ON THE DISTRIBUTION OF VERY SHORT CHARACTER SUMS

PAWEŁ NOSAL

ABSTRACT. We establish a central limit theorem of $\frac{1}{\sqrt{h_p}}\sum_{X<n\leq X+h_p}\left(\frac{n}{p}\right)$ for almost all the primes p, with X uniformly random in $[g(p)]$, $g(p)$ an arbitrary divergent function growing slower than any power of p, provided $(\log h_p)/(\log g(p)) \to 0, h_p \to \infty$ as $p \to \infty$. This improves the recent results of Basak, Nath and Zaharescu, who established this for $g(p) = (\log p)^A, A > 1$. We also use the best currently available tools to expand the original central limit theorem of Davenport and Erdős for all the primes to a shorter interval of starting points.
In this paper we exploit a Selberg's sieve argument, recently used by Harper, an intersection result due to Evertse and Silverman and some consequences of the Weil bound on general character sums.

1. INTRODUCTION

Let q be a prime number. Obviously, the least quadratic residue modulo any prime is 1. No such information is in general offered for quadratic non-residues (further denoted NQR) and the only trivial observation that we can make is that the first NQR needs to be a prime smaller than q (by complete multiplicativity of Dirichlet characters and their orthogonality relation). Detection of the least NQR is one of the main motivations for study of sums of the type

$$\sum_{n\leq h}\chi_q(n),$$

where χ_q is the Dirichlet character realized by the Legendre symbol $(\frac{n}{q})$. Being able to non-trivially bound short character sums immediately gives us a way to detect the least quadratic non-residues, simply by finding the first h when

$$\sum_{n\leq h}\chi_q(n) < \lfloor h \rfloor.$$

In general, the study of partial sums of sequences has an absolutely critical role in analytic number theory, as via standard techniques, like Perron's formula, these are directly connected to the behaviour of their Dirichlet series. In this paper the object of our attention are character sums with varying starting point

$$S_h(x,q) := \sum_{x<n\leq x+h}\chi_q(n),$$

where q, the modulus of the character is huge and $h = o(q)$. Note that since $\chi(n)$ is periodic with period q, there is no need to consider sums with $h > q$.

1.1. **Remark on notation.** Throughout the paper we adopt the usual asymptotic notation. We write $f \ll g$, in the Vinogradov notation, or $f = O(g)$ in the *big O* notation to mean $|f(x)| \leq Cg(x)$, for some constant $C > 0$, as $x \to \infty$. Similarly, we write $f(x) = o(g(x))$ to signify that $|(f/g)(x)| \to 0$ as $x \to \infty$. For any $X \in \mathbb{R}, k \in \mathbb{N}_{\geq 0}$ we shall also write

$$[X] := \{1, 2, ..., \lfloor X \rfloor\}$$

and

$$[X]^k := \{x \in \mathbb{R}^k \text{ with coefficients in } [X]\}.$$

The author is supported by the Warwick Mathematics Institute Centre for Doctoral Training, and gratefully acknowledges funding from the Heilbronn Institute for Mathematical Research (HIMR) and the UK Engineering and Physical Sciences Research Council under the "Additional Funding Programme for Mathematical Sciences" (Grant number: EP/V521917/1).

Additionally, for convenience of comparison we establish notation closely resembling that of [3]. Let $\mathcal{P}$ be the set of primes, $\eta \geq 0.01$ be an arbitrary constant, h_q depends on a prime $q \in \mathcal{P}$. Given a subset of primes $\mathcal{B} \subseteq \mathcal{P}$ we say it is η-sufficient if

$$|(\mathcal{P}\backslash\mathcal{B}) \cap [X, X + X^{\eta}]| = o\left(\frac{X^{\eta}}{\log X}\right), \tag{1}$$

holds for all X sufficiently large. We will say $\mathcal{B}$ is η-strong if for some $\delta > 0$

$$|(\mathcal{P}\backslash\mathcal{B}) \cap [X, X + X^{\eta}]| \leq \frac{X^{\eta}}{(\log X)^{1+\delta}}$$

holds for all X sufficiently large. Now, let $g(q)$ be any function that diverges to infinity. Further, for any real number λ define

$$\mathcal{M}_{g,q,h_q}(\lambda) := \frac{1}{g(q)} \left| 1 \leq m \leq g(q); S_{h_q}(m, q) \leq \lambda\sqrt{h_q} \right|.$$

Finally, we define the extended Rademacher random multiplicative function:

Definition 1. *Extended Rademacher random multiplicative functions*
Let $(f(q))_{q\in\mathcal{P}}$ be a set of independent random variables indexed by primes, such that for any $q \in \mathcal{P}$, $f(q) = \pm 1$ with equal probability. Then, for any $n \in \mathbb{N}$ define

$$f(n) := \prod_{q^{\alpha_q}||n} f(q)^{\alpha_q},$$

and $f(1) = 1$, so that $f(n)$ is completely multiplicative. We call $f(n)$ the extended Rademacher random multiplicative function.

1.2. **Historical results and motivation.** The first non-trivial result on distribution of character sums was proven independently by Pólya and Ivan Vinogradov:

$$\sum_{n\leq h} \chi(n) \ll \sqrt{q} \log q, \tag{2}$$

where χ is any non-principal Dirichlet character of conductor q. In the paper of Vinogradov, he also came up with a method, which provided a character sum bound for Legendre symbols like the one above, allows us to get an additional saving for the size of the least NQR. He showed that the least NQR is of order

$$\ll q^{\frac{1}{2\sqrt{e}}} (\log q)^2$$

and conjectured that its actual size is

$$\ll_{\epsilon} q^{\epsilon}, \tag{3}$$

for all $\epsilon > 0$. Following this, the next improvement was due to Davenport and Erdős [9], who removed a $2 - 1/\sqrt{e}$ factors of $\log q$ from Vinogradov's bound of least quadratic NQR. After that, Burgess [7] improved the character sum bound to

$$\sum_{n\leq H} \chi(n) \ll q^{1/4} \log q$$

and used Vinogradov's trick to obtain the last unconditional improvement of the general bound for the least NQR:

$$\ll_{\epsilon} q^{\frac{1}{4\sqrt{e}}+\epsilon},$$

for any fixed ϵ. Improving this bound remains an important problem in analytic number theory. Worth mentioning are the conditional results of Ankeny [1], who showed that under the Generalized Riemann Hypothesis the least NQR is $\ll (\log q)^2$, and Lamzouri et al. [16] who made this bound explicit with implied constant equal to one. However, in this paper we will be concerned with another type of results, first of which appeared in the aforementioned paper of Davenport and Erdős [9]. They establish, that if X is uniformly random $\in \{0, 1, ..., q - 1\}$ and

$$\frac{\log h_q}{\log q} \to 0, h_q \to \infty$$

with q, then the normalized sums of lengths h_q with a starting point at X converge in distribution to a standard normal

$$\frac{\sum_{X<n\leq X+h_q}\chi_q(n)}{\sqrt{h_q}} \to_D \mathcal{N}(0,1), \tag{4}$$

as the conductor $q\to\infty$ and χ_q is the Legendre symbol. This type of central limit theorem results are the main focus of our paper. In this setting, short character sums were studied by Chan, Choi and Zaharescu [8] who established a multidimensional version of (4), Lamzouri [15], who extended it to more general Dirichlet characters, Harper [14], who showed that for some extremal ranges of H, (4) is true for almost all the primes, but **not** all the primes, and Basak, Nath and Zaharescu [3] who showed (4) for almost all the primes, with X uniformly random in interval of length $(\log q)^A, A>1$. All of these results employ the method of moments to establish desired distributional convergence and it will be the same in our case. We note in passing, that long character sums with $h_q\approx q$ have been studied in depth for example in [5] or [11–13]. This paper was motivated by the aforementioned work of Basak, Nath and Zaharescu [3] on the distribution of short character sums with moving starting point and the paper of Harper [14]. Authors of [3] managed to prove that if one takes the limit in equation (4) along any but a small, badly behaved set of primes, then one can prove suitable central limit theorem while having the starting point chosen uniformly at random from a set of length $(\log q)^A, A>1$.

Theorem 0. *(Basak-Nath-Zaharescu, 2023) Let $\eta\in(1/2,1], A>1$ be arbitrary constants. For each prime q choose an integer h_q such that $h_q\to\infty$ and $\frac{\log h_q}{\log\log q}\to 0$ as $q\to\infty$. Then, there exists an η-strong set $\mathcal{B}$ of primes such that for any $\lambda\in\mathbb{R}$*

$$\lim_{\substack{q\in\mathcal{B}\\ q\to\infty}} \mathcal{M}_{(\log q)^A,q,h_q}(\lambda) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\lambda} e^{-t^2/2}dt.$$

In our work, we streamline the proof, extend their results, and additionally present an up to date central limit theorem result for all the primes.

1.3. **Statement of results and discussion of improvements.** In the main result of the paper we improve on the above.

Theorem 1. *Let g be an increasing, divergent function such that*

- $g'(c)=O\left((g(Q))^{0.99}Q^{-0.01}\right), \forall c\in[Q,Q+Q^{0.01}]$ *as* $Q\to\infty$,
- $g(Q)\ll Q^{\epsilon}, \forall\epsilon>0$.

For each prime q, choose h_q such that $\log h_q/(\log g(q))\to 0$ and $h_q\to\infty$ as $q\to\infty$. Then, there exists a set $\mathcal{B}$ of primes such that $\mathcal{B}$ is η-sufficient for all $\eta\in[0.01,1]$ and for any $\lambda\in\mathbb{R}$,

$$\lim_{\substack{q\to\infty\\ q\in\mathcal{B}}} \mathcal{M}_{g,q,h_q}(\lambda) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\lambda} e^{-t^2/2}dt.$$

Especially, we can choose $g(q)=(\log q)^A, 0<A\leq 1$, not permitted by Theorem 0. In this case, one can also recover from our proof that $\mathcal{B}$ can be taken to be η-strong, so the above extends Theorem 0 to this missing range. We would like to note that the lower bound for the parameter, $\eta=0.01$ is just chosen for convenience. Indeed, we will see that one could conduct all of our proofs if we let $\eta=a, a\in(0,1]$ be any fixed positive number. After proving Theorem 1 one can use it to prove the following Corollary.

Corollary 1. *Let g be an increasing, divergent function such that*

- $g'(c)=O\left((g(Q))^{0.99}Q^{-0.01}\right), \forall c\in[Q,Q+Q^{0.01}]$ *as* $Q\to\infty$,
- $g(Q)\ll Q^{\epsilon}, \forall\epsilon>0$.

For each prime q, choose h_q such that $\log h_q/(\log g(q)) \to 0$ and $h_q \to \infty$ as $q \to \infty$. Then, there exists a set $\mathcal{B}$ of primes such that for any $\lambda \in \mathbb{R}$,

$$\lim_{\substack{q\to\infty \\ q\in\mathcal{B}}} \mathcal{M}_{g,q,h_q}(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} e^{-t^2/2} dt$$

and

$$\lim_{X\to\infty} \frac{|\mathcal{B} \cap [X, X+X^{\eta}|]}{|\mathcal{P} \cap [X, X+X^{\eta}]|} = 1, \text{ for all } \eta \in [0.525, 1].$$

Similarly as with Theorem 1, this recovers and extends on [3, Corollary 1.2]. The improvement to the method of [3] in our paper comes from an observation, that one can pass from an average of sums of Legendre symbols over primes in intervals of length Q^{η}, to expectation of these sums with Legendre symbol replaced with extended Rademacher random multiplicative functions. This fact, established by Harper in [14, Number Theory Result 3] (for $\eta = 1$), can be proved using appropriate Selberg's sieve weights. This allows us to save a factor of $(\log Q)$ lost in argument of [3], clean up the analysis (as $f(n)$ are perfectly orthogonal, so nicer to work with than Legendre symbols), and allows much shorter ranges for the averages over primes ($\eta \leq 0.5$). In the last part of the paper we gather some technical tools to extend the distributional results of [9], to allow for the starting point to be chosen uniformly at random from $[g(q)]$, where $g(q)$ is any function satisfying $Q^{1/2} \log Q = o(g(Q)^{1-\epsilon})$ for any $\epsilon > 0$.

Theorem 2. *Let $g(Q)$ be any increasing function satisfying $g(Q)^{1-\epsilon} \gg Q^{1/2} \log Q$ for any $\epsilon > 0$ and $q(Q) \leq Q, \forall Q$. For any prime q let $h_q \to \infty$ and $\frac{\log h_q}{\log(g(q))} \to 0$ as $q \to \infty$. Then with $\mathcal{M}_{g,q,h_q}(\lambda)$ defined as before, we have*

$$\mathcal{M}_{g,q,h_q}(\lambda) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} e^{-t^2/2} dt, \text{ as } q \to \infty, \tag{5}$$

for any fixed λ.

1.4. **Discussion on further improvements.** Our paper establishes the central limit theorem for character sums with starting point chosen uniformly at random from $[g(q)]$, for all the primes with $q^{1/2+\epsilon} \ll g(q) \ll q$ and for almost all the primes for practically arbitrary divergent $g(q) \ll q^{\epsilon}$, for all $\epsilon > 0$. We believe that our method would give similar kind of improvements, and at least a $\log q$ factor saving in another work of Basak, Nath and Zaharescu [4]. When it comes to results on all the primes, it remains an open problem to show that a central limit theorem holds with $g(q) = O(q^{1/2-c})$ for some fixed $c > 0$.

2. Ingredients

In this section we mention two results, which are necessary for us to establish Theorem 1 and Corollary 1. The first ingredient of our proof is based on the following result of Evertse and Silverman [10]:

Number Theory Fact 1. *(Evertse-Silverman, 1986)*
Set the following notation:

- *K is an algebraic number field of degree m.*
- *S is a finite set of places of K, containing the infinite places.*
- *$s = \#S$.*
- *R_S is the ring of S- integers of K.*
- *$f(X) \in R_S[X]$ is a polynomial of degree d with discriminant $disc(f) \in R_S^*$.*
- *L/K is an extension of degree M.*
- *$\kappa_n(L)$ is the n-rank of the ideal class group of L.*

For $n \geq 2$, let

$$V(R_S, f, n) = \{x \in R_S : f(x) \in K^{*n}\},$$

where K^{*n} denotes the elements of K which are perfect n'th powers.

(a) Let $n \geq 3, d \geq 2$ and assume that L contains at least two zeros of f. Then

$$\#V(R_S, f, n) \leq 17^{M(6m+s)} n^{2Ms+\kappa_n(L)}.$$

(b) Let $d \geq 3$, and assume that L contains at least three zeros of f. Then,

$$\#V(R_S, f, n) \leq 7^{M(4m+9s)} 4^{\kappa_2(L)}.$$

As noted by the authors of [3], this implies the following Corollary, which allows us to put uniform bounds on the number of solutions to equations that we will be working with further in the paper.

Corollary 2. *Let*

$$f(X) = X(X+\alpha_1)...(X+\alpha_k),$$

where $0 < \alpha_1 < ... < \alpha_k$ are integers and $k \geq 3$. Then the number of integers solutions to the equation $Y^2 = f(X)$ is $\leq 7^{13+9\omega(D(f))}$, where $\omega(D(f))$ denotes the number of distinct primes dividing the discriminant $D(f)$.

The second ingredient, needed for the proof of the second part of Corollary 1 is the strongest currently known statement establishing the order of primes in short intervals due to Baker, Harman and Pintz [2].

Number Theory Fact 2. *(Baker-Harman-Pintz, 2001)*
Let Q be sufficiently large. Then for all $\eta \in [0.525, 1]$ we have

$$\pi(Q + Q^\eta) - \pi(Q) \gg \frac{Q^\eta}{\log Q}. \tag{6}$$

3. Improvements

We start from the following version of [14, Number Theory Fact 3], which is a main component separating our approach from the one in [3]. As in [14], [9] and [3], we only need upper bounds for the "variance" terms that will arise in next section, instead of asymptotic equalities. We get them by employing the Selberg's sieve weights. This allows us to salvage the factor of $\log Q$ that was lost in [3]. Intuitively, the fact below is telling us, that when averaged over a set of prime moduli, the characters behave like the extended Rademacher random multiplicative functions.

Number Theory Fact 3. *Let $f(n)$ be an extended Rademacher random multiplicative function, $\eta \in [0.01, 1]$. Then, uniformly for any large $Q, N \leq Q^\eta$ and any complex coefficients $(a_n)_{n \leq N}$ we have*

$$\frac{\log Q}{Q^\eta} \sum_{Q \leq q \leq Q+Q^\eta} \left| \sum_{n \leq N} a_n \left(\frac{n}{q}\right) \right|^2 \ll \mathbb{E} \left| \sum_{n \leq N} a_n f(n) \right|^2 + \frac{1}{Q^{\eta/2}} \left(\sum_{n \leq N} |a_n| \sqrt{s(n)} \right)^2,$$

where the sum on the left hand side is over primes q and $s(n)$ denotes the squarefree part of n.

Proof. The proof is primarily an adaptation of the proof [14, Number Theory Fact 3], restricting it to short intervals. The same argument appeared earlier in [18, Lemma 9]. The main idea is to use the sieve theoretic machinery to upper bound the left hand side, in a way where we obtain some non-trivial cancellations by switching from sum over primes in short intervals, to sum over essentially all numbers. This allows us to use Pólya-Vinogradov, while keeping contributions coming from perfect squares under control. Note that for prime q and $n \leq q$, we have $\left(\frac{n}{q}\right) = \left(\frac{s(n)}{q}\right)$. Hence, we can rewrite the inner sum on the left hand side as

$$\sum_{\substack{s \leq N \\ s \text{ squarefree}}} \left(\frac{s}{q}\right) \sum_{\substack{s(n)=s \\ n \leq N}} a_n.$$

At the same time, note that from the definition of Rademacher random multiplicative function it follows that

$$\mathbb{E} \left| \sum_{n \leq N} a_n f(n) \right|^2 = \mathbb{E} \left| \sum_{\substack{s \leq N \\ s \text{ squarefree}}} f(s) \sum_{\substack{s(n)=s \\ n \leq N}} a_n \right|^2 = \sum_{\substack{s \leq N \\ s \text{ squarefree}}} \left| \sum_{\substack{s(n)=s \\ n \leq N}} a_n \right|^2.$$

Now, rewriting the left hand side of the original sum we have

$$\frac{\log Q}{Q^\eta}\sum_{Q\le q\le Q+Q^\eta}\left|\sum_{n\le N}a_n\left(\frac{n}{q}\right)\right|^2=\frac{\log Q}{Q^\eta}\sum_{Q\le q\le Q+Q^\eta}\sum_{\substack{s_1,s_2\le N\\ s\text{ squarefree}}}\left(\frac{s_1s_2}{q}\right)\left(\sum_{\substack{s(n)=s_1\\ n\le N}}a_n\right)\overline{\left(\sum_{\substack{s(n)=s_2\\ n\le N}}a_n\right)}. \tag{7}$$

We proceed by introduction of similar Selberg's sieve weights as [14], with slight difference coming from the fact that our short intervals can be a lot shorter than the ones there. We can make a choice of a sequence λ_e such that $\sum_{e|q}\lambda_e\ge\mathbb{I}\{p\mid q,\implies p\ge Q^{\eta/4}\}$ for all q, $\lambda_e=0$, for all $e>Q^{\eta/2}$, $\sum_{Q<q\le Q+Q^\eta}\sum_{e|q}\lambda_e\ll\frac{Q^\eta}{\log Q}$ and $\sum_e|\lambda_e|\ll\frac{Q^{\eta/2}}{(\log Q)^2}$. Using them, we upper bound (7) by

$$\frac{\log Q}{Q^\eta}\sum_{\substack{Q\le q\le Q+Q^\eta\\ q\text{ odd}}}\left(\sum_{e|q}\lambda_e\right)\sum_{\substack{s_1,s_2\le N\\ \text{squarefree}}}\left(\sum_{\substack{n\le N\\ s(n)=s_1}}a_n\right)\overline{\left(\sum_{\substack{n\le N\\ s(n)=s_2}}a_n\right)}\left(\frac{s_1s_2}{q}\right). \tag{8}$$

The main contribution in this kind of sums comes from the terms where s_1s_2 (and so n_1n_2) is a perfect square. Since s_1,s_2 are squarefree, this is counted by the diagonal summands $s_1=s_2$. Notice that we can bound these by the expectation of the sum over the random multiplicative functions and the sieve weights allow us to preserve essentially the density of the primes.

$$\sum_{\substack{s\le N\\ \text{squarefree}}}\left|\sum_{\substack{n\le N\\ s(n)=s}}a_n\right|^2\frac{\log Q}{Q^\eta}\sum_{\substack{Q\le q\le Q+Q^\eta\\ q\text{ odd}}}\left(\sum_{e|q}\lambda_e\right)\left(\frac{s^2}{q}\right)\ll\sum_{\substack{s\le N\\ \text{squarefree}}}\left|\sum_{\substack{n\le N\\ s(n)=s}}a_n\right|^2=\mathbb{E}\left|\sum_{n\le N}a_nf(n)\right|^2.$$

This step could be done without the help of sieve-theory, but it is necessary in bounding the contributions for $s_1\ne s_2$. In that case, s_1s_2 is not a perfect square and we want to use the cancellations coming from the fact that $(\frac{s_1s_2}{q})$ is a non-principal Dirichlet character of a conductor at most $4s_1s_2$. To get these cancellations, sieve theory allows us to pass from sums over primes to sums over essentially all integers in a given interval, at which point we can use the Pólya-Vinogradov bound. Hence, the off-diagonal contributions can be bounded by

$$\ll\frac{\log Q}{Q^\eta}\sum_{\substack{e\le Q^{\eta/2}\\ e\text{ odd}}}|\lambda_e|\sum_{\substack{s_1\ne s_1\le N\\ \text{squarefree}}}\left|\sum_{\substack{n\le N\\ s(n)=s_1}}a_n\right|\left|\sum_{\substack{n\le N\\ s(n)=s_2}}a_n\right|\left|\sum_{\substack{Q\le q\le Q+Q^\eta\\ q\text{ odd}\\ e|q}}\left(\frac{s_1s_2}{q}\right)\right|$$

$$\ll\frac{\log Q}{Q^\eta}\sum_{\substack{e\le Q^{\eta/2}\\ e\text{ odd}}}|\lambda_e|\sum_{\substack{s_1\ne s_2\le N\\ \text{squarefree}}}\left|\sum_{\substack{n\le N\\ s(n)=s_1}}a_n\right|\left|\sum_{\substack{n\le N\\ s(n)=s_2}}a_n\right|\sqrt{s_1s_2}\log(s_1s_2).$$

Since our weights satisfy $\sum_e|\lambda_e|\ll\frac{Q^{\eta/2}}{(\log Q)^2}$ we can check that this expression is suitably small. □

Remark. *Note that a more suitable choice of sieve weights would allow us to save a higher power of Q in the error term, but this will be enough for our goals.*

Notice that Number Theory Fact 3 immediately gives us the following bound

$$\frac{\log Q}{Q^\eta}\sum_{Q\le q\le Q+Q^\eta}\left|\sum_{n\le N}a_n\left(\frac{n}{q}\right)\right|^2\ll\mathbb{E}\left|\sum_{n\le N}a_nf(n)\right|^2+\frac{N}{Q^{\eta/2}}\left(\sum_{n\le N}|a_n|\right)^2. \tag{9}$$

This allows us to give improvements to [3, Lemma 3.2], which we rewrite in terms of an average of the primes in the interval $Q, Q+Q^{\eta}$. The point, is that the right hand side would then tend to zero. As we will see in the proof of Theorem 1, it is enough to prove this and subsequent results for $\eta = 0.01$, which we have chosen as our lower bound on η.

Lemma 1. *Let r be a fixed positive integer. Let g be an increasing, divergent function such that*

- $g'(c) = O\left((g(Q))^{0.99}Q^{-0.01}\right), \forall c \in [Q, Q+Q^{0.01}]$ *as* $Q \to \infty$,
- $g(Q) \ll Q^{\delta}, \forall \delta > 0$.

Let $r \leq h \leq (g(Q+Q^{0.01}))^{\frac{1}{2500r^2}}$. Then we have

$$\frac{\log Q}{Q^{0.01}} \sum_{\substack{Q \leq q \leq Q+Q^{0.01} \\ q \text{ prime}}} \left(\frac{1}{g(Q)} \sum_{1 \leq m \leq g(Q)} S_h^{2r}(m,q) - \mu_{2r}(h-\theta r)^r \right)^2 \ll (g(Q))^{-1+1/300} \tag{10}$$

and

$$\frac{\log Q}{Q^{0.01}} \sum_{\substack{Q \leq q \leq Q+Q^{0.01} \\ q \text{ prime}}} \left(\frac{1}{g(Q)} \sum_{1 \leq m \leq g(Q)} S_h^{2r-1}(m,q) \right)^2 \ll (g(Q))^{-1+1/300}, \tag{11}$$

where μ_{2r} are the $2r'$th moments of the standard normal and $\theta := \theta(h,r) \in [0,1]$ is defined in the proof.

Proof. We shall only prove the lemma for $2r$, the second statement follows the same argument. Our proof uses the same ides as the proof in [3], but first passing from the relevant sum to expectations. Note that the term $-\mu_{2r}(h-\theta r)^r$ can be taken to be the coefficient of $\left(\frac{1}{q}\right)$. Using (9) for $\eta = 0.01$ we have, for f being the extended Rademacher random multiplicative function

$$\frac{\log Q}{Q^{0.01}} \sum_{\substack{Q \leq q \leq Q+Q^{0.01} \\ q \text{ prime}}} \left(\frac{1}{g(Q)} \sum_{1 \leq m \leq g(Q)} S_h^{2r}(m,q) - \mu_{2r}(h-\theta r)^r \right)^2$$

$$\ll \mathbb{E}\left[\left| \frac{1}{g(Q)} \sum_{1 \leq m \leq g(Q)} \sum_{\alpha \in [h]^{2r}} f\left(\prod_{i=1}^{2r} m + \alpha_i \right) - \mu_{2r}(h-\theta r)^r \right|^2 \right] + \frac{(g(Q)+h)^{2r}}{Q^{0.005}} \left(\sum_{n \leq (g(Q)+h)^{2r}} |a_n| \right)^2,$$

with constants a_n implied by construction. We will want the right hand side of the above to go to zero, as that would imply further that, as $Q \to \infty$, we will be able to prove the gaussian behaviour for almost all the primes in the interval, using method of moments argument. First, we deal with bounding the expectation term. It is equal to

$$\frac{1}{(g(Q))^2} \mathbb{E}\left[\sum_{1 \leq m_1, m_2 \leq g(Q)} \sum_{\alpha, \beta \in [h]^{2r}} f\left(\prod_{i=1}^{2r} (m_1 + \alpha_i)(m_2 + \beta_i) \right) \right] \tag{12}$$

$$- 2\mu_{2r}(h-\theta r)^r \mathbb{E}\left[\frac{1}{g(Q)} \sum_{1 \leq m \leq g(Q)} \sum_{\alpha \in [h]^{2r}} f\left(\prod_{i=1}^{2r} (m + \alpha_i) \right) \right] + \mu_{2r}^2 (h-\theta r)^{2r}. \tag{13}$$

Starting now, we proceed to conduct the same analysis as [3]. Note however, that unlike Dirichlet characters, random multiplicative functions are perfectly orthogonal, which makes the analysis much easier. We start with the second expectation and proceed accordingly. In both cases, we will show that as in the case of Dirichlet characters, the most significant contributions to the sum will come from these n's which are perfect squares. Hence, we fix $\alpha \in [h]^{2r}$ and estimate for how many $m's$ in our range $\prod_{i=1}^{2r}(m+\alpha_i)$ is a perfect square. For this purpose consider the following algorithm:

Algorithm 1. *Consider α_1, the first coordinate of α. Then, in order, go through the remaining coordinates of α until we find α_j such that $\alpha_j = \alpha_i$. If no such α_j exists, proceed to do the same with α_2 and so on, until we go through all coordinates of the vector. If such coordinate exists proceed to delete the two coordinates from the vector, and repeat the procedure for the next non-deleted coordinate of α. The algorithm terminates when we reach the last non-deleted coordinate of α. Let us call the algorithm Ψ and the resulting new vector $\gamma = \Psi(\alpha) \in [h]^{2k}, k \leq r$.*

Notice that for any $m, \prod_{i=1}^{2r}(m+\alpha_i)$ is a perfect square iff $\prod_{i=1}^{2k}(m+\gamma_i)$ is one. Consider the three following cases:

- Assume $k \geq 2$. We want to calculate the number of $m \in \{1, ..., g(Q)\}$ such that $\prod_{i=1}^{2k}(m+\gamma_i)$ is a perfect square. Let $f(m) = \prod_{i=1}^{2k}(m+\gamma_i)$. Corollary 2 says that given distinct γ_i (as in our case), the amount of solutions to the diophantine equation
$$Y^2 = f(m) \tag{14}$$
is less than $7^{13+9\omega(D(f))}$, where ω is the distinct prime divisor function and $D(f)$ is the discriminant of the polynomial f. We have the well known bound (see for example [19])
$$\omega(n) \leq \frac{\log n}{\log\log n}(1+o(1))$$
Note that the RHS of the inequality is an increasing function of n. The discriminant of $f(m)$ in our case is given by $\prod_{i<j}^{2k}(\gamma_i - \gamma_j)^2$. Therefore, we can say that the amount of solutions to (14) is bounded by
$$\begin{aligned} &\ll e^{9\log 7\omega(D(f))} \ll \left(e^{(\log D(f))}\right)^{9\log 7/\log\log D(f)} \\ &= \left(\prod_{i<j}^{2k}(\gamma_i-\gamma_j)^2\right)^{9\log 7/\log\log(\prod_{i<j}^{2k}(\gamma_i-\gamma_j)^2)} \ll \left(h^{4r^2}\right)^{9\log 7/\log\log(h^{4r^2})} \\ &\ll_\epsilon (g(Q))^\epsilon, \end{aligned}$$
which we get recalling $h \leq g(Q+Q^{0.01})^{1/2500r^2}$, for any $\epsilon > 0$.
- Assume $k = 1$. In this case we want to count the number of solutions to the diophantine equation
$$Y^2 = (m+\alpha_i)(m+\alpha_j),$$
and we use an elementary divisor function argument to find a bound of sufficient order. Without loss of generality assume $\alpha_i < \alpha_j$. Let $D = m+\alpha_i, Z = \alpha_j - \alpha_i$. Then, we want to find to find the number of integer solutions to
$$Y^2 = D(D+Z) \iff 4Y^2 + Z^2 = (2D+Z)^2 \iff Z^2 = (2D+Z-2Y)(2D+Z+2Y).$$
Therefore, notice that given Z, D, Y must be such that $2D+Z-2Y = d, 2D+Z+2Y = Z^2/d$, for some divisor d of Z^2. Since for any such divisor, the above gives a unique solution, the number of such pairs is trivially bounded by $\tau(Z^2) \ll h^\epsilon \ll (g(Q))^\epsilon$, for any $\epsilon > 0$.
- Assume $k = 0$. In this case, $\prod_{i=1}^{2r}(m+\alpha_i)$ is a perfect square for all m.

Write $K(r,h)$ for the number of vectors arising in case 3. Now, notice that if n is not a perfect square we have
$$\mathbb{E}[a_n f(n)] = a_n \mathbb{E}[f(n)] = a_n \mathbb{E}[f(s)] = 0,$$
where s is the squarefree part of n. Therefore we have
$$\mathbb{E}\left[\frac{1}{g(Q)} \sum_{1\leq m\leq g(Q)} \sum_{\alpha\in[h]^{2r}} f\left(\prod_{i=1}^{2r} m+\alpha_i\right)\right] = K(r,h) + O(g(Q)^{-1+\epsilon} g(Q)^{2r/2500r^2}), \tag{15}$$

for any $\epsilon > 0$. Therefore, we need to estimate $K(r,h)$. This is the amount of $\alpha \in [h]^{2r}$ such that $\Psi(\alpha)$ vanishes. Notice that there are $\binom{h}{r}$ ways to choose exactly r numbers less than or equal to h without repetitions. Out of these numbers, we have r options as the choice for the first coordinate of α and $2r-1$ choices for the coordinate where it will appear again. Proceeding that way says that there are at least

$$\binom{h}{r} \cdot r \cdot (2r-1) \cdot (r-1) \cdot (2r-3) \cdot ... \cdot 2 \cdot 3 \cdot 1 = \binom{h}{r} \cdot r! \cdot (2r-1) \cdot (2r-3) ... \cdot 1 = h \cdot (h-1) \cdot ... \cdot (h-r+1) \cdot \mu_{2r}$$

ways of choosing α such that $\Psi(\alpha)$ vanishes, so $K(r,h) \geq \mu_{2r} h(h-1)...(h-r+1)$.
At the same time, note that if for that to happen, we need to have at most r distinct pairs of numbers from $\{1,2,...,h\}$ to form α. The number of ways of choosing them is h^r and after the choice, there is at most $(2r-1) \cdot (2r-3) \cdot ... \cdot 1$ ways of arranging them as r pairs. Therefore, we can see that $K(r,h) \leq \mu_{2r} h^r$. Hence we have, that for any h, r there exists $\theta := \theta(h,r) \in [0,1]$ so that

$$K(r,h) = \mu_{2r}(h-\theta r)^r. \tag{16}$$

The proof proceeds similarly when estimating the first expectation, but we need to modify the algorithm slightly as well as deal with some differences which arise in cases 1 and 3 of the algorithm. For convenience, we shall use the same notation. We shall now deal with the first expectation term in a similar way, so we want to count the contribution of the perfect squares $\prod_{i=1}^{2r}(m_1+\alpha_i)(m_2+\beta_i)$. As before, we do it through counting m_1 which form perfect squares for fixed vectors $\alpha, \beta \in [h]^{2r}$ and fixed m_2. Consider the following algorithm:

Algorithm 2. *Let $m := m_2 - m_1$ and let $\xi = (m+\beta_1, m+\beta_2, ..., m+\beta_{2r})$. Then, adjoin the vectors α, ξ to create the vector $(\alpha_1, ..., \alpha_{2r}, \xi_1, ..., \xi_{2r}) \in [m+h]^{4r}$. Then, proceed with the same argument and map Ψ as before, but this time the vector is of length $4r$, and the output is vector $\Psi((\alpha,\xi)) =: \gamma \in [m+h]^{2k}, k \leq 2r$. Notice that $\prod_{i=1}^{2r}(m_1+\alpha_i)(m_2+\beta_i)$ is a square iff $\prod_{i=1}^{2k}(m_1+\gamma_i)$ is.*

As before, we follow with case by case analysis.

- Assume $k \geq 2$. Then, again by Corollary 2 we can bound the number of solutions to the diophantine equation

  $$Y^2 = f(m_1) = (m_1+\gamma_1)...(m_1+\gamma_{2k}).$$

  As above, we use Number Theory Fact 2 to bound the number of these solutions by

  $$\begin{aligned} &\ll e^{9\log 7 \omega(D(f))} \ll \left(e^{(\log D(f))}\right)^{9\log 7/\log\log D(f)} \\ &\ll \left((g(Q)+h)^{8r^2} h^{8r^2}\right)^{9\log 7/\log\log\left((g(Q)+h)^{8r^2}h^{8r^2}\right)} \\ &\ll_{\epsilon,r} (g(Q))^{\epsilon}, \end{aligned}$$

  for any $\epsilon > 0$.
- Assume $k = 1$. This works exactly as before, so amount of solutions m_1 to this equation can also be bounded by $(g(Q))^{\epsilon}$.
- Assume $k = 0$. In this case, note that if $|m| \geq h+1$ then $\prod_{i=1}^{2r}(m_1+\alpha_i)$ and $\prod_{i=1}^{2r}(m_2+\beta_i)$ both are perfect squares, each with at most r distinct coefficients (since $k=0$). Therefore, the number of vectors α, β for which both products are perfect squares for all pairs of m_1, m_2 is exactly $\mu_{2r}^2(h-\theta r)^{2r}$. If $|m| \leq h$ then even if we assume that there is a solution for any vectors α, β and any m_2, but for each such choice there is at most $O(h)$ choices for m_1.

Now, notice similarly to before, by orthogonality of Rademacher random multiplicative functions that the expectation for non-square products will have null contribution. Using (15) and summing over all possible

α, β for each case discussed after Algorithm 2 implies that we can bound the original expectation (12) by

$$\begin{aligned}\mu_{2r}^2(h-\theta r)^{2r} &+ O\left((g(Q))^{-1+\epsilon+4(r^2+1)/2500r^2}\right)\\ &- 2\mu_{2r}^2(h-\theta r)^{2r} + O\left((g(Q)^{-1+\epsilon+3r^2/2500r^2}\right)\\ &+ \mu_{2r}^2(h-\theta r)^{2r} \ll (g(Q))^{-1+1/300}.\end{aligned}$$

This concludes dealing with bounds on expectations. What is left now is to complete the proof of our Lemma is to show that the term

$$\frac{(g(Q)+h)^{2r}}{Q^{0.005}}\left(\sum_{n\leq (g(Q)+h)^{2r}}|a_n|\right)^2 \tag{17}$$

goes to zero at least as quickly. Notice that for $n>1, |a_n|$ can be explicitly described as the count of representations of n as a product $\prod_{i=1}^{2r}(m+\alpha_i)$ for $1\leq m\leq g(Q), \alpha\in[h]^{2r}$, later normalised by $\frac{1}{g(Q)}$. Since every choice of α and m represents some n, their sum (still averaged over all m) is easily seen to be

$$\sum_{1<n\leq (g(Q)+h)^{2r}}|a_n| = h^{2r}.$$

Additionally, we pick up the $-\mu_{2r}(h-\theta r)^r$ term for $n=1$, so that

$$\frac{(g(Q)+h)^{2r}}{Q^{0.005}}\left(\sum_{n\leq (g(Q)+h)^{2r}}|a_n|\right)^2 \ll \frac{(g(Q)+h)^{2r}(g(Q))^{2r/(2500r^2)}}{Q^{0.005}} \ll \frac{(g(Q))^{2r(1+1/(2500r^2))}}{Q^{0.005}},$$

which is of sufficiently small order. This finishes our argument for $2r$. Notice that in case $2r-1$, the reduction algorithms never give $k=0$ as $2r-1$ is odd and so the term with μ doesn't arise, which proves the lemma fully. □

This implies the following Corollary, which can be viewed as an equivalent to Chebyshev's inequality.

Corollary 3. *Let r be a fixed positive integer. Let g be an increasing, divergent function such that*

- $g'(c) = O\left((g(Q))^{0.99}Q^{-0.01}\right), \forall c\in[Q, Q+Q^{0.01}]$ *as* $Q\to\infty$,
- $g(Q)\ll Q^{\delta}, \forall\delta>0$.

For each prime $q\in[Q,Q+Q^{0.01}]$ choose h_q such that

$$r\leq h_q\leq g(q)^{\frac{1}{2500r^2}}.$$

Define

$$\mathcal{E}_1(Q;g,r) := \left\{q\in\mathcal{P}\cap[Q,Q+Q^{0.01}] : \left|g(Q)^{-1}\sum_{m\leq g(Q)}S_{h_q}^{2r}(m,q) - \mu_{2r}(h_q-\theta_q r)^r\right| \geq \frac{1}{g(q)^{1/8}}\right\}$$

and

$$\mathcal{E}_2(Q;g,r) := \left\{q\in\mathcal{P}\cap[Q,Q+Q^{0.01}] : \left|g(Q)^{-1}\sum_{m\leq g(Q)}S_{h_q}^{2r-1}(m,q)\right| \geq \frac{1}{g(q)^{1/8}}\right\}.$$

Then

$$|\mathcal{E}_1\cup\mathcal{E}_2| \ll \frac{Q^{0.01}}{(\log Q)g(Q)^{74/100}}.$$

Proof. We have

$$\sum_{q\in\mathcal{E}_1} g(q)^{1/4}\left(\frac{1}{g(Q)}\sum_{1\le m\le g(Q)} S_{h_q}^{2r}(m,q)-\mu_{2r}(h_q-\theta_q r)^r\right)^2$$
$$\ll g(Q)^{1/4}\sum_{\substack{Q\le q\le Q+Q^{0.01}\\ q \text{ prime}}}\left(\frac{1}{g(Q)}\sum_{1\le m\le g(Q)} S_{h_q}^{2r}(m,q)-\mu_{2r}(h_q-\theta_q r)^r\right)^2.$$
$$= g(Q)^{1/4}\sum_{h\le g(Q+Q^{0.01})^{1/(2500r^2)}}\sum_{\substack{Q\le q\le Q+Q^{0.01}\\ q \text{ prime}\\ h_q=h}}\left(\frac{1}{g(Q)}\sum_{1\le m\le g(Q)} S_{h}^{2r}(m,q)-\mu_{2r}(h-\theta r)^r\right)^2$$
$$\ll g(Q)^{1/4}g(Q)^{1/(2500r^2)}\sum_{\substack{Q\le q\le Q+Q^{0.01}\\ q \text{ prime}}}\left(\frac{1}{g(Q)}\sum_{1\le m\le g(Q)} S_{h}^{2r}(m,q)-\mu_{2r}(h-\theta r)^r\right)^2$$
$$\ll \frac{Q^{0.01}}{\log Q g(Q)^{74/100}},$$

where the last line follows in an obvious way from Lemma 1. At the same time,

$$\sum_{q\in\mathcal{E}_1} g(q)^{1/4}\left(\frac{1}{g(q)}\sum_{1\le m\le g(q)} S_{h}^{2r}(m,q)-\mu_{2r}(h_q-\theta_q r)^r\right)^2 \gg |\mathcal{E}_1|,$$

which implies

$$|\mathcal{E}_1|\ll \frac{Q^{\eta}}{(\log Q)g(Q)^{74/100}}.$$

The same result for $\mathcal{E}_2$ implies the Corollary. □

4. Proof of Theorem 1 and Corollary 1

In this section we put it all together in a method of moments argument to prove Theorem 1.

Proof. Consider $\eta\in[0.01,1]$. We begin with an observation, that if a subset $\mathcal{B}\subseteq\mathcal{P}$ of primes is η_1-sufficient, it is also η_2-sufficient for any $\eta_1<\eta_2\le 1$. One can prove it with a simple covering argument of the set $[X,X+X^{\eta_2}]$ with shorter intervals of form $[Q,Q+Q^{\eta_1}]$. Therefore, it is enough for our purposes to prove the theorem when $\eta=0.01$. For each $k\in\mathbb{N}$ set $R_k=2^k$. By Corollary 3 there exists an increasing sequence of integers Q_k such that for each k, for all $Q\ge Q_k$ and all $r\le R_k$

$$\left|\frac{1}{g(Q)}\sum_{m\le g(Q)} S_{h_q}^{2r}(m,q)-\mu_{2r}(h_q-\theta_q r)^r\right|<\frac{1}{g(q)^{1/8}} \tag{18}$$

and

$$\left|\frac{1}{g(Q)}\sum_{m\le g(Q)} S_{h_q}^{2r-1}(m,q)\right|<\frac{1}{g(q)^{1/8}}, \tag{19}$$

for all but

$$\le \frac{Q^{0.01}}{(\log Q)g(Q)^{7/10}}$$

many primes in the interval $[Q,Q+Q^{0.01}]$. To apply the Corollary, we use the assumption that $h_q\to\infty$ as $q\to\infty$ of Theorem 1. Notice that we used some factors of $g(Q)$ to get rid of Vinogradov notation in place of an inequality, so with choice $g(Q)=(\log Q)^A, A>0$ we could preserve the property of being η-strong

of [3]. For any $Q_k < Q_{k+1}$ let $\mathcal{B}_k$ be the set of primes $q \in [Q_k, Q_{k+1}]$, such that $q \in [N, N + N^{0.01}]$ for some $N \in [Q_k, Q_{k+1}]$ and (18),(19) hold with $Q = N$ for all $r \leq R_k$. Then for all $q \in \mathcal{B}_k$, and $j \in \mathbb{N}$ we have that there exists N as described before, such that

$$g(q)^{-1} \sum_{m \leq g(q)} \Big((h_q)^{-1/2} S_{h_q}(m,q)\Big)^j$$
$$= (g(N) + g'(c)(q-N))^{-1} \sum_{m \leq g(N)+g'(c)(q-N)} \Big((h_q)^{-1/2} S_{h_q}(m,q)\Big)^j,$$

for some $c \in (N, N + N^{0.01})$. By Corollary 3, assumptions on g', (18) and (19), this is equal to

$$= g(N)^{-1}(1 + o(g(N)^{-0.01}))^{-1} \sum_{m \leq g(N)(1+o(g(N)^{-0.01}))} \Big((h_q)^{-1/2} S_{h_q}(m,q)\Big)^j \tag{20}$$

$$= g(N)^{-1}(1 + o(1))^{-1} \sum_{m \leq g(N)} \Big((h_q)^{-1/2} S_{h_q}(m,q)\Big)^j + o(1) \to \mu_j, \tag{21}$$

as $k \to \infty$, where μ_j is the j'th moment of a standard Gaussian random variable. Define

$$\mathcal{B} := \cup_{k=1}^{\infty} \mathcal{B}_k.$$

We can immediately see that by construction, $\mathcal{B}$ is an 0.01-sufficient set of prime, so also η-sufficient for all $\eta \in [0.01, 1]$. Consider a sequence $(q_n) \in \mathcal{B}$. Then, notice that by (20) we have for any fixed positive j

$$g(q_n)^{-1} \sum_{m \leq g(q_n)} \Big((h_{q_n})^{-1/2} S_{h_{q_n}}(m,q_n)\Big)^j \to \mu_j, \tag{22}$$

as $n \to \infty$. For convenience of presentation define

$$\mathcal{N}_q(s) := \big|\big\{1 \leq m \leq g(q) : S_{h_q}(m,q) \leq s\big\}\big|.$$

This is a non-decreasing function of s, constant except for discontinuities at certain integral values of s. Also

$$\mathcal{N}_q(s) = \begin{cases} 0 & \text{for} \quad s < -h_q \\ \lfloor g(q) \rfloor & \text{for} \quad s > h_q. \end{cases}$$

We can recognise

$$\mathcal{M}_{g,q,h_q}(\lambda) = \frac{1}{g(q)} \mathcal{N}(\lambda h_q^{1/2}).$$

Note that for any q, h_q this defines a natural cumulative distribution function Ψ_q on the real line, equipped with a probability measure. The goal now, is to show that as $n \to \infty$ for $(q_n) \in \mathcal{B}$, this distribution converges weakly to a Gaussian. Now, rewriting (22) with the new notation gives

$$g(q_n)^{-1} \sum_{s=-h_{q_n}}^{h_{q_n}} (h_{q_n}^{-1/2} s)^j \left(\mathcal{N}_{q_n}(s) - \mathcal{N}_{q_n}(s-1)\right) = \int_{-\infty}^{\infty} t^j d\Psi_{q_n} \to \mu_j,$$

as $n \to \infty$ for any fixed positive j. This implies, that the random variable X coming from the distribution associated to the law of distribution $\mathbb{P}(X \leq \lambda) = M_{g,q,h_q}(\lambda)$ has identical moments to those of the standard normal. From a well-known probabilistic result, we know that if moments of a sequence of random variables converge to moments of the standard normal, this completely determines their limiting distribution as that of a standard normal. In fact, this is true also in more general setting, where the limiting moments are finite and unique to some probability distribution. Therefore, we have the convergence of cumulative distribution functions

$$\mathcal{M}_{g,q_n,h_{q_n}}(\lambda) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} e^{-t^2/2} dt, \tag{23}$$

which completes the proof of Theorem 1. To prove Corollary 1 simply apply Number Theory Fact 2. □

5. Results for all the primes

For the remainder of this section let us forget about previous assumptions and let $g(q)$ be an arbitrary increasing, divergent function. Davenport and Erdős [9] and Lamzouri [15] establish (4) for all primes provided $h_q \to \infty$, but $h_q = q^{o(1)}$, with the starting points of the sum being uniform in $\{1, ..., q\}$, while our method establishes (4) for almost all the primes, with $h_q \to \infty, h_q = g(q)^{o(1)}$ and the intervals of size $g(q)$, for any function $g(q)$ as in our results. In fact, if $g(q)$ is small enough (e.g. of size $(\log q)^{1/2}$), then there exist primes q for which $(\frac{n}{q}) = 1$ for all $n \leq g(q)$, which makes it impossible to improve Theorem 1 to cover the whole set of primes. This suggests, that there is some threshold for the size of the interval $g(q)$, where (4) changes from convergence for almost all, to all the primes. This problem takes us back to the reason why we even considered averaging over primes in intervals: To get distributional results for sums as short as ours, there is the need to employ cancellations coming from changes in sign of $(\frac{n}{q})$, when n is not a perfect square and q varies. If we consider just one prime q, we are forced to understand the structure and get necessarily good cancellations in sums of the form

$$\sum_{1\leq m\leq g(q)} S_{h_q}^{2r}(m) = \sum_{1\leq m\leq g(q)} \sum_{\alpha_1=1}^{h_q} \cdots \sum_{\alpha_{2r}=1}^{h_q} \left(\frac{(m+\alpha_1)(m+\alpha_2)...(m+\alpha_{2r})}{q}\right), \tag{24}$$

where

$$S_{h_q}(x) := \sum_{x\leq n\leq x+h_q} \left(\frac{n}{q}\right).$$

If we consider fixed $\alpha \in [h_q]^{2r}$, as in the proof of Lemma 1, we see that the main contribution comes from α, which all components appear inside the vector an even amount of times, in which case $f(m) := \prod_{i=1}^{2r}(m+\alpha_i)$ is a square of a polynomial. This is case 3 of either of the algorithms of that proof. On the other hand, when dealing with the case when this product is not of this form, there is no prime averaging to come to our rescue, but we still need to show that

$$\sum_{1\leq m\leq g(q)}{}' S_{h_q}^{2r}(m) = \sum_{1\leq m\leq g(q)}{}' \sum_{\alpha_1=1}^{h_q} \cdots \sum_{\alpha_{2r}=1}^{h_q} \left(\frac{(m+\alpha_1)(m+\alpha_2)...(m+\alpha_{2r})}{q}\right) = o(g(q)h_q^r),$$

where the sums above are over $\alpha \in [h_q]^{2r}$ for which $f(m)$, as defined above is not a square of a polynomial. In the original paper of Davenport and Erdős, they use the Weil bound [21], which allows them to show that if $g(q) = q$, the above holds. We have

$$\sum_{1\leq m\leq q} \left(\frac{(m+n_1)(m+n_2)...(m+n_{2r})}{q}\right) = O(q^{1/2}),$$

for these n's that do not contribute to case 3 in any of algorithms in Lemma 1. In this section, we use a fact about incomplete character sums, that follows from the Weil bound [21] and show that the result on distribution of sums of Legendre character for all the primes, holds true also if $g(q) \gg (q^{1/2}\log q)^{1/(1-\epsilon)}$, for any $\epsilon > 0$. The main ingredient that we use is [17, Theorem 2].

Number Theory Fact 4. *(Mauduit - Sárközy, 1997)*
Let q be a prime number, χ a non-principal character modulo q of order d, and $f(x) \in \mathbb{F}_q[x]$ have degree k and a factorization

$$f(x) = b(x-x_1)^{d_1}(x-x_2)^{d_2}...(x-x_s)^{d_s},$$

with $x_i \neq x_j$ for $i \neq j$ in $\overline{\mathbb{F}_q}$. Then, if $(d, d_1, d_2, ..., d_s) = 1$, then for any real numbers $X, 0 < Y \leq q$ we have

$$\left|\sum_{X<n\leq X+Y} \chi(f(n))\right| < 9kq^{1/2}\log q. \tag{25}$$

The proof of the Theorem 2 is essentially the same method of moments proof as that of [9] and is much simpler than the one presented for almost all the primes before.

5.1. **Proof of Theorem 2.**

Proof. The main idea of the method of moments remains unchanged, except this time we do not allows ourselves to average over different primes. We follow the original argument due to Davenport and Erdős [9]. First, for any $r \in \mathbb{N}_{>0}$ and any sufficiently large prime q we want to show that

$$\frac{1}{g(q)} \sum_{m \leq g(q)} S_{h_q}^{2r}(m,q) - \mu_{2r}(h-\theta_2 r)^r \tag{26}$$

and

$$\frac{1}{g(q)} \sum_{m \leq g(q)} S_{h_q}^{2r-1}(m,q) \tag{27}$$

can be bounded sufficiently well in terms of some function of q. Using that $h_q \to \infty$ as $q \to \infty$ we can assume that $r \leq h$. We first deal with the case when the exponent of the sum is $2r$. Write

$$\sum_{m \leq g(q)} S_{h_q}^{2r} = \sum_{m \leq g(q)} \sum_{\alpha_1=1}^{h_q} \cdots \sum_{\alpha_{2r}=1}^{h_q} \left(\frac{(m+\alpha_1)(m+\alpha_2)...(m+\alpha_{2r})}{q} \right)$$

First, consider only these $\alpha \in [h_q]^{2r}$ for which the product above is a square of a polynomial in $\mathbb{F}_q$. These are exactly the α that constitute the case 3 of the algorithms described in the proof of Lemma 1. For these α, using the argument of Lemma 1, there exists $\theta_1 \in [0,1]$ such that

$$\sum_{m \leq g(q)} \left(\frac{(m+\alpha_1)(m+\alpha_2)...(m+\alpha_{2r})}{q} \right) = g(q) - \theta_1 r. \tag{28}$$

As in the proof of Lemma 1, we estimate the number of $\alpha \in [h_q]^{2r}$ like this to be $K(r,h_q) = \mu_{2r}(h_q - \theta_2 r)^r$ for some $\theta_2 \in [0,1]$. For any other α under consideration, we perform on it the algorithm of Lemma 1. If we define $\gamma \in [h_q]^{2k}$ to be the newly acquired vector with $k \leq r$, the problem becomes that of dealing with estimating the sum

$$\sum_{m \leq g(q)} \left(\frac{(m+\gamma_1)(m+\gamma_2)...(m+\gamma_{2k})}{q} \right). \tag{29}$$

If we write

$$f(m) = (m+\gamma_1)(m+\gamma_2)...(m+\gamma_{2k})$$

then (29) is susceptible to the Number Theory Fact 4, which gives us

$$\left| \sum_{m \leq g(q)} \left(\frac{(m+\alpha_1)(m+\alpha_2)...(m+\alpha_{2r})}{q} \right) \right| = \left| \sum_{m \leq g(q)} \left(\frac{(m+\gamma_1)(m+\gamma_2)...(m+\gamma_{2k})}{q} \right) \right| < 18kq^{1/2} \log q. \tag{30}$$

Therefore, combining (28), (30) and the triangle inequality we get

$$\sum_{m \leq g(q)} S_{h_q}^{2r}(m,q) - \mu_{2r}(h_q - \theta_2 r)^r g(q) \ll h_q^{2r} q^{1/2} \log q. \tag{31}$$

Similarly, since in the case when the exponent of the sum $S_h(m,q)$ is $2r-1$, the product

$$f(m) = (m+\alpha_1)(m+\alpha_2)...(m+\alpha_{2r-1})$$

is never a square of a polynomial in $\mathbb{F}_p$, the second part of the above argument shows

$$\sum_{m \leq g(q)} S_{h_q}^{2r-1}(m,q) \ll h_q^{2r-1} q^{1/2} \log q. \tag{32}$$

Therefore, combining (31) and (32) we have that for all $r \in \mathbb{N}_{>0}$ we have

$$\frac{1}{g(q)} \sum_{m \leq g(q)} (h_q^{-1/2} S_h(m,q))^r = \mu_r + O(g(q)^{-1} h_q^{r/2} q^{1/2} \log q), \tag{33}$$

where μ_r is the r'th moment of the standard normal, so that as long as $h_q^r q^{1/2} \log q = o(g(q))$, we let $q \to \infty$ to give

$$\frac{1}{g(q)} \sum_{m \leq g(q)} (h_q^{-1/2} S_h(m,q))^r \to \mu_r,$$

as $q \to \infty$, which by method of moments finishes the proof. □

One might hope that somehow the state of the art of distributional results should reflect the state of bounds on the actual size of the least NQR, in which case we should be able to deal with $g(q) \approx q^{1/4}$, following the Burgess' bound [7]. In the argument as above, an improvement of this type would be used in (30). Currently, it is bounded using Number Theory Fact 4, which relies on the Weil bound. Unfortunately, generalising the results of Burgess to a case where the sum of characters runs over $f(n)$ instead of n, where f is a squarefree polynomial is not an easy task. One can find notable results that are applicable to our situation in the works of Bourgain and Chang [6] and Pierce and Xu [20]. Unfortunately, both of these papers give bounds for (30) which grow in powers of $g(q)$ along with the degree of the polynomial f. This would possibly allow us to set $g(q) = q^{1/2-o(1)}$, but extending Theorem 2 to $g(q)$ of smaller order remains a hard problem.

Acknowledgements. The author would like to thank his supervisor Adam Harper for recommending this problem, guidance and comments on the previous versions of this paper.

References

1. Nesmith Ankeny, *The least quadratic non residue*, Annals of Mathematics **55** (1952), 65–72.
2. Roger. C. Baker, Glyn Harman, and János Pintz, *The Difference Between Consecutive Primes, II*, Proceedings of the London Mathematical Society **83** (2001), no. 3, 532–562.
3. Debmalya Basak, Kunjakanan Nath, and Alexandru Zaharescu, *Gaussian phenomena for small quadratic residues and non-residues*, Transactions of the American Mathematical Society **376** (2023), no. 5, 3695–3724 (English (US)), Publisher Copyright: © 2023 American Mathematical Society.
4. ———, *Poisson behavior for chains of small quadratic non-residues*, International Mathematics Research Notices **2024** (2023), 3356–3390.
5. Jonathan Bober, Leo Goldmakher, Andrew Granville, and Dimitris Koukoulopoulos, *The frequency and the structure of large character sums*, J. Eur. Math. Soc. 20 **no.7** (2017), 1759–1818.
6. Jean Bourgain and Mei-Chu Chang, *On a multilinear character sum of Burgess*, Comptes Rendus Mathematique **348** (2010), no. 3, 115–120.
7. David A. Burgess, *The distribution of quadratic residues and non-residues*, Mathematika **4** (1957), no. 2, 106–112.
8. O-Yeat Chan, Geumlan Choi, and Alexandru Zaharescu, *A multidimensional version of a result of Davenport-Erdős.*, Journal of Integer Sequences [electronic only] **6** (2003), no. 2, Art. 03.2.6, 9 p., electronic only–Art. 03.2.6, 9 p., electronic only (eng).
9. Harold Davenport and Paul L. Erdős, *The distribution of quadratic and higher residues*, Publ. Math. Debrecen **2** (1952), 252–265.
10. Jan-Hendrik Evertse and Joseph H. Silverman, *Uniform bounds for the number of solutions to $Y^n = f(X)$*, Mathematical Proceedings of the Cambridge Philosophical Society **100** (1986), no. 2, 237–248.
11. Andrew Granville and Kannan Soundararajan, *Large character sums*, Journal of the American Mathematical Society **14** (1999), 365–397.
12. ———, *Large character sums: Pretentious characters and the Pólya-Vinogradov theorem*, Journal of the American Mathematical Society **20** (2005), 357–384.
13. ———, *Large character sums: Burgess's theorem and zeros of L-functions*, arXiv: Number Theory (2015).
14. Adam J Harper, *A note on character sums over short moving intervals*, Journal of the Institute of Mathematics of Jussieu **24** (2025), no. 4, 1395–1427.
15. Youness Lamzouri, *The distribution of short character sums*, Mathematical Proceedings of the Cambridge Philosophical Society **155** (2011), 207 – 218.
16. Youness Lamzouri, Xiannan Li, and Kannan Soundararajan, *Conditional bounds for the least quadratic non-residue and related problems*, Mathematics of Computation **84** (2015), no. 295, 2391–2412.

17. Christian Mauduit and András Sárközy, *On finite pseudorandom binary sequences I: Measure of pseudorandomness, the Legendre symbol*, Acta Arithmetica **82** (1997), no. 4, 365–377 (eng).
18. Hugh L. Montgomery and Robert C. Vaughan, *Mean Values of Character Sums*, Canadian Journal of Mathematics **31** (1979), no. 3, 476–487.
19. ———, *Multiplicative Number Theory I: Classical Theory*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2006.
20. Lillian B. Pierce and Junyan Xu, *Burgess bounds for short character sums evaluated at forms*, Algebra & Number Theory **14** (2020), no. 7, 1911–1951.
21. André Weil, *Sur les courbes algébriques et les variétés qui s'en déduisent*, 1948.

University of Warwick, Mathematics Institute, Zeeman Building, University of Warwick, Coventry CV4 7AL
Email address: Pawe.Nosal@warwick.ac.uk
URL: https://sites.google.com/view/pawelnosalmaths