Research Article

Benjamin Justus

The distribution of quadratic residues and non-residues in the Goldwasser–Micali type of cryptosystem. II

Abstract: We provide three statistical laws concerning the limit distribution of quadratic residues and quadratic non-residues in $\mathbb{Z}/N\mathbb{Z}$, where $N = pq$ is an RSA modulus used in the Goldwasser–Micali cryptosystem.

Keywords: Quadratic residuosity problem, Goldwasser–Micali cryptosystem, quadratic residues distribution

MSC 2010: 94A60, 11N64, 60G50

Benjamin Justus: Klosterhof 7, 82405 Wessobrunn, Germany, e-mail: benjaminjustus@gmail.com

Communicated by: Spyros S. Magliveras

1 Introduction

The present paper continues to study the *quadratic residuosity problem*. Given a composite integer N, and a positive integer a relative prime to N, the quadratic residuosity problem is to decide whether a is a quadratic residue or a quadratic non-residue modulo N (i.e. whether or not $x^2 = a \bmod N$ has a solution). The question we investigate in the present article is:

- Are there any statistical laws that govern the distribution of quadratic residues and non-residues in $\mathbb{Z}/N\mathbb{Z}$ where N is a large RSA modulus?

We are here of course not completely aimless in posing such a question. Indeed one is motivated to study the problem by facts already known in the case of finite fields $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. When the base field is $\mathbb{F}_p$, the quadratic character associated with $\mathbb{F}_p$ is the Legendre symbol $\chi = (\frac{\cdot}{p})$. In such a case, χ can be viewed as a pseudo-quadratic residue characteristic function because $\chi(n)$ takes on the values $1, -1$ for n being a quadratic residue or non-residue, respectively. The values of the character sum

$$S(\chi; H, x) := \sum_{x<n\le x+H} \chi(n)$$

thus encode information about how the quadratic residues and non-residues are distributed in $\mathbb{F}_p$. Concerning the distribution of $S(\chi; H, x)$ when χ is quadratic, Davenport and Erdös [3] first proved the following interesting result:

$$\frac{1}{p}\left|\left\{0 \le x \le p-1 : \frac{S(\chi; H, x)}{\sqrt{H}} \le \lambda\right\}\right| \to \frac{1}{2\pi}\int_{-\infty}^{\lambda} \exp\left(-\frac{x^2}{2}\right)dx$$

provided $\log H = o(\log p)$ and $p, H \to \infty$. This result can be interpreted as follows: as x runs over elements of $\mathbb{F}_p$, the values of the character sum $S(\chi; H, x)$ tend to a Gaussian distribution of mean 0 and variance H.

As $\chi(n)$ takes on the values $1, -1$ more or less randomly with equal probability $1/2$, one may suspect that Gaussian distribution behaviors exist for sums involving a much larger class of functions. Indeed, let f be a random multiplicative function whose values at prime arguments are independent random variables taking on the values $1, -1$ with equal probability $1/2$. Extend the domain of f to the square-free integers by the multiplicative property of f. It is recently established that values of such sums $\sum f(n)$ have Gaussian distribution behaviors under suitable conditions [1, 5]. The result of Davenport and Erdös can also be generalized to include non-real character sums and exponential sums involving multiplicative as well as additive characters [10, 12].

The aim of the present paper is to establish an analogous distribution law in the setting $\mathbb{Z}/N\mathbb{Z}$ where N is a large RSA modulus.[1] It turns out that an extension to the setting $\mathbb{Z}/N\mathbb{Z}$ is not as obvious as it seems. First, unlike the $\mathbb{F}_p$ case, one does not have in the ring $\mathbb{Z}/N\mathbb{Z}$ a Dirichlet character χ which is capable of discerning whether a positive integer is a quadratic residue or non-residue modulo N. In the present paper, following the work of [9], we rely on the following quadratic residue characteristic function Φ:

$$\Phi(n) := \frac{1}{4}(1 + \chi_p(n))(1 + \chi_q(n)).$$

Here p, q are the prime factors of N, and χ_p and χ_q are the quadratic characters modulo p, q, respectively. If $\gcd(n, N) = 1$, then $\Phi(n) = 1$ if n is a quadratic residue and $\Phi(n) = 0$ if n is a non-residue modulo N. Similarly to short sums in $\mathbb{F}_p$, define

$$S(\Phi; H, x) := \sum_{x<n\le x+H} \Phi(n).$$

In order to study the distribution properties of $S(\Phi; H, x)$, we are led to consider the probabilistic model $X_1 + \cdots + X_H$ where the random variables $X_1, \ldots, X_H$ take on the values $1, 0$ with the corresponding probability $\delta, 1-\delta$, respectively, where $0 < \delta < 1$. This probabilistic model closely adheres to the reality: In $\mathbb{Z}/N\mathbb{Z}$, approximately $\delta = 1/4$ of the ring elements are quadratic residues, and $1 - \delta = 3/4$ of them are quadratic non-residues. Our analysis shows that such a probabilistic model $X_1 + \cdots + X_H$ gives rise to a Gaussian distribution of mean $H\delta$ and variance $H\delta(1-\delta)$, see Lemma 5.2.

Our study for the distribution of $S(\Phi; H, x)$ relies on a strategy of Lamzouri [10], as his method is quite effective in producing the main term of the distribution as well as the rate of convergence. The method is based on drawing connection between the probabilistic model and the moments defined as

$$M(r) := \frac{1}{N}\sum_{x=1}^{N} S(\Phi; H, x)^r.$$

Consequently, we are able to establish that $S(\Phi; H, x)$ has a Gaussian distribution of mean $\mu = H/4$ and variance $\sigma^2 = 3H/16$ provided $H \log H = o(\log N)$ and $H, N \to \infty$.

Theorem 1.1. *Let N be a large RSA modulus. Let two real numbers $a \le b$ be given. If both $\frac{\log N}{H \log H} \to \infty$ and $H \to \infty$, then*

$$\frac{1}{N}\left|\left\{1 \le x \le N : a \le \frac{S(\Phi; H, x)}{\sqrt{3H/16}} \le b\right\}\right| = \frac{1}{\sqrt{2\pi}}\int_a^b \exp\Big(-\frac{(x-\sqrt{H/3})^2}{2}\Big)dx + O\Big((b-a+1)\Big(H^{-1/6} + \sqrt{\frac{H\log H}{\log N}}\Big)\Big).$$

Remark 1.2. The term $\tilde{\mu} = \sqrt{H/3}$ inside the Gaussian kernel is the normalized mean (i.e. $\tilde{\mu} = \mu/\sigma$). Properly speaking, Theorem 1.1 says the normalized sum $S(\Phi; H, x)/\sqrt{3H/16}$ behaves Gaussian with mean $\sqrt{H/3}$ and variance 1.

Remark 1.3. The valid range for H in Theorem 1.1 is $H \log H = o(\log N)$, this is slightly weaker than the range obtained by Davenport and Erdös in the setting $\mathbb{Z}/p\mathbb{Z}$, i.e. $\log H = o(\log p)$. The slight loss of range is due to an asymmetry inherent in the probabilistic model. The same calculation would produce an error term in Theorem 1.1 without $\sqrt{H}$ in the symmetric setting, see Proposition 7.2 for details.

The second question we investigate in this paper has to do with how the quadratic residues and non-residues modulo N are distributed in $\mathbb{Z}/m\mathbb{Z}$, when m is small compared to N. There are two interesting aspects of this problem. We first introduce some notations. Define

$$\mathcal{Q}(R; a, m) := \{n = a + rm : 1 \le r \le R, \text{ and } n \text{ is a quadratic residue mod } N\},$$
$$\mathcal{NQ}(R; a, m) := \{n = a + rm : 1 \le r \le R, \text{and } n \text{ is a quadratic non-residue mod } N\}.$$

1 RSA modulus: $N = pq$ with $1 < p \ne q \le cN^{1/2}$ where $c > 0$.

Thus $|\mathcal{Q}|$ (resp. $|\mathcal{NQ}|$) counts the number of quadratic residues (resp. non-residues) in an arithmetic progression, say $a \pmod m$, where $1 \le a \le m$, and m is a modulus lying in a suitable range. Since the residues and non-residues are uniformly distributed in the interval $[1, N]$, one may suspect that the quadratic residues and non-residues should distribute uniformly over the residue classes modulo m. This indeed turns out to be the case, see Corollary 1.5.

Theorem 1.4. *Let N be a large RSA modulus. Suppose $\lambda > 0$ is a real number such that $\min(p, q) > \lambda N^{1/2}$. Let r, m, A, R be positive integers such that $A \le m, m \le \lambda N^{1/2}, mA(R+1) < \lambda N^{1+1/(2r)}$. Then we have*

$$\sum_{a \le A} |\mathcal{Q}(R; a, m)| = \frac{1}{4} AR + \mathrm{O}(E), \qquad \sum_{a \le A} |\mathcal{NQ}(R; a, m)| = \frac{3}{4} AR + \mathrm{O}(E),$$

where the error terms satisfy

$$E \ll_r A^{1-\frac{1}{r}} R^{1-\frac{1}{2r}} N^{\frac{r+1}{4r^2}} (\log N)^{1+\frac{1}{2r}}$$

if the following conditions hold:

$$A \ll N^{1/(2r)} \log N, \tag{1A}$$

$$R \gg N^{\frac{r-1}{2r(2r-1)}} (\log N)^{-\frac{1}{2r-1}}, \tag{1B}$$

$$\frac{AR^{1/2}}{N^{\frac{r+1}{4r}} (\log N)^{r+\frac{1}{2}}} \to \infty \quad \text{as } N \to \infty. \tag{1C}$$

Moreover, we have

$$E \ll_r A^{1-\frac{1}{2r}} R^{1-\frac{1}{2r}} N^{\frac{1}{4r}} (\log N)^{1+\frac{1}{2r}}$$

if the following condition holds:

$$\frac{AR}{N^{1/2} (\log N)^{2r+1}} \to \infty \quad \text{as } N \to \infty. \tag{2A}$$

The proof of Theorem 1.4 uses Fourier analysis tools among other things. This method turns out to be quite effective in dealing with bounding character sums on arithmetic progressions, and was introduced by Friedlander and Iwaniec [4]. Theorem 1.4 shows that, for example in the range $A \ll N^{1/(2r)} \log N$, the asymptotic formula holds in the range $AR^{1/2} \gg N^{1/4+\epsilon}$ (when r is large in condition (1C)). This is essentially the nontrivial bound range from the Burgess method. We should also remark that one could obtain a comparable result using a mean value Burgess estimate recently obtained by Heath-Brown [6] and Shao [15]. However these results do not provide the logarithm factor for the error terms in Theorem 1.4. And Heath-Brown's mean value estimate gives rise to a slightly weaker asymptotic range $AR^{1/3} \gg N^{1/4+\epsilon}$.

By letting $A = 1, 2, \ldots, m$ successively in Theorem 1.4, we immediately deduce that the quadratic residues (resp. non-residues) are uniformly distributed in the residues classes modulo m.

Corollary 1.5. *Let r, m, R be positive integers and N a large RSA modulus such that $m \ll N^{1/(2r)} \log N$ and $m^2(R+1) < \lambda N^{1+1/(2r)}$, and*

$$\frac{R}{N^{\frac{r+1}{2r}} (\log N)^{2r+1}} \to \infty \quad \text{as } N \to \infty.$$

Then, we have uniformly for $1 \le a \le m$,

$$\lim_{N \to \infty} \frac{|\mathcal{Q}(R; a, m)|}{\sum_{a \le m} |\mathcal{Q}(R; a, m)|} = \frac{1}{m}.$$

The same estimate also holds for $|\mathcal{NQ}(R; a, m)|$.

One may also ask how the cardinalities of the quadratic residues and non-residues are distributed in $\mathbb{Z}/m\mathbb{Z}$. This line of question was investigated by Lamzouri and Zaharescu [11] in the setting of $\mathbb{F}_p$. Let $1 \le k \le N$. Define

$$\mathcal{R}(k) := \{1 \le n \le k : n \text{ is a quadratic residue mod } N\},$$
$$\mathcal{N}(k) := \{1 \le n \le k : n \text{ is a quadratic non-residue mod } N\},$$

and

$$\Psi_{\mathcal{R}}(N;m,a) := \frac{1}{N}\big|\{1 \le k \le N : |\mathcal{R}(k)| \equiv a \bmod m\}\big|,$$
$$\Psi_{\mathcal{N}}(N;m,a) := \frac{1}{N}\big|\{1 \le k \le N : |\mathcal{N}(k)| \equiv a \bmod m\}\big|.$$

Theorem 1.6. *Let N be a large RSA modulus. Then for any integer $m \ge 2$ with $m = o(N^{1/2})$, we have*

$$\sum_{a=0}^{m-1}\Big(\Psi_{\mathcal{Q},(\mathcal{N}\mathcal{Q})}(N;m,a) - \frac{1}{m}\Big)^2 \ll \frac{m}{\log N}.$$

Consequently, if $2 \le m \ll (\log N)^{1/3}$, we have uniformly for all $0 \le a \le m-1$,

$$\Psi(N;m,a)_{\mathcal{Q},(\mathcal{N}\mathcal{Q})} = \frac{1}{m} + O\Big(\sqrt{\frac{m}{\log N}}\Big).$$

Theorem 1.6 shows that in the range $m \ll (\log N)^{1/3}$, the quantities $|\mathcal{R}(k)|, |\mathcal{N}(k)|$ are uniformly distributed in $\mathbb{Z}/m\mathbb{Z}$. During the proof of Theorem 1.6, we are led to consider a slightly different probabilistic model than the one used by Lamzouri and Zaharescu, namely: $X_1 + \cdots + X_H \pmod m$ where $X_1, \dots, X_H$ are independent random variables taking on the values 1 and 0 with the corresponding probability δ and $1-\delta$, respectively, where $0 < \delta < 1$. The lack of symmetry in the model creates a number of technical difficulties including several weighted arithmetic sums which do not exist in the symmetric setting. Otherwise, we are more or less able to follow the method of [11] in deducing Theorem 1.6.

The rest of the paper is organized as follows. Section 2 contains several technical lemmas that are used throughout the paper. Theorem 1.4 is proved in Section 3. Section 4 provides bounds involving the quadratic residue characteristic function. Section 5 studies the probabilistic model $X_1 + \cdots + X_H$. Section 6 studies the probabilistic model $X_1 + \cdots + X_H \pmod m$. Sections 7 and 9 provide links for the probabilistic models introduced earlier and the main theorems. Theorems 1.1 and 1.6 are proved in Sections 8 and 10, respectively.

Convention. The phrase "large RSA modulus" precisely means $N = pq$ with $1 < p \ne q \le cN^{1/2}$ where $c > 0$, and furthermore $N \to \infty$.

2 Technical lemmas

The notation $e_m(x) = \exp(\frac{2\pi i x}{m})$ is used throughout the paper. We also use $(\nu_1, \dots, \nu_k) \in \{0,1\}^k$ to indicate that a binary vector $\nu = (\nu_1, \dots, \nu_k)$ is sampled from $\{0,1\}^k$.

Lemma 2.1. *Let k, r be positive integers such that $1 \le k \le r$. The number of positive integer solutions of the linear equation $x_1 + x_2 + \cdots + x_k = r$ is $\binom{r-1}{k-1}$.*

Proof. We proceed by induction on k. The case $k = 1$ is clear. Let $k \le r' + 1$. Assume now the linear equation $x_1 + x_2 + \cdots + x_{k-1} = r'$ has $\binom{r'-1}{k-2}$ positive integer solutions. Applying the induction hypothesis, we get

$$\sum_{x_1+\cdots+x_k=r} 1 = \sum_{x_k=1}^{r-k+1} \sum_{x_1+\cdots+x_{k-1}=r-x_k} 1 = \sum_{x_k=1}^{r-k+1}\binom{r-x_k-1}{k-2} = \sum_{d=k-2}^{r-2}\binom{d}{k-2} = \binom{r-1}{k-1}. \qquad \square$$

Lemma 2.2. *Let $x > 0$ be a real number, and r a positive integer. Then*

$$\sum_{k=1}^{r}\binom{r}{k}\binom{r-1}{k-1}x^k = x^r + O(4^r).$$

Proof. Using the identity $\binom{r}{k} = \frac{r}{k}\binom{r-1}{k-1}$, we have

$$\sum_{k=1}^{r}\binom{r}{k}\binom{r-1}{k-1}x^k = \sum_{k=1}^{r}\frac{k}{r}\binom{r}{k}^2 x^k = \sum_{k=0}^{r-1}\frac{r-k}{r}\binom{r}{r-k}^2 x^{r-k}$$
$$= x^r\sum_{k=0}^{r-1}\Big(1-\frac{k}{r}\Big)\binom{r}{k}^2 x^{-k} = x^r\Big(1+\sum_{k=1}^{r-1}\Big(1-\frac{k}{r}\Big)\binom{r}{k}^2 x^{-k}\Big).$$

The sum over k can be bounded as follows:

$$\sum_{k=1}^{r-1}\Big(1-\frac{k}{r}\Big)\binom{r}{k}^2 x^{-k} < \sum_{k=1}^{r}\binom{r}{k}^2 x^{-k} \ll x^{-r}\binom{2r}{r} \le x^{-r}4^r. \qquad \square$$

Lemma 2.3. *Let real numbers $t, \lambda \neq 0$ be given. Then we have*

$$\sum_{(\nu_1,\dots,\nu_k)\in\{0,1\}^k}\lambda^{\nu_1+\cdots+\nu_k}e_m\big(t(\nu_1+\cdots+\nu_k)\big) = \big(1+\lambda e_m(t)\big)^k.$$

Proof. We split the sum into residue classes modulo m according to the values $\nu_1+\cdots+\nu_k \bmod m$. Thus

$$\sum_{(\nu_1,\dots,\nu_k)\in\{0,1\}^k}\lambda^{\nu_1+\cdots+\nu_k}e_m\big(t(\nu_1+\cdots+\nu_k)\big) = \sum_{b=0}^{m-1}e_m(tb)\sum_{\substack{(\nu_1,\dots,\nu_k)\in\{0,1\}^k\\ \nu_1+\cdots+\nu_k\equiv b \bmod m}}\lambda^{\nu_1+\cdots+\nu_k}.$$

Note that for $k \ge n$, we have $\nu_1+\cdots+\nu_k = n$ if and only if among $\nu_1,\dots,\nu_k$, n of them have the value 1 and the rest $(k-n)$ variables have the value 0. Furthermore, the number of binary vectors $(\nu_1,\dots,\nu_k)$ which have n "1" coordinates is $\binom{k}{n}$. Hence, the sum becomes

$$\sum_{b=0}^{m-1}e_m(tb)\sum_{l\ge 0}\binom{k}{b+lm}\lambda^{b+lm} = \sum_{b=0}^{m-1}\sum_{l\ge 0}\binom{k}{b+lm}e_m(t(b+lm))\lambda^{b+lm} = \big(1+\lambda e_m(t)\big)^k. \qquad \square$$

Lemma 2.4. *Let positive integers k, a, m be given such that $0 \le a < m \le k$. Let x be a real number. Define*

$$S(x;a,m,k) := \sum_{\substack{n\equiv a \bmod m\\ n\le k}}\binom{k}{n}x^n = \sum_{l\ge 0}\binom{k}{lm+a}x^{lm+a}.$$

Then we have

$$S(x;a,m,k) = \begin{cases}\frac{1}{m}\sum_{l=0}^{m-1}e_m(-al)(1+xe_m(l))^k, & \text{if } x\neq 0,\\ 0, & \text{if } x = 0.\end{cases}$$

Proof. The evaluation of $S(1;a,m,k)$ is a classical identity due to Ramus [13]. The general case is treated in [8, Theorem 1]. $\square$

Lemma 2.5. *Let real number $\lambda \ge 0$, and positive integers $m \ge 3, k, t$ be given. Then*

$$\big(1+\lambda e_m(t)\big)^k \begin{cases} = 1, & \text{if } \lambda = 0,\\ = (1+\lambda)^k, & \text{if } t \equiv 0 \bmod m,\\ \ll (1+\lambda)^k\exp\big(-\frac{2\pi^2\lambda k}{3(\lambda+1)m^2}\big), & \text{if } \lambda, t \neq 0.\end{cases}$$

Proof. The cases $\lambda = 0$ and $t \equiv 0$ are trivial. In the remaining case, we make a change of variable $\lambda \mapsto \frac{\delta}{1-\delta}$ with $0 < \delta < 1$. This is permissible because the function $f(x) = \frac{x}{1-x}$ is strictly increasing and maps the interval $(0,1)$ onto $(0,\infty)$. Therefore for $t \equiv 1, 2, \dots, m-1$,

$$\Big|1+\frac{\delta}{1-\delta}e_m(t)\Big| = (1-\delta)^{-1}\big|1-\delta+\delta e_m(t)\big| = (1-\delta)^{-1}\big|1-2\delta+\delta(1+e_m(t)\big|$$
$$= (1-\delta)^{-1}\Big|1-2\delta+\delta e_m\Big(\frac{t}{2}\Big)\Big(e_m\Big(\frac{-t}{2}\Big)+e_m\Big(\frac{t}{2}\Big)\Big)\Big|$$
$$\le (1-\delta)^{-1}\Big(1-2\delta+2\delta\cos\Big(\frac{\pi}{m}\Big)\Big)$$
$$\le (1-\delta)^{-1}\Big(1-2\delta+2\delta\Big(1-\frac{\pi^2}{3m^2}\Big)\Big) = (1-\delta)^{-1}\Big(1-\frac{2\pi^2\delta}{3m^2}\Big),$$

where we have used in the second last step the inequality $\cos(x) \le 1 - x^2/3$ for $0 \le x \le \pi/2$. Thus for $m \ge 3$, we have

$$\begin{aligned}\left|1 + \frac{\delta}{1-\delta} e_m(t)\right|^k &\le (1-\delta)^{-k}\Big(1 - \frac{2\pi^2\delta}{3m^2}\Big)^k \\ &= (1-\delta)^{-k} \exp\Big(k \log\Big(1 - \frac{2\pi^2\delta}{3m^2}\Big)\Big) \le (1-\delta)^{-k} \exp\Big(-\frac{2\pi^2 k\delta}{3m^2}\Big).\end{aligned}$$

Changing δ back to λ via the transformation $\delta \mapsto \frac{\lambda}{1+\lambda}$ proves the lemma. □

Lemma 2.6. *Let m, N be positive integers such that $m = o(N^{1/2})$. Let $0 < \delta < 6/\pi^2 \approx 0.6079$. Given any $\epsilon > 0$, we have for large N*

$$\sum_{t=1}^{m-1}\sum_{n=1}^{N-1} (N-n)\big(1 - \delta + \delta e_m(t)\big)^n \ll_{\delta,\epsilon} N^{2-\epsilon} m^{2\epsilon}.$$

Proof. We have

$$\begin{aligned}\sum_{t=1}^{m-1}\sum_{n=1}^{N-1} (N-n)\big(1 - \delta + \delta e_m(t)\big)^n &= \sum_{t=1}^{m-1}\sum_{n=1}^{N-1} (N-n)\Big(1 - 2\delta + \delta e_m\Big(\frac{t}{2}\Big)\Big(e_m\Big(\frac{-t}{2}\Big) + e_m\Big(\frac{t}{2}\Big)\Big)\Big)^n \\ &\ll \sum_{1\le t\le m/2}\sum_{n=1}^{N-1} (N-n)\Big(1 - 2\delta + 2\delta\Big|\cos\Big(\frac{\pi t}{m}\Big)\Big|\Big)^n.\end{aligned}$$

Now using the inequality $\cos(x) \le 1 - x^2/3$, $0 \le x \le \pi/2$ gives the upper bound

$$\begin{aligned}\sum_{1\le t\le m/2}\sum_{n=1}^{N-1} (N-n)\Big(1 - 2\delta + 2\delta\Big(1 - \frac{\pi^2 t^2}{3m^2}\Big)\Big)^n &\le N \sum_{1\le t\le m/2}\sum_{n=1}^{N-1}\Big(1 - \frac{2\delta\pi^2 t^2}{3m^2}\Big)^n \\ &= N \sum_{1\le t\le m/2}\sum_{n=1}^{N-1} \exp\Big(n\log\Big(1 - \frac{2\delta\pi^2 t^2}{3m^2}\Big)\Big) \\ &\le N \sum_{1\le t\le m/2}\sum_{n=1}^{N-1} \exp\Big(-\frac{2n\delta\pi^2 t^2}{3m^2}\Big) \\ &\ll N \sum_{n=1}^{N-1}\int_1^{m/2} \exp\Big(-\frac{2n\delta\pi^2 t^2}{3m^2}\Big)dt \\ &\ll_\delta N \sum_{n=1}^{N-1} \frac{m^2}{n}\exp\Big(-\frac{2n\delta\pi^2}{3m^2}\Big) \\ &\ll_\epsilon N \sum_{n=1}^{N-1}\Big(\frac{m^2}{n}\Big)^\epsilon \exp\Big(-\frac{n\delta}{m^2}\Big) \ll_\epsilon N^{2-\epsilon} m^{2\epsilon}.\end{aligned}$$

□

Lemma 2.7. *Let $b, A \ge 2$ be positive integers. Let f be the cut-off function defined as*

$$f(x) = \begin{cases} \min(x, 1, A+1-x), & \text{if } 0 \le x \le A+1, \\ 0, & \text{otherwise.}\end{cases}$$

Denote the Fourier transform of f by $\hat{f}$. Then we have

$$\int_{-\infty}^{\infty} |\hat{f}(t)|dt = \frac{4}{\pi^2}\log A + O(1), \qquad \int_{-\infty}^{\infty} |\hat{f}(t/b)|\frac{dt}{b} = \frac{4}{\pi^2}\log A + O(1).$$

Proof. We shall prove only the first estimate. The proof of the second estimate is similar. It is known that $|\hat{f}(t)| = (\pi t)^{-2}|\sin(\pi t)\sin(\pi A t)|$, which gives

$$\int_{-\infty}^{\infty} |\hat{f}(t)|dt = \int_{-\infty}^{\infty} \frac{|\sin(\pi t)\sin(\pi A t)|}{(\pi t)^2}dt = 2\int_0^{\infty} \frac{|\sin(\pi t)\sin(\pi A t)|}{(\pi t)^2}dt.$$

We next break the integral into two sub-intervals $[0,1]$ and $[1,\infty]$. Thus,

$$\int_0^\infty \frac{|\sin(\pi t)\sin(\pi At)|}{(\pi t)^2}dt = \int_0^1 \frac{|\sin(\pi t)\sin(\pi At)|}{(\pi t)^2}dt + \mathrm{O}(1) = \int_0^1 \frac{|\sin(\pi At)|}{\pi t}dt + \mathrm{O}(1)$$

after replacing $\sin(\pi t)/(\pi t) = 1 + \mathrm{O}(t^2)$ in the integrand above. The last integral is

$$\int_0^1 \frac{|\sin(\pi At)|}{t}dt \sim \frac{2}{\pi}\log A.$$

To see this, we make a change of variable and break the new range into subintervals of the type $[(k-1)\pi, k\pi]$. We get

$$I(A) := \int_0^1 \frac{|\sin(\pi At)|}{t}dt = \int_0^{\pi A} \frac{|\sin(t)|}{t}dt = \sum_{k=1}^{A}\int_{(k-1)\pi}^{k\pi} \frac{|\sin(t)|}{t}dt.$$

The upper-bound can be achieved as follows:

$$\begin{aligned} I(A) &= \int_0^\pi \frac{|\sin(t)|}{t}dt + \sum_{k=2}^{A}\int_{(k-1)\pi}^{k\pi} \frac{|\sin(t)|}{t}dt \\ &\le C + \sum_{k=2}^{A}\frac{1}{(k-1)\pi}\int_{(k-1)\pi}^{k\pi}|\sin(t)|dt = C + \frac{2}{\pi}\sum_{k=1}^{A-1}\frac{1}{k} \le \frac{2}{\pi}\log A + \mathrm{O}(1). \end{aligned}$$

For the lower bound we have

$$I(A) \ge \sum_{k=1}^{A}\frac{1}{k\pi}\int_{(k-1)\pi}^{k\pi}|\sin(t)|dt = \frac{2}{\pi}\sum_{k=1}^{A}\frac{1}{k} = \frac{2}{\pi}\log A + \mathrm{O}(1). \qquad \square$$

Lemma 2.8. *Let p a large prime number. Let a, m, R be positive integers such that* $\gcd(p, am) = 1$. *Define the set*

$$S := \{1 \le r \le R : a + rm \equiv 0 \pmod p\}.$$

Then we have

$$|S| = \frac{R}{p} + \mathrm{O}(\log p).$$

Proof. Recall the orthogonal relation

$$\frac{1}{p}\sum_{t=0}^{p-1} e_p(tn) = \begin{cases} 1, & \text{if } n \equiv 0 \bmod p, \\ 0, & \text{otherwise.} \end{cases}$$

Thus, we have

$$|S| = \frac{1}{p}\sum_{r\le R}\sum_{l=0}^{p-1} e_p(l(a+rm)) = \frac{1}{p}\sum_{r\le R}\sum_{l=0}^{p-1} e_p(lrm).$$

The double sum above has a contribution R/p when $l = 0$. In the remaining range $1 \le l \le p-1$, we have

$$\sum_{r\le R}\sum_{l=1}^{p-1} e_p(lrm) = \sum_{l=1}^{p-1}\frac{e_p(lm(R+1))}{(1-e_p(lm))} \ll \sum_{l=1}^{p-1}\frac{1}{\left|\sin\frac{\pi lm}{p}\right|}.$$

Now $\{lm\}$, $l = 1, \dots, p-1$, runs through the non-zero residue classes of $\mathbb{Z}/p\mathbb{Z}$ because $\gcd(p, m) = 1$. We thus have

$$|S| \ll \frac{1}{p}\sum_{l=1}^{p-1}\frac{1}{\left|\sin\frac{\pi l}{p}\right|} \ll \log p.$$

The second bound is well known, see for instance the chapter on the Pólya–Vinogradov inequality in [2]. $\square$

Lemma 2.9. *Let N be a large RSA modulus. Suppose $\lambda > 0$ is a real number such that $\min(p,q) > \lambda N^{1/2}$. Let A, R, m be positive integers such that $A \le m \le \lambda N^{1/2}$. Define the set E*

$$E := \{(a,r) : 1 \le a \le A,\ 1 \le r \le R,\ \gcd(N, a+rm) > 1\}.$$

Then we have

$$|E| \le \frac{AR}{p} + \frac{AR}{q} + O(A \log N).$$

Proof. Notice that $\gcd(N, a+rm) > 1$ if and only if $p|(a+rm)$ or $q|(a+rm)$. Consider the sets

$$E_{a,l} := \{1 \le r \le R : a + rm \equiv 0 \pmod{l}\}$$

with $1 \le a \le A$ and $l \in \{p,q\}$. In view of Lemma 2.8, we have $|E_{a,l}| = R/l + O(\log l)$, and

$$E = \bigcup_{a=1}^{A} \bigcup_{l=p,q} E_{a,l}.$$

Since $A \le m \le \lambda N^{1/2}$, we have $E_{a_1,l} \cap E_{a_2,l} = \varnothing$ for $a_1 \ne a_2$. Consequently, any intersection of more than two distinct of these sets is empty. Furthermore, since

$$E_{a_1,p} \cap E_{a_2,q} = \{r : r \equiv -a_1/m \pmod{p} \text{ and } r \equiv -a_2/m \pmod{q}\},$$

we have $|E_{a_1,p} \cap E_{a_2,q}| \le 1$ for $1 \le a_1, a_2 \le A$. By the inclusion-exclusion principle, it follows that

$$|E| = \sum_{a=1}^{A} \sum_{l \in \{p,q\}} |E_{a,l}| - \sum_{a_1=1}^{A} \sum_{a_2=1}^{A} |E_{a_1,p} \cap E_{a_2,q}| \le \sum_{a=1}^{A} \sum_{l \in \{p,q\}} |E_{a,l}| = \frac{AR}{p} + \frac{AR}{q} + O(A \log N). \qquad \square$$

Lemma 2.10. *Let $M, T \ge 1, L \ge 1$ be integers. Suppose $LT \ll N$ for large N. Then the cardinality of the set*

$$\{(a_1, a_2, \lambda_1, \lambda_2) : a_1\lambda_2 \equiv a_2\lambda_1 \pmod{N},\ M < a_1, a_2 \le M + T,\ 1 \le \lambda_1, \lambda_2 \le L\}$$

has a bound $\ll LT \log L$.

Proof. See [7, Lemma 12.7]. $\square$

Lemma 2.11. *Let A, R, m, N be positive integers such that $A \le m$ and $m(R+1) < N$. Define the set*

$$S := \{(a_1, a_2, r_1, r_2) : a_1 + r_1 m \equiv a_2 + r_2 m \pmod{N},\ -A < a_1, a_2 \le A,\ 1 \le r_1, r_2 \le R\}.$$

Then we have

$$|S| = \begin{cases} 2RA, & \text{if } 1 \le A \le \frac{m}{2}, \\ 2RA + 2(R-1)(2A-m), & \text{if } \frac{m}{2} < A \le m. \end{cases}$$

Proof. Since $1 \le a_i + r_i m \le (R+1)m < N$, the congruence $a_1 + r_1 m \equiv a_2 + r_2 m \pmod{N}$ is equivalent to the equality

$$a_1 - a_2 = (r_2 - r_1)m. \tag{2.1}$$

If $r_1 = r_2$, the equality (2.1) implies $a_1 = a_2$. Thus there are $2RA$ solutions satisfying (2.1) in this case. On the other hand, if $|r_1 - r_2| \ge 2$, then (2.1) has no solutions because $|a_1 - a_2| < 2m$ in the range $-A < a_1, a_2 \le A$. In the last case, when $|r_1 - r_2| = 1$, we have

$$\left|\{(a_1, a_2) : |a_1 - a_2| = m,\ -A < a_1, a_2 \le A\}\right| = \begin{cases} 0, & \text{if } 1 \le A \le \frac{m}{2}, \\ 2(2A-m), & \text{if } \frac{m}{2} < A \le m. \end{cases}$$

Furthermore, there are $2(R-1)$ pairs (r_1, r_2) such that $|r_1 - r_2| = 1$. $\square$

Lemma 2.12. *Let N be a large RSA modulus. Suppose $\lambda > 0$ is a real number such that $\min(p,q) > \lambda N^{1/2}$. Let A, B, R, m be positive integers such that $A \le m$, $B \le \lambda N^{1/2}$, and $m(R+1)B < N$. Define the set*

$$T := \Big\{(a_1,a_2,b_1,b_2,r_1,r_2) : \frac{a_1+r_1m}{b_1} \equiv \frac{a_2+r_2m}{b_2} \pmod N, \\ -A < a_1,a_2 \le A,\ 1 \le b_1,b_2 \le B,\ 1 \le r_1,r_2 \le R\Big\}.$$

Then we have $|T| \ll \min(ARB\log N + ARB^2, mRB\log N)$.

Proof. First, notice that the elements of B are invertible because of the condition $B \le \lambda N^{1/2}$. Furthermore, we have $1 \le a_i + r_i m \le (R+1)m < N$. Hence given any pair of (λ_1,λ_2) with $1 \le \lambda_1,\lambda_2 \le (R+1)m$, the system of equations $a_i + r_i m = \lambda_i$ with $i = 1,2$ has at most one solution in (a_1,a_2,r_1,r_2). Therefore in view of Lemma 2.10, $|T|$ is at most

$$\big|\{(\lambda_1,\lambda_2,b_1,b_2) : \lambda_1 b_2 \equiv \lambda_2 b_1 \pmod N,\ 1 \le \lambda_1,\lambda_2 \le m(R+1),\ 1 \le b_1,b_2 \le B\}\big| \ll mRB\log N.$$

On the other hand, the relation $(a_1+r_1m)b_2 \equiv (a_2+r_2m)b_1$ implies that $a_1 + r_1m \equiv a_2 + r_2m$ if and only if $b_1 \equiv b_2$. Therefore in view of Lemma 2.11, we have $|T| \ll ARB$ in the case $a_1 + r_1m \equiv a_2 + r_2m$. In the case $a_1 + r_1m \not\equiv a_2 + r_2m$ and $r_1 = r_2 = r$, we have that both $a_1 + r_1m$ and $a_2 + r_2m$ belong to the interval $(-A+rm, A+rm]$. Thus for a fixed r, the number of quadruples (a_1,a_2,b_1,b_2) such that $(a_1+rm)b_2 \equiv (a_2+rm)b_1$ is bounded by

$$\big|\{(\lambda_1,\lambda_2,b_1,b_2) : b_2\lambda_1 \equiv b_1\lambda_2 \pmod N,\ -A+rm < \lambda_1,\lambda_2 \le A+rm,\ 1 \le b_1,b_2 \le B\}\big| \ll AB\log N.$$

Therefore, $|T| \ll ARB\log N$ in this case. The last scenario when $a_1 + r_1m \not\equiv a_2 + r_2m$ and $r_1 \not\equiv r_2$. For a fixed triple (a_1,r_1,b_1), the number of triples (a_2,r_2,b_2) such that $(a_1+r_1m)b_2 \equiv (a_2+r_2m)b_1$ is at most B. To see this, notice that the set

$$\{a_2 + r_2m : -A < a_2 \le A,\ 1 \le r_2 \le R,\ r_2 \ne r_1\} = \bigcup_{r\ne r_1}(-A+rm, A+rm]$$

has at most one intersection with the singleton element $(a_1+r_1m)b_2/b_1$. This means for a fixed b_2, the number of pairs (a_2,r_2) such that $(a_1+r_1m)b_2/b_1 = a_2 + r_2m$ is at most one. Finally, $|T| \ll ARB^2$ in this case. □

Lemma 2.13. *Let $\chi \pmod N$ be a primitive character of conductor N of order h. Let $f(x) \in \mathbb{Z}[x]$ be a polynomial written as*

$$f(x) = \prod_{k=1}^{s}(x+a_k)^{d_k},$$

where d_k are positive integers and a_k are any integers. Define

$$\Delta = \prod_{i\ne j}\prod(a_i - a_j).$$

Suppose the condition

$$(d_1,\dots,d_s,h) = 1$$

is satisfied, then we have

$$\Big|\sum_{x \pmod N}\chi(f(x))\Big| \le (s-1)^{\omega(N)}(\Delta,N)^{1/2}N^{1/2}.$$

where $\omega(N)$ is the number of divisors of N.

Proof. See [7, Corollary 12.12]. □

Lemma 2.14. *Let N be a large RSA modulus, χ_N the associated quadratic character (Jacobi symbol). Suppose $\lambda > 0$ is a real number such that $\min(p,q) > \lambda N^{1/2}$. Let r be a positive integer and $t \in \mathbb{R}$. Then we have for $C \le \lambda N^{1/2}$,*

$$\sum_{u \bmod N}\Big|\sum_{c\le C} e(ct)\chi_N(u+c)\Big|^{2r} \le NC^r + (2r-1)^2N^{1/2}C^{2r}.$$

Proof. We have

$$\sum_{u \bmod N} \Big|\sum_{c\le C} e(ct)\chi_N(u+c)\Big|^{2r} = \sum_{u \bmod N} \Big(\sum_{1\le c_1,c_2\le C} e\big((c_1-c_2)t\big)\chi_N(u+c_1)\chi_N(u+c_2)\Big)^r.$$

Clearly, if $c_1 = c_2$, the contribution of the above sum is NC^r. When $c_1 \neq c_2$, the above sum is at most

$$\sum_{\substack{1\le c_1^1,\dots,c_1^r\le C\\ (c_1^1,\dots,c_1^r)\neq(c_2^1,\dots,c_2^r)}} \sum_{1\le c_2^1,\dots,c_2^r\le C} \Big|\sum_{u \bmod N} \chi_N\Big(\prod_{i=1}^{r}(u+c_1^i)(u+c_2^i)\Big)\Big| \le (2r-1)^2 N^{1/2} C^{2r}$$

by the virtue of Lemma 2.13. Notice that $(\Delta, N) = 1$ because $|c^i_{\{1,2\}} - c^j_{\{1,2\}}| < \min(p,q)$. Moreover, the condition $(d_1,\dots,d_s,2) = 1$ in Lemma 2.13 is satisfied because there exists at least one $d_i = 1$. □

3 Proof of Theorem 1.4

Let n be a positive integer such that $(n,N) = 1$, where $N = pq$ is an RSA modulus. Recall n is a quadratic residue modulo N if and only if $\chi_p(n) = \chi_q(n) = 1$. Thus

$$\frac{(1+\chi_p(n))(1+\chi_q(n))}{4} = \begin{cases} 1, & \text{if } n \text{ is a quadratic residue modulo } N,\\ 0, & \text{if } n \text{ is a quadratic non-residue modulo } N.\end{cases}$$

Therefore, we have in view of Lemma 2.9

$$\begin{aligned}\sum_{a\le A}|\mathcal{Q}(R;a,m)| &= \frac{1}{4}\sum_{a\le A}\sum_{r\le R}\big(1+\chi_p(a+rm)\big)\big(1+\chi_q(a+rm)\big) + \mathrm{O}\Big(\frac{AR}{\min(p,q)} + A\log N\Big)\\ &= \frac{1}{4}RA + \frac{1}{4}\big(S_p(A,R) + S_q(A,R) + S_N(A,R)\big) + \mathrm{O}\big(ARN^{-1/2} + A\log N\big), \end{aligned} \tag{3.1}$$

where the error term comes from counting those terms $\{a+rm\}_{a,r}$ such that $\gcd(a+rm,N) > 1$. The character sums

$$S_{\{p,q,N\}}(A,R) = \sum_{a\le A}\sum_{r\le R}\chi_{\{p,q,N\}}(a+rm)$$

require a non-trivial bound. We bound below only $S_N(A,R)$, the treatments for $S_{\{p,q\}}(A,R)$ are similar. Using the cut-off function introduced in Lemma 2.7, we have

$$S_N(A,R) = \frac{1}{A}\sum_{-A<a\le A}\sum_{r\le R}\sum_{d\le A} f(a+d)\chi_N(a+d+rm). \tag{3.2}$$

Let $A = BC$, with $B \ll N^{1/2-1/(4r)}$ and $C \le \lambda N^{1/(2r)}$ where λ is as defined in Lemma 2.14. The precise values of the parameters B, C will be chosen later. We use shifts of the type $d = bc$ with $1 \le b \le B$, $1 \le c \le C$. Averaging (3.1) over b, c gives

$$\begin{aligned} S_N(A,R) &= \frac{1}{A}\sum_{-A<a\le A}\sum_{r\le R}\sum_{b\le B}\sum_{c\le C} f(a+bc)\chi_N(a+bc+rm)\\ &\le \frac{1}{A}\sum_{-A<a\le A}\sum_{r\le R}\sum_{b\le B}\int_{-\infty}^{\infty}|\hat{f}(y/b)|\frac{dy}{b}\Big|\sum_{c\le C} e(ct)\chi_N(a+bc+rm)\Big| \end{aligned}$$

for some $t \in \mathbb{R}$. We may bound the Fourier integral by Lemma 2.7, and then use Hölder's inequality to obtain

$$\begin{aligned} S_N(A,R) &\ll \frac{\log N}{A}\sum_{-A<a\le A}\sum_{r\le R}\sum_{b\le B}\Big|\sum_{c\le C} e(ct)\chi_N(a/b+c+rm/b)\Big|\\ &= \frac{\log N}{A}\sum_{u \bmod N}\lambda(u)\Big|\sum_{c\le C} e(ct)\chi_N(u+c)\Big|\\ &\le \frac{\log N}{A}\Big(\sum_u \lambda(u)\Big)^{1-1/r}\Big(\sum_u \lambda(u)^2\Big)^{1/(2r)}\Big(\sum_u\Big|\sum_{c\le C} e(ct)\chi_N(u+c)\Big|^{2r}\Big)^{1/(2r)}, \end{aligned}$$

where $\lambda(u)$ counts the frequency of values in arithmetic progressions which is represented by a residue class modulo N, i.e.

$$\lambda(u) = \#\{(a,b,r) : (a+rm)/b \equiv u \,(\mathrm{mod}\, N),\ -A < a \le A,\ 1 \le b \le B,\ 1 \le r \le R\}.$$

Clearly,

$$\sum_{u \bmod N} \lambda(u) \le 2ABR \quad \text{and} \quad \sum_{u} \lambda(u)^2 = |T|,$$

where the set T is as defined in Lemma 2.12. Therefore in view of Lemma 2.12 and Lemma 2.14, we have that

$$S_N(A,R) \ll_r \begin{cases} \frac{\log N}{A}(ABR)^{1-1/r}(ARB\log N)^{1/(2r)}(NC^r + N^{1/2}C^{2r})^{1/(2r)}, & \text{if } B \ll \log N, \\ \frac{\log N}{A}(ABR)^{1-1/r}(ARB^2\log N)^{1/(2r)}(NC^r + N^{1/2}C^{2r})^{1/(2r)}, & \text{otherwise.} \end{cases}$$

Choose $B = \lambda^{-1}AN^{-1/(2r)}$ and $C = \lambda N^{1/(2r)}$ giving the error terms as purported in Theorem 1.4.

Finally, notice that in the range $B \ll \log N$ (i.e. $A \ll N^{1/(2r)}\log N$)

$$ARN^{-1/2} \ll A^{1-\frac{1}{r}}R^{1-\frac{1}{2r}}N^{\frac{r+1}{4r^2}}(\log N)^{1+\frac{1}{2r}}$$

is equivalent to the statement

$$A^2R \ll N^{r+\frac{1}{2}+\frac{1}{2r}}(\log N)^{2r+1}.$$

Now, the conditions $mA(R+1) < \lambda N^{1+1/(2r)}$ and $A \le m \le \lambda N^{1/2}$ imply that

$$A^2R \le mAR \ll N^{\frac{1}{2}}N^{1+\frac{1}{2r}} \ll N^{\frac{3}{2}+\frac{1}{2r}} \ll N^{r+\frac{1}{2}+\frac{1}{2r}}(\log N)^{2r+1}.$$

Furthermore, the conditions

$$R \gg N^{\frac{r-1}{2r(2r-1)}}(\log N)^{-\frac{1}{2r-1}} \quad \text{and} \quad A \le m \ll N^{\frac{1}{2}}$$

imply that

$$A\log N \ll A^{1-\frac{1}{r}}R^{1-\frac{1}{2r}}N^{\frac{r+1}{4r^2}}(\log N)^{1+\frac{1}{2r}}.$$

This proves the theorem in view of (3.1) for $\sum|\mathcal{Q}|$ in the range $A \ll N^{1/(2r)}\log N$. In the remaining range, a similar argument gives

$$ARN^{-1/2} + A\log N \ll A^{1-\frac{1}{2r}}R^{1-\frac{1}{2r}}N^{\frac{1}{4r}}(\log N)^{1+\frac{1}{2r}}.$$

We omit the details here. Finally, notice that $\sum|\mathcal{NQ}|$ follows because of the relation

$$\sum_{a\le A}|\mathcal{Q}(R;a,m)| + \sum_{a\le A}|\mathcal{NQ}(R;a,m)| + \mathrm{O}(ARN^{-1/2} + A\log N) = AR.$$

4 A character sum involving χ_p and χ_q

Recall the quadratic residue characteristic function

$$\Phi(n) := \frac{1}{4}\big(1+\chi_p(n)\big)\big(1+\chi_q(n)\big).$$

Proposition 4.1. *Let $N = pq$ be a large RSA modulus. Let T be a non-empty subset of the set $\{0,1,\dots,N\}$ such that*

$$\gcd\Big(\prod_{\substack{i,j\in T\\ i\ne j}}\prod(i-j), N\Big) = 1.$$

Then in the range $1 \le |T| < \frac{1}{4}\log_2 N$, we have

$$\sum_{n=1}^{N}\prod_{i\in T}\Phi(n+i)^{d_i} = \frac{N}{4^{|T|}} + \mathrm{O}\Big(\frac{|T|N^{3/4}}{2^{|T|}}\Big), \tag{4.1}$$

where $d_i, i \in T$ are any positive integer weights.

Proof. Assume without loss of generality $d_i = 1$ since $\Phi(n+i)^{d_i} = \Phi(n+i)$ as long as $\gcd(n+i, N) = 1$, and the number $n \le N$ such that the sequence $\{\Phi(n+i)\}_{i\in T}$ contains a term that is not co-prime with N being bounded by $|T|N^{1/2}$ (see [9, Lemma 3.2]). Therefore,

$$\begin{aligned}\sum_{n=1}^{N}\prod_{i\in T}\Phi(n+i) &= \frac{1}{4^{|T|}}\sum_{n=1}^{N}\prod_{i\in T}(1+\chi_p(n+i))(1+\chi_q(n+i))\\ &= \frac{1}{4^{|T|}}\sum_{n=1}^{N}\Big(1+\sum_{\substack{\Delta\subset T\\ \Delta\neq\emptyset}}\chi_p\Big(\prod_{i\in\Delta}(n+i)\Big)\Big)\Big(1+\sum_{\substack{\Delta\subset T\\ \Delta\neq\emptyset}}\chi_q\Big(\prod_{i\in\Delta}(n+i)\Big)\Big)\\ &= \frac{1}{4^{|T|}}\Big(N+\sum_{\substack{\Delta\subset T\\ \Delta\neq\emptyset}}\sum_{n=1}^{N}\chi_{\{p,q,N\}}\Big(\prod_{i\in\Delta}(n+i)\Big)+\sum_{\substack{\Delta\subset T\\ \Delta\neq\emptyset}}\sum_{\substack{\Delta'\subset T\\ \Delta'\neq\emptyset\\ \Delta\neq\Delta'}}S(\Delta,\Delta')\Big),\end{aligned} \tag{4.2}$$

where

$$S(\Delta,\Delta') := \sum_{n=1}^{N}\chi_p\Big(\prod_{i\in\Delta}(n+i)\Big)\chi_q\Big(\prod_{i\in\Delta'}(n+i)\Big).$$

In (4.2), the character sums involving χ_p, χ_q can be bounded using the Weil bound and the sum involving χ_N bounded by Lemma 2.13. We have

$$\sum_{\substack{\Delta\subset T\\ \Delta\neq\emptyset}}\sum_{n=1}^{N}\chi_p\Big(\prod_{i\in\Delta}(n+i)\Big) \le qp^{1/2}|\Delta|2^{|T|} \ll N^{3/4}|T|2^{|T|}, \tag{4.3}$$

$$\sum_{\substack{\Delta\subset T\\ \Delta\neq\emptyset}}\sum_{n=1}^{N}\chi_q\Big(\prod_{i\in\Delta}(n+i)\Big) \le pq^{1/2}|\Delta'|2^{|T|} \ll N^{3/4}|T|2^{|T|}, \tag{4.4}$$

$$\sum_{\substack{\Delta\subset T\\ \Delta\neq\emptyset}}\sum_{n=1}^{N}\chi_N\Big(\prod_{i\in\Delta}(n+i)\Big) \le N^{1/2}|\Delta|^2 2^{|T|} \le N^{1/2}|T|^2 2^{|T|}. \tag{4.5}$$

To bound $S(\Delta,\Delta')$, we write $n = n_1 q + n_2 p$ where n_1 runs over a complete set of residues modulo p and n_2 runs over a complete set of residues modulo q. Therefore, using the notations

$$f_1(x) = \prod_{i\in\Delta}(x+i), \quad f_2(x) = \prod_{i\in\Delta'}(x+i),$$

we have

$$\begin{aligned}S(\Delta,\Delta') &= \sum_{\substack{n_1 \bmod p\\ n_2 \bmod q}}\chi_p(f_1(n_1q+n_2p))\chi_q(f_2(n_1q+n_2p))\\ &= \sum_{\substack{n_1 \bmod p\\ n_2 \bmod q}}\chi_p(f_1(n_1q))\chi_q(f_2(n_2p)) = \sum_{n_1 \bmod p}\chi_p(f_1(n_1))\sum_{n_2 \bmod q}\chi_q(f_2(n_2)) \le |\Delta||\Delta'|p^{1/2}q^{1/2}\end{aligned}$$

after applying the Weil bound to the factored complete character sums above. Finally,

$$\sum_{\substack{\Delta\subset T\\ \Delta\neq\emptyset}}\sum_{\substack{\Delta'\subset T\\ \Delta'\neq\emptyset\\ \Delta\neq\Delta'}}S(\Delta,\Delta') \ll N^{1/2}4^{|T|}|T|^2. \tag{4.6}$$

The proposition is proved in view of (4.2)–(4.6). □

Proposition 4.1 has a few nice consequences. In particular, it implies that in the range $|T| \ll \log N$, the pattern of $(0,1)$ derived from the quadratic residue characteristic function $\Phi(\cdot)$ of length $|T|$ tends to a uniform distribution. The following corollary answers a question previously raised in [9].

Corollary 4.2. *Let N be a large RSA modulus. Let s be a positive integer satisfying $1 \le s < \frac{1}{4}\log_4 N$. For any binary vector $(\nu_0, \dots, \nu_{s-1}) \in \{0,1\}^s$, define the set*

$$\mathcal{D} := \{1 \le n \le N : \Phi(n+i) = \nu_i \text{ for all } 0 \le i \le s-1\}.$$

Then we have

$$|\mathcal{D}| = \frac{N}{4^s} + \mathrm{O}(N^{3/4}\log N).$$

Proof. Similarly to the proof of [9, Theorem 1.1], we have

$$|\mathcal{D}| = \sum_{n=1}^{N}\prod_{i\in T}\Phi(n+i) + \mathrm{O}(N^{1/2}\log N),$$

where the error term comes from counting the number $1 \le n \le N$ such that $\{\Phi(n+i)\}$, $0 \le i \le s-1$, contains a term that is not co-prime with N. The main term above can be estimated by Proposition 4.1. □

5 Probabilistic model

In this section, $X_1, \dots, X_H$ are independent random variables taking on the values 1 and 0 with the corresponding probability δ and $1-\delta$, respectively, where $0 < \delta < 1$. Define Z_H as sum of the random variables:

$$Z_H := X_1 + \cdots + X_H.$$

To ease notations, we denote the normalized random variables as

$$\tilde{t}, (\widetilde{X}, \widetilde{Z_H}) := \frac{t, (X, Z_H)}{\sqrt{H\delta(1-\delta)}}.$$

Lemma 5.1. *Let r be a non-negative integer. Then we have $\mathbb{E}(Z_H^r) = (\delta H)^r$.*

Proof. We have

$$\begin{aligned}\mathbb{E}(Z_H^r) &= \mathbb{E}\big((X_1 + \cdots + X_H)^r\big)\\ &= \mathbb{E}\Big(\sum_{1\le n_1,\dots,n_r\le H} X_{n_1}\cdots X_{n_r}\Big) = \sum_{1\le n_1,\dots,n_r\le H}\mathbb{E}(X_{n_1}\cdots X_{n_r})\\ &= \sum_{1\le n_1,\dots,n_r\le H}\mathbb{E}(X_{n_1})\cdots\mathbb{E}(X_{n_r}) = \delta^r\sum_{1\le n_1,\dots,n_r\le H}1 = (\delta H)^r.\end{aligned}$$

□

We next compute the characteristic function of the normalized random variable $\widetilde{Z_H}$. The following lemma says that $\widetilde{Z_H}$ has the mean $\sqrt{H\delta/(1-\delta)}$ and variance 1. Without normalization, Z_H has the mean $H\delta$ and variance $H\delta(1-\delta)$.

Lemma 5.2. *Let H be large. Then uniformly for $u \le |F(\delta)|^{-1/3}H^{1/6}$, we have*

$$\mathbb{E}\big(\exp(iu\widetilde{Z_H})\big) := \mathbb{E}\Big(\exp\Big(\frac{iuZ_H}{\sqrt{H\delta(1-\delta)}}\Big)\Big) = \exp\Big(-\frac{u^2}{2}\Big)\Big(1 + \mathrm{O}\Big(|F(\delta)|\frac{u^3}{\sqrt{H}}\Big)\Big)\cdot\exp\Big(iu\sqrt{\frac{H\delta}{1-\delta}}\Big),$$

where

$$F(\delta) = \frac{1}{(\delta(1-\delta))^{3/2}}\Big(\frac{\delta^2}{2} - \frac{\delta}{6} - \frac{\delta^3}{3}\Big).$$

Proof. First, for a random variable X taking on the values 1 and 0 with the corresponding probability δ and $1-\delta$, respectively, we have

$$\begin{aligned}\mathbb{E}\big(\exp(iu\widetilde{X})\big) &= \delta\exp(i\tilde{u}) + (1-\delta) = \delta\cos(\tilde{u}) + (1-\delta) + i\delta\sin(\tilde{u})\\ &= 1 - \frac{u^2}{2H(1-\delta)} + i\delta\sin(\tilde{u}) + \mathrm{O}_\delta\Big(\frac{u^4}{H^2}\Big),\end{aligned} \tag{5.1}$$

where we have expanded $\cos(\tilde{u})$ in power series in the last line. On the other hand, since the random variables X_i are independent, we have, using (5.1),

$$\mathbb{E}\big(\exp(iu\widetilde{Z_H})\big) = \mathbb{E}\big(\exp(iu\widetilde{X})\big)^H = \exp\big(H\log(1+z)\big),$$

where

$$z = -\frac{u^2}{2H(1-\delta)} + i\delta\sin(\tilde{u}) + \mathrm{O}_\delta\Big(\frac{u^4}{H^2}\Big).$$

Next using the series expansion $\log(1+z) = z - z^2/2 + z^3/3 + \mathrm{O}(z^4)$ for $|z| < 1$ and $\sin(\tilde{u}) = \tilde{u} + \mathrm{O}(\tilde{u}^3/6)$ gives

$$\begin{aligned}\mathbb{E}\big(\exp(iu\widetilde{Z_H})\big) &= \exp\Big(H\Big(-\frac{u^2}{2H} + i\frac{\delta u}{\sqrt{H\delta(1-\delta)}} + \mathrm{O}\Big(|F(\delta)|\frac{u^3}{H^{3/2}}\Big)\Big)\Big)\\ &= \exp\Big(-\frac{u^2}{2}\Big)\Big(1 + \mathrm{O}\Big(|F(\delta)|\frac{u^3}{\sqrt{H}}\Big)\Big)\cdot\exp\Big(iu\sqrt{\frac{H\delta}{1-\delta}}\Big).\end{aligned}$$

□

6 Probabilistic model modulo m

In this section, we study the probabilistic model Z_H modulo m. Let $X_1, \dots, X_H$ be independent random variables taking on the values 1 and 0 with the corresponding probability δ and $1-\delta$, respectively, where $0 < \delta < 1$. The first lemma shows that Z_H is uniformly distributed in $\mathbb{Z}/m\mathbb{Z}$ for large enough H.

Lemma 6.1. *Let positive integers $a, m \geq 3$ be given such that $0 \leq a < m$. Then $Z_H \bmod m$ approaches uniform distribution on $\mathbb{Z}/m\mathbb{Z}$ after an expected running time H with the condition $H\delta/m^2 \to \infty$.*

Proof. This is proved in [8, Theorem 2]. □

In order to prove Theorem 1.6, we also need to know quantitatively the probability when Z_H is equal to $a \pmod m$. To this end, define

$$\Phi_{\mathrm{rand}}(N; m, a) := \frac{1}{N}\big|\{1 \leq k \leq N : Z_k \equiv a \bmod m\}\big|.$$

We next study the variance of $\Phi_{\mathrm{rand}}(N; m, a) - 1/m$. The following proposition generalizes [11, Proposition 1], which has a symmetric setting (i.e. $\delta = 1/2$).

Proposition 6.2. *Let $m \geq 2$ be a positive integer. Then for $m = \mathrm{o}(N^{1/2})$ we have*

$$\sum_{a=0}^{m-1}\mathbb{E}\Big(\Big(\Phi_{\mathrm{rand}}(N; m, a) - \frac{1}{m}\Big)^2\Big) \ll_\delta \frac{m}{N}.$$

Proof. Similarly to the proof of [11, Proposition 1], we have

$$\begin{aligned}\sum_{a=0}^{m-1}\mathbb{E}\Big(\Big(\Phi_{\mathrm{rand}}(N; m, a) - \frac{1}{m}\Big)^2\Big) &= \sum_{a=0}^{m-1}\sum_{(\nu_1,\dots,\nu_N)\in\{0,1\}^N}\delta^{\nu_1+\cdots+\nu_N}(1-\delta)^{N-(\nu_1+\cdots+\nu_N)}\Big(\frac{1}{N}\sum_{\substack{1\leq j\leq N\\ \nu_1+\cdots+\nu_j\equiv a \bmod m}} 1 - \frac{1}{m}\Big)^2\\ &= \frac{(1-\delta)^N}{(mN)^2}\sum_{a=0}^{m-1}\sum_{(\nu_1,\dots,\nu_N)\in\{0,1\}^N}\Big(\frac{\delta}{1-\delta}\Big)^{\nu_1+\cdots+\nu_N}\Big(m\sum_{\substack{1\leq j\leq N\\ \nu_1+\cdots+\nu_j\equiv a \bmod m}} 1 - N\Big)^2.\end{aligned}$$

Using the orthogonal relation, the inner sum can be rewritten and gives

$$\begin{aligned}&\frac{(1-\delta)^N}{(mN)^2}\sum_{a=0}^{m-1}\sum_{(\nu_1,\dots,\nu_N)\in\{0,1\}^N}\Big(\frac{\delta}{1-\delta}\Big)^{\nu_1+\cdots+\nu_N}\Big(\sum_{j=1}^{N}\sum_{t=0}^{m-1}e_m\big(t(\nu_1+\cdots+\nu_j-a)\big) - N\Big)^2\\ &\quad= \frac{(1-\delta)^N}{(mN)^2}\sum_{(\nu_1,\dots,\nu_N)\in\{0,1\}^N}\Big(\frac{\delta}{1-\delta}\Big)^{\nu_1+\cdots+\nu_N}\sum_{a=0}^{m-1}A(N,m,a),\end{aligned}\tag{6.1}$$

where

$$\begin{aligned}
\sum_{a=0}^{m-1} A(N,m,a) &= \sum_{a=0}^{m-1}\Big(\sum_{j=1}^{N}\sum_{t=0}^{m-1} e_m\big(t(\nu_1+\cdots+\nu_j-a)\big)-N\Big)^2\\
&= \sum_{a=0}^{m-1}\Big(\sum_{j=1}^{N}\sum_{t=1}^{m-1} e_m\big(t(\nu_1+\cdots+\nu_j-a)\big)\Big)^2\\
&= \sum_{a=0}^{m-1}\sum_{1\le r,s\le m-1} e_m\big(a(s-r)\big)\sum_{1\le j,k\le N} e_m\big(r(\nu_1+\cdots+\nu_j)-s(\nu_1+\cdots+\nu_k)\big)\\
&= m\sum_{t=1}^{m-1}\sum_{1\le j,k\le N} e_m\big(t(\nu_1+\cdots+\nu_j)-t(\nu_1+\cdots+\nu_k)\big)\\
&= m(m-1)N + m\sum_{t=1}^{m-1}\sum_{1\le j\ne k\le N} e_m\big(t(\nu_1+\cdots+\nu_j)-t(\nu_1+\cdots+\nu_k)\big)\\
&= m(m-1)N + m\sum_{t=1}^{m-1}\sum_{1< j< k\le N} e_m\big(t(\nu_j+\nu_{j+1}+\cdots+\nu_k)\big)+e_m\big(-t(\nu_j+\nu_{j+1}+\cdots+\nu_k)\big). \qquad (6.2)
\end{aligned}$$

Thus in view of (6.1) and (6.2), we have

$$\sum_{a=0}^{m-1}\mathbb{E}\Big(\Big(\Phi_{\mathrm{rand}}(N;m,a)-\frac{1}{m}\Big)^2\Big) = I + II. \qquad (6.3)$$

Using Lemma 2.3, the first term I is

$$\begin{aligned}
I &= \frac{m(m-1)N(1-\delta)^N}{(mN)^2}\sum_{(\nu_1,\dots,\nu_N)\in\{0,1\}^N}\Big(\frac{\delta}{1-\delta}\Big)^{\nu_1+\cdots+\nu_N}\\
&= \frac{(m-1)(1-\delta)^N}{mN}\Big(1+\frac{\delta}{1-\delta}\Big)^N = \frac{m-1}{mN} < \frac{1}{N}. \qquad (6.4)
\end{aligned}$$

And the second term II is

$$\begin{aligned}
II = \frac{(1-\delta)^N}{mN^2}&\sum_{(\nu_1,\dots,\nu_N)\in\{0,1\}^N}\Big(\frac{\delta}{1-\delta}\Big)^{\nu_1+\cdots+\nu_N}\\
&\cdot\sum_{t=1}^{m-1}\sum_{1< j< k\le N} e_m\big(t(\nu_j+\nu_{j+1}+\cdots+\nu_k)\big)+e_m\big(-t(\nu_j+\nu_{j+1}+\cdots+\nu_k)\big).
\end{aligned}$$

In order to evaluate II, we are going to group the inner sum into residue classes modulo m according to the values $\nu_j+\cdots+\nu_k \bmod m$, and then change the order of summation. We first make the following observation. Let $1\le n<N$. Given an N-tuple $(\nu_1,\dots,\nu_N)\in\{0,1\}^N$, there are exactly $(N-n)$ sub-blocks of length n of the form $(\nu_j,\nu_{j+1},\dots,\nu_{j+n-1})$. Furthermore, given an n-tuple $(\nu_j,\nu_{j+1},\dots,\nu_{j+n-1})$ such that $\nu_j+\nu_{j+1}+\cdots+\nu_{j+n-1}\equiv a \bmod m$, the number of vectors $(\mu_1,\dots,\mu_N)\in\{0,1\}^N$ such that $\mu_j+\mu_{j+1}+\cdots+\mu_{j+n-1}\equiv a \bmod m$ is

$$\sum_{l\ge 0}\binom{n}{a+lm}\sum_{(\nu_1,\dots,\nu_{N-n})\in\{0,1\}^{N-n}} 1 = 2^{N-n}\sum_{l\ge 0}\binom{n}{a+lm}.$$

Therefore,

$$\begin{aligned}
II &= \frac{(1-\delta)^N}{mN^2}\sum_{t=1}^{m-1}\sum_{a=0}^{m-1}\big(e_m(ta)+e_m(-ta)\big)\\
&\quad\cdot\sum_{n=1}^{N-1}(N-n)\sum_{l\ge 0}\binom{n}{a+lm}\Big(\frac{\delta}{1-\delta}\Big)^{a+lm}\sum_{(\nu_1,\dots,\nu_{N-n})\in\{0,1\}^{N-n}}\Big(\frac{\delta}{1-\delta}\Big)^{\nu_1+\cdots+\nu_{N-n}}\\
&= \frac{(1-\delta)^N}{mN^2}\sum_{t=1}^{m-1}\sum_{a=0}^{m-1}\big(e_m(ta)+e_m(-ta)\big)\sum_{n=1}^{N-1}(N-n)S\Big(\frac{\delta}{1-\delta};a,m,n\Big)\Big(1+\frac{\delta}{1-\delta}\Big)^{N-n}.
\end{aligned}$$

Expanding $S(\frac{\delta}{1-\delta}; a, m, n)$ using Lemma 2.4 leads to

$$\begin{aligned}
&\frac{(1-\delta)^N}{mN^2}\sum_{t=1}^{m-1}\sum_{a=0}^{m-1}(e_m(ta)+e_m(-ta))\sum_{n=1}^{N-1}\frac{N-n}{m}\Big(1+\frac{\delta}{1-\delta}\Big)^{N-n}\sum_{l=0}^{m-1}e_m(-al)\Big(1+\frac{\delta}{1-\delta}e_m(l)\Big)^n\\
&\quad=\frac{(1-\delta)^N}{mN^2}\sum_{t=1}^{m-1}\sum_{n=1}^{N-1}\sum_{l=0}^{m-1}\frac{N-n}{m}\Big(1+\frac{\delta}{1-\delta}\Big)^{N-n}\Big(1+\frac{\delta}{1-\delta}e_m(l)\Big)^n\sum_{a=0}^{m-1}(e_m(a(t-l))+e_m(-a(t+l)))\\
&\quad=\frac{(1-\delta)^N}{mN^2}\sum_{t=1}^{m-1}\sum_{n=1}^{N-1}(N-n)\Big(1+\frac{\delta}{1-\delta}\Big)^{N-n}\Big\{\Big(1+\frac{\delta}{1-\delta}e_m(t)\Big)^n+\Big(1+\frac{\delta}{1-\delta}e_m(-t)\Big)^n\Big\}\\
&\quad=\frac{1}{mN^2}\sum_{t=1}^{m-1}\sum_{n=1}^{N-1}(N-n)\{(1-\delta+\delta e_m(t))^n+(1-\delta+\delta e_m(-t))^n\}.
\end{aligned}$$

Finally, we apply Lemma 2.6 (with the choice $\epsilon = 1$) to bound the double sum above. This gives

$$II \ll_\delta \frac{m}{N}. \tag{6.5}$$

The proposition is proved in view of (6.3), (6.4) and (6.5). □

7 Link for Theorem 1.1

This section provides links between the probabilistic model introduced in Section 5 and the moment of the short sum $S(\Phi; H, x)$. Define the moment function

$$M(r) := \frac{1}{N}\sum_{x=1}^{N} S(\Phi; H, x)^r.$$

Lemma 7.1. *Let N be a large RSA modulus. Suppose $\lambda > 0$ is a real number such that $\min(p, q) > \lambda N^{1/2}$. Let r, H be positive integers such that $H \le \lambda N^{1/2}$ and $1 \le r < 1/4 \log_4 N$. Then as $H \to \infty$,*

$$M(r) = \Big(\frac{H}{4}\Big)^r\big(1 + \mathrm{O}\big((16/H)^r + 4^r N^{-1/4}\big)\big).$$

Proof. We have

$$\begin{aligned}
M(r) &= \frac{1}{N}\sum_{x=1}^{N}\Big(\sum_{x<n\le x+H}\Phi(n)\Big)^r = \frac{1}{N}\sum_{x=1}^{N}\sum_{x<n_1,\dots,n_r\le x+H}\Phi(n_1)\cdots\Phi(n_r)\\
&= \frac{1}{N}\sum_{1\le y_1,\dots,y_r\le H}\sum_{x=1}^{N}\prod_{i=1}^{r}\Phi(x+y_i).
\end{aligned}$$

Given any r-tuple $(y_1, \dots, y_r) \in [1, H]^r$, we can write

$$\prod_{i=1}^{r}\Phi(x+y_i) = \Phi(x+y_{i_1})^{d_{i_1}}\Phi(x+y_{i_2})^{d_{i_2}}\cdots\Phi(x+y_{i_k})^{d_{i_k}},$$

where the elements of k-tuple $(y_{i_1}, \dots, y_{i_k})$ are pair-wise distinct, and d_i are positive integers such that $d_{i_1} + \cdots + d_{i_k} = r$. Therefore, $M(r)$ becomes

$$\frac{1}{N}\sum_{k=1}^{r}\sum_{\substack{\{i_1,\dots,i_k\}\subset\{1,\dots,r\}\\ 1\le y_{i_1},\dots,y_{i_k}\le H}}\sum_{d_{i_1}+\cdots+d_{i_k}=r}\sum_{x=1}^{N}\Phi(x+y_{i_1})^{d_{i_1}}\cdots\Phi(x+y_{i_k})^{d_{i_k}}. \tag{7.1}$$

The sum over x in (7.1) can be bounded using Proposition 4.1, and the sum over d_i can be bounded using Lemma 2.1. Thus,

$$\begin{aligned} M(r) &= \frac{1}{N}\sum_{k=1}^{r}\binom{r}{k}\binom{r-1}{k-1}H^k(N/4^k + \mathrm{O}(kN^{3/4}/2^k)) \\ &= \sum_{k=1}^{r}\binom{r}{k}\binom{r-1}{k-1}(H/4)^k + \mathrm{O}\Big(N^{-1/4}\sum_{k=1}^{r}\binom{r}{k}\binom{r-1}{k-1}H^k\Big) \\ &= (H/4)^r + \mathrm{O}(4^r + N^{-1/4}H^r), \end{aligned}$$

where we have used Lemma 2.2 for bounding the binomial sum above. □

We next compute the characteristic function of the distribution of the normalized sum which is defined as

$$\psi(u) := \frac{1}{N}\sum_{x=1}^{N}\exp\big(iu\widetilde{S(\Phi;H,x)}\big) = \frac{1}{N}\sum_{x=1}^{N}\exp\Big(iu\frac{S(\Phi;H,x)}{\sqrt{H\delta(1-\delta)}}\Big).$$

Proposition 7.2. *Fix $\delta = 1/4$. Let N be a large RSA modulus. Let H, T be positive integers such that $H \le \lambda N^{1/2}$ and $T < 1/4\log_4 N$. Then we have uniformly for $u \le |F(1/4)|^{-1/3}H^{1/6} \approx 1.732H^{1/6}$,*

$$\begin{aligned} \psi(u) = \exp\Big(-\frac{u^2}{2}\Big)\Big(1 + \mathrm{O}\Big(\frac{u^3}{\sqrt{H}}\Big)\Big)\cdot\exp\Big(iu\sqrt{\frac{H}{3}}\Big) \\ + \mathrm{O}\Big((1+16u/\sqrt{3H})^T + N^{-1/4}(1+4u\sqrt{H/3})^T + \frac{(u\sqrt{H/3})^T}{T!}\Big). \end{aligned}$$

Proof. Truncating the exponential series up to the T-th term gives

$$\begin{aligned} \frac{1}{N}\sum_{x=1}^{N}\exp\big(iu\widetilde{S(\Phi;H,x)}\big) &= \frac{1}{N}\sum_{x=1}^{N}\Big(\sum_{r=0}^{T-1}\frac{(iu\widetilde{S(\Phi;H,x)})^r}{r!} + \mathrm{O}\Big(\frac{u^T\widetilde{S(\Phi;H,x)}^T}{T!}\Big)\Big) \\ &= \sum_{r=0}^{T-1}\frac{(iu)^rM(r)}{r!(H\delta(1-\delta))^{r/2}} + \mathrm{O}\Big(\frac{u^TM(T)}{T!(H\delta(1-\delta))^{T/2}}\Big). \end{aligned} \tag{7.2}$$

Using Lemma 7.1, the error is bounded by

$$\frac{u^TM(T)}{T!(H\delta(1-\delta))^{T/2}} \ll \frac{u^T(H\delta)^T}{T!(H\delta(1-\delta))^{T/2}} = \frac{(u\sqrt{H/3})^T}{T!}. \tag{7.3}$$

In view of Lemma 7.1 and Lemma 5.1, the main term in (7.2) becomes

$$\begin{aligned} \sum_{r=0}^{T-1}\frac{(iu)^rM(r)}{r!(H\delta(1-\delta))^{r/2}} &= \sum_{r=0}^{T-1}\frac{(iu)^r\mathbb{E}(Z_H^r)}{r!(H\delta(1-\delta))^{r/2}}\big(1+\mathrm{O}\big((16/H)^r + 4^rN^{-1/4}\big)\big) \\ &= \mathbb{E}\Big(\sum_{r=0}^{T-1}\frac{(iu\widetilde{Z_H})^r}{r!}\Big) + \mathrm{O}\Big(\sum_{r=0}^{T-1}\frac{(16u/H)^r\mathbb{E}(Z_H^r)}{r!(H\delta(1-\delta))^{r/2}} + N^{-1/4}\sum_{r=0}^{T-1}\frac{u^r4^r\mathbb{E}(Z_H^r)}{r!(H\delta(1-\delta))^{r/2}}\Big). \end{aligned} \tag{7.4}$$

Using the power series expansion for $\exp(ix)$, we see that the first term in (7.4) is

$$\mathbb{E}\big(\exp(iu\widetilde{Z_H})\big) + \mathrm{O}\Big(\frac{u^T\mathbb{E}(Z_H^T)}{T!(H\delta(1-\delta))^{T/2}}\Big) = \mathbb{E}\big(\exp(iu\widetilde{Z_H})\big) + \mathrm{O}\Big(\frac{(u\sqrt{H/3})^T}{T!}\Big)$$

because by Lemma 5.1

$$\frac{u^T\mathbb{E}(Z_H^T)}{T!(H\delta(1-\delta))^{T/2}} \ll \frac{u^T(H\delta)^T}{T!(H\delta(1-\delta))^{T/2}} = \frac{(u\sqrt{H/3})^T}{T!}.$$

The first error term in (7.4) is

$$\sum_{r=0}^{T-1}\frac{(16u/H)^r\mathbb{E}(Z_H^r)}{r!(H\delta(1-\delta))^{r/2}} = \sum_{r=0}^{T-1}\frac{(16u/\sqrt{3H})^r}{r!} \le \Big(1+\frac{16u}{\sqrt{3H}}\Big)^T.$$

The second error term in (7.4) is

$$N^{-1/4}\sum_{r=0}^{T-1}\frac{u^r4^r\mathbb{E}(Z_H^r)}{r!(H\delta(1-\delta))^{r/2}} \ll N^{-1/4}\sum_{r=0}^{T-1}\frac{(4u\sqrt{H/3})^r}{r!} \ll N^{-1/4}(1+4u\sqrt{H/3})^T.$$

The proposition is proved in view of (7.4) and Lemma 5.2. □

8 Proof of Theorem 1.1

A key element in the proof of Theorem 1.1 is a smooth approximation for the signum function. The original idea is due to Selberg [14], and the method is applied effectively by Lamzouri [10] in establishing a two-dimensional Gaussian distribution for short complex character sums. In deriving Theorem 1.1, the arguments are essentially those of Lamzouri's adapted to our situation.

Let $1_{a,b}$ be the characteristic function of the interval $[a, b]$.

Lemma 8.1. *Let real numbers $t > 0, a, b$ be given such that $a \le b$. Then*

$$1_{a,b}(x) = \operatorname{Im} \int_0^t G\Big(\frac{u}{t}\Big) \exp(2\pi i u x) f_{a,b}(u) \frac{du}{u} + \mathrm{O}\Big(\frac{\sin^2(\pi t(x-a))}{(\pi t(x-a))^2} + \frac{\sin^2(\pi t(x-b))}{(\pi t(x-b))^2} \Big),$$

where the functions G, $f_{a,b}$ are defined as

$$G(u) = \frac{2u}{\pi} + 2(1-u)u\cot(\pi u) \quad \text{for } u \in [0,1],$$
$$f_{a,b}(u) = \frac{\exp(-2\pi i a u) - \exp(-2\pi i b u)}{2}.$$

Furthermore, $G(u)$ is differentiable and $0 \le G(u) \le 2/\pi$ for $0 \le u \le 1$. The function $f_{a,b}$ satisfies the bound

$$|f_{a,b}(u)| \le \pi u |b-a|.$$

Proof. See [10, Lemma 4.1 and the subsequent discussion]. □

The following lemma generalizes [10, Lemma 4.2].

Lemma 8.2. *Let δ be real number such that $0 < \delta < 1$. Let t be a large positive number. Then, uniformly for all real numbers $a < b$,*

$$\operatorname{Im} \int_0^t G\Big(\frac{u}{t}\Big) \exp\Big(-\frac{(2\pi u)^2}{2}\Big) \exp\Big(2\pi i u \sqrt{\frac{H\delta}{1-\delta}}\Big) f_{a,b}(u) \frac{du}{u} = \frac{1}{\sqrt{2\pi}} \int_a^b \exp\Big(-\frac{(x - \sqrt{H\delta/(1-\delta)})^2}{2}\Big) dx + \mathrm{O}\Big(\frac{1}{t}\Big).$$

Proof. Let X be a Gaussian random variable with mean $\mu = \sqrt{H\delta/(1-\delta)}$ and variance $\sigma^2 = 1$. We have

$$\frac{1}{\sqrt{2\pi}} \int_a^b \exp\Big(-\frac{(x - \sqrt{H\delta/(1-\delta)})^2}{2}\Big) dx = P(X \in [a,b]) = \mathbb{E}(1_{a,b}(X)). \tag{8.1}$$

Using Lemma 8.1, we have

$$\mathbb{E}(1_{a,b}(X)) = \operatorname{Im} \int_0^t G\Big(\frac{u}{t}\Big) \mathbb{E}\big(\exp(2\pi i u X)\big) f_{a,b}(u) \frac{du}{u} + \mathrm{O}\Big(\mathbb{E}\Big(\frac{\sin^2(\pi t(X-a))}{(\pi t(X-a))^2} + \frac{\sin^2(\pi t(X-b))}{(\pi t(X-b))^2} \Big)\Big). \tag{8.2}$$

The error terms in (8.2) are known to be $\ll 1/t$ (see the proof of [10, Lemma 4.2]). Furthermore for the Gaussian random variable X, we have

$$\mathbb{E}\big(\exp(2\pi i u X)\big) = \exp(2\pi i \mu u) \exp\Big(-\frac{(2\pi\sigma u)^2}{2}\Big) = \exp\Big(2\pi i u \sqrt{\frac{H\delta}{1-\delta}}\Big) \exp\Big(-\frac{(2\pi u)^2}{2}\Big). \tag{8.3}$$

This proves the lemma after putting (8.2) and (8.3) in (8.1). □

Proof of Theorem 1.1. Let $\delta = 1/4$ be fixed. Let T be a positive integer, and $t \ge 1$ be a real number such that

$$t = \min\Big(\frac{1}{2\pi} H^{1/6}, \frac{1}{4}\sqrt{\frac{\log N}{H \log H}}\Big) \quad \text{and} \quad T = [t^2 H].$$

Using Lemma 8.1, we have

$$\frac{1}{N}\sum_{x=1}^{N} 1_{a,b}(S(\widetilde{\Phi;H,x})) = \operatorname{Im}\int_0^t G\Big(\frac{u}{t}\Big)\psi(2\pi u) f_{a,b}(u)\frac{du}{u} + \mathrm{O}\big(I(t,a)+I(t,b)\big). \tag{8.4}$$

The main term in (8.4) can be expanded using Proposition 7.2:

$$\operatorname{Im}\int_0^t G\Big(\frac{u}{t}\Big)\psi(2\pi u) f_{a,b}(u)\frac{du}{u} = \operatorname{Im}\int_0^t G\Big(\frac{u}{t}\Big)\exp\Big(-\frac{(2\pi u)^2}{2}\Big)\exp\Big(2\pi i u\sqrt{\frac{H}{3}}\Big) f_{a,b}(u)\frac{du}{u} + \mathrm{O}(E),$$

where

$$E = (b-a)\int_0^t \exp\Big(-\frac{(2\pi u)^2}{2}\Big)\frac{(2\pi u)^3}{\sqrt{H}} + \big(1+32\pi u/\sqrt{3H}\big)^T + N^{-1/4}\big(1+8\pi u\sqrt{H/3}\big)^T + \frac{(2\pi u\sqrt{H/3})^T}{T!}du.$$

Thus in view of Lemma 8.2, we have

$$\begin{aligned}\frac{1}{N}\sum_{x=1}^{N} 1_{a,b}(S(\widetilde{\Phi;H,x})) &= \frac{1}{\sqrt{2\pi}}\int_a^b \exp\Big(-\frac{(x-\sqrt{H/3})^2}{2}\Big)dx + \mathrm{O}\Big(\frac{1}{t}\Big)\\ &\quad + \mathrm{O}\Big((b-a+1)\Big(\frac{1}{\sqrt{H}} + \Big(\frac{64t^2}{\sqrt{H}}\Big)^T + N^{-1/4}(16t\sqrt{H})^{T+1} + \frac{(4t\sqrt{H})^{T+1}}{T!}\Big)\Big).\end{aligned} \tag{8.5}$$

The error terms $I(t,a), I(t,b)$ in (8.4) can be bounded similarly as done in the proof of [10, Theorem 1]. For $l = a, b$, we have

$$\begin{aligned} I(t,l) &= \frac{1}{N}\sum_{x=1}^{N}\frac{\sin^2(\pi t(S(\widetilde{\Phi;H,x})-l))}{(\pi t(x-l))^2}\\ &= \frac{1}{N}\sum_{x=1}^{N}\frac{2}{t^2}\int_0^t (t-v)\cos\big(2\pi v(S(\widetilde{\Phi;H,x})-l)\big)dv\\ &= \operatorname{Re}\frac{1}{N}\sum_{x=1}^{N}\frac{2}{t^2}\int_0^t (t-v)\exp(-2\pi i l v)\exp\big(2\pi i v S(\widetilde{\Phi;H,x})\big)dv\\ &= \operatorname{Re}\frac{2}{t^2}\int_0^t (t-v)\exp(-2\pi i l v)\psi(2\pi v)dv\\ &\ll \frac{1}{t}\int_0^t \exp\Big(-\frac{(2\pi u)^2}{2}\Big) + \Big(1+\frac{32\pi u}{\sqrt{3H}}\Big)^T + N^{-1/4}\Big(1+8\pi u\sqrt{\frac{H}{3}}\Big)^T + \frac{(2\pi u\sqrt{H/3})^T}{T!}du\\ &\ll \frac{1}{t}\Big(1+\Big(\frac{64t^2}{\sqrt{H}}\Big)^T + N^{-1/4}(16t\sqrt{H})^{T+1} + \frac{(4t\sqrt{H})^{T+1}}{T!}\Big). \end{aligned} \tag{8.6}$$

In view of (8.4), (8.5) and (8.6), thus far we have obtained

$$\frac{1}{N}\sum_{x=1}^{N} 1_{a,b}(S(\widetilde{\Phi;H,x})) = \frac{1}{\sqrt{2\pi}}\int_a^b \exp\Big(-\frac{(x-\sqrt{H/3})^2}{2}\Big)dx + \mathrm{O}(E_1),$$

where the error term E_1 is

$$E_1 = (b-a+1)\Big(\frac{1}{t} + \frac{1}{\sqrt{H}} + \Big(\frac{64t^2}{\sqrt{H}}\Big)^T + N^{-1/4}(16t\sqrt{H})^{T+1} + \frac{(4t\sqrt{H})^{T+1}}{T!}\Big). \tag{8.7}$$

Now the Stirling approximation gives

$$\frac{(4t\sqrt{H})^{T+1}}{T!} \ll \frac{4^T T^{\frac{T+1}{2}}}{T!} \ll \frac{1}{\sqrt{T}} \ll \frac{1}{t}.$$

For large H, we have

$$\Big(\frac{64t^2}{\sqrt{H}}\Big)^T \ll \Big(\frac{1}{\sqrt[6]{H}}\Big)^T \ll \frac{1}{\sqrt[6]{H}}.$$

Moreover,

$$N^{-1/4}\big(16t\sqrt{H}\big)^{T+1} \ll N^{-1/32}$$

because for large H,

$$16^{T+1} \ll 16^{\frac{\log N}{16\log H}} = \exp\Big(\frac{\log 16\log N}{16\log H}\Big) \ll \exp\Big(\frac{\log N}{32}\Big) = N^{1/32},$$

and as $\log N/\log H$ becomes large,

$$t^{T+1} \le H^{\frac{1}{6}(T+1)} \le H^{\frac{1}{6}(\frac{\log N}{16\log H}+1)} \ll N^{1/64}, \qquad (\sqrt{H})^{T+1} \le H^{\frac{1}{2}(\frac{\log N}{16\log H}+1)} \ll H^{\frac{3\log N}{64\log H}} = N^{3/64}.$$

To conclude, we have the error term (8.7):

$$E \ll (b-a+1)(1/t+N^{-1/32}) \ll (b-a+1)\Big(H^{-1/6}+\sqrt{\frac{H\log H}{\log N}}\Big). \qquad \square$$

9 Link for Theorem 1.6

This section provides links between the probabilistic model introduced in Section 6 and Theorem 1.6. The goal of this section is the proof of Proposition 9.3. The result extends [11, Proposition 3.1] to the setting $\mathbb{Z}/N\mathbb{Z}$. Let $\vec{v} = (v_1, \dots, v_s) \in \{0,1\}^s$. Define

$$\Delta(\vec{v}) := \big\{1 \le i \le s : v_i = 1,\ \vec{v} = (v_1, \dots, v_s)\big\},$$
$$\overline{\Delta}(\vec{v}) := \big\{1 \le i \le s : v_i = 0,\ \vec{v} = (v_1, \dots, v_s)\big\}.$$

Basically, $\Delta(\vec{v})$ is the index set where the vector $\vec{v}$ has the 1 coordinates, and $\overline{\Delta}(\vec{v})$ is the index set where the vector $\vec{v}$ has the 0 coordinates. Define

$$D(N;\vec{v},s) := \big\{1 \le n \le N : \Phi(n+j) = v_j \text{ for all } 0 \le j \le s-1\big\},$$

where $\Phi(\cdot)$ is the quadratic residue characteristic function.

Lemma 9.1. *Let $n \le N$ and s be positive integers. Then the number n such that the sequence $\{\Phi(n), \Phi(n+1), \dots, \Phi(n+s-1)\}$ contains a term not equal to 0 or 1 is $\ll sN^{1/2}$.*

Proof. Notice that $\Phi(n) \neq 0, 1$ if p or q divides n. The bound for this exception set

$$\mathcal{D}_0 := \{0 \le n < N : \text{there exists } i \text{ such that } 0 \le i \le s-1 \text{ and } \chi_N(n+i) = 0\}$$

is known, see [9, Lemma 3.2] for a precise estimate. $\square$

Lemma 9.2. *Let $1 \le s < \frac{1}{16}\log_2 N$ be a positive integer. Then we have*

$$|D(N;\vec{v},s)| = N\Big(\frac{1}{4}\Big)^{|\Delta(\vec{v})|}\Big(\frac{3}{4}\Big)^{|\overline{\Delta}(\vec{v})|}\big(1+O(N^{-1/16})\big).$$

Proof. First, notice that if $\Phi(n+j) \in \{0,1\}$, then

$$(1-v_j)+(2v_j-1)\Phi(n+j) = \begin{cases} 1, & \text{if } \Phi(n+j) = v_j, \\ 0, & \text{if } \Phi(n+j) \neq v_j. \end{cases}$$

Thus $(1-v_j)+(2v_j-1)\Phi(n+j)$ is a characteristic function on the sequence v_j, $0 \le j < s$, provided there are only 0, 1-terms in the sequence $\{\Phi(n), \dots, \Phi(n+s-1)\}$. The number of exceptions (i.e. sequences $\{\Phi(n), \dots, \Phi(n+s-1)\}$ containing a non-0, 1-term) is bounded by $sN^{1/2}$ by Lemma 9.1. Therefore,

$$\begin{aligned}
|D(N;\vec{v},s)| &= \sum_{n=1}^{N}\prod_{j=0}^{s-1}\big(1-v_j+(2v_j-1)\Phi(n+j)\big)+\mathrm{O}(sN^{1/2})\\
&= \sum_{n=1}^{N}\Big(\prod_{j\in\Delta(\vec{v})}\Phi(n+j)\Big)\Big(\prod_{j\in\overline{\Delta}(\vec{v})}\big(1-\Phi(n+j)\big)\Big)+\mathrm{O}(N^{1/2}\log N)\\
&= \sum_{n=1}^{N}\Big(\prod_{j\in\Delta(\vec{v})}\Phi(n+j)\Big)\Big(1+\sum_{\substack{T\subset\overline{\Delta}(\vec{v})\\ T\neq\emptyset}}(-1)^{|T|}\prod_{j\in T}\Phi(n+j)\Big)+\mathrm{O}(N^{1/2}\log N)\\
&= \sum_{n=1}^{N}\Big(\prod_{j\in\Delta(\vec{v})}\Phi(n+j)\Big)+\sum_{\substack{T\subset\overline{\Delta}(\vec{v})\\ T\neq\emptyset}}(-1)^{|T|}\sum_{n=1}^{N}\Big(\prod_{j\in\Delta(\vec{v})\cup T}\Phi(n+j)\Big)+\mathrm{O}(N^{1/2}\log N)\\
&= I+II+\mathrm{O}(N^{1/2}\log N). \qquad (9.1)
\end{aligned}$$

The asymptotic evaluation of I, II can be derived based on Proposition 4.1:

$$I = \sum_{n=1}^{N}\prod_{j\in\Delta(\vec{v})}\Phi(n+j) = \frac{N}{4^{|\Delta(\vec{v})|}}+\mathrm{O}(sN^{3/4}), \qquad (9.2)$$

$$\begin{aligned}
II &= \sum_{\substack{T\subset\overline{\Delta}(\vec{v})\\ T\neq\emptyset}}(-1)^{|T|}\sum_{n=1}^{N}\prod_{j\in\Delta(\vec{v})\cup T}\Phi(n+j) = \sum_{\substack{T\subset\overline{\Delta}(\vec{v})\\ T\neq\emptyset}}(-1)^{|T|}\Big(\frac{N}{4^{|\Delta(\vec{v})|+|T|}}+\mathrm{O}(sN^{3/4})\Big)\\
&= \frac{N}{4^{|\Delta(\vec{v})|}}\sum_{\substack{T\subset\overline{\Delta}(\vec{v})\\ T\neq\emptyset}}\frac{(-1)^{|T|}}{4^{|T|}}+\mathrm{O}(s2^sN^{3/4}) = \frac{N}{4^{|\Delta(\vec{v})|}}\Big(\Big(1-\frac{1}{4}\Big)^{|\overline{\Delta}(\vec{v})|}-1\Big)+\mathrm{O}(s2^sN^{3/4})\\
&= N\Big(\frac{1}{4}\Big)^{|\Delta(\vec{v})|}\Big(\frac{3}{4}\Big)^{|\overline{\Delta}(\vec{v})|}-\frac{N}{4^{|\Delta(\vec{v})|}}+\mathrm{O}(s2^sN^{3/4}). \qquad (9.3)
\end{aligned}$$

Finally, in view of (9.1), (9.2), and (9.3), we have

$$|D(N;\vec{v},s)| = N\Big(\frac{1}{4}\Big)^{|\Delta(\vec{v})|}\Big(\frac{3}{4}\Big)^{|\overline{\Delta}(\vec{v})|}+\mathrm{O}(s2^sN^{3/4}) = N\Big(\frac{1}{4}\Big)^{|\Delta(\vec{v})|}\Big(\frac{3}{4}\Big)^{|\overline{\Delta}(\vec{v})|}\big(1+\mathrm{O}(N^{-1/16})\big).$$

The error term above is obtained by observing that $(\frac{1}{4})^{|\Delta(\vec{v})|}(\frac{3}{4})^{|\overline{\Delta}(\vec{v})|}$ has the absolute minimum value $1/4^s$. □

Proposition 9.3. *Let N be a large RSA modulus. Let $1 \le s < \frac{1}{16}\log_2 N$ be a positive integer. Then for any non-negative function $h : \{0,1\}^s \to \mathbb{R}^+$, we have*

$$\frac{1}{N}\sum_{n=1}^{N}h\big(\Phi(n),\dots,\Phi(n+s-1)\big) = \mathbb{E}\big(h(X_1,\dots,X_s)\big)\big(1+\mathrm{O}(N^{-1/16})\big),$$

where $X_1, \dots, X_s$ are independent random variables taking the values 1, 0 with probability $\frac{1}{4}, \frac{3}{4}$, respectively.

Proof. In view of Lemma 9.2, we have

$$\begin{aligned}
\frac{1}{N}\sum_{n=1}^{N}h\big(\Phi(n),\dots,\Phi(n+s-1)\big) &= \frac{1}{N}\sum_{\vec{v}=(v_1,\dots,v_s)\in\{0,1\}^s}h(v_1,\dots,v_s)|D(N;\vec{v},s)|\\
&= \sum_{\vec{v}=(v_1,\dots,v_s)\in\{0,1\}^s}h(v_1,\dots,v_s)\Big(\frac{1}{4}\Big)^{|\Delta(\vec{v})|}\Big(\frac{3}{4}\Big)^{|\overline{\Delta}(\vec{v})|}\big(1+\mathrm{O}(N^{-1/16})\big)\\
&= \mathbb{E}\big(h(X_1,\dots,X_s)\big)\big(1+\mathrm{O}(N^{-1/16})\big). \qquad \square
\end{aligned}$$

10 Proof of Theorem 1.6

We closely follow the proof of [11, Theorem 2]. Let

$$Y(k,a) = \begin{cases} 1, & \text{if } |\Omega(k)| \equiv a \bmod m, \\ 0, & \text{otherwise.} \end{cases}$$

For any integer $j \geq 1$, we have

$$\Psi(N;m,a) = \frac{1}{N}\sum_{n=1}^{N} Y(n,a) = \frac{1}{N}\sum_{n=1}^{N} Y(n+j,a) + \mathrm{O}\Big(\frac{j}{N}\Big).$$

Therefore,

$$\begin{aligned}\sum_{a=0}^{m-1}\Big(\Psi(N;m,a) - \frac{1}{m}\Big)^2 &= \sum_{a=0}^{m-1}\Big(\frac{1}{NL}\sum_{j=1}^{L}\sum_{n=1}^{N} Y(n+j,a) - \frac{1}{m} + \mathrm{O}\Big(\frac{L}{N}\Big)\Big)^2 \\ &= \sum_{a=0}^{m-1}\Big(\frac{1}{NL}\sum_{j=1}^{L}\sum_{n=1}^{N} Y(n+j,a) - \frac{1}{m}\Big)^2 + \mathrm{O}\Big(\frac{mL}{N}\Big),\end{aligned}$$

where $L = [1/16\log_2 N]$. Applying the Cauchy–Schwarz inequality gives the upper bound

$$\sum_{a=0}^{m-1}\frac{1}{N}\sum_{n=1}^{N}\Big(\frac{1}{L}\sum_{j=1}^{L} Y(n+j,a) - \frac{1}{m}\Big)^2 + \mathrm{O}\Big(\frac{mL}{N}\Big) = \sum_{n=1}^{N}\frac{1}{N}h\big(\Phi(n+1),\dots,\Phi(n+L)\big) + \mathrm{O}\Big(\frac{mL}{N}\Big)$$

where the function $h : \{0,1\}^L \to \mathbb{R}^+$ is defined as

$$h(v_1,\dots,v_L) = \sum_{b=0}^{m-1}\Big(\frac{1}{L}\,\big|1 \leq j \leq L : v_1 + \cdots + v_j \equiv b \bmod m\big|\Big)^2.$$

Finally, in view of Proposition 9.3 and Proposition 6.2, we have

$$\begin{aligned}\sum_{a=0}^{m-1}\Big(\Psi(N;m,a) - \frac{1}{m}\Big)^2 &\leq \mathbb{E}\big(h(X_1,\dots,X_s)\big)\big(1 + \mathrm{O}(N^{-1/16})\big) + \mathrm{O}\Big(\frac{mL}{N}\Big) \\ &= \sum_{a=0}^{m-1}\mathbb{E}\Big(\Big(\Phi_{\mathrm{rand}}(L;m,a) - \frac{1}{m}\Big)^2\Big)\big(1 + \mathrm{O}(N^{-1/16})\big) + \mathrm{O}\Big(\frac{mL}{N}\Big) \\ &\ll \frac{m}{L} + \frac{mL}{N} \ll \frac{m}{\log N}.\end{aligned}$$

Notations

N	RSA modulus: $N = pq$ with $1 < p, q \leq cN^{1/2}$ where $c > 0$
$\phi(\cdot)$	Euler's totient function
$e_m(x)$	$\exp(\frac{2\pi i x}{m})$
$\chi_p(\cdot)$, $\chi_q(\cdot)$	Legendre symbols: $(\frac{\cdot}{p})$, $(\frac{\cdot}{q})$
$\Phi(\cdot)$	Quadratic residue characteristic function: $\frac{1}{4}(1+\chi_p(n))(1+\chi_q(n))$
$S(\Phi;H,x)$	Short character sum: $\sum_{x<n\leq x+H}\Phi(n)$
$M(r)$	r-th moment: $\frac{1}{N}\sum_{x=1}^{N} S(\Phi;H,x)^r$
$S(x;a,m,k)$	$\sum_{n\equiv a \bmod m,\, n\leq k}\binom{k}{n}x^n$
$S_{N,p,q}(T,T')$	$\sum_{n=1}^{N}\chi_p(\prod_{i\in T}(n+i))\chi_q(\prod_{i\in T'}(n+i))$

X_i	Random variable
$\mathbb{E}(X)$	expectation of the random variable X
Z_H	$X_1 + \cdots + X_H$
$\widetilde{Z_H}, \widetilde{X}, \tilde{u}$	Normalized variables
$\Phi_{\text{rand}}(N; m, a)$	$\frac{1}{N}\big\lvert\{1 \le k \le N : Z_k \equiv a \bmod m\}\big\rvert$
$\vec{v}$	Binary vector: $\vec{v} = (v_1, \dots, v_k) \in \{0,1\}^k$
$\Delta(\vec{v})$	$\{1 \le i \le s : v_i = 1,\ \vec{v} = (v_1, \dots, v_s)\}$
$\overline{\Delta}(\vec{v})$	$\{1 \le i \le s : v_i = 0,\ \vec{v} = (v_1, \dots, v_s)\}$
$D(N; \vec{v}, s)$	$\{1 \le n \le N : \Phi(n+j) = v_j \text{ for all } 0 \le j \le s-1\}$
$\mathcal{Q}(R; a, q)$	$\big\lvert\{n = a + rq : 1 \le r \le R, \text{ and } n \text{ is a quadratic residue mod } N\}\big\rvert$
$\mathcal{NQ}(R; a, q)$	$\big\lvert\{n = a + rq : 1 \le r \le R, \text{ and } n \text{ is a quadratic non-residue mod } N\}\big\rvert$
$\mathcal{R}(k)$	$\{1 \le n \le k : n \text{ is a quadratic residue mod } N\}$
$\mathcal{N}(k)$	$\{1 \le n \le k : n \text{ is a quadratic non-residue mod } N\}$
$\Psi_{\mathcal{R}}(N; m, a)$	$\frac{1}{N}\big\lvert\{1 \le k \le N : \lvert\mathcal{R}(k)\rvert \equiv a \bmod m\}\big\rvert$
$\Psi_{\mathcal{N}}(N; m, a)$	$\frac{1}{N}\big\lvert\{1 \le k \le N : \lvert\mathcal{N}(k)\rvert \equiv a \bmod m\}\big\rvert$

Acknowledgement: The author particularly wishes to thank Lillian Pierce for helpful discussions concerning the character sum bound in Section 4.

References

[1] S. Chatterjee and K. Soundararajan, Random multiplicative functions in short intervals, *Int. Math. Res. Not. IMRN* (2012), no. 3, 479–492.
[2] H. Davenport, *Multiplicative Number Theory*, 2nd ed., Graduate Texts in Math. 74, Springer, New York, 1980.
[3] H. Davenport and P. Erdös, The distribution of quadratic and higher residues, *Publ. Math. Debrecen* **2** (1952), 252–265.
[4] J. Friedlander and H. Iwaniec, Estimates for character sums, *Proc. Amer. Math. Soc.* **119** (1993), no. 2, 365–372.
[5] A. J. Harper, On the limit distributions of some sums of a random multiplicative function, *J. Reine Angew. Math.* **678** (2013), 95–124.
[6] D. R. Heath-Brown, Burgess's bounds for character sums, in: *Number Theory and Related Fields*, Springer Proc. Math. Stat. 43, Springer, New York (2013), 199–213.
[7] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, Amer. Math. Soc. Colloq. Publ. 53, American Mathematical Society, Providence, 2004.
[8] B. Justus, An extension of ramus' identity with applications, *Šiauliai Math. Semin.* **8** (2013), 109–115.
[9] B. Justus, The distribution of quadratic residues and non-residues in the Goldwasser–Micali type of cryptosystem, *J. Math. Cryptol.* **8** (2014), 115–140.
[10] Y. Lamzouri, The distribution of short character sums, *Math. Proc. Cambridge Philos. Soc.* **155** (2013), no. 2, 207–218.
[11] Y. Lamzouri and A. Zaharescu, Randomness of character sums modulo m, *J. Number Theory* **132** (2012), no. 12, 2779–2792.
[12] K.-H. Mak and A. Zaharescu, The distribution of values of short hybrid exponential sums on curves over finite fields, *Math. Res. Lett.* **18** (2011), no. 1, 155–174.
[13] C. Ramus, Solution generale d'un probleme d'analyse combinatoire, *J. Reine Angew. Math.* **11** (1834), 353–355.
[14] A. Selberg, Old and new conjectures and results about a class of Dirichlet series, in: *Proceedings of the Amalfi Conference on Analytic Number Theory* (Maiori 1989), University of Salerno (1992), 367–385.
[15] X. Shao, Character sums over unions of intervals, *Forum Math.* (2014), DOI 10.1515/forum-2013-0080.

Received January 31, 2014; accepted May 4, 2015.