A)}} \exp\left(-\frac{1}{2}\mathbf{x}^T A^{-1}\mathbf{x}\right).\end{aligned} \tag{4.4}$$

Moreover, since $|a_{jk}| \leq 1/(2r)$ for $j \neq k$ then

$$\left|\sum_{1\leq j\neq k\leq r} a_{jk}t_jt_k\right| \leq \frac{1}{2r}\left(\sum_{j=1}^r |t_j|\right)^2 \leq \frac{1}{2}\sum_{j=1}^r t_j^2,$$

by the Cauchy–Schwarz inequality. This implies

$$\mathbf{t}^T A\mathbf{t} = \sum_{j=1}^r \sum_{k=1}^r a_{jk}t_jt_k \geq \frac{1}{2}\sum_{j=1}^r t_j^2. \tag{4.5}$$

Hence, we get

$$\begin{aligned}(2\pi)^{-r} & \int_{\|\mathbf{t}\|>R} \exp\left(-\frac{1}{2}\mathbf{t}^T A\mathbf{t}\right) d\mathbf{t} \\ &\leq (2\pi)^{-r} \int_{\|\mathbf{t}\|>R} \exp\left(-\frac{1}{4}\|\mathbf{t}\|^2\right) d\mathbf{t} \ll \exp\left(-\frac{R^2}{5}\right),\end{aligned}$$

which in view of (4.4) completes the proof. $\square$

Proof of Theorem 1.1. To lighten the notation we shall write δ_q for $\delta_{q;a_1,\dots,a_r}$ and μ_q for $\mu_{q;a_1,\dots,a_r}$. Let $R = 3\sqrt{\mathrm{Var}(q)}\log q$. First, using Lemma 2.3 we derive

$$\delta_q = \int_{y_1>y_2>\cdots>y_r} d\mu_q(y_1,\dots,y_r) = \int_{\substack{y_1>y_2>\cdots>y_r\\ |\mathbf{y}|_\infty\le R}} d\mu_q(y_1,\dots,y_r) + O\left(\exp\left(-2\log^2 q\right)\right). \tag{4.6}$$

Next, we apply the Fourier inversion formula to the measure μ_q to obtain

$$\int_{\substack{y_1>y_2>\cdots>y_r\\ |\mathbf{y}|_\infty\le R}} d\mu_q(y_1,\dots,y_r) = (2\pi)^{-r}\int_{\substack{y_1>y_2>\cdots>y_r\\ |\mathbf{y}|_\infty\le R}}\int_{\mathbf{s}\in\mathbb{R}^r} \mathrm{e}^{\mathrm{i}(s_1y_1+\cdots+s_ry_r)}\hat{\mu}_q(s_1,\dots,s_r)\,d\mathbf{s}\,d\mathbf{y}.$$

Since the Fourier transform $\hat{\mu}_q(s_1,\dots,s_r)$ is rapidly decreasing, we shall deduce that the main contribution to the integral over $\mathbb{R}^r$ of $\mathrm{e}^{\mathrm{i}(s_1y_1+\cdots+s_ry_r)}\hat{\mu}_q(s_1,\dots,s_r)$ comes from a small ball centered at 0. Indeed, we infer from Proposition 3.2 that

$$\int_{\mathbf{s}\in\mathbb{R}^r} \mathrm{e}^{\mathrm{i}(s_1y_1+\cdots+s_ry_r)}\hat{\mu}_q(s_1,\dots,s_r)d\mathbf{s} = \int_{\|\mathbf{s}\|\le\epsilon} \mathrm{e}^{\mathrm{i}(s_1y_1+\cdots+s_ry_r)}\hat{\mu}_q(s_1,\dots,s_r)d\mathbf{s} + O\left(\exp\left(-2\log^2 q\right)\right),$$

where $\epsilon = 3(\mathrm{Var}(q))^{-1/2}\log q$. Hence we obtain

$$\delta_q = (2\pi)^{-r}\int_{\substack{y_1>y_2>\cdots>y_r\\ |\mathbf{y}|_\infty\le R}}\int_{\|\mathbf{s}\|\le\epsilon} \mathrm{e}^{\mathrm{i}(s_1y_1+\cdots+s_ry_r)}\hat{\mu}_q(s_1,\dots,s_r)\,d\mathbf{s}\,d\mathbf{y} + O\left(\exp\left(-\log^2 q\right)\right), \tag{4.7}$$

since $R^r \ll \exp(r\log q)$. Now, we make the change of variables

$$t_j := \sqrt{\mathrm{Var}(q)}s_j \text{ and } x_j := \frac{y_j}{\sqrt{\mathrm{Var}(q)}}, \text{ for all } 1\le j\le r$$

to obtain

$$\delta_q = (2\pi)^{-r}\int_{\substack{x_1>x_2>\cdots>x_r\\ |\mathbf{x}|_\infty\le 3\log q}}\int_{\|\mathbf{t}\|\le 3\log q} \mathrm{e}^{\mathrm{i}(t_1x_1+\cdots+t_rx_r)}\hat{\mu}_q \times\left(\frac{t_1}{\sqrt{\mathrm{Var}(q)}},\dots,\frac{t_r}{\sqrt{\mathrm{Var}(q)}}\right)d\mathbf{t}d\mathbf{x} + O\left(\exp\left(-\log^2 q\right)\right). \tag{4.8}$$

Replacing $\hat{\mu}_q\left(\frac{t_1}{\sqrt{\mathrm{Var}(q)}},\dots,\frac{t_r}{\sqrt{\mathrm{Var}(q)}}\right)$ by the approximation (4.1) that we derived in Proposition 3.1 yields

$$\delta_q = (2\pi)^{-r}\int_{\substack{x_1>x_2>\cdots>x_r\\ |\mathbf{x}|_\infty\le 3\log q}}\int_{\|\mathbf{t}\|\le 3\log q} \mathrm{e}^{\mathrm{i}(t_1x_1+\cdots+t_rx_r)}\exp\left(-\frac{1}{2}\mathbf{t}^T\mathcal{C}\mathbf{t}\right)d\mathbf{t}d\mathbf{x} + E_1,$$

where

$$E_1 \ll q^{-1/3}(\log q)^{3r} \ll q^{-1/4},$$

since $d(q) = q^{o(1)}$ and $\mathbf{t}^T\mathcal{C}\mathbf{t} \geq 0$ by (4.5). Furthermore, applying Lemma 4.3 we derive

$$\delta_q = \frac{1}{(2\pi)^{r/2}\sqrt{\det(\mathcal{C})}} \int_{\substack{x_1>x_2>\cdots>x_r\\ |\mathbf{x}|_\infty \leq 3\log q}} \exp\left(-\frac{1}{2}\mathbf{x}^T\mathcal{C}^{-1}\mathbf{x}\right) d\mathbf{x} + O\left(q^{-1/4}\right). \tag{4.9}$$

Since $\mathcal{C}_{jk} = B_q(a_j, a_k)/\mathrm{Var}(q) \ll (\log q)^{-1}$ for $j \neq k$ by Lemma 2.2, there exists an absolute constant $\alpha_0 > 0$ such that $\mathcal{C} \in \mathcal{M}_r(\beta)$ with $\beta = \alpha_0/\log q$. Therefore, appealing to Lemma 4.2 we obtain

$$\begin{aligned}
\mathbf{x}^T\mathcal{C}^{-1}\mathbf{x} &= \left(1 + O\left(\frac{r^2}{\log^2 q}\right)\right)\sum_{j=1}^r x_j^2 + O\left(\frac{1}{\log q}\left(\sum_{j=1}^r |x_j|\right)^2\right)\\
&= \left(1 + O\left(\frac{r}{\log q}\right)\right)\|\mathbf{x}\|^2,
\end{aligned}$$

which follows from the Cauchy–Schwarz inequality. Hence we deduce

$$-\frac{1}{2}\left(1 + \frac{\alpha_1 r}{\log q}\right)\|\mathbf{x}\|^2 \leq -\frac{1}{2}\mathbf{x}^T\mathcal{C}^{-1}\mathbf{x} \leq -\frac{1}{2}\left(1 - \frac{\alpha_1 r}{\log q}\right)\|\mathbf{x}\|^2, \tag{4.10}$$

for some absolute constant $\alpha_1 > 0$. This implies

$$\begin{aligned}
&\int_{\substack{x_1>x_2>\cdots>x_r\\ |\mathbf{x}|_\infty > 3\log q}} \exp\left(-\frac{1}{2}\mathbf{x}^T\mathcal{C}^{-1}\mathbf{x}\right) d\mathbf{x}\\
&\qquad \leq \int_{|\mathbf{x}|_\infty > 3\log q} \exp\left(-\frac{1}{4}\|\mathbf{x}\|^2\right) d\mathbf{x} \ll \exp\left(-\log^2 q\right).
\end{aligned}$$

Inserting this estimate in (4.9) and using Lemma 4.1 we get

(4.11)

$$\delta_q = \left(1 + O\left(\frac{r^2}{\log^2 q}\right)\right)\frac{1}{(2\pi)^{r/2}}\int_{x_1>x_2>\cdots>x_r} \exp\left(-\frac{1}{2}\mathbf{x}^T\mathcal{C}^{-1}\mathbf{x}\right) d\mathbf{x} + O\left(q^{-1/4}\right).$$

Let κ be a real number such that $|\kappa| \leq \alpha_1 r/\log q$. Since the function $\|\mathbf{x}\|^2$ is symmetric in the variables $\{x_j\}_{1\leq j\leq r}$ we obtain

$$\begin{aligned}
&\frac{1}{(2\pi)^{r/2}}\int_{x_1>x_2>\cdots>x_r} \exp\left(-\frac{1}{2}(1+\kappa)\|\mathbf{x}\|^2\right) d\mathbf{x}\\
&\qquad = \frac{1}{r!}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(1+\kappa)y^2\right) dy\right)^r\\
&\qquad = \frac{1}{r!(1+\kappa)^{r/2}} = \frac{1}{r!}\exp\left(O\left(\frac{r^2}{\log q}\right)\right).
\end{aligned} \tag{4.12}$$

The theorem follows upon combining this estimate with (4.10) and (4.11). □

Proof of Theorem 1.2. The result can be obtained by proceeding along the same lines as the proof of Theorem 1.1, except that we make a different choice of parameters in this case. Indeed, choosing $R = 5\sqrt{\mathrm{Var}(q) r \log r}$ and using Lemma 2.3 and Proposition 3.2, we obtain analogously to (4.8)

$$\begin{aligned}\delta_q = (2\pi)^{-r} &\int_{\substack{x_1>x_2>\cdots>x_r\\ |\mathbf{x}|_\infty \le 5\sqrt{r\log r}}} \int_{\|\mathbf{t}\|\le 3\log q} \mathrm{e}^{\mathrm{i}(t_1x_1+\cdots+t_rx_r)} \hat{\mu}_q \\ &\times \left(\frac{t_1}{\sqrt{\mathrm{Var}(q)}}, \dots, \frac{t_r}{\sqrt{\mathrm{Var}(q)}}\right) d\mathbf{t} d\mathbf{x} \\ &+ O\left(\exp\left(-4r\log r\right)\right).\end{aligned} \tag{4.13}$$

Moreover, we infer from (4.1) that

$$\delta_q = (2\pi)^{-r} \int_{\substack{x_1>x_2>\cdots>x_r\\ |\mathbf{x}|_\infty \le 5\sqrt{r\log r}}} \int_{\|\mathbf{t}\|\le 3\log q} \mathrm{e}^{\mathrm{i}(t_1x_1+\cdots+t_rx_r)} \exp\left(-\frac{1}{2}\mathbf{t}^T\mathcal{C}\mathbf{t}\right) d\mathbf{t} d\mathbf{x} + E_2, \tag{4.14}$$

where

$$\begin{aligned}E_2 \ll \frac{d(q)\log^3 q}{\sqrt{q}} (2\pi)^{-r} &\int_{\substack{x_1>x_2>\cdots>x_r\\ |\mathbf{x}|_\infty \le 5\sqrt{r\log r}}} d\mathbf{x} \int_{\|\mathbf{t}\|\le 3\log q} \\ &\times \exp\left(-\frac{1}{2}\mathbf{t}^T\mathcal{C}\mathbf{t}\right) d\mathbf{t} + \exp\left(-4r\log r\right).\end{aligned} \tag{4.15}$$

Note that

$$\begin{aligned}\int_{\substack{x_1>x_2>\cdots>x_r\\ |\mathbf{x}|_\infty \le 5\sqrt{r\log r}}} d\mathbf{x} = \frac{1}{r!}\int_{|\mathbf{x}|_\infty \le 5\sqrt{r\log r}} d\mathbf{x} &= \frac{(10\sqrt{r\log r})^r}{r!} \\ &= \exp\left(-\frac{r\log r}{2} + O(r\log\log r)\right),\end{aligned}$$

by Stirling's formula. On the other hand, it follows from (4.5) that

$$\frac{1}{(2\pi)^r}\int_{\|\mathbf{t}\|\le 3\log q} \exp\left(-\frac{1}{2}\mathbf{t}^T\mathcal{C}\mathbf{t}\right) d\mathbf{t} \le \frac{1}{(2\pi)^r}\int_{\mathbf{t}\in\mathbb{R}^r} \exp\left(-\frac{\|\mathbf{t}\|^2}{4}\right) d\mathbf{t} = \frac{1}{\pi^{r/2}}.$$

Therefore, inserting these estimates in (4.15) and using the classical bound $d(q) = \exp\left(O(\log q/\log\log q)\right)$ we deduce

$$E_2 \ll \exp\left(-\frac{1}{2}(\log q + r\log r) + O\left(\frac{\log q}{\log\log q} + r\log\log r\right)\right) + \exp\left(-4r\log r\right).$$

Continuing along the same line as in the proof of Theorem 1.1, we obtain analogously to (4.11)

$$\delta_q = \left(1 + O\left(\frac{r^2}{\log^2 q}\right)\right)\frac{1}{(2\pi)^{r/2}}\int_{x_1>x_2>\cdots>x_r} \exp\left(-\frac{1}{2}\mathbf{x}^T\mathcal{C}^{-1}\mathbf{x}\right) d\mathbf{x} + E_3, \tag{4.16}$$

where

$$E_3 \ll \exp\left(-\frac{1}{2}(\log q + r\log r) + O\left(\frac{\log q}{\log\log q} + r\log\log r\right)\right) + \exp\left(-4r\log r\right).$$

Furthermore, it follows from (4.10) and (4.12) that

$$\begin{aligned}\frac{1}{(2\pi)^{r/2}} \int_{x_1>x_2>\cdots>x_r} & \exp\left(-\frac{1}{2}\mathbf{x}^T\mathcal{C}^{-1}\mathbf{x}\right) d\mathbf{x} \\ &= \frac{1}{r!} \exp\left(O\left(\frac{r^2}{\log q}\right)\right) \\ &= \exp\left(-r\log r + r + O\left(\log r + \frac{r^2}{\log q}\right)\right),\end{aligned}$$

by Stirling's formula. Inserting this estimate in (4.16) completes the proof. □

Proof of Theorem 1.3. Since $\mu_{q;a_1,\dots,a_r}$ is absolutely continuous with respect to the Lebesgue measure, it follows from (1.1) that

$$\delta_{q;a_1,\dots,a_{r-1}} = \delta_{q;a_r,a_1,\dots,a_{r-1}} + \delta_{q;a_1,a_r,\dots,a_{r-1}} + \cdots + \delta_{q;a_1,\dots,a_{r-1},a_r}.$$

Hence, if $2 \le s < r \le \phi(q)$ are positive integers then

$$\max_{(a_1,\dots,a_r)\in\mathcal{A}_r(q)} \delta_{q;a_1,\dots,a_r} < \max_{(b_1,\dots,b_s)\in\mathcal{A}_s(q)} \delta_{q;b_1,\dots,b_s}. \tag{4.17}$$

On the other hand, using Theorem 1.2 with $s = [(1-\epsilon/2)\log q/\log\log q]$, we get

$$\max_{(b_1,\dots,b_s)\in\mathcal{A}_s(q)} \delta_{q;b_1,\dots,b_s} = \exp\left(-s\log s + s + O\left(\log s + \frac{r^2}{\log q}\right)\right) \ll_\epsilon \frac{1}{q^{1-\epsilon}}.$$

The theorem follows upon combining this inequality with (4.17). □

Acknowledgments

The author was supported by a postdoctoral fellowship from the Natural Sciences and Engineering Research Council of Canada.

References

[1] C. Bays and R.H. Hudson, *The mean behavior of primes in arithmetic progressions.* J. Reine Angew. Math. **296** (1977), 80–99.

[2] C. Bays and R.H. Hudson, *The cyclic behavior of primes in the arithmetic progressions modulo* 11. J. Reine Angew. Math. **339** (1983), 215–220.

[3] H. Davenport, *Multiplicative number theory.* Graduate Texts in Mathematics, **74**. Springer–Verlag, New York, 2000.

[4] A. Feuerverger and G. Martin, *Biases in the Shanks–Rényi prime number race.* Exp. Math. **9**(4) (2000), 535–570.

[5] D. Fiorilli and G. Martin, *Inequities in the Shanks–Rényi prime number race: an asymptotic formula for the densities.* 76 pages. To appear in J. Reine Angew. Math.

[6] K. Ford and S. Konyagin, *The prime number race and zeros of L-functions off the critical line.* Duke Math. J. **113**(2) (2002), 313–330.

[7] K. Ford and S. Konyagin, *The prime number race and zeros of L-functions off the critical line. II.* Proceedings of the Session in Analytic Number Theory and Diophantine Equations, 40 pp., Bonner Math. Schriften, 360, Univ. Bonn, Bonn, 2003.

[8] J. Kaczorowski, *A contribution to the Shanks–Rényi race problem.* Quart. J. Math. Oxford Ser. (2) **44**(176) (1993), 451–458.

[9] J. Kaczorowski, *Results on the distribution of primes.* J. Reine Angew. Math. **446** (1994), 89–113.

[10] J. Kaczorowski, *On the Shanks–Rényi race problem.* Acta Arith. **74**(1) (1996), 31–46.

[11] S. Knapowski and P. Turán, *Comparative prime-number theory. I.* Acta Math. Acad. Sci. Hungar. **13** (1962), 299–314; II. 13 (1962), 315–342; III. 13 (1962), 343–364; IV. 14 (1963), 31–42; V. 14 (1963), 43–63; VI. 14 (1963), 65–78; VII. 14 (1963), 241–250; VIII. 14 (1963), 251–268.

[12] Y. Lamzouri, *Prime number races with three or more competitors*. 38 pages. arXiv:1101.0836.

[13] Y. Lamzouri, *Large deviations of the limiting distribution in the Shanks–Rényi prime number race*. 20 pages. To appear in Math. Proc. Cambridge Philos. Soc.

[14] J.E. Littlewood, *Distribution des nombres premiers*. C. R. Acad. Sci. Paris **158** (1914), 1869–1872.

[15] G. Martin, *Asymmetries in the Shanks–Rényi prime number race*. Number Theory for the Millennium, II (Urbana, IL, 2000), 403–415, A K Peters, Natick, MA, 2002.

[16] M. Rubinstein and P. Sarnak, *Chebyshev's bias*. Experiment. Math. **3**(3) (1994), 173–197.

Department of Mathematics and Statistics
York University
4700 Keele Street, Toronto, ON, M3J1P3
Canada
E-mail address: `lamzouri@mathstat.yorku.ca`