Some recent interactions of probability and number theory

Corentin Perret-Gentil (Centre de recherches mathématiques, Montréal, Canada)

Around eighty years after its birth, the field of probabilistic number theory continues to see very interesting developments. On the occasion of a thematic program on the subject that took place last May in Montréal, we give a brief survey of a (far from exhaustive) selection of recent advances.

1 Introduction

While many number theoretical questions are statistical in nature, *probabilistic number theory* is usually understood as the use of probabilistic techniques or ideas in number theory [44], analogously to what analysis is to analytic number theory. The association of the two latter was notably proved fruitful in Hadamard and de la Valllée Poussin's 1896 proof of the prime number theorem

$$\pi(x) \sim \frac{x}{\log x} \qquad (x \to \infty), \tag{1}$$

that gives an asymptotic for the number $\pi(x)$ of prime numbers $p \le x$.

The origins of probabilistic number theory can be traced back to Turán's new proof [47], in 1934, of the result by Hardy and Ramanujan from 1917 on the normal order of ω, the "number of prime divisors" function: for any $\varepsilon > 0$,

$$|\omega(n) - \log\log n| \le (\log\log n)^{1/2+\varepsilon}$$

for almost all integers n (that is, the proportion of integers $n \le N$ such that this does not hold goes to 0 as $N \to \infty$). Turán's argument can be seen as based on Chebyshev's inequality

$$\mathrm{Prob}\left(|X - \mathbb{E}(X)| \ge \alpha\sqrt{\mathrm{Var}(X)}\right) \le 1/\alpha^2,$$

for any $\alpha > 0$ and X a random variable with finite expected value $\mathbb{E}(X)$ and variance $\mathrm{Var}(X)$. However, as Elliott reports [11, Ch. 12], Turán did not realize it before a letter of the probabilist Mark Kac. It is relevant to recall here that, at the time of Turán's paper, Kolmogorov's axiomatization of probability theory had just been published, in 1933.

Still according to [11], Kac asked Turán whether he could compute higher asymptotic moments of ω, maybe suggesting that ω had a Gaussian limiting probability distribution. Using the concept of independent random variables and the central limit theorem[1], Erdős and Kac ([12], 1940) proved this and strengthened the Hardy–Ramanujan theorem by showing that

$$f(n) = \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \qquad (n \in \mathbb{N})$$

has standard normal limiting probability distribution, i.e.

$$\frac{|\{n \le x : f(n) \le z\}|}{x} \xrightarrow{x\to\infty} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-t^2/2}dt =: \Phi(z)$$

for any $z \in \mathbb{R}$. Along with the Erdős–Wintner theorem (1939) on limiting distributions of additive functions on the integers, this can be seen as the beginning of probabilistic number theory. We refer the reader to W. Schwarz's survey [44] for an account of the main developments that followed.

Herein, we would like to give a brief survey of a (far from exhaustive) selection of recent works that use concepts such as martingales, suprema of Gaussian and log-correlated processes, orderings of weakly correlated random variables, normal approximations, large deviation estimates, comparison inequalities, and random Fourier series, to obtain significant results or insights in number theory.

2 Random multiplicative functions

We recall that the Möbius function $\mu : \mathbb{N} \to \{-1, 0, 1\}$ is defined by $\mu(n) = (-1)^{\omega(n)}$ if n is squarefree and $\mu(n) = 0$ otherwise. It is multiplicative, and the prime number theorem (1) is equivalent to

$$M_\mu(x) := \sum_{n\le x}\mu(n) = o(x) \quad \text{and} \quad \sum_{n\le x}\frac{\mu(n)}{n} = o(1)$$

as $x \to \infty$. The summatory function $M_\mu(x)$ is called Mertens' function, and it turns out that the Riemann hypothesis is equivalent to $M_\mu(x) = O(x^{1/2+\varepsilon})$ for any $\varepsilon > 0$. On the other hand, we have

$$\overline{\lim_{x\to\infty}} M_\mu(x)/\sqrt{x} > 0 \quad \text{and} \quad \underline{\lim_{x\to\infty}} M_\mu(x)/\sqrt{x} < 0,$$

and an unpublished conjecture of Gonek states that

$$\overline{\lim_{x\to\infty}} \frac{M_\mu(x)}{\sqrt{x}(\log\log\log x)^{5/4}} = \pm A$$

for some constant $A > 0$ (see [40]).

As a heuristic for sums of multiplicative functions such as $M_\mu(x)$, Wintner ([48], 1944) studied sums of *Rademacher random multiplicative functions* f: for $f(p)$ a sequence of independent identically distributed (iid) random variables indexed by primes, uniform on $\{-1, 1\}$, we let

$$f(n) := \begin{cases} \prod_{p|n} f(p) & n \text{ squarefree} \\ 0 & \text{otherwise.} \end{cases}$$

Wintner showed that for any $\varepsilon > 0$,

$$M_f(x) := \sum_{n\le x} f(n) \ll_\varepsilon x^{1/2+\varepsilon} \qquad \text{almost surely}$$

as $x \to \infty$, most recently improved by Lau–Tenenbaum–Wu ([37], 2013) to $M_f(x) \ll_\varepsilon \sqrt{x}(\log\log x)^{2+\varepsilon}$ almost surely, which is to be compared with the law of the iterated logarithm

$$\overline{\lim_{n\to\infty}} \frac{\sum_{i=1}^n X_i}{\sqrt{2n\log\log n}} = 1 \qquad \text{almost surely,}$$

when X_i are iid with mean 0 and variance 1 (e.g. if no multiplicative structure were imposed on f).

On the other hand, Halász ([20], 1982) showed that there exists a constant $B > 0$ such that

$$M_f(x) \neq O\Big(\sqrt{x}e^{-B\sqrt{\log\log x \log\log\log x}}\Big) \quad \text{almost surely.} \quad (2)$$

Alternatively, one may also define $f(p)$ to be uniform on the unit circle in $\mathbb{C}$, giving rise to *Steinhaus random multiplicative functions*.

2.1 Martingales and normal distributions of $M_f^{(k)}(x)$

Given the considerations above about a random multiplicative function f, it would be interesting to obtain information on the limiting distribution of $M_f(x)$ as $x \to \infty$. A simpler object is obtained by restricting the sum to integers having a fixed small number $k \geq 1$ of prime factors, that is

$$M_f^{(k)}(x) := \sum_{\substack{n\leq x\\ \omega(n)=k}} f(n).$$

When k is small, there is not as much multiplicative dependency among the values of f in the sum, and the problem may be more manageable than for the full sum $M_f(x)$. For example, when $k = 1$, the limiting distribution of $M_f^{(1)}(x)/\sqrt{x}$ is standard normal by the central limit theorem.

In 2009, Hough [29] showed:

Theorem 1 ([29]). *For f Rademacher, $z \in \mathbb{R}$ and $k = k(x) = o(\log\log\log x)$, we have a normal limiting distribution:*

$$\mathbb{P}\left(\frac{M_f^{(k)}(x)}{\mathbb{E}\left(M_f^{(k)}(x)^2\right)} \leq z\right) \xrightarrow{x\to\infty} \Phi(z). \quad (3)$$

The proof proceeds, classically, by the method of moments. An important idea is that if n has few prime factors, then it should have a large one.

In 2013, Harper [23] significantly extended the range allowed for k by another method:

Theorem 2 ([23]). *For f Rademacher, $z \in \mathbb{R}$ and $k = k(x) = o(\log\log x)$, the limiting normal distribution (3) still holds.*

His idea, starting from an insight of Blei and Janson ([3], 2004), is to identify a *martingale difference sequence* and apply the central limit theorem for those due to McLeish ([39], 1974). Indeed, we can decompose $M^{(k)}(x) = \sum_{p\leq x} M_p^{(k)}(x)$, where

$$M_p^{(k)}(x) \quad := \quad f(p) \sum_{\substack{n\leq x/p\\ \omega(n)=k-1\\ P(n)<p}} f(n),$$

for $P(n)$ the largest prime factor of n. By the linearity of expectation, it follows that

$$\mathbb{E}\left(M_p^{(k)}(x) \mid f(\ell) \quad (\ell < p \text{ prime})\right) = 0,$$

so that $\left(M_p^{(k)}(x)\right)_p$ is a martingale difference sequence with respect to the filtration $(\sigma(\{f(\ell) : \ell \leq p \text{ prime}\}))_p$. Theorem 2 is then reduced to verifying the hypotheses of McLeish's result, which amounts to number theoretical estimates that constitute most of the paper.

Using a version of Stein's method for normal approximation developed by Chatterjee ([9], 2008), Chatterjee and Soundararajan ([10], 2012) obtained a similar result for sums of f in short intervals.

Note that the range $k = \omega(n) = o(\log\log n)$ of Theorem 2 falls just short of the size of a typical integer given by the Erdős-Kac theorem (see Section 1). One may wonder how large k may be while keeping a normal limiting distribution (3). In the same article [23], Harper also gave the following negative result:

Theorem 3 ([23]). *Let $0 < \varepsilon < A$. The limiting normal distribution (3) does* not *hold if $\varepsilon \log\log x \leq k = k(x) \leq A \log\log x$.*

This is proved by showing that the expectation of a thresholded second moment does not converge to what it should, through a conditioning argument that allows a good estimation.

2.2 Lower bounds for suprema of Gaussian processes, and omega results for $M_f(x)$

Another interesting result by Harper ([21], 2013) is the following strong improvement to Halász's negative result (2):

Theorem 4 ([21]). *For f Rademacher and $\varepsilon > 0$, we have*

$$M_f(x) \neq O\Big(\sqrt{x}(\log\log x)^{-2.5+\varepsilon}\Big) \quad \textit{almost surely.}$$

The main input is new general bounds for suprema of Gaussian processes. Indeed, Halász proof of (2) shows that almost certain lower bounds on $|M_f(x)|$ can be obtained from lower bounds on

$$\sup_{t\geq 1} \exp\Big(S(t,x,f) - \log t - \log\log(t+2)/2\Big), \text{ where}$$
$$S(t,x,f) := \sum_{p\leq x} f(p)\frac{\cos(t\log p)}{p^{1/2+1/\log x}}.$$

Using a multivariate central limit theorem, $f(p)$ can be replaced by a sequence $g(p)$ of independent standard Gaussians. Moreover, t can essentially be assumed to lie in a finite set T.

To analyze the resulting process $(S(t,x,g))_{t\in T}$, Harper develops general lower bounds for upper tail probabilities

$$\mathbb{P}\Big(\max_{t\in T} Z(t) \geq u\Big), \text{ with } Z(t) \text{ jointly standard normal,}$$

which can be non trivial even when u is of moderate size, while existing results require u to be very large. The strategy is to first decompose and condition the probability, and then apply several comparison inequalities, along with the known distribution of the maximum of a Brownian motion. This yields lower bounds on the resulting probabilities that depend on the correlations between the $Z(t)$, which are estimated in the case of the process above.

2.3 Moments of random multiplicative functions

Using the results of Harper [21] mentioned in the previous section, Harper, Nikeghbali and Radziwiłł ([27], 2015) obtained lower bounds on the moments

$$N_f(x,k) := \mathbb{E}\left|M_f(x)\right|^k$$

of $M_f(x)$, improving on results of Bondarenko–Seip ([6], 2016):

Theorem 5 ([27]). *For f Rademacher or Steinhaus, as $x \to \infty$,*

$$N_f(x,1) \gg \sqrt{x}(\log\log x)^{-3+o(1)}. \tag{4}$$

In particular, for $k \in [0,1]$, $N_f(x,2k) \gg x^k(\log\log x)^{-6+o(1)}$.

A first application of the lower bounds from [21] gives that (4) holds for infinitely many x, and a more delicate argument yields the theorem.

The authors also compute certain moments asymptotically[2], relying on a general result of La Bretèche (2001) on mean values of multiplicative functions:

Theorem 6 ([27]). *Let $k \geq 1$ be an integer. There exist explicit constants $C_k, D_k > 0$ such that, as $x \to \infty$:*

1. *For f Steinhaus,*

$$N_f(x,2k) \sim C_k x^k (\log x)^{(k-1)^2}. \tag{5}$$

2. *For f Rademacher and $k \geq 3$,*

$$\mathbb{E}\left(M_f(x)^k\right) \sim D_k x^{k/2} (\log x)^{k(k-3)/2}.$$

From Theorems 5 and 6, they make the following guesses for the remaining moments:

Conjecture 7 ([27]). *For f Steinhaus and $k \in \mathbb{R}^+$, there exists a constant $C_k > 0$ such that*

- *if $k \geq 1$, the asymptotic (5) still holds;*
- *if $k \in [0,1]$, then $N_f(x,2k) \sim C_k x^k$ as $x \to \infty$.*

In particular, the case $k = 1$ would disprove Helson's conjecture ([28], 2010) that $M_f(x)$ exhibits more than square-root cancellation:

Conjecture 8 ([28]). *For f Steinhaus, as $x \to \infty$, $N_f(x,1) = o(\sqrt{x})$.*

This conjecture would be surprising from the point of view of a number theoretical model, but it can be motivated as follows (see also [27, pp. 2–3], [28]): by definition of f, the statement is equivalent to

$$\lim_{T\to\infty} \int_0^T \left|\sum_{n\leq x} n^{-it}\right| dt = o(\sqrt{x}).$$

A first insight is that the inner sum is a multiplicative analogue of the Dirichlet kernel $\sum_{n\leq x} e^{2\pi i n t}$, whose L^1 norm on $[0,2\pi]$ is $\ll \log x$. A second one is that Bondarenko, Heap and Seip ([5], 2015) showed that

$$\lim_{T\to\infty} \int_0^T \left|\sum_{n\leq x} n^{-1/2-it}\right| dt \ll (\log x)^{1/4+o(1)},$$

which is also stronger than square-root cancellation.

In two recent preprints [24, 25], Harper announced explicit formulas for all the moments $\mathbb{E}\left|M_f(x)\right|^k$, with $k \in \mathbb{R}^+$ and f Rademacher or Steinhaus. In particular, for f Steinhaus,

$$N_f(x,1) \asymp \sqrt{x}(\log\log x)^{-1/4},$$

which (for $k = 1$ and f Steinhaus) proves Helson's conjecture 8 and disproves Conjecture 7.

The first step in the computation of the moments is a careful passage to Euler products, reducing to the consideration of expected values of the form

$$\mathbb{E}\left(\left[\int_{-1/2}^{1/2} |F_x(1/2+it)|^2 dt\right]^{k/2}\right), \quad F_x(s) = \prod_{p\leq x}\left(1-\frac{f(p)}{p^s}\right)^{-1},$$

where the random variables $\left(\log|F_x(1/2+it)|\right)_{|t|\leq 1/2}$ are approximately Gaussian and have logarithmic covariance structure. Writing

$$|F_x(1/2+it)|^2 = e^{2h(t)} \text{ with } h(t) = \log|F_x(1/2+it)|,$$

the second step draws links to *critical multiplicative chaos* to analyze these random Euler products.

3 Maximum of the zeta function on bounded intervals

Little is known about the maximum modulus

$$M(T) := \max_{0\leq t\leq T} |\zeta(1/2+it)|$$

of the Riemann zeta function on an initial interval of the critical line. The Lindelöf hypothesis (hence the Riemann hypothesis) implies that

$$M(T) \ll \exp\left(C\frac{\log T}{\log\log T}\right) \qquad \text{as } T \to \infty$$

for some constant $C > 0$. A conjecture of Farmer, Gonek and Hughes ([13], 2007) states that

$$M(T) = \exp\left(\left(\frac{1}{\sqrt{2}} + o(1)\right)\sqrt{\log T \log\log T}\right). \tag{6}$$

Alternatively, one may also consider bounded intervals, for which Fyodorov, Hiary and Keating ([18, 17]) have proposed, based on numerical evidence and links with statistical mechanics the following conjecture:

Conjecture 9. *For t sampled uniformly in $[0,T]$,*

$$\max_{\substack{u\in\mathbb{R}\\ |u-t|\leq 1}} \log|\zeta(1/2+iu)| = \log\log T - \frac{3}{4}\log\log\log T + X_T,$$

where X_T is a random variable converging weakly to an explicit distribution as $T \to \infty$.

Note that by Selberg's central limit theorem ([45], 1946), for $t \in [0,T]$ uniform,

$$\frac{\log|\zeta(1/2+it)|}{\sqrt{(1/2)\log\log T}}$$

converges in law to a standard normal random variable. Moreover, the maximum of independent normal random variables with mean 0 and variance $\frac{1}{2}\log\log T$ is

$$\log\log T - \frac{1}{4}\log\log\log T + O(1),$$

which is a summand $-\frac{1}{2}\log\log\log T$ away from Conjecture 9; the former would account for the non-independence. There

is a multivariate version of Selberg's theorem by Bourgade ([7], 2010), with logarithmic correlations, but this is not enough to make this heuristic rigorous (see [2, p. 4]).

From the analysis of log-correlated random processes (more particularly *branching random walks*), Arguin, Belius, Harper, Radziwiłł and Soundararajan have recently progressed towards Conjecture 9, or an analogue in a random model. This will be the subject of the following sections.

3.1 Supremum of log-correlated Gaussian random variables and leading-order term for a random model

As for conjecture (6) of Farmer–Gonek–Hughes, Conjecture 9 is based on modelling $|\zeta(1/2+is)|$ by the characteristic polynomial of a random unitary matrix.

In 2013, Harper [22] obtained the leading term $\log\log T$ for an analogous model based on random Euler products. The motivation for the model is the following, adapted from an argument of Soundararajan based on the work of Selberg:

Proposition 10 ([22])**.** *Under the Riemann hypothesis, for $T \geq 1$ large enough, there exists $H \subset [T, T+1]$ of relative measure ≥ 0.99 such that for all $t \in H$,*

$$\log|\zeta(1/2+it)| = \mathrm{Re}\left(\sum_{p\leq T}\frac{1}{p^{1/2+it}}\frac{\log(T/p)}{\log T}\right) + O(1).$$

Since $(p^{-it})_{p \text{ prime}}$, for $t \in [0,T]$ uniform, converges as $T \to \infty$, in the sense of finite distributions, to $(U_p)_{p \text{ prime}}$ for U_p iid uniform on the unit circle, this suggest the model

$$\tilde{M}_1(T) := \max_{h\in[0,1]} \sum_{p\leq T} \frac{\mathrm{Re}(U_p p^{-ih})}{p^{1/2}} \tag{7}$$

for $\max_{h\in[0,1]}\log|\zeta(1/2+i(t+h))|$ when t is uniform in $[0,T]$. Harper's main result is then essentially the following:

Theorem 11 ([22])**.** *As $T\to\infty$, $\tilde{M}_1(T) = (1+o_P(1))\log\log T$, where $o_P(1)$ stands for convergence to 0 in probability.*

More precisely, it is also shown that the second-order term should lie between $-2\log\log\log T$ and $-\frac{1}{4}\log\log\log T$. The upper bounds use tail bounds for sums of independent random variables of Talagrand ([46], 1995), while the lower bound uses the bounds from [21] mentioned in Section 2.2.

3.2 Branching random walks and Conjecture 9 for the model

Through a connection with branching random walks, Arguin, Belius and Harper ([2], 2017) managed to obtain the second-order term for the model (7), thus improving Theorem 11:

Theorem 12 ([2])**.** *As $T\to\infty$, we have*

$$\begin{aligned}\tilde{M}_1(T) &= \log\log T - \frac{3}{4}\log\log\log T \\ &\quad + o_P(\log\log\log T),\end{aligned}$$

where the error converges to 0 in probability when divided by $\log\log\log T$.

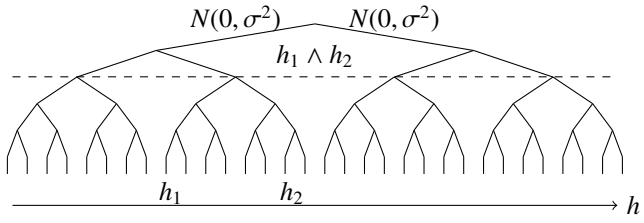

Figure 1: Branching Brownian motion.

Let us assume that $\log T = 2^n$ for some integer $n \geq 1$, and for every $h \in [0,1]$, let $X_n(h) = \sum_{p\leq e^{2^n}} \mathrm{Re}(U_p p^{-ih})p^{-1/2}$. One can compute the covariances explicitly and check that for $h, h' \in [0,1]$,

$$\mathbb{E}\left(X_n(h)X_n(h')\right) \approx \frac{1}{2}\begin{cases}(-1)\log|h-h'| & : |h-h'| \geq 2^{-n} \\ n\log 2 & : |h-h'| < 2^{-n}.\end{cases}$$

Decomposing

$$X_n(h) = \sum_{i=0}^{n} Y_i(h) \quad \text{with} \quad Y_i(h) := \sum_{2^{i-1}<p\leq 2^i} \mathrm{Re}(U_p p^{-ih})p^{-1/2}$$

and letting $h \wedge h' = \lfloor -\log(|h-h'|)/\log 2 \rfloor$, we have that $Y_i(h)$ and $Y_i(h')$ are almost perfectly correlated with variance $\sigma^2 \approx \log(2)/2$ if $i \leq h\wedge h'$, and almost perfectly correlated if $i > h\wedge h'$. Moreover, the variation of the $X_n(h)$ is captured by 2^n equally spaced $h \in [0,1]$.

This is similar to a branching random walk (or branching Brownian motion) on a binary tree of depth n, where iid Gaussian random variables Y_i with mean 0 and variance σ^2 are attached to every edge, and each of the 2^n leaves is associated with the unique random walk on the edges from the root. Indeed, $X_n(h)$ and $X_n(h')$ would correspond to leaves with lowest common ancestor at height $h\wedge h'$ (see Figure 1).

Bramson ([8], 1978) determined the maximum of a branching Brownian motion to be roughly

$$cn - \frac{3\sigma^2}{2c}\log n, \quad \text{where} \quad c = \sigma\sqrt{2\log 2}.$$

When $\sigma^2 = \log(2)/2$, this gives precisely the leading and subleading order terms predicted by Conjecture 9.

Working on this analogy, Arguin, Belius and Harper obtained Theorem 12 using a method of Kistler ([32], 2015) to handle processes like $X_n(h)$ that may not have an exact tree structure and where the Y_i may not be exactly Gaussian. Two particular pieces of information required by the method are large deviation estimates and Berry–Esseen approximations (to compare the Y_i to Gaussians). We refer the reader to [2, pp. 5–7] for a detailed sketch of the strategy.

3.3 Branching random walks and leading-order term for the actual zeta function

In collaboration with Bourgade, Radziwiłł and Soundararajan, Arguin and Belius ([1], 2016) successfully adapted these ideas to obtain the leading-order term in Conjecture 9, that is for the actual zeta function[3]:

Theorem 13 ([1]). *For $\varepsilon > 0$ and t uniform in $[T, 2T]$,*

$$\left| \max_{\substack{u \in \mathbb{R} \\ |u-t| \le 1}} \log|\zeta(1/2+iu)| - \log\log T \right| < \varepsilon$$

almost surely, as $T \to \infty$.

The implied upper bound follows from a Sobolev-type inequality and known bounds on the moments of ζ. For the lower bound, the first step is to reduce to understanding the maximum of finite Dirichlet series $\sum_{p \le X} p^{-\sigma+iu}$, where $\sigma \approx 1/2$; this is done using ideas from the alternative proof by Radziwiłł and Soundararajan ([41], 2015) of Selberg's central limit theorem. Then, an approximate branching random walk is identified and studied with Kistler's method, as in [2] with the model arising from Proposition 10.

4 Biases in prime number races

The prime number theorem in arithmetic progressions, a quantitative version of Dirichlet's theorem on primes in arithmetic progressions, states that if $q \ge 1$ and $a \in (\mathbb{Z}/q)^\times$, then

$$\pi(x,q,a) := \frac{|\{p \le x : p \equiv a \pmod q\}|}{\pi(x)} \sim \frac{1}{\varphi(q)}$$

as $x \to \infty$, where $\pi(x)$ is the number of primes $p \le x$. In other words, primes equidistribute in admissible congruence classes.

In particular, if $a_1, a_2 \in (\mathbb{Z}/q)^\times$, then $\frac{\pi(x,q,a_1)}{\pi(x,q,a_2)} \to 1$. One may wonder whether $\pi(x,q,a_1) > \pi(x,q,a_2)$ for infinitely many x, and if the two orderings are equally likely. Chebyshev observed in 1853 that, surprisingly, we have $\pi(x,4,3) > \pi(x,4,1)$ most of the time.

More generally, for $n, q \ge 2$ fixed and

$$\boldsymbol{a} \in \mathcal{A}_n(q) := \{\boldsymbol{a} \in ((\mathbb{Z}/q)^\times)^n : a_1, \dots, a_n \text{ distinct}\},$$

we can study the $x \ge 2$ such that the ordering

$$\pi(x,q,a_1) > \pi(x,q,a_2) > \cdots > \pi(x,q,a_n) \tag{8}$$

holds (a "Shanks–Rényi prime number race").

In a breakthrough work, Rubinstein and Sarnak ([43], 1994) showed that, conditionally on the General Riemann Hypothesis (GRH) and the $\mathbb{Q}$-linear independence of non-negative imaginary parts of non-trivial zeros of Dirichlet L-functions (conjecture LI), the ordering (8) happens for infinitely many x, actually for a positive logarithmic density[4]

$$\delta(\boldsymbol{a},q) := \lim_{X\to\infty} \frac{1}{\log X} \int_2^X \delta_{(8)\text{ holds}} \frac{dx}{x}.$$

Under these conjectures, they confirm Chebyshev's observation by showing that $\delta(3,1,4) = 0.9959\dots$ In general, they give an explicit expression for the densities $\delta(\boldsymbol{a},q)$, a criterion for the symmetry of the density function (in which cases the races are unbiased, i.e. $\delta(\boldsymbol{a},q) = 1/n!$ for all $\boldsymbol{a} \in \mathcal{A}_n(q)$), and show that the biases dissolve as $q \to \infty$, that is

$$\lim_{q\to\infty} \max_{\boldsymbol{a}\in\mathcal{A}_n(q)} |\delta(\boldsymbol{a},q) - 1/n!| = 0 \tag{9}$$

when $n \ge 2$ is fixed. To do so, the main step is the following:

Theorem 14 ([43]). *As $X \to \infty$,*

$$(x \in [2,X]) \mapsto \left(\frac{\log x}{\sqrt{x}} (\varphi(q)\pi(x,q,a_i) - \pi(x)) \right)_{1\le i\le n}$$

has limiting distribution given by the random vector

$$\begin{aligned} X_{q,a_1,\dots,a_n} &= (X(q,a_1),\dots,X(q,a_n)), \text{ where} \\ X(q,a) &= -C_q(a) \\ &\quad + \sum_{\substack{\chi \pmod q \\ \chi\neq\chi_0}} \sum_{\gamma_\chi > 0} \frac{2\operatorname{Re}(\chi(a)U(\gamma_\chi))}{\sqrt{1/4+\gamma_\chi^2}}, \\ C_q(a) &= -1 + \sum_{\substack{b^2\equiv a \pmod q \\ 1\le b\le q}} 1, \end{aligned}$$

for $U(\gamma_\chi)$ iid on the unit circle in $\mathbb{C}$ and γ_χ running over the non-negative zeros of $L(1/2+i\gamma_\chi,\chi)$.

We refer the reader to [19] for a survey of subsequent results.

In the remainder of this section, we will direct our attention to recent advances in *prime number races with many contestants*, that is when n is allowed to grow to infinity with q, instead of being fixed. All the results stated will be conditional on the GRH and LI conjectures.

4.1 Prime races with many contestants

Feuerverger and Martin ([14], 2000) conjectured that (9) still holds when $n = n(q) \to \infty$, $n \le \varphi(q)$, i.e. the biases still dissolve.

In 2012, Lamzouri [35] obtained a first uniform version of (9) in a certain range, namely:

Theorem 15 ([35]). *If $2 \le n \le \sqrt{\log q}$ and $\boldsymbol{a} \in \mathcal{A}_n(q)$, then*

$$\delta(\boldsymbol{a},q) = \frac{1}{n!}\left(1 + O\left(\frac{n^2}{\log q}\right)\right).$$

Note that the second summand of $X(q,a)$ above is given as a weighted sum of independent random variables, and $C_q(a)$ can essentially be ignored. The idea behind Theorem 15 is to approximate $X_{q,a_1,\dots,a_n}$ as a multivariate normal random variable (through a quantitative central limit theorem), with an estimated covariance matrix, and to then directly estimate the resulting density function.

Concerning larger ranges of n, an unpublished conjecture of Ford and Lamzouri states that there should actually be a transition when $n = (\log q)^{1+o(1)}$:

Conjecture 16. *Let $\varepsilon > 0$, $q \ge 2$ be large enough, and $n = n(q)$.*

1. *If $2 \le n \le (\log q)^{1-\varepsilon}$, then (9) holds.*
2. *If $(\log q)^{1+\varepsilon} \le n \le \varphi(q)$, then there are extreme biases, namely there exist $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{A}_n(q)$ such that*

$$n!\delta(\boldsymbol{a},q) \to 0 \quad \text{and} \quad n!\delta(\boldsymbol{b},q) \to \infty.$$

A stronger version of the first part of the conjecture was proved last year by Harper and Lamzouri ([26], 2018):

Theorem 17 ([26]). *If $2 \le n \le (\log q)/(\log\log q)^4$, then (9) holds. More precisely, for any $\boldsymbol{a} \in \mathcal{A}_n(q)$,*

$$\delta(\boldsymbol{a}, q) = \frac{1}{n!}\left(1 + O\left(\frac{n(\log n)^4}{\log q}\right)\right).$$

This follows the same strategy as Theorem 15 above (performing the normal approximation with a 2009 result of Reinert–Röllin [42]), but with better estimates for the covariances through harmonic analysis.

4.2 Orderings of weakly correlated normal random variables and leader in prime races

From there, we can also study the "leader" in a prime number race, i.e. for $\boldsymbol{a} \in \mathcal{A}_n(q)$, the logarithmic density $\delta_1(\boldsymbol{a}, q)$ of the $x \ge 2$ such that

$$\pi(x, q, a_1) > \pi(x, q, a_2), \dots, \pi(x, q, a_n). \tag{10}$$

An application of Theorem 17 shows that $\delta_1(\boldsymbol{a}, q) \to 1/n$ if $2 \le n = n(q) = o(\log q/(\log\log q)^4)$, i.e. each contestant has an equal chance of being the leader in this range. However, Harper and Lamzouri showed that this can be significantly extended with more involved arguments:

Theorem 18 ([26]). *We have $\delta_1(\boldsymbol{a}, q) \to 1/n$ as soon as $2 \le n = n(q) \le \varphi(q)^{1/32}$.*

After the normal approximation and the estimation of the covariances, this is derived from a general result on the ordering of weakly correlated jointly normal random variables that Harper and Lamzouri establish:

Theorem 19 ([26]). *For $n \ge 2$ and $\varepsilon > 0$, let $X_1, \dots, X_n$ be jointly normal random variables, each with mean zero and variance one. Let $r_{i,j} = \mathbb{E}(X_i X_j)$ denote the covariances, and assume that $|r_{i,j}| \le \varepsilon$ whenever $i \ne j$. Then*

$$\left|P\left(X_1 > \max_{2\le i\le n} X_i\right) - \frac{1}{n}\right| \ll_\varepsilon n^{-100} + n^{-1.99}\sum_{2\le i\le n}|r_{1,i}| + n^{-2.99}\sum_{2\le i<j\le n}|r_{i,j}|.$$

If the X_i were independent, then the probability would be exactly $1/n$. An important input in the proof of Theorem 19 is the use of the normal comparison result of Li–Shao ([38], 2002) to compare $X_1, \dots, X_n$ to independent normal random variables. However, as the probabilities may be small with respect to rather large bounds on the covariances (as is the case for Theorem 18), this alone may yield a trivial bound. To overcome this, the authors note that if the X_i were independent, then

$$\max_{2\le j\le n} X_j = \sqrt{(2 - o(1))\log n} \quad \text{with high probability,}$$

while

$$P\left(X_1 > \sqrt{(2 - o(1))\log n}\right) = \frac{1}{n^{1-o(1)}}.$$

In other words, the small probability $1/n$ that X_1 is the leader is mostly caused by the event that X_1 is large enough to be so. Conditioning on the latter gives bounds which are more achievable, and using Slepian's comparison inequality (which is single-sided unlike that of Li–Shao, but always non-trivial) allows one to conclude the argument.

Using similar ideas, Harper and Lamzouri also obtain estimates for the logarithmic density $\delta_k(\boldsymbol{a}, q)$ (for $k < n$) of the $x \ge 2$ such that

$$\pi(x, q, a_1) > \cdots > \pi(x, q, a_k) > \max_{k+1\le j\le n} \pi(x, q, a_j). \tag{11}$$

4.3 Normal approximation and extreme biases

In the direction of the second part of Conjecture 16, Harper and Lamzouri ([26], 2018) gave one of the first results (along with work of Fiorilli [15]) where biases do *not* dissolve asymptotically:

Theorem 20 ([26]). *Let $\varepsilon > 0$. There exists a constant $c_\varepsilon > 0$ such that if $\varphi(q)^\varepsilon \le n \le \varphi(q)$, there exists $\boldsymbol{a} \in \mathcal{A}_n(q)$ with*

$$\delta(\boldsymbol{a}, q) < (1 - c_\varepsilon)/n!.$$

The idea is to get biases for $\delta_k(\boldsymbol{a}, q)$, which then gives biases for $\delta(\boldsymbol{a}, q)$ by summing over permutations of the components of $\boldsymbol{a}$. The ordering (11) corresponds to the ordering

$$X(q, a_1) > \cdots > X(q, a_k) > \max_{k+1\le j\le n} X(q, a_j) \tag{12}$$

of the random variables in Theorem 14. If this holds, then $X(q, a_1), \dots, X(q, a_k)$ (after renormalization) should all be larger than $\approx \sqrt{2\log n}$ with high probability, as in the previous section. If $(a_1, \dots, a_k) \in \mathcal{A}_k(q)$ is chosen so that $X(q, a_1), \dots, X(q, a_k)$ have (maximum) pairwise correlations $\approx -\frac{\log 2}{\log q}$, this introduces a bias of size $\approx \frac{k\log n}{\log q}$ inside the exponential of the density function, after normal approximation as in Theorem 17.

However, note that the biases in Theorem 20 are always close to and smaller than 1, unlike the extreme biases predicted by Conjecture 16. In a recent preprint, Ford, Harper and Lamzouri [16] improve on this by showing that the second part of the conjecture holds, actually as soon as $n/\log q \to \infty$. More precisely:

Theorem 21 ([16]). *There exists an absolute constant $C > 0$ such that if $1 \ll n \le \varphi(q)$, there exist $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{A}_n(q)$ with*

$$\delta(\boldsymbol{a}, q) \le \exp\left(-\frac{\min(n, \varphi(q)^{1/50})}{C\log q}\right)\frac{1}{n!},$$
$$\delta(\boldsymbol{b}, q) \ge \exp\left(\frac{\min(n, \varphi(q)^{1/50})}{C\log q}\right)\frac{1}{n!}.$$

This follows the same strategy as the one sketched above for Theorem 20, with two main improvements: to get extreme biases, the parameter k is allowed to grow to the order of magnitude of n instead of being fixed; to get small and large biases, the situation (12) is replaced by a slightly different one. One of the issues that arises is that the typical Berry–Esseen type errors in the normal approximation of $X_{q,a_1,\dots,a_n}$ are too large with respect to the main term. To overcome this, the authors develop a multivariate "moderate deviation" estimate for sums of independent random variables, from the Lindeberg replacement strategy.

5 Random Fourier series and paths of partial exponential sums

We conclude with works on exponential sums that also involve some deep results from algebraic geometry.

An exponential sum of fundamental interest in number theory is the Kloosterman sum

$$\mathrm{Kl}_{2,p}(a) = \frac{1}{\sqrt{p}} \sum_{x \in \mathbb{F}_p^\times} \exp\left(\frac{2\pi i(ax+\overline{x})}{p}\right) \quad (a \in \mathbb{F}_p^\times).$$

These are a real numbers, and Weil's proof of the Riemann hypothesis for curves over finite fields (1948) shows that they lie in $[-2,2]$. In practice, partial Kloosterman sums

$$\mathrm{Kl}_{2,p}(a,t) = \frac{1}{\sqrt{p}} \sum_{0 \le x \le tp} \exp\left(\frac{2\pi i(ax+\overline{x})}{p}\right) \in \mathbb{C},$$

for $t \in [0,1)$, are just as interesting. When t is not an integer multiple of $1/p$, let us replace $\mathrm{Kl}_{2,p}(a,t)$ by a linear interpolation of the values at the two closest integers. For every $a \in \mathbb{F}_p^\times$, this gives a continuous path $t \mapsto \mathrm{Kl}_{2,p}(a,t)$ made of straight lines, that ends at $\mathrm{Kl}_{2,p}(a) \in [-2,2]$ (see Figure 2). In the 1980s, such paths of partial exponential sums were studied by Lehmer, Dekking–Mendès France, Loxton, and Deshouillers.

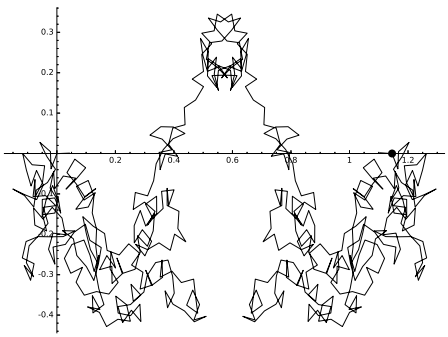

Figure 2: Partial sums of the Kloosterman sum $\mathrm{Kl}_{2,547}(1)$**.**

For the Kloosterman sum, with the uniform measure on $\mathbb{F}_p^\times$, we get a stochastic process

$$(K_p(t))_{t\in[0,1)},$$

whose limiting distribution (as $p \to \infty$) was recently studied by Kowalski and Sawin ([33], 2014). To do so, they define the random Fourier series

$$K(t) = \sum_{h \in \mathbb{Z}} \frac{\exp(2\pi i h z) - 1}{2\pi i h} \mathrm{ST}_h \qquad (t \in [0,1]),$$

where $(\mathrm{ST}_h)_{h\in\mathbb{Z}}$ are independent random variables distributed according to the Sato-Tate measure $\frac{1}{2\pi}\sqrt{4-x^2}$ on $[-2,2]$. Their main result is the following:

Theorem 22 ([33])**.** *1. $K(t)$ converges almost surely and in law, taking symmetric partial sums. The limit, as a random function, is almost surely continuous. For any $t \in [0,1)$, $\mathbb{E}(K(t)) = 0$ and $\mathrm{Var}(K(t)) = t$.*

2. In the sense of convergence of finite distributions,

$$\left(K_p(t)\right)_{t\in[0,1)} \xrightarrow[fd]{p\to\infty} \left(K(t)\right)_{t\in[0,1)}.$$

The proof of the second part uses the method of moments. The latter are estimated asymptotically from the work of Katz [30, 31], relying in particular on Deligne's generalization of the Riemann hypothesis over finite fields.

One could also ask for the stronger result of convergence in law as $C([0,1))$-valued random variables. It turns out that this is linked to important conjectures on short exponential sums. Using Prokhorov's theorem (that is, by checking Kolmogorov's tightness criterion), Kowalski and Sawin show the convergence in law unconditionally for Birch sums

$$\mathrm{Bi}_p(a,t) = \frac{1}{\sqrt{p}} \sum_{n \le t} \exp\left(\frac{2\pi i(ax+x^3)}{p}\right) \quad (a \in \mathbb{F}_p^\times)$$

and for a two-dimensional domain variant of $\mathrm{Kl}_{2,p}$.

This yields interesting applications, such as bounds for the probability of large values of partial Kloosterman sums and partial Birch sums, which is analogous to the recent results of Bober, Goldmakher, Granville and Koukoulopoulos ([4], 2018) for Dirichlet characters. For example, we have:

Theorem 23 ([33])**.** *For $A > 0$, let*

$$L(A) = \lim_{p\to\infty} \frac{|\{a \in \mathbb{F}_p^\times : \max_{0\le t<p} |\mathrm{Bi}_p(a,t)| > A\}|}{p-1}.$$

There exists a constant $c > 0$ such that for any $A > 0$,

$$c^{-1}\exp(-\exp(Ac)) \le L(A) \le c\exp(-\exp(A/c)).$$

This follows from Theorem 22, with elementary arguments for the lower bound, and general tail bounds on sums of martingale difference sequences for the upper bound (see also Section 2.1).

Theorem 23 was very recently improved in a preprint of Lamzouri [36], in particular by obtaining upper and lower bounds in a uniform range for A with respect to p, and of roughly the same order of magnitude. The lower bound holds similarly for Kloosterman sums, while the upper bounds is conditional on certain bounds on short Kloosterman sums.

Finally, the support of the random Fourier series $K(t)$ was computed by Kowalski and Sawin in a subsequent work ([34], 2017), with further arithmetic applications to Kloosterman sums having all their partial sums small.

Notes

1. A proof using the method of moments was given by Halberstam in 1955.
2. This was also obtained independently by Granville and Soundararajan (unpublished), and by Heap and Lindqvist (2016).
3. The same result was obtained independently at the same time by J. Najnudel, under the Riemann hypothesis.
4. However, the natural density does not exist, disproving a conjecture of Knapowksi and Turán.

Acknowledgements The author thanks Lucile Devin, Javier Fresán, Adam Harper and Youness Lamzouri for helpful comments on the manuscript.

References

[1] Louis-Pierre Arguin, David Belius, Paul Bourgade, Maksym Radziwiłł, and K. Soundararajan. Maximum of the Riemann zeta function

on a short interval of the critical line. *Comm. Pure and Applied Math.*, 2016.

[2] Louis-Pierre Arguin, David Belius, and Adam J. Harper. Maxima of a randomized Riemann zeta function, and branching random walks. *Ann. Appl. Probab.*, 27(1):178–215, 2017.

[3] Ron Blei and Svante Janson. Rademacher chaos: tail estimates versus limit theorems. *Ark. Mat.*, 42(1):13–29, 2004.

[4] Jonathan Bober, Leo Goldmakher, Andrew Granville, and Dimitris Koukoulopoulos. The frequency and the structure of large character sums. *J. Eur. Math. Soc.*, 2018.

[5] Andriy Bondarenko, Winston Heap, and Kristian Seip. An inequality of Hardy-Littlewood type for Dirichlet polynomials. *J. Number Theory*, 150:191–205, 2015.

[6] Andriy Bondarenko and Kristian Seip. Helson's problem for sums of a random multiplicative function. *Mathematika*, 62(1):101–110, 2016.

[7] Paul Bourgade. Mesoscopic fluctuations of the zeta zeros. *Probab. Theory Related Fields*, 148(3-4):479–500, 2010.

[8] Maury D. Bramson. Maximal displacement of branching Brownian motion. *Comm. Pure Appl. Math.*, 31(5):531–581, 1978.

[9] Sourav Chatterjee. A new method of normal approximation. *Ann. Probab.*, 36(4):1584–1610, 2008.

[10] Sourav Chatterjee and Kannan Soundararajan. Random multiplicative functions in short intervals. *Int. Math. Res. Not. IMRN*, 2012(3):479–492, 2012.

[11] P. D. T. A. Elliott. *Probabilistic number theory. II*, volume 240 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin-New York, 1980. Central limit theorems.

[12] Paul Erdös and Mark Kac. The Gaussian law of errors in the theory of additive number theoretic functions. *Amer. J. Math.*, 62:738–742, 1940.

[13] David W. Farmer, Steven M. Gonek, and Christopher P. Hughes. The maximum size of L-functions. *J. Reine Angew. Math.*, 609:215–236, 2007.

[14] Andrey Feuerverger and Greg Martin. Biases in the Shanks-Rényi prime number race. *Experiment. Math.*, 9(4):535–570, 2000.

[15] Daniel Fiorilli. Highly biased prime number races. *Algebra Number Theory*, 8(7):1733–1767, 2014.

[16] Kevin Ford, Adam J. Harper, and Youness Lamzouri. Extreme biases in prime number races with many contestants. 2017. Preprint, arXiv:1711.08539.

[17] Yan V. Fyodorov, Ghaith A. Hiary, and Jonathan P. Keating. Freezing transition, characteristic polynomials of random matrices, and the riemann zeta function. *Phys. Rev. Lett.*, 108, 2012.

[18] Yan V. Fyodorov and Jonathan P. Keating. Freezing transitions and extreme values: random matrix theory, and disordered landscapes. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 372(2007), 2014.

[19] Andrew Granville and Greg Martin. Prime number races. *Amer. Math. Monthly*, 113(1):1–33, 2006.

[20] Gábor Halász. On random multiplicative functions. In *Hubert Delange colloquium (Orsay, 1982)*, volume 83 of *Publ. Math. Orsay*, pages 74–96. Univ. Paris XI, Orsay, 1983.

[21] Adam J. Harper. Bounds on the suprema of Gaussian processes, and omega results for the sum of a random multiplicative function. *Ann. Appl. Probab.*, 23(2):584–616, 2013.

[22] Adam J. Harper. A note on the maximum of the Riemann zeta function, and log-correlated random variables. 2013. arXiv:1304.0677.

[23] Adam J. Harper. On the limit distributions of some sums of a random multiplicative function. *J. Reine Angew. Math.*, 2013(678), 2013.

[24] Adam J. Harper. Moments of random multiplicative functions, I: low moments, better than squareroot cancellation, and critical multiplicative chaos. 2017. Preprint, arXiv:1703.06654.

[25] Adam J. Harper. Moments of random multiplicative functions, II: high moments. 2018. Preprint, arXiv:1804.04114.

[26] Adam J. Harper and Youness Lamzouri. Orderings of weakly correlated random variables, and prime number races with many contestants. *Probab. Theory Related Fields*, 170(3-4):961–1010, 2018.

[27] Adam J. Harper, Ashkan Nikeghbali, and Maksym Radziwiłł. A note on Helson's conjecture on moments of random multiplicative functions. In Carl Pomerance and Michael Th. Rassias, editors, *Analytic Number Theory*, pages 145–169. Springer International Publishing, Cham, 2015.

[28] Henry Helson. Hankel forms. *Studia Math.*, 198(1):79–84, 2010.

[29] Bob Hough. Summation of a random multiplicative function on numbers having few prime factors. *Math. Proc. Cambridge Philos. Soc.*, 150(2):193–214, 2011.

[30] Nicholas M. Katz. *Gauss sums, Kloosterman sums, and monodromy Groups*, volume 116 of *Annals of Math. Studies*. Princeton University Press, 1988.

[31] Nicholas M. Katz. *Exponential sums and differential equations*, volume 124 of *Annals of Mathematical Studies*. Princeton University Press, 1990.

[32] Nicola Kistler. Derrida's random energy models. From spin glasses to the extremes of correlated random fields. In *Correlated random systems: five different methods*, volume 2143 of *Lecture Notes in Math.*, pages 71–120. Springer, 2015.

[33] Emmanuel Kowalski and William F. Sawin. Kloosterman paths and the shape of exponential sums. *Compos. Math.*, 152(7), 2016.

[34] Emmanuel Kowalski and William F. Sawin. On the support of the Kloosterman paths. 2017. Preprint, arXiv:1709.05192.

[35] Youness Lamzouri. The Shanks-Rényi prime number race with many contestants. *Math. Res. Lett.*, 19(3), 2012.

[36] Youness Lamzouri. On the distribution of the maximum of cubic exponential sums. 2018. Preprint, arXiv:1802.09701.

[37] Yuk-Kam Lau, Gérald Tenenbaum, and Jie Wu. On mean values of random multiplicative functions. *Proc. Amer. Math. Soc.*, 141(2):409–420, 2013.

[38] Wenbo V. Li and Qi-Man Shao. A normal comparison inequality and its applications. *Probability Theory and Related Fields*, 122(4):494–508, April 2002.

[39] Don L. McLeish. Dependent central limit theorems and invariance principles. *Ann. Probability*, 2:620–628, 1974.

[40] Nathan Ng. The distribution of the summatory function of the Möbius function. *Proc. London Math. Soc. (3)*, 89(2):361–389, 2004.

[41] Maksym Radziwiłł and K Soundararajan. Selberg's central limit theorem for $|\log \zeta(1/2 + it)|$. *L'Enseignement Mathématique*, 2018. To appear, arXiv:1509.06827.

[42] Gesine Reinert and Adrian Röllin. Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *The Annals of Probability*, 37(6):2150–2173, November 2009.

[43] Michael Rubinstein and Peter Sarnak. Chebyshev's bias. *Experiment. Math.*, 3(3):173–197, 1994.

[44] Wolfgang Schwarz. Some highlights from the history of probabilistic number theory. In *Probability and number theory—Kanazawa 2005*, volume 49 of *Adv. Stud. Pure Math.*, pages 367–419. Math. Soc. Japan, Tokyo, 2007.

[45] Atle Selberg. Contributions to the theory of the Riemann zeta-function. *Arch. Math. Naturvid.*, 48(5):89–155, 1946.

[46] Michel Talagrand. The missing factor in Hoeffding's inequalities. *Ann. Inst. H. Poincaré Probab. Statist.*, 31(4):689–702, 1995.

[47] Paul Turán. On a Theorem of Hardy and Ramanujan. *J. London Math. Soc.*, 9(4):274–276, 1934.

[48] Aurel Wintner. Random factorizations and Riemann's hypothesis. *Duke Math. J.*, 11:267–275, 1944.

Corentin Perret-Gentil [corentin.perretgentil@gmail.com] received his PhD at ETH Zürich in 2016, and is now a postdoctorall fellow at the Centre de recherches mathématiques in Montréal, Canada, working in analytic number theory,