

THE DISTRIBUTION OF VALUES OF $\frac{L'}{L}(1/2 + \epsilon, \chi_D)$

ALIA HAMIEH AND RORY MCCLENAGAN

ABSTRACT. We determine the limiting distribution of the family of values $\frac{L'}{L}(1/2 + \epsilon, \chi_D)$ as D varies over fundamental discriminants. Here, $0 < \epsilon < \frac{1}{2}$, and χ_D is the real character associated with D . Moreover, we also establish an upper bound for the rate of convergence of this family to its limiting distribution. As a consequence of this result, we derive an asymptotic bound for the small values of $\left| \frac{L'}{L}(1/2 + \epsilon, \chi_D) \right|$.

1. INTRODUCTION

Many mathematicians have studied the distribution of values of L -functions in the critical strip. Some of the earliest results on this topic are due to Bohr-Jessen ([2], [3]) and Jessen-Wintner [15]. These authors obtain the distribution function of $\log \zeta(\sigma + it)$ for a fixed $\sigma > \frac{1}{2}$ and established several analytic properties of this function. Another influential result on this topic is Selberg's central limit theorem [25] which states that values $\log \zeta(\frac{1}{2} + it)$ have an approximately two-dimensional Gaussian distribution. Distribution problems for several other families of L -functions have been considered from various points of view over the last 70 years. Consider for example the family of L -functions associated with real quadratic characters χ_D where $\chi_D(n)$ is the Kronecker symbol $(\frac{D}{n})$. Chowla and Erdos [4] proved that the family $\{L(\sigma, \chi_D) : D > 0, D \equiv 0, 1 \pmod{4}\}$, for a fixed $\sigma > \frac{3}{4}$, admits a continuous and strictly increasing asymptotic distribution function. Elliott also considered this particular family of L -values in a series of papers in the 1970's, thereby improving on the previous body of work. One of Elliott's results in this direction is the following theorem [7, Theorem 1].

Theorem 1.1. *Let σ_0 be a real number that satisfies $\frac{1}{2} + (\log \log \log N)^{-\frac{1}{2}} \leq \sigma_0 \leq 1$. There exist distribution functions $F(s, z)$ so that the estimate*

$$\frac{1}{\pi(N)} \#\{p \leq N : p \text{ prime}, |L(s, \chi_p)| \leq e^z\} = F(s, z) + O((\log \log \log N)^{-2}), \quad \text{as } N \rightarrow \infty,$$

holds uniformly for all s in $R_N = \{s : \sigma_0 \leq \sigma \leq \frac{5}{4}, |\Im(s)| < N^{\frac{1}{13}(2\sigma-1)}\}$ and for all real numbers z . For each value of s , the function $F(s, z)$ is infinitely differentiable with respect to z . Moreover, the characteristic function $\varphi(s, \tau)$ of $F(s, z)$ has the form

$$\varphi(s, \tau) = \prod_{p \text{ prime}} \frac{1}{2} \left(\exp(-i\tau |\log(1 + p^{-s})|) + \exp(-i\tau |\log(1 - p^{-s})|) \right),$$

Date: September 21, 2021.

2010 Mathematics Subject Classification. 11R42, 11M38, 11M41.

Key words and phrases. value-distribution, logarithmic derivatives of L -functions, quadratic characters.

Research of the first author is partially supported by NSERC Discovery Grant. Research of the second author was partially supported by NSERC Undergraduate Summer Research Awards.

and satisfies the bound $\varphi(s, \tau) \ll \exp\left(-\frac{c_1|\tau|^{\frac{1}{\sigma}}}{2\sigma-1}\right)$ for all $\sigma > \frac{1}{2}$.

In an important paper [8], Granville and Soundararajan studied the distribution of large values of $L(1, \chi_D)$ as D varies over all fundamental discriminants. One of their results implies that the proportion of fundamental discriminants D with $|D| \leq x$ such that $L(1, \chi_D) \geq e^\gamma \tau$ decays doubly exponentially in $\tau = \log \log x$. In [8], the authors compare the distribution of the values of $L(1, \chi_D)$ with the distribution of the probabilistic model $L(1, X)$ for some carefully chosen random variable X .

The idea of comparing the distribution of values of $L(1, \chi_D)$ to a random model precedes [8]. For instance, it appears in the work of Elliott [5, 7] where he reduced the problem to a probability problem concerning sums of independent random variables on a finite probability space.

Lamzouri explored this line of research even further and established a framework for studying the distribution of large values of various families of L -functions inside the critical strip (see for example [17] and [18]). In [19], Lamzouri studied the distribution of large values of $\frac{L'}{L}(1, \chi_D)$. These values have great arithmetic significance as they are directly related to the values of the Euler-Kronecker constants of the quadratic fields $\mathbb{Q}(\sqrt{D})$. In fact, the distribution of values of logarithmic derivatives of Dirichlet L -functions in the critical strip was initiated by Ihara and Matsumoto (see for example [11], [12], [13] and [14]). Their approach, however, does not follow the probabilistic framework employed in [8], [10], [17], [18] and [19] among other papers. Instead, it is based on classical results such as Lévy's continuity theorem and Jessen-Wintner theory of infinite convolutions of distribution functions.

Following the method employed in [14], Mourtada and Murty proved the following result (see [22, Theorem 2]).

Theorem 1.2. *Let $\sigma > \frac{1}{2}$, and assume the GRH. Let $\mathcal{F}(N)$ denote the set of fundamental discriminants in the interval $[-N, N]$. Then, there exists a probability density function M_σ , such that*

$$\lim_{N \rightarrow \infty} \frac{1}{|\mathcal{F}(N)|} \left| \left\{ D \in \mathcal{F}(N); \frac{L'}{L}(\sigma, \chi_D) \leq z \right\} \right| = \int_{-\infty}^z M_\sigma(t) dt.$$

Moreover, the characteristic function $\varphi_{F_\sigma}(y)$ of the asymptotic distribution function $F_\sigma(z) = \int_{-\infty}^z M_\sigma(t) dt$ is given by

$$\varphi_{F_\sigma}(y) = \prod_p \left(\frac{1}{p+1} + \frac{p}{2(p+1)} \exp\left(-\frac{iy \log p}{p^\sigma - 1}\right) + \frac{p}{2(p+1)} \exp\left(\frac{iy \log p}{p^\sigma + 1}\right) \right).$$

The purpose of this paper is to revisit this problem and strengthen Theorem 1.2 by removing the dependence on the GRH and providing an explicit error term. To this end, we follow the approach of Lamzouri [19] and employ some ideas from Lamzouri, Lester, and Radziwiłł [21] to compare the distribution of $\frac{L'}{L}(\sigma, \chi_D)$ to that of a probabilistic random model constructed using the independent random variables introduced in [8] (see (2) below).

Following the typographical convention in [16], we will use sans-serif fonts, such as X , to denote arithmetic random variables, and more standard fonts, such as X , for abstract random variables. Using the same letter will usually indicate that the random variable X is a model of the arithmetic quantity X .

Define $\mathcal{F}(N)$ as the set of fundamental discriminants D with $|D| \leq N$, and set

$$(1) \quad \begin{aligned} \mathbf{X}_{n,N} : \mathcal{F}(N) &\rightarrow \{-1, 0, 1\}, \\ D &\mapsto \chi_D(n). \end{aligned}$$

Let $\{X_p\}_{p \text{ prime}}$ be the sequence of independent random variables given by

$$(2) \quad \mathbb{P}(X_p = a) = \begin{cases} \frac{p}{2(p+1)} & \text{if } a = \pm 1, \\ \frac{1}{p+1} & \text{if } a = 0. \end{cases}$$

Furthermore, for any positive integer n , define $X_n = \prod_{p|n} X_p^{\nu_p(n)}$, where $\nu_p(n)$ is the p -adic valuation of n . The random variables X_n satisfy

$$(3) \quad \mathbb{E}(X_n) = \begin{cases} \prod_{p|n} \left(\frac{p}{p+1}\right) & \text{if } n \text{ is a square} \\ 0 & \text{otherwise.} \end{cases}$$

The sequence $X = \{X_n\}_{n \in \mathbb{N}}$ was first introduced in [8] for the purpose of studying the distribution of the extreme values of $L(1, \chi_D)$ as D varies over all fundamental discriminants.

For an odd prime p , consider $\chi_D(p)$ for $D \in \mathbb{Z}/p^2\mathbb{Z}$. Since D is a fundamental discriminant, we know that the residue class corresponding to p^2 is not contained in $\mathcal{F}(N)$. For the remaining $p^2 - 1$ residue classes $\chi_D(p) = 0$ for $p - 1$ of them (whenever D is a multiple of p). The values -1 and 1 on the other hand should occur equally often amongst the remaining $p^2 - p$ residue classes.

This suggests that the random model X should be a good model for the arithmetic sequence $\{\mathbf{X}_{n,N}\}_{n \in \mathbb{N}}$. In fact, one can prove that for all $k \in \mathbb{N}$, we have

$$(4) \quad \lim_{N \rightarrow \infty} \mathbb{E}_N[\mathbf{X}_{n,N}^k] = \mathbb{E}[X_n^k],$$

where $\mathbb{E}_N[\mathbf{X}_{n,N}^k] = \frac{1}{\mathcal{F}(N)} \sum_{D \in \mathcal{F}(N)} X_{n,N}^k(D)$. This follows from Lemma 4.4 in Section 4 using the complete multiplicativity of X_n and $X_{n,N}$.

Fix ϵ with $0 < \epsilon < \frac{1}{2}$. The objects of interest in this paper are the values $\frac{L'}{L}(1/2 + \epsilon, \chi_D)$ as D varies over fundamental discriminants. Formally, we want to analyze the limiting distribution as $N \rightarrow \infty$ of the arithmetic random variables

$$\begin{aligned} L_{\epsilon,N} : \mathcal{F}(N) &\rightarrow \mathbb{R}, \\ D &\mapsto \frac{L'}{L}(1/2 + \epsilon, \chi_D). \end{aligned}$$

For $\Re(s) > 1$, we have

$$(5) \quad \frac{L'}{L}(s, \chi_D) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \chi_D(n) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \mathbf{X}_{n,N}(D).$$

In view of (4) and (5), we introduce the abstract \mathbb{R} -valued random variable

$$(6) \quad L_\epsilon = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{\frac{1}{2} + \epsilon}} X_n.$$

Using the orthogonality relation (3) and applying the Menshov-Rademacher theorem [16, Theorem B.10.5], we see that the random series on the right hand side of (6) is almost surely convergent, and thus L_ϵ is well-defined.

More generally, let $\tau > \frac{1}{2}$, and let $U_\tau = \{s \in \mathbb{C} : \Re(s) > \tau\}$. It follows from the Menshov-Rademacher theorem that the random series

$$(7) \quad \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} X_n$$

is almost surely convergent on U_τ , and so it defines a holomorphic function there. We also consider the random series

$$(8) \quad \sum_p \frac{(\log p) X_p}{p^s - X_p},$$

which, by Kolmogorov's theorem [16, Theorem B.10.1], is almost surely convergent on U_τ , and so it defines a holomorphic function there. One could easily verify that the series (7) and (8) are equal for all s with $\Re(s) > 1$. By analytic continuation, we see that

$$(9) \quad \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} X_n = \sum_p \frac{(\log p) X_p}{p^s - X_p}$$

almost surely in U_τ . In particular, we have

$$(10) \quad L_\epsilon = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{\frac{1}{2}+\epsilon}} X_n = \sum_p \frac{(\log p) X_p}{p^{\frac{1}{2}+\epsilon} - X_p}.$$

Consider the distribution functions

$$F_{\epsilon,N}(z) = \mathbb{P}_N(L_{\epsilon,N} \leq z) := \frac{1}{|\mathcal{F}(N)|} \left| \left\{ D \in \mathcal{F}(N) : \frac{L'}{L} (1/2 + \epsilon, \chi_D) \leq z \right\} \right|,$$

and

$$F_\epsilon(z) = \mathbb{P}(L_\epsilon \leq z),$$

for all $z \in \mathbb{R}$. Recall that $L_{\epsilon,N}$ is said to converge in distribution to L_ϵ if

$$(11) \quad \lim_{N \rightarrow \infty} F_{\epsilon,N}(z) = F_\epsilon(z),$$

for every continuity point z of F_ϵ . Not only are we interested in establishing (11), but we are also interested in determining how well the distribution of L_ϵ approximates that of $L_{\epsilon,N}$. More precisely, the main result of this paper is the following theorem.

Theorem 1.3. *Fix $0 < \epsilon < \frac{1}{2}$. Then $F_{\epsilon,N}$ converges in distribution to F_ϵ which possesses a smooth density function. The characteristic function of F_ϵ has the form*

$$(12) \quad \varphi_{F_\epsilon}(\tau) = \mathbb{E}(\exp(i\tau L_\epsilon)) = \prod_p \left(\frac{1}{p+1} + \frac{p}{2(p+1)} \left[\exp\left(-i\tau \frac{\log p}{p^{\frac{1}{2}+\epsilon} - 1}\right) + \exp\left(i\tau \frac{\log p}{p^{\frac{1}{2}+\epsilon} + 1}\right) \right] \right).$$

Furthermore, as $N \rightarrow \infty$, we have

$$\|F_{\epsilon,N} - F_\epsilon\|_\infty \ll_\epsilon \left(\frac{\log \log N}{\log N} \right)^{\frac{1}{2}+\epsilon}.$$

Using this theorem, we derive the following asymptotic bound for the small values of $\left| \frac{L'}{L}(1/2 + \epsilon, \chi_D) \right|$.

Corollary 1.4. *Let $m_N = \min_{D \in \mathcal{F}(N)} \left(\left| \frac{L'}{L}(1/2 + \epsilon, \chi_D) \right| \right)$. As $N \rightarrow \infty$, we have*

$$m_N \ll \left(\frac{\log \log N}{\log N} \right)^{\frac{1}{2} + \epsilon}.$$

Proof. Let $\eta = \eta(N)$ be a positive parameter which will be chosen so that $\eta(N) \rightarrow 0$ as $N \rightarrow \infty$. Let

$$\Psi_N(\eta) = \left| \left\{ D \in \mathcal{F}(N) : \left| \frac{L'}{L}(1/2 + \epsilon, \chi_D) \right| \leq \eta \right\} \right|.$$

By Theorem 1.3, we have

$$\frac{\Psi_N(\eta)}{|\mathcal{F}(N)|} = \mathbb{P}(L_\epsilon \in [-\eta, \eta]) + O\left(\left(\frac{\log \log N}{\log N} \right)^{\frac{1}{2} + \epsilon} \right).$$

Let $f_\epsilon(x)$ be the smooth density function associated with F_ϵ . By [16, Proposition B.10.8] applied to the random series $\sum_p \frac{(\log p) X_p}{p^{\frac{1}{2} + \epsilon - X_p}}$, we know that $f_\epsilon(0) > 0$. It follows that

$$\mathbb{P}(L_\epsilon \in [-\eta, \eta]) = \int_{-\eta}^{\eta} f_\epsilon(x) dx \gg \eta.$$

Choosing $\eta = C \left(\frac{\log \log N}{\log N} \right)^{\frac{1}{2} + \epsilon}$ for some large enough positive constant C yields

$$\frac{\Psi_N(\eta)}{|\mathcal{F}(N)|} \gg \left(\frac{\log \log N}{\log N} \right)^{\frac{1}{2} + \epsilon}.$$

Hence, we get $m_N \ll \left(\frac{\log \log N}{\log N} \right)^{\frac{1}{2} + \epsilon}$ as desired. □

The corollary above is an analogue of [20, Theorem 1.1] where the authors investigate the small values of $\left| \frac{L'}{L}(1, \chi) \right|$ for non-principal Dirichlet characters χ modulo q , as $q \rightarrow \infty$ over the primes.

Organization. The structure of the paper is as follows. In Section 2, we prove Proposition 2.2 which provides a version of Berry-Esseen inequality based on the method of moments. In Section 3, we show how we use Proposition 2.2 to deduce Theorem 1.3 from two key results; namely, Theorem 3.1 and Proposition 3.2. The former is a result relating the moments of the random model L_ϵ and the arithmetic model $L_{\epsilon, N}$. The latter is a decay bound on the moments of the random model L_ϵ . In Section 4, we collect some key lemmas, allowing for streamlined proofs of these two key results. In Section 5, we prove Theorem 3.1. In Section 6, we prove Proposition 3.2.

Conventions and Notation.

- Given two functions $f(x)$ and $g(x)$, we shall interchangeably use the notation $f(x) = O(g(x))$ and $f(x) \ll g(x)$ to mean there exists $M > 0$ such that $|f(x)| \leq M|g(x)|$ for all sufficiently large x . We write $f(x) \asymp g(x)$ to mean that the estimates $f(x) \ll g(x)$ and $g(x) \ll f(x)$ hold simultaneously.

- Throughout the paper ϵ denotes a fixed positive constant with $0 < \epsilon < \frac{1}{2}$.
- The letter p will always be used to denote a prime number.
- The capital letter F is used for distribution functions and the characteristic function of a distribution function F is denoted by φ_F .
- We denote by $\mathcal{F}(N)$ the set of all fundamental discriminants D with $|D| \leq N$.
- For a subset $\mathcal{A}(N)$ of $\mathcal{F}(N)$, we set $\mathbb{P}_N(\mathcal{A}(N)) = \frac{|\mathcal{A}(N)|}{|\mathcal{F}(N)|}$.
- For an arithmetic random variable Y_N on $\mathcal{F}(N)$, we denote by $\mathbb{E}_N(Y_N)$ the average

$$\frac{1}{|\mathcal{F}(N)|} \sum_{D \in \mathcal{F}(N)} Y_N(D).$$

We also use the notation $\mathbb{E}_N(\mathbf{1}_{\mathcal{A}(N)^c} Y_N)$ to denote the average $\frac{1}{|\mathcal{F}(N)|} \sum_{D \in \mathcal{F}(N) \setminus \mathcal{A}(N)} Y_N(D)$.

Acknowledgements. The authors would like to thank Amir Akbary and Edward Dobrowolski for helpful comments and discussions related to this work.

2. BERRY-ESSEEN INEQUALITY

There are two main tools used to prove convergence in distribution. The first tool is Lévy's continuity theorem, which relates convergence in distribution of a given sequence of distribution functions to point-wise convergence of the corresponding characteristic functions. The second tool is the method of moments which instead relies on proving the convergence of all of the integral moments of the random variables. Both of these methods are non-quantitative in their original forms. However, with some additional assumptions, we can reformulate both of these results in a quantitative format. For instance, we have the following effective analogue of Lévy's continuity theorem (see [26, page 431]).

Proposition 2.1. *Let $\{Y_N\}_{N=1}^\infty$ and Y be real-valued random variables. Let F_N and F denote the corresponding distribution functions, and let φ_{F_N} and φ_F denote the corresponding characteristic functions. Suppose that F is absolutely continuous with bounded density. Then we have*

$$(13) \quad \|F_N - F\|_\infty \ll \frac{1}{T(N)} + \int_{-T(N)}^{T(N)} \left| \frac{\varphi_{F_N}(\tau) - \varphi_F(\tau)}{\tau} \right| d\tau,$$

for any $T(N) \rightarrow \infty$.

This type of result, which uses effective point-wise convergence of characteristic functions to attain effective convergence in distribution, is sometimes referred to as a *Berry-Esseen theorem* (although some authors reserve this term for the specific case in which the limiting distribution is normal). In [21], the authors utilized this approach effectively in combination with Beurling-Selberg functions to obtain an improved upper bound on the discrepancy between the distribution of $\zeta(s)$ on the line $\Re(s) = \frac{1}{2} + \epsilon$ and that of its random model (See [21, Theorem 1.1]).

The following proposition is a version of (13) based on the method of moments, and it provides the main probabilistic tool which allows us to attain the discrepancy bound in Theorem 1.3

Proposition 2.2. *Let $\{Y_N\}_{N=1}^\infty$ and Y be real-valued random variables for which all moments exist and satisfy*

$$\mathbb{E}(|Y_N|^k)^{1/k}, \mathbb{E}(|Y|^k)^{1/k} = o(k), \quad \text{as } k \rightarrow \infty.$$

Let F_N and F denote the distribution functions of Y_N and Y respectively and suppose that φ_F is absolutely integrable. Let $m(N, k)$ be some positive function such that uniformly for $k \in \mathbb{Z}^+$

$$|\mathbb{E}(Y_N^k) - \mathbb{E}(Y^k)|^{1/k} \ll m(N, k), \quad \text{as } N \rightarrow \infty.$$

Suppose that there exists some function $M(N) \rightarrow \infty$ such that uniformly for $k > \log M(N)$, we have

$$(14) \quad \frac{m(N, k)}{k} \ll \frac{1}{M(N)}, \quad \text{as } N \rightarrow \infty,$$

and uniformly for $k \leq \log M(N)$, we have

$$(15) \quad M(N) \ll \log \frac{1}{m(N, k)^k}, \quad \text{as } N \rightarrow \infty.$$

Then $\{Y_N\}_N$ converges to Y in distribution, F has a smooth density function, and

$$\|F_N - F\|_\infty \ll \frac{1}{M(N)}.$$

Proof. By Proposition 2.1, for any $T(N) \rightarrow \infty$ we have

$$\|F_N - F\|_\infty \ll \frac{1}{T(N)} + \int_{-T(N)}^{T(N)} \left| \frac{\varphi_{F_N}(\tau) - \varphi_F(\tau)}{\tau} \right| d\tau.$$

Recall that $\varphi_{F_N}(\tau) = \mathbb{E}(\exp(i\tau Y_N))$ and $\varphi_F(\tau) = \mathbb{E}(\exp(i\tau Y))$. We have

$$\begin{aligned} \int_{-T(N)}^{T(N)} \left| \frac{\varphi_{F_N}(\tau) - \varphi_F(\tau)}{\tau} \right| d\tau &= \int_{-T(N)}^{T(N)} \left| \sum_{k=1}^{\infty} \frac{\mathbb{E}(Y_N^k) - \mathbb{E}(Y^k)}{k!} (i\tau)^{k-1} \right| d\tau \\ &\ll \int_{-T(N)}^{T(N)} \sum_{k=1}^{\infty} \frac{(C_1 m(N, k))^k}{k!} \tau^{k-1} d\tau, \end{aligned}$$

for some absolute positive constant $C_1 > 0$. Interchanging summation and integration yields

$$\int_{-T(N)}^{T(N)} \left| \frac{\varphi_{F_N}(\tau) - \varphi_F(\tau)}{\tau} \right| d\tau \ll \sum_{k=1}^{\infty} \frac{(C_1 m(N, k))^k}{k \cdot k!} T(N)^k.$$

Using Stirling's formula, $\sqrt{2\pi} k^{k+\frac{1}{2}} e^{-k} \leq k! \leq e k^{k+\frac{1}{2}} e^{-k}$ for all $k \in \mathbb{N}$, we get

$$\int_{-T(N)}^{T(N)} \left| \frac{\varphi_{F_N}(\tau) - \varphi_F(\tau)}{\tau} \right| d\tau \ll \sum_{k=1}^{\infty} \frac{1}{k^{3/2}} \left(\frac{e C_1 m(N, k)}{k} T(N) \right)^k.$$

It follows that

$$\begin{aligned} \|F_N - F\|_\infty &\ll \frac{1}{T(N)} + \sum_{k=1}^{\infty} \frac{1}{k^{3/2}} \left(\frac{e C_1 m(N, k)}{k} T(N) \right)^k \\ (16) \quad &\ll \frac{1}{T(N)} + \sum_{k \leq \log M(N)} \frac{1}{k^{3/2}} \left(e C_1 \frac{m(N, k)}{k} M(N) \right)^k + \sum_{k > \log M(N)} \left(e C_1 \frac{m(N, k)}{k} M(N) \right)^k. \end{aligned}$$

Choose $T(N) = CM(N)$ for some positive constant C to be determined later. By (14), we have that for sufficiently large n ,

$$\sup_{k > \log M(N)} \frac{m(N, k)}{k} \leq \frac{C_2}{M(N)},$$

for some absolute positive constant C_2 . Hence,

$$(17) \quad \sum_{k > \log M(N)} \left(eC_1 \frac{m(N, k)}{k} T(N) \right)^k \ll \sum_{k > \log M(N)} \left(eCC_1C_2 \right)^k \ll \frac{1}{M(N)},$$

provided that $C < \frac{1}{e^2C_1C_2}$. Notice that (15) implies that there exists $C_3 > 0$ such that

$$\sup_{k \leq \log M(N)} m(N, k)^k \ll e^{-C_3M(N)}, \quad \text{as } N \rightarrow \infty.$$

This and another application of Stirling's formula implies

$$(18) \quad \sum_{k \leq \log M(N)} \frac{1}{k^{3/2}} \left(eC_1 \frac{m(N, k)}{k} T(N) \right)^k \ll e^{-C_3M(N)} \sum_{k=0}^{\infty} \frac{(CC_1M(N))^k}{k!} \ll e^{(CC_1-C_3)M(N)} \ll \frac{1}{M(N)},$$

provided that $CC_1 < C_3$. Choosing $0 < C < \min(\frac{1}{e^2C_1C_2}, \frac{C_3}{C_1})$ and combining (16), (17) and (18) yield the desired result. \square

3. PROOF OF THEOREM 1.3

The proof of Theorem 1.3 is accomplished in two parts. The first part consists of proving that the large moments $\mathbb{E}_N(1_{\mathcal{E}(N)^c} L_{\epsilon, N}^k)$ of $L_{\epsilon, N}$ defined as the average of $\frac{L'}{L}(1/2 + \epsilon, \chi_D)^k$ over $D \in \mathcal{F}(N) \setminus \mathcal{E}(N)$ can be approximated by the corresponding moments $\mathbb{E}(L_\epsilon^k)$ of the random model L_ϵ . Here $\mathcal{E}(N)$ is an exceptional set of fundamental discriminants such that $|\mathcal{E}(N)| = O(N^{1-c})$ for some $c > 0$. More precisely, we prove the following theorem.

Theorem 3.1. *There exists a set of fundamental discriminants $\mathcal{E}(N) \subset \mathcal{F}(N)$ with $\mathbb{P}_N(\mathcal{E}(N)) = O(N^{-c})$ for some $c > 0$, such that uniformly for $k \in \mathbb{Z}^+$, we have*

$$|\mathbb{E}_N(1_{\mathcal{E}(N)^c} L_{\epsilon, N}^k) - \mathbb{E}(L_\epsilon^k)|^{1/k} \ll \frac{\log N}{N^{\frac{\epsilon^2(\epsilon+3)}{12k}}}.$$

Furthermore, this holds when $\mathcal{E}(N)$ is replaced by any $\mathcal{E}_*(N) \supset \mathcal{E}(N)$ as long as $\mathbb{P}_N(\mathcal{E}_*(N)) \asymp \mathbb{P}_N(\mathcal{E}(N))$.

The second part of the proof of Theorem 1.3 consists of using the Berry-Essen inequality described in Proposition 2.2 to relate the distribution functions $F_{\epsilon, N}$ and F_ϵ to the moments $\mathbb{E}_N(1_{\mathcal{E}(N)^c} L_{\epsilon, N}^k)$ and $\mathbb{E}(L_\epsilon^k)$. This allows us to get an upper bound on the rate of convergence of $F_{\epsilon, N}$ to F_ϵ .

We require the following two propositions in order to verify that the conditions of Proposition 2.2 are satisfied.

Proposition 3.2. *As $k \rightarrow \infty$, we have $\mathbb{E}(|L_\epsilon|^k)^{1/k} \ll k^{\frac{1}{2}-\epsilon}$.*

Proposition 3.3 (Lemma 4 of [22]). *As $|\tau| \rightarrow \infty$, we have $\varphi_{F_\epsilon}(\tau) \ll \exp(-C|\tau|^{\frac{1}{2+\epsilon}})$, for some positive constant C that depends only on ϵ .*

A proof of Proposition 3.3 can be found in [22]. The reader is referred to Section 5 and Section 6 for the proofs of Theorem 3.1 and Proposition 3.2 respectively.

Finally, we need the following result which is inspired by [21, Lemma 3.4] and follows from Theorem 3.1 and Proposition 3.2.

Lemma 3.4. *There exists a constant $B = B(\epsilon) > 0$ such that*

$$\mathbb{P}_N \left[|\mathbf{1}_{\mathcal{E}(N)^c} \mathbf{L}_{\epsilon, N}| \geq \left(\frac{\log N}{\log \log N} \right)^{\frac{1}{2}-\epsilon} \right] \ll \exp \left(-B \frac{\log N}{\log \log N} \right).$$

Proof. Markov's inequality (see, for example, [1, Eq. 5.31]) gives

$$(19) \quad \mathbb{P}_N \left[|\mathbf{1}_{\mathcal{E}(N)^c} \mathbf{L}_{\epsilon, N}| \geq \left(\frac{\log N}{\log \log N} \right)^{\frac{1}{2}-\epsilon} \right] \leq \left(\mathbb{E}_N (|\mathbf{1}_{\mathcal{E}(N)^c} \mathbf{L}_{\epsilon, N}|^k)^{1/k} \left(\frac{\log \log N}{\log N} \right)^{\frac{1}{2}-\epsilon} \right)^k,$$

for any positive integer k . By Theorem 3.1 and Proposition 3.2, we have

$$(20) \quad \mathbb{E}_N (\mathbf{1}_{\mathcal{E}(N)^c} \mathbf{L}_{\epsilon, N}^k)^{1/k} \leq C \frac{\log N}{N^{\frac{\epsilon^2(\epsilon+3)}{12k}}} + k^{\frac{1}{2}-\epsilon},$$

for some positive constant C . Choosing $k = \frac{\log N}{A \log \log N}$ with $A > \max(C^{\frac{1}{\frac{1}{2}-\epsilon}}, \frac{\frac{1}{2}+\epsilon}{\epsilon^2(\epsilon+3)})$ and combining (19) and (20) yield the desired result. \square

Proof of Theorem 1.3. Let $\mathcal{E}_*(N) = \mathcal{E}(N) \cup \{D \in \mathcal{F}(N) : |\frac{L}{L}(1/2 + \epsilon, \chi_D)| \geq (\frac{\log N}{\log \log N})^{\frac{1}{2}-\epsilon}\}$. By Lemma 3.4, we have $\mathbb{P}_N(\mathcal{E}_*(N)) \asymp \mathbb{P}_N(\mathcal{E}(N))$. Thus, we may apply Theorem 3.1 to obtain

$$(21) \quad |\mathbb{E}_N((\mathbf{1}_{\mathcal{E}_*(N)^c} \mathbf{L}_{\epsilon, N})^k) - \mathbb{E}(L_\epsilon^k)|^{1/k} \ll \frac{\log N}{N^{\delta/k}},$$

where $\delta = \epsilon^2(\epsilon + 3)/12 > 0$. On the other hand, by our definition of $\mathcal{E}_*(N)$ and Proposition 3.2 we have

$$(22) \quad |\mathbb{E}_N((\mathbf{1}_{\mathcal{E}_*(N)^c} \mathbf{L}_{\epsilon, N})^k) - \mathbb{E}(L_\epsilon^k)|^{1/k} \ll k^{\frac{1}{2}-\epsilon} + \left(\frac{\log N}{\log \log N} \right)^{\frac{1}{2}-\epsilon}.$$

Since the first term on the right hand side of (22) is dominant as long as $k \gg \frac{\log N}{\log \log N}$, we combine (21) and (22) to get

$$|\mathbb{E}_N((\mathbf{1}_{\mathcal{E}_*(N)^c} \mathbf{L}_{\epsilon, N})^k) - \mathbb{E}(L_\epsilon^k)|^{1/k} \ll m(N, k),$$

where

$$m(N, k) = \begin{cases} \frac{\log N}{N^{\delta/k}} & \text{if } k \leq \frac{\delta}{\frac{1}{2}+\epsilon} \frac{\log N}{\log \log N}, \\ k^{\frac{1}{2}-\epsilon} & \text{if } k > \frac{\delta}{\frac{1}{2}+\epsilon} \frac{\log N}{\log \log N}. \end{cases}$$

Observe that

$$\sup_{k \gg \log \log N} \frac{m(N, k)}{k} \ll \left(\frac{\log \log N}{\log N} \right)^{\frac{1}{2}+\epsilon},$$

and

$$\inf_{k \ll \log \log N} k \log \frac{1}{m(N, k)} \gg \inf_{k \ll \log \log N} k \log \frac{N^{\delta/k}}{\log N} \gg \log N.$$

It follows that conditions (14) and (15) are satisfied. Since the characteristic function of the random model is absolutely integrable by Proposition 3.3, we can apply Proposition 2.2 to get

$$\|F_{\epsilon, N; \star} - F_\epsilon\|_\infty \ll \left(\frac{\log \log N}{\log N} \right)^{\frac{1}{2} + \epsilon},$$

where $F_{\epsilon, N; \star}$ is the distribution function corresponding to $\mathbf{1}_{\mathcal{E}_\star(N)^c} L_{\epsilon, N}$. Finally, combining this with Lemma 3.4 implies

$$\|F_{\epsilon, N} - F_\epsilon\|_\infty \ll \exp\left(-B \frac{\log N}{\log \log N}\right) + \|F_{\epsilon, N; \star} - F_\epsilon\|_\infty \ll \left(\frac{\log \log N}{\log N} \right)^{\frac{1}{2} + \epsilon},$$

as desired. □

4. PRELIMINARY LEMMAS

Recall that if $\Re(s) > 1$, we have

$$\left(\frac{L'}{L}(s, \chi_D) \right)^k = (-1)^k \sum_{n=1}^{\infty} \frac{\Lambda_k(n)}{n^s} \chi_D(n),$$

where

$$\Lambda_k(n) = \sum_{\substack{n_1, n_2, \dots, n_k \geq 1 \\ n_1 n_2 \dots n_k = n}} \Lambda(n_1) \Lambda(n_2) \dots \Lambda(n_k),$$

and it satisfies

$$(23) \quad \Lambda_k(n) \leq \left(\sum_{m|n} \Lambda(m) \right)^k = (\log n)^k.$$

For $0 < \sigma < 1$, this upper bound along with an application of partial summation yields

$$(24) \quad \sum_{n \leq \lambda} \frac{\Lambda_k(n)}{n^\sigma} \ll (\log \lambda)^k \lambda^{1-\sigma}.$$

In what follows, we collect several basic lemmata that are required in the sequel.

Lemma 4.1 (Corollary 5.3 of [23]). *Consider the Dirichlet series $\alpha(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ with abscissa of absolute convergent σ_a and abscissa of convergence σ_c . Fix some $\sigma_0 > \sigma_c$. Choose c such that $c > \max(0, \sigma_a - \sigma_0)$. Let $\lambda > 0$ be non-integral. Then,*

$$\sum_{n \leq \lambda} \frac{a_n}{n^{\sigma_0}} = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \alpha(\sigma_0 + s) \frac{\lambda^s}{s} ds + E$$

where

$$E \ll \sum_{\lambda/2 < n < 2\lambda} \frac{|a_n|}{n^{\sigma_0}} \min\left(1, \frac{\lambda}{T|\lambda - n|}\right) + \frac{4^c + \lambda^c}{T} \sum_{n=1}^{\infty} \frac{|a_n|}{n^{\sigma_0+c}}.$$

Lemma 4.2 (Lemma 2.2 of [19]). *Suppose that $L(s, \chi_D)$ is nonzero for $\Re(s) > \sigma_0$ and $|\Im(s)| \leq |t| + 1$. Then, for $\sigma > \sigma_0$, we have*

$$\frac{L'}{L}(\sigma + it, \chi_D) \ll \frac{\log(|D|(|t| + 2))}{\sigma - \sigma_0}.$$

Lemma 4.3 (Theorem 3 of [9]). *Let $N(v, \tau, \chi_D)$ denote the number of zeros of $L(s, \chi_D)$ in the rectangle $R(v, \tau) = \{s \in \mathbb{C} : \frac{1}{2} + v < \Re(s) \leq 1, |\Im(s)| \leq \tau\}$. Then for any $\delta > 0$, we have*

$$\sum_{D \in \mathcal{F}(N)} N(v, \tau, \chi_D) \ll_{\delta} (N\tau)^{\delta} N^{\frac{3-6v}{3-2v}} \tau^{\frac{4-4v}{3-2v}}.$$

In view of this lemma, if we let $\mathcal{E}(v, \tau; N)$ denote the set of $D \in \mathcal{F}(N)$ for which $\frac{L'}{L}(s, \chi_D)$ has at least one pole in $R(v, \tau)$, then

$$(25) \quad |\mathcal{E}(v, \tau; N)| \ll_{\delta} (N\tau)^{\delta} N^{\frac{3-6v}{3-2v}} \tau^{\frac{4-4v}{3-2v}}.$$

Finally, the following lemma serves a crucial role as a bridge from the arithmetic random setting into the abstract probabilistic setting.

Lemma 4.4. *For sufficiently large N , we have $\mathbb{E}_N(\mathbf{X}_{n,N}) - \mathbb{E}(X_n) \ll N^{-\frac{1}{2}} n^{\frac{1}{4}} \log n$.*

Proof. By definition, we have

$$(26) \quad \mathbb{E}_N(\mathbf{X}_{n,N}) = \frac{1}{|\mathcal{F}(N)|} \sum_{D \in \mathcal{F}(N)} \chi_D(n),$$

where

$$|\mathcal{F}(N)| = \frac{6}{\pi^2} N + O\left(N^{\frac{1}{2} + \epsilon}\right).$$

In fact, if $n = m^2$, then we have the following standard estimate (see for example [19, page 640])

$$\sum_{D \in \mathcal{F}(N)} \chi_D(m^2) = \sum_{\substack{D \in \mathcal{F}(N) \\ (D, m)=1}} 1 = \frac{6}{\pi^2} N \prod_{p|m} \left(\frac{p}{p+1}\right) + O(N^{\frac{1}{2}} \tau(m)),$$

where $\tau(m)$ is the divisor function. Combining this with (3) yields

$$\mathbb{E}_N(\mathbf{X}_{n,N}) - \mathbb{E}(X_n) \ll N^{-\frac{1}{2}} \tau(\sqrt{n}),$$

provided that n is a perfect square. By [8, lemma 4.1], we have

$$\sum_{D \in \mathcal{F}(N)} \chi_D(n) \ll N^{\frac{1}{2}} n^{\frac{1}{4}} \log n,$$

for non-square n . This implies that $\mathbb{E}_N(\mathbf{X}_{n,N}) - \mathbb{E}(X_n) \ll N^{-\frac{1}{2}} n^{\frac{1}{4}} \log n$ if n is not a perfect square. \square

5. PROOF OF THEOREM 3.1

The point of departure in proving Theorem 3.1 is approximating integral powers of $\frac{L'}{L}(1/2+\epsilon, \chi_D)$ by short Dirichlet polynomials.

Let d and e be two positive constants such that $d < e < \epsilon$. We set

$$\mathcal{E}_{d,e}(N) = \{D \in \mathcal{F}(N) : L(s, \chi_D) = 0 \text{ for some } s \in R_{d,e}\},$$

where

$$R_{d,e} = \{s \in \mathbb{C} : \frac{1}{2} + (\epsilon - e) < \Re(s) \leq 1, |\Im(s)| \leq \lambda^{\frac{1}{2} - (\epsilon - d)} + 1\}.$$

For simplicity, we suppress the subscripts from our notation and set $\mathcal{E}(N) = \mathcal{E}_{d,e}(N)$ and $R = R_{d,e}$. It follows from (25) that for any $\delta > 0$, we have

$$(27) \quad \mathbb{P}_N(\mathcal{E}(N)) \ll \frac{1}{N^b},$$

with

$$b = \frac{4}{3 - 2(\epsilon - e)} \left[(\epsilon - e) - (1 - (\epsilon - e)) \left(\frac{1}{2} - (\epsilon - d) \right) \left(\frac{\log \lambda}{\log N} \right) \right] - \delta.$$

Proposition 5.1. *Suppose λ satisfies $\log \lambda \ll \log N$. Then, for all $D \in \mathcal{F}(N) \setminus \mathcal{E}(N)$ and $k \in \mathbb{Z}^+$, we have*

$$\left(\frac{L'}{L} \left(\frac{1}{2} + \epsilon, \chi_D \right) \right)^k = (-1)^k \sum_{n \leq \lambda} \frac{\Lambda_k(n)}{n^{\frac{1}{2} + \epsilon}} \chi_D(n) + O \left(\frac{(C \log N)^{k+1}}{\lambda^d} \right),$$

for some positive constant C .

Proof. Assume throughout that $D \in \mathcal{F}(N) \setminus \mathcal{E}(N)$. Lemma 4.1 gives

$$(28) \quad \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \left(-\frac{L'}{L} \left(\frac{1}{2} + \epsilon + s, \chi_D \right) \right)^k \frac{\lambda^s}{s} ds = \sum_{n \leq \lambda} \frac{\Lambda_k(n) \chi_D(n)}{n^{\frac{1}{2} + \epsilon}} + \Psi_1 + \Psi_2,$$

where

$$\Psi_1 \ll \sum_{\lambda/2 < n < 2\lambda} \frac{\Lambda_k(n)}{n^{\frac{1}{2} + \epsilon}} \min \left(1, \frac{\lambda}{T|n - \lambda|} \right) \quad \text{and} \quad \Psi_2 \ll \frac{4^c + \lambda^c}{T} \sum_{n=1}^{\infty} \frac{\Lambda_k(n)}{n^{\frac{1}{2} + \epsilon + c}}.$$

We fix $c = \frac{1}{2} - \epsilon + 1/\log \lambda$ and assume without loss of generality that $\lambda \in \mathbb{Z} + \frac{1}{2}$. Using (23) we get

$$\Psi_1 \ll \frac{\lambda^{\frac{1}{2} - \epsilon} (\log(2\lambda))^k}{T} \sum_{n \leq 2\lambda} \frac{1}{|n - \lambda|} \ll \frac{\lambda^{\frac{1}{2} - \epsilon} (\log(2\lambda))^{k+1}}{T}$$

and

$$\Psi_2 \ll \frac{\lambda^{\frac{1}{2} - \epsilon}}{T} \left(\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{1+1/\log \lambda}} \right)^k \ll \frac{\lambda^{\frac{1}{2} - \epsilon}}{T} (\log \lambda)^k.$$

We now shift the line of integration in (28) from $\Re(s) = c$ to $\Re(s) = -d$. Since $D \notin \mathcal{E}(N)$, the integrand has only a simple pole at $s = 0$. By the residue theorem, we get

$$\left(-\frac{L'}{L} \left(\frac{1}{2} + \epsilon, \chi_D \right) \right)^k = \sum_{n \leq \lambda} \frac{\Lambda_k(n)}{n^{\frac{1}{2} + \epsilon}} \chi_D(n) + \Psi_1 + \Psi_2 + \Psi_3 + \Psi_4,$$

where

$$\Psi_3 + \Psi_4 = \left(\int_{c+iT}^{-d+iT} + \int_{-d+iT}^{-d-iT} + \int_{-d-iT}^{c-iT} \right) \left(-\frac{L'}{L} \left(\frac{1}{2} + \epsilon + s, \chi_D \right) \right)^k \frac{\lambda^s}{s} ds.$$

Here, Ψ_3 denotes the first and third integral, and Ψ_4 denotes the second integral. Applying Lemma 4.2 gives

$$\frac{L'}{L} \left(\frac{1}{2} + \epsilon + \sigma + it, \chi_D \right) \ll \frac{\log(|D|(|t| + 2))}{\sigma + e} \leq \frac{\log(N(|t| + 2))}{\sigma + e},$$

for $|t| \leq T$. It follows that

$$\Psi_3 \ll \int_{-d}^c \left| \frac{L'}{L} \left(\frac{1}{2} + \epsilon + \sigma + iT, \chi_D \right) \right|^k \frac{\lambda^\sigma}{T} d\sigma \ll \frac{\lambda^c}{T \log \lambda} \left(\frac{\log(N(T+2))}{e-d} \right)^k.$$

Similarly,

$$\Psi_4 \ll \int_{-T}^T \left| \frac{L'}{L} \left(\frac{1}{2} + \epsilon - d + it, \chi_D \right) \right|^k \left| \frac{\lambda^{-d+it}}{-d+it} \right| dt \ll \frac{1}{\lambda^d} \left(\frac{\log(N(T+2))}{e-d} \right)^k \log(1 + T/d).$$

If we set $E = \Psi_1 + \Psi_2 + \Psi_3 + \Psi_4$, then

$$E \ll \frac{\lambda^{\frac{1}{2}-\epsilon} (\log(2\lambda))^{k+1}}{T} + (\log(N(T+2)))^{k+1} \left[\frac{\lambda^{\frac{1}{2}-\epsilon}}{T \log \lambda} + \frac{1}{\lambda^d} \right].$$

Choosing $T = \lambda^{\frac{1}{2}-(\epsilon-d)}$ and assuming $\log \lambda \ll \log N$ yields the desired result. \square

Proof of Theorem 3.1. Choose λ such that $\log \lambda \ll \log N$. Using Proposition 5.1, (24) and (27) gives

$$\mathbb{E}_N(\mathbf{1}_{\mathcal{E}(N)^c} \mathbf{L}_{\epsilon, N}^k) = \sum_{n \leq \lambda} \frac{\Lambda_k(n)}{n^{\frac{1}{2}+\epsilon}} \mathbb{E}_N(X_{n, N}) + O \left[\frac{(C \log N)^{k+1}}{\lambda^d} + \frac{\lambda^{\frac{1}{2}-\epsilon} (\log \lambda)^k}{N^b} \right].$$

Note that this equation still holds if $\mathcal{E}(N)$ is replaced by some larger exceptional set $\mathcal{E}_*(N)$ as long as $\mathbb{P}_N(\mathcal{E}_*(N)) \asymp \mathbb{P}_N(\mathcal{E}(N))$. We apply Lemma 4.4 to obtain

$$\mathbb{E}_N(\mathbf{1}_{\mathcal{E}(N)^c} \mathbf{L}_{\epsilon, N}^k) = \mathbb{E} \left(\sum_{n \leq \lambda} \frac{\Lambda_k(n)}{n^{\frac{1}{2}+\epsilon}} X_n \right) + O \left[(C \log N)^{k+1} \left(\frac{1}{\lambda^d} + \frac{\lambda^{\frac{1}{2}-\epsilon}}{N^b} + \frac{\lambda^{3/4-\epsilon}}{N^{1/4}} \right) \right].$$

The orthogonality property of X_n (see (3)) implies

$$\mathbb{E} \left(\sum_{n > \lambda} \frac{\Lambda_k(n)}{n^{\frac{1}{2}+\epsilon}} X_n \right) \ll \frac{(2 \log \lambda)^k}{\lambda^{1+2\epsilon}},$$

which can clearly be neglected. Thus,

$$\mathbb{E}_N(\mathbf{1}_{\mathcal{E}(N)^c} \mathbf{L}_{\epsilon, N}^k) = \mathbb{E}(L_\epsilon^k) + O \left(\frac{(C \log N)^{k+1}}{N^t} \right),$$

where

$$t = \min \left[d \frac{\log \lambda}{\log N}, \min \left(b, \frac{1}{4} \left(1 - \frac{\log \lambda}{\log N} \right) \right) - \frac{\log \lambda}{\log N} \left(\frac{1}{2} - \epsilon \right) \right].$$

Choosing δ sufficiently small, $d = \epsilon - (3 - \sqrt{9 - 6\epsilon})/2$, e sufficiently close to d , and

$$\frac{\log \lambda}{\log N} = \min \left(\frac{6}{7} \frac{4(\epsilon - d)}{(3 - 2(\epsilon - d))(1 - 2(\epsilon - d))}, \frac{1}{3 - 4(\epsilon - d)} \right),$$

gives $t > \epsilon^2(\epsilon + 3)/12$. □

6. PROOF OF PROPOSITION 3.2

Proof of Proposition 3.2. Using (10) and applying Minkowski's inequality yield

$$(29) \quad \mathbb{E}(|L_\epsilon|^k)^{1/k} \leq \mathbb{E} \left[\left| \sum_p \frac{\log p}{p^{\frac{1}{2}+\epsilon}} X_p \right|^k \right]^{1/k} + \mathbb{E} \left[\left| \sum_{p^j, j \geq 2} \log p \left(\frac{X_p}{p^{\frac{1}{2}+\epsilon}} \right)^j \right|^k \right]^{1/k}.$$

The second sum on the right hand side of (29) is

$$\mathbb{E} \left[\left| \sum_p \log p \frac{X_p^2}{p^{1+2\epsilon} - p^{\frac{1}{2}+\epsilon} X_p} \right|^k \right]^{1/k} \ll \sum_p \frac{\log p}{p^{1+2\epsilon}} \ll 1.$$

We split the first sum on the right hand side of (29) at some y which we determine later to get

$$\begin{aligned} \mathbb{E}(|L_\epsilon|^k)^{1/k} &\ll \mathbb{E} \left[\left| \sum_{p \leq y} \frac{\log p}{p^{\frac{1}{2}+\epsilon}} X_p \right|^k \right]^{1/k} + \mathbb{E} \left[\left| \sum_{p > y} \frac{\log p}{p^{\frac{1}{2}+\epsilon}} X_p \right|^k \right]^{1/k} + 1 \\ &\ll \sum_{p \leq y} \frac{\log p}{p^{\frac{1}{2}+\epsilon}} + \mathbb{E} \left[\left(\sum_{p > y} \frac{\log p}{p^{\frac{1}{2}+\epsilon}} X_p \right)^{2k} \right]^{1/2k} + 1, \end{aligned}$$

by Minkowski's inequality and the Cauchy-Schwartz inequality. Partial summation and the prime number theorem give $\sum_{p \leq y} \frac{\log p}{p^{\frac{1}{2}+\epsilon}} \ll y^{\frac{1}{2}-\epsilon}$. Observe that

$$(30) \quad \begin{aligned} \mathbb{E} \left[\left(\sum_{p > y} \frac{\log p}{p^{\frac{1}{2}+\epsilon}} X_p \right)^{2k} \right]^{1/2k} &= \mathbb{E} \left[\sum_{\substack{p_1, \dots, p_k > y \\ q_1, \dots, q_k > y}} \frac{\log p_1 \dots \log p_k \log q_1 \dots \log q_k}{(p_1 \dots p_k q_1 \dots q_k)^{\frac{1}{2}+\epsilon}} X_{p_1 \dots p_k} X_{q_1 \dots q_k} \right]^{1/2k} \\ &\leq \left[\sum_{p_1, \dots, p_k > y} \frac{(\log p_1 \dots \log p_k)^2}{(p_1 \dots p_k)^{1+2\epsilon}} \sum_{\substack{q_1, \dots, q_k > y \\ q_1 \dots q_k = p_1 \dots p_k}} 1 \right]^{1/2k}, \end{aligned}$$

where the last inequality follows from the orthogonality of X_n (see (3)). The innermost sum in (30) counts the number of permutations on the set $\{1, \dots, k\}$, which is just $k!$. An application of Stirling's formula then yields

$$\mathbb{E} \left[\left(\sum_{p > y} \frac{\log p}{p^{\frac{1}{2}+\epsilon}} X_p \right)^{2k} \right]^{1/2k} \ll \sqrt{k} \left(\sum_{p > y} \frac{\log p}{p^{1+2\epsilon}} \right)^{\frac{1}{2}} \ll \frac{\sqrt{k}}{y^\epsilon}.$$

Choosing $y = k$ gives the desired result. □

REFERENCES

- [1] Patrick Billingsley, *Probability and measure*, Anniversary Edition, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., New York, 2012.
- [2] Harald Bohr and Børge Jessen, *Über die Werteverteilung der Riemannschen Zetafunktion, erste Mitteilung*, Acta Math. **54** (1930), no. 1, 1–35 (German).
- [3] ———, *Über die Werteverteilung der Riemannschen Zetafunktion*, Acta Math. **58** (1932), no. 1, 1–55 (German).
- [4] Sarvadaman Chowla and Paul Erdős, *A theorem on the distribution of the values of L -functions*, J. Indian Math. Soc. (N.S.) **15** (1951), 11–18.
- [5] P. D. T. A. Elliott, *On the distribution of the values of Dirichlet L -series in the half-plane $\sigma > \frac{1}{2}$* , Nederl. Akad. Wetensch. Proc. Ser. A **74**=Indag. Math. **33** (1971), 222–234.
- [6] ———, *On the distribution of $\arg L(s, \chi)$ in the half-plane $\sigma > \frac{1}{2}$* , Acta Arith. **20** (1972), 155–169.
- [7] ———, *On the distribution of the values of quadratic L -series in the half-plane $\sigma > \frac{1}{2}$* , Invent. Math. **21** (1973), 319–338.
- [8] Andrew Granville and Kannan Soundararajan, *The distribution of values of $L(1, \chi_d)$* , Geom. Funct. Anal. **13** (2003), no. 5, 992–1028.
- [9] Roger D. Heath-Brown, *A mean value estimate for real character sums*, Acta Arith. **72** (1995), no. 3, 235–275.
- [10] Tetsuya Hattori and Kohji Matsumoto, *A limit theorem for Bohr-Jessen’s probability measures of the Riemann zeta-function*, J. Reine Angew. Math. **507** (1999), 219–232.
- [11] Yasutaka Ihara, *On the Euler-Kronecker constants of global fields and primes with small norms*, Algebraic geometry and number theory, Progr. Math., vol. 253, Birkhäuser Boston, Boston, MA, 2006, pp. 407–451.
- [12] ———, *The Euler-Kronecker invariants in various families of global fields*, Arithmetics, geometry, and coding theory (AGCT 2005), Sémin. Congr., vol. 21, Soc. Math. France, Paris, 2010, pp. 79–102 (English, with English and French summaries).
- [13] Yasutaka Ihara and Kohji Matsumoto, *On certain mean values and the value-distribution of logarithms of Dirichlet L -functions*, Q. J. Math. **62** (2011), no. 3, 637–677.
- [14] ———, *On $\log L$ and L'/L for L -functions and the associated “ M -functions”: connections in optimal cases*, Mosc. Math. J. **11** (2011), no. 1, 73–111, 182 (English, with English and Russian summaries).
- [15] Børge Jessen and Aurel Wintner, *Distribution functions and the Riemann zeta function*, Trans. Amer. Math. Soc. **38** (1935), no. 1, 48–88.
- [16] Emmanuel Kowalski, *Lecture Notes: An Introduction to Probabilistic Number Theory*, ETH Zurich, Version of March 13, 2020. <https://people.math.ethz.ch/~kowalski/probabilistic-number-theory.pdf>.
- [17] Youness Lamzouri, *Distribution of values of L -functions at the edge of the critical strip*, Proc. Lond. Math. Soc. (3) **100** (2010), no. 3, 835–863.
- [18] ———, *On the distribution of extreme values of zeta and L -functions in the strip $\frac{1}{2} < \sigma < 1$* , Int. Math. Res. Not. IMRN **23** (2011), 5449–5503.
- [19] ———, *The distribution of Euler-Kronecker constants of quadratic fields*, J. Math. Anal. Appl. **432** (2015), no. 2, 632–653.
- [20] Youness Lamzouri and Alessandro Languasco, *Small values of $|L'/L(1, \chi)|$* , Experimental Mathematics (to appear).
- [21] Youness Lamzouri, Stephen Lester, and Maksym Radziwiłł, *Discrepancy bounds for the distribution of the Riemann zeta-function and applications*, Journal d’Analyse Mathématique **139** (2019), no. 2, 453–494.
- [22] Mariam Mourtada and V. Kumar Murty, *Distribution of values of $L'/L(\sigma, \chi_D)$* , Mosc. Math. J. **15** (2015), no. 3, 497–509, 605 (English, with English and Russian summaries).
- [23] Hugh L. Montgomery and Robert C. Vaughan, *Multiplicative Number Theory I: Classical Theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2006.
- [24] E. Stankus, *Distribution of Dirichlet L -functions with real characters in the half-plane $\operatorname{Re} s > 1/2$* , Litovsk. Mat. Sb. **15** (1975), no. 4, 199–214, 249 (Russian, with Lithuanian and English summaries). MR0406956
- [25] Kai-Man Tsang, *The distribution of the values of the Riemann zeta-function*, ProQuest LLC, Ann Arbor, MI, 1984. Thesis (Ph.D) – Princeton University.

- [26] Gérald Tenenbaum, *Introduction to analytic and probabilistic number theory*, 3rd ed., Graduate Studies in Mathematics, vol. 163, American Mathematical Society, Providence, RI, 2015. Translated from the 2008 French edition by Patrick D. F. Ion.

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF NORTHERN BRITISH COLUMBIA, 3333
UNIVERSITY WAY, PRINCE GEORGE, BC V2N4Z9, CANADA

E-mail address: alia.hamieh@unbc.ca

E-mail address: mcclenaga@unbc.ca