Large deviations of sums of random variables

Andrew Granville\textsuperscript{a, b, 1} and Youness Lamzouri\textsuperscript{c}

\textsuperscript{a} Département de Mathématiques et Statistique, Université de Montréal, CP 6128 succ Centre-Ville, Montréal, QC H3C 3J7, Canada
\textsuperscript{b} Department of Mathematics, University College London, Gower Street, London WC1E 6BT, England
\textsuperscript{c} Institut Élie Cartan de Lorraine, Université de Lorraine, BP 70239, 54506 Vandoeuvre-lès-Nancy Cedex, France
(e-mail: andrew.granville@umontreal.ca; youness.lamzouri@univ-lorraine.fr)

Dedicated to the memory of Jonas Kubilius

Abstract. In this paper, we investigate the large deviations of sums of weighted random variables that are approximately independent, generalizing and improving some results of Montgomery and Odlyzko. We are motivated by examples arising from number theory, including the sequences $p^t$, $\chi(p)$, $\chi_d(p)$, $\lambda_f(p)$, and $\text{Kl}_q(a-n, b)$, where $p$ ranges over the primes, $t$ varies in a large interval, $\chi$ varies among all characters modulo $q$, $\chi_d$ varies over quadratic characters attached to fundamental discriminants $|d| \leq x$, $\lambda_f(n)$ are the Fourier coefficients of holomorphic cusp forms $f$ of (a large) weight $k$ for the full modular group, and $\text{Kl}_q(a, b)$ are the normalized Kloosterman sums modulo a large prime $q$, where $a, b$ vary in $(\mathbb{F}_q)^\times$.

MSC: 11M26, 11M06, 60F10

Keywords: zeta functions, distributions, moment generating function

1 Introduction

In this paper, we give estimates for the probability of extremely large deviations of sums of weighted random variables that are approximately independent. Our results generalize and improve results in a wide variety of questions of interest to analytic number theory (arguably in the spirit of some of Kubilius’ work). In this section, we will carefully formulate seemingly obscure independence hypotheses on distributions of sums of weighted random variables and then state results for such distributions near to the high end of their feasible range. Then, in the next two sections, we exhibit applications of our results to many examples considered in the literature.

Let $A = \{1 \leq q_1 \leq q_2 \leq \cdots\}$ be a given infinite sequence of real numbers that is fairly well behaved in that if $x \geq 2^\beta$, then

$$A(x) := \#\{n: q_n \leq x\} = \alpha \int_{\beta}^x \left(\log \frac{t}{\beta}\right)^{A-1} \left(1 + O\left(\frac{1}{(\log t)^B}\right)\right) dt$$

1 The author is partially supported by grants from NSERC (Canada) and by European Research Council grant, agreement No. 670239.
for some constants \( \alpha, \beta > 0 \) and \( A \geq 0 \), where \( B > \max\{A, 1\} \), and \( \beta' = (2 + \beta)^2 \). In several number theory applications, we are interested in the distribution of large values of

\[
H_X(Q) := \sum_{q_n \leq Q} \frac{X(n)}{q_n},
\]

where \( Q \) is large (perhaps \( \infty \)), and the \( X(n) \) are real-valued random variables that are "independent enough". The asymptotics for \( A(x) \) implies that the corresponding harmonic sum

\[
H(x) := \sum_{q_n \leq x} \frac{1}{q_n}
\]
diverges as \( x \to \infty \), so the sum in (1.1) can get arbitrarily large if the \( X(n) \) are supported, say, on \([−1, 1]\), and \( Q \to \infty \).

For examples, we might take:

- The \( q_n \) to be the positive integers \((\alpha = 1, \alpha \beta = 1, A = 1)\);
- The \( q_n \) to be the primes \((\alpha = 1, \alpha \beta = 1, A = 0)\); or even
- The \( q_n \) to be the absolute values of the zeros of the Riemann zeta-function \((\alpha = 1/2\pi, \alpha \beta = 1, A = 2)\), counting each zero and its conjugate.\(^2\)

In each of these cases, \( B \) may be taken arbitrarily large.

What do we mean by the random variables being "independent enough"? Since \( 1 \) and \( \pi \) are linearly independent over the rationals, we can deduce that the pair \((e^{it}, e^{2\pi it})\) is equidistributed in \( U^2 \) (where \( U \) is the unit circle) as \( t \) varies over a long segment over the real line; that is, \( e^{it} \) and \( e^{2\pi it} \) act more or less independently, despite being obviously multiplicatively dependent. However, we might want to assume that they appear to be independent on a not too long interval, and also with several real numbers (in place of 1 and \( \pi \)) that we believe are linearly independent over \( \mathbb{Q} \).

Our goal is to determine the probability that the sum in (1.1) is \( > V \), with \( V \) large, that is, to determine

\[
\Phi_Q(V) = \Phi_Q(X, V) := P(H_X(Q) > V),
\]

the probability that \( H_X(Q) > V \). This will depend on several parameters, most importantly, on \( A \), and some invariants of the probability distributions of our random variables.

### 1.1 Families of random variables

Our goal is to model a sequence of random variables \( X(n) \) that are "independent enough" by a sequence of independent random variables \( \mathbb{Y}(n) \), for which we determine the distribution of large values of \( H_{\mathbb{Y}}(Q) \). We then show that in fact the model provides the correct probability for \( H_X(Q) \) in an appropriate range. To explicitly model the \( X(n) \) by the \( \mathbb{Y}(n) \), we will assume the following explicit connection:

**Approximate independence hypothesis AIH(\( L, Q \))**

The random variables \( \{X(n)\}_{q_n \leq Q} \) satisfy AIH(\( L, Q \)) if there exist independent random variables \( \{\mathbb{Y}(n)\}_{q_n \leq Q} \) such that

\[
E(X(n_1) \cdots X(n_\ell)) = E(\mathbb{Y}(n_1) \cdots \mathbb{Y}(n_\ell)) + O(e^{-L})
\]

for all \( \ell \leq L \) and all choices of \( n_1, \ldots, n_\ell \leq A(Q) \).\(^3\)

---

\(^2\) We can work with any individual \( L \)-function or appropriate families of \( L \)-functions since they also satisfy the hypothesis with \( A = 2 \).

\(^3\) Throughout, \( E(T) \) denotes the expected value of a random variable \( T \).
It is convenient to divide the $\mathbb{Y}(n)$ by an appropriate constant (which can be incorporated into the $q_n$) and reorder so that each $\mathbb{Y}(n)$ is supported on $[-1,1]$ with $E(\mathbb{Y}(n)) = 0$. It will also be useful if the $\mathbb{Y}(n)$ “converge” to a well-behaved universal distribution $\mathbb{Y}$ as $n \to \infty$:

Convergent sequence of random variables hypothesis CRVH($\mathbb{Y}$)

The independent random variables $\{\mathbb{Y}(n)\}_{n \geq 1}$ satisfy CRVH($\mathbb{Y}$) if

- Each $\mathbb{Y}(n)$ is supported on $[-1,1]$ with $E(\mathbb{Y}(n)) = 0$;
- There exists an absolute constant $c > 0$ such that

$$P\left(\mathbb{Y}(n) > 1 - \frac{1}{t}\right) \gg \exp\left(-ce^{\sqrt{t}}\right) \quad \text{for all } t \geq 1, \tag{1.2}$$

where the implicit constant is absolute;
- There exist independent random variables $\{\mathbb{Y}_n\}_{n \geq 1}$, each distributed identically to $\mathbb{Y}$, such that $\mathbb{Y}(n)$ depends on $\mathbb{Y}_n$ with $\mathbb{Y}(n) = \mathbb{Y}_n + O(1/(\log n)^2)$.

We can deduce that if this holds, then $\mathbb{Y}$ is supported on $[-1,1]$ with $E(\mathbb{Y}) = 0$ and satisfies (1.2). We now define some invariants associated with $\mathbb{Y}$: Let

$$f_\mathbb{Y}(t) := \begin{cases} \log E(e^{t\mathbb{Y}}) & \text{if } 0 \leq t < 1, \\ \log E(e^{t\mathbb{Y}}) - t & \text{if } t \geq 1, \end{cases} \quad \text{with } \eta_\mathbb{Y} = \int_0^\infty f_\mathbb{Y}(u) \frac{du}{u^2}. \tag{1.3}$$

We will show that this last integral converges in Lemma 8 because of the hypotheses on $\mathbb{Y}$.

1.2 The diagonal sum

Applying partial summation to the counting function $A(\cdot)$, we find that (see Lemma 6) if $y \geq 4\beta'$, then

$$H(y) = \begin{cases} \frac{\alpha}{4}(\log \frac{y}{\beta})^A + C_A + o_{y \to \infty}(1) & \text{if } A > 0, \\ \alpha \log \log \frac{y}{\beta} + C_A + o_{y \to \infty}(1) & \text{if } A = 0, \tag{1.4} \end{cases}$$

for some constant $C_A$, since we assumed that $B > A$.

If $\Phi_Q(V) > 0$, then $V \leq H(Q)$, which implies that $W(V) \leq (1 + o(1)) \log(Q/\beta)$ if we define

$$W = W(V) := \begin{cases} \left(\frac{\alpha}{A}(V - C_A)\right)^{1/A} & \text{if } A > 0, \\ \exp\left(\frac{1}{A}(V - C_A)\right) & \text{if } A = 0, \end{cases}$$

for $V \geq 1 + C_A$.

We obtain estimates for the distribution of large values of $H_\mathbb{Y}(Q) = \sum_{q_n \leq Q} \mathbb{Y}(n)/q_n$, which are more precise than those of Montgomery and Odlyzko [16], who obtained upper and lower bounds for $\Phi_Q(\mathbb{Y}, V)$ in the case where $\mathbb{Y}(n)$ are independent and identically distributed but considered more general weights $q_n$.

**Theorem 1.** Let $Q$ be large and suppose that $\{\mathbb{Y}(n)\}_{q_n \leq Q}$ is a sequence of independent random variables satisfying CRVH($\mathbb{Y}$). Then

$$\Phi_Q(\mathbb{Y}, V) = \exp\left(-\alpha \beta e^{-\eta_\mathbb{Y} - 1} W^{A - 1} e^W \left(1 + O\left(\frac{\sqrt{\log W}}{W^{1/2}} + \frac{1}{W(B - 1/2)}\right)\right)\right)$$

---

4 That is, let $\mu_n = E(\mathbb{Y}(n))$ and $m_n := \max \{\mathbb{Y}(n) - \mu_n\}$, and then we work with $\mathbb{Y}'(n) := (\mathbb{Y}(n) - \mu_n)/m_n$, so that $\sum_{q_n \leq Q} \mathbb{Y}(n)/q_n = \sum_{q_n \leq Q} \mathbb{Y}'(n)/q_n + \mu$, where $q_n = q_n/m_n$ and $\mu := \sum_{q_n \leq Q} \mu_n/q_n$. 

Lith. Math. J., 61(3):345–372, 2021.
for those values of $V$ for which

$$c \leq W = W(V) \leq \log Q - \log \log Q - \Theta'$$

for some suitably large constants $\Theta' > 0$ and $c \geq 1$.

It is arguably surprising that the specifics of the probability distribution $Y$ only seem to affect the constant term $e^{-\eta V}$ in $\Phi_Q(Y, V)$. All other parts of this formula are determined from the asymptotics for $A(x)$, except for the constant $C_A$ (which appears in the definition of $W(V)$), which largely depends on the small $q_n$.

We now give our main result, which is applicable to problems in analytic number theory, by starting with a sequence of random variables $X(n)$ that can be appropriately approximated by a sequence of independent random variables $Y(n)$.

**Theorem 2.** Let $Q$ be large. Suppose that the random variables $\{X(n)\}_{n \leq Q}$ satisfy AIH(L, Q), so they are approximated by the independent random variables $\{Y(n)\}_{n \leq Q}$ with $(\log Q)^2A^2 \leq L \leq A(Q)$. Moreover, suppose that $\{Y(n)\}_{n \leq Q}$ satisfy CRVH($Y$). Then

$$\Phi_Q(X, V) = \exp\left(-\alpha\beta e^{-\eta V}W^{-1}e^W\left(1 + O\left(\frac{\log W}{W^{1/2}} + \frac{1}{W^{(B-1)/2}}\right)\right)\right),$$

where

$$c \leq W = W(V) \leq \log L - A \log \log Q - 1_{A=0} \cdot \log \log Q - \Theta'$$

for some suitably large constants $\Theta' > 0$ and $c \geq 1$. ($1_C$ is the indicator function for an event $C$.)

**Remark 1.** If AIH(L, Q) holds for $L = A(Q) \sim \alpha Q(\log(Q/\beta))^{A-1}$, then the range for $W$ is

$$1 \ll W(V) \ll \log Q - \log \log Q - 1_{A=0} \cdot \log \log Q - \Theta'',$$

much like in Theorem 1.

If $A > 1$, then $W = ((A/\alpha)V)^{1/A} + O(1/V^{1-1/A})$, so we may replace $W$ by $W' := ((A/\alpha)V)^{1/A}$ in Theorem 2 at the cost of an additional error term of $O(1/W^{A-1}) = o(1)$). Therefore the result becomes

$$\log \Phi_Q(V) \sim -\kappa V\alpha^{-1},$$

where $\kappa = \beta A e^{-\eta^{-1}}$.

If $A = 1$, then we are at liberty to select $\beta$ since its contribution can be incorporated into the constant $C_A$ (by changing $C_A$ to $C_A' = C_A - \alpha \log \beta$ in (1.4)), and it is most convenient to let $\beta = 1$. Therefore the result becomes

$$\log \Phi_Q(V) \sim -\kappa V/\alpha,$$

where $\kappa = \alpha e^{-\eta^{-1}} - C_A/\alpha$.

Under the same assumptions in Theorem 2, we will establish that the sum $H_X(Q)$ converges in distribution to that of some random variable $W$. We will also prove that the characteristic function of $W$ is rapidly decreasing, which implies that $W$ is absolutely continuous and has a uniformly bounded probability distribution function. Furthermore, we will use the Berry–Esseen inequality to bound the rate of convergence.

**Theorem 3.** Under the same assumptions as in Theorem 2, there exists an absolutely continuous random variable $W$ such that $H_X(Q)$ converges in distribution to $W$ as $Q \to \infty$. Moreover, we have

$$\sup_{V \in \mathbb{R}} |\Phi_Q(X, V) - P(W > V)| \ll \frac{H(Q)}{L} + \frac{(\log Q)^{(A-1)/3}}{Q^{1/3}}.$$

The plan of the paper is as follows. In the next section, we describe some examples of sums of independent random variables $H_Y(Q)$ for which Theorem 1 holds. In Section 3, we exhibit sequences of random variables $X(n)$ that satisfy hypothesis AIH and present some applications of our results. In Section 4, we describe some estimates for the harmonic sum $H(Q)$, which will be useful in our subsequent work. In Section 5, we
investigate the cumulant generating function of $H_Y(Q)$. Theorem 1 will be proved in Section 6. Finally, in Section 7, we explore how to use hypothesis AIH and prove Theorems 2 and 3.

The methods that we use may be unsurprising to the experts, and the results that we obtain rely on the unusually strong hypotheses. However, we justify these remarks by noting that our results are applicable to a very wide selection of number theory distribution problems, and the ranges in our theorems are longer than in the literature on relevant questions, extending very close to the boundary of what is possible. Example 3 in the next section highlights the value of these results, improving the range in what is believed about the distribution of the error term in the prime number theorem, certainly a very well-studied problem.

2 Examples of sums of random variables $H_Y(Q)$ where Theorem 1 holds

Example 1 [The integers]. We are interested in the large values of

$$\sum_{n \geq 1} \frac{Y(n)}{n},$$

where $\{Y(n)\}_{n \geq 1}$ is a sequence of independent random variables satisfying CRVH and converging in distribution to a random variable $Y$, which is supported on $[-1, 1]$ with $E(Y) = 0$, and for which (1.2) holds. An important example is $Y(n) = \text{Re}(X(n))$ where for each $n$, $X(n)$ is uniformly distributed on the unit circle. Taking $\alpha = 1, \beta = 1, A = 1, a$ and $Q \to \infty$ in Theorem 1, for all real numbers $V > 0$, we obtain

$$\Phi(Y, V) = P\left(\sum_{n \geq 1} \frac{Y(n)}{n} > V\right) = \exp\left(-\left(1 + o(1)\right)e^{V - \gamma - \eta V - 1}\right),$$

where $\gamma$ is the Euler–Mascheroni constant.

Example 2 [The primes]. Let $\{Y(p)\}_p$ be a sequence of independent random variables indexed by the primes and satisfying CRVH($Y$). Then, taking $\alpha = 1, \beta = 1, A = 0, a$ and $Q \to \infty$ in Theorem 1, we obtain

$$\Phi(Y, V) = P\left(\sum_p \frac{Y(p)}{p} > V\right) = \exp\left(-\left(1 + o(1)\right)e^{e^{V - C - V + C - \eta V - 1}}\right),$$

where

$$C := \lim_{Q \to \infty} \left(\sum_{p \leq Q} \frac{1}{p} - \log \log Q\right) = 0.2614972128 \ldots .$$  \hspace{1cm} (2.1)

By renormalizing we obtain

$$P\left(\prod_p e^{Y(p)/p} > e^{CT}\right) = \Phi(Y, \log T + C) = \exp\left(-\left(1 + o(1)\right)e^{T - \eta T - 1}\right).$$

In particular, this result applies to the sum

$$\text{Re}\left(\sum_{p \text{ prime}} \frac{X(p)}{p}\right),$$

Lith. Math. J., 61(3):345–372, 2021.
where each $\chi(p)$ is independently uniformly distributed on the unit circle. However, $\prod_p e^{\chi(p)/p}$ is not exactly the same as

$$L(1, \chi) := \prod_p \left( 1 - \frac{\chi(p)}{p} \right)^{-1},$$

which is used to model the values of $\zeta(1 + it)$ and $L(1, \chi)$, though it differs by no more than a bounded (multiplicative) constant. Therefore, up to that constant, (2.2) as each $\chi(p)$ is independently uniformly distributed on the unit circle models the distribution of values of $\log |\zeta(1 + it)|$ as $t$ varies, as well as the distribution of values of $\log |L(1, \chi)|$ as we vary over all of the nonprincipal characters mod $q$ as $q \to \infty$, up to a constant. Moreover, (2.2) as each $\chi(p)$ is independently uniformly distributed at each of the $m$th roots of unity, models the distribution of values of $\log |L(1, \chi)|$ as we vary over the nonprincipal characters of order $m$, over moduli $q$ for which $m$ is a possible order, up to a constant.

To deduce an estimate for $P(|L(1, \chi)| > T)$, we can show that when $|L(1, \chi)|$ is large, so is $\frac{\zeta(1)}{(1 - \frac{1}{p})} e^{-\gamma}\frac{1}{p}$. Therefore we can model the error term in the weighted prime number theorem by the

$$-\frac{\zeta(1)}{(1 - \frac{1}{p})} e^{-\gamma}\frac{1}{p} = o(1).$$

The effect is to replace the constant $C$ above by $\gamma \approx 0.5772156649 \ldots$ to obtain

$$P\left( |L(1, \chi)| > e^\gamma T \right) = \exp\left( -\left( 1 + o(1) \right) \frac{e^{T - \eta_\chi - 1}}{T} \right),$$

a result that was obtained unconditionally by Granville and Soundararajan [7] by a different method. In the next subsection, we will see how to use our methods to reprove this unconditionally.

**Example 3 [Sums over zeros of the Riemann zeta-function].** The explicit formula [3, Chap. 17, (9)] for $\psi(x) := \sum_{n \leq x} \Lambda(n)$ implies that, under the assumption of the Riemann hypothesis, the error term in the weighted prime number theorem is

$$\psi(x) = -\sum_{\rho=1/2+i\gamma : \zeta(\rho)=0} \frac{x^{\rho}}{\rho} + o(1).$$

If $\zeta(1/2 + i\gamma) = 0$, then $\zeta(1/2 - i\gamma) = 0$. Otherwise, we believe that the $\gamma$ are completely linearly independent and even that the $x^{\rho}$, with $\gamma > 0$, are truly independent. This means that we can analyze the value of this sum by replacing each $x^{\rho}$ by a random variable that is uniformly distributed on the unit circle. However, this sum is not quite in the correct form since the $\rho$ are complex numbers, but we can simply adjust our random variable by the angle $-|\rho|/p$. Therefore we can model the error term in the weighted prime number theorem by the values of $2 \Re \left( \sum_{n \geq 1} \frac{\chi(\gamma_n)}{|\rho_n|} \right)$, where $0 < \gamma_1 < \cdots$ are the successive ordinates of the zeros $1/2 + i\gamma_n$ of $\zeta(s)$. Taking $\alpha = 1/2\pi$, $\alpha \beta = 1$, and $A = 2$ in Theorem 1, we obtain

$$P\left( 2 \sum_{n \geq 1} \frac{\Re(\chi(n))}{|\rho_n|} > T \right) = \exp\left( -e^{-\eta_\chi - 1} (\sqrt{2\pi T} + O(\sqrt{\log T})) \exp(\sqrt{2\pi T}) \right).$$

\[5\] Similarly, $\exp(\Re(\sum_{p \text{prime}} \chi(p)/p)))$ differs from $|L(1, \chi)|$ by a bounded (multiplicative) constant.
Montgomery and Odlyzko obtained the $\sqrt{2\pi}$ parts of this formula but were unable to show the asymptotic behavior of $\log \Phi(V)$ (see [17, (13.48)]). This probability is around $1/u$ when

$$T = \frac{1}{2\pi} \left( \log \log u - \log \log \log u + \eta \zeta + 1 + o(1) \right)^2.$$  

Montgomery and Odlyzko then used their estimates to predict the largest error term in the prime number theorem:

$$\limsup_{x \to \infty} \frac{\psi(x) - x}{\sqrt{x} \log \log x} = \frac{1}{2\pi}$$

and

$$\liminf_{x \to \infty} \frac{\psi(x) - x}{\sqrt{x} \log \log x} = -\frac{1}{2\pi}$$

taking $x = e^u$, which is further supported by calculations, and an alternative perspective in Monach’s thesis [14] (see also [17, (15.25), (15.26)]). Our work suggests, more precisely, that we can perhaps find arbitrarily large $x$ with

$$\psi(x) - x = \pm \frac{1}{2\pi} \sqrt{x} \left( \log \log x \log \log \log x + \eta \zeta + 1 + o(1) \right)^2$$

but no larger.

### 2.1 A more complicated example: The values of an $L$-function

Here we have

$$Z(p) = -\log \left( 1 - \frac{\mathbb{V}_p}{p} \right),$$

where each $\mathbb{V}_p$ is uniformly distributed on the unit circle or perhaps uniformly distributed on the $m$th roots of unity for some integer $m \geq 2$. We have to begin by recentering these random variables. First, note that

$$E(Z(p)) = \sum_{k \geq 1} \frac{E(\mathbb{V}_{p,k}^{*})}{kp^k} = \sum_{k \geq 1} \frac{1}{kp^k} = -\frac{1}{m} \log \left( 1 - \frac{1}{p^m} \right)$$

writing $m = \infty$ if $\mathbb{V}_p$ is uniformly distributed on the unit circle. This implies that

$$-\log \left( 1 + \frac{1}{p} \right) + \frac{1}{m} \log \left( 1 - \frac{1}{p^m} \right) \leq \text{Re}(Z(p)) - E(Z(p)) \leq -\log \left( 1 - \frac{1}{p} \right) + \frac{1}{m} \log \left( 1 - \frac{1}{p^m} \right).$$

The upper bound is at least as large as the lower bound in absolute value (with equality when $m = 2$), and the upper bound is attained. Therefore define

$$\mathbb{V}(p) := \frac{\text{Re}(Z(p)) - E(Z(p))}{-\log(1-1/p) - E(Z(p))} = \text{Re} \left( \mathbb{V}_p \right) + O \left( \frac{1}{p} \right).$$

This sequence of random variables exactly satisfies CRVH($\mathbb{V}$). Now define $q_p$ by

$$\frac{1}{q_p} := -\log \left( 1 - \frac{1}{p} \right) - E(Z(p)),$$

so that $q_p = p + O(1)$, and

$$\sum_{p \leq Q} \frac{\mathbb{V}(p)}{q_p} = -\sum_{p \leq Q} \text{Re} \left( \log \left( 1 - \frac{\mathbb{V}_p}{p} \right) \right) - \frac{1}{m} \log \zeta(m) + O \left( \frac{1}{Q} \right).$$

Lith. Math. J., 61(3):345–372, 2021.
This means that \( \alpha = \alpha \beta = 1 \), and \( A = 0 \) with
\[
C_A = \lim_{y \to \infty} - \sum_{p \leq y} \left( 1 - \frac{1}{p} \right) - \sum_{p \leq y} E(Z(p)) - \log \log y = \gamma - \frac{1}{m} \log \zeta(m).
\]
Therefore by Theorem 1 we have
\[
P \left( \left| \prod_{p \leq Q} \left( 1 - \frac{Z(p)}{p} \right)^{-1} \right| > e^{\gamma} W \right) = \exp \left( - \frac{e^{W-\gamma-1}}{W} \left( 1 + O \left( \frac{\sqrt{\log W}}{W^{1/2}} \right) \right) \right)
\]
in the range \( 1 \ll W \ll \log Q - \log \log Q - \Theta' \). This corresponds to understanding the distribution of the values of \( \zeta(1 + it) \) as \( t \) varies, as well as the values of \( L(1, \chi) \) as \( \chi \) varies over all characters mod \( q \) for a large modulus \( q \).

Remark 2. This can be generalized rather nicely. Fix \( m > 1 \) (including \( m = \infty \)) and let \( C_m \) be the set of \( m \)th roots of unity. Let \( f_p(t) = t + \sum_{k \geq 1} a_k p t^k \in \mathbb{R}[t] \) with \( f_p(0) = 0 \), which converges absolutely in \( |t| < 1/p \). Let \( \mu_{p,m} = \sum_{m|k \geq 1} a_k p t^k \), and let \( F_p(t) = f_p(t/p) \). Assume that the maximum of \( |\Re(F_p(z) - \mu_{p,m})| \) over the \( m \)th roots occurs at \( z = 1 \).

Let \( Z(p) = F_p(Y_p) \). Then \( 1/q_p := F_p(1) - \mu_{p,m} \) with \( Y(p) = q_p(\Re(Z(p)) - \mu_{p,m}) \). Proceeding as before, we obtain
\[
P \left( \left| \prod_{p \leq Q} e^{Z(p)} \right| > e^{\gamma} W \right) = \exp \left( - \frac{e^{W-\gamma-1}}{W} \left( 1 + O \left( \frac{\sqrt{\log W}}{W^{1/2}} \right) \right) \right)
\]
in the range \( 1 \ll W \ll \log Q - \log \log Q - \Theta' \), where we have
\[
\prod_{p \leq y} e^{F_p(1)} \sim e^{\gamma} \log y.
\]

We can rewrite this in terms of the distance from the maximum:
\[
P \left( \left| \prod_{p \leq Q} e^{Z(p)-F_p(1)} \right| > \frac{W}{\log Q} \right) = \exp \left( - \frac{e^{W-\gamma-1}}{W} \left( 1 + O \left( \frac{\sqrt{\log W}}{W^{1/2}} \right) \right) \right).
\]
This corresponds to understanding the distribution of the values of \( L(1, \chi) \) as \( \chi \) varies over the characters of order \( m \) and conductor up to some large \( x \)—the main results in [6] yield a result like this when \( m = 2 \).

3 Examples of sequences satisfying AIH

3.1 The sequence \( p^it \)

For \( \Re(s) > 1 \), we let \( G(s) = \sum_{p \leq y} 1/p^s \). Then we have \( \zeta(s) = \exp(G(s)) R(s) \), where \( R \) is analytic in the half-plane \( \Re(s) > 1/2 \). Therefore, using Lemma 2.2 of [7], which is derived using the classical zero density estimates for \( \zeta(s) \) (see, e.g., [21, Thm. 9.19A]), we have
\[
G(1 + it) = \sum_{p \leq (\log T)^{C_0}} \frac{1}{p^{1+it}} + O \left( \frac{1}{\log T} \right)
\]
for all \( t \in [T, 2T] \) except a set of measure \( T^{4/5} \), where \( C_0 \) is an absolute constant, and \( T \) is large.
Large deviations of sums of random variables

We let $X(p) = \text{Re}(p^{it})$ for $t$ varying from $T$ to $2T$. Then we take $\mathcal{Y}(p) = \text{Re}(\mathcal{Y}_p)$, where each $\mathcal{Y}_p$ is a random variable uniformly distributed on the unit circle. We will see that this sequence satisfies the approximate independence hypothesis AIH$(L, Q)$ with $Q = (\log T)^{C_0}$ and $L \sim (C_0 \log \log T)$.

**Lemma 1.** Let $X(p) = \text{Re}(p^{it})$ and define $E_T(f(t)) = (1/T) \int_T^{2T} f(t) \, dt$. Suppose that $Q = (\log T)^{C_0}$, where $C_0$ is a positive constant. Let $L$ be such that $(e^2 Q)^{L/2} \leq T$. For all primes $p_1, \ldots, p_L \leq Q$ and any $\ell \leq L$, we have

$$E_T(X(p_1) \cdots X(p_{\ell})) = E(\mathcal{Y}(p_1) \cdots \mathcal{Y}(p_{\ell})) + O(e^{-L}),$$

where $\mathcal{Y}(p) = \cos \theta_p$ for each prime $p$, and $\{\theta_p\}$ is a sequence of independent random variables uniformly distributed on $[0, 2\pi]$.

**Proof.** Let $N = p_1 \cdots p_L$. The left-hand side above is

$$\frac{1}{T} \int_T^{2T} \left( \prod_{j=1}^{\ell} \frac{p_j^{it} + p_j^{-it}}{2} \right) \, dt = \frac{1}{2^\ell} \sum_{ab=N} \frac{1}{T} \int_T^{2T} \left( \frac{a}{b} \right)^{it} \, dt.$$

The terms with $a = b$ give exactly the right-hand side (as can be seen by expanding the right-hand side in an analogous way), so the difference between the two sides is

$$\leq \max_{ab=N, a \neq b} \left| \frac{1}{T} \int_T^{2T} \left( \frac{a}{b} \right)^{it} \, dt \right| = \max_{ab=N, a \neq b} \left| \left( \frac{a}{b} \right)^{it} \right| \frac{1}{T} \int_T^{2T} \left| \log \left( \frac{a}{b} \right) \right| \, dt \ll \max_{ab=N, a \neq b} \frac{1}{T} \left| \log \left( \frac{a}{b} \right) \right| \frac{1}{T} \int_T^{2T} \left| \log \left( \frac{a}{b} \right) \right| \, dt.$$

This is $\ll 1/T$ unless $a, b \approx \sqrt{N}$, in which case it is $\ll b/(T|b - a|) \ll \sqrt{N}/T \ll Q^{L/2}/T \ll e^{-L}$ as $N \leq Q^L$ and $e^{LQ^{L/2}} \leq T$. \(\square\)

Using this result in Theorem 2 with $Q = (\log T)^{C_0}$ and $L \sim (C_0 \log \log T)$, we deduce that

$$\frac{1}{T} \text{meas}\left\{ t \in [T, 2T]: \left| \exp\left(G(1 + it)\right)\right| > e^C V \right\} = \exp\left(-\left(1 + o(1)\right) \frac{e^{V - \Theta(1)}}{V}\right)$$

for $V$ in the range $1 \ll V \ll \log_3 T - \log_3 T - \Theta$, where here and throughout $\log_k$ is the $k$th iterate of the natural logarithm, $C$ is defined in (2.1), $\Theta$ is a large constant, and $\mathbb{Y} = \text{Re}(\mathbb{W})$ where $\mathbb{W}$ is uniformly distributed on the unit circle. Granville and Soundararajan [7] proved a similar result (in a slightly wider range) for $|\zeta(1 + it)|$. Note that in accordance with Section 2.1, we can obtain the analogue of (3.1) for $|\zeta(1 + it)|$ by proving the analogous AIH hypothesis for the random variables

$$X(p) = \frac{-\text{Re}\log(1 - 1/p^{1+it})}{-\log(1 - 1/p)}$$

(instead of $\text{Re}(p^{it})$), which follows along the same lines of Lemma 1 but is slightly more technical, so we omit it.

### 3.2 The sequence $\chi(p)$ for characters $\chi \pmod{q}$

This sequence is similar to the above example of $p^{it}$. We let $X(p) = \text{Re}(\chi(p))$ for $\chi$ varying among all characters $\chi \pmod{q}$, where $q$ is a large integer. Then we let $\mathcal{Y}(p) = \text{Re}(\mathcal{Y}_p)$ where each $\mathcal{Y}_p$ is a random variable uniformly distributed on the unit circle.

We will see that this sequence satisfies the approximate independence hypothesis AIH$(L, Q)$ with $Q = (\log q)^{C_0}$ and any $L$ such that $Q^L < q$. 

Lith. Math. J., 61(3):345–372, 2021.
Lemma 2. Let $X(p) = \text{Re}(\chi(p))$ and define $E_q(f(\chi)) = (1/\phi(q)) \sum_{\chi \pmod q} f(\chi)$. Suppose that $Q = (\log q)^{C_0}$, where $C_0$ is a positive constant. Take $L$ such that $Q^L < q$. For all primes $p_1, \ldots, p_\ell \leq Q$ and any $\ell \leq L$, we have

$$E_q(X(p_1) \cdots X(p_\ell)) = E(Y(p_1) \cdots Y(p_\ell)), \quad (3.2)$$

where $Y(p) = \cos \theta_p$ for each prime $p$, and $\{\theta_p\}$ is a sequence of independent random variables uniformly distributed on $[0, 2\pi]$.

**Proof.** Let $N = p_1 \cdots p_\ell$. The left-hand side of (3.2) is

$$\frac{1}{\phi(q)} \sum_{\chi \pmod q} \left( \prod_{j=1}^{\ell} \frac{\chi(p_j) + \overline{\chi}(p_j)}{2} \right) = \frac{1}{2^\ell} \sum_{ab=N} \frac{1}{\phi(q)} \sum_{\chi \pmod q} \chi\left(\frac{a}{b}\right).$$

As in the proof of Lemma 1, the contribution of the diagonal terms with $a = b$ equals

$$E(Y(p_1) \cdots Y(p_\ell)).$$

Moreover, since $Q^\ell < q$, we have $a, b < q$, and hence by the orthogonality relations of the characters the contribution of the off-diagonal terms $a \neq b$ is 0. □

**Remark 3.** We observe that the same proof works to show that the sequence $\chi(p)$ satisfies AIH$(L, Q)$ under the same assumptions on $Q$ and $L$, where $\chi$ varies over the set of primitive characters $\pmod q$, and $q$ is a large prime. In this case, we should add the error term $O(1/q)$ to the right-hand side of (3.2).

We also remark that we might extend the range $Q^L < q$ by averaging instead over all moduli $q \leq X$ for a large number $X$.

As in the previous section, we can apply this result together with Theorem 2 to investigate the distribution of $\text{Re}(G_\chi(1))$, where

$$G_\chi(s) = \sum_p \frac{\chi(p)}{p^s}, \quad (3.3)$$

for $\text{Re}(s) > 1$. Indeed, Lemma 8.2 of [5], together with the classical zero density estimates for the family of Dirichlet $L$-functions (see, e.g., [15, Thm. 12.1]), implies that for a positive constant $C_0$, we have

$$G_\chi(1) = \sum_{p \leq (\log q)^{C_0}} \frac{\chi(p)}{p} + O\left(\frac{1}{\log q}\right) \quad (3.4)$$

for all nonprincipal characters modulo a large prime $q$, except for a set of size $\ll \sqrt{q}$. Therefore, combining Lemma 2 with Theorem 2 with $Q = (\log q)^{C_0}$ and $L = [\log q/(C_0 \log \log q)] - 1$, we deduce that

$$\frac{1}{q-1} \left| \{ \chi \pmod q : |\exp(G_\chi(1))| > e^{C V} \right| = \exp\left( - (1 + o(1)) \frac{e^{V - \eta_C - 1}}{V} \right) \quad (3.5)$$

for $V$ in the range $1 \ll V \ll \log_2 q - \log_3 q - \log_4 q - \Theta$, where $C$ is defined in (2.1), $\Theta$ is a large constant, and $Y = \text{Re}(W)$, where $W$ is uniformly distributed on the unit circle. Granville and Soundararajan [7] proved the analogue of (3.5) for $|L(1, \chi)|$. 
3.3 Values of the quadratic characters at the primes

Here we consider the sequence \( \chi_d(p) \), where \( \chi_d \) is the Kronecker symbol modulo \( d \), and \( d \) varies over fundamental discriminants \( |d| \leq x \). The associated independent random variables are such that \( \mathbb{Y}(p) = 1 \) or \(-1\) each with probability \( p/(2(p+1)) \) and \( \mathbb{Y}(p) = 0 \) with probability \( 1/(p+1) \).

In this case, using the estimates of Graham and Ringrose [4] on character sums to smooth moduli, we can obtain a nontrivial improvement on the range where the AIH is valid. Indeed, we will see that this sequence satisfies AIH(\( L, Q \)) with \( Q = (\log x)^C_1 \) and any \( L \) such that \( Q^L \leq x^{C_2 \log x} \), where \( C_1 \) is an arbitrary positive constant, and \( C_2 > 0 \) is a suitably small constant depending on \( C_1 \). Note that by using the Pólya–Vinogradov or the Burgess inequalities instead, we would obtain a range of the form \( Q^L \leq x^{C_3} \) for some positive constant \( C_3 \).

**Lemma 3.** Let \( x \) be large, and let \( \mathcal{F}(x) \) denote the set of fundamental discriminants \( d \) with \( |d| \leq x \). Let \( \chi(p) = \chi_d(p) \) and define \( E_x(f(d)) = (1/|\mathcal{F}(x)|) \sum_{d \in \mathcal{F}(x)} f(d) \). Suppose that \( Q = (\log x)^C_1 \) for some positive constant \( C_1 \). Take \( L \) such that \( Q^L < x^{C_2 \log x} \) for some suitably small constant \( C_2 > 0 \) depending on \( C_1 \). For all primes \( p_1, \ldots, p_{\ell} \leq Q \) and any \( \ell \leq L \), we have

\[
E_x(\chi(p_1) \cdots \chi(p_{\ell})) = E(\mathbb{Y}(p_1) \cdots \mathbb{Y}(p_{\ell})) + O(e^{-L}),
\]

where \( \mathbb{Y}(p) \) is a sequence of independent random variables taking the values \( 1 \) or \(-1\) with probability \( p/(2(p+1)) \) and the value \( 0 \) with probability \( 1/(p+1) \).

To prove this result, we need the following lemma.

**Lemma 4.** Let \( x \) be large, and let \( n \geq 1 \) be an integer. Suppose that all prime factors of \( n \) are below \( y \), where \( y \leq x \) is a real number. Let \( k \geq 1 \) be any integer and put \( K = 2^k \). Then

\[
\frac{1}{|\mathcal{F}(x)|} \sum_{d \in \mathcal{F}(x)} \chi_d(n) \begin{cases} = \prod_{p|n}(\frac{p}{p+1}) + O(x^{-1/4}) & \text{if } n \text{ is a square,} \\ \ll x^{-k/(8K)} \prod_{p|n}(1 + \frac{1}{p^{1/3}n^{1/(7K)}\tau(n)^{1/2}/K} & \text{otherwise,} \end{cases}
\]

where \( \tau(n) \) is the divisor function.

**Proof.** When \( n \) is a square, the asymptotics for the character sum is standard since in this case \( \sum_{d \in \mathcal{F}} \chi_d(n) \) corresponds to the number of fundamental discriminants \( d \) that are coprime to \( n \) and such that \( |d| \leq x \). On the other hand, in the case where \( n \) is not a square, the stated bound corresponds to Lemma 4.2 of [6], which is a consequence of Theorem 5 of Graham and Ringrose [4]. \( \square \)

**Proof of Lemma 3.** Let \( N = p_1 \cdots p_{\ell} \). First, note that for any prime \( p \) and positive integer \( a \), we have

\[
E(\mathbb{Y}^a(p)) = \begin{cases} \frac{p}{p+1} & \text{if } a \text{ is even,} \\ 0 & \text{if } a \text{ is odd.} \end{cases}
\]

Hence we deduce that

\[
E(\mathbb{Y}(p_1) \cdots \mathbb{Y}(p_{\ell})) = \begin{cases} \prod_{p|N}(\frac{p}{p+1}) & \text{if } N \text{ is a square,} \\ 0 & \text{otherwise.} \end{cases}
\]

Let \( k = \lceil \log_3 x / \log 2 \rceil \) such that \( K = 2^k \in ((\log_2 x)/2, \log_2 x] \). Then Lemma 4 implies that

\[
E_x(\chi(p_1) \cdots \chi(p_{\ell})) = \frac{1}{|\mathcal{F}(x)|} \sum_{d \in \mathcal{F}(x)} \chi_d(N) = E(\mathbb{Y}(p_1) \cdots \mathbb{Y}(p_{\ell})) + E_1(N),
\]

Lith. Math. J., 61(3):345–372, 2021.
where
\[ E_1(N) \ll \begin{cases} x^{-1/4} & \text{if } N \text{ is a square}, \\ x^{-k/(8K)} \prod_{p|N} \left(1 + \frac{1}{p-1-k/(8K)}\right) N^{1/(7K)} \tau(N)^{k^2/K} & \text{otherwise}. \end{cases} \]
Therefore, if \( N \) is a square, then we obtain the desired bound \( E_1(N) \ll e^{-L} \). Otherwise,
\[ \prod_{p|N} \left(1 + \frac{1}{p-1-k/(8K)}\right) \ll \tau(N) \ll 2^L \]
and
\[ N^{1/(7K)} \ll Q^{\ell/(7K)} \ll Q^{L/(7K)} \ll x^{2C_2 \log x}/(7 \log_2 x). \]
Therefore
\[ E_1(N) \ll x^{-(\log x)/((8 \log 2) \log_3 x)} 4^L (\log x)^{C_1} x (2C_2 \log_3 x)/(7 \log_2 x) < e^{-L} \]
since \( L \leq (C_2 \log x \log_3 x)/(C_1 \log_2 x) \), provided that \( (2/7 + (1 + \log 4)/C_1)/C_2 < 1/(8 \log 2) \).

As an application, we can deduce from this result together with the analogue of (3.4) for quadratic characters (which follows from Lemma 8.2 of [5] together with the zero density result of Heath-Brown [8] for Dirichlet \( L \)-functions attached to quadratic characters) and Theorem 2 with \( Q = (\log x)^{C_1} \) and \( L = [C_2 \log x \log_3 x]/(\log_2 x) \) that
\[ \frac{1}{|F(x)|} |\{d \in F(x) : \exp(G_{\chi_d}(1)) > e^{CV}\}| = \exp \left(-\left(1 + o(1)\right) \frac{e^{V - \eta - 1}}{V}\right) \tag{3.6} \]
for \( V \) in the range \( 1 \ll V \ll \log_2 x - \log_3 x - \Theta \), where \( G_{\chi} \) is defined in (3.3), \( C \) is defined in (2.1), \( \Theta \) is a large constant, and \( V \) takes the values 1 and \( -1 \) with equal probability \( 1/2 \). Granville and Soundararajan [6] proved the analogue of (3.6) for \( |L(1,\chi_d)| \) in a slightly wider range for \( V \). Again, we can recover their result (in our range of \( V \)) by modifying the sequence that satisfies the AIH from \( \chi_d(p) \) to
\[ X(p) = \frac{-\log(1 - \chi_d(p)/p)}{-\log(1 - 1/p)}. \]

### 3.4 Fourier coefficients of holomorphic cusp forms at the primes

Let \( \mathcal{H}_k \) be the set of Hecke eigencuspforms of weight \( k \) for the full modular group \( SL(2, \mathbb{Z}) \). We will assume that the weight \( k \) is large and note that \( \mathcal{H}_k \) contains about \( k/12 \) forms. Given \( f \in \mathcal{H}_k \), we write its Fourier expansion as
\[ f(z) = \sum_{n=1}^{\infty} \lambda_f(n) n^{(k-1)/2} e^{2\pi i n z}, \]
where we have normalized the Fourier coefficients so that the Hecke eigenvalues \( \lambda_f(n) \) satisfy Deligne’s bound \( |\lambda_f(n)| \ll \tau(n) \). In view of the Petersson trace formula (see (3.9)), it is convenient to use the “harmonic weights”
\[ \omega_f := \frac{\Gamma(k-1)}{(4\pi)^{k-1} \langle f, f \rangle}, \]
where \( \langle f, g \rangle \) denotes the Petersson inner product. In particular, we have

\[
\sum_{f \in \mathcal{H}_k} \omega_f = 1 + O\left(k^{-5/6}\right),
\]

and hence the harmonic weight is close to the natural weight \( 1/|\mathcal{H}_k| \) on average. These facts are standard and may be found in [9]. We now define

\[
E_k(g(f)) := \frac{1}{|\mathcal{H}_k|} \sum_{f \in \mathcal{H}_k} \omega f g(f),
\]

where

\[
\sum_{f \in \mathcal{H}_k} g(f) = \sum_{f \in \mathcal{H}_k} \omega f g(f), \quad \text{and} \quad |S|_h = \sum_{f \in S} 1 \quad \text{for} \ S \subseteq \mathcal{H}_k.
\]

**Lemma 5.** Let \( \lambda_f(p) = \lambda_f(p)/2 \). Suppose that \( Q = (\log k)^{C_0} \) for some positive constant \( C_0 \). Take \( L \) such that \( Q^L \leq k^2/10 \). For all primes \( p_1, \ldots, p_\ell \leq Q \) and any \( \ell \leq L \), we have

\[
E_k(\mathcal{X}(p_1) \cdots \mathcal{X}(p_\ell)) = E(\mathcal{Y}(p_1) \cdots \mathcal{Y}(p_\ell)) + O\left(e^{-L}\right),
\]

where \( \mathcal{Y}(p) = \cos \theta_p \) and \( \theta_p \) is a sequence of independent random variables distributed according to the Sato–Tate measure \( d\mu_{ST} = (2/\pi)(\sin \theta)^2 \, d\theta \) on \([0, \pi]\).

**Proof.** Let \( p \) be a prime number, and let \( a \) a positive integer. We start by expressing \( \lambda_f(p)^a \) in terms of \( \lambda_f(p^b) \) for \( 0 \leq b \leq a \) (note that \( \lambda_f(n) \) is not completely multiplicative). By Deligne’s bound \( |\lambda_f(p)| \leq 2 \) we may write \( \lambda_f(p) = 2 \cos \theta_f(p) \) for some \( \theta_f(p) \in [0, \pi] \). Recall that for any \( b \geq 0 \), we have that \( \lambda_f(p^b) = \sin((b + 1)\theta_f(p))/\sin \theta_f(p) \). We now use the fact that the functions \( \{S_n\}_{n \geq 0} \) defined by

\[
S_n(\theta) := \frac{\sin((n + 1)\theta)}{\sin \theta}
\]

form an orthonormal basis of \( L^2([0, \pi], \mu_{ST}) \). This implies

\[
\lambda_f(p)^a = 2^a \left(\cos \theta_f(p)\right)^a = 2^a \sum_{0 \leq b \leq a} C_a(b) \frac{\sin((b + 1)\theta_f(p))}{\sin \theta_f(p)} = 2^a \sum_{0 \leq b \leq a} C_a(b) \lambda_f(p^b),
\]

where the coefficients \( C_a(b) \) are defined by

\[
C_a(b) := \frac{2}{\pi} \int_0^\pi (\cos \theta)^a \sin \theta \sin((b + 1)\theta) \, d\theta.
\]

Note that (3.8) is equivalent to the standard Hecke relations

\[
\lambda_f(m)\lambda_f(n) = \sum_{d|(m,n)} \lambda_f\left(\frac{mn}{d^2}\right).
\]

Lith. Math. J., 61(3):345–372, 2021.
We now write \( p_1 \cdots p_\ell = q_1^{a_1} \cdots q_r^{a_r} \), where \( q_1, \ldots, q_r \) are distinct primes, and \( a_1 + \cdots + a_r = \ell \). Hence from (3.8) it follows that
\[
\mathbb{E}_k(\mathbb{X}(p_1) \cdots \mathbb{X}(p_\ell)) = \frac{1}{|H_k|} \sum_{f \in H_k} \prod_{j=1}^{\ell} \left( \cos \theta_f(q_j) \right)^{a_j} = \frac{1}{|H_k|} \sum_{f \in H_k} \prod_{j=1}^{r} \left( \sum_{0 \leq b_j \leq a_j} C_{a_j}(b_j) \lambda_f(q_j^{b_j}) \right)
\]
\[
= \sum_{0 \leq b_1 \leq a_1} \sum_{0 \leq b_2 \leq a_2} \cdots \sum_{0 \leq b_r \leq a_r} \prod_{j=1}^{r} C_{a_j}(b_j) \frac{1}{|H_k|} \sum_{f \in H_k} \lambda_f(q_1^{b_1} \cdots q_r^{b_r}).
\]

To estimate the inner sum, we use the following version of the Petersson trace formula, which follows from Lemma 2.1 of [19] together with (3.7):
\[
\frac{1}{|H_k|} \sum_{f \in H_k} \lambda_f(m) = 1_{m=1} + O(k^{-5/6})
\]
for all \( m \leq k^2/10 \). Since \( Q_L \leq k^2/10 \), we deduce that
\[
\mathbb{E}_k(\mathbb{X}(p_1) \cdots \mathbb{X}(p_\ell)) = \prod_{j=1}^{r} C_{a_j}(0) + E_1,
\]
where
\[
E_1 \ll k^{-5/6} 2^r \prod_{j=1}^{r} (a_j + 1) \ll k^{-5/6} 4^\ell \ll e^{-L},
\]
since \( a + 1 \leq 2^a \) and \( |C_a(b)| \leq 2 \) for all \( a, b \). The result follows upon noting that
\[
\prod_{j=1}^{r} C_{a_j}(0) = \prod_{j=1}^{r} \mathbb{E}(\mathbb{Y}^{a_j}(q_j)) = \mathbb{E}(\mathbb{Y}(p_1) \cdots \mathbb{Y}(p_\ell)).
\]

As an application, we will estimate the distribution of large values of \( \text{Re}(G_f(1)) \), where
\[
G_f(s) = \sum_p \frac{\lambda_f(p)}{p^s}
\]
for \( \text{Re}(s) > 1 \). In this case, Lemma 4.2 of [2], together with the zero density estimates of Kowalski and Michel (see [10, Thm. 4]), implies that for all \( f \in H_k \),
\[
G_f(1) = \sum_{p \leq (\log k)^{C_0}} \frac{\lambda_f(p)}{p} + O\left( \frac{1}{\log k} \right),
\]
except for a set of cardinality \( \ll k^{9/10} \), where \( C_0 \) is a positive constant. Therefore it follows from Lemma 5 together with Theorem 2 with \( Q = (\log k)^{C_0} \) and \( L \sim 2 \log k/(C_0 \log \log k) \) that
\[
\frac{1}{|H_k|} \left| \left\{ f \in H_k : \exp(G_f(1)) > (e^V)^2 \right\} \right|_h = \exp\left( -\left( 1 + o(1) \right) \frac{e^{V-\eta V-1}}{V} \right) \quad (3.10)
\]
Large deviations of sums of random variables

for $V$ in the range $1 \ll V \leq \log_2 k - \log_3 k - \log_4 k - \Theta$, where $C$ is defined in (2.1), $\Theta$ is a large constant, and $Y = \cos \theta$, where $\theta$ is distributed according to the Sato–Tate measure on $[0, \pi]$. Liu, Royer, and Wu [13] proved the analogue of (3.10) for $|L(1, f)|$, where $L(s, f)$ is the $L$-function attached to the eigencuspsform $f$. We can recover their result by modifying the sequence that satisfies the AIH from $\lambda_f(p)/2 = \cos \theta_f(p)$ to

$$\mathcal{X}(p) = \frac{-\log(1 - e^{i\theta_f(p)/p}) - \log(1 - e^{-i\theta_f(p)/p})}{-2 \log(1 - 1/p)}.$$ 

3.5 Kloosterman sums

Let $q$ be a large prime. The classical (normalized) Kloosterman sums modulo $q$ are defined by

$$\text{Kl}_q(a, b) := \frac{1}{\sqrt{q}} \sum_{n \in \mathbb{F}_q^*} e_q(\alpha n + b \overline{\alpha}),$$

where $e_q(z) = \exp(2\pi iz/q)$, and $\overline{\alpha}$ is the multiplicative inverse of $\alpha$ modulo $q$.

The analogue of Lemma 2.1 of [12] for Kloosterman sums (which follows in a similar way from Proposition 4.2 of [18]; see also Proposition 9.9 of [1] for a more general result) implies that for all positive integers $1 \leq \ell \leq (\log q)/2$ and all $n_1, \ldots, n_\ell \in \mathbb{F}_q$, we have

$$\frac{1}{(q-1)^2} \sum_{(a, b) \in (\mathbb{F}_q^*)^2} \text{Kl}_q(a - n_1, b) \cdots \text{Kl}_q(a - n_\ell, b) = E(\mathbb{Y}(n_1) \cdots \mathbb{Y}(n_k)) + O\left(\frac{2^\ell \ell}{\sqrt{q}}\right),$$

(3.11)

where we have $\mathbb{Y}(n) = 2 \cos \theta_n$ for all $n$, where $\theta_n$ are independent random variables with Sato–Tate distribution on $[0, \pi]$. This shows that the sequence $\mathcal{X}(n) = \text{Kl}_q(a - n, b)/2$ satisfies the AIH with $Q = q$ and $L \sim (\log q)/(2 \log 2 + 2)$.

The asymptotics (3.11) is proved using deep ingredients from algebraic geometry, including Deligne’s equidistribution theorem, which generalizes the Riemann hypothesis over finite fields, and Katz’s work on monodromy groups. This estimate was a key ingredient in investigating the distribution of partial sums of Kloosterman sums in [1] and [11].

4 Estimates for the sum $H(s)$

In this section, we obtain estimates for $H(s)$, which will be useful in our subsequent work. We start with the following lemma, which easily follows from partial summation and the growth rate of $A(x)$.

**Lemma 6.** If $s \geq 4/\beta'$, then

$$H(s) = \begin{cases} \frac{\alpha}{4} (\log s)^4 + C_A - \alpha (\log \beta)(\log s)^{A-1} + O((\log s)^{\max\{A-2, A-B\}}) & \text{if } A > 0, \\ \alpha \log \log s + C_A - \alpha (\log \beta)(\log s)^{A-1} + O((\log s)^{\max\{-2, -B\}}) & \text{if } A = 0. \end{cases}$$

**Proof.** Let $E(x) = A(x) - \alpha \int_2^x (\log(t/\beta))^{A-1} \, dt$. By partial summation we have

$$H(s) = \sum_{q_n \leq s} \frac{1}{q_n} + \int_2^s \frac{dA(t)}{t} = \alpha \int_2^s \frac{(\log(t/\beta))^{A-1} \, dt}{t} + \int_2^s \frac{dE(t)}{t} + \sum_{1 \leq q_n \leq s} \frac{1}{q_n}.$$
Since \( E(t) \ll t(\log t)^{A-1-B} \) and \( B > A \), the integral \( \int_2^s \frac{dE(t)}{t} \) is convergent, and hence we get

\[
H(s) = \begin{cases} \frac{\alpha}{A}(\log \frac{s}{\beta})^A + C_A + \mathcal{E}_1(s) & \text{if } A > 0, \\ \alpha \log \alpha + C_A + \mathcal{E}_1(s) & \text{if } A = 0, \end{cases}
\]

where

\[
\mathcal{E}_1(s) \ll \int_s^\infty \frac{dt}{t(\log t)^{B-A+1}} \ll (\log s)^{A-B}.
\]

The result follows from expanding the term \( \log \frac{s/\beta}{\alpha} \) on the right-hand side in each case. \( \square \)

Our next result gives estimates for the difference \( H(u) - H(v) \), which are needed to prove Theorems 1 and 2.

**Lemma 7.** Let \( u, v \geq 2 \), and put \( \tau = \log u - \log v \). If \( |\tau| \leq (\log v)/2 \), then

\[
H(u) - H(v) = \alpha \tau (\log v)^{A-1} + O((\tau + \tau^2)(\log v)^{A-2} + (\log v)^{A-B}).
\]

**Proof.** By definition we have

\[
H(u) - H(v) = \int_u^v \frac{dA(t)}{t} = \alpha \int_u^v \frac{(\log(t/\beta))^A}{\beta} \frac{dt}{t} + O((\log v)^{A-B}).
\]

Let \( L = \log(v/\beta) \). If \( A > 0 \), then here the main term is

\[
\frac{\alpha}{A}(L + \tau)^A - L^A = \alpha \tau L^{A-1} + O(\tau^2 L^{A-2}) = \alpha \tau (\log v)^{A-1} + O((\tau + \tau^2) L^{A-2}).
\]

If \( A = 0 \), then the main term is

\[
\alpha (\log(L + \tau) - \log L) = \alpha \frac{\tau}{L} + O \left( \frac{\tau^2}{L^2} \right) = \alpha \frac{\tau}{\log v} + O \left( \frac{\tau + \tau^2}{L^2} \right). \quad \square
\]

## 5 Estimates for the cumulant generating function of \( H_Y(Q) \)

In this section, we assume that the independent random variables \( \{Y(n)\}_{n \leq Q} \) satisfy CRVH\( (Y) \). For every complex number \( s \), we denote the cumulant generating function for the sum \( H_Y(Q) \), by

\[
K_Q(s) = K_Q(Y, s) := \log \mathbb{E}(e^{sH_Y(Q)}) = \log \mathbb{E}(e^{s\sum_{n \leq Q} Y(n)/q_n}).
\]

By the independence of the \( Y(n) \) we have

\[
K_Q(s) = \sum_{q_n \leq Q} \log \mathbb{E}(e^{sY(n)/q_n}) = sH(s) + \sum_{q_n \leq Q} f_{Y(n)} \left( \frac{s}{q_n} \right), \quad (5.1)
\]

extending the definition of the \( f_{Y(n)} \) in (1.3) to complex \( s \), the two cases corresponding to \( |s| < 1 \) and \( |s| \geq 1 \). In this section, we will show that \( \lim_{Q \to \infty} K_Q(s) \) exists for all \( s \in \mathbb{C} \), and we will estimate \( K_Q(s) \) in a large range of \( s \) in terms of \( Q \). These will be important ingredients in the proofs of Theorems 1 and 3.

We need the following estimates for the function \( f_{Y(n)} \).
Lemma 8. For each $n$, we have uniformly

$$f_{\mathbb{Y}(n)}(t) \ll \begin{cases} \frac{t^2}{\log^2(2t)}, & 0 \leq t < 1, \\ \frac{t}{\log^2(2t)}, & t \geq 1, \end{cases}$$

(which implies that the integral defining $\eta_{\mathbb{Y}} := \int_0^\infty f_{\mathbb{Y}}(u)/u^2 \, du$ converges), and

$$f'_{\mathbb{Y}(n)}(t) \ll \begin{cases} t, & 0 < t < 1, \\ \frac{1}{\log^2(2t)}, & t > 1. \end{cases}$$

Proof. We start by proving (5.2). Let $s$ be a complex number such that $|s| < 1$. Using the Taylor expansion of the exponential, we have

$$\mathbb{E}(e^{s\mathbb{Y}(n)}) = \mathbb{E}(1 + s\mathbb{Y}(n) + O(|s|^2\mathbb{Y}(n)^2)) = 1 + O(|s|^2)$$

since $\mathbb{E}(\mathbb{Y}(n)) = 0$ and $|\mathbb{Y}(n)| \leq 1$. This implies

$$\log \mathbb{E}(e^{s\mathbb{Y}(n)}) \ll |s|^2,$$

which gives the desired estimate for $f_{\mathbb{Y}(n)}(t)$ for $0 \leq t < 1$.

We now suppose that $t \geq 1$. Let $N > 1$ be a parameter to be chosen. Then by (1.2) we have

$$\mathbb{E}(e^{t\mathbb{Y}(n)}) \geq \mathbb{P}(\mathbb{Y}(n) > 1 - \frac{1}{N})e^{t(1-1/N)} \gg \exp\left(t \left(1 - \frac{1}{N}\right) - ce^{\sqrt{N}}\right).$$

Choosing $N = ((1/2) \log(2t/c^2))^2$ and using that $\mathbb{Y}(n) \leq 1$, we obtain

$$\exp\left(t - \frac{5t}{\log^2(2t/c^2)}\right) \ll \mathbb{E}(e^{t\mathbb{Y}(n)}) \ll e^t,$$

from which the desired estimate for $f_{\mathbb{Y}(n)}(t)$ follows in this case.

Next, we establish (5.3). Note that $f_{\mathbb{Y}(n)}$ is differentiable on $(0, \infty) \setminus \{1\}$ and

$$f'_{\mathbb{Y}(n)}(t) = \begin{cases} \frac{\mathbb{E}(\mathbb{Y}(n)e^{t\mathbb{Y}(n)})}{\mathbb{E}(e^{t\mathbb{Y}(n)})}, & 0 < t < 1, \\ \frac{\mathbb{E}(\mathbb{Y}(n)e^{t\mathbb{Y}(n)})}{\mathbb{E}(e^{t\mathbb{Y}(n)})} - 1, & t > 1. \end{cases}$$

As before, in the case $0 < t < 1$ the estimate of $f'_{\mathbb{Y}(n)}(t)$ follows from the Taylor expansions $\mathbb{E}(e^{t\mathbb{Y}(n)}) = 1 + O(t^2)$ and

$$\mathbb{E}(\mathbb{Y}(n)e^{t\mathbb{Y}(n)}) = \mathbb{E}(\mathbb{Y}(n) + t\mathbb{Y}^2(n) + O(t^2|\mathbb{Y}^3(n)|)) = t\mathbb{E}(\mathbb{Y}^2(n)) + O(t^2).$$

We now suppose that $t > 1$, and let $N > 1$ be a parameter to be chosen. Let $\mathcal{B}$ be the event $\mathbb{Y}(n) > 1 - 1/N$, and let $\mathcal{B}^c$ be its complement. Then we have

$$\mathbb{E}(\mathbb{Y}(n)e^{t\mathbb{Y}(n)}) = \mathbb{E}(1_{\mathcal{B}} \cdot \mathbb{Y}(n)e^{t\mathbb{Y}(n)}) + \mathbb{E}(1_{\mathcal{B}^c} \cdot \mathbb{Y}(n)e^{t\mathbb{Y}(n)})$$

$$\geq \left(1 - \frac{1}{N}\right)\mathbb{E}(1_{\mathcal{B}} \cdot e^{t\mathbb{Y}(n)}) + O(e^{t(1-1/N)}),$$

Lith. Math. J., 61(3):345–372, 2021.
since $\mathbb{E}(1_{B^c} \cdot e^{\mathcal{Y}(n)}) \leq e^{t(1-1/N)}$. Moreover, this implies that
\begin{equation}
\mathbb{E}(\mathcal{Y}(n)e^{\mathcal{Y}(n)}) \geq \left(1 - \frac{1}{N}\right)\mathbb{E}(e^{\mathcal{Y}(n)}) + O(e^{t(1-1/N)}).
\end{equation}

We choose $N = (\log 2t/c^2)^2/6$. Then from (5.5) it follows that
\[
e^{t(1-1/N)} \ll \frac{\mathbb{E}(e^{\mathcal{Y}(n)})}{t}.
\]
Inserting this estimate into (5.6) and using the bound $\mathcal{Y}(n) \leq 1$ give
\[
1 - \frac{1}{N} + O\left(\frac{1}{t}\right) \leq \frac{\mathbb{E}(\mathcal{Y}(n)e^{\mathcal{Y}(n)})}{\mathbb{E}(e^{\mathcal{Y}(n)})} \leq 1.
\]
In this case, this implies the desired estimate for $f_Y'(n)$.

**Corollary 1.** For each complex number $s$, the limit $\lim_{Q \to \infty} K_Q(s)$ exists and equals $K(s)$, say. Moreover, $K_Q(s) = K(s) + O(|s|^2(\log Q)^{A-1}/Q)$ if $Q > |s|$.

**Proof.** If $|s| < Q < q_n \leq Z$, then $|s|/q_n < 1$, and so
\[
\log \mathbb{E}(e^{s\mathcal{Y}(n)/q_n}) \ll \left(\frac{|s|}{q_n}\right)^2
\]
by (5.4). Therefore, by (5.1),
\[
K_Z(s) - K_Q(s) = \sum_{Q < q_n \leq Z} \log \mathbb{E}(e^{s\mathcal{Y}(n)/q_n}) \ll \sum_{Q < q_n \leq Z} \left(\frac{|s|}{q_n}\right)^2 \ll |s|^2 (\log Q)^{A-1}.
\]
by the growth rate of $A(x)$. This is $o_{Q \to \infty}(1)$, and so $K_Q(s)$ form a Cauchy sequence and therefore have a limit. Then the same calculation implies the claimed estimate. □

Next, we wish to estimate the cumulant generating function $K_Q(s)$ in a large range of $s$ in terms of $Q$.

**Proposition 1.** For any real number $s$ in the range $3 \leq s \leq Q/\log Q$, we have
\[
K_Q(s) = sH(s) + \left(\alpha \eta_Y + O\left(\frac{\log \log s}{\log s}\right)\right)s(\log s)^{A-1}.
\]

**Proof.** By (5.1) it suffices to show that for $3 \leq s \leq Q/\log Q$, we have
\[
\sum_{q_n \leq Q} f_{\mathcal{Y}(n)} \left(\frac{s}{q_n}\right) = \left(\alpha \eta_Y + O\left(\frac{\log \log s}{\log s}\right)\right)s(\log s)^{A-1}.
\]
To estimate this sum, we will split it into three parts: $q_n \leq s^{1/2}$, $s^{1/2} < q_n < s \log s$, and $s \log s \leq q_n \leq Q$. By (5.2) and Lemma 6 the contribution of the first part is
\[
\sum_{q_n \leq s^{1/2}} f_{\mathcal{Y}(n)} \left(\frac{s}{q_n}\right) \ll \frac{s}{\log^2 s} \sum_{q_n \leq s^{1/2}} \frac{1}{q_n} \ll s(\log s)^{A-2} \log \log s.
\]
Moreover, using (5.2), we deduce that the contribution of the last part is
\[
\sum_{s \log s \leq q_n < Q} f_Y(n) \left( \frac{s}{q_n} \right) \ll s^2 \sum_{s \log s \leq q_n < Q} \frac{1}{q_n^2} = s^2 \int_{s \log s}^{Q} \frac{dA(t)}{t^2} \ll s \left( \log s \right)^{A-2}
\]
by the growth rate of \( A(x) \). Hence we deduce that
\[
\sum_{q_n \leq Q} f_Y(n) \left( \frac{s}{q_n} \right) = \sum_{s^{1/2} < q_n < s \log s} f_Y(n) \left( \frac{s}{q_n} \right) + O \left( s \left( \log s \right)^{A-2} \log \log s \right).
\]
By the third hypothesis of CRVH(\( \mathbb{Y} \)) we see that for all \( t \),
\[
E \left( e^{tY(n)} \right) = E \left( e^{tY(n) + O(t/\log n)^2} \right) = E(e^{tY}) e^{O(t/(\log n)^2)},
\]
so that \( f_Y(n)(t) = f_Y(t) + O(t/(\log n)^2) \). Therefore
\[
\sum_{s^{1/2} < q_n < s \log s} f_Y(n) \left( \frac{s}{q_n} \right) = \sum_{s^{1/2} < q_n < s \log s} f_Y \left( \frac{s}{q_n} \right) + O \left( \sum_{s^{1/2} < q_n < s \log s} \frac{s}{q_n (\log n)^2} \right).
\]
For this last sum, we note that if \( s^{1/2} < q_n < s \log s \), then \((\log n)^2 \asymp (\log s)^2\), so
\[
\sum_{s^{1/2} < q_n < s \log s} \frac{s}{q_n (\log n)^2} \ll \frac{s}{(\log s)^2} H(s \log s) \ll s \left( \log s \right)^{A-2} \log s
\]
by Lemma 6, and so
\[
\sum_{q_n \leq Q} f_Y(n) \left( \frac{s}{q_n} \right) = \sum_{s^{1/2} < q_n < s \log s} f_Y \left( \frac{s}{q_n} \right) + O \left( s \left( \log s \right)^{A-2} \log \log s \right). \tag{5.7}
\]
Now we have that, for \( E(x) = A(x) - \alpha \int_2^x (\log (t/\beta)) A-1 \, dt \ll x (\log x)^{A-1-B} \),
\[
\sum_{s^{1/2} < q_n < s \log s} f_Y \left( \frac{s}{q_n} \right) = \int_{s^{1/2}}^{s \log s} f_Y \left( \frac{s}{t} \right) dA(t) = \alpha \int_{s^{1/2}}^{s \log s} f_Y \left( \frac{s}{t} \right) \left( \log \frac{t}{\beta} \right)^{A-1} dt + \mathcal{E}_1, \tag{5.8}
\]
where
\[
\mathcal{E}_1 = \int_{s^{1/2}}^{s \log s} f_Y \left( \frac{s}{t} \right) dE(t) = \left[ E(t) f_Y \left( \frac{s}{t} \right) \right]_{s^{1/2}}^{s \log s} + \int_{s^{1/2}}^{s \log s} \frac{f_Y \left( \frac{s}{t} \right) E(t) dt}{t^2}
\]
\[
\ll f_Y \left( \left( \log s \right)^{-1} s \left( \log s \right)^{A-B} + f_Y \left( s^{1/2} \right) s^{1/2} \left( \log s \right)^{A-1-B} + \int_{s^{1/2}}^{s \log s} \frac{f_Y \left( \frac{s}{t} \right)}{t^2} t \left( \log t \right)^{A-1-B} dt
\]
\[
\ll s \left( \log s \right)^{A-B-2} + s \left( \log s \right)^{A-1-B} \int_{s^{1/2}}^{s \log s} \left| \frac{f_Y \left( s/t \right)}{t} \right| dt
\]

Lith. Math. J., 61(3):345–372, 2021.
by (5.2). To bound this last integral, we use (5.3). This gives
\[
\int_{s^{1/2}}^{s \log s} \frac{|f_Y(s/t)|}{t} \, dt \ll \int_{s^{1/2}}^{s} \frac{1}{t \log(2s/t)^2} \, dt + \int_{s}^{s \log s} \frac{s \log s}{t^2} \, dt \ll \int_{2}^{2s^{1/2}} \frac{1}{v \log v^2} \, dv + 1 \ll 1
\]
by the change of variable \( v = 2s/t \). Inserting these estimates into (5.8) gives
\[
\sum_{s^{1/2} \leq q_n \leq s \log s} f_Y\left( \frac{s}{q_n} \right) = \alpha \int_{s^{1/2}}^{s \log s} f_Y\left( \frac{s}{t} \right) \left( \log \frac{t}{\beta} \right)^{A-1} \, dt + O(s \log s)^{A-1-B}. \tag{5.9}
\]
Using the change of variables \( u = s/t \), we obtain
\[
\int_{s^{1/2}}^{s \log s} f_Y\left( \frac{s}{t} \right) \left( \log \frac{t}{\beta} \right)^{A-1} \, dt = s \int_{s^{1/2}}^{s \log s} f_Y\left( \frac{s}{u} \right) \left( \log \frac{s}{u \beta} \right)^{A-1} \, du
\]
\[
= s \left( \log \frac{s}{\beta} \right)^{A-1} \int_{s^{1/2}}^{s \log s} f_Y\left( \frac{s}{u} \right) \left( 1 + O\left( \frac{\log u}{\log s} \right) \right) \, du.
\]
Moreover, using (5.2), we get
\[
\int_{0}^{\left( \log s \right)^{-1}} \frac{f_Y(u)}{u^2} \, du + \int_{s^{1/2}}^{\infty} \frac{f_Y(u)}{u^2} \, du \ll \int_{0}^{\left( \log s \right)^{-1}} 1 \, du + \int_{s^{1/2}}^{\infty} \frac{1}{u \log u} \, du \ll \frac{1}{\log s}
\]
and
\[
\int_{s^{1/2}}^{\infty} \frac{f_Y(u) |\log u|}{u^2} \, du \ll \int_{1}^{\left( \log s \right)^{-1}} |\log u| \, du + \int_{1}^{\infty} \frac{1}{u \log u} \, du \ll \log \log s.
\]
This gives
\[
\int_{s^{1/2}}^{s \log s} f_Y\left( \frac{s}{t} \right) \left( \log \frac{t}{\beta} \right)^{A-1} \, dt = \eta_Y s (\log s)^{A-1} \left( 1 + O\left( \frac{\log \log s}{\log s} \right) \right).
\]
Combining this estimate with (5.7) and (5.9) completes the proof. \( \square \)

6 The main theorem for the sum of independent random variables: Proof of Theorem 1

The key technical result in this paper is the following result.

**Theorem 4.** Let \( Q \) be large, and let \( \{ Y(n) \}_{n \leq Q} \) be a sequence of independent random variables satisfying \( \text{CRVH}(Y) \). Given \( V \) in the range
\[
c \leq V \leq R := \begin{cases} \frac{\alpha}{A} (\log Q)^A + C_A - (\alpha \log Q + \Theta)(\log Q)^{A-1} & \text{if } A > 0, \\ \alpha \log Q + C_A - (\alpha \log Q + \Theta)(\log Q)^{-1} & \text{if } A = 0, \end{cases}
\]

where \( \Theta > 0 \) and \( c > 1 \) are suitably large constants, choose \( Z = Z(V) \) as the smallest number \( z \) for which \( H(z) \geq V \). Then we have
\[
\Phi_Q(Y, V) = \exp \left( -\alpha e^{-\eta Y} - 1 Z(\log Z)^{A-1} \left( 1 + O\left( \sqrt{\mathcal{E}(z)} \right) \right) \right),
\]
where \( \mathcal{E}(Z) \) is defined by
\[
\mathcal{E}(Z) := \frac{\log \log Z}{\log Z} + \frac{1}{(\log Z)^{B-1}}.
\] (6.2)

**Proof.** We will use the saddle-point method to prove Theorem 4. For \( V \geq 1 \), we define
\[
s(V) := e^{-\eta Y} Z(V).
\]
The definition of \( Z \) implies that \( H(Z) + O(1/Z) = V \leq R \). Therefore by inserting the estimate for \( H(Z) \) in Lemma 6 into (6.1) we deduce that
\[
s(V) \ll \frac{Q}{\log Q},
\]
and so for any \( s \ll s(V) \), we have
\[
K_Q(s) = sH(s) + (\alpha \eta Y + O(\mathcal{E}(s(V)))) s L^{A-1}
\] (6.3)
by Proposition 1, where for the rest of this proof, we write \( \mathcal{L} := \log s(V) \).

For \( u > 0 \), we have
\[
\exp (K_Q(u)) = - \int_{-\infty}^{\infty} e^{ut} d\Phi_Q(t) = u \int_{-\infty}^{\infty} e^{ut} \Phi_Q(t) dt.
\] (6.4)

We will now show (in (6.9)) that for \( u = s \), by far the largest part of the integral on the right-hand side occurs for \( t \) in a short interval around \( s \):

Let \( 0 < \delta < 1 \) be a small parameter to be chosen, and put \( S_1 = e^\delta s \) with \( s = s(V) \). Let \( \Delta = \delta \alpha \mathcal{L}^{A-1} \). Then from (6.3) and (6.4) it follows that
\[
\int_{V+\Delta}^{\infty} e^{st} \Phi_Q(t) dt \leq \exp ((s - S_1)(V + \Delta)) \int_{V+\Delta}^{\infty} e^{St} \Phi_Q(t) dt
\]
\[
\leq \exp ((s - S_1)(V + \Delta) + S_1 H(S_1) + (\alpha \eta Y + O(\mathcal{E}(s))) S_1 L^{A-1}).
\] (6.5)

Now by Lemma 7 we have
\[
S_1 H(S_1) + \alpha \eta Y S_1 L^{A-1} = e^\delta (sH(s) + \alpha \eta Y s L^{A-1}) + \alpha e^\delta s L^{A-1} + O(s L^{A-1} \mathcal{E}(s)).
\]

Furthermore, by Lemma 7 we have
\[
V = H(Z) + O\left( \frac{1}{Z} \right) = H(s) + (\alpha (\eta Y + 1) + O(\mathcal{E}(s))) L^{A-1}. \] (6.6)

---

6 The constant \( c \) above is chosen to be large enough that \( Z \) is large enough so that \( \mathcal{E}(Z) \) is well defined.
Inserting these estimates into (6.5), we get
\[
\int_{V+\Delta}^{\infty} e^{st} \Phi_Q(t) \, dt \leq \exp \left( sH(s) + \alpha \eta_Y s \mathcal{L}^{A-1} + (1 + \delta - e^{\delta}) \alpha s \mathcal{L}^{A-1} + O(s \mathcal{L}^{A-1} \mathcal{E}(s)) \right).
\]

Therefore, choosing \( \delta = B_0 \sqrt{\mathcal{E}(s)} \) for a suitably large constant \( B_0 \) and using (6.3), we deduce that
\[
\int_{V+\Delta}^{\infty} e^{st} \Phi_Q(t) \, dt \leq e^{-s \mathcal{L}^{A-1} \mathcal{E}(s)} \int_{-\infty}^{\infty} e^{st} \Phi_Q(t) \, dt. \tag{6.7}
\]

A similar calculation shows that by putting \( S_2 = e^{-\delta} s \) we obtain
\[
\int_{-\infty}^{V-\Delta} e^{st} \Phi_Q(t) \, dt \leq \exp \left( (s - S_2)(V - \Delta) \right) \int_{-\infty}^{V-\Delta} e^{S_2 t} \Phi_Q(t) \, dt
\]
\[
\leq \exp \left( (s - S_2)(V - \Delta) + S_2 H(S_2) + (\alpha \eta_Y + O(\mathcal{E}(s))) S_2 \mathcal{L}^{A-1} \right)
\]
\[
= \exp \left( sH(s) + \alpha \eta_Y s \mathcal{L}^{A-1} + (1 - \delta - e^{-\delta}) \alpha s \mathcal{L}^{A-1} + O(s \mathcal{L}^{A-1} \mathcal{E}(s)) \right),
\]
and hence by our choice of \( \delta \) and (6.3) we deduce
\[
\int_{-\infty}^{V-\Delta} e^{st} \Phi_Q(t) \, dt \leq e^{-s \mathcal{L}^{A-1} \mathcal{E}(s)} \int_{-\infty}^{\infty} e^{st} \Phi_Q(t) \, dt. \tag{6.8}
\]

Combining the bounds (6.7) and (6.8) with (6.3) and (6.4) gives
\[
\int_{V-\Delta}^{V+\Delta} e^{st} \Phi_Q(t) \, dt = \exp \left( sH(s) + \alpha \eta_Y s \mathcal{L}^{A-1} \left( 1 + O(\mathcal{E}(s)) \right) \right), \tag{6.9}
\]
as claimed.

Now \( \Phi_Q(t) \) is nonincreasing as a function of \( t \) by definition, and so we may bound the above integral as follows:
\[
e^{sV + O(s \Delta)} \Phi_Q(V + \Delta) \leq \int_{V-\Delta}^{V+\Delta} e^{st} \Phi_Q(t) \, dt \leq e^{sV + O(s \Delta)} \Phi_Q(V - \Delta).
\]

Inserting these bounds into (6.9) and using (6.6), we obtain
\[
\Phi_Q(V + \Delta) \leq \exp \left( -\alpha s \mathcal{L}^{A-1} \left( 1 + O(\sqrt{\mathcal{E}(s)}) \right) \right) \leq \Phi_Q(V - \Delta). \tag{6.10}
\]

Let \( s^+ \) and \( s^- \) be defined by \( s^\pm = e^{-\eta_Y} Z(V \pm \Delta) \). Then it follows by two applications of Lemma 7 that
\[
V \pm \Delta = H(Z(V \pm \Delta)) + O(1)
\]
\[
= H(s^\pm) + \alpha(\eta_Y + 1)(\log s^\pm)^{A-1} + O\left( (\log s^\pm)^{\max\{A-2, A-B\}} \right)
\]
\[
= H(s) + \alpha t^\pm \mathcal{L}^{A-1} + \alpha(\eta_Y + 1)(\log s + t^\pm)^{A-1} + O\left( (\log s^\pm)^{\max\{A-2, A-B\}} \right),
\]
where \( t^\pm = \log s^\pm - \log s \). By (6.6) we deduce that
\[
\pm \Delta = \alpha t^\pm \mathcal{L}^{A-1} + \alpha (\eta t^\pm + 1) \left( (\log s + t^\pm)^{A-1} - \mathcal{L}^{A-1} \right) + O(\mathcal{E}(s^+) (\log s^+)^{A-1}).
\]
Since \( \Delta = o(\mathcal{L}^{A-1}) \), this implies that \( t^\pm = o(1) \), and so
\[
\alpha t^\pm \mathcal{L}^{A-1} = \pm \Delta + O(\mathcal{E}(s) \mathcal{L}^{A-1}).
\]
Since \( \Delta = \delta \alpha \mathcal{L}^{A-1} \), this implies that \( t^\pm = \delta + O(\mathcal{E}(s)) \ll \sqrt{\mathcal{E}(s)} \), and thus
\[
s^\pm (\log s^+)^{A-1} = s \mathcal{L}^{A-1} (1 + O(\sqrt{\mathcal{E}(s)})). \tag{6.11}
\]
Finally, using the upper bound of (6.10) with \( V + \Delta \) instead of \( V \) together with this last estimate gives
\[
\Phi_Q(V) \geq \exp \left( -\alpha s^+ (\log s^+)^{A-1} \left( 1 + O(\sqrt{\mathcal{E}(s)} \right) \right) = \exp \left( -\alpha s \mathcal{L}^{A-1} (1 + O(\sqrt{\mathcal{E}(s)})) \right).
\]
A similar lower bound follows from the lower of (6.10) with \( V - \Delta \) instead of \( V \), together with (6.11).

**Deduction of Theorem 1.** To deduce Theorem 1 from Theorem 4, we express \( Z = Z(V) \) as an explicit function in terms of \( V \) using the definition of \( Z \) and Lemma 6: If \( A > 0 \), then we have
\[
V - C_A = \frac{\alpha}{A} (\log Z)^A \left( 1 - \frac{A \log \beta}{\log Z} + O((\log Z)^{\max\{-2,-B\}}) \right).
\]
Rearranging this gives
\[
W = \log Z \left( \frac{1}{1 - \frac{A \log \beta}{\log Z} + O((\log Z)^{\max\{-2,-B\}})} \right)^{1/A} = \log Z - \log \beta + O((\log Z)^{\max\{-1,1-B\}}),
\]
and hence by exponentiating we obtain
\[
Z = \beta e^W (1 + O(W^{\max\{-1,1-B\}})). \tag{6.12}
\]
If \( A = 0 \), then, analogously,
\[
V - C_A = \alpha \log \log Z - \frac{\alpha \log \beta}{\log Z} + O\left( \frac{1}{(\log Z)^{\min\{2,B\}}} \right),
\]
so that
\[
W = (\log Z) \exp \left( -\frac{\log \beta}{\log Z} + O\left( \frac{1}{(\log Z)^{\min\{2,B\}}} \right) \right) = \log Z - \log \beta + O\left( \frac{1}{(\log Z)^{\min\{1,B-1\}}} \right),
\]
and hence (6.12) holds by exponentiating.

The estimate claimed for \( \Phi_Q(V) \) now follows by substituting (6.12) into the estimate for \( \Phi_Q(V) \) in Theorem 4.

We also deduce the range for \( W \) from the range for \( V \) in Theorem 4 by an argument similar to that above with any \( \Theta' > \Theta/\alpha \).
7 Using the AIH: Proofs of Theorems 2 and 3

We start by using the AIH to relate the moment generating function of \( H_X(s) \) to that of \( H_Y(s) = \sum_{q_n \leq Q} \mathbb{Y}(n)/q_n \), where \( \mathbb{Y}(n) \) are independent random variables.

**Proposition 2.** Assume \( \text{AIH}(L, Q) \). For all complex numbers \( s \) such that \( |s| \leq L/(10H(Q)) \), we have

\[
\mathbb{E}(e^{sH_X(Q)}) = \mathbb{E}(e^{sH_Y(Q)}) + O(e^{-L/2}).
\]

**Proof.** First, note that

\[
\mathbb{E}(e^{sH_X(Q)} - e^{sH_Y(Q)}) = \sum_{k \leq L} \frac{s^k}{k!} \mathbb{E}(H_X(Q)^k - H_Y(Q)^k) + \text{Err},
\]

where since \( |H_X(Q)|, |H_Y(Q)| \leq H(Q) \), we have

\[
\text{Err} \leq 2 \sum_{k > L} \frac{|s|^k H(Q)^k}{k!} \ll \sum_{k > L} \left( \frac{3|s|H(Q)}{L} \right)^k \ll e^{-L}
\]

by Stirling’s formula and our assumption on \( s \). We now expand the moments on the right-hand side of (7.1) and use the AIH to get

\[
\mathbb{E}(H_X(Q)^k - H_Y(Q)^k) = \sum_{q_n, \ldots, q_{nk} \leq Q} \mathbb{E}\left(\prod_{n=1}^{k} \mathbb{X}(n) - \prod_{n=1}^{k} \mathbb{Y}(n)\right) \ll H(Q)^k e^{-L}.
\]

Inserting this estimate into (7.1) allows us to bound the sum on the right-hand side by

\[
\ll e^{-L} \cdot \sum_{k \leq L} \frac{(|s|H(Q))^k}{k!} \ll e^{s|H(Q)|} e^{-L} \leq e^{-L/2}
\]

by our assumption on \( s \), and the result follows. \( \square \)

Combining this result with Corollary 1, we prove that the characteristic function of \( H_X(Q) \) converges as \( Q \to \infty \) to the characteristic function of some random variable.

**Proposition 3.** Assume \( \text{AIH}(L, Q) \) and \( \text{CRVH}(\mathbb{Y}) \). There exists a function \( \varphi(t) \) such that \( \mathbb{E}(e^{itH_X(Q)}) \) converges pointwise to \( \varphi(t) \) on \( \mathbb{R} \). Moreover, for all real numbers \( |t| \leq \min(L/(10H(Q)), Q^{1/2} \log Q)^{-A-1/2} \), we have

\[
\mathbb{E}(e^{itH_X(Q)}) = \varphi(t) + O\left( e^{-L/2} + \frac{t^2(\log Q)^{A-1}}{Q} \right)
\]

(7.2)

(so that \( \mathbb{E}(e^{itH_X(Q)}) \) converges uniformly to \( \varphi(t) \) on any finite interval). Furthermore, the function \( \varphi \) satisfies

\[
\varphi(t) \ll \exp(-C_0|t| (\log |t|)^A - 2)
\]

(7.3)

for all real numbers \( |t| \geq 2 \), where \( C_0 \) is an absolute constant.
Proof. Let \( \varphi(t) = e^{K(it)} \), where \( K(s) = \lim_{Q \to \infty} K_Q(Y, s) \) from Corollary 1. Then it follows from this result that for \( |t| \leq Q^{1/2} (\log Q)^{-(A-1)/2} \), we have

\[
\varphi(t) = \exp\left(K_Q(Y, it) + O\left(\frac{t^2 (\log Q)^{A-1}}{Q}\right)\right) = \mathbb{E}(e^{iH_t(Q)}) \left(1 + O\left(\frac{t^2 (\log Q)^{A-1}}{Q}\right)\right)
\]

Combining this estimate with Proposition 2 gives (7.2).

Hence it remains to prove the bound (7.3). We will prove that this bound is satisfied by \( \mathbb{E}(e^{iH_t(Q)}) \) and then use (7.4) to deduce the same bound for \( \varphi(t) \). Let \( |t| \geq 2 \) be a real number, and assume that \( Q \) is large so that \( Q > t^2 \). Since the \( Y(n) \) are independent, we have

\[
|\mathbb{E}(e^{iH_t(Q)})| = \left|\prod_{q_n \leq Q} \mathbb{E}(e^{iY(n)/q_n})\right| \leq \prod_{n: |t| \log |t| < q_n < t^2} |\mathbb{E}(e^{iY(n)/q_n})|.
\]

Now, since \( q_n > |t| \), we can use the Taylor expansion of the exponential to get

\[
\mathbb{E}(e^{iY(n)/q_n}) = \mathbb{E}\left(1 + \frac{it}{q_n} Y(n) - \frac{t^2}{2q_n^2} Y^2(n) + O\left(\frac{t^3}{q_n^3}\right)\right) = 1 - \frac{t^2}{2q_n^2} \mathbb{E}(Y^2(n)) + O\left(\frac{t^3}{q_n^3}\right)
\]

as each \( |Y(n)| \leq 1 \). Furthermore, from (1.2) it follows that there exists an absolute constant \( c_1 > 0 \) such that \( \mathbb{E}(Y^2(n))^2 \geq c_1 \) for all \( n \geq 1 \). Therefore we deduce that

\[
|\mathbb{E}(e^{iH_t(Q)})| \leq \exp\left(-\frac{c_1 t^2}{2} \sum_{|t| \log |t| < q_n < t^2} \frac{1}{q_n^2} + O\left(\frac{t^3}{q_n^3}\sum_{|t| \log |t| < q_n < t^2} \frac{1}{q_n^3}\right)\right).
\]

The asymptotics for \( A(x) \) now implies that

\[
\sum_{|t| \log |t| < q_n < t^2} \frac{1}{q_n^2} \ll \frac{(\log |t|)^{A-2}}{|t|} \quad \text{and} \quad \sum_{|t| \log |t| < q_n < t^2} \frac{1}{q_n^3} \ll \frac{(\log |t|)^{A-3}}{t^2},
\]

and so

\[
\mathbb{E}(e^{iH_t(Q)}) \ll \exp\left(-C_0 |t| (\log |t|)^{A-2}\right)
\]

for some positive constant \( C_0 \). By letting \( Q \to \infty \) in (7.4) we deduce that the same estimate holds for \( \varphi(t) \), which completes the proof. \( \Box \)

Proof of Theorem 3. First, note that \( \mathbb{E}(e^{iH_t(Q)}) \) converges uniformly to \( \varphi(t) \) on any finite interval by (7.2). This shows that \( \varphi(t) \) is continuous at \( t = 0 \), since \( \mathbb{E}(e^{iH_t(Q)}) \) is continuous at \( t = 0 \). Therefore it follows from Lévy’s continuity theorem that \( \varphi(t) \) is the characteristic function of some random variable \( \mathbb{W} \) and that \( H_{\mathbb{W}}(Q) \) converges in distribution to \( \mathbb{W} \). Furthermore, since \( \varphi(t) = \mathbb{E}(e^{it\mathbb{W}}) \) is rapidly decreasing by (7.3), \( \mathbb{W} \) is absolutely continuous and has a bounded probability density function on \( \mathbb{R} \). Therefore it follows from

Lith. Math. J., 61(3):345–372, 2021.
the Berry–Esseen inequality (see [20, Sec. II.7.6]) that for all $T > 0$, we have

$$
\sup_{V \in \mathbb{R}} |\Phi_Q(X, V) - \mathbb{P}(\mathbb{W} > V)| \ll \frac{1}{T} + \int_{-T}^{T} \frac{|\mathbb{E}(e^{itH_Q}) - \varphi(t)|}{t} \, dt.
$$

Finally, choosing $T = \min(L/(10H(Q)), Q^{1/3}(\log Q)^{-A(A-1)/3})$ and using (7.2) give

$$
\sup_{V \in \mathbb{R}} |\Phi_Q(X, V) - \mathbb{P}(\mathbb{W} > V)|
\ll \frac{1}{T} + \frac{T^2(\log Q)^{A-1}}{Q} \ll \frac{1}{T} \ll \frac{H(Q)}{L} + \frac{(\log Q)^{(A-1)/3}}{Q^{1/3}},
$$

as desired. □

We now prove the analogue of Theorem 4 in the case of a sum of approximately independent random variables. Since the proof is similar, we will only indicate the main occurring differences.

**Theorem 5.** Let $Q$ be large. Suppose that the random variables $\{X(n)\}_{q_n \leq Q}$ satisfy $AIH(L, Q)$, approximated by the independent random variables $\{Y(n)\}_{q_n \leq Q}$, with $(\log Q)^{2A} \leq L \leq A(Q)$. Moreover, suppose that the $\{Y(n)\}_{q_n \leq Q}$ satisfy $CRVH(Y)$. Suppose that $V$ is in the range

$$
c \leq V \leq R(L, Q) := \left\{ \begin{array}{ll}
\frac{A}{4}(\log L)^A + C_A - (\alpha A \log \log Q + \Theta)(\log L)^{A-1} & \text{if } A > 0, \\
\alpha \log \log L + C_A - (\alpha \log \log Q + \Theta)(\log L)^{-1} & \text{if } A = 0,
\end{array} \right.
$$

(7.5)

where $\Theta > 0$ and $c > 1$ are suitably large constants. Then

$$
\Phi_Q(V) = \exp\left(-\alpha e^{-\eta - 1} Z(\log Z)^{A-1}(1 + O(\sqrt{\mathcal{E}(Z)}))\right),
$$

where $\mathcal{E}(Z)$ is defined in (6.2).

**Proof.** We modify the proof of Theorem 4. The key issue is that we wish to employ Proposition 2, which requires that $s \ll L/H(Q)$.

Now $L \leq A(Q) - \alpha Q(\log Q)^{A-1}$, and

$$
H(Q) = \sum_{q_n \leq Q} \frac{1}{q_n} \sim \left\{ \begin{array}{ll}
\frac{A}{4}(\log Q)^A & \text{if } A > 0, \\
\alpha \log \log Q & \text{if } A = 0,
\end{array} \right.
$$

by Lemma 6. Therefore $L/H(Q) \leq (A + o(1))Q/\log Q$, and so

$$
\min\left\{ \frac{Q}{\log Q}, \frac{L}{10H(Q)} \right\} \gg \frac{L}{H(Q)} \gg \left\{ \begin{array}{ll}
\frac{L}{\log Q} & \text{if } A > 0, \\
\frac{L}{\log \log Q} & \text{if } A = 0.
\end{array} \right.
$$

(7.6)

By hypothesis we have $V \leq R(L, Q)$, so that $H(Z) \leq R(L, Q) + O(1/Z)$. We can again use Lemma 6 to evaluate $H(Z)$, and using the definition of $R(L, Q)$ in (7.5), we deduce that if $A > 0$, then $Z \ll L/(\log Q)^A$, and if $A = 0$, then $Z \ll L/\log \log Q$. We deduce from (7.6) that, in either case,

$$
s(V) \ll Z \ll \min\left\{ \frac{Q}{\log Q}, \frac{L}{10H(Q)} \right\}.
$$
and therefore we can apply Proposition 1 for any $s \gg s(V)$, and we complete the proof exactly as in the proof of Theorem 4. □

Proof of Theorem 2. We deduce Theorem 2 from Theorem 5 exactly as we deduce Theorem 1 from Theorem 4. The only difference is determining the range for $W$, which follows from the analogous calculation. □

Acknowledgment. Thanks to Nadia Sidorova for useful conversations and one of the referees for a very careful reading of the paper, which highlighted some necessary minor corrections.

References

1. P. Autissier, D. Bonolis, and Y. Lamzouri, The distribution of the maximum of partial sums of Kloosterman sums and other trace functions, *Compos. Math.*, 2021 (to appear).

2. J. Cogdell and P. Michel, On the complex moments of symmetric power $L$-functions at $s = 1$, *Int. Math. Res. Not.*, 31:1561–1617, 2004.

3. H.M. Davenport, *Multiplicative Number Theory*, 3rd ed., Grad. Texts Math., Vol. 74, Springer-Verlag, New York, 2000. Revised and with a preface by H.L. Montgomery.

4. S.W. Graham and C.J. Ringrose, Lower bounds for least quadratic non-residues, in B.C. Berndt, H.G. Diamond, H. Halberstam, and A. Hildebrand (Eds.), *Analytic Number Theory*, Prog. Math., Vol. 85, Birkhäuser, Boston, MA, 1990, pp. 269–309.

5. A. Granville and K. Soundararajan, Large character sums, *J. Am. Math. Soc.*, 14(2):365–397, 2001.

6. A. Granville and K. Soundararajan, The distribution of values of $L(1, \chi_d)$, *Geom. Funct. Anal.*, 13(5):992–1028, 2003.

7. A. Granville and K. Soundararajan, Extreme values of $|\zeta(1 + it)|$, in *The Riemann Zeta Function and Related Themes: Papers in Honour of Professor K. Ramachandra*, Ramanujan Math. Soc. Lect. Notes Ser., Vol. 2, Ramanujan Mathematical Society, Mysore, India, 2006, pp. 65–80.

8. D.R. Heath-Brown, A mean value estimate for real character sums, *Acta Arith.*, 72(3):235–275, 1995.

9. H. Iwaniec, *Topics in Classical Automorphic Forms*, Grad. Stud. Math., Vol. 17, AMS, Providence, RI, 1997.

10. E. Kowalski and P. Michel, The analytic rank of $J_0(q)$ and zeros of automorphic $L$-functions, *Duke Math. J.*, 100(3):503–542, 1999.

11. E. Kowalski and W. Sawin, Kloosterman paths and the shape of exponential sums, *Compos. Math.*, 152(7):1489–1516, 2016.

12. Y. Lamzouri, On the distribution of the maximum of cubic exponential sums, *J. Inst. Math. Jussieu*, 19(4):1259–1286, 2020.

13. J.Y. Liu, E. Royer, and J. Wu, On a conjecture of Montgomery–Vaughan on extreme values of automorphic $L$-functions at $1$, in *Anatomy of Integers*, CRM Proc. Lect. Notes, Vol. 46, AMS, Providence, RI, 2008, pp. 217–245.

14. W.R. Monach, *Numerical Investigations of Several Problems in Number Theory*, PhD thesis, University of Michigan, Ann Arbor, MI, 1980.

15. H.L. Montgomery, *Topics in Multiplicative Number Theory*, Lect. Notes Math., Vol. 227, Springer-Verlag, Berlin, New York, 1971.

16. H.L. Montgomery and A.M. Odlyzko, Large deviations of sums of independent random variables, *Acta Arith.*, 49(4):427–434, 1988.

17. H.L. Montgomery and R.C. Vaughan, *Multiplicative Number Theory. I: Classical Theory*, Camb. Stud. Adv. Math., Vol. 97, Cambridge Univ. Press, Cambridge, 2007.
18. C. Perret-Gentil, Gaussian distribution of short sums of trace functions over finite fields, *Math. Proc. Camb. Philos. Soc.*, **163**(3):385–422, 2017.

19. Z. Rudnick and K. Soundararajan, Lower bounds for moments of $L$-functions: Symplectic and orthogonal examples, in *Multiple Dirichlet Series, Automorphic Forms, and Analytic Number Theory*, Proc. Symp. Pure Math., Vol. 75, AMS, Providence, RI, 2006, pp. 293–303.

20. G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, 3rd ed., Grad. Stud. Math., Vol. 163, AMS, Providence, RI, 2015.

21. E.C. Titchmarsh, *The Theory of the Riemann Zeta-Function*, 2nd ed., Clarendon Press, Oxford Univ. Press, New York, 1986. Edited and with a preface by D.R. Heath-Brown.