

THE DISTRIBUTION OF VALUES OF ZETA AND L-FUNCTIONS

KANNAN SOUNDARARAJAN

ABSTRACT

We survey recent progress on understanding the distribution of values of zeta and L -functions. In particular, we discuss the problem of moments of $|\zeta(\frac{1}{2} + it)|$ and moments of central L -values in families, where the last 25 years have seen a conjectural understanding of the asymptotics of these moments, together with progress in obtaining good upper and lower bounds in many situations.

MATHEMATICS SUBJECT CLASSIFICATION 2020

Primary 11M06; Secondary 11M26, 11M50, 60F05, 60F10

KEYWORDS

Zeta and L -functions, value distribution, moments, extreme values

This article concerns the distribution of values of the Riemann zeta-function and related L -functions. We begin with a brief discussion of L -values at the edge of the critical strip, which give information on arithmetic invariants such as class numbers. The remainder of the article is concerned with the value distribution of $\zeta(\frac{1}{2} + it)$ and the distribution of central values in families of L -functions. The typical behavior of $\zeta(\frac{1}{2} + it)$ is described by a fundamental theorem of Selberg (discussed in Section 2) which asserts that $\log \zeta(\frac{1}{2} + it)$ is distributed like a complex Gaussian with prescribed mean and variance. Analogues of Selberg's theorem for central values in families of L -functions were conjectured by Keating and Snaith, and we motivate these conjectures and the progress towards them in Section 3. Section 4 begins our treatment of the problem of understanding the moments of $|\zeta(\frac{1}{2} + it)|$ and analogous questions for central L -values. While this is a classical topic, going back to work of Hardy and Littlewood, it is only in the last 25 years that even a good conjectural understanding of the problem has emerged. The Keating–Snaith conjectures for the asymptotics of moments were first developed by pursuing an analogy between values of the zeta function and the values of the characteristic polynomial of large random matrices. These conjectures are described in Section 5, which also shows how the problem of understanding moments is tied up with understanding the *large deviations* range in Selberg's theorem. Progress towards the moment conjectures (see Section 6) has been of three types: (i) understanding asymptotics for small moments in a number of examples, (ii) obtaining lower bounds of the correct order of magnitude (which are known in many cases), and (iii) obtaining in great generality upper bounds of the correct order of magnitude assuming the Generalized Riemann Hypothesis. In Section 7 we discuss what is known about the maximal size of $|\zeta(\frac{1}{2} + it)|$ and central L -values, and speculate on what the truth might be. Finally, in Section 8 we consider briefly an intriguing problem of Fyodorov–Hiary–Keating on understanding the “local maximum” of $|\zeta(\frac{1}{2} + it)|$ for t in intervals of length 1, which is closely connected to problems in *branching Brownian motion* and *Gaussian multiplicative chaos*.

1. VALUES AT THE EDGE OF THE CRITICAL STRIP

It was already observed by Gauss and Dirichlet that certain special values of L -functions encode interesting arithmetic information. Recall that a *discriminant* is an integer $d \equiv 0, 1 \pmod{4}$, and d is called a *fundamental discriminant* if d/m^2 is not a discriminant for any divisor m^2 of d larger than 1. Fundamental discriminants are in one-to-one correspondence with discriminants of quadratic fields $\mathbb{Q}(\sqrt{d})$. Associated to a fundamental discriminant d is the Kronecker–Legendre symbol $\chi_d(n) = \left(\frac{d}{n}\right)$, which is a primitive Dirichlet character $(\text{mod } |d|)$. For example, if p is an odd prime then either p or $-p$ is a fundamental discriminant (depending on whether p is 1 or 3 $(\text{mod } 4)$), and in either case the associated quadratic character is the familiar Legendre symbol $(\text{mod } p)$. Associated to the primitive character χ_d is the Dirichlet L -function

$$L(s, \chi_d) = \sum_{n=1}^{\infty} \frac{\chi_d(n)}{n^s} = \prod_p \left(1 - \frac{\chi_d(p)}{p^s}\right)^{-1}.$$

Although $d = 1$ is permitted in our definition of fundamental discriminants (and corresponds to the Riemann zeta-function), it is an anomalous case and we shall mainly be interested in fundamental discriminants $d \neq 1$. Like the Riemann zeta-function, the Dirichlet L -function $L(s, \chi_d)$ converges absolutely for $\operatorname{Re}(s) > 1$, extends analytically to the entire complex plane (unlike $\zeta(s)$, there is no pole at $s = 1$ here), and satisfies a functional equation connecting values at s to values at $1 - s$. The nontrivial zeros of $L(s, \chi_d)$ lie in the *critical strip* $0 < \operatorname{Re}(s) < 1$, with the Generalized Riemann Hypothesis (GRH) predicting that they lie on the critical line $\operatorname{Re}(s) = \frac{1}{2}$. For background on Dirichlet L -functions see Davenport [50], and for a general comprehensive treatment of analytic number theory (including information on many other families of L -functions that will be considered here) see Iwaniec and Kowalski [92].

In this family of quadratic Dirichlet L -functions, the values $L(1, \chi_d)$ (lying at the edge of the critical strip) are of great arithmetical interest. A key step in Dirichlet's proof that there are infinitely many primes in arithmetic progressions involves showing that $L(1, \chi_d) \neq 0$. Dirichlet established this by finding a beautiful connection between $L(1, \chi_d)$ and the group of equivalence classes of binary quadratic forms of discriminant d which had earlier been studied by Gauss. For example, if d is a negative fundamental discriminant, then Dirichlet's class number formula states that

$$L(1, \chi_d) = \frac{2\pi}{w} \frac{h(d)}{\sqrt{|d|}},$$

where $h(d)$ is a positive integer, namely the class number of the imaginary quadratic field $\mathbb{Q}(\sqrt{d})$, and w counts the number of roots of unity in $\mathbb{Q}(\sqrt{d})$ (so that $w = 2$ for $d < -4$, and $w = 4$ for $d = -4$, and $w = 6$ for $d = -3$). The special case $d = -4$ of the Dirichlet class number formula is widely familiar as the Madhava–Leibniz–Gregory series $1 - 1/3 + 1/5 - 1/7 + \dots = \pi/4$. Another classical connection to these special L -values arises in the Gauss–Legendre three squares theorem. If n is a square-free integer with $n \equiv 3 \pmod{8}$, then the number of ways of writing n as a sum of three squares, $r(n)$, equals $24h(-n)$; a result known to Gauss, together with variants when $n \equiv 1, 2 \pmod{4}$.

These connections motivate the study of the distribution of the values $L(1, \chi_d)$. Here are some natural questions that arise. If fundamental discriminants d are chosen uniformly with $|d| \leq X$, (i) what is the statistical distribution of the values $L(1, \chi_d)$, and (ii) what are the largest and smallest possible values of $L(1, \chi_d)$? As we shall see, the problem of the statistical distribution of $L(1, \chi_d)$ can be understood quite precisely, but there are still large gaps in our understanding of the extreme values.

Let us begin with the simpler situation of $L(2, \chi_d)$ where both the Dirichlet series and Euler product in the definition of $L(s, \chi_d)$ converge absolutely. If the values $\chi_d(p)$ are known for all primes $p \leq z$ then

$$\left| L(2, \chi_d) - \prod_{p \leq z} \left(1 - \frac{\chi_d(p)}{p^2} \right)^{-1} \right| \leq \sum_{n > z} \frac{1}{n^2} = O\left(\frac{1}{z}\right).$$

The value $\chi_d(p) = \left(\frac{d}{p}\right)$ is determined by $d \pmod{p}$ for odd primes p , and for $p = 2$ the value of $\chi_d(p)$ is determined by $d \pmod{8}$. Thus, by the Chinese Remainder Theorem, the

values of $\chi_d(p)$ for $p \leq z$ are determined by d modulo $4 \prod_{p \leq z} p$. One way to view this is as a kind of *almost periodicity*: if two fundamental discriminants d_1 and d_2 are congruent modulo $4 \prod_{p \leq z} p$ then $L(2, \chi_{d_1}) = L(2, \chi_{d_2}) + O(1/z)$.

If p is an odd prime and X is large, then a little calculation shows that a proportion $\frac{1}{p+1}$ of the fundamental discriminants d with $|d| \leq X$ are multiples of p (this is essentially the proportion of square-free integers that are multiples of p) and $\chi_d(p) = 0$ here. The remaining proportion $\frac{p}{p+1}$ of fundamental discriminants are evenly split among the possible values $\chi_d(p) = 1$ or -1 . Pleasantly, it turns out that for $p = 2$ also a proportion $\frac{1}{3}$ of the fundamental discriminants $|d| \leq X$ satisfy each of the three cases $\chi_d(2) = 0, 1$ or -1 . Moreover, the Chinese Remainder Theorem tells us that for different primes p , the values $\chi_d(p)$ are distributed “independently” of each other, at least if we restrict to primes $p \leq z$ with $\prod_{p \leq z} p$ being small in comparison with X . This motivates us to define for prime numbers p , independent random variables $\mathbb{X}(p)$ taking the values 0 with probability $1/(p+1)$ and the values ± 1 with probability $p/(2(p+1))$. Then the distribution of $\prod_{p \leq z} (1 - \chi_d(p)/p^2)^{-1}$ is the same as the distribution of the random Euler product $\prod_{p \leq z} (1 - \mathbb{X}(p)/p^2)^{-1}$. Letting $z \rightarrow \infty$, we have described the distribution of $L(2, \chi_d)$ as being precisely the distribution of $\prod_p (1 - \mathbb{X}(p)/p^2)^{-1}$.

The story for extreme values is also clear:

$$\frac{\zeta(4)}{\zeta(2)} = \prod_p \left(1 + \frac{1}{p^2}\right)^{-1} \leq \prod_p \left(1 - \frac{\chi_d(p)}{p^2}\right)^{-1} = L(2, \chi_d) \leq \prod_p \left(1 - \frac{1}{p^2}\right)^{-1} = \zeta(2).$$

Moreover, we may find values $L(2, \chi_d)$ arbitrarily close to $\zeta(4)/\zeta(2)$ by choosing d with $\chi_d(p) = -1$ for all primes $p \leq z$, and we may find values arbitrarily close to $\zeta(2)$ by choosing d with $\chi_d(p) = 1$ for all primes $p \leq z$.

Let us now turn to the distribution of $L(1, \chi_d)$ where there is a similar story but with some added complications since the series and product defining $L(s, \chi_d)$ are no longer absolutely convergent. For example, one can show that if $z \leq (\log X)^{10}$ then $L(1, \chi_d) = \prod_{p \leq z} (1 - \chi_d(p)/p)^{-1} + O(1/z^{\frac{1}{4}})$ for all but $O(X/z^{\frac{1}{4}})$ of the fundamental discriminants $|d| \leq X$. This again may be viewed as a kind of almost periodicity: allowing z to tend slowly to infinity with X , for almost all pairs of discriminants d_1 and d_2 with $d_1 \equiv d_2 \pmod{4 \prod_{p \leq z} p}$ one has $L(1, \chi_{d_1}) \approx L(1, \chi_{d_2})$.

For primes p , let $\mathbb{X}(p)$ denote the random variables described earlier, and extend \mathbb{X} to all integers using (complete) multiplicativity; thus, if $n = p_1^{e_1} \cdots p_k^{e_k}$ then $\mathbb{X}(n) = \mathbb{X}(p_1)^{e_1} \cdots \mathbb{X}(p_k)^{e_k}$. This is an example of a *random multiplicative function*, and we may correspondingly consider the random L -function

$$L(s, \mathbb{X}) = \sum_{n=1}^{\infty} \frac{\mathbb{X}(n)}{n^s} = \prod_p \left(1 - \frac{\mathbb{X}(p)}{p^s}\right)^{-1}. \tag{1.1}$$

Both the series and product above converge almost surely provided $\operatorname{Re}(s) > \frac{1}{2}$; this follows essentially from the fact that the variance of $\sum_p \mathbb{X}(p)/p^s$ is $\sum_p \frac{p}{p+1} \frac{1}{p^{2\operatorname{Re}(s)}}$, which is a convergent sum when $\operatorname{Re}(s) > \frac{1}{2}$. In particular, the random Euler product $L(1, \mathbb{X})$ converges almost surely, and the values $L(1, \chi_d)$ are distributed like $L(1, \mathbb{X})$. We may see this by first

approximating most $L(1, \chi_d)$ by $\prod_{p \leq z} (1 - \chi_d(p)/p)^{-1}$, noting that this truncated Euler product is distributed exactly like $\prod_{p \leq z} (1 - \mathbb{X}(p)/p)^{-1}$, and finally letting $z \rightarrow \infty$.

Let us state the result discussed above more precisely. Given any $\tau > 0$, the proportion of fundamental discriminants $|d| \leq X$ with $L(1, \chi_d) \geq e^\gamma \tau$ tends as $X \rightarrow \infty$ to $\text{Prob}(L(1, \mathbb{X}) > e^\gamma \tau)$. Here γ is Euler's constant, and we have normalized in this fashion in view of Mertens's theorem $\prod_{p \leq z} (1 - 1/p)^{-1} \sim e^\gamma \log z$. If τ is large, and we seek values of $L(1, \chi_d)$ larger than $e^\gamma \tau$, the most likely way in which such large values arise is when $\chi_d(p) = 1$ for all primes up to about e^τ . Similarly, the proportion of fundamental discriminants $|d| \leq X$ with $L(1, \chi_d) < \zeta(2)/(e^\gamma \tau)$ tends as $X \rightarrow \infty$ to $\text{Prob}(L(1, \mathbb{X}) < \zeta(2)/(e^\gamma \tau))$. The normalization here is made in view of $\prod_{p \leq z} (1 + 1/p)^{-1} \sim \zeta(2)/(e^\gamma \log z)$. The distribution of $L(1, \mathbb{X})$ is continuous—it is more natural to think of the distribution of $\log L(1, \mathbb{X})$ which is smooth—and its tails $\text{Prob}(L(1, \mathbb{X}) > e^\gamma \tau)$ and $\text{Prob}(L(1, \mathbb{X}) < \zeta(2)/(e^\gamma \tau))$ decay double exponentially, behaving like $\exp(-(1 + o(1))e^{\tau - C_1}/\tau)$ for a suitable constant C_1 (see [71]). With high likelihood one has $1/10 \leq L(1, \mathbb{X}) \leq 10$, although there is a small positive probability of finding arbitrarily large or arbitrarily small values.

The qualitative results mentioned above were obtained by Chowla and Erdős [39], and with some uniformity in τ by Elliott [58]. The question of uniformity in τ is studied in more detail by Montgomery and Vaughan [126], and Granville and Soundararajan [71], with the aim of understanding the extreme values of $L(1, \chi_d)$. By “uniformity in τ ,” we mean the problem of allowing τ to depend on X while still guaranteeing that the proportion of $|d| \leq X$ with $L(1, \chi_d) > e^\gamma \tau$ is comparable to the tail probability that $L(1, \mathbb{X}) > e^\gamma \tau$ (and similarly for small values of $L(1, \chi_d)$). In view of the double exponential decay of the tails of the distribution of $L(1, \mathbb{X})$ mentioned above, the largest viable range for uniformity in τ is $\tau \leq \tau_{\max} + \varepsilon$, with $\tau_{\max} = \log \log X + \log \log \log X + C_1$ and any fixed $\varepsilon > 0$ —at this point one has $\text{Prob}(L(1, \mathbb{X}) > e^\gamma \tau_{\max}) < 1/X$. The results in [71] show excellent agreement between the distribution of $L(1, \chi_d)$ and the probabilistic model $L(1, \mathbb{X})$ in almost the entire viable range. These results suggest the following conjectures on the extreme values of $L(1, \chi_d)$:

$$\max_{|d| \leq X} L(1, \chi_d) = e^\gamma (\tau_{\max} + o(1)), \quad \text{and} \quad \min_{|d| \leq X} L(1, \chi_d) = \zeta(2)/(e^\gamma (\tau_{\max} + o(1))). \tag{1.2}$$

In [71] it is shown that there are values of $L(1, \chi_d)$ nearly as large as the conjecture in (1.2) (for example, assuming the truth of GRH one can find values as large as $e^\gamma (\tau_{\max} - C)$ for some constant C) and values almost as small as in (1.2). However, as we shall discuss next, there are large gaps in our understanding of why the extreme values cannot be much larger or smaller.

How large can z be such that for some fundamental discriminant $|d| \leq X$ one has $\chi_d(p) = 1$ for all primes $p \leq z$? This problem is intimately related to finding large values of $L(1, \chi_d)$. Correspondingly, the problem of finding small values of $L(1, \chi_d)$ may be thought of as wanting $\chi_d(p) = -1$ for all primes $p \leq z$. We noted already that the values of $\chi_d(p)$ for $p \leq z$ may be determined by knowing $d \pmod{4 \prod_{p \leq z} p}$. The prime number theorem gives $\prod_{p \leq z} p = e^{z(1+o(1))}$, so that with $z = \frac{1}{2} \log X$ (say) we can find $|d| \leq X$ with any

given signs $\chi_d(p)$ for $p \leq z$ —for example we may make them all 1, or all -1 . If we think of the probabilistic model \mathbb{X} which treats $\chi_d(p)$ as essentially being a “coin toss” we may expect that the primes up to about $z = \log X \log \log X$ (there are about $\log X$ primes below this z) may take any prescribed signs. This dovetails nicely with the conjectured size of extreme values in (1.2), since (in the case of large values) $\prod_{p \leq z} (1 - 1/p)^{-1} \sim e^\gamma \log z \approx e^\gamma (\log \log X + \log \log \log X)$. For primes p larger than about $\log X \log \log X$, we expect randomness to kick in, and to find an equal number of positive and negative values of $\chi_d(p)$.

Our current knowledge is very far from these probabilistic considerations. Given a prime ℓ , Vinogradov conjectured that the least quadratic nonresidue (mod ℓ) lies below $C(\varepsilon)\ell^\varepsilon$ for some constant $C(\varepsilon)$. That is, there must be a prime $p \leq C(\varepsilon)\ell^\varepsilon$ with $(\frac{p}{\ell}) = -1$, which is a weak version of the prediction from the random model that there exists such p with $p \leq C \log \ell \log \log \ell$ for some constant C . Toward Vinogradov’s conjecture, we know, as a consequence of the Burgess bounds for character sums, that the least quadratic nonresidue lies below $\ell^{1/(4\sqrt{e})+o(1)}$ (see [29]), and no improvement over this exponent has been made in more than 50 years. In terms of $L(1, \chi_d)$, the work towards Vinogradov’s conjecture may be used to show that (see [70, 168])

$$L(1, \chi_d) < \left(\frac{1}{4} \left(2 - \frac{2}{\sqrt{e}} \right) + o(1) \right) \log |d|.$$

This is far from the conjecture in (1.2), and even an improvement in the constant above would be significant and lead to an improvement on the bound for the least quadratic nonresidue (see also [20, 72, 169] for related work).

Even less is known about the problem of bounding the least prime p such that p is a quadratic residue (mod ℓ). To give a sense of the interest of this problem, we note that if $\ell \equiv 3 \pmod{4}$ is a prime, then the imaginary quadratic field $\mathbb{Q}(\sqrt{-\ell})$ has class number 1 if and only if $(\frac{p}{\ell}) = -1$ for all $p < (1 + \ell)/4$. For such a prime ℓ , the polynomial $n^2 + n + (1 + \ell)/4$ takes prime values for $0 \leq n < (\ell - 3)/4$. Euler’s famous polynomial $n^2 + n + 41$ is the largest example of this phenomenon, corresponding to the prime $\ell = 163$ for which the first 12 primes (the primes below 41) are all quadratic nonresidues. Toward this problem, we know that the least prime quadratic residue (mod ℓ) lies below $C(\varepsilon)\ell^{\frac{1}{4}+\varepsilon}$ for any $\varepsilon > 0$ (see [82]), but with a constant $C(\varepsilon)$ that is *ineffective* (meaning the proof only shows the existence of $C(\varepsilon)$, but without any way to compute it, even in principle). This is related to Siegel’s ineffective lower bound (see [50]): for any $\varepsilon > 0$ there exists $C(\varepsilon) > 0$ with

$$L(1, \chi_d) > C(\varepsilon)|d|^{-\varepsilon}.$$

Thus our knowledge of small values of $L(1, \chi_d)$ is even further from the conjecture in (1.2).

If we assume the truth of GRH, then much better results are known. On GRH, the least quadratic nonresidue (mod ℓ) can be shown to be $< (\log \ell)^2$, and the least prime quadratic residue also lies below $(1 + o(1))(\log \ell)^2$ (see [114]). Moreover, for any fundamental discriminant d , one has

$$L(1, \chi_d) \sim \prod_{p \leq (\log |d|)^2} \left(1 - \frac{\chi_d(p)}{p} \right)^{-1}, \tag{1.3}$$

so that the extreme values of $L(1, \chi_d)$ over all $|d| \leq X$ are bounded above by $(2 + o(1))e^{\gamma} \tau_{\max}$ and below by $(\frac{1}{2} + o(1))\zeta(2)/(e^{\gamma} \tau_{\max})$. There is still a gap between these GRH bounds and the probabilistic conjecture in (1.2), but now one is off only by a factor of 2, corresponding to the expectation based on the random model that in (1.3) we only need to take the product over primes $p \leq (\log |d|)$ in order to approximate $L(1, \chi_d)$.

To summarize our discussion, the values of $L(1, \chi_d)$ have an almost periodic structure in d , and these values may be accurately modeled by random Euler products. The random model gives a satisfactory description of the statistical distribution of $L(1, \chi_d)$. It also makes predictions on the largest and smallest possible values of $L(1, \chi_d)$, but there is a large gap between these predictions and our current unconditional knowledge, and even assuming GRH there is still a factor of 2 at issue.

Similar results may be established for the distribution at the edge of the critical strip for values in other families of L -functions. For example, consider the distribution of $\zeta(1 + it)$, where t is chosen uniformly from $[T, 2T]$ with $T \rightarrow \infty$. These values may be modeled by the random Euler product

$$\zeta(s, \mathbb{X}) = \prod_p \left(1 - \frac{\mathbb{X}(p)}{p^s}\right)^{-1} = \sum_{n=1}^{\infty} \frac{\mathbb{X}(n)}{n^s}, \tag{1.4}$$

where the random variables $\mathbb{X}(p)$ are independent for different primes p , and are all chosen uniformly from the unit circle $\{|z| = 1\}$, and extended to random variables $\mathbb{X}(n)$ over all natural numbers n by multiplicativity. As before, the product and series both converge almost surely when $\text{Re}(s) > \frac{1}{2}$. Then the statistical distribution of $\zeta(1 + it)$ is identical to that of $\zeta(1, \mathbb{X})$ (equivalently, of $\zeta(1 + iy, \mathbb{X})$ for any real y). We can also formulate an *almost periodicity* result: For any $\varepsilon > 0$, we can find a sequence of *almost periods* τ_n , with $\tau_n \rightarrow \infty$ and $|\tau_{n+1} - \tau_n|$ bounded, such that for T sufficiently large (in terms of any fixed almost period τ) one has $|\zeta(1 + it + i\tau) - \zeta(1 + it)| < \varepsilon$ for almost all $t \in [T, 2T]$. The sequence of almost periods are found by requiring $p^{i\tau} \approx 1$ for all primes p up to some point. For a study of the distribution of $\zeta(1 + it)$, with a focus on uniformity, see Lamzouri [109].

There is an extensive literature concerned with distribution at the edge of the critical strip, and we end this section with references to some further examples. We motivated our discussion of $L(1, \chi_d)$ with the class number formula, which (for negative fundamental discriminants) shows that $2\sqrt{|d|}L(1, \chi_d)/(2\pi)$ is quantized to be an integer. This raises questions on the *granularity* of the distribution of $L(1, \chi_d)$, and shows that in very short scales there must be arithmetic deviations from the random model. These questions are related to the problem of understanding how many imaginary quadratic fields there are with any given class number (see [88, 111, 162]). For positive fundamental discriminants, the class number formula relates $L(1, \chi_d)$ to the product of the class number and the regulator which cannot in general be separated from each other. One way to get around this problem is to order the real quadratic fields by the size of their regulator rather than by discriminant, and this ordering has a pleasing interpretation in terms of lengths of closed geodesics on the hyperbolic surface $\text{PSL}(2, \mathbb{Z}) \backslash \mathbb{H}$. The study of $L(1, \chi_d)$, or the class number $h(d)$, when d is ordered in this way was initiated by Sarnak [152]; it is closely related to specializing dis-

criminants d in suitable quadratic sequences (for example, of the form $4n^2 + 1$, or $n^2 + 4$), and for recent investigations see [49, 110, 145]. For a small sample of investigations in other families of L -functions, see [40, 56, 117, 118, 123].

2. SELBERG'S CENTRAL LIMIT THEOREM

In the previous section we discussed the distribution of values of L -functions at the edge of the critical strip. In fact, similar results hold for the value distribution inside the critical strip, but keeping to the right of the critical line. As an illustration, consider the problem of the distribution of values of $\zeta(\sigma + it)$ where $\frac{1}{2} < \sigma \leq 1$ is fixed, and t is chosen uniformly from $[T, 2T]$ with $T \rightarrow \infty$. The random $\zeta(s, \mathbb{X})$ defined in (1.4) still converges when $\text{Re}(s) = \sigma > \frac{1}{2}$, and one can show that $\zeta(\sigma + it)$ is distributed like $\zeta(\sigma, \mathbb{X})$. To give a very brief indication of the proof, one can show that for any parameter $1 \leq N \leq T$,

$$\frac{1}{T} \int_T^{2T} \left| \zeta(\sigma + it) - \sum_{n \leq N} \frac{1}{n^{\sigma+it}} \right|^2 dt = O\left(\sum_{n > N} \frac{1}{n^{2\sigma}}\right) = O(N^{1-2\sigma}), \quad (2.1)$$

which parallels

$$\mathbb{E} \left[\left| \zeta(\sigma, \mathbb{X}) - \sum_{n \leq N} \frac{\mathbb{X}(n)}{n^\sigma} \right|^2 \right] = \sum_{n > N} \frac{1}{n^{2\sigma}} = O(N^{1-2\sigma}).$$

Since $\sigma > \frac{1}{2}$, the term $N^{1-2\sigma}$ tends to 0 provided N tends to infinity with T , and for such N it follows that for most $t \in [T, 2T]$ one has $\zeta(\sigma + it) \approx \sum_{n \leq N} n^{-\sigma+it}$. If now N tends slowly to infinity with T , then we can show that $\sum_{n \leq N} n^{-\sigma+it}$ is distributed like $\sum_{n \leq N} \mathbb{X}(n)/n^\sigma$, by matching the moments of both quantities, for example. This is a classical result (see Chapter XI of [170]), and a recent quantitative study has been made in [113].

As with the distribution of $\zeta(1 + it)$, there is an almost periodic structure in the values of $\zeta(\sigma + it)$. The partial sums $\sum_{n \leq N} n^{-\sigma-it}$ clearly have an almost periodic structure—if $n^{i\tau} \approx 1$ for all $n \leq N$, then τ will be an almost period for these partial sums—and as we noted above $\zeta(\sigma + it)$ can often be approximated by such partial sums.

For $\frac{1}{2} < \sigma \leq 1$, the values $\zeta(\sigma, \mathbb{X})$ are distributed densely in the complex plane; indeed, for any given complex number z and any $\varepsilon > 0$, with positive probability (depending on z and ε) one has $|\zeta(\sigma, \mathbb{X}) - z| < \varepsilon$. This is not hard to show, starting with the fact that $\log \zeta(\sigma, \mathbb{X})$ is essentially $\sum_p \mathbb{X}(p)/p^\sigma$. It follows that the set $\{\zeta(\sigma + it) : t \in \mathbb{R}\}$ is dense in \mathbb{C} . A related striking *universality* result of Voronin [171] states that if f is any nonvanishing continuous function in $|z| \leq r$ with $0 < r < \frac{1}{4}$, then there exist arbitrarily large values $t \in \mathbb{R}$ such that $|\zeta(\frac{3}{4} + it + z) - f(z)| < \varepsilon$ for all $|z| \leq r$. In other words, the zeta function in a disc of radius r around $\frac{3}{4} + it$ can be made to mimic any given analytic function that does not take the value 0. The value 0 must be excluded in view of the Riemann Hypothesis! There are more precise versions of this result, but we do not pursue this direction further, pointing instead to [10, 106, 112] for recent related work.

We now turn to the distribution of values of $\zeta(\frac{1}{2} + it)$, which forms the main focus of this article. The random Euler product $\zeta(s, \mathbb{X})$ defined in (1.4) does not converge for $s = \frac{1}{2}$.

Indeed, there is no almost periodic structure to the values $\zeta(\frac{1}{2} + it)$, and on the critical line the zeta-function cannot typically be understood simply from a knowledge of p^{it} for small primes p . Instead, we have the following fundamental result of Selberg.

Theorem 2.1 (Selberg [155, 156]). *If T is large, and t is chosen uniformly from $[T, 2T]$, then $\log \zeta(\frac{1}{2} + it)$ is distributed like a complex Gaussian with mean 0 and variance $\log \log T$. In particular, $\operatorname{Re}(\log \zeta(\frac{1}{2} + it))$ and $\operatorname{Im}(\log \zeta(\frac{1}{2} + it))$ are distributed like real Gaussians with mean 0 and variance $\frac{1}{2} \log \log T$.*

To clarify normalizations, we recall that a standard complex Gaussian (of mean 0 and variance 1) has density $\frac{1}{\pi} e^{-|z|^2}$, and that its real and imaginary part are independent real Gaussians with mean 0 and variance $\frac{1}{2}$. Selberg’s theorem gives that for any fixed box \mathcal{B} in the complex plane, as $T \rightarrow \infty$ one has

$$\frac{1}{T} \operatorname{meas} \left\{ T \leq t \leq 2T, \frac{\log \zeta(\frac{1}{2} + it)}{\sqrt{\log \log T}} \in \mathcal{B} \right\} \rightarrow \frac{1}{\pi} \int_{x+iy \in \mathcal{B}} e^{-x^2-y^2} dx dy.$$

In Selberg’s theorem we may omit the countably many zeros of $\zeta(s)$ where the logarithm is not defined. For t not equalling the ordinate of a zero of $\zeta(s)$, the argument of $\zeta(\frac{1}{2} + it)$ (that is, $\operatorname{Im}(\log \zeta(\frac{1}{2} + it))$) is defined by continuous variation along the straight lines from 2 (where the argument is taken to be zero) to $2 + it$ and thence to $1/2 + it$.

Here is a striking illustration of the difference between the value distributions of $\zeta(\frac{1}{2} + it)$ and $\zeta(\sigma + it)$ for $1 \geq \sigma > \frac{1}{2}$. Typically, $|\zeta(\sigma + it)|$ is of constant size, for example, taking values between 1/2 and 2 with positive probability. On the other hand, Selberg’s theorem implies that for any fixed V and large T ,

$$\frac{1}{T} \operatorname{meas} \left\{ T \leq t \leq 2T, \frac{\log |\zeta(\frac{1}{2} + it)|}{\sqrt{\frac{1}{2} \log \log T}} \geq V \right\} \sim \frac{1}{\sqrt{2\pi}} \int_V^\infty e^{-x^2/2} dx, \quad (2.2)$$

so that $|\zeta(\frac{1}{2} + it)|$ is large (say, $> \exp(\varepsilon \sqrt{\log \log T})$) nearly half the time, or $|\zeta(\frac{1}{2} + it)|$ is small (below $\exp(-\varepsilon \sqrt{\log \log T})$) nearly half the time. We noted earlier that the set $\{\zeta(\sigma + it) : t \in \mathbb{R}\}$ is dense in the complex plane. It is rare to find values of $\zeta(\frac{1}{2} + it)$ of constant size, and whether the set $\{\zeta(\frac{1}{2} + it) : t \in \mathbb{R}\}$ is dense in \mathbb{C} remains an intriguing open problem. This question was raised first by Ramachandra; for partial progress, see [108].

The argument principle, together with the functional equation for $\zeta(s)$ and Stirling’s formula, may be used to show that $N(t)$, the number of zeros of $\zeta(s)$ with real part between 0 and 1 and imaginary part between 0 and t , satisfies

$$N(t) = \frac{t}{2\pi} \log \frac{t}{2\pi} - \frac{t}{2\pi} + \frac{7}{8} + S(t) + O\left(\frac{1}{t}\right), \quad \text{where } S(t) = \frac{1}{\pi} \arg \zeta\left(\frac{1}{2} + it\right). \quad (2.3)$$

Thus Selberg’s theorem for $\operatorname{Im}(\log \zeta(\frac{1}{2} + it))$ shows that the remainder term in the asymptotic formula for $N(t)$ has Gaussian fluctuations.

We now give a brief, oversimplified, description of the ideas behind Selberg’s theorem; we caution the reader that some statements below should be taken as merely indicative, and not interpreted as being literally correct. Taking logarithms in the Euler product for $\zeta(s)$,

we may write

$$\log \zeta(s) = \sum_{p,k} \frac{1}{k p^{ks}} = \sum_{n=2}^{\infty} \frac{\Lambda(n)}{\log n} \frac{1}{n^s},$$

where the sums above are over prime powers p^k , and $\Lambda(n)$ is the von Mangoldt function which equals $\log p$ if $n = p^k$ and 0 otherwise. The series above converges absolutely when $\operatorname{Re}(s) > 1$, and it certainly does not converge on the critical line $\operatorname{Re}(s) = \frac{1}{2}$. Nevertheless, we might hope that a truncated sum over prime powers might serve as an approximation to $\log \zeta(s)$ (thinking of $s = \frac{1}{2} + it$ with $T \leq t \leq 2T$). This forms the first step in Selberg's argument, who finds an expression of the form

$$\log \zeta(s) = \sum_{2 \leq n \leq x} \frac{\Lambda(n)}{n^s \log n} + Z_x(s), \quad (2.4)$$

where $Z_x(s)$ is a remainder term that may be thought of as the contribution from zeros ρ of $\zeta(s)$ with $|\rho - s| \leq 1/\log x$. By a complicated argument, Selberg showed how the sum over zeros may in turn also be bounded in terms of sums over primes, and thus shown to be small on average. An alternative argument of Bombieri and Hejhal [21] avoids some of Selberg's difficulties by bounding the average values of $Z_x(s)$ instead of seeking point-wise bounds. Nevertheless, these arguments are technically involved; they are simpler if the Riemann hypothesis is assumed, but can be established unconditionally by relying on a subtle zero-density estimate for zeros of $\zeta(s)$ near the critical line (established by Selberg). Although we have not made the relation (2.4) precise, we give a couple of remarks that may be helpful in thinking about such relations. Firstly, one can think of such relations as variants of the explicit formula connecting zeros and primes. Secondly, in addition to the Euler product, the zeta function possesses a Hadamard product over its zeros

$$s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s) = e^{Bs} \prod_{\rho} \left(1 - \frac{s}{\rho}\right) e^{s/\rho}, \quad (2.5)$$

where the product is over all nontrivial zeros of the zeta-function, and B is a constant. The relation (2.4) has the flavor of a hybrid Euler–Hadamard product (see [69] for work in this direction), taking some primes and some zeros, and it is natural to expect an inverse relationship (or uncertainty principle) between the number of primes that one must take versus the number of zeros that are needed.

Returning to the argument, in the range $x \leq T$, the remainder term $Z_x(s)$ in (2.4) is typically of size $O(\log T/\log x)$; this corresponds to the expected number of zeros of $\zeta(s)$ within $1/\log x$ of $\frac{1}{2} + it$. If we choose $x = T^{1/(\log \log T)^{\frac{1}{4}}}$, for example, then $\log T/\log x = (\log \log T)^{\frac{1}{4}}$ is small in comparison to the typical expected size of $\log \zeta(s)$, which is $\sqrt{\log \log T}$, and therefore the remainder term is negligible. In other words, with this choice of x , the proof of Selberg's theorem reduces to establishing the Gaussian nature of

$$\sum_{2 \leq n \leq x} \frac{\Lambda(n)}{\log n} \frac{1}{n^s} = \sum_{p \leq x} \frac{1}{p^s} + \frac{1}{2} \sum_{p \leq \sqrt{x}} \frac{1}{p^{2s}} + \sum_{\substack{p^k \leq x \\ k \geq 3}} \frac{1}{k p^{ks}}. \quad (2.6)$$

The contribution from prime powers p^k with $k \geq 3$ is $O(1)$ and may be omitted. The contribution from the squares of primes is also negligible; it is $\frac{1}{2} \sum_{p \leq \sqrt{x}} 1/p^{1+2it}$ which behaves roughly like $\frac{1}{2} \log \zeta(1 + 2it)$ and so is of constant size typically. We are left with the contribution of just the primes, which we may understand by computing moments. If k and ℓ are any natural numbers then, for large T ,

$$\frac{1}{T} \int_T^{2T} \left(\sum_{p \leq x} \frac{1}{p^{1/2+it}} \right)^k \left(\sum_{p \leq x} \frac{1}{p^{1/2-it}} \right)^\ell dt = \begin{cases} (1 + o(1))k!(\log \log T)^k & \text{if } k = \ell, \\ o(T) & \text{if } k \neq \ell. \end{cases} \quad (2.7)$$

These moments match asymptotically the moments of a complex Gaussian with mean 0 and variance $\log \log T$, from which Selberg's theorem would follow.

To give a justification for (2.7), we discuss an orthogonality relation for Dirichlet polynomials, which we shall find useful in the sequel. Roughly speaking, integrals over $[T, 2T]$ may be thought of as possessing T "harmonics" that can distinguish between the functions $f_n(t) = n^{it}$ for natural numbers n going up to about T . More precisely, suppose Φ is a smooth function approximating the indicator function of $[1, 2]$. Then, if $\max(M, N) \leq T/\log T$,

$$\begin{aligned} \int \sum_{m \leq M} a(m)m^{it} \overline{\sum_{n \leq N} b(n)n^{it}} \Phi\left(\frac{t}{T}\right) dt &= \sum_{m=n} a(m)\overline{b(n)} T \widehat{\Phi}(0) \\ &\quad + \sum_{m \neq n} a(m)\overline{b(n)} T \widehat{\Phi}\left(T \log \frac{n}{m}\right) \\ &\sim T \widehat{\Phi}(0) \sum_{m=n} a(m)\overline{b(n)}, \end{aligned} \quad (2.8)$$

where the contribution of the "off-diagonal" terms $m \neq n$ is negligible because $T|\log(m/n)| \gg T|m-n|/|m+n| \geq T/(M+N)$ is large and the Fourier transform $\widehat{\Phi}$ decays rapidly.

Write $(\sum_{p \leq x} 1/p^{1/2+it})^k = \sum_{n \leq x^k} a_k(n)/n^{1/2+it}$, so that $a_k(n) = 0$ unless n has exactly k prime factors. If n has prime factorization $p_1^{e_1} \cdots p_r^{e_r}$ with $e_1 + \cdots + e_r = k$ then $a_k(n) = k!/(e_1! \cdots e_r!)$. Then an application of (2.8) shows that the moment in (2.7) is

$$\sim T \sum_{m=n \leq x^k} \frac{a_k(n)a_\ell(n)}{n}.$$

If $k \neq \ell$ then either $a_k(n)$ or $a_\ell(n)$ must be zero, and this case of (2.7) follows. If $k = \ell$, then the diagonal terms are dominated by integers with k distinct prime factors, and so the above is

$$\sim Tk! \sum_{n \leq x^k} \frac{a_k(n)}{n} = Tk! \left(\sum_{p \leq x} \frac{1}{p} \right)^k \sim Tk!(\log \log x)^k,$$

and since $\log \log x$ and $\log \log T$ are close, the other case in (2.7) follows.

This concludes our sketch of the ideas behind Selberg's theorem. Two alternative approaches that work for $\log |\zeta(\frac{1}{2} + it)|$ are given in [115, 142]. These avoid the subtle zero-density estimates near the critical line, and it would be of interest to extend such approaches to $\text{Im}(\log \zeta(\frac{1}{2} + it))$.

3. ANALOGUES OF SELBERG'S THEOREM IN FAMILIES OF L-FUNCTIONS

Selberg's theorem discussed above applies not only to the Riemann zeta-function, but more generally to a large class of L -functions. For example, in [156] Selberg introduced what is now known as the *Selberg class* of L -functions, which formalizes some of the observed properties of automorphic L -functions and is expected to coincide with this class. For a primitive L -function in the Selberg class (or, if one prefers, for a cuspidal automorphic L -function for $\mathrm{GL}_n(\mathbb{Q})$), one expects that $\log L(\frac{1}{2} + it)$ with $T \leq t \leq 2T$ is distributed like a complex Gaussian with mean 0 and variance $\log \log T$. The key ingredient needed to make this precise is an analogue of the zero-density estimate close to the critical line, and this is known for GL_1 and GL_2 ; in the general case, GRH must be assumed (see [21, 156] for more details).

Interesting differences arise when we consider analogues of Selberg's theorem for central values in families of L -functions. There are three categories into which families of L -functions fall, and we illustrate these with examples. Unlike Selberg's *Theorem*, the analogous central limit theorems that we formulate in these families are still *conjectural*, and these conjectures were first formulated by Keating and Snaith [101].

Unitary families. A typical example is the family of all Dirichlet characters $\chi \pmod{q}$, with q a large prime (for simplicity). The question is to understand the distribution of $\log L(\frac{1}{2}, \chi)$ as χ ranges over all primitive characters $\chi \pmod{q}$ (if q is prime, this is equivalent to χ not being the principal character). We must discard potential characters with $L(\frac{1}{2}, \chi) = 0$, but in fact it is conjectured that $L(\frac{1}{2}, \chi) \neq 0$ for all Dirichlet L -functions. This situation is expected to be exactly as in Selberg's theorem, and the Keating–Snaith conjecture for this family states that for large q the distribution of $\log L(\frac{1}{2}, \chi)$ is approximately a complex Gaussian with mean 0 and variance $\log \log q$. In particular, $\log |L(\frac{1}{2}, \chi)|$ is (conjecturally) distributed like a real Gaussian with mean 0 and variance $\frac{1}{2} \log \log q$, so that (like $|\zeta(\frac{1}{2} + it)|$) roughly half the time $|L(\frac{1}{2}, \chi)|$ is as large as $\exp(\varepsilon \sqrt{\log \log q})$ and the other half of the time it is as small as $\exp(-\varepsilon \sqrt{\log \log q})$.

Another example of this type is the family of twists by Dirichlet characters of a fixed newform f . The family $\zeta(\frac{1}{2} + it)$ with $T \leq t \leq 2T$ may also be thought of as an example of a unitary family.

Symplectic families. Consider the family of quadratic Dirichlet L -functions $L(s, \chi_d)$, where d ranges over fundamental discriminants with $|d| \leq X$. The values $L(\frac{1}{2}, \chi_d)$ are real, and GRH predicts that they are all nonnegative (else there would be a real zero of $L(s, \chi_d)$ between $1/2$ and 1). Further, the values $L(\frac{1}{2}, \chi_d)$ are all expected to be nonzero (a conjecture of Chowla, which is a special case of the belief that $L(\frac{1}{2}, \chi) \neq 0$ for all Dirichlet characters χ). The Keating–Snaith conjecture for this family predicts that the values $\log L(\frac{1}{2}, \chi_d)$ are distributed like a real Gaussian with mean $\frac{1}{2} \log \log X$ and variance $\log \log X$. Since the mean is positive, the values of $L(\frac{1}{2}, \chi_d)$ are (conjecturally) of typical size $(\log X)^{\frac{1}{2} + o(1)}$.

Orthogonal families. These families arise naturally in the context of modular forms, and we give a couple of prototypical examples. Let k be an even integer, and consider the family \mathcal{H}_k of all weight k modular forms for the full modular group $\mathrm{SL}_2(\mathbb{Z})$ that are also eigenfunctions of all Hecke operators. Associated to such a form f is its L -function, which we normalize so that the functional equation connects values at s to $1 - s$:

$$\Lambda(s, f) = (2\pi)^{-s} \Gamma\left(s + \frac{k-1}{2}\right) L(s, f) = i^k \Lambda(1-s, f).$$

In the case $k \equiv 2 \pmod{4}$, the sign of this functional equation is -1 , and all the central values $L(\frac{1}{2}, f)$ are zero. In the case $k \equiv 0 \pmod{4}$, the sign of the functional equation is $+1$, and we ask for the distribution of $L(\frac{1}{2}, f)$ (or, in keeping with Selberg's theorem, $\log L(\frac{1}{2}, f)$). In this situation, a remarkable result of Waldspurger [173] (see also [105] for an explicit version) relates these central L -values to the squares of Fourier coefficients of a half-integer weight modular form associated to f (namely its Shimura correspondent). As a byproduct, we know that $L(\frac{1}{2}, f)$ is nonnegative, and it is conjectured never to be zero. The Keating–Snaith conjectures predict that for large $k \equiv 0 \pmod{4}$, the values $\log L(\frac{1}{2}, f)$ are distributed like a real Gaussian with mean $-\frac{1}{2} \log \log k$ and variance $\log \log k$. Since the mean is negative, the values $L(\frac{1}{2}, f)$ in this family are typically small, of size $(\log k)^{-\frac{1}{2}+o(1)}$.

A related example is to fix a newform f , and to consider the family of quadratic twists of f . Once again normalizing so that the functional equation connects s and $1 - s$, our interest is in the central values $L(\frac{1}{2}, f \times \chi_d)$, where d runs over fundamental discriminants $|d| \leq X$ with d coprime to the level of f for simplicity. As in the previous example, half of these twists will have a functional equation with $-$ sign (where the central L -value vanishes), and we restrict attention to the complementary case when the sign is $+$. Again Waldspurger's formula shows that the central L -values are nonnegative, but it is possible for these values to be 0. For example, if f corresponds to an elliptic curve, then the Birch–Swinnerton-Dyer conjectures predict that the central value is zero when the quadratic twist of this elliptic curve has positive rank (and the rank must also be even when the sign of the functional equation is $+$). However, one expects that typically $L(\frac{1}{2}, f \times \chi_d) \neq 0$, and the Keating–Snaith conjectures predict further that the distribution of $\log L(\frac{1}{2}, f \times \chi_d)$ (where $|d| \leq X$ is coprime to the level of f and the twist has $+$ sign of the functional equation) is that of a real Gaussian with mean $-\frac{1}{2} \log \log X$ and variance $\log \log X$.

The classification of families into unitary, symplectic, and orthogonal is based on the philosophy of Katz and Sarnak [98] which connects (conjecturally) the distribution of low lying zeros in these families to the distribution of eigenvalues near 1 of large random matrices chosen from the corresponding classical groups—we shall discuss these links to random matrix theory later. We now give heuristic reasons to explain the three different Keating–Snaith conjectures, point out the obstructions to making these precise, and describe the partial progress that has been made.

Recall that in (2.4) we considered approximations to $\log \zeta(\frac{1}{2} + it)$ by Dirichlet series over prime powers of a flexible length x . In (2.6) we saw that for $\zeta(\frac{1}{2} + it)$ the contribution of prime powers p^k with $k \geq 3$ is bounded, and the contribution from prime squares is also typically small. Finally, the distribution of the sums over primes could be understood

by computing moments. We now consider analogues of this calculation for the families discussed above, and the key difference in the orthogonal and symplectic cases will arise in the contribution of squares of primes.

Let us first look at the unitary family of Dirichlet characters (mod q) with q a large prime. Suppose that we have an approximation of the form

$$\log L\left(\frac{1}{2}, \chi\right) \approx \sum_{n \leq x} \frac{\Lambda(n)}{\sqrt{n} \log n} \chi(n) = \sum_{p \leq x} \frac{\chi(p)}{\sqrt{p}} + \frac{1}{2} \sum_{p \leq \sqrt{x}} \frac{\chi(p)^2}{p} + O(1). \quad (3.1)$$

A typical character $\chi \pmod{q}$ is not quadratic; χ^2 is then a nonprincipal character and the sum over prime squares above is typically of bounded size, behaving a lot like $\log L(1, \chi^2)$. We are left with the sum over primes, and if x is a small power of q , then we can understand the moments of this sum (much as in (2.7)) using the orthogonality relation for the characters (mod q) (in place of (2.8)). This gives a heuristic justification for the Keating–Snaith conjectures in this family, and the missing ingredient is the very first step which may fail badly, for example, if $L(\frac{1}{2}, \chi) = 0$ for many characters $\chi \pmod{q}$.

Consider next the symplectic example of quadratic Dirichlet L -functions $L(s, \chi_d)$ with d ranging over fundamental discriminants $|d| \leq X$. Suppose that an approximation as in (3.1) holds. Since χ_d is a quadratic character, note that the squares of primes in (3.1) have $\chi_d(p)^2 = 1$ (ignoring the primes p that divide d), and so these terms contribute

$$\frac{1}{2} \sum_{p \leq \sqrt{x}} \frac{1}{p} \sim \frac{1}{2} \log \log x \sim \frac{1}{2} \log \log X,$$

if x is a small power of X . Thus the prime square terms account for the mean of $\log L(\frac{1}{2}, \chi_d)$ being $\sim \frac{1}{2} \log \log X$ in the Keating–Snaith conjectures. If x is a small power of X , we may compute the moments of the sum over primes:

$$\sum_{|d| \leq X} \left(\sum_{p \leq x} \frac{\chi_d(p)}{\sqrt{p}} \right)^k = \sum_{p_1, \dots, p_k \leq x} \frac{1}{\sqrt{p_1 \cdots p_k}} \sum_{|d| \leq X} \left(\frac{d}{p_1 \cdots p_k} \right).$$

The inner sum over d may be viewed as a character sum (mod $p_1 \cdots p_k$). This character is principal if $p_1 \cdots p_k$ is a square, and we get a main term here, while if $p_1 \cdots p_k$ is not a square we may expect the character sum to cancel out (and this can be justified if x^k is small in comparison to X). The product $p_1 \cdots p_k$ can be a square only if k is even, and the primes p_1, \dots, p_k can be paired off into $k/2$ equal pairs. With a little calculation, this shows that the moments of the sum over primes match the moments of a real Gaussian with mean 0 and variance $\sum_{p \leq x} 1/p \sim \log \log X$. Taking into account the shift in mean arising from the prime square terms, this gives a heuristic justification for the Keating–Snaith conjecture.

Finally, let us look at the orthogonal family of quadratic twists of a newform in the case where the sign of the functional equation is $+$. The L -function $L(s, f \times \chi_d)$ is given by an Euler product, the p th factor of which (for a prime p not dividing the level of the form) takes the shape

$$\left(1 - \frac{\alpha_p \chi_d(p)}{p^s}\right)^{-1} \left(1 - \frac{\beta_p \chi_d(p)}{p^s}\right)^{-1},$$

where $\alpha_p \beta_p = 1$ and $\alpha_p + \beta_p = \lambda(p)$ is the normalized Hecke eigenvalue of f (normalized so that the Deligne bound gives $|\lambda(p)| \leq 2$). The logarithm of this Euler factor is

$$\sum_{k=1}^{\infty} (\alpha_p^k + \beta_p^k) \frac{\chi_d(p^k)}{k p^{ks}},$$

and in analogy with (2.4), (2.6), (3.1), we may hope to approximate $\log L(\frac{1}{2}, f \times \chi_d)$ by

$$\begin{aligned} & \sum_{p \leq x} \frac{(\alpha_p + \beta_p) \chi_d(p)}{\sqrt{p}} + \frac{1}{2} \sum_{p \leq \sqrt{x}} \frac{(\alpha_p^2 + \beta_p^2) \chi_d(p)^2}{p} + O(1) \\ &= \sum_{p \leq x} \frac{\lambda(p) \chi_d(p)}{\sqrt{p}} + \frac{1}{2} \sum_{p \leq \sqrt{x}} \frac{\lambda(p)^2 - 2}{p} + O(1). \end{aligned} \tag{3.2}$$

If the discriminants d go up to size X , and x is a small power of X , then the distribution of $\sum_{p \leq x} \lambda(p) \chi_d(p) / \sqrt{p}$ may be determined by computing moments (similarly to the discussion for $L(\frac{1}{2}, \chi_d)$). The prime terms in (3.2) are distributed like a real Gaussian with mean 0 and variance

$$\sum_{p \leq x} \frac{\lambda(p)^2}{p} \sim \log \log x \sim \log \log X, \tag{3.3}$$

by Rankin–Selberg theory. In view of (3.3), the prime square terms in (3.2) contribute

$$\frac{1}{2} \sum_{p \leq \sqrt{x}} \frac{\lambda(p)^2 - 2}{p} \sim -\frac{1}{2} \log \log \sqrt{x} \sim -\frac{1}{2} \log \log X.$$

This justifies the Keating–Snaith conjecture for this family.

In all these heuristics, it is the first step of connecting $\log L(\frac{1}{2})$ to sums over prime powers that is a serious stumbling-block. Indeed, if $L(\frac{1}{2})$ is zero (or if there is a zero very close to $\frac{1}{2}$) for many elements in the family, then the Keating–Snaith conjectures would not hold. This problem does not arise in the continuous Selberg theorem, since the points t with $\frac{1}{2} + it$ very close to a zero of $\zeta(s)$ have small measure and thus do not affect the distribution.

The problem of nonvanishing of L -functions has been investigated extensively, but in general it remains a challenge to show that almost all elements in a family have nonzero central value. More often, progress towards this problem focusses on showing that a positive proportion of L -functions in a family have nonzero central value. To give a few examples: in the family of Dirichlet characters $\chi \pmod{q}$, Khan and Ngo [103] have shown that at least $\frac{3}{8}$ of these characters have $L(\frac{1}{2}, \chi) \neq 0$; in the family of quadratic Dirichlet L -functions, Soundararajan [161] shows that a proportion at least $\frac{7}{8}$ of such central values are nonzero; in the family \mathcal{H}_k of all Hecke eigenforms of weight $k \equiv 0 \pmod{4}$ for the full modular group, with $k \leq K$, Iwaniec and Sarnak [93] show that at least $\frac{1}{2}$ of the central values are non-zero, and improving this proportion (in a certain sense) would have consequences for the existence of Landau–Siegel zeros of Dirichlet L -functions.

There are some situations where, for deep algebraic reasons, one can show that most central values in a family are nonzero, but these arguments do not appear to control the size of the central value, or to deal with the possibility that there might be a zero very

near $\frac{1}{2}$. For example, Chinta [38] (following work of Rohrlich [147]) has shown that if E is an elliptic curve over \mathbb{Q} then for all but $O(q^{\frac{7}{8}})$ of the Dirichlet characters mod q (with q a large prime) one has $L(\frac{1}{2}, E \times \chi) \neq 0$. This exploits the fact (established by Shimura) that if χ^σ is a Galois conjugate of the character χ , then the vanishing of $L(\frac{1}{2}, E \times \chi)$ is equivalent to the vanishing of $L(\frac{1}{2}, E \times \chi^\sigma)$ (the algebraic parts of these L -values are Galois conjugate). Another example where algebraic techniques are very successful concerns the family of quadratic twists of an elliptic curve. In special cases, Smith [160] has shown that the (algebraic) rank of quadratic twists of elliptic curves is typically 0 (when the sign of the functional equation is $+$) or 1 (when the sign is $-$). The Birch–Swinnerton-Dyer conjecture (on which there has been a lot of progress in the cases of rank 0 and 1) would then yield Goldfeld’s conjecture that the central L -values are typically nonzero when the sign of the functional equation is $+$.

If there is a zero at or very near $\frac{1}{2}$, we might expect that its effect is to make $|L(\frac{1}{2})|$ unusually small. This observation was made in Soundararajan [164], where it was shown (assuming GRH) that $\log |L(\frac{1}{2})|$ can be bounded from above using Dirichlet series over prime powers of flexible length; we shall discuss this in more detail in Section 6. It was also observed in [164] that one could (assuming a suitable GRH) establish a one sided version of the Keating–Snaith conjecture, showing that the frequency with which $\log |L(\frac{1}{2})| \geq \text{Mean} + \lambda \sqrt{\text{Var}}$ is bounded above by the expected Gaussian $\frac{1}{\sqrt{2\pi}} \int_{\lambda}^{\infty} e^{-x^2/2} dx$; here λ is a fixed real number, the size of the family is assumed to grow. Further, if one knew that most elements in the family did not have a zero near $\frac{1}{2}$ (which, for example, would follow from the “one level density” conjectures in Katz and Sarnak [98]) then the Keating–Snaith conjecture for $\log |L(\frac{1}{2})|$ would follow.

Such one sided central limit theorems were first made precise (and unconditional) by Hough [89] in certain families of L -functions. Hough’s approach relies on knowledge of a zero density estimate putting most low lying zeros of L -functions in the family close to the critical line—an analogue of Selberg’s zero density estimate for the zeta function, mentioned in Section 2. For example, Hough’s approach would work for $\log |L(\frac{1}{2}, \chi)|$ in the unitary family of Dirichlet characters $\chi \pmod{q}$, or $\log |L(\frac{1}{2}, \chi_d)|$ in the symplectic family of quadratic Dirichlet L -functions, or in the orthogonal family $\log L(\frac{1}{2}, f)$ where f ranges over Hecke eigenforms of weight $k \equiv 0 \pmod{4}$ for the full modular group.

An alternative approach to this half of the Keating–Snaith conjectures is developed in Radziwiłł and Soundararajan [141]. This method is arguably simpler and also more widely applicable, relying only on knowledge of the first moment “+ epsilon” in the family, and avoiding zero density estimates (which require knowledge of the second moment “+ epsilon”). In [141] the method is illustrated for the family of quadratic twists of an elliptic curve (with positive sign of the functional equation), where the zero density estimates required in Hough’s approach are not known. Conjecturally, the central values in this family (when nonzero) measure (after accounting for quantities such as Tamagawa factors that are relatively easy to understand) the size of the Tate–Shafarevich group for the twisted elliptic curve. The Keating–Snaith conjecture thus predicts that the sizes of Tate–Shafarevich groups in the family of quadratic twists have a log normal distribution, with prescribed means and

variance (see Conjecture 1 in [141]). The method applies to quadratic twists of any newform (holomorphic or Maass form), and thus (by Waldspurger’s formula) also gives information on the size of Fourier coefficients of half-integer weight modular forms, establishing that these are typically a little bit smaller than the conjectured Ramanujan bounds.

Another application where this method works is to the problem of the fluctuations of a quantum observable for the modular surface. Let ψ denote a fixed even Hecke–Maass form for the full modular group, and let ϕ_j denote an even Hecke–Maass form with eigenvalue λ_j . The problem is to understand $\mu_j(\psi) = \int_{\mathrm{PSL}_2(\mathbb{Z}) \backslash \mathbb{H}} \psi(z) |\phi_j(z)|^2 \frac{dx dy}{y^2}$ for large eigenvalue λ_j . For generic hyperbolic surfaces, it has been suggested in the physics literature [57] that similar quantum fluctuations have a Gaussian distribution. In the case of the modular group, $|\mu_j(\psi)|^2$ is related to the central value $L(\frac{1}{2}, \psi \times \phi_j \times \phi_j)$, so that the Keating–Snaith conjectures predict that it is in fact $\log |\mu_j(\psi)|$ (rather than $\mu_j(\psi)$ itself) that has a normal distribution. A one sided central limit theorem for $\log |\mu_j(\psi)|$ is obtained in Siu [159], and in particular it follows that $\lambda_j^{\frac{1}{4}} |\mu_j(\psi)| = o(1)$ for almost all eigenfunctions ϕ_j .

We have already discussed that the problem of nonvanishing of central L -values is a barrier to obtaining lower bounds towards the Keating–Snaith conjectures. There are two analytic techniques that produce a positive proportion of nonzero central values of L -functions in families: (i) the mollifier method, which is unconditional and relies on knowledge of two moments (“+ epsilon”) and (ii) understanding 1-level densities of low lying zeros, which is conditional on GRH and is not always guaranteed to yield a nonzero proportion. Both of these methods may be refined to permit an understanding of the typical size of nonzero L -values that are produced [165]. Here are two such sample results. In the family of quadratic Dirichlet L -functions, where we know [161] that $\frac{7}{8}$ of the fundamental discriminants $|d| \leq X$ satisfy $L(\frac{1}{2}, \chi_d) \neq 0$, we may establish that for any interval (α, β) of \mathbb{R} and large X ,

$$\begin{aligned} & \#\left\{ |d| \leq X : \frac{\log |L(\frac{1}{2}, \chi_d)| - \frac{1}{2} \log \log X}{\sqrt{\log \log X}} \in (\alpha, \beta) \right\} \\ & \geq \left(\frac{7}{8} \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-x^2/2} dx + o(1) \right) \#\{ |d| \leq X \}. \end{aligned}$$

In the family of quadratic twists of a fixed newform f with positive sign of the functional equation, on GRH it is known that a proportion $\geq \frac{1}{4}$ of such L -values are nonzero (see [85]), and we may refine this to yield (with $\mathcal{E}(X)$ denoting the set of fundamental discriminants $|d| \leq X$ with the quadratic twist of f has positive sign)

$$\begin{aligned} & \#\left\{ d \in \mathcal{E}(X) : \frac{\log L(\frac{1}{2}, f \times \chi_d) + \frac{1}{2} \log \log X}{\sqrt{\log \log X}} \in (\alpha, \beta) \right\} \\ & \geq \left(\frac{1}{4} \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-x^2/2} dx + o(1) \right) |\mathcal{E}(X)|. \end{aligned}$$

Finally, we mention recent work of Bui et al. [26] which considers a variant of the Keating–Snaith conjectures when L -values are counted with suitable weights (which depend on “mollified L -values”).

4. MOMENTS OF ZETA AND L-FUNCTIONS

A classical problem, going back to Hardy and Littlewood, asks for an understanding of the moments of $\zeta(\frac{1}{2} + it)$,

$$M_k(T) = \int_0^T \left| \zeta\left(\frac{1}{2} + it\right) \right|^{2k} dt, \quad (4.1)$$

where k is a natural number. Hardy and Littlewood established that $M_1(T) \sim T \log T$ (see [170]), and this was later refined by Ingham who showed that

$$M_1(T) = \int_0^T \left| \zeta\left(\frac{1}{2} + it\right) \right|^2 dt = T \log \frac{T}{2\pi} + (2\gamma - 1)T + E(T), \quad (4.2)$$

with $E(T) = O(T^{\frac{1}{2}} \log T)$, with a further refinement in Balasubramanian [13] yielding $E(T) = O(T^{\frac{1}{3} + \varepsilon})$. Ingham also established an asymptotic for the fourth moment: $M_2(T) \sim \frac{1}{2\pi^2} T(\log T)^4$, which was refined by Heath-Brown [81] to

$$M_2(T) = \int_0^T \left| \zeta\left(\frac{1}{2} + it\right) \right|^4 dt = TP_4(\log T) + O(T^{\frac{7}{8} + \varepsilon}), \quad (4.3)$$

for a polynomial P_4 of degree 4 with leading coefficient $1/(2\pi^2)$.

Despite much effort, these remain the only two cases in which an asymptotic formula for $M_k(T)$ is known. To explain why, we recall that Hardy and Littlewood gave an “approximate functional equation” (in fact, Riemann’s unpublished notes had a more precise version, known now as the Riemann–Siegel formula)

$$\zeta\left(\frac{1}{2} + it\right) \approx \sum_{n \leq \sqrt{|t|}/2\pi} \frac{1}{n^{\frac{1}{2} + it}} + e^{i\vartheta(t)} \sum_{n \leq \sqrt{|t|}/2\pi} \frac{1}{n^{\frac{1}{2} - it}}, \quad (4.4)$$

where $e^{i\vartheta(t)} = \pi^{it/2} \Gamma((\frac{1}{2} - it)/2) / (\pi^{-it/2} \Gamma((\frac{1}{2} + it)/2))$ is the ratio of Γ -factors in the functional equation for $\zeta(s)$. Thus $\zeta(\frac{1}{2} + it)$ can be approximated by two Dirichlet polynomials of length about $\sqrt{|t|}$. We saw in (2.8) that the mean square of Dirichlet polynomials of length up to T could be evaluated, with the diagonal terms making the dominant contribution. This permits the evaluation of the second moment (4.2) with Ingham’s bound on the remainder term $E(T)$ (we have not discussed the cross terms that arise in squaring (4.4) but these turn out to be negligible). Similarly, we can approximate $\zeta(\frac{1}{2} + it)^2$ by two Dirichlet polynomials of length about $|t|/2\pi$, and this leads to Ingham’s asymptotic for $M_2(T)$, although the more precise form in (4.3) requires further ideas. When $k \geq 3$, the complexity of $\zeta(\frac{1}{2} + it)^k$ becomes too great; to approximate it, we require Dirichlet polynomials of length about $|t|^{k/2}$ (which is now larger than $|t|$), and (2.8) is no longer sufficient to estimate the mean-square of such long Dirichlet polynomials. Let $d_k(n)$ denote the k -divisor function, which arises as the Dirichlet series coefficients of $\zeta(s)^k = \sum_{n=1}^{\infty} d_k(n)/n^s$ (valid for $\text{Re}(s) > 1$). One new problem that arises when considering higher moments involves the correlations

$$\sum_{n \leq x} d_k(n) d_k(n + h). \quad (4.5)$$

One would like asymptotics for such quantities, uniformly in a range for h , and while this problem has been solved for $k = 2$ (and underlies the precise asymptotics given in (4.3)), when $k = 3$ or larger, asymptotics for the quantity in (4.5) remain unknown (even in the case $h = 1$).

Indeed, until the late 1990s it was not even clear what the conjectural asymptotics for $M_k(T)$ should be. However, in the last 25 years, much progress has been made in understanding conjecturally the nature of these moments, obtaining lower bounds of the correct conjectured value (for all positive real k), and obtaining complementary upper bounds of the correct order conditional on the Riemann Hypothesis. Similar progress has been made for moments in a number of different families of L -functions. We shall discuss these conjectures and the progress towards them in the following sections, but first give some motivation for considering such moments.

One motivation for considering the moments of $\zeta(s)$ is that they capture information about the large values of $|\zeta(\frac{1}{2} + it)|$. The Lindelöf hypothesis that $|\zeta(\frac{1}{2} + it)| \ll_\epsilon (1 + |t|)^\epsilon$ (which is a consequence of RH) is equivalent to the bound $M_k(T) \ll_{k,\epsilon} T^{1+\epsilon}$ for all $k \in \mathbb{N}$. From the approximate functional equation (4.4) it follows that $|\zeta(\frac{1}{2} + it)| \ll (1 + |t|)^{\frac{1}{4}}$, a bound known as the *convexity bound*. Going beyond the convexity bound involves showing cancelation in the exponential sums in (4.4), and has remained an active problem from its initiation by Weyl, and Hardy and Littlewood who showed early on that $|\zeta(\frac{1}{2} + it)| \ll (1 + |t|)^{\frac{1}{6}+\epsilon}$ (see [170] and the best current exponent may be found in [25]). Sharp moment estimates encode Lindelöf bounds on average, and in some cases can also yield pointwise subconvexity estimates. For example, we note that Ingham’s bound $E(T) \ll T^{\frac{1}{2}} \log T$ (for the error term in the second moment (4.2)) implies that $\int_T^{T+1} |\zeta(\frac{1}{2} + it)|^2 dt \ll T^{\frac{1}{2}} \log T$ from which the convexity bound $|\zeta(\frac{1}{2} + it)| \ll |t|^{\frac{1}{4}+\epsilon}$ may be deduced. Similarly Balasubramanian’s improved estimate for $E(T)$ implies the Hardy–Littlewood–Weyl subconvexity bound $|\zeta(\frac{1}{2} + it)| \ll (1 + |t|)^{\frac{1}{6}+\epsilon}$. Similarly, Ingham’s asymptotic for the fourth moment yields the convexity bound, while the more precise result (4.3) of Heath-Brown gives a subconvexity bound for $\zeta(s)$. As a third example of bounds for moments that encode good pointwise bounds, we mention Heath-Brown’s [80] estimate for the 12th moment

$$\int_0^T \left| \zeta\left(\frac{1}{2} + it\right) \right|^{12} dt \ll T^{2+\epsilon},$$

which again contains the bound $|\zeta(\frac{1}{2} + it)| \ll |t|^{\frac{1}{6}+\epsilon}$.

Ingham’s work on the fourth moment of $\zeta(\frac{1}{2} + it)$ is also crucial in establishing “zero density estimates” which are bounds for the number of potential exceptions to the Riemann hypothesis. These have arithmetic applications, for example, playing a key role in showing that the prime number theorem holds in short intervals: $\pi(x + h) - \pi(x) \sim h / \log x$ provided $x^{\frac{7}{12}+\epsilon} < h \leq x$. A sharp bound for the sixth moment (for instance) would lead to improvements in zero-density results and in the application to the prime number theorem. We refer to Chapter 10 of [92] for a discussion of these themes.

There is a large body of work studying analogous problems for moments of central values in families of L -functions, and in many cases asymptotics for small moments are

known. We give a few examples here, and discuss some more in Section 6. Two motivations for studying such questions are (i) the problem of showing that many central values are nonzero, which can be attacked analytically if we know two moments with a little room to spare (we gave a few examples of such results in the previous section), and (ii) obtaining subconvexity bounds for L -functions (there is a vast literature here, and we content ourselves to pointing to earlier surveys on this topic [63, 94, 119, 128] and to Nelson [130, 132] for very recent progress).

The unitary family of Dirichlet characters (mod q) (for a large prime q) is closest in spirit to $\zeta(\frac{1}{2} + it)$, but there are still some differences. It is easy to evaluate the second moment $\sum_{\chi \pmod{q}}^* |L(\frac{1}{2}, \chi)|^2$ (where the $*$ indicates that the sum is restricted to primitive characters) and, in analogy with (4.1), this is, $\sim q \log q$. The fourth moment can also be evaluated, and in analogy with Ingham's result, Heath-Brown [83] established that $\sum_{\chi \pmod{q}}^* |L(\frac{1}{2}, \chi)|^4 \sim \frac{1}{2\pi^2} q (\log q)^4$. However, an analogue of (4.3), obtaining lower order terms in the asymptotic formula with a "power saving" in the error term, proved substantially more difficult, and was first achieved in the work of Young [174]. Higher moments remain unknown, although one can make progress by averaging over q (see Section 6). Another natural unitary family is the twists of a fixed Hecke eigenform by Dirichlet characters (mod q). The complexity of the second moment in this family is naively comparable to the fourth moment of Dirichlet L -functions, but there are further formidable difficulties. An extensive discussion of this problem, with variants and applications, may be found in the memoir of Blomer et al. [18].

In the symplectic case of quadratic Dirichlet L -functions, the first three moments $\sum_{|d| \leq x} L(\frac{1}{2}, \chi_d)^k$ are known (see [96, 161], and for interesting work on the error term in the cubic moment see [55, 175]), and the asymptotics in these cases ($k = 1, 2, 3$) take the shape of $x P_k(\log x)$ for a polynomial P_k of degree $k(k+1)/2$. We shall explain in the next section how this ties in with the Keating–Snaith conjecture for the distribution of $\log L(\frac{1}{2}, \chi_d)$. The techniques behind evaluating these moments also establish that a proportion at least $\frac{7}{8}$ of these values are nonzero (see [161]).

As an example of an orthogonal family, consider the set \mathcal{H}_k of Hecke eigenforms for the full modular group with large weight $k \equiv 0 \pmod{4}$. Here the moments $\sum_{f \in \mathcal{H}_k} L(\frac{1}{2}, f)^r$ may be evaluated for $r = 1, 2$, and if an extra averaging over $K \leq k \leq 2K$ is taken, then in the cases $r = 3$ and 4 also (this follows from the techniques in [93]). The asymptotic answers here are of the shape $|\mathcal{H}_k| P_r(\log k)$ for a polynomial P_r of degree $r(r-1)/2$. A sharp bound for the third moment (without an average in k) is established in Peng [136]; this permits a subconvexity bound $L(\frac{1}{2}, f) \ll k^{\frac{1}{3} + \varepsilon}$, which is comparable in strength to the Hardy–Littlewood–Weyl subconvexity bound for $\zeta(\frac{1}{2} + it)$. An analogous cubic moment (with such a subconvexity bound) has been studied in the case of Maass forms by Ivic [91]; interestingly, these cubic moments are also connected by a beautiful formula of Motohashi [127] to the fourth moment of $\zeta(\frac{1}{2} + it)$. Substantial progress has been made towards obtaining estimates for the fifth moment for modular forms (in the weight and level aspects) and in finding "reciprocity relations" among the fourth moments in different families; see [19, 102, 104].

We mention one more striking example: the work of Conrey–Iwaniec [47] gives sharp estimates for the cubic moment of $L(\frac{1}{2}, f \times \chi)$ where f runs over modular forms of level dividing q (an odd square-free integer) and χ denotes the quadratic character (mod q). This gives a good Weyl-type subconvexity bound for such L -values, and an analogous calculation for Maass forms gives Weyl-type subconvexity bounds for quadratic Dirichlet L -functions (improving upon classical results of Burgess). Further spectacular work in this direction may be found in Petrow and Young [137], and Nelson [131].

5. CONJECTURES FOR THE ASYMPTOTICS OF MOMENTS

Before discussing in detail the moments on the critical line, let us consider the moments on the line $\text{Re}(s) = \sigma > \frac{1}{2}$. We mentioned in Section 2 that $\zeta(\sigma + it)$ is distributed like the random object $\zeta(\sigma, \mathbb{X})$ defined in (1.4). We may therefore expect that for any $k \in \mathbb{N}$ and as $T \rightarrow \infty$,

$$\frac{1}{T} \int_0^T |\zeta(\sigma + it)|^{2k} dt \sim \mathbb{E}[|\zeta(\sigma, \mathbb{X})|^{2k}] = \sum_{n=1}^{\infty} \frac{d_k(n)^2}{n^{2\sigma}}, \quad (5.1)$$

since $\zeta(\sigma, \mathbb{X})^k = \sum_{n=1}^{\infty} d_k(n)\mathbb{X}(n)/n^\sigma$ with $d_k(n)$ being the k -divisor function (the series converges almost surely for $\sigma > \frac{1}{2}$). When $\sigma > 1$, it is clear that (5.1) holds (indeed, for any real number k), since the values $|\zeta(\sigma + it)|$ lie in a compact subset of $(0, \infty)$ and the distributions match. The case $\sigma = 1$ is more delicate, but with a little more effort one can justify (5.1) here as well. Moving now into the critical strip, there is no known value of $\frac{1}{2} < \sigma < 1$ where the asymptotic (5.1) is known to hold for all $k \in \mathbb{N}$. Indeed, such a result would imply that $|\zeta(\sigma + it)| \ll |t|^\varepsilon$, which remains unknown for any $\frac{1}{2} < \sigma < 1$. However, if one is willing to assume RH, then it is possible to approximate $\zeta(\sigma + it)^k$ by short Dirichlet polynomials, and then (5.1) follows for all real numbers k .

Returning to moments on the critical line, as mentioned previously, asymptotic formulae for $M_k(T)$ are known only in the cases $k = 1$ and 2. But, using (5.1) as a guide, we may guess the order of magnitude of $M_k(T)$. The series on the right-hand side of (5.1) diverges when $\sigma = \frac{1}{2}$, but we might consider truncating that sum around size T . It is easy to show that for any real number k ,

$$\sum_{n \leq T} \frac{d_k(n)^2}{n} \sim \frac{a_k}{\Gamma(k^2 + 1)} (\log T)^{k^2}, \quad (5.2)$$

where

$$a_k = \prod_p \left(1 - \frac{1}{p}\right)^{k^2} \left(\sum_{a=0}^{\infty} \frac{d_k(p^a)^2}{p^a}\right). \quad (5.3)$$

Thus one might guess that for all positive real numbers k , $M_k(T) \sim C_k T (\log T)^{k^2}$ for some constant C_k . Conrey and Ghosh suggested that it might be instructive to write C_k as $g_k a_k / \Gamma(k^2 + 1)$, and expected that the unknown factor g_k might have nice properties

(for example, that g_k would be a natural number when k is a natural number). The Hardy–Littlewood asymptotics for the second moment (see (4.2)) is in keeping with this conjecture, and gives $g_1 = 1$. Similarly, Ingham’s result on the fourth moment (see (4.3)) yields $g_2 = 2$.

Another way to guess at the order of magnitude for $M_k(T)$ arises from extrapolations of Selberg’s central limit theorem. If X is a random variable that is normally distributed with mean μ and variance σ^2 , then for any real number t we have

$$\begin{aligned} \mathbb{E}[e^{tX}] &= \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} \exp\left(tu - \frac{(u - \mu)^2}{2\sigma^2}\right) du \\ &= e^{t\mu + t^2\sigma^2/2} \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} \exp\left(-\frac{(u - \mu - t\sigma^2)^2}{2\sigma^2}\right) du = e^{t\mu + t^2\sigma^2/2}. \end{aligned} \quad (5.4)$$

Further, the dominant contribution above comes from values of X that are about $\mu + t\sigma^2 + O(\sigma)$. Selberg’s theorem tells us that $\log |\zeta(\frac{1}{2} + it)|$ is distributed like a Gaussian with mean 0 and variance $\sim \frac{1}{2} \log \log T$. The calculation in (5.4) therefore suggests that

$$M_k(T) = \int_0^T \exp\left(2k \log \left| \zeta\left(\frac{1}{2} + it\right) \right| \right) dt = T \exp\left((2k)^2 \frac{\frac{1}{2} \log \log T}{2}\right) = T(\log T)^{k^2}.$$

Moreover, the dominant contribution to the $(2k)$ th moment should arise from values of $\zeta(\frac{1}{2} + it)$ of size $(\log T)^k$ and the set on which such values are attained has measure about $T/(\log T)^{k^2}$. We should clarify that Selberg’s theorem is concerned with typical values of $\log |\zeta(\frac{1}{2} + it)|$, which are on the scale of $\sqrt{\log \log T}$, whereas the moments $M_k(T)$ are concerned with the large deviations regime where $\log |\zeta(\frac{1}{2} + it)|$ is of size $k \log \log T$. In this regime Selberg’s result does not immediately apply, and indeed we should expect some deviations from the Gaussian, which are reflected in the constant C_k appearing in the conjecture for $M_k(T)$ (see [60, 139]). Later we shall discuss a coarse version of Selberg’s theorem in this large deviations regime [164], conditional on RH, which leads to good (conditional) upper bounds for $M_k(T)$. To give an analogy, both $\omega(n)$ (the number of distinct prime factors of n) and $\log d(n)/\log 2$ (with $d(n)$ being the divisor function) are additive functions that are distributed (if n is chosen uniformly in $[1, N]$) like a Poisson random variable with parameter $\log \log N$; this is the Erdős–Kac theorem (noting that Poisson with large parameter approximates a Gaussian). This suggests that both $\sum_{n \leq N} 2^{\omega(n)}$ and $\sum_{n \leq N} d(n)$ are on the scale of $N \log N$, but the constants involved in the asymptotics are not immediate (and are different in the two cases).

Neither of the two heuristics given above makes a prediction for the constant $C_k = a_k g_k / \Gamma(k^2 + 1)$. Indeed, until the 1990s there was no clear conjecture as to the value of g_k for any $k \neq 1, 2$. Then Conrey and Ghosh [44, 45], based on an earlier conjecture of Balasubramanian, Conrey, and Heath-Brown [15], advanced the conjecture that $g_3 = 42$. A little later Conrey and Gonek [46], based on conjectures on the asymptotics of divisor correlation sums (as in (4.5)), arrived again at the conjecture that $g_3 = 42$ (see Ng [133] for further work on making this precise), while also advancing the conjecture that $g_4 = 24024$. These methods did not extend to produce conjectures for larger k , and the problem once again seemed stuck. A great advance was made when Keating and Snaith [100],

using ideas from random matrix theory, arrived at the following remarkable conjecture for $M_k(T)$ for all positive real numbers k .

Conjecture 5.1 (Keating and Snaith). *For any positive real number k , as $T \rightarrow \infty$, we have $M_k(T) \sim g_k \frac{a_k}{\Gamma(k^2+1)} T(\log T)^{k^2}$, with*

$$g_k = \Gamma(k^2 + 1) \frac{G(1 + k)^2}{G(1 + 2k)},$$

where G is the Barnes G -function. In particular, if $k \in \mathbb{N}$ then

$$g_k = (k^2)! \prod_{j=0}^{k-1} \frac{j!}{(k + j)!},$$

so that $g_1 = 1$, $g_2 = 2$, $g_3 = 42$, and $g_4 = 24024$.

We recall that the Barnes G -function is an entire function of order 2 which satisfies the functional equation $G(z + 1) = \Gamma(z)G(z)$ with the normalization $G(1) = 1$. Thus for a natural number n , one has $G(n) = \prod_{j=0}^{n-2} j!$.

The key insight of Keating and Snaith was to quantify and develop in the context of value distribution problems a conjectural connection between the distribution of zeros of the Riemann zeta function and the distribution of eigenvalues of large random matrices. Nearly 50 years back, Montgomery [124] initiated a study of the spacings between the ordinates of zeros of the Riemann zeta function, and a chance conversation with Dyson revealed that his partial results on this question matched corresponding statistics in the study of spacings between eigenvalues of large random matrices. Assuming RH for clarity, let $\gamma_1 \leq \gamma_2 \leq \dots$ denote the sequence of nonnegative ordinates of zeros of $\zeta(s)$ (written with multiplicity), so that from (2.3) it follows that $\gamma_n \sim 2\pi n / \log n$. The question then is to determine the distribution (as $n \rightarrow \infty$) of $(\gamma_{n+1} - \gamma_n)(\log \gamma_n) / (2\pi)$, which has been normalized to have mean spacing 1. For example, with what frequency does this normalized spacing lie in a given interval $(\alpha, \beta) \subset (0, \infty)$? One way to express the (amazing!) conjectured answer is as follows. Consider a random element g drawn from the unitary group $U(N)$ with respect to the Haar measure dg (normalized so that $U(N)$ has volume 1). Each such g has eigenvalues $e^{i\theta_1}, \dots, e^{i\theta_N}$ with the angles ordered $0 \leq \theta_1 \leq \theta_2 \leq \dots \leq \theta_N < 2\pi$, and consider the spacings $(\theta_{n+1} - \theta_n)N / (2\pi)$ (normalized to have average approximately 1). Average this spacing distribution over the whole group $U(N)$, and finally let $N \rightarrow \infty$. For example, we could count the frequency with which $(\theta_{n+1} - \theta_n)N / (2\pi)$ lies in (α, β) , average that frequency over $U(N)$, and take the limiting frequency as $N \rightarrow \infty$. The model that we have described is known as the *Circular Unitary Ensemble* (CUE), and the same distribution for nearest neighbor spacings arises in other models of random matrices such as the Gaussian Unitary Ensemble (GUE).

Theoretical support for this link between zeros of $\zeta(s)$ and random matrix theory arose first with Montgomery's calculation of the *pair correlation* of zeros (in certain ranges) mentioned earlier, and this was generalized to general n -level correlations in the work of Rudnick and Sarnak [148]. Experimental support for this link comes from extensive computations

of Odlyzko [134] who considered the spacing distribution of about 175 million zeros around the 10^{20} th zero (which occurs at height around 1.5×10^{19}), and found an astonishingly close match between the empirical data and the predicted answer. Yet, Odlyzko's data found that the numerical data did not match closely some other statistics for $\zeta(s)$ such as Selberg's theorem on $\log \zeta(\frac{1}{2} + it)$. One might attribute such deviations to the slow growth of the variance $\log \log T$, which even at height 10^{19} is only about 3.7, but Keating and Snaith [100] suggested a much more insightful explanation. They posited that properties of $\zeta(\frac{1}{2} + it)$ for t around a specific height T may be modeled by analogous objects for random matrices of a specific size N , refining the expectation that the large T and large N limits coincide. The relation between N and T is suggested by the average spacing between the zeros at height T , which is about $(2\pi)/\log(T/2\pi)$ by (2.3), and the average spacing between eigenvalues, which is about $(2\pi)/N$. Setting these equal, we arrive at the correspondence $N \approx \log(T/2\pi)$. The analogue of the zeta function, which is determined by its zeros, is the characteristic polynomial of a random matrix, which is determined in a similar fashion by its eigenvalues. Keating and Snaith determined the distribution of $\log \det(I - ge^{-i\theta})$ for a random matrix $g \in U(N)$, and found that in the large N limit this tends to a complex Gaussian with mean 0 and variance $\log N$ (analogously to Selberg's theorem), but there are lower order terms that are significant for finite N . The range of Odlyzko's computations, $T \approx 1.5 \times 10^{19}$, corresponds to matrices of size $N = 42$, and Keating and Snaith found an excellent fit between Odlyzko's numerical data for $\log \zeta(\frac{1}{2} + it)$ and the distribution of $\log \det(I - ge^{-i\theta})$ for random $g \in U(42)$ (see Figure 1 in [100]).

Returning to the moments, one might now hope to understand the asymptotic behavior of $M_k(T)$ by computing the analogous moments in the context of $U(N)$: namely

$$\int_{g \in U(N)} \frac{1}{2\pi} \int_0^{2\pi} |\det(I - ge^{-i\theta})|^{2k} d\theta dg = \int_{g \in U(N)} |\det(I - g)|^{2k} dg. \quad (5.5)$$

By the Weyl integration formula expressing the measure dg in terms of the eigenvalues of g , this equals the multiple integral

$$\frac{1}{(2\pi)^N N!} \int_0^{2\pi} \cdots \int_0^{2\pi} \left| \prod_{j=1}^N (1 - e^{i\theta_j}) \right|^{2k} \prod_{1 \leq j < m \leq N} |e^{i\theta_j} - e^{i\theta_m}|^2 d\theta_1 \cdots d\theta_N. \quad (5.6)$$

It turns out that the integral in (5.6) may be evaluated exactly using a remarkable formula of Selberg [154] (see [62] for many developments arising from the Selberg integral) and it equals

$$\prod_{j=1}^N \frac{\Gamma(j)\Gamma(2k+j)}{(\Gamma(j+k))^2} \sim g_k \frac{N^{k^2}}{\Gamma(k^2+1)}, \quad (5.7)$$

where g_k is as in Conjecture 5.1, and the asymptotic holds for large N . The constant g_k has an intriguing combinatorial interpretation as the number of standard Young tableaux of shape $k \times k$ (that is, the number of ways of filling a $k \times k$ array with the numbers $1, \dots, k^2$ such that the entries along each row and column are in increasing order). See [28, 53, 99] for related combinatorial discussions, and [42, 67] for discussions on the divisibility properties of g_k and related constants.

This calculation motivates Conjecture 5.1, but note that no primes appear in the random matrix model, and so the constant a_k must be “put in by hand.” Here we note that the Euler product for a_k in (5.3) arises naturally upon considering $\mathbb{E}[|(1 - X(p)/\sqrt{p})|^{-2k}] = \sum_{a=0}^{\infty} d_k(p^a)^2/p^a$ with $X(p)$ chosen uniformly from the unit circle. Thus the constant a_k may be thought of as arising from a version of the random Euler product, while the g_k term arises from the local behavior of zeros of the zeta function. For an exploration of Conjecture 5.1 along these lines, developing a hybrid Euler–Hadamard product, see the work of Gonek, Hughes, and Keating [69]. We mentioned earlier the analogy with determining asymptotics for multiplicative functions such as $k^{\omega(n)}$ or $d_k(n)$, and here the known asymptotic formulae (going back to Landau, Selberg, and Delange) factor as a “local” product over primes together with a “global” term determined by the Poisson behavior of $\omega(n)$; for an interesting discussion of this analogy, see [95].

Random matrix theory also informs our understanding of moments of central values of L -functions in families. While the distribution of spacings between zeros at large height for any given L -function is expected to follow the same law that we described for $\zeta(s)$ (see [148]), the distribution of the zeros close to the central point $\frac{1}{2}$ can vary depending on the particular family. Based on analogies with the function field case, Katz and Sarnak [98] found (conjecturally) that the distribution of zeros near $\frac{1}{2}$ in families of L -functions fell into the three categories *unitary*, *symplectic*, and *orthogonal* (which we discussed in Section 3), and that these distributions matched the distribution of the eigenvalues close to 1 of large random matrices chosen from $U(N)$, $USp(2N)$, or $SO(2N)$ (or $SO(2N + 1)$ depending on the sign of the functional equation). To give an illustration of the Katz–Sarnak conjectures, consider the family of quadratic Dirichlet L -functions $L(s, \chi_d)$ as d ranges over fundamental discriminants, which is expected to have symplectic symmetry. The density of zeros of $L(s, \chi_d)$ near $\frac{1}{2}$ is about $(\log |d|)/(2\pi)$, and a sample question is to understand the distribution of $\gamma_1 \frac{\log |d|}{2\pi}$ where γ_1 is the least nonnegative ordinate of a zero of $L(s, \chi_d)$. To describe the conjectured answer, consider a random matrix $g \in USp(2N)$ (chosen with respect to Haar measure normalized to have total volume 1) and write its eigenvalues as $e^{\pm i\theta_1}, e^{\pm i\theta_2}, \dots, e^{\pm i\theta_N}$ with $0 \leq \theta_1 \leq \dots \leq \theta_N \leq \pi$. Then as d varies over fundamental discriminants $|d| \leq X$ with $X \rightarrow \infty$, the distribution of $\gamma_1 \frac{\log |d|}{2\pi}$ is identical to the limiting distribution of $\theta_1 \frac{2N}{2\pi}$ for randomly chosen $g \in USp(2N)$ as $N \rightarrow \infty$.

Conrey and Farmer [42] proposed that the moments of central values of L -functions in families are also dictated by the symmetry type in the Katz–Sarnak conjectures. In particular, the analogue of the factor g_k should depend only on the symmetry type and not on the particular family, whereas the analogue of the factor a_k will be sensitive to the particular family (in a straightforward way). This was developed further by Keating and Snaith [101], who modeled properties of the central L -values by the characteristic polynomial $\det(I - ge^{-i\theta})$ evaluated at $\theta = 0$, with the size parameter N of the random matrix ensemble chosen to match with the density of zeros in the family. Indeed, it is a consideration of the behavior of $\log \det(I - g)$ in $USp(2N)$ or $SO(2N)$ that informed their conjectures for the analogues of Selberg’s theorem in symplectic and orthogonal families (discussed in Section 3).

Just as extrapolating Selberg’s theorem allows us to guess the order of magnitude of moments of $\zeta(s)$, the Keating–Snaith lognormality conjectures, together with the calculation in (5.4), give an understanding of the order of magnitude of moments in families. For example, in the symplectic example of moments of $L(\frac{1}{2}, \chi_d)$ with $|d| \leq X$, since $\log L(\frac{1}{2}, \chi_d)$ is conjectured to be normal with mean $\sim \frac{1}{2} \log \log X$ and variance $\sim \log \log X$, the calculation in (5.4) suggests that $\sum_{|d| \leq X} L(\frac{1}{2}, \chi_d)^k$ is of size $X(\log X)^{\frac{k(k+1)}{2}}$. Similarly, in the orthogonal case of Hecke eigenforms $f \in \mathcal{H}_k$, since $\log L(\frac{1}{2}, f)$ is expected to be normal with mean $\sim -\frac{1}{2} \log \log k$ and variance $\sim \log \log k$, the moments $\sum_{f \in \mathcal{H}_k} L(\frac{1}{2}, f)^r$ may be expected to be of order $k(\log k)^{\frac{r(r-1)}{2}}$.

Further, by considering moments of $\det(I - g)$ in the appropriate matrix group, Keating and Snaith [101] formulated analogues of Conjecture 5.1 in families of L -functions. For example, in the family of quadratic Dirichlet L -functions $L(s, \chi_d)$, the analogue of the constant g_k is predicted by considering

$$\begin{aligned} \int_{g \in \mathrm{USp}(2N)} \det(I - g)^k dg &= 2^{2Nk} \prod_{j=1}^N \frac{\Gamma(1 + N + j)\Gamma(1/2 + k + j)}{\Gamma(1/2 + j)\Gamma(1 + k + N + j)} \\ &\sim f_k \frac{N^{k(k+1)/2}}{\Gamma(k(k+1)/2 + 1)}. \end{aligned}$$

This calculation again reduces to the Selberg integral, and the constant f_k may be expressed in terms of the Barnes G -function. If k is a natural number then f_k takes the pleasant form $(k(k+1)/2)! / \prod_{j=1}^k (2j-1)!!$. After incorporating an analogue of the constant a_k in (5.3), which here is (with $\mathbb{X}(p)$ denoting the random variables modeling quadratic characters discussed in Section 1)

$$\begin{aligned} &\prod_p \left(1 - \frac{1}{p}\right)^{\frac{k(k+1)}{2}} \mathbb{E} \left[\left(1 - \frac{\mathbb{X}(p)}{\sqrt{p}}\right)^{-k} \right] \\ &= \prod_p \left(1 - \frac{1}{p}\right)^{\frac{k(k+1)}{2}} \left(\frac{p}{2(p+1)} \left(\left(1 + \frac{1}{\sqrt{p}}\right)^{-k} + \left(1 - \frac{1}{\sqrt{p}}\right)^{-k} \right) + \frac{1}{p+1} \right), \end{aligned}$$

we arrive at a conjecture for the moments of $L(\frac{1}{2}, \chi_d)$, which matches the known asymptotics for the first three moments.

The Keating–Snaith conjectures identify the leading order term in the asymptotics for moments, but there will be lower order terms (just a logarithm smaller) which are not identified. We may see this already in the asymptotics for the second and fourth moments of $\zeta(\frac{1}{2} + it)$ (see (4.2) and (4.3)), and other examples in families given in Section 4. Identifying such lower order terms is of interest because the leading order constant in Conjecture 5.1, $a_k g_k / \Gamma(k^2 + 1)$ tends rapidly to zero as k grows, and so for the ranges of T in which numerical investigations may be carried out, the lower order terms may dominate the eventual main term. When k is a positive integer, Conrey et al. [43] conjectured that $M_k(T) = \int_0^T P_k(\log t / 2\pi) dt + O(T^{1-\delta})$ (for some $\delta > 0$, and perhaps even any $\delta < \frac{1}{2}$ is permissible) for a polynomial P_k of degree k^2 with leading coefficient $a_k g_k / (k^2!)$, and they gave a “recipe” for determining all the coefficients of P_k . Their recipe predicts the full main

term for integral moments in many families of L -functions, but it remains open to give an asymptotic expansion when k is not an integer. The paper [43] also gives numerical evidence towards the full moment conjecture, and further data may be found in [86]. A related approach via *multiple Dirichlet series* is described in the work of Diaconu, Goldfeld, and Hoffstein [54] who develop conjectures for the integral moments of quadratic Dirichlet L -functions (which are in agreement with [43]).

We give a brief illustration of the recipe from [43] in the unitary family of Dirichlet L -functions $\chi \pmod{q}$ with q a large prime. For simplicity, we consider only even characters (thus $\chi(-1) = 1$), where the functional equation reads $\Lambda(s, \chi) = (q/\pi)^{s/2} \Gamma(s/2) L(s, \chi) = \varepsilon_\chi \Lambda(1-s, \bar{\chi})$ with ε_χ satisfying $|\varepsilon_\chi| = 1$ and $\varepsilon_\chi \varepsilon_{\bar{\chi}} = 1$. Let $\underline{\alpha} = (\alpha_1, \dots, \alpha_k)$, and $\underline{\beta} = (\beta_1, \dots, \beta_k)$ denote two k -tuples of complex numbers (thought of as small), and we also find it convenient to write $\alpha_{k+j} = \beta_j$ and think of $(\underline{\alpha}, \underline{\beta})$ as the $2k$ -tuple $(\alpha_1, \dots, \alpha_{2k})$. Instead of considering $|L(\frac{1}{2}, \chi)|^{2k}$ directly, we work with

$$\Lambda(\chi; \underline{\alpha}, \underline{\beta}) := \prod_{j=1}^k \Lambda\left(\frac{1}{2} + \alpha_j, \chi\right) \Lambda\left(\frac{1}{2} - \beta_j, \bar{\chi}\right)$$

and finally let all the parameters α_j and β_j tend to zero (which would then equal $|L(\frac{1}{2}, \chi)|^{2k}$ multiplied by the constant $(q/\pi)^{k/2} \Gamma(1/4)^{2k}$). Permuting the k entries in $\underline{\alpha}$ or the k entries in $\underline{\beta}$ does not change $\Lambda(\chi; \underline{\alpha}, \underline{\beta})$. Less obviously, it turns out that $\Lambda(\chi; \underline{\alpha}, \underline{\beta})$ is invariant under any permutation of the $2k$ -entries in $(\underline{\alpha}, \underline{\beta})$; this is because any such permutation must change some ℓ of the α 's to β 's and a corresponding number of β 's to α 's and 2ℓ applications of the functional equation (ℓ of them with a factor ε_χ and ℓ with a factor $\varepsilon_{\bar{\chi}}$) justify the claim. Thus any conjecture that we propose for $\sum_\chi \Lambda(\chi; \underline{\alpha}, \underline{\beta})$ must satisfy this S_{2k} symmetry.

Now if $\text{Re}(s)$ is large, expanding the L -functions into their Dirichlet series, we may write

$$\begin{aligned} \prod_{j=1}^k \Lambda(s + \alpha_j, \chi) \Lambda(s - \beta_j, \chi) &= \prod_{j=1}^k \left(\frac{q}{\pi}\right)^{s + \frac{\alpha_j - \beta_j}{2}} \Gamma\left(\frac{s + \alpha_j}{2}\right) \Gamma\left(\frac{s - \beta_j}{2}\right) \\ &\quad \times \sum_{m,n=1}^{\infty} \frac{\sigma(m; \underline{\alpha})}{m^s} \chi(m) \frac{\sigma(n; -\underline{\beta})}{n^s} \bar{\chi}(n), \end{aligned} \tag{5.8}$$

where $\sigma(m; \underline{\alpha}) = \sum_{m=m_1 \cdots m_k} m_1^{-\alpha_1} \cdots m_k^{-\alpha_k}$ and similarly $\sigma(n; -\underline{\beta}) = \sum_{n=n_1 \cdots n_k} n_1^{\beta_1} \cdots n_k^{\beta_k}$, so that if $\alpha_i = \beta_i = 0$ these would simply be the k -divisor function. We average this over all the even characters mod q (omitting the trivial character), and hypothesize that only the diagonal terms $m = n$ survive this averaging. This is of course not justified, but is similar to the first heuristic we gave in this section for the order of magnitude of moments. After a computation with Euler products, these terms give (for the sum over m, n in (5.8))

$$\sum_{n=1}^{\infty} \frac{\sigma(n; \underline{\alpha}) \sigma(n; -\underline{\beta})}{n^{2s}} = \mathcal{A}(s; \underline{\alpha}, \underline{\beta}) \prod_{j,\ell=1}^k \zeta(2s + \alpha_j - \beta_\ell), \tag{5.9}$$

where \mathcal{A} is given by an Euler product that converges absolutely in $\text{Re}(s) > \frac{1}{2} - \delta$ if α_j, β_j are small enough. This factor \mathcal{A} is similar to the a_k appearing in (5.3). Evaluating this at

$s = \frac{1}{2}$, we arrive at a candidate for the average value of $\Lambda(\chi; \underline{\alpha}, \underline{\beta})$, namely

$$\mathcal{C}(\underline{\alpha}, \underline{\beta}) = \prod_{j=1}^k \left(\frac{q}{\pi}\right)^{\frac{1+\alpha_j-\beta_j}{2}} \Gamma\left(\frac{\frac{1}{2} + \alpha_j}{2}\right) \Gamma\left(\frac{\frac{1}{2} - \beta_j}{2}\right) \mathcal{A}\left(\frac{1}{2}; \underline{\alpha}, \underline{\beta}\right) \prod_{j,\ell=1}^k \zeta(1 + \alpha_j - \beta_\ell). \tag{5.10}$$

The candidate answer $\mathcal{C}(\underline{\alpha}, \underline{\beta})$ is invariant when the entries of $\underline{\alpha}$ are permuted, or when the entries of $\underline{\beta}$ are permuted, but does not have the S_{2k} symmetry we require of being allowed to permute the $2k$ -entries of $(\underline{\alpha}, \underline{\beta})$. The beautifully simple answer proposed in [43] is to symmetrize $\mathcal{C}(\underline{\alpha}, \underline{\beta})$ by summing over all $\binom{2k}{k}$ cosets of $S_{2k}/(S_k \times S_k)$,

$$\sum_{\pi \in S_{2k}/S_k \times S_k} \mathcal{C}(\pi(\underline{\alpha}, \underline{\beta})). \tag{5.11}$$

While the expression in (5.10) has singularities whenever $\alpha_j = \beta_\ell$, the symmetrized expression in (5.11) turns out to be regular when $|\alpha_j|, |\beta_j|$ are small. Now setting $\alpha_1 = \dots = \alpha_k = \beta_1 = \dots = \beta_k = 0$, we arrive at the conjectured answer for the average of $|L(\frac{1}{2}, \chi)|^{2k}$. The leading term matches the Keating–Snaith conjecture, but now we also have the full polynomial of degree k^2 .

To end our discussion of the moment conjectures, we mention recent work of Conrey and Keating [41] which aims to give a heuristic derivation of the moment conjectures of $\zeta(s)$ from correlations of divisor functions (as in [46] for the sixth and eighth moments). It would be of interest to develop their work in other families of L -functions. Sawin [153] develops a heuristic approach based on representation theory which (conditional on some hypotheses) recovers the recipe in Conrey et al. [43] in the function field setting (with a fixed field of constants).

6. PROGRESS TOWARDS UNDERSTANDING THE MOMENTS

In Section 4 we gave a number of examples where asymptotics for low moments are known, and all of these are in agreement with the conjectures described in the previous section. A rule of thumb suggests that an asymptotic for a moment may be computed if there are more elements in the family compared to the complexity of approximating the required power of the L -value (what we have informally called the complexity can be thought of as the square-root of the *analytic conductor*, see [94]). For example, as we saw in (4.4) $\zeta(\frac{1}{2} + it)$ may be approximated by (two) Dirichlet polynomials of length about \sqrt{t} , allowing for the calculation of the second and fourth moments. This rule of thumb is only a rough guide, and can be difficult to attain. For example, the fourth moment of Dirichlet L -functions mod q (evaluated in [174]), or the mean square of twists of a modular form by Dirichlet characters mod q (see [18, 107]) may seem of comparable difficulty to the fourth moment of the zeta function, but the first two problems turn out to be substantially harder. The largest moment that may be computed by this rule of thumb recovers the convexity bound for the L -value, and so there is great interest in going beyond this range, either by shrinking suitably the family over which we average, or by adding an extra short Dirichlet polynomial to the moment.

From the viewpoint of verifying the moment conjectures (for example to check the constants 42 and 24024 appearing in the sixth and eighth moments) one might look for large families where the complexity is still small. The family of primitive Dirichlet characters $\chi \pmod{q}$ ranging over all moduli $q \leq Q$ is a good example, where the size of the family is about Q^2 whereas the complexity of such $L(\frac{1}{2}, \chi)$ is about \sqrt{Q} . This suggests the possibility of evaluating the sixth and eighth moments in this family, and indeed the large sieve gives a quick upper bound of the correct order of magnitude for these moments (see [98]). By developing an asymptotic version of the large sieve, Conrey, Iwaniec, and Soundararajan [48] obtained an asymptotic formula for

$$\sum_{q \leq Q} \sum_{\chi \pmod{q}}^b \int_{-\infty}^{\infty} \left| \Lambda\left(\frac{1}{2} + iy, \chi\right) \right|^6 dy, \quad (6.1)$$

where $\Lambda(s, \chi) = (q/\pi)^{s/2} \Gamma(s/2) L(s, \chi)$ denotes the completed L -function, and the b indicates a sum over even primitive characters χ . Here the averaging over y is a technical defect, needed for the proof, which (owing to the rapid decay of the Γ -function) may be thought of as an integral over essentially a bounded range of y . This asymptotic formula verified the predicted constant $g_3 = 42$ in this instance, and, moreover, [48] obtained a similar asymptotic formula with shifts $(\alpha_1, \alpha_2, \alpha_3)$ and $(\beta_1, \beta_2, \beta_3)$ which verified the recipe of [43] in this situation and yielded the full polynomial of degree 9 in $\log Q$ for (6.1). Chandee and Li [35] tackle the analogue of (6.1) for the eighth moment, and obtain an asymptotic formula conditional on the Generalized Riemann Hypothesis. Their work confirmed that $g_4 = 24024$ in this instance, but they could only verify the leading order term in the asymptotic and not the full polynomial of degree 16. Forthcoming work of Chandee, Li, Matomäki, and Radziwiłł (see [138] for an announcement) removes the imperfection of the average over y in (6.1) for the sixth moment while still obtaining the full asymptotic formula with power saving. They also obtain the leading order behavior of the eighth moment without invoking GRH, and without the integral in y .

The family of newforms of a fixed weight k for the group $\Gamma_1(q)$ with q a large prime offers another instance of a large family where the complexity (or analytic conductor) remains small. These correspond to newforms for $\Gamma_0(q)$ with character $\chi \pmod{q}$. This is a family of about q^2 elements, and is unitary since almost all of the characters $\chi \pmod{q}$ are not real. The complexity of the L -values is about size \sqrt{q} , and we may hope to address the sixth and eighth moments. Chandee and Li [34] give an asymptotic for the sixth moment analogous to (6.1) in this family (confirming again $g_3 = 42$), and obtain in [33] a good upper bound for the eighth moment. It would be of interest to find further examples of families where one can compute higher moments, and in particular to obtain such examples of symplectic and orthogonal families. The recent work of Nelson [139] on subconvexity for automorphic L -functions raises the hope that one might be able to compute high moments in $GL(n)$ families for suitably large n .

In addition to examples where asymptotics for moments are known, substantial progress has been made in obtaining upper and lower bounds of the conjectured order of

magnitude in a good deal of generality. Summarizing the work of many researchers, here is our knowledge of such bounds for the moments of $\zeta\left(\frac{1}{2} + it\right)$.

Theorem 6.1. *Let $k > 0$ and $T \geq e$ be real numbers. Then there are positive constants c_k and C_k such that*

$$c_k T (\log T)^{k^2} \leq \int_0^T \left| \zeta\left(\frac{1}{2} + it\right) \right|^{2k} dt \leq C_k T (\log T)^{k^2}.$$

Here the lower bound holds unconditionally for all k , while the upper bound holds unconditionally in the range $0 < k \leq 2$, and the upper bound holds assuming the truth of the Riemann Hypothesis for all $k > 2$.

We shall now discuss this result and its extensions in families of L -functions. The discussion splits naturally into three parts (i) lower bounds for moments, (ii) unconditional upper bounds for moments, and (iii) upper bounds assuming RH or GRH.

The lower bound stated in Theorem 6.1 was first established by Ramachandra [143, 144] in the case when $2k$ is a natural number. This was then extended by Heath-Brown [84] to the case when k is any positive rational number, but the constants c_k in his result depended upon the height of the rational number k , so that the method did not extend to irrational k . Further, the techniques in these works were specific to the “ t -aspect” and did not extend to moments in families of L -functions. Rudnick and Soundararajan [149, 150] developed an alternative approach, which worked in general families. For example, their method would show that $\sum_{|d| \leq X} |L(\frac{1}{2}, \chi_d)|^k \geq c_k X (\log X)^{k(k+1)/2}$ for all rational $k \geq 1$ and a suitable positive constant c_k , which again did not vary continuously with k but depended on the height of the rational number k . This was further refined by Radziwiłł and Soundararajan [140], who obtained the lower bounds in Theorem 6.1 for all real $k \geq 1$ with e^{-30k^4} being a permissible value for c_k if T is large. A further round of simplification is carried out in Heap and Soundararajan [79], which also gives the lower bound in Theorem 6.1 for real $0 < k \leq 1$.

The story for lower bounds may be encapsulated in the following broad principle. Whenever we can compute the mean value of $L(\frac{1}{2})$ multiplied by short Dirichlet polynomials in a family, we can obtain lower bounds of the right order of magnitude for the moments $|L(\frac{1}{2})|^k$ for any real $k \geq 1$. Of course, in general Hölder’s inequality will give lower bounds for higher moments in terms of smaller moments, but those would not be of the conjectured order of magnitude since the exponent of the logarithm in the moment conjectures is quadratic in k . If we can also compute the mean value of $|L(\frac{1}{2})|^2$ multiplied by short Dirichlet polynomials, then we can obtain lower bounds of the right order of magnitude for the moments $|L(\frac{1}{2})|^k$ in the range $0 < k \leq 1$ as well. It may seem puzzling why the problem for small k should require more information than for large k , but in fact this is natural. Consider letting $k \rightarrow 0^+$. Then the moments $|L(\frac{1}{2})|^k$ essentially pick up whether $L(\frac{1}{2})$ is zero or not, so that lower bounds for the small moments encode lower bounds for non-vanishing. The analytic methods for producing nonzero values of $L(\frac{1}{2})$ (the *mollifier method*) rely on knowledge of the first two moments in the family (with a little room to spare). Thus we may

establish (using the methods of either [140] or [79]) that for all real $k > 0$,

$$\sum_{\chi \pmod{q}} \left| L\left(\frac{1}{2}, \chi\right) \right|^{2k} \gg_k q (\log q)^{k^2}, \tag{6.2}$$

where q is a large prime, and that

$$\sum_{|d| \leq X} \left| L\left(\frac{1}{2}, \chi\right) \right|^k \gg_k X (\log X)^{\frac{k(k+1)}{2}}. \tag{6.3}$$

In the family of quadratic twists of a fixed Hecke eigenform f , we only have access to the first moment and not the second, and therefore we only know in the range $k \geq 1$ that

$$\sum_{|d| \leq X} L\left(\frac{1}{2}, f \times \chi_d\right)^k \gg_k X (\log X)^{\frac{k(k-1)}{2}}. \tag{6.4}$$

We now turn to the unconditional upper bounds in Theorem 6.1, which were established in the special cases $k = 1/n$ or $k = 1 + 1/n$ (for natural numbers n) by Heath-Brown [84] and Bettin, Chandee, and Radziwiłł [17], respectively. Then in Heap, Radziwiłł, and Soundararajan [78] the bound was established for all $0 < k \leq 2$, as an illustration of an upper bound principle (complementing the one for lower bounds above) enunciated by Radziwiłł and Soundararajan [141]. Whenever we can compute a moment $|L(\frac{1}{2})|^k$ (usually with k being a positive integer) together with flexibility to introduce a short Dirichlet polynomial, we can obtain upper bounds of the conjectured order of magnitude for the moments $|L(\frac{1}{2})|^r$ for all $0 < r \leq k$. Thus one can obtain complementary upper bounds in (6.2) for $k \leq 1$ (with more effort, using Young’s work [174], this could perhaps be extended to the range $k \leq 2$), matching upper bounds in (6.3) in the range $k \leq 2$ (if one knew the positivity of $L(\frac{1}{2}, \chi_d)$ this would also follow in the range $k \leq 3$ and it would be interesting to attain that range unconditionally), and for the family in (6.3) for $k \leq 1$ (this is the example carried out in [141]).

The conditional bounds in Theorem 6.1 originated from work of Soundararajan [164] who established (assuming RH) the nearly sharp bound $M_k(T) \ll_{k,\varepsilon} T(\log T)^{k^2+\varepsilon}$. This was then refined in the beautiful work of Harper [77] to its present sharp form. The method is very general and applies in any family where we are able to compute the mean values of short Dirichlet polynomials. Thus (assuming GRH in the relevant families) one can obtain upper bounds of the correct order of magnitude for all nonnegative k in the examples (6.2), (6.3), and (6.4).

The main idea behind the conditional upper bounds in Theorem 6.1 is that on RH (or GRH) one can obtain an *upper bound* for $\log |\zeta(\frac{1}{2} + it)|$ (or more generally the logarithm of central L -values) just in terms of sums over primes. This is related to the ideas behind Selberg’s central limit theorem and the one sided versions for L -values that we discussed in Sections 2 and 3. A barrier to approximating $\log |\zeta(\frac{1}{2} + it)|$ by a suitable Dirichlet polynomial is the presence of zeros near $\frac{1}{2} + it$; the crucial point is that these zeros should only make $|\zeta(\frac{1}{2} + it)|$ smaller, so that such Dirichlet polynomials could serve as an upper bound.

One way to see this is to note that RH is equivalent to the property that, with $s = \sigma + it$,

$$|\xi(s)| = |s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s)| = \prod_p \left| 1 - \frac{s}{p} \right|$$

is an increasing function of σ in $\sigma \geq \frac{1}{2}$ for any fixed t . This permits bounding $|\zeta(\frac{1}{2} + it)|$ in terms of $|\zeta(\sigma_0 + it)|$ for any $\sigma_0 > \frac{1}{2}$, and one can adapt Selberg's ideas to approximate $\log |\zeta(\sigma_0 + it)|$. In this manner, it was shown in [164] that for $T \leq t \leq 2T$ and any $2 \leq x \leq T^2$ one has, assuming RH and with $\sigma_0 = \frac{1}{2} + \frac{1}{\log x}$,

$$\log \left| \zeta \left(\frac{1}{2} + it \right) \right| \leq \operatorname{Re} \sum_{2 \leq n \leq x} \frac{\Lambda(n)}{n^{\sigma_0 + it} \log n} \frac{\log(x/n)}{\log x} + \frac{\log T}{\log x} + O \left(\frac{1}{\log x} \right). \quad (6.5)$$

Analogous bounds hold for $\log |L(\frac{1}{2})|$ if a corresponding GRH is assumed.

The usefulness of (6.5) lies in its flexibility with choosing the parameter x . If x is suitably small, then the distribution of the sum (which is essentially $\operatorname{Re} \sum_{p \leq x} 1/p^{\frac{1}{2} + it}$) in (6.5) can be understood accurately by studying its moments (as we discussed in Sections 2 and 3), but we lose some information in the $\log T / \log x$ term. Here it is also useful to split the sum over p into different ranges (say $p \leq z$ and $z < p \leq x$); for small ranges of p , more moments may be computed so that a finer understanding of the sum is possible, while for the larger ranges the slow growth of the variance (which is roughly $\sum_{z \leq p \leq x} 1/p \sim \log \frac{\log x}{\log z}$) permits a good understanding with fewer moments. In this way [164] established a coarse version of Selberg's central limit theorem in the large deviations regime, showing that in the range $\sqrt{\log \log T} \leq V = o(\log \log T \log \log \log T)$ one has

$$\operatorname{meas} \left\{ T \leq t \leq 2T : \log \left| \zeta \left(\frac{1}{2} + it \right) \right| \geq V \right\} \ll T \exp \left(- \frac{V^2}{\log \log T} (1 + o(1)) \right). \quad (6.6)$$

As we mentioned in Section 5, the $2k$ -th moment of zeta should be dominated by values of $|\zeta(\frac{1}{2} + it)|$ of size $(\log T)^k$, and the (6.6) shows that this set has measure $\ll T(\log T)^{-k^2 + o(1)}$, which yields $M_k(T) \ll T(\log T)^{k^2 + \varepsilon}$.

Harper's sharp upper bound for $M_k(T)$ builds on some of these ideas, but deals directly with the moments rather than going through the intermediary of the large deviations in Selberg's theorem (6.6). Instead there is an elaborate decomposition of the sum over primes in (6.5) into many ranges, and then the exponentials of such sums are handled by approximating these by suitable truncations of their Taylor expansion. Similar ideas were developed independently around the same time in [141] for bounding small moments unconditionally, and the recent paper [79] develops these ideas in the context of lower bounds. Thus the proofs of all three aspects of Theorem 6.1 have a unified feel, and the spirit of the proofs may be described as thinking in terms of Euler products but performing computations by replacing Euler products by short Dirichlet series obtained from their Taylor expansions. These proofs were also influenced by ideas from sieve theory. For example, in analogy with (6.5) we may note that $\omega(n)$ (the number of prime factors of n) may be bounded above by $\sum_{p|n, p \leq y} 1 + (\log n) / \log y$ for any y , and this could be used to give upper bounds for the mean-value of $d_k(n)$ (which is roughly $k^{\omega(n)}$) in short intervals.

The ideas behind obtaining conditional bounds for moments have found diverse applications. Soundararajan and Young [166] used such bounds for “shifted moments” (see also [32]) to obtain an asymptotic formula (on GRH) for the second moment of quadratic twists of an eigenform $\sum_{|d|\leq X} L(\frac{1}{2}, f \times \chi_d)^2$. This is a tantalizing problem, which falls within the purview of the rule of thumb described at the beginning of this section, but an unconditional asymptotic has so far been elusive. A similar problem is to compute the asymptotic for the fourth moment of quadratic Dirichlet L -functions $\sum_{|d|\leq X} L(\frac{1}{2}, \chi_d)^4$, and recently Shen [157] has extended the method in [166] to obtain (on GRH) such an asymptotic. Analogues of these two problems over function fields have been established in [27, 61], and since GRH is known in this setting, the corresponding results hold unconditionally.

In a very different direction, Lester and Radziwiłł [116] showed on GRH that the Fourier coefficients of half-integer weight Hecke cusp forms exhibit a positive proportion of sign changes as we range over fundamental discriminants. Among the many innovations in their beautiful proof, is an application of the ideas discussed above to obtain sharp upper bounds for the second mollified moment of quadratic twists of the Shimura correspondent of the given half-integer weight form. This realization that sharp upper bounds for the second mollified moment suffice has led to another striking result in the work of David, Florea, and Lalin [51], who show that a positive proportion of L -functions attached to cubic characters (in the function field setting) have nonzero central value. Two other recent applications include Zenz [176] to bounding the L^4 norm of Hecke eigenforms of large weight k for the full modular group, and Shubin [158] to bounding the variance of lattice points on the sphere in random small spherical caps. See [66, 121, 122] for further examples.

7. EXTREME VALUES

In Sections 2 and 3 we discussed the typical size of $|\zeta(\frac{1}{2} + it)|$ and central values of L -functions, which are governed by Selberg’s central limit theorem, and the analogous Keating–Snaith conjectures. In Sections 5 and 6 we discussed how the moment problem aims for an understanding of the *large deviations* range of values of $|\zeta(\frac{1}{2} + it)|$ (or $|L(\frac{1}{2})|$). We now discuss the maximal size of $|\zeta(\frac{1}{2} + it)|$ (for $0 \leq t \leq T$) and analogous problems in families of L -functions.

As we mentioned in Section 4, our unconditional knowledge is far from the Lindelöf hypothesis that $|\zeta(\frac{1}{2} + it)| \ll (1 + |t|)^\epsilon$, and for general L -functions already the subconvexity problem poses formidable difficulties. In 1924 Littlewood established that the Riemann Hypothesis implies the Lindelöf hypothesis in the quantitative form

$$\left| \zeta\left(\frac{1}{2} + it\right) \right| \ll \exp\left(\frac{C \log |t|}{\log \log |t|}\right) \tag{7.1}$$

for some constant C . The estimate (6.5) yields such a result, upon taking $x = (\log t)^2$ there, and bounding the sum over prime powers trivially. This strategy was optimized in [36] which showed that one may take any $C > \frac{\log 2}{2}$ in (7.1). Apart from this refinement of the constant C , no improvement has been made over Littlewood’s estimate. Corresponding results hold for

general L -functions, and explicit versions of such bounds (which are useful in computational applications) may be found in [31].

Complementing (7.1), one may ask for lower bounds on $\max_{T \leq t \leq 2T} |\zeta(\frac{1}{2} + it)|$. Recall that in Section 1 we discussed the extreme values of zeta and L -functions at the edge of the critical strip, and already there was a gap in our knowledge between the extreme values that may be exhibited and the bounds that follow from GRH (see the discussion surrounding (1.2) and (1.3)). This gap becomes much more pronounced on the critical line. By using lower bounds for integer moments of $\zeta(\frac{1}{2} + it)$, with attention to the uniformity in k , Balasubramanian and Ramachandra [16] (optimized in [14]) established that

$$\begin{aligned} \max_{T \leq t \leq 2T} \left| \zeta\left(\frac{1}{2} + it\right) \right| &\geq \max_k \left(\frac{1}{T} \int_T^{2T} \left| \zeta\left(\frac{1}{2} + it\right) \right|^{2k} dt \right)^{\frac{1}{2k}} \\ &\gg \max_{k \in \mathbb{N}} \left(\sum_{n \leq T} \frac{d_k(n)^2}{n} \right)^{\frac{1}{2k}} = \exp\left((B + o(1)) \frac{\sqrt{\log T}}{\sqrt{\log \log T}} \right), \end{aligned} \quad (7.2)$$

with $B \approx 0.53$. With the development of lower bounds for moments in families of L -functions (discussed in Section 6), such bounds also became available for central L -values. However, a different *resonance method* developed in [163] has proved to be still more efficient. The main idea in [163] is to find a Dirichlet polynomial $R(t) = \sum_n r(n)n^{-it}$ which “resonates” with $\zeta(\frac{1}{2} + it)$ and picks out its large values. This is based on computing

$$I_1 = \int_T^{2T} |R(t)|^2 dt \quad \text{and} \quad I_2 = \int_T^{2T} \zeta\left(\frac{1}{2} + it\right) |R(t)|^2 dt, \quad (7.3)$$

and noting that

$$\max_{T \leq t \leq 2T} \left| \zeta\left(\frac{1}{2} + it\right) \right| \geq \frac{|I_2|}{I_1}. \quad (7.4)$$

If the resonator Dirichlet polynomial $R(t)$ is short, in the sense that $r(n) = 0$ unless $n \leq T^{1-\varepsilon}$, then I_1 and I_2 in (7.3) may be evaluated asymptotically, and these quantities give two quadratic forms in the unknown coefficients $r(n)$. The ratio of these two quadratic forms is maximized in [163], yielding

$$\max_{T \leq t \leq 2T} \left| \zeta\left(\frac{1}{2} + it\right) \right| \geq \exp\left((1 + o(1)) \frac{\sqrt{\log T}}{\sqrt{\log \log T}} \right). \quad (7.5)$$

While this is only a little bit better than (7.2), the method also yields lower bounds on the measure of the set on which large values are attained:

$$\text{meas} \left\{ t \in [T, 2T] : \left| \zeta\left(\frac{1}{2} + it\right) \right| \geq e^V \right\} \gg \frac{T}{(\log T)^4} \exp\left(-10 \frac{V^2}{\log \frac{\log T}{8V^2 \log V}} \right), \quad (7.6)$$

uniformly for $3 \leq V \leq \frac{1}{5} \sqrt{\log T / \log \log T}$. There is some scope to improve such bounds, especially when V is of size $C \log \log T$, where one would like to match the upper bound in (6.6) which would be in keeping with Selberg’s theorem (see [79] for more precise results when $V \leq (2 - \varepsilon) \log \log T$). The estimate (7.6) shows that large values on the scale of (7.5) occur fairly often (on a set of measure $\geq T^{1-C/\log \log T}$) suggesting that still larger values

might exist. Furthermore, the resonance method extends readily to families of L -functions, and thus we may show (for example) that

$$\max_{X \leq |d| \leq 2X} L\left(\frac{1}{2}, \chi_d\right) \geq \exp\left(\left(\frac{1}{\sqrt{5}} + o(1)\right) \frac{\sqrt{\log X}}{\sqrt{\log \log X}}\right), \quad (7.7)$$

and that, for any Hecke eigenform f ,

$$\max_{X \leq |d| \leq 2X} L\left(\frac{1}{2}, f \times \chi_d\right) \geq \exp\left(c \frac{\sqrt{\log X}}{\sqrt{\log \log X}}\right), \quad (7.8)$$

for a suitable positive constant c . Indeed, the large values in (7.7) and (7.8) are attained for more than $X^{1-\varepsilon}$ discriminants d with $X \leq |d| \leq 2X$. By Waldspurger's formula, the large values produced in (7.8) show that fundamental Fourier coefficients of half-integer weight eigencuspforms must get large, and the resonance method has been adapted in [73] to show that this holds more generally for half-integer weight cusp forms (not necessarily an eigenform). Another application of this resonance method may be found in the work of Milicevic [120] who obtains large values of Hecke–Maass cusp forms on arithmetic hyperbolic surfaces.

Bondarenko and Seip [23] recently made a breakthrough on this problem, by exhibiting still larger values of $|\zeta(\frac{1}{2} + it)|$. The key ingredient is a beautiful result on *GCD sums* or *Gál sums*: The problem is to find

$$\max_{|\mathcal{N}|=N} \sum_{m,n \in \mathcal{N}} \frac{(m,n)}{\sqrt{mn}}, \quad (7.9)$$

where the maximum is over all N element subsets of the natural numbers. This elegant combinatorial problem turns out to be closely related to maximizing the ratio of quadratic forms (see [2])

$$\max_{|\mathcal{N}|=N} \sup_{\mathbf{x} \in \mathbb{C}^N \neq \mathbf{0}} \left(\sum_{m,n \in \mathcal{N}} x_m \bar{x}_n \frac{(m,n)}{\sqrt{mn}} \right) / \left(\sum_n |x_n|^2 \right). \quad (7.10)$$

Bondarenko and Seip [22, 23] established that the maximum in (7.9) (and also (7.10)) lies between

$$N \exp\left((1 - \varepsilon) \frac{\sqrt{\log N \log \log \log N}}{\sqrt{\log \log N}}\right) \quad \text{and} \quad N \exp\left((7 + \varepsilon) \frac{\sqrt{\log N \log \log \log N}}{\sqrt{\log \log N}}\right),$$

De la Bretèche and Tenenbaum [52] refined this to show that the maximums in (7.9) and (7.10) equal

$$N \exp\left(\left(2\sqrt{2} + o(1)\right) \frac{\sqrt{\log N \log \log \log N}}{\sqrt{\log \log N}}\right). \quad (7.11)$$

The relevance of the bounds for (related) GCD sums to large values of $|\zeta(\sigma + it)|$ was first appreciated by Aistleitner [1] who showed that for fixed $\sigma \in (\frac{1}{2}, 1)$ and T large one has (for some $c_\sigma > 0$)

$$\max_{0 < t \leq T} |\zeta(\sigma + it)| \geq \exp\left(\frac{c_\sigma (\log T)^{1-\sigma}}{(\log \log T)^\sigma}\right),$$

which improved upon earlier applications of the resonance method (see [87, 172]) but only matched the results obtained by Montgomery [125] using very different ideas (see also [3] for large values on the 1-line, and [4] for analogous results for Dirichlet L -functions). On the critical line, Bondarenko and Seip [23] obtained a substantial improvement over the previously known large values of $|\zeta(\frac{1}{2} + it)|$ (see (7.5)) by establishing that

$$\max_{0 < t \leq T} \left| \zeta\left(\frac{1}{2} + it\right) \right| \geq \exp\left((c + o(1)) \frac{\sqrt{\log T \log \log \log T}}{\sqrt{\log \log T}} \right), \quad (7.12)$$

for a positive constant c (in [23] $c = 1/\sqrt{2}$ is permissible, while [52] allows for the improved $c = \sqrt{2}$). The key insight is that in the resonance method one can choose “long resonators” where $R(t)$ is no longer constrained to be a short Dirichlet polynomial ($r(n) = 0$ unless $n \leq T^{1-\varepsilon}$) but instead $R(t)$ is allowed to have $T^{1-\varepsilon}$ nonzero coefficients $r(n)$ so long as these are *positive*. This leads to an optimization problem closely related to the GCD/Gál sums discussed above, and permits the stronger bound in (7.12). Why is it possible to take such long resonators? Consider a smooth nonnegative function Φ whose Fourier transform $\widehat{\Phi}$ is also nonnegative; for example, we could take $\Phi(t) = e^{-t^2/2}$. In place of I_1 and I_2 in (7.3) consider the smoothed integrals

$$\int_{-\infty}^{\infty} |R(t)|^2 \Phi(t/T) dt \quad \text{and} \quad \int_{-\infty}^{\infty} \zeta\left(\frac{1}{2} + it\right) |R(t)|^2 \Phi(t/T) dt. \quad (7.13)$$

Replacing $\zeta(\frac{1}{2} + it)$ with its approximation $\sum_{k \leq T} k^{-\frac{1}{2}-it}$, the second quantity above is approximately

$$\begin{aligned} & \sum_{k \leq T} \frac{1}{\sqrt{k}} \sum_{m,n} r(m)r(n) \int_{-\infty}^{\infty} \left(\frac{n}{mk}\right)^{it} \Phi(t/T) dt \\ & = T \sum_{k \leq T} \frac{1}{\sqrt{k}} \sum_{m,n} r(m)r(n) \widehat{\Phi}(T \log(n/mk)). \end{aligned}$$

Since m and n may be much larger than T , we are unable to restrict just to the “diagonal terms” $n = mk$, but the crucial point is that the positivity of $\widehat{\Phi}$, the resonator coefficients $r(m)$, $r(n)$, and the “coefficients of ζ ” (namely, the function taking 1 on all positive integers) all allow us to keep any terms that we please on the right side above, and ignore other contributions. In this way, one can get a satisfactory lower bound for the ratio of the quantities in (7.13), without needing to evaluate each of these integrals. The restriction on the number of terms allowed in the resonator arises from the fact that $\sum_{k \leq T} k^{-\frac{1}{2}-it}$ is a poor approximation to $\zeta(\frac{1}{2} + it)$ if t is small. These small values of t are unavoidable because the condition that $\widehat{\Phi}$ is nonnegative forces $\Phi(0)$ to be strictly positive.

Unlike the resonance method which applies in great generality, there are (at present) limitations on when the Bondarenko–Seip method of using long resonators applies. In the first place, as we noted above small t must be included, and therefore the maximum in (7.12) is over $t \in [0, T]$ (this can be refined to the interval $[T^\beta, T]$ for any $\beta < 1$ at the cost of weakening the constant c in (7.12)), rather than the dyadic intervals $[T, 2T]$ seen in (7.5). More significantly, the method requires the positivity of the Dirichlet series coefficients of

the L -functions in question (analogously to ζ having coefficients 1), and also the positivity of the right-hand side of any orthogonality relation or trace formula (analogously to $\widehat{\Phi}$ being nonnegative). Apart from $\zeta(s)$, there is one other example in which the Bondarenko–Seip method has been successfully implemented, and this is the work of de la Breteche and Tenenbaum [52] which produces large values of $|L(\frac{1}{2}, \chi)|$ as χ varies over Dirichlet characters (mod q) with q a large prime. To illustrate the subtleties involved, we note that [52] exhibits large values of $|L(\frac{1}{2}, \chi)|$ for *even* characters χ , but the method does not work for *odd* character. This is because in the even case the orthogonality relation

$$\sum_{\substack{\chi \pmod{q} \\ \chi \text{ even}}} \chi(a) = \begin{cases} \frac{\phi(q)}{2} & \text{if } a \equiv \pm 1 \pmod{q}, \\ 0 & \text{otherwise} \end{cases}$$

involves only nonnegative terms on the right-hand side, whereas this is not the situation for odd characters

$$\sum_{\substack{\chi \pmod{q} \\ \chi \text{ odd}}} \chi(a) = \begin{cases} \pm \frac{\phi(q)}{2} & \text{if } a \equiv \pm 1 \pmod{q}, \\ 0 & \text{otherwise.} \end{cases}$$

In particular, the results in (7.7) and (7.8) remain the best currently known, and it would be of great interest to see if the Bondarenko–Seip method could be extended to more general situations.

There is a vast gulf between the conditional upper bounds for $|\zeta(\frac{1}{2} + it)|$ in (7.1) and the large values exhibited in (7.12), and it is natural to ask which of these is closer to the truth. Already in Section 1 we saw a gap (of a factor of 2) between the extreme values of $L(1, \chi_d)$ that may be exhibited (see (1.2)) and the conditional bounds on these extreme values (see (1.3)). There the probabilistic models suggested that the extreme values exhibited in (1.2) represented the truth, and on the critical line too we expect the large values exhibited in (7.12) to be closer to the truth than the bounds in (7.1). For example, if we use Selberg’s central limit theorem as a guide and extrapolate, then the measure of $t \in [0, T]$ with $|\zeta(\frac{1}{2} + it)| \geq e^V$ may be expected to be $\ll T \exp(-(1 + o(1))V^2 / \log \log T)$ (confer (6.6)). If $V = (1 + \varepsilon)\sqrt{\log T \log \log T}$, this measure becomes $\leq T^{-\varepsilon}$, but one can show that if $|\zeta(\frac{1}{2} + it)|$ attains its maximum for $t \in [0, T]$ at $t = t_0$ then in an interval $|t - t_0| \leq c / \log T$ its values are at least of size $\frac{1}{2}|\zeta(\frac{1}{2} + it_0)|$ (see Lemma 2.2 of [59]). This suggests that

$$\max_{t \in [0, T]} \left| \zeta\left(\frac{1}{2} + it\right) \right| \leq \exp\left((1 + o(1))\sqrt{\log T \log \log T}\right).$$

Farmer, Gonek, and Hughes [59] have conjectured that even this overestimates the true size of the maximum, and that possibly

$$\max_{0 \leq t \leq T} \left| \zeta\left(\frac{1}{2} + it\right) \right| = \exp\left(\left(\frac{1}{\sqrt{2}} + o(1)\right)\sqrt{\log T \log \log T}\right). \quad (7.14)$$

To give one indication of why this might hold, consider (6.5) which gives (on RH) an upper bound for $\log |\zeta(\frac{1}{2} + it)|$ in terms of essentially a sum over primes going up to x , accepting an error term of size $\log T / \log x$. If we choose $x = \exp(\sqrt{\log T})$ then this error term is

negligible, and now $\operatorname{Re} \sum_{p \leq x} 1/p^{\frac{1}{2}+it}$ behaves like a Gaussian with mean 0 and variance $\frac{1}{2} \sum_{p \leq x} 1/p \sim \frac{1}{2} \log \log x \sim \frac{1}{4} \log \log T$. Extrapolating this Gaussian behavior, we arrive at the conjectured behavior in (7.14). The conjecture in [59] is based upon a more careful analysis of the hybrid Euler–Hadamard formula developed in [69], which decomposes $\log |\zeta(\frac{1}{2} + it)|$ into terms arising from both primes and zeros in suitable ranges. Probabilistic models for both these terms are analyzed (with the contribution of zeros being modeled using random matrix theory), and the conjecture (7.14) is consistent with many different ways of splitting into primes and zeros. Similar conjectures may be formulated in other families of L -functions, and for example [59] conjectures that

$$\max_{|d| \leq X} L\left(\frac{1}{2}, \chi_d\right) = \exp((1 + o(1))\sqrt{\log X \log \log X}), \quad (7.15)$$

which again is a little smaller (by a factor $\sqrt{2}$ in the exponent) than what might be guessed from extrapolating the Keating–Snaith conjectures for $\log L(\frac{1}{2}, \chi_d)$.

As we discussed in Section 4, one motivation for studying the moments of $|\zeta(\frac{1}{2} + it)|$ is to gain an understanding of its extreme values. In order to do so, one would need an understanding of how $M_k(T)$ behaves with uniformity in k , and a first step might be to examine the asymptotic behavior of the constants a_k and g_k appearing in Conjecture 5.1. One can show that $\log a_k \sim -k^2 \log(2e^\gamma \log k)$, and that $\log g_k \sim k^2 \log(k/4\sqrt{e})$ (see [46]), so that it may seem tempting to speculate that for $T \geq 10$ (say) and uniformly for all $k \geq 2$ one has (for some positive constant c)

$$T \left(\frac{c \log T}{k \log k}\right)^{k^2} \leq \int_0^T \left|\zeta\left(\frac{1}{2} + it\right)\right|^{2k} dt \leq T(\log T)^{k^2}.$$

But there is a curious paradox, and the upper and lower bounds above are inconsistent! If the upper bound above holds uniformly, then it follows that

$$\max_{0 \leq t \leq T} \left|\zeta\left(\frac{1}{2} + it\right)\right| \leq \exp((1 + o(1))\sqrt{\log T \log \log T}).$$

Whereas if the lower bound above holds uniformly, then one must have

$$\max_{0 \leq t \leq T} \left|\zeta\left(\frac{1}{2} + it\right)\right| \geq \exp(C \log T / \log \log T)$$

for some positive constant C . This is an instance where the leading order asymptotic in the moment conjecture does not capture the full story, and one should look instead at the recipe in [43] which (for natural numbers k) gives the entire (conjectural) polynomial P_k of degree k^2 . An analysis of this full moment conjecture suggests that the uniform upper bound stated above might hold: thus, for $T \geq 10$ and natural numbers $k \geq 1$ we conjecture that

$$\int_0^T \left|\zeta\left(\frac{1}{2} + it\right)\right|^{2k} dt \leq T(\log T)^{k^2}. \quad (7.16)$$

In other words, we guess that $\log |\zeta(\frac{1}{2} + it)|$ is *sub-Gaussian* (when thinking of the frequency of its large values), and this gives a weaker version of the Farmer, Gonek, Hughes conjecture (7.14).

While we have confined our discussion above to large values of $|\zeta(\frac{1}{2} + it)|$, or equivalently $\text{Re}(\log \zeta(\frac{1}{2} + it))$, similar considerations apply also to $\text{Im}(\log \zeta(\frac{1}{2} + it))$; see for example [24, 30, 68].

8. FYODOROV–HIARY–KEATING CONJECTURE

A fascinating set of problems has emerged recently with the work of Fyodorov and Keating [65], and Fyodorov, Hiary, and Keating [64], who initiated a study of the distribution of “local maxima” of the Riemann zeta function. More precisely, if t is chosen uniformly from $[T, 2T]$, what is the distribution of

$$\max_{0 \leq h \leq 1} \left| \zeta\left(\frac{1}{2} + it + ih\right) \right|?$$

Although it does not make much of a difference, [64] considers the maximum over intervals of length 2π instead of 1 since this has a natural analogue in random matrix theory. If a matrix g is chosen randomly from $U(N)$ (with respect to Haar measure), what is the distribution of

$$\max_{\theta \in [0, 2\pi)} |\det(I - g e^{-i\theta})|?$$

In the context of $\zeta(\frac{1}{2} + it)$, one initial motivation for considering this problem was that it might shed new light on the global maximum over the long interval $[0, T]$ (discussed in the previous section). While the distribution of the local maxima leads to striking new and subtle phenomena involving the local correlations of the zeta function, it does not seem to inform the behavior of the global maximum.

Conjecture 8.1 (Fyodorov–Hiary–Keating [64]). *For any real number y , as $T \rightarrow \infty$ one has*

$$\frac{1}{T} \text{meas} \left\{ T \leq t \leq 2T : \max_{0 \leq h \leq 1} \left| \zeta\left(\frac{1}{2} + it + ih\right) \right| \leq e^y \frac{\log T}{(\log \log T)^{\frac{3}{4}}} \right\} \rightarrow F(y), \quad (8.1)$$

where the cumulative distribution function F satisfies $F(y) \rightarrow 0$ as $y \rightarrow -\infty$, and satisfies $1 - F(y) \sim C y e^{-2y}$ as $y \rightarrow \infty$ for some constant $C > 0$. In particular, for any function $g(T)$ tending to infinity with T , one has

$$\begin{aligned} &\text{meas} \left\{ T \leq t \leq 2T : \left| \max_{0 \leq h \leq 1} \log \left| \zeta\left(\frac{1}{2} + it + ih\right) \right| - \log \log T + \frac{3}{4} \log \log \log T \right| \leq g(T) \right\} \\ &\sim T. \end{aligned} \quad (8.2)$$

Let us first explain what is striking and unexpected about this conjecture. Roughly speaking, in an interval of length 1 we may think of the zeta function as being determined by about $\log T$ values—this is about the number of zeros we expect to find in such an interval, and we may guess that if $|t_1 - t_2| \leq 1/\log T$ then $\log |\zeta(\frac{1}{2} + it_1)|$ and $\log |\zeta(\frac{1}{2} + it_2)|$ are about the same. Selberg’s theorem tells us that the values $\log |\zeta(\frac{1}{2} + it)|$ are distributed like a normal variable with mean 0 and variance $\frac{1}{2} \log \log T$. Thus a first guess for the distribution of $\max_{0 \leq h \leq 1} \log |\zeta(\frac{1}{2} + it + ih)|$ might be that it behaves like the maximum of about $\log T$ independently drawn normal random variables with mean 0 and variance $\frac{1}{2} \log \log T$. The

maximum of N independent normal variables with mean 0 and variance 1 is very sharply concentrated around $\sqrt{2 \log N} - (\log \sqrt{4\pi \log N})/\sqrt{2 \log N}$ (the precise distribution is known as the Gumbel distribution, and has been extensively studied in view of its enormous significance in practical assessments of the risk of rare events). After scaling by the standard deviation $\sqrt{\frac{1}{2} \log \log T}$ in Selberg's theorem, this naive model would indicate that $\max_{0 \leq h \leq 1} \log |\zeta(\frac{1}{2} + it + ih)|$ should typically be around

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

In contrast, Conjecture 8.1 predicts that $\max_{0 \leq h \leq 1} |\zeta(\frac{1}{2} + it)|$ is usually a bit smaller, of size $(\log T)/(\log \log T)^{\frac{3}{4}}$. There is also a subtle difference in the decay of $1 - F(y)$ in (8.1), which is predicted to decay like ye^{-2y} , whereas the Gumbel distribution would have predicted a decay rate of e^{-2y} .

The flaw in the naive heuristic presented above is that nearby values of the zeta function are not independent, but are correlated. Suppose t is randomly chosen from $[T, 2T]$ and $0 \leq h \leq 1$, and consider the covariance of $\log |\zeta(\frac{1}{2} + it)|$ and $\log |\zeta(\frac{1}{2} + it + ih)|$. As in our discussion of Selberg's theorem in Section 2, we may often approximate these values by corresponding sums over primes $\text{Re} \sum_{p \leq x} 1/p^{\frac{1}{2} + it}$ and $\text{Re} \sum_{p \leq x} 1/p^{\frac{1}{2} + it + ih}$ with x a suitable small power of T . If p is small in comparison to $e^{1/h}$ then $p^{ih} \approx 1$, and the corresponding terms in our prime sums are strongly correlated. The terms with p much larger than $e^{1/h}$ are largely uncorrelated, since as p varies in such large ranges p^{ih} will become equidistributed on the unit circle. Thus one may see that

$$\begin{aligned} \frac{1}{T} \int_T^{2T} \log \left| \zeta \left(\frac{1}{2} + it \right) \right| \log \left| \zeta \left(\frac{1}{2} + it + ih \right) \right| dt &\sim \frac{1}{2} \sum_{p \leq x} \frac{\cos(h \log p)}{p} \\ &\sim \frac{1}{2} \log \min(h^{-1}, \log T). \end{aligned} \quad (8.3)$$

This correlation structure of nearby values must be taken into account when trying to predict the behavior of local maxima.

To gain a rough idea of how to model the local behavior of $\log |\zeta(\frac{1}{2} + it)|$, put for each $1 \leq k \leq \log \log T - 1$,

$$\mathcal{P}_k(t) = \text{Re} \sum_{e^{e^{k-1}} \leq p \leq e^{e^k}} \frac{1}{p^{1/2 + it}}, \quad (8.4)$$

so that we may think of $\log |\zeta(\frac{1}{2} + it)|$ as something like $\sum_k \mathcal{P}_k(t)$. Each $\mathcal{P}_k(t)$ is distributed like a normal random variable with mean 0 and variance $\sim \frac{1}{2} \sum_{e^{e^{k-1}} \leq p \leq e^{e^k}} 1/p \sim \frac{1}{2}$. Moreover, for different values of k , the sums $\mathcal{P}_k(t)$ involve primes in disjoint ranges, and therefore behave independently of each other. Notice further that if $|t_1 - t_2| \leq e^{-k}$ then $\mathcal{P}_k(t_1)$ and $\mathcal{P}_k(t_2)$ are more or less the same. Thus instead of modeling $\log |\zeta(\frac{1}{2} + it)|$ in intervals of length 1 by about $\log T$ independent samples of a normal random variable, we are led to the following more nuanced model. For each k , let P_k denote any one of about e^k independent drawings of a normal random variable with mean 0 and variance $\frac{1}{2}$. Then $\log |\zeta(\frac{1}{2} + it)|$ in an interval of length 1 is modeled by all the possibilities for $\sum_k P_k$.

The model described above has been analyzed in the probability literature surrounding branching random walks and branching Brownian motion. Consider a particle starting at time 0 and moving as a standard Brownian motion. At time t , with probability e^{-t} the particle might split into two particles, that move according to independent standard Brownian motions starting from that position. These particles may again split (independently of each other) at a future time, giving rise to more daughter particles, and so on. After time T , how is the maximum value of all these particles distributed? This problem was resolved by Bramson who established that the maximum is almost surely $\sqrt{2}(T - \frac{3}{4} \log T) + O(1)$. Notice the $\frac{3}{4}$ term here, which exactly parallels the $\frac{3}{4}$ terms appearing in Conjecture 8.1!

In recent years there has been a lot of progress towards understanding Conjecture 8.1. In [7] Arguin, Belius, and Harper considered $\max_{0 \leq h \leq 1} \operatorname{Re} \sum_{p \leq T} X(p)/p^{\frac{1}{2}+ih}$ where the $X(p)$'s are independent random variables chosen uniformly on the unit circle (a randomized model for $\log |\zeta(\frac{1}{2} + it + ih)|$), and established that almost surely this is $\log \log T - (\frac{3}{4} + o(1)) \log \log \log T$. Najnudel [129] established that on RH the set of $t \in [T, 2T]$ with $\max_{0 \leq h \leq 1} |\zeta(\frac{1}{2} + it + ih)| = (\log T)^{1+o(1)}$ has measure $\sim T$. Independently this result was also established unconditionally by Arguin, Belius, Bourgade, Radziwiłł, and Soundararajan [6]. A lovely exposition of Conjecture 8.1 and the results mentioned so far may be found in Harper's Bourbaki seminar [75]. Still more recently, Harper [76] established that if t is not in an exceptional subset of $[T, 2T]$ with measure $o(T)$, then

$$\max_{0 \leq h \leq 1} \log \left| \zeta \left(\frac{1}{2} + it + ih \right) \right| \leq \log \log T - \frac{3}{4} \log \log \log T + O(\log \log \log \log T),$$

so that at least in one direction, the difference between the naive constant $\frac{1}{4}$ and the refined prediction $\frac{3}{4}$ could be established. Independently Arguin, Bourgade, and Radziwiłł [8] established the sharper result that for any $y \geq 1$,

$$\frac{1}{T} \operatorname{meas} \left\{ t \in [T, 2T] : \max_{0 \leq h \leq 1} \left| \zeta \left(\frac{1}{2} + it + ih \right) \right| > \frac{e^y \log T}{(\log \log T)^{\frac{3}{4}}} \right\} \leq C y e^{-2y},$$

for some constant C . This beautiful result establishes part of Conjecture 8.1, and the decay in y above matches (up to constants) the conjectured behavior of $1 - F(y)$. There has also been substantial progress toward the analogue of Conjecture 8.1 in random matrix theory; see [5, 37, 135].

Instead of considering the maximum of the zeta function in intervals of length 1, one may also examine other "local moments" $\int_0^1 |\zeta(\frac{1}{2} + it + ih)|^\beta dh$. This was already suggested in [64], who conjectured that a transition in the behavior of these local moments occurs at the critical exponent $\beta = 2$ —for $\beta < 2$ these local moments are typically of size $(\log T)^{\beta^2/4}$ (the size of the global moment $\frac{1}{T} \int_T^{2T} |\zeta(\frac{1}{2} + it)|^\beta dt$), whereas for $\beta > 2$ they are typically of size $(\log T)^{\beta-1}$ corresponding to the largest value of zeta in that interval (about size $\log T$) which might be expected to occur on an interval of length about $1/\log T$. For work in this direction see [9, 11, 76]. We mention a lovely result of Harper [76] for the critical exponent $\beta = 2$:

$$\frac{1}{T} \int_T^{2T} \left(\frac{1}{\log T} \int_0^1 \left| \zeta \left(\frac{1}{2} + it + ih \right) \right|^2 dh \right)^{\frac{1}{2}} dt \ll \frac{1}{(\log \log T)^{\frac{1}{4}}}.$$

A simple application of Cauchy's inequality together with the second moment of $\zeta(\frac{1}{2} + it)$ shows that the above quantity is $\ll 1$, and the fact that it is a little bit smaller is a reflection of the correlation structure of nearby values of $\zeta(s)$ that also underlies Conjecture 8.1.

The ideas discussed here are closely connected to what is termed *Gaussian multiplicative chaos*, which was initiated by Kahane [97], and which has been extensively studied in the probability literature [146]. In number theory, these ideas are closely related to the study of mean values of random multiplicative functions. We content ourselves with giving a few pointers to surveys and related work [12, 74, 151, 167].

ACKNOWLEDGMENTS

I am grateful to Brian Conrey, Jon Keating, Emmanuel Kowalski, Vivian Kuperberg, Maksym Radziwiłł, Matt Tyler, and Max Xu for their careful reading and many valuable suggestions.

FUNDING

This work was partially supported by grants from the National Science Foundation, and a Simons Investigator Award from the Simons Foundation.

REFERENCES

- [1] C. Aistleitner, Lower bounds for the maximum of the Riemann zeta function along vertical lines. *Math. Ann.* **365** (2016), no. 1–2, 473–496.
- [2] C. Aistleitner, I. Berkes, and K. Seip, GCD sums from Poisson integrals and systems of dilated functions. *J. Eur. Math. Soc. (JEMS)* **17** (2015), no. 6, 1517–1546.
- [3] C. Aistleitner, K. Mahatab, and M. Munsch, Extreme values of the Riemann zeta function on the 1-line. *Int. Math. Res. Not. IMRN* **22** (2019), 6924–6932.
- [4] C. Aistleitner, K. Mahatab, M. Munsch, and A. Peyrot, On large values of $L(\sigma, \chi)$. *Q. J. Math.* **70** (2019), no. 3, 831–848.
- [5] L.-P. Arguin, D. Belius, and P. Bourgade, Maximum of the characteristic polynomial of random unitary matrices. *Comm. Math. Phys.* **349** (2017), no. 2, 703–751.
- [6] L.-P. Arguin, D. Belius, P. Bourgade, M. Radziwiłł, and K. Soundararajan, Maximum of the Riemann zeta function on a short interval of the critical line. *Comm. Pure Appl. Math.* **72** (2019), no. 3, 500–535.
- [7] L.-P. Arguin, D. Belius, and A. J. Harper, Maxima of a randomized Riemann zeta function, and branching random walks. *Ann. Appl. Probab.* **27** (2017), no. 1, 178–215.
- [8] L.-P. Arguin, P. Bourgade, and M. Radziwiłł, The Fyodorov–Hiary–Keating conjecture. I. 2020, arXiv:2007.00988.
- [9] L.-P. Arguin, F. Ouimet, and M. Radziwiłł, Moments of the Riemann zeta function on short intervals of the critical line. 2021, arXiv:1901.04061.

- [10] B. Bagchi, *Statistical behaviour and universality properties of the Riemann zeta-function and other allied Dirichlet series*. Ph.D. thesis, Indian Statistical Institute, Kolkata, 1981.
- [11] E. C. Bailey and J. P. Keating, On the moments of the moments of $\zeta(1/2 + it)$. *J. Number Theory* **223** (2021), 79–100.
- [12] E. C. Bailey and J. P. Keating, Maxima of log-correlated fields: some recent developments. 2021, arXiv:2106.15141.
- [13] R. Balasubramanian, An improvement on a theorem of Titchmarsh on the mean square of $|\zeta(\frac{1}{2} + it)|$. *Proc. Lond. Math. Soc. (3)* **36** (1978), no. 3, 540–576.
- [14] R. Balasubramanian, On the frequency of Titchmarsh’s phenomenon for $\zeta(s)$. IV. *Hardy-Ramanujan J.* **9** (1986), 1–10.
- [15] R. Balasubramanian, J. B. Conrey, and D. R. Heath-Brown, Asymptotic mean square of the product of the Riemann zeta-function and a Dirichlet polynomial. *J. Reine Angew. Math.* **357** (1985), 161–181.
- [16] R. Balasubramanian and K. Ramachandra, On the frequency of Titchmarsh’s phenomenon for $\zeta(s)$. III. *Proc. Indian Acad. Sci., Sect. A, Phys. Sci.* **86** (1977), no. 4, 341–351.
- [17] S. Bettin, V. Chandee, and M. Radziwiłł, The mean square of the product of the Riemann zeta-function with Dirichlet polynomials. *J. Reine Angew. Math.* **729** (2017), 51–79.
- [18] V. Blomer, É. Fouvry, E. Kowalski, P. Michel, D. Milićević, and W. Sawin, The second moment theory of families of L-functions. 2019, arXiv:1804.01450.
- [19] V. Blomer and R. Khan, Twisted moments of L-functions and spectral reciprocity. *Duke Math. J.* **168** (2019), no. 6, 1109–1177.
- [20] J. W. Bober and L. Goldmakher, Pólya–Vinogradov and the least quadratic non-residue. *Math. Ann.* **366** (2016), no. 1–2, 853–863.
- [21] E. Bombieri and D. A. Hejhal, On the distribution of zeros of linear combinations of Euler products. *Duke Math. J.* **80** (1995), no. 3, 821–862.
- [22] A. Bondarenko and K. Seip, GCD sums and complete sets of square-free numbers. *Bull. Lond. Math. Soc.* **47** (2015), no. 1, 29–41.
- [23] A. Bondarenko and K. Seip, Large greatest common divisor sums and extreme values of the Riemann zeta function. *Duke Math. J.* **166** (2017), no. 9, 1685–1701.
- [24] A. Bondarenko and K. Seip, Extreme values of the Riemann zeta function and its argument. *Math. Ann.* **372** (2018), no. 3–4, 999–1015.
- [25] J. Bourgain, Decoupling, exponential sums and the Riemann zeta function. *J. Amer. Math. Soc.* **30** (2017), no. 1, 205–224.
- [26] H. M. Bui, N. Evans, S. Lester, and K. Pratt, Weighted central limit theorems for central values of L-functions. 2021, arXiv:2109.06829.
- [27] H. M. Bui, A. Florea, J. P. Keating, and E. Roditty-Gershon, Moments of quadratic twists of elliptic curve L-functions over function fields. *Algebra Number Theory* **14** (2020), no. 7, 1853–1893.

- [28] D. Bump and A. Gamburd, On the averages of characteristic polynomials from classical groups. *Comm. Math. Phys.* **265** (2006), no. 1, 227–274.
- [29] D. A. Burgess, The distribution of quadratic residues and non-residues. *Mathematika* **4** (1957), 106–112.
- [30] E. Carneiro, V. Chandee, and M. B. Milinovich, Bounding $S(t)$ and $S_1(t)$ on the Riemann hypothesis. *Math. Ann.* **356** (2013), no. 3, 939–968.
- [31] V. Chandee, Explicit upper bounds for L -functions on the critical line. *Proc. Amer. Math. Soc.* **137** (2009), no. 12, 4049–4063.
- [32] V. Chandee, On the correlation of shifted values of the Riemann zeta function. *Q. J. Math.* **62** (2011), no. 3, 545–572.
- [33] V. Chandee and X. Li, The eighth moment of Dirichlet L -functions. *Adv. Math.* **259** (2014), 339–375.
- [34] V. Chandee and X. Li, The sixth moment of automorphic L -functions. *Algebra Number Theory* **11** (2017), no. 3, 583–633.
- [35] V. Chandee and X. Li, The 8th moment of the family of $\Gamma_1(q)$ -automorphic L -functions. *Int. Math. Res. Not. IMRN* **22** (2020), 8443–8485.
- [36] V. Chandee and K. Soundararajan, Bounding $|\zeta(\frac{1}{2} + it)|$ on the Riemann hypothesis. *Bull. Lond. Math. Soc.* **43** (2011), no. 2, 243–250.
- [37] R. Chhaibi, T. Madaule, and J. Najnudel, On the maximum of the $C\beta E$ field. *Duke Math. J.* **167** (2018), no. 12, 2243–2345.
- [38] G. Chinta, Analytic ranks of elliptic curves over cyclotomic fields. *J. Reine Angew. Math.* **544** (2002), 13–24.
- [39] S. Chowla and P. Erdős, A theorem on the distribution of the values of L -functions. *J. Indian Math. Soc. (N.S.)* **15** (1951), 11–18.
- [40] J. Cogdell and P. Michel, On the complex moments of symmetric power L -functions at $s = 1$. *Int. Math. Res. Not.* **31** (2004), 1561–1617.
- [41] B. Conrey and J. P. Keating, Moments of zeta and correlations of divisor-sums: V. *Proc. Lond. Math. Soc. (3)* **118** (2019), no. 4, 729–752.
- [42] J. B. Conrey and D. W. Farmer, Mean values of L -functions and symmetry. *Int. Math. Res. Not.* **17** (2000), 883–908.
- [43] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein, and N. C. Snaith, Integral moments of L -functions. *Proc. Lond. Math. Soc. (3)* **91** (2005), no. 1, 33–104.
- [44] J. B. Conrey and A. Ghosh, Mean values of the Riemann zeta-function. III. In *Proceedings of the Amalfi conference on analytic number theory (Maiori, 1989)*, pp. 35–59, Univ. Salerno, Salerno, 1992.
- [45] J. B. Conrey and A. Ghosh, A conjecture for the sixth power moment of the Riemann zeta-function. *Int. Math. Res. Not.* **15** (1998), 775–780.
- [46] J. B. Conrey and S. M. Gonek, High moments of the Riemann zeta-function. *Duke Math. J.* **107** (2001), no. 3, 577–604.
- [47] J. B. Conrey and H. Iwaniec, The cubic moment of central values of automorphic L -functions. *Ann. of Math. (2)* **151** (2000), no. 3, 1175–1216.

- [48] J. B. Conrey, H. Iwaniec, and K. Soundararajan, The sixth power moment of Dirichlet L -functions. *Geom. Funct. Anal.* **22** (2012), no. 5, 1257–1288.
- [49] A. Dahl and Y. Lamzouri, The distribution of class numbers in a special family of real quadratic fields. *Trans. Amer. Math. Soc.* **370** (2018), no. 9, 6331–6356.
- [50] H. Davenport, *Multiplicative number theory. Second edn.* Grad. Texts in Math. 74, Springer, New York–Berlin, 1980.
- [51] C. David, A. Florea, and M. Lalin, Non-vanishing for cubic L -functions. 2020, arXiv:2006.15661.
- [52] R. de la Bretèche and G. Tenenbaum, Sommes de Gál et applications. *Proc. Lond. Math. Soc. (3)* **119** (2019), no. 1, 104–134.
- [53] P.-O. Dehaye, Combinatorics of lower order terms in the moment conjectures for the Riemann zeta function. 2012, arXiv:1201.4478.
- [54] A. Diaconu, D. Goldfeld, and J. Hoffstein, Multiple Dirichlet series and moments of zeta and L -functions. *Compos. Math.* **139** (2003), no. 3, 297–360.
- [55] A. Diaconu and I. Whitehead, On the third moment of $L(\frac{1}{2}, \chi_d)$ II: the number field case. *J. Eur. Math. Soc. (JEMS)* **23** (2021), no. 6, 2051–2070.
- [56] W. Duke, Extreme values of Artin L -functions and class numbers. *Compos. Math.* **136** (2003), no. 1, 103–115.
- [57] B. Eckhardt, S. Fishman, J. Keating, O. Agam, J. Main, and K. Müller, Approach to ergodicity in quantum wave functions. *Phys. Rev. E* **52** (1995), 5893–5903.
- [58] P. D. T. A. Elliott, On the distribution of the values of quadratic L -series in the half-plane $\sigma > \frac{1}{2}$. *Invent. Math.* **21** (1973), 319–338.
- [59] D. W. Farmer, S. M. Gonek, and C. P. Hughes, The maximum size of L -functions. *J. Reine Angew. Math.* **609** (2007), 215–236.
- [60] A. Fazzari, A weighted central limit theorem for $\log |\zeta(1/2 + it)|$. *Mathematika* **67** (2021), no. 2, 324–341.
- [61] A. Florea, The fourth moment of quadratic Dirichlet L -functions over function fields. *Geom. Funct. Anal.* **27** (2017), no. 3, 541–595.
- [62] P. J. Forrester and S. O. Warnaar, The importance of the Selberg integral. *Bull. Amer. Math. Soc. (N.S.)* **45** (2008), no. 4, 489–534.
- [63] J. B. Friedlander, Bounds for L -functions. In *Proceedings of the international congress of mathematicians, Vol. 1, 2 (Zürich, 1994)*, pp. 363–373, Birkhäuser, Basel, 1995.
- [64] Y. V. Fyodorov, G. Hiary, and J. Keating, Freezing transition, characteristic polynomials of random matrices, and the Riemann zeta function. *Phys. Rev. Lett.* **108** (2012), 170601.
- [65] Y. V. Fyodorov and J. Keating, Freezing transitions and extreme values: random matrix theory, and disordered landscapes. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **372** (2014), no. 2007, 20120503.
- [66] P. Gao and L. Zhao, Bounds for moments of cubic and quartic Dirichlet L -functions. 2021, arXiv:2104.09909.

- [67] J. Germain, Factorization of constants involved in conjectural moments of zeta-functions. *Integers* **8** (2008), A47, 18.
- [68] D. A. Goldston and S. M. Gonek, A note on $S(t)$ and the zeros of the Riemann zeta-function. *Bull. Lond. Math. Soc.* **39** (2007), no. 3, 482–486.
- [69] S. M. Gonek, C. P. Hughes, and J. P. Keating, A hybrid Euler–Hadamard product for the Riemann zeta function. *Duke Math. J.* **136** (2007), no. 3, 507–549.
- [70] A. Granville and K. Soundararajan, Upper bounds for $|L(1, \chi)|$. *Q. J. Math.* **53** (2002), no. 3, 265–284.
- [71] A. Granville and K. Soundararajan, The distribution of values of $L(1, \chi_d)$. *Geom. Funct. Anal.* **13** (2003), no. 5, 992–1028.
- [72] A. Granville and K. Soundararajan, Large character sums: Burgess’s theorem and zeros of L -functions. *J. Eur. Math. Soc. (JEMS)* **20** (2018), no. 1, 1–14.
- [73] S. Gun, W. Kohlen, and K. Soundararajan, Large Fourier coefficients of half-integer weight modular forms. 2020, arXiv:2004.14450.
- [74] A. J. Harper, Moments of random multiplicative functions, I: Low moments, better than square-root cancellation, and critical multiplicative chaos. *Forum Math. Pi* **8** (2020), e1, 95.
- [75] A. J. Harper, The Riemann zeta function in short intervals [after Najnudel, and Arguin, Belius, Bourgade, Radziwiłł and Soundararajan]. *Astérisque* **422** (2020). Séminaire Bourbaki. Vol. 2018/2019. Exposés 1151–1165, 391–414.
- [76] A. J. Harper, On the partition function of the Riemann zeta function, and the Fyodorov–Hiary–Keating conjecture. 2019, arXiv:1906.05783.
- [77] A. J. Harper, Sharp conditional bounds for moments of the Riemann zeta function. 2013, arXiv:1305.4618.
- [78] W. Heap, M. Radziwiłł, and K. Soundararajan, Sharp upper bounds for fractional moments of the Riemann zeta function. *Q. J. Math.* **70** (2019), no. 4, 1387–1396.
- [79] W. Heap and K. Soundararajan, Lower bounds for moments of zeta and L -functions revisited. 2020, arXiv:2007.13154.
- [80] D. R. Heath-Brown, The twelfth power moment of the Riemann zeta-function. *Quart. J. Math. Oxf. Ser. (2)* **29** (1978), no. 116, 443–462.
- [81] D. R. Heath-Brown, The fourth power moment of the Riemann zeta function. *Proc. Lond. Math. Soc. (3)* **38** (1979), no. 3, 385–422.
- [82] D. R. Heath-Brown, On a paper of Baker and Schinzel. *Acta Arith.* **35** (1979), no. 2, 203–207.
- [83] D. R. Heath-Brown, The fourth power mean of Dirichlet’s L -functions. *Analysis* **1** (1981), no. 1, 25–32.
- [84] D. R. Heath-Brown, Fractional moments of the Riemann zeta function. *J. Lond. Math. Soc. (2)* **24** (1981), no. 1, 65–78.
- [85] D. R. Heath-Brown, The average analytic rank of elliptic curves. *Duke Math. J.* **122** (2004), no. 3, 591–623.

- [86] G. A. Hiary and A. M. Odlyzko, The zeta function on the critical line: numerical evidence for moments and random matrix theory models. *Math. Comp.* **81** (2012), no. 279, 1723–1752.
- [87] T. Hilberdink, An arithmetical mapping and applications to Ω -results for the Riemann zeta function. *Acta Arith.* **139** (2009), no. 4, 341–367.
- [88] S. Holmin, N. Jones, P. Kurlberg, C. McLeman, and K. Petersen, Missing class groups and class number statistics for imaginary quadratic fields. *Exp. Math.* **28** (2019), no. 2, 233–254.
- [89] B. Hough, The distribution of the logarithm in an orthogonal and a symplectic family of L -functions. *Forum Math.* **26** (2014), no. 2, 523–546.
- [90] M. N. Huxley, The large sieve inequality for algebraic number fields. II. Means of moments of Hecke zeta-functions. *Proc. Lond. Math. Soc. (3)* **21** (1970), 108–128.
- [91] A. Ivić, On sums of Hecke series in short intervals. *J. Théor. Nombres Bordeaux* **13** (2001), no. 2, 453–468.
- [92] H. Iwaniec and E. Kowalski, *Analytic number theory*. Amer. Math. Soc. Colloq. Publ. 53, American Mathematical Society, Providence, RI, 2004.
- [93] H. Iwaniec and P. Sarnak, The non-vanishing of central values of automorphic L -functions and Landau–Siegel zeros. *Israel J. Math. (A)* **120** (2000), 155–177.
- [94] H. Iwaniec and P. Sarnak, Perspectives on the analytic theory of L -functions. *Geom. Funct. Anal. Special Volume, Part II* (2000), 705–741.
- [95] J. Jacod, E. Kowalski, and A. Nikeghbali, Mod-Gaussian convergence: new limit theorems in probability and number theory. *Forum Math.* **23** (2011), no. 4, 835–873.
- [96] M. Jutila, On the mean value of $L(\frac{1}{2}, \chi)$ for real characters. *Analysis* **1** (1981), no. 2, 149–161.
- [97] J.-P. Kahane, Sur le chaos multiplicatif. *Ann. Sci. Math. Québec* **9** (1985), no. 2, 105–150.
- [98] N. M. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*. Amer. Math. Soc. Colloq. Publ. 45, American Mathematical Society, Providence, RI, 1999.
- [99] J. P. Keating, B. Rodgers, E. Roditty-Gershon, and Z. Rudnick, Sums of divisor functions in $\mathbb{F}_q[t]$ and matrix integrals. *Math. Z.* **288** (2018), no. 1–2, 167–198.
- [100] J. P. Keating and N. C. Snaith, Random matrix theory and $\zeta(1/2 + it)$. *Comm. Math. Phys.* **214** (2000), no. 1, 57–89.
- [101] J. P. Keating and N. C. Snaith, Random matrix theory and L -functions at $s = 1/2$. *Comm. Math. Phys.* **214** (2000), no. 1, 91–110.
- [102] R. Khan, The fifth moment of Hecke L -functions in the weight aspect. *Math. Proc. Cambridge Philos. Soc.* **168** (2020), no. 3, 543–566.
- [103] R. Khan and H. T. Ngo, Nonvanishing of Dirichlet L -functions. *Algebra Number Theory* **10** (2016), no. 10, 2081–2091.

- [104] E. M. Kiral and M. Young, The fifth moment of modular L -functions. *J. Eur. Math. Soc. (JEMS)* **23** (2021), no. 1, 237–314.
- [105] W. Kohnen and D. Zagier, Values of L -series of modular forms at the center of the critical strip. *Invent. Math.* **64** (1981), no. 2, 175–198.
- [106] E. Kowalski, Bagchi’s theorem for families of automorphic forms. In *Exploring the Riemann zeta function*, pp. 181–199, Springer, Cham, 2017.
- [107] E. Kowalski, P. Michel, and W. Sawin, Bilinear forms with Kloosterman sums and applications. *Ann. of Math. (2)* **186** (2017), no. 2, 413–500.
- [108] E. Kowalski and A. Nikeghbali, Mod-Gaussian convergence and the value distribution of $\zeta(\frac{1}{2} + it)$ and related quantities. *J. Lond. Math. Soc. (2)* **86** (2012), no. 1, 291–319.
- [109] Y. Lamzouri, The two-dimensional distribution of values of $\zeta(1 + it)$. *Int. Math. Res. Not. IMRN* **2008** (2008), rnn 106, 48 pp.
- [110] Y. Lamzouri, Large moments and extreme values of class numbers of indefinite binary quadratic forms. *Mathematika* **63** (2017), no. 2, 564–586.
- [111] Y. Lamzouri, On the average of the number of imaginary quadratic fields with a given class number. *Ramanujan J.* **44** (2017), no. 2, 411–416.
- [112] Y. Lamzouri, S. Lester, and M. Radziwiłł, An effective universality theorem for the Riemann zeta function. *Comment. Math. Helv.* **93** (2018), no. 4, 709–736.
- [113] Y. Lamzouri, S. Lester, and M. Radziwiłł, Discrepancy bounds for the distribution of the Riemann zeta-function and applications. *J. Anal. Math.* **139** (2019), no. 2, 453–494.
- [114] Y. Lamzouri, X. Li, and K. Soundararajan, Conditional bounds for the least quadratic non-residue and related problems. *Math. Comp.* **84** (2015), no. 295, 2391–2412.
- [115] A. Laurinčikas, A limit theorem for the Riemann zeta-function on the critical line. II. *Litov. Mat. Sb.* **27** (1987), no. 3, 489–500.
- [116] S. Lester and M. Radziwiłł, Signs of Fourier coefficients of half-integral weight modular forms. *Math. Ann.* **379** (2021), no. 3–4, 1553–1604.
- [117] X. Li, Upper bounds on L -functions at the edge of the critical strip. *Int. Math. Res. Not. IMRN* **4** (2010), 727–755.
- [118] W. Luo, Values of symmetric square L -functions at 1. *J. Reine Angew. Math.* **506** (1999), 215–235.
- [119] P. Michel and A. Venkatesh, Equidistribution, L -functions and ergodic theory: on some problems of Yu. Linnik. In *International congress of mathematicians. Vol. II*, pp. 421–457, Eur. Math. Soc., Zürich, 2006.
- [120] D. Milićević, Large values of eigenfunctions on arithmetic hyperbolic surfaces. *Duke Math. J.* **155** (2010), no. 2, 365–401.
- [121] M. B. Milinovich, Upper bounds for moments of $\zeta'(\rho)$. *Bull. Lond. Math. Soc.* **42** (2010), no. 1, 28–44.
- [122] M. B. Milinovich and N. Ng, Lower bounds for moments of $\zeta'(\rho)$. *Int. Math. Res. Not. IMRN* **12** (2014), 3190–3216.

- [123] G. Molteni, Upper and lower bounds at $s = 1$ for certain Dirichlet series with Euler product. *Duke Math. J.* **111** (2002), no. 1, 133–158.
- [124] H. L. Montgomery, The pair correlation of zeros of the zeta function. In *Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972)*, pp. 181–193, American Math Society, 1973.
- [125] H. L. Montgomery, Extreme values of the Riemann zeta function. *Comment. Math. Helv.* **52** (1977), no. 4, 511–518.
- [126] H. L. Montgomery and R. C. Vaughan, Extreme values of Dirichlet L -functions at 1. In *Number theory in progress, Vol. 2 (Zakopane–Kościelisko, 1997)*, pp. 1039–1052, de Gruyter, Berlin, 1999.
- [127] Y. Motohashi, *Spectral theory of the Riemann zeta-function*. Cambridge Tracts in Math. 127, Cambridge University Press, Cambridge, 1997.
- [128] R. Munshi, The subconvexity problem for L -functions. In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. II. Invited lectures*, pp. 363–376, World Sci. Publ., Hackensack, NJ, 2018.
- [129] J. Najnudel, On the extreme values of the Riemann zeta function on random intervals of the critical line. *Probab. Theory Related Fields* **172** (2018), no. 1–2, 387–452.
- [130] P. D. Nelson, Bounds for standard L -functions. 2021, arXiv:2109.15230.
- [131] P. D. Nelson, Eisenstein series and the cubic moment for $\mathrm{PGL}(2)$. 2020, arXiv:1911.06310.
- [132] P. D. Nelson, Spectral aspect subconvex bounds for $U_{n+1} \times U_n$. 2021, arXiv:2012.02187.
- [133] N. Ng, The sixth moment of the Riemann zeta function and ternary additive divisor sums. *Discrete Anal.* **6** (2021), 60.
- [134] A. M. Odlyzko, The 10^{20} -th zero of the Riemann zeta function and 175 million of its neighbors. 1992, <http://www.dtc.umn.edu/~odlyzko/unpublished/zeta.10to20.1992.pdf>.
- [135] E. Paquette and O. Zeitouni, The maximum of the CUE field. *Int. Math. Res. Not. IMRN* **16** (2018), 5028–5119.
- [136] Z. Peng, *Zeros and central values of automorphic L -functions*. Ph.D. thesis, Princeton University, 2001.
- [137] I. Petrow and M. P. Young, The Weyl bound for Dirichlet L -functions of cube-free conductor. *Ann. of Math. (2)* **192** (2020), no. 2, 437–486.
- [138] M. Radziwiłł, High moments of Dirichlet L -functions. In *Automorphic forms and arithmetic*, pp. 2520–2521, Oberwolfach Rep. 40, EMS Press, 2017.
- [139] M. Radziwiłł, Large deviations in Selberg’s central limit theorem. 2011, arXiv:1108.5092.
- [140] M. Radziwiłł and K. Soundararajan, Continuous lower bounds for moments of zeta and L -functions. *Mathematika* **59** (2013), no. 1, 119–128.
- [141] M. Radziwiłł and K. Soundararajan, Moments and distribution of central L -values of quadratic twists of elliptic curves. *Invent. Math.* **202** (2015), no. 3, 1029–1068.

- [142] M. Radziwiłł and K. Soundararajan, Selberg’s central limit theorem for $\log |\zeta(1/2 + it)|$. *Enseign. Math.* **63** (2017), no. 1–2, 1–19.
- [143] K. Ramachandra, Some remarks on the mean value of the Riemann zeta function and other Dirichlet series. I. *Hardy-Ramanujan J.* **1** (1978), 15.
- [144] K. Ramachandra, Some remarks on the mean value of the Riemann zeta function and other Dirichlet series. II. *Hardy-Ramanujan J.* **3** (1980), 1–24.
- [145] N. Raulf, Limit distribution of class numbers for discriminants in progressions and fundamental discriminants. *Int. J. Number Theory* **12** (2016), no. 5, 1237–1258.
- [146] R. Rhodes and V. Vargas, Gaussian multiplicative chaos and applications: a review. *Probab. Surv.* **11** (2014), 315–392.
- [147] D. E. Rohrlich, On L -functions of elliptic curves and cyclotomic towers. *Invent. Math.* **75** (1984), no. 3, 409–423.
- [148] Z. Rudnick and P. Sarnak, Zeros of principal L -functions and random matrix theory. *Duke Math. J.* **81** (1996), no. 2, 269–322.
- [149] Z. Rudnick and K. Soundararajan, Lower bounds for moments of L -functions. *Proc. Natl. Acad. Sci. USA* **102** (2005), no. 19, 6837–6838.
- [150] 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*, pp. 293–303, Proc. Sympos. Pure Math. 75, Amer. Math. Soc., Providence, RI, 2006.
- [151] E. Saksman and C. Webb, The Riemann zeta function and Gaussian multiplicative chaos: statistics on the critical line. *Ann. Probab.* **48** (2020), no. 6, 2680–2754.
- [152] P. Sarnak, Class numbers of indefinite binary quadratic forms. *J. Number Theory* **15** (1982), no. 2, 229–247.
- [153] W. Sawin, A representation theory approach to integral moments of L -functions over function fields. *Algebra Number Theory* **14** (2020), no. 4, 867–906.
- [154] A. Selberg, Remarks on a multiple integral. *Norsk Mat. Tidsskr.* **26** (1944), 71–78.
- [155] A. Selberg, Contributions to the theory of the Riemann zeta-function. *Arch. Math. Naturvid.* **48** (1946), no. 5, 89–155.
- [156] A. Selberg, Old and new conjectures and results about a class of Dirichlet series. In *Proceedings of the Amalfi conference on analytic number theory (Maiori, 1989)*, pp. 367–385, Univ. Salerno, Salerno, 1992.
- [157] Q. Shen, The fourth moment of quadratic Dirichlet L -functions. *Math. Z.* **298** (2021), no. 1–2, 713–745.
- [158] A. Shubin, Variance estimates in Linnik’s problem. 2021, arXiv:2108.00726.
- [159] H. C. Siu, *Value distribution of automorphic forms in a family*. Ph.D. thesis, Stanford University, 2016.
- [160] A. Smith, 2^∞ -Selmer groups, 2^∞ -class groups, and Goldfeld’s conjecture. 2017, arXiv:1702.02325.

- [161] K. Soundararajan, Nonvanishing of quadratic Dirichlet L -functions at $s = \frac{1}{2}$. *Ann. of Math. (2)* **152** (2000), no. 2, 447–488.
- [162] K. Soundararajan, The number of imaginary quadratic fields with a given class number. *Hardy-Ramanujan J.* **30** (2007), 13–18.
- [163] K. Soundararajan, Extreme values of zeta and L -functions. *Math. Ann.* **342** (2008), no. 2, 467–486.
- [164] K. Soundararajan, Moments of the Riemann zeta function. *Ann. of Math. (2)* **170** (2009), no. 2, 981–993.
- [165] K. Soundararajan, Value distribution of L -functions. In *Automorphic forms and arithmetic*, pp. 2528–2529, Oberwolfach Rep. 40, EMS Press, 2017.
- [166] K. Soundararajan and M. P. Young, The second moment of quadratic twists of modular L -functions. *J. Eur. Math. Soc. (JEMS)* **12** (2010), no. 5, 1097–1116.
- [167] K. Soundararajan and A. Zaman, A model problem for multiplicative chaos in number theory. 2021, arXiv:2108.07264.
- [168] P. J. Stephens, Optimizing the size of $L(1, \chi)$. *Proc. Lond. Math. Soc. (3)* **24** (1972), 1–14.
- [169] T. Tao, The Elliott–Halberstam conjecture implies the Vinogradov least quadratic nonresidue conjecture. *Algebra Number Theory* **9** (2015), no. 4, 1005–1034.
- [170] E. C. Titchmarsh, *The theory of the Riemann zeta-function. Second edn.* The Clarendon Press Oxford University Press, New York, 1986.
- [171] S. M. Voronin, A theorem on the distribution of values of the Riemann zeta-function. *Dokl. Akad. Nauk SSSR* **221** (1975), no. 4, 771.
- [172] S. M. Voronin, Lower bounds in Riemann zeta-function theory. *Izv. Akad. Nauk SSSR Ser. Mat.* **52** (1988), no. 4, 882–892, 896.
- [173] J.-L. Waldspurger, Sur les coefficients de Fourier des formes modulaires de poids demi-entier. *J. Math. Pures Appl. (9)* **60** (1981), no. 4, 375–484.
- [174] M. P. Young, The fourth moment of Dirichlet L -functions. *Ann. of Math. (2)* **173** (2011), no. 1, 1–50.
- [175] M. P. Young, The third moment of quadratic Dirichlet L -functions. *Selecta Math. (N.S.)* **19** (2013), no. 2, 509–543.
- [176] P. Zenz, Sharp bound for the fourth moment of holomorphic Hecke cusp forms. 2021, arXiv:2108.13868.

KANNAN SOUNDARARAJAN

Department of Mathematics, Stanford University, Stanford, CA, USA,
ksound@stanford.edu

