

Finite Size Effects for Spacing Distributions in Random Matrix Theory: Circular Ensembles and Riemann Zeros

By Folkmar Bornemann, Peter J. Forrester, and Anthony Mays

According to Dyson's threefold way, from the viewpoint of global time reversal symmetry, there are three circular ensembles of unitary random matrices relevant to the study of chaotic spectra in quantum mechanics. These are the circular orthogonal, unitary, and symplectic ensembles, denoted COE, CUE, and CSE, respectively. For each of these three ensembles and their thinned versions, whereby each eigenvalue is deleted independently with probability $1 - \xi$, we take up the problem of calculating the first two terms in the scaled large N expansion of the spacing distributions. It is well known that the leading term admits a characterization in terms of both Fredholm determinants and Painlevé transcendents. We show that modifications of these characterizations also remain valid for the next to leading term, and that they provide schemes for high precision numerical computations. In the case of the CUE, there is an application to the analysis of Odlyzko's data set for the Riemann zeros, and in that case, some further statistics are similarly analyzed.

1. Introduction

Random matrix theory is of importance for both its conceptual and its predictive powers. At a conceptual level, there is the threefold way

Address for correspondence: A. Mays, Department of Mathematics and Statistics, ARC Centre of Excellence for Mathematical, and Statistical Frontiers, The University of Melbourne, Victoria 3010, Australia; e-mail: anthony.mays@unimelb.edu.au

classification of global time reversal symmetries of Hamiltonians for chaotic quantum systems [1], later extended to 10 by the inclusion of chiral and topological symmetries [2]. For each of the 10 Hamiltonians, there is an ensemble of random Hermitian matrices, which in turn are the Hermitian part of the 10 matrix Lie algebras associated with the infinite families of symmetric spaces. The symmetric spaces themselves give rise to 10 ensembles of random unitary matrices [3].

To see how the formulation of ensembles leads to a predictive statement, consider for definiteness the threefold way classification of Hamiltonians for chaotic quantum systems. The basic hypothesis, initiated by Wigner [4] and Dyson [1], and refined by Bohigas et al. [5], is that a quantum mechanical system for which the underlying classical mechanics is chaotic, will, for large energies have the same statistical distribution as the bulk scaled eigenvalues of the relevant ensemble of Hermitian matrices. The particular ensemble is determined by the presence or absence of a time reversal symmetry, with the former consisting of two subcases depending on the time reversal operator T being such that $T^2 = 1$, or $T^2 = -1$; see, for example [6, chapter 1]. This becomes predictive upon the specification of the statistical distributions for the bulk scaled eigenvalues of the random matrix ensemble. Thus, according to the hypothesis, these same distributions will be exhibited by the point process formed from the highly excited energy levels of a chaotic quantum system. The point process can be realized in laboratory experiments of nuclear excitations (see, e.g., [7]), and microwave billiards [8], among other examples.

Arguably the most spectacular example of both the conceptual and predictive powers of random matrix theory shows itself in relation to the Riemann zeta function $\zeta(s)$. The celebrated Riemann hypothesis asserts that all the complex zeros of $\zeta(s)$ are of the form $s = \frac{1}{2} \pm iE$, with $E > 0$. The Montgomery–Odlyzko law asserts that for large E these zeros—referred to as the Riemann zeros—have the same statistical properties as the bulk eigenvalues of large Hermitian random matrices with complex elements. However, from the threefold way, the latter is the ensemble of random Hermitian matrices giving the statistical properties of the large energy levels of a chaotic quantum mechanical system without a time reversal symmetry. The implied relationship between the Riemann zeros and quantum chaos is consistent with and extends the Hilbert–Pólya conjecture [9] asserting that the Riemann zeros correspond to the eigenvalues of some (unknown) unbounded self-adjoint operator; see [10, 11] for contemporary research on this topic.

To realize the predictive consequences of this link between two seemingly unrelated quantities—the Riemann zeros and chaotic quantum Hamiltonians without time reversal symmetry—requires a list of the Riemann zeros for large E . Such a list has been provided by Odlyzko [12], who in a celebrated

numerical computation has provided the high precision evaluation of the 10^{20} -th Riemann zero and over 70 million of its neighbors. From this data set, the veracity of the Montgomery–Odlyzko law can be tested. Primary statistical quantities for this purpose include the two-point correlation function $\langle \rho_{(2)}(x, x + s) \rangle$ —which measures the density of zeros a distance s from x , with x itself averaged over a window of zeros which itself is large, but still small relative to x —and the spacing distribution $p(k; s)$ for the event that the k -th next (in consecutive order) neighbors¹ are a distance s apart. As displayed in [12], at a graphical level, the agreement between the Riemann zero data and the predictions of random matrix theory is seemingly exact.

Although zeros of order 10^{20} are huge on an absolute scale, it turns out that convergence to random matrix predictions occurs on a scale where $\log E$ plays the role of N , giving rise to potentially significant finite size effects. Such effects were first considered in the work of Keating and Snaith [13] in their study of the statistical properties of the value distribution of the logarithm of the Riemann zeta function on $\text{Re } s = 1/2$ with $|s|$ large. Quite unexpectedly, they showed that the finite size corrections could be understood by introducing as a model for the Riemann zeros the eigenvalues of Haar distributed unitary random matrices from $U(N)$, instead of complex Hermitian random matrices. It had been known since the work of Dyson [14] that the eigenvalues for Haar distributed unitary random matrices from $U(N)$ and complex Hermitian random matrices with Gaussian entries have the same limiting statistical distribution. However, for finite N , they are fundamentally different, with the former being rotationally invariant while the spectral density of the latter is given by the Wigner semicircle law and thus has edge effects.

Odlyzko’s work on the computation of large Riemann zeros is ongoing. In [15], he announced a data set beginning with the 10^{23} -rd zero and its next 10^9 neighbors. Taking advantage of the great statistical accuracy offered by this data set, the finite size correction for the deviation of the empirical nearest neighbor spacing distribution and the limiting random matrix prediction was displayed, and shown to have structure. To explain the functional form of this correction, in keeping with the work of Keating and Snaith, Bogomolny and collaborators [16] were able to present analytic evidence that the correction term could again be understood in terms of a model of eigenvalues from Haar distributed $U(N)$ matrices.

More explicitly, let $p^{U(N)}(0; s)$ refer to the spacing distribution between consecutive eigenvalues for Haar distributed unitary random matrices, with the angles of the eigenvalues rescaled to have mean spacing unity. The study

¹In an increasingly ordered list the k -th next neighbor of the item x_j is item x_{j+k+1} .

[16] calls for the functional form of $r_2(0; s)$ in the large N expansion

$$p^{U(N)}(0; s) = p_2(0; s) + \frac{1}{N^2}r_2(0; s) + \dots \tag{1}$$

Here the subscript “2” on the RHS is the label from Dyson’s threefold way in the case of no time reversal symmetry. As a straightforward application of the pioneering work of Mehta [17] and Gaudin [18], Dyson [14] showed that

$$p_2(0; s) = \frac{d^2}{ds^2} \det(\mathbb{I} - \mathbb{K}_s), \tag{2}$$

where \mathbb{K}_s is the integral operator on $(0, s)$ with kernel

$$K(x, y) = \frac{\sin \pi(x - y)}{\pi(x - y)}. \tag{3}$$

Twenty years later, this Fredholm determinant evaluation was put in the context of Painlevé theory by the Kyoto school of Jimbo et al. [19], who showed

$$\det(\mathbb{I} - \xi \mathbb{K}_s) = \exp \int_0^{\pi s} \frac{\sigma^{(0)}(t; \xi)}{t} dt, \tag{4}$$

where σ satisfies the differential equation (an example of the Painlevé V equation in sigma form)

$$(t\sigma'')^2 + 4(t\sigma' - \sigma)(t\sigma' - \sigma + (\sigma')^2) = 0 \tag{5}$$

with small t boundary conditions

$$\sigma^{(0)}(t; \xi) = -\frac{\xi}{\pi}t - \frac{\xi^2}{\pi^2}t^2 + O(t^3). \tag{6}$$

Note that the parameter ξ introduced in (4) only enters in the characterization through the boundary condition; we refer to [6, chapter 8] for background theory relating to the Painlevé equations as they occur in random matrix theory.

In [16], the leading correction term $r_2(0; s)$ was estimated using a numerical extrapolation from the tabulation of $p^{U(N)}(0; s) - p_2(0; s)$. In [20] the task of finding analytic forms for $r_2(0; s)$ was addressed. One first notes that analogous to the formula (2), for finite N ,

$$p^{U(N)}(0; s) = \frac{d^2}{ds^2} \det(\mathbb{I} - \mathbb{K}_s^N), \tag{7}$$

where \mathbb{K}_s^N is the integral operator on $(0, s)$ with kernel

$$K^N(x, y) = \frac{\sin \pi(x - y)}{N \sin(\pi(x - y)/N)}. \tag{8}$$

Expanding now for large N gives

$$K^N(x, y) = K(x, y) + \frac{1}{N^2}L(x, y) + O\left(\frac{1}{N^4}\right), \tag{9}$$

where the leading correction term in this expansion is given explicitly by

$$L(x, y) = (\pi(x - y)/6) \sin \pi(x - y). \tag{10}$$

It was observed in [20] (cf. Lemma 1 below) that this expansion lifts to

$$\det(\mathbb{I} - \mathbb{K}_s^N) = \det(\mathbb{I} - \mathbb{K}_s) + \frac{1}{N^2}\Omega(\mathbb{K}_s) : \mathbb{L}_s + O\left(\frac{1}{N^4}\right), \tag{11}$$

where \mathbb{K}_s is as in (3) and \mathbb{L}_s is the integral operator on $(0, s)$ with kernel L ; the leading correction is now given by the operator expression

$$\Omega(\mathbb{K}_s) : \mathbb{L}_s = -\det(\mathbb{I} - \mathbb{K}_s)\text{tr}((\mathbb{I} - \mathbb{K}_s)^{-1}\mathbb{L}_s), \tag{12}$$

which is *linear* in \mathbb{L} . Substituting in (7) and comparing to (1) we get that

$$r_2(0; s) = \frac{d^2}{ds^2}\Omega(\mathbb{K}_s) : \mathbb{L}_s. \tag{13}$$

In Section 2.2 of this paper, the problem addressed is how to use such formulas to provide a high precision numerical tabulation of $r_2(k; s)$. In the case $k = 1$, we exhibit the resulting functional form in Odlyzko’s data set for the Riemann zeros; see Section 4.

An expression for the large N expansion of the Fredholm determinant in (7) involving Painlevé transcendents was also given in [20]. Thus, it was shown that

$$\begin{aligned} &\det(\mathbb{I} - \xi\mathbb{K}_s^N) \\ &= \exp\left(\int_0^{\pi s} \frac{\sigma^{(0)}(t; \xi)}{t} dt\right) \left(1 + \frac{1}{N^2} \int_0^{\pi s} \frac{\sigma^{(1)}(t; \xi)}{t} dt + O\left(\frac{1}{N^4}\right)\right). \end{aligned} \tag{14}$$

As noted above, $\sigma^{(0)}(t; \xi)$ satisfies the particular Painlevé V equation in sigma form (5) with boundary condition (6), with only the latter depending on ξ . The function $\sigma^{(1)}(t; \xi)$ was characterized as the solution of the second-order, linear differential equation

$$A(t; \xi)y'' + B(t; \xi)y' + C(t; \xi)y = D(t; \xi), \tag{15}$$

where, with $\sigma = \sigma^{(0)}(t; \xi)$,

$$A(t; \xi) = 2t^2\sigma'',$$

$$B(t; \xi) = -8\sigma'\sigma + 12t(\sigma')^2 + 8t(t\sigma' - \sigma),$$

$$C(t; \xi) = -4(\sigma')^2 - 8(t\sigma' - \sigma),$$

$$D(t; \xi) = -\frac{4}{3}t^2\sigma'' \left(\sigma - t\sigma' - \frac{t^2}{2}\sigma'' \right) - \frac{4}{3}(t\sigma' - \sigma)(3\sigma^2 + 2t\sigma(t - \sigma') - 2t^2\sigma'(t + \sigma')),$$

and fulfilling the $t \rightarrow 0^+$ boundary condition

$$\sigma^{(1)}(t; \xi) = - \left(t^4 \frac{\xi^2}{9\pi^2} + t^5 \frac{5\xi^3}{36\pi^3} + O(t^6) \right). \tag{16}$$

Substituting (14) in (7) and comparing to (1) we read off that

$$r_2(0; s) = \frac{d^2}{ds^2} \exp \left(\int_0^{\pi s} \frac{\sigma^{(0)}(t; 1)}{t} dt \right) \int_0^{\pi s} \frac{\sigma^{(1)}(t; 1)}{t} dt. \tag{17}$$

By using a nested power series method to solve the differential equation (15) numerically, this expression was used in [20] to determine the graphical form of $r_2(0; s)$, and also to determine its small and large s expansions. However, the operator approach advocated in this paper is simpler and more straightforward to use numerically.

Random unitary matrices from $U(N)$ chosen with Haar measure form one of Dyson’s three circular ensembles of unitary random matrices coming from the considerations of the threefold way, and in this context is referred to as the circular unitary ensemble (CUE). The other two circular ensembles are the circular orthogonal ensemble (COE) of symmetric unitary matrices and the circular symplectic ensemble (CSE) of self dual unitary matrices. We will show that our methods can be adapted to provide a systematic investigation of the leading correction in the large N expansion of spacing distributions for each of these examples.

1.1. Outline of the paper

After some preparatory material on the expansion of operator determinants and numerical methods for evaluating the corresponding terms in Section 2, we begin in Section 3 by extending the study initiated in [20] on this problem as it applies to CUE matrices. The first question addressed is the computation of the leading order correction term in the large N expansion of $p^{U(N)}(k; s) - p_2(k; s)$, or equivalently $p^{CUE_N}(k; s) - p_2(k; s)$, in terms of an integral operator formula extending (13). We then do the same in the circumstance that each eigenvalue has been deleted independently at random with probability $(1 - \xi)$. Such a thinning procedure, well known in the theory of point processes (see, e.g., [21]), was introduced into random matrix theory by Bohigas and Pato [22], and has been applied to Odlyzko’s data set for the Riemann zeros in [20]. Continuing with the point process perspective, next we consider the setting of a translationally

invariant one-dimensional point process, normalized to have average spacing unity, superimposed with a Poisson point process. Our interest is in the distribution of the minimum distance from each Poisson point to a point in the original process. In an ensemble setting, this is equivalent to computing the minimum distance from the origin of a point in the original process. For CUE matrices, the distribution can be expressed in a form analogous to (7), which allows the statistic to be expanded for large N . In the final subsection of Section 3, the statistic for the nearest neighbor spacing, that is, the minimum of the spacing distance between left and right neighbors in the CUE is similarly studied.

In Section 4, our results are compared against empirical findings from Odlyzko's data set for the Riemann zeros. There we will use a unique $O(N^{-3})$ correction of the CUE correlation kernel that was observed in the study [16] to match the corresponding $O(N^{-3})$ terms of the two- and three-point correlation functions. Including this term in the numerical calculations systematically improves the fit to the empirical data.

The study of spacing distributions for matrices from the COE and CSE is more involved than for CUE matrices. The reason is that the latter is based on the single integral operator \mathbb{K}_s^N with kernel (8), whereas the pathway to tractable expressions in the case of the COE and CSE makes use of inter-relations between gap probabilities. The necessary theory is covered in the first part of Section 5. In the second part of Section 5, we give the leading correction term for the large N expansion of

$$p_\xi^{\text{C}\beta\text{E}}(k; s) - p_{\beta, \xi}(k; s),$$

where $\beta = 1$ for the COE and $\beta = 4$ for the CSE, in terms of a characterization analogous to (14) and (15). The parameter ξ specifies the thinning probability.

1.2. A note on sampling sizes of empirical data

In this paper, we discuss finite size effects up to an error $O(N^{-4})$, with N in the Odlyzko data set of Riemann zeros being effectively $N \approx 10$. We choose the same order of magnitude of N for the simulations of the circular ensembles. Now, by the law of large numbers, sampling is known to introduce a statistical error of about $O(M^{-1/2})$, where M denotes the sample size. Hence, pushing the sampling error to the same order of magnitude as the remaining finite size error thus requires a sampling size of $M = N^8 \approx 10^8$. This was actually the choice for our simulations and is well matched by the Odlyzko data set of a little more than $10^9 \approx 10^8 \cdot N$ zeros. However, observing structure also in the $O(N^{-4})$ remainder term would hence require to increase the sampling size by at least two to four orders of magnitude. This is only possible with parallel processing and a subtle

memory management of the resulting gigantic data sets (the raw material provided by Odlyzko is already about 22 GB in size).

2. Expansions of determinants and their numerical evaluation

2.1. Expansions of operator determinants

The integral operator formulas of this paper are based on the following *folklore* lemma, which we prove here for ease of reference.

LEMMA 1. *Let J be a bounded interval and let $K_h(x, y)$, $K(x, y)$, $L(x, y)$ be continuously differentiable kernels on $J \times J$. If the expansion*

$$K_h(x, y) = K(x, y) + hL(x, y) + O(h^2)$$

holds uniformly up to the first derivatives as $h \rightarrow 0$, then it lifts as an expansion $\mathbb{K}_h = \mathbb{K} + h\mathbb{L} + O(h^2)$ of the induced integral operators on $L^2(J)$ in trace-class norm. If $1 \notin \text{Spec}(\mathbb{K})$, the operator determinant expands as

$$\det(\mathbb{I} - \mathbb{K}_h) = \det(\mathbb{I} - \mathbb{K}) + h\Omega(\mathbb{K}) : \mathbb{L} + O(h^2) \tag{18a}$$

with the leading correction term given by²

$$\Omega(\mathbb{K}) : \mathbb{L} = -\det(\mathbb{I} - \mathbb{K})\text{tr}((\mathbb{I} - \mathbb{K})^{-1}\mathbb{L}), \tag{18b}$$

an expression that is linear in \mathbb{L} .

Proof: On bounded intervals J , continuously differentiable kernels $K(x, y)$ induce integral operators \mathbb{K} on $L^2(J)$ that are trace-class: using the fundamental theorem of calculus, one can represent them straightforwardly as the sum of a rank-one operator and a product of two Hilbert–Schmidt operators, see, for example, [23, p. 879]. By the same construction the expansion lifts from kernels in C^1 -norm to operators in trace-class norm. Now, using the multiplicativity of the operator determinant we get that

$$\det(\mathbb{I} - \mathbb{K}_h) = \det(\mathbb{I} - \mathbb{K}) \det(\mathbb{I} - h(\mathbb{I} - \mathbb{K})^{-1}\mathbb{L} + O(h^2)).$$

Truncating the series definition of the operator determinant in terms of exterior products, see, for example, [24, eq. (3.5)],

$$\det(\mathbb{I} + z\mathbb{T}) = 1 + \sum_{n=1}^{\infty} z^n \text{tr} \bigwedge^n \mathbb{T},$$

²Note that $\Omega(\mathbb{K}) = -\det(\mathbb{I} - \mathbb{K})(\mathbb{I} - \mathbb{K})^{-1}$ is the Fréchet derivative of the map $\mathbb{K} \mapsto \det(\mathbb{I} - \mathbb{K})$ under the identification of the dual of the space of trace-class operators with the space of bounded operators, see [24, corollary 5.2]. The derivative of that map in the direction of the bounded operator \mathbb{L} is thus given by $\Omega(\mathbb{K}) : \mathbb{L} = \text{tr}(\Omega(\mathbb{K})\mathbb{L})$. By analytic continuation this interpretation of the expression $\Omega(\mathbb{K}) : \mathbb{L}$ extends well to the case that $1 \in \text{Spec}(\mathbb{K})$.

at the second term, we obtain

$$\det(\mathbb{I} - h(\mathbb{I} - \mathbb{K})^{-1}\mathbb{L} + O(h^2)) = 1 - h\text{tr}((\mathbb{I} - \mathbb{K})^{-1}\mathbb{L}) + O(h^2),$$

which completes the proof. ■

2.2. Numerical evaluation of operator terms

The computations for all figures of this paper are based on the numerical evaluation of operator terms such as $\det(\mathbb{I} - \mathbb{K})$ or $\Omega(\mathbb{K}) : \mathbb{L}$. In fact, there is an extension of the numerical method introduced in [23] for the operator determinant that facilitates the efficient highly accurate evaluation of such terms with exponential convergence³ if the kernels extend analytically into the complex domain. Given a quadrature method

$$\int_J f(x) dx \approx \sum_{j=1}^n f(x_j)w_j$$

with *positive* weights w_j , we associate to an integral operator \mathbb{K} on $L^2(J)$ with the kernel $K(x, y)$ the *Nyström matrix*

$$\mathbb{K}_w = (K(x_j, x_k)w_k)_{j,k=1}^n.$$

For compact intervals J the best choice for smooth f is Gauss–Legendre quadrature, for which recently a super-fast algorithm of $O(n)$ complexity has been found [25]; it converges exponentially fast if f extends analytically into the complex domain.

As was shown in [23, theorem 6.2], these convergence properties are inherited by the determinant of the Nyström matrix as an approximation of the operator determinant, namely,

$$\det(\mathbb{I} - \mathbb{K}) \approx \det(\mathbb{I} - \mathbb{K}_w),$$

where \mathbb{I} denotes the identity operator in $L^2(J)$ on the left and the $n \times n$ identity matrix on the right. In particular, the convergence is exponential if the kernel K extends analytically into the complex domain. Now, the same method of replacing the integral operator by the associated Nyström matrix applies to the approximation of quite general operator terms.

We give four examples:

- The trace is approximated straightforwardly by the quadrature formula as

$$\text{tr}(\mathbb{K}) = \int_J K(x, x) dx \approx \sum_{j=1}^n K(x_j, x_j)w_j = \text{tr}(\mathbb{K}_w).$$

³There is a constant $c > 0$ such that the approximation error is $O(e^{-cn})$, where n denotes the dimension parameter of the method (e.g., the number of quadrature points).

- Operator products are approximated by matrix products, that is,

$$(\mathbb{K}\mathbb{L})_w \approx \mathbb{K}_w \mathbb{L}_w.$$

This follows from

$$\begin{aligned} ((\mathbb{K}\mathbb{L})_w)_{jk} &= w_k \int_J K(x_j, x') L(x', x_k) dx' \\ &\approx \sum_{l=1}^n K(x_j, x_l) w_l L(x_l, x_k) w_k = \sum_{l=1}^n (\mathbb{K}_w)_{jl} (\mathbb{L}_w)_{lk}. \end{aligned}$$

- The resolvent kernel $R(x, x')$ of the operator $(\mathbb{I} - \mathbb{K})^{-1}$ satisfies the integral equation

$$u(x) = \int_J R(x, x') u(x') dx' - \int_J \int_J K(x, z) R(z, x') u(x') dx' dz$$

for all u . Application of the quadrature rule gives with $u_j = u(x_j)$

$$u_j \approx \sum_{k=1}^n R(x_j, x_k) w_k \cdot u_k - \sum_{l,k=1}^n K(x_j, x_l) w_l \cdot R(x_l, x_k) w_k \cdot u_k,$$

that is, $\mathbb{I} \approx (\mathbb{I} - \mathbb{K}_w)(\mathbb{I} - \mathbb{K})^{-1}_w$ and, therefore,

$$((\mathbb{I} - \mathbb{K})^{-1})_w \approx (\mathbb{I} - \mathbb{K}_w)^{-1}.$$

- By combining all the examples so far, we get

$$\begin{aligned} \Omega(\mathbb{K}) : \mathbb{L} &= -\det(\mathbb{I} - \mathbb{K}) \text{tr}((\mathbb{I} - \mathbb{K})^{-1} \mathbb{L}) \\ &\approx -\det(\mathbb{I} - \mathbb{K}_w) \text{tr}((\mathbb{I} - \mathbb{K}_w)^{-1} \mathbb{L}_w) = \Omega(\mathbb{K}_w) : \mathbb{L}_w. \end{aligned}$$

In all the above, the convergence properties are inherited from the underlying quadrature formula; in particular, the convergence is exponential if the kernels extend analytically to the complex domain. The proof is straightforward in the first two examples and follows from the theory of the Nyström method for integral equations in the third one; the fourth is a combination of all the previous results.

The numerical derivatives with respect to the s and z variables of the generating functions for the spacing distributions are computed in exactly the same fashion as discussed in [26], that is, based on Chebyshev expansions with respect to s and contour integration with respect to z .

The actual implementation and use of this methodology within the Matlab toolbox of [26] is described in the Appendix of this paper.

3. Spacing distributions for the CUE

3.1. Preliminaries

The sequence of angles, to be referred to as eigenangles, specifying the eigenvalues of a unitary random matrix from any of the three circular ensembles forms a point process on $(-\pi, \pi]$ with uniform density $N/2\pi$. Due to the rotational symmetry, the spacing distributions of the angles, $p^{C\beta E}(k; s)$ with $\beta = 1, 2, 4$ for the COE, CUE, CSE, respectively, and $k = 0, 1, \dots$, are thus independent of the absolute location of the eigenvalues. If there is an eigenangle at x and at $x + s$, we may rotate all the angles so that x becomes the origin, and speak of a spacing of size s . Let us do this, and also normalize the angles so that the mean spacing is unity.

Fundamental to the study of spacing distributions $p^{C\beta E}(k; s)$ are the so-called gap probabilities $E^{C\beta E}(l; s)$. The latter specify the probability that an interval of size s contains exactly l eigenangles for $C\beta E$ matrices. Specifically, let us introduce the generating functions

$$p^{C\beta E}(s; z) := \sum_{n=0}^N (1 - z)^n p^{C\beta E}(n; s), \tag{19}$$

$$E^{C\beta E}(s; z) := \sum_{n=0}^N (1 - z)^n E^{C\beta E}(n; s). \tag{20}$$

Of course the naming on the LHS's represent an abuse of notation, due to the same functional forms also appearing on the RHS; however, the appearance of the generating function symbol z is enough to distinguish the different quantities. We have the relationship between generating functions (see, e.g., [6, proposition 8.1])

$$p^{C\beta E}(s; z) = \frac{1}{z^2} \frac{d^2}{ds^2} E^{C\beta E}(s; z), \tag{21}$$

or equivalently the formula

$$p^{C\beta E}(n; s) = \frac{d^2}{ds^2} \sum_{j=0}^n (n - j + 1) E^{C\beta E}(j; s). \tag{22}$$

Also fundamental is the relationship between the generating function (20) and the k -point correlation functions $\rho_{(n)}^{C\beta E}(x_1, \dots, x_n)$. Thus, we have (see, e.g., [6, proposition 9.1.1])

$$E^{C\beta E}(s; z) = 1 + \sum_{n=1}^N \frac{(-z)^n}{n!} \int_0^s dx_1 \cdots \int_0^s dx_n \rho_{(n)}^{C\beta E}(x_1, \dots, x_n). \tag{23}$$

The eigenvalues for the CUE form a determinantal point process, while the eigenvalues for the COE and CSE form Pfaffian point processes; see, for example, [6, chapters 5 and 6]. Specifically, in the case of the CUE, we have

$$\rho_{(n)}^{\text{CUE}}(x_1, \dots, x_n) = \det [K^N(x_j, x_l)]_{j,l=1,\dots,k}, \tag{24}$$

where K^N is given by (8), while the analogous formula for the COE and CSE involves a Pfaffian of a 2×2 antisymmetric kernel function. This structural difference distinguishes the case of the CUE and thus makes it simpler.

3.2. Integral operator formulas

Substituting (24) in (23), and making use of a standard expansion formula in the theory of Fredholm integral operators (see, e.g., [24, eq. (3.14)]) gives

$$E^{\text{CUE}}(s; z) = \det(\mathbb{I} - z\mathbb{K}_s^N). \tag{25}$$

Now substituting the expansion (11) in (25) and substituting the result in (21) we read off the large N expansion to second order.

PROPOSITION 1. *We have*

$$p^{\text{CUE}}(s; z) = p_2(s; z) + \frac{1}{N^2}r_2(s; z) + \mathcal{O}\left(\frac{1}{N^4}\right), \tag{26}$$

where

$$p_2(s; z) = \frac{1}{z^2} \frac{d^2}{ds^2} \det(\mathbb{I} - z\mathbb{K}_s) \tag{27}$$

and

$$r_2(s; z) = \frac{1}{z} \frac{d^2}{ds^2} \Omega(z\mathbb{K}_s) : \mathbb{L}_s. \tag{28}$$

In particular

$$p^{\text{CUE}}(0; s) = p_2(s; z) \Big|_{z=1} + \frac{1}{N^2}r_2(s; z) \Big|_{z=1} + \mathcal{O}\left(\frac{1}{N^4}\right), \tag{29}$$

and

$$p^{\text{CUE}}(1; s) = -\frac{d}{dz} p_2(s; z) \Big|_{z=1} - \frac{1}{N^2} \frac{d}{dz} r_2(s; z) \Big|_{z=1} + \mathcal{O}\left(\frac{1}{N^4}\right). \tag{30}$$

3.3. Thinning

The operation of thinning applied to a point process refers to the procedure of independently deleting each point in a sample with probability $(1 - \xi)$, $0 < \xi \leq 1$. With the k -point correlation function before thinning being given

by $\rho_{(k)}(x_1, \dots, x_k)$, the k -point correlation function after thinning is simply $\xi^k \rho_{(k)}(x_1, \dots, x_k)$, due to the independence. It thus follows from (23), that the generating function for the gap probability in the presence of thinning, $E_\xi^{\text{C}\beta\text{E}}(s; z)$ say, is related to the generating function for the gap probability without thinning by

$$E_\xi^{\text{C}\beta\text{E}}(s; z) = E^{\text{C}\beta\text{E}}(s; \xi z). \tag{31}$$

Specializing further to the CUE, we can substitute (25) in the RHS of this expression to conclude

$$E_\xi^{\text{CUE}}(s; z) = \det(\mathbb{I} - z\xi\mathbb{K}_s^N). \tag{32}$$

The analogues of (21) and (22) for the spacing probabilities are

$$p_\xi^{\text{C}\beta\text{E}}(s; z) = \frac{1}{\xi z^2} \frac{d^2}{ds^2} E^{\text{C}\beta\text{E}}(s; \xi z),$$

$$p_\xi^{\text{C}\beta\text{E}}(n; s) = \frac{1}{\xi} \frac{d^2}{ds^2} \sum_{j=0}^n (n - j + 1) E_\xi^{\text{C}\beta\text{E}}(j; s). \tag{33}$$

Using the first of these, together with (25) and the expansion (11) we can write down the analogue of Proposition 1.

PROPOSITION 2. *We have*

$$p_\xi^{\text{CUE}}(s; z) = p_{2,\xi}(s; z) + \frac{1}{N^2} r_{2,\xi}(s; z) + \mathcal{O}\left(\frac{1}{N^4}\right), \tag{34}$$

where

$$p_{2,\xi}(s; z) = \frac{1}{\xi z^2} \frac{d^2}{ds^2} \det(\mathbb{I} - \xi z\mathbb{K}_s) \tag{35}$$

and

$$r_{2,\xi}(s; z) = \frac{1}{z} \frac{d^2}{ds^2} \Omega(\xi z\mathbb{K}_s) : \mathbb{L}_s. \tag{36}$$

In particular

$$p_\xi^{\text{CUE}}(0; s) = p_{2,\xi}(s; 1) + \frac{1}{N^2} r_{2,\xi}(s; 1) + \mathcal{O}\left(\frac{1}{N^4}\right), \tag{37}$$

and

$$p_\xi^{\text{CUE}}(1; s) = -\frac{d}{dz} p_{2,\xi}(s; z) \Big|_{z=1} - \frac{1}{N^2} \frac{d}{dz} r_{2,\xi}(s; z) \Big|_{z=1} + \mathcal{O}\left(\frac{1}{N^4}\right). \tag{38}$$

Implementation of these formulas according to the method of Section 2.2 will be carried out in Section 4, in the context of a comparison with the corresponding statistics for Odlyzko’s data set of the Riemann zeros.

3.4. Minimum spacing distribution from a randomly chosen origin

The quantity $E^{C\beta E}(0; (-s, s))$ is the probability that there are no eigenvalues in the interval $(-s, s)$ of the scaled $C\beta E$. Differentiating this quantity with respect to s gives the probability density function $p_{\text{Origin}}^{C\beta E}(s)$ for the scaled eigenangle closest to the origin,

$$p_{\text{Origin}}^{C\beta E}(s) = -\frac{d}{ds} E^{C\beta E}(0; (-s, s)).$$

In the case $\beta = 2$ with thinning, it follows from (32), as well as the translation invariance of the sine kernel, that

$$p_{\text{Origin},\xi}^{\text{CUE}}(s) = -\frac{d}{ds} \det(\mathbb{I} - \xi \mathbb{K}_{2s}^N). \tag{39}$$

Use of (14) gives a characterization of the leading two terms of the large N expansion of this quantity.

PROPOSITION 3. *We have*

$$p_{\text{Origin},\xi}^{\text{CUE}}(s) = p_{\text{Origin},\xi}(s) + \frac{1}{N^2} r_{\text{Origin},\xi}(s) + O\left(\frac{1}{N^4}\right), \tag{40}$$

where in terms of integral operators

$$\begin{aligned} p_{\text{Origin},\xi}(s) &= -\frac{d}{ds} \det(\mathbb{I} - \xi \mathbb{K}_{2s}), \\ r_{\text{Origin},\xi}(s) &= -\xi \frac{d}{ds} \Omega(\xi \mathbb{K}_{2s}) : \mathbb{L}_{2s}. \end{aligned} \tag{41}$$

Alternatively, in terms of $\sigma^{(0)}$ and $\sigma^{(1)}$ defined in (14)

$$p_{\text{Origin},\xi}(s) = -\frac{d}{ds} \exp\left(\int_0^{2\pi s} \frac{\sigma^{(0)}(t; \xi)}{t} dt\right) \tag{42}$$

and

$$r_{\text{Origin},\xi}(s) = -\frac{d}{ds} \left(\int_0^{2\pi s} \frac{\sigma^{(1)}(t; \xi)}{t} dt\right) \exp\left(\int_0^{2\pi s} \frac{\sigma^{(0)}(t; \xi)}{t} dt\right). \tag{43}$$

To simulate this quantity in translationally invariant empirical data such as the CUE itself, one can actually draw the origin uniformly and independently within the interval for which the data are defined (i.e., one superimposes a Poisson point process defining the origin). To prepare for corresponding computations within a large set of Riemann zeros, this was done, as a proof of concept, for CUE matrices in Figs. 1 and 2.

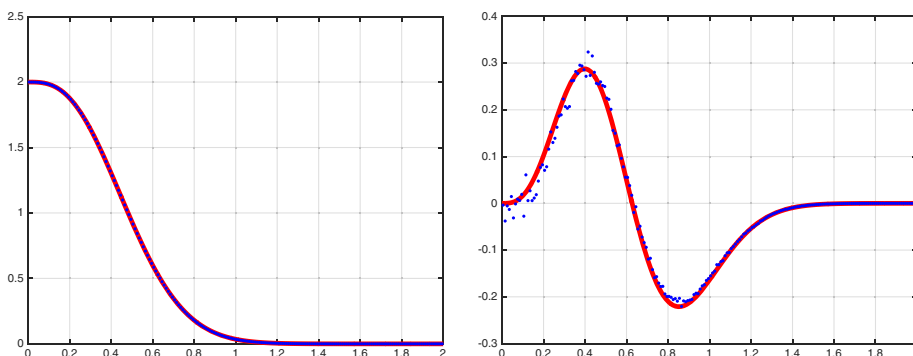


Figure 1. Minimum spacing from a randomly chosen origin: simulation versus theory for finite size CUE, no thinning ($\xi = 1$). Left panel: a histogram of empirical data from CUE_N with $N = 20$ scaled to unit mean spacing, computed using a bin size of 0.01 and 10^8 samples from CUE_N , for each of which 1,000 samples of a uniformly distributed random origin were drawn (blue dots); the large N limit $p_{\text{Origin},\xi}(s)$ (red solid line). Right panel: the simulation data minus $p_{\text{Origin},\xi}(s)$ scaled by N^2 (blue dots); the leading correction term $r_{\text{Origin},\xi}(s)$ (red solid line).

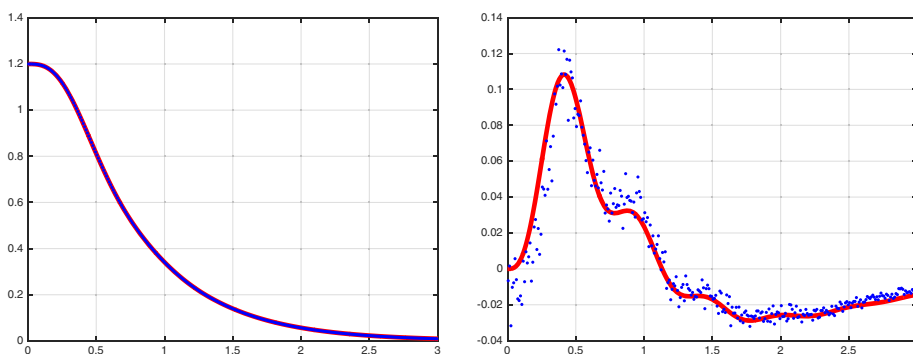


Figure 2. As in Fig. 1 but with thinning $\xi = 0.6$.

3.5. Nearest neighbor spacing distribution

A variant of the spacing distribution between consecutive eigenvalues is the spacing distribution between nearest neighbor eigenvalues [27]. For this, at each eigenvalue, one measures the smallest of the spacings to the eigenvalue immediately to the left, and the spacing immediately to the right. In a theoretical ensemble formulation, the system is to be conditioned so that there is an eigenvalue at the origin. Consider in particular the CUE with eigenangles scaled to have unit spacing. With $\rho_{(n),\xi}^{\text{CUE},\theta=0}$ denoting the n -point correlation function for the conditioned system in the presence of thinning,

we have

$$\begin{aligned} \rho_{(n),\xi}^{\text{CUE},\theta=0}(\theta_1, \dots, \theta_n) &= \xi^n \rho_{(n+1)}^{\text{CUE}}(\theta_1, \dots, \theta_n, 0) \\ &= \xi^n \det \left(\frac{\sin \pi(\theta_j - \theta_k)}{N \sin(\pi(\theta_j - \theta_k)/N)} - \frac{\sin \pi \theta_j}{N \sin \pi \theta_j / N} \frac{\sin \pi \theta_k}{N \sin \pi \theta_k / N} \right)_{j,k=1,\dots,n}, \end{aligned} \tag{44}$$

where the second line follows from (24) and (8), together with elementary row operations applied to the determinant.

Let the entry of the determinant in (44) be denoted by $\tilde{K}^N(\theta_j, \theta_k)$, and denote the corresponding integral operator supported on $(-s, s)$ by $\tilde{\mathbb{K}}_{(-s,s)}^N$. Analogous to (39), the scaled nearest neighbor spacing in the presence of thinning is then

$$p_{\text{nn},\xi}^{\text{CUE}}(s) = -\frac{d}{ds} \det \left(\mathbb{I} - \xi \tilde{\mathbb{K}}_{(-s,s)}^N \right). \tag{45}$$

By observing

$$\tilde{K}^N(x, y) = K^N(x, y) - K^N(x, 0)K^N(0, y),$$

we can read off from (45), upon expanding the kernel to order $1/N^2$, integral operator formulas for the leading two terms analogous to (41).

PROPOSITION 4. *In terms of the notation (3) and (10), let*

$$\begin{aligned} K^{\text{nn}}(x, y) &= K(x, y) - K(x, 0)K(0, y), \\ L^{\text{nn}}(x, y) &= L(x, y) - L(x, 0)K(0, y) - K(x, 0)L(0, y). \end{aligned} \tag{46}$$

Denote by $\mathbb{K}_{(-s,s)}^{\text{nn}}$ and $\mathbb{L}_{(-s,s)}^{\text{nn}}$, the integral operators on $(-s, s)$ with kernels $K^{\text{nn}}(x, y)$ and $L^{\text{nn}}(x, y)$, respectively. We have

$$p_{\text{nn},\xi}^{\text{CUE}}(s) = p_{\text{nn},\xi}(s) + \frac{1}{N^2} r_{\text{nn},\xi}(s) + \mathcal{O}\left(\frac{1}{N^4}\right), \tag{47}$$

where

$$\begin{aligned} p_{\text{nn},\xi}(s) &= -\frac{d}{ds} \det \left(\mathbb{I} - \xi \mathbb{K}_{(-s,s)}^{\text{nn}} \right), \\ r_{\text{nn},\xi}(s) &= -\xi \frac{d}{ds} \Omega \left(\xi \mathbb{K}_{(-s,s)}^{\text{nn}} \right) : \mathbb{L}_{(-s,s)}^{\text{nn}}. \end{aligned} \tag{48}$$

Remark 1. In [27] a Painlevé transcendent evaluation of $p_{\text{nn},\xi}(s)$ has been given; see also [6, §9.5.2].

4. Application to the statistics of Riemann zeros

4.1. The two-point correlation function

Starting with the 1973 seminal work of Montgomery, there have been significant developments on a deep (conjectural) connection between the statistics of the fluctuation properties of the zeros of the Riemann zeta function on the critical axis $\text{Re } s = 1/2$ and those of the eigenangles of CUE_N , supported by large-scale numerical calculations based on the extensive tables of Riemann zeros provided by Odlyzko, see, for example, [10, 12, 13, 15, 16, 20, 27] and the literature cited therein. Answering a question of Odlyzko [15], who had posed the challenge of understanding the structure in the numerical difference graph between the Riemann zeta spacing distribution for a set of zeros of large height and the (conjectured) asymptotics, finite size effects of this statistic were studied in [16]. This study gave a precise quantitative association of Riemann zeros $1/2 + iE$ at height E to a (formal) size N of the corresponding CUE_N . Using the Hardy–Littlewood conjecture of the distribution of prime pairs, Bogomonly and Keating [28] had earlier given an analytic expression for the pair correlation of the Riemann zeros. The authors of [16] expanded this for large height E , with the local density normalized to unity, to obtain

$$R_2^{\text{RZ}}(s) = 1 - \frac{\sin^2(\pi s)}{\pi^2 s^2} - \frac{\Lambda \sin^2(\pi s)}{\pi^2} \bar{\rho}^{-2} - \frac{Qs \sin(2\pi s)}{2\pi^2} \bar{\rho}^{-3} + O(\bar{\rho}^{-4}), \quad (49)$$

where

$$\bar{\rho} = \frac{\log(E/2\pi)}{2\pi}$$

denotes the smooth asymptotic density of zeros at E and Λ , Q are the following constants (with γ_n the Stieltjes constants and \sum_p summation over the primes)⁴:

$$\Lambda = \gamma_0^2 + 2\gamma_1 + \sum_p \frac{\log^2 p}{(p-1)^2} = 1.57315\ 10713\ 24955 \dots,$$

$$Q = \sum_p \frac{\log^3 p}{(p-1)^2} = 2.31584\ 63849\ 58803 \dots$$

⁴The highly accurate numerical values were obtained by applying the Euler–Maclaurin formula to the Stieltjes constants,

$$\gamma_n = \lim_{m \rightarrow \infty} \left(\sum_{k=1}^m \frac{\log^n k}{k} - \frac{\log^{n+1} m}{m+1} \right),$$

and the method of [29, p. 2] to the sums over the primes.

On the other hand, the pair correlation function

$$R_2^{\text{CUE}_N}(s) = \rho_{(2)}(0, s) = \begin{vmatrix} K^N(0, 0) & K^N(0, s) \\ K^N(s, 0) & K^N(s, s) \end{vmatrix}$$

of CUE_N , normalized to mean spacing unity, expands by (9) as

$$R_2^{\text{CUE}}(s) = \rho_{(2)}(x, x + s) = 1 - \frac{\sin^2(\pi s)}{\pi^2 s^2} - \frac{\sin^2(\pi s)}{3} N^{-2} + O(N^{-4}). \quad (50)$$

By matching the first correction terms of (49) and (50) the authors of [16] got as effective dimension N of a CUE_N at height E

$$N = \frac{\pi}{\sqrt{3\Lambda}} \bar{\rho} = \frac{1}{\sqrt{12\Lambda}} \log\left(\frac{E}{2\pi}\right).$$

This way one has the (conjectural) approximation

$$R_2^{\text{RZ}}(s) = R_2^{\text{CUE}}(s) + O(N^{-3}).$$

4.2. Exterior rescaling of the leading correction terms

As noted in [16] a rescaling of the s -variable in the leading correction term of the $\bar{\rho}$ -expansion of the two-point functions absorbs the $O(\bar{\rho}^{-3})$ term, respectively, the $O(N^{-3})$ term:

$$\begin{aligned} R_2^{\text{RZ}}(s) &= 1 - \frac{\sin^2(\pi s)}{\pi^2 s^2} - \frac{\Lambda \sin^2(\pi \alpha s)}{\pi^2} \bar{\rho}^{-2} + O(\bar{\rho}^{-4}) \\ &= 1 - \frac{\sin^2(\pi s)}{\pi^2 s^2} - \frac{\sin^2(\pi \alpha s)}{3} N^{-2} + O(N^{-4}), \end{aligned}$$

where

$$\alpha = 1 + \frac{Q}{2\pi \Lambda \bar{\rho}} = 1 + \frac{Q}{\Lambda \log(E/2\pi)} = 1 + \frac{Q}{\Lambda \sqrt{12\Lambda}} N^{-1}.$$

That is, one gets an $O(N^{-4})$ approximation of $R_2^{\text{RZ}}(s)$ by expanding $R_2^{\text{CUE}}(s)$ into powers of N^{-2} and rescaling the s -variable of the leading correction term by α .

It was suggested in [16] to apply the same procedure, that is, α -rescaling the s -variable of the leading correction term of a CUE_N -based expansion, to improve the fit even for more general statistics, such as the spacing distribution originally considered by Odlyzko. In [20], it was successfully applied to improve the fit of the CUE leading correction term to the spacing distribution in the presence of thinning.

4.3. Interior rescaling of the leading correction terms

The numerical data, physical arguments and mathematical conjectures surrounding the fluctuations of the Riemann zeros indicate that their statistics are asymptotically well approximated, to more than just leading order, by some correlation kernel K_{RZ}^N . Consequently, to justify the mechanism of the exterior rescaling introduced in the last paragraph, [16, p. 10748] suggested to pull-back the absorption of the $O(\bar{\rho}^{-3})$ term from the two-point function into the expansion of such an assumed kernel. In this paper, we will use such an expansion to improve systematically the CUE_N fit of the fluctuation properties of the Riemann zeros statistics from $O(N^{-3})$ to $O(N^{-4})$.

By modifying the expansion (9) of the CUE correlation kernel K^N to include an $O(N^{-3})$ term,

$$K_{\text{RZ}}^N(x, y) = K(x, y) + L(x, y)N^{-2} + M(x, y)N^{-3} + O(N^{-4}), \tag{51}$$

the task is thus to identify the unknown kernel $M(x, y)$ from the $O(N^{-3})$ term in (49). By translation invariance and symmetry of the kernel K_{RZ}^N and by the normalization $K_{\text{RZ}}^N(x, x) = 1$ to mean spacing unity, we get the form

$$M(x, y) = \mu(x - y)$$

of a convolution kernel with a *symmetric* function μ satisfying $\mu(0) = 0$. A brief calculation results in the corresponding two-point function

$$R_2^{\text{RZ}}(s) = 1 - \frac{\sin^2(\pi s)}{\pi^2 s^2} - \frac{\sin^2(\pi s)}{3} N^{-2} - \frac{2\mu(s) \sin(\pi s)}{\pi s} N^{-3} + O(N^{-4}).$$

Matching with (49) gives, cf. the first equality in [16, eq. (28)],

$$\mu(s) = \eta \frac{\pi^2 s^2 \cos(\pi s)}{6}, \quad \eta = \frac{Q}{\Lambda \sqrt{3\Lambda}}.$$

In the appendix of [16], it is shown that this expansion of the kernel allows one to consistently reproduce a conjectural expansion of the three-point function (which was obtained from a calculation published much later in [30]). This is further evidence that the kernel expansion (51) induces, quite generally, the $O(N^{-3})$ correction terms in the determinantal point process fluctuation statistics.

Now, the correction kernel M can be absorbed easily by an *interior* rescaling of the leading correction kernel L , namely, by introducing

$$L_{\text{RZ}}(x, y) = \frac{\pi(x - y) \sin(\pi \bar{\alpha}(x - y))}{6}, \tag{52}$$

which expands as $L_{\text{RZ}}(x, y) = L(x, y) + M(x, y)N^{-1} + O(N^{-2})$, where⁵

$$\bar{\alpha} = 1 + \eta N^{-1} = 1 + \frac{Q}{\pi \Lambda \bar{\rho}} = 1 + \frac{2Q}{\Lambda \log(E/2\pi)} = 2\alpha - 1. \tag{53}$$

To summarize, we have

$$K_{\text{RZ}}^N(x, y) = K(x, y) + L_{\text{RZ}}(x, y)N^{-2} + O(N^{-4}), \tag{54}$$

and thus, get the following recipe to go from the leading correction term of a CUE_N fluctuation statistics to the corresponding one of the Riemann zeros: simply replace the leading correction kernel L by its interior rescaling L_{RZ} .

In our experiments, the most pronounced difference between interior and exterior rescaling can be observed in the statistics shown in the right panel of Fig. 6.

4.4. The Odlyzko data set

The largest data set of Riemann zeros currently provided by Odlyzko, which was announced already in [15], consists of the 1 041 719 075 consecutive zeros starting with zero number

$$10^{23} + 985\,531\,550.$$

The first of them has height

$$13\,066\,434\,408\,793\,621\,120\,027.39614\,65854\dots,$$

while the last one has height

$$13\,066\,434\,408\,793\,754\,462\,591.63384\,74590\dots$$

Because of the logarithmic dependence on the height, even within that large data set, the smooth density $\bar{\rho}$, the effective dimension N , and the scaling parameters α , $\bar{\alpha}$ remain essentially *constant* to 15 digits precision,

$$\bar{\rho} = 7.81235\,22019\,1727\dots,$$

$$N = 11.29759\,09009\,547\dots,$$

$$\alpha = 1.02999\,00807\,6719\dots,$$

$$\bar{\alpha} = 1.05998\,01615\,3438\dots$$

⁵Note that [16, eq. (28)] contains a miscalculation by claiming that $\bar{\alpha}$ would be the same as α .

Table 1
Riemann Zeros Statistics Shown in Figs. 3–8.

Spacing	Density	CUE Case	$\xi = 1$	$\xi = 0.6$
First next neighbor	$p_{\xi}^{\text{RZ}}(1; s)$	Section 3.3	Fig. 3	Fig. 4
Random origin	$p_{\text{Origin}, \xi}^{\text{RZ}}(s)$	Section 3.4	Fig. 5	Fig. 6
Nearest neighbor	$p_{\text{nn}, \xi}^{\text{RZ}}(s)$	Section 3.5	Fig. 7	Fig. 8

In this table, we denote by p_{\dots}^{RZ} those probability densities for the Riemann zeros that are defined analogously to the CUE case p_{\dots}^{CUE} .

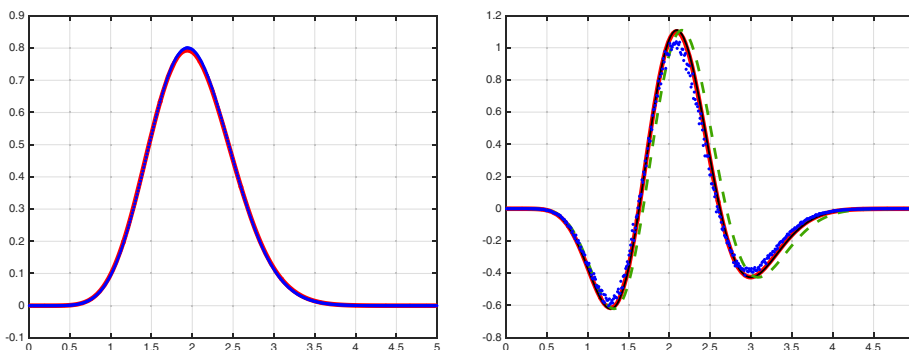


Figure 3. First next neighbor spacing: Riemann zeros data versus random-matrix-based prediction with effective dimension $N = \log(E/2\pi)/\sqrt{12\Lambda} \approx 11.3$, no thinning ($\xi = 1$). Left panel: a histogram of the Odlyzko data set with bin size 0.01 (blue dots); the large N limit $p_{2,\xi}(1; s)$ (red solid line). Right panel: the Riemann zero data minus $p_{2,\xi}(1; s)$ scaled by N^2 (blue dots); the leading correction term $r_{2,\xi}(1; s)$ with interior rescaling (red solid line), exterior rescaling (thin black line), and no rescaling (dashed green line).

From this data set, we extracted various spacing statistics,⁶ with ($\xi = 0.6$) and without ($\xi = 1$) thinning, and compared them up to the leading correction with the theoretical prediction in terms of operator determinants $\det(\mathbb{I} - \mathbb{K})$ and their corrections $\Omega(\mathbb{K}) : \mathbb{L}$ as obtained in Section 3 for the CUE. As detailed in Table 1, in Figs. 3–8, the Riemann zero data are shown as blue dots, the leading order term as a red solid line, and the correction terms

⁶The 0-th next neighbor spacing can be found in [15, Figs. 1 and 2] (no thinning and no theoretical prediction of the leading correction term), in [16, Figs. 1–3] (no thinning, exterior rescaling only), and in [20, Figs. 9 and 10] (exterior rescaling only). Because in this case, there is no difference visible between interior and exterior rescaling, we refrained from showing the results once more.

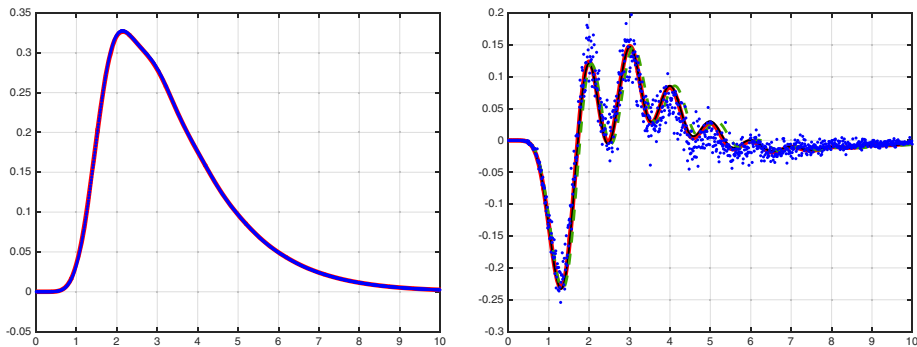


Figure 4. As in Fig. 3 but with thinning $\xi = 0.6$.

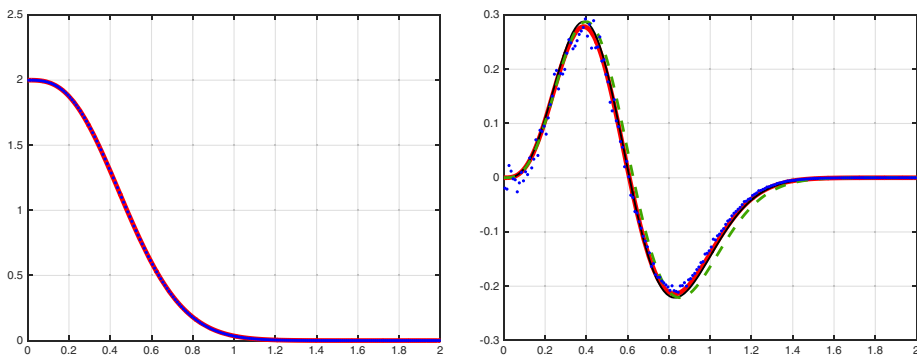


Figure 5. Minimum spacing from a randomly chosen origin: Riemann zeros data versus random-matrix-based prediction with effective dimension $N = \log(E/2\pi)/\sqrt{12\Delta} \approx 11.3$, no thinning ($\xi = 1$). Left panel: a histogram of the Odlyzko data set with bin size 0.01 (blue dots), where the approximately 10^9 zeros were broken into 1,024 sets of nearly equal size, for each of which 10^7 samples of a uniformly distributed random origin were drawn; the large N limit $p_{\text{Origin},\xi}(s)$ (red solid line). Right panel: the Riemann zero data minus $p_{\text{Origin},\xi}(s)$ scaled by N^2 (blue dots); the leading correction term $r_{\text{Origin},\xi}(s)$ with interior rescaling (red solid line), exterior rescaling (thin black line), and no rescaling (dashed green line).

- with interior rescaling $L \rightarrow L_{RZ}$ as a red solid line,
- with exterior rescaling $s \rightarrow \alpha s$ as a thin black line,
- without rescaling as a green dashed line.

In particular, with (interior) rescaling the statistics of the Riemann zeros listed in Table 1 are in excellent agreement up to the leading correction term.

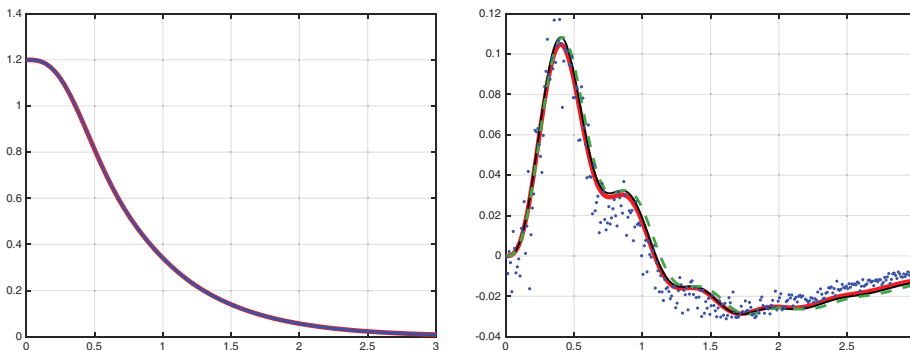


Figure 6. As in Fig. 5 but with thinning $\xi = 0.6$.

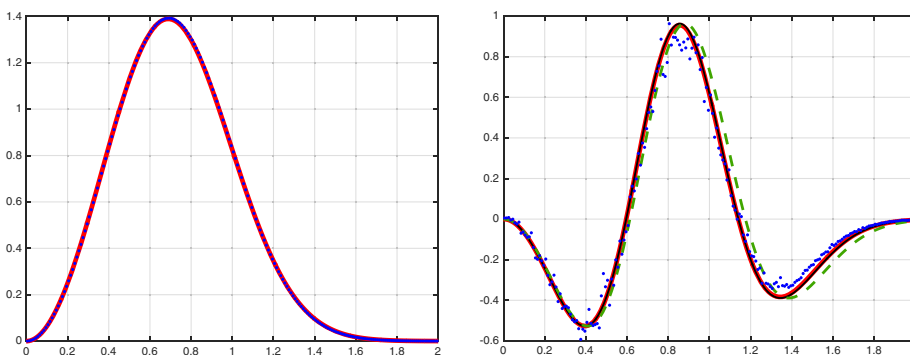


Figure 7. Nearest neighbor spacing: Riemann zeros data versus random-matrix-based prediction with effective dimension $N = \log(E/2\pi)/\sqrt{12\Lambda} \approx 11.3$, no thinning ($\xi = 1$). Left panel: a histogram of the Odlyzko data set with bin size 0.01 (blue dots); the large N limit $p_{nn,\xi}(s)$ (red solid line). Right panel: the Riemann zero data minus $p_{nn,\xi}(s)$ scaled by N^2 (blue dots); the leading correction term $r_{nn,\xi}(s)$ with interior rescaling (red solid line), exterior rescaling (thin black line), and no rescaling (dashed green line).

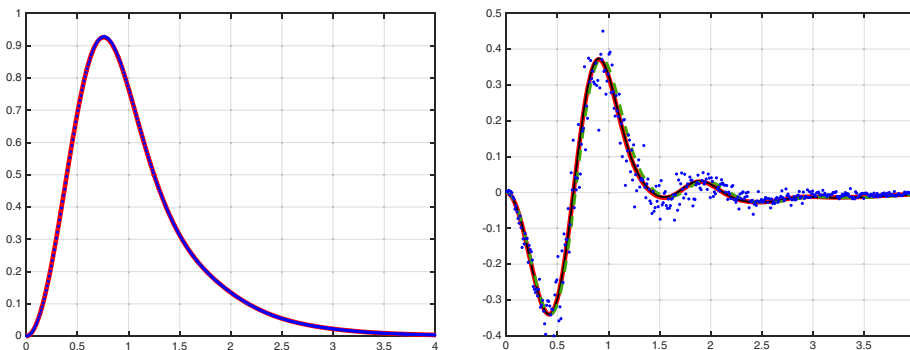


Figure 8. As in Fig. 7 but with thinning $\xi = 0.6$.

5. Spacing distributions for the COE and CSE

5.1. Generating functions

As already mentioned, the fact that the eigenvalues of COE matrices form a Pfaffian point process tells us that the counterpart to the formula (24) involves a Pfaffian of a 2×2 antisymmetric kernel function. Substituting this in (23), the Pfaffian analogue of the identity implying (25) (see, e.g., [6, eq. (6.32)]) allows for a Pfaffian analogue in the case of the COE to be given. However, the resulting formula is not of the same practical utility as (25) because the 2×2 kernel function is not analytic, which in turn means that the numerical quadrature methods of Section 2.2 have poor convergence properties.

Fortunately, there is a second option. This presents itself due to fundamental interrelations between gap probabilities for the COE and CUE. These also involve the gap probabilities for Haar distributed orthogonal matrices from the classical groups $O^+(N)$ and $O^-(N)$. In reference to the latter, eigenvalues on the real axis and, thus, corresponding to $\theta = 0$ or π are to be disregarded, and only the eigenvalues with angles in $(0, \pi)$ are to be considered (the remaining eigenvalues occur at the negative of these angles, i.e., the complex conjugates). The first interrelation of interest is between the generating functions for the gap probabilities in the CUE, and in Haar-distributed real orthogonal matrices [31, 32] and [6, eq. (8.127)]

$$\mathcal{E}^{\text{CUE}}((-\theta, \theta); z) = \mathcal{E}^{O^-(2\lfloor(N+1)/2\rfloor+1)}((0, \theta); z) \mathcal{E}^{O^+(2\lfloor N/2\rfloor+1)}((0, \theta); z). \tag{55}$$

With ME denoting a particular matrix ensemble, we have used the symbol \mathcal{E}^{ME} instead of E^{ME} to indicate that in this expression no scaling of the eigenangles has been imposed; this convention will be followed below. In fact, we only use the scaled quantity for ensembles $\text{ME} = C\beta E_N$, which are translationally invariant with uniform density $N/2\pi$. The unscaled and scaled quantities are then related by

$$\mathcal{E}^{C\beta E_N}((-2\pi s/N, 2\pi s/N); z) = E^{C\beta E_N}(2s; z).$$

To present the next interrelation, introduce the generating functions

$$E^{\text{COE}, \pm}((-\theta, \theta); z) := \sum_{n=0}^N (1-z)^n (E^{\text{COE}}(2n, (0, \theta)) + E^{\text{COE}}(2n \pm 1, (0, \theta))). \tag{56}$$

With $\text{COE} \cup \text{COE}$ denoting the point process of $2N$ eigenvalues on the circle that results by superimposing the eigenvalue sequences of two independent COE matrices, a result conjectured by Dyson [33] and proved by Gunson [34] gives that

$$\text{alt}(\text{COE} \cup \text{COE}) = \text{CUE}, \tag{57}$$

where the operation “alt” refers to the operation of observing every second eigenvalue only. As a consequence of this, one has that [33,35]

$$E^{\text{CUE}}((-\theta, \theta); z) = E^{\text{COE},-}((-\theta, \theta); z)E^{\text{COE},+}((-\theta, \theta); z). \tag{58}$$

According to (55), it follows from (58) that the gap probabilities for the COE are related to those for $O^+(n)$ and $O^-(n)$ with n suitably chosen, but this alone does not determine the former. To be able to do this, additional interrelationships between generating functions are required. In the case $N \mapsto 2N$ and thus, N even, the additional interrelationships have been given in [32] according to the generating function identity

$$\mathcal{E}^{\text{COE},\pm}((-\theta, \theta); z) \Big|_{N \mapsto 2N} = \mathcal{E}^{O^\pm(2N+1)}((0, \theta); z). \tag{59}$$

Note that (59) substituted in (58) is consistent with (55) in the case N even. The identity (59) has very recently [36] been shown to be a corollary of the identities between eigenvalue distributions

$$\text{even } |\text{COE}_{2N}| = O^+(2N + 1), \quad \text{odd } |\text{COE}_{2N}| = O^-(2N + 1). \tag{60}$$

Here, the notation COE_{2N} refers to the eigenvalue distribution of $2N \times 2N$ COE matrices, while $|\text{COE}_{2N}|$ refers to the distribution in the circumstance that the eigenvalues with angles $-\pi < \theta < 0$ are reflected in the real axis by $\theta \mapsto -\theta$, and thus, all eigenvalues have angles between 0 and π . The operation even (odd) refers to observing only those eigenvalues that occur an even (odd) number of places from $\theta = 0$, reading anticlockwise.

Substituting (59) in (56), and setting $\bar{z} = 2z - z^2$ so that $1 - \bar{z} = (1 - z)^2$ we obtain, after some minor manipulation, a known closed formula for the gap probabilities of COE_{2N} in terms of the gap probabilities for $O^\pm(2N + 1)$ [6, eq. (8.150)]

$$\mathcal{E}^{\text{COE}_{2N}}((-\theta, \theta); z) = \frac{(1 - z)\mathcal{E}^{O^+(2N+1)}((0, \theta); \bar{z}) + \mathcal{E}^{O^-(2N+1)}((0, \theta); \bar{z})}{2 - z}. \tag{61}$$

The analogue of (60) for COE_{2N-1} allows us to deduce the analogue of (61). The required formulas were not known until [36] and consequently this is a new result.

PROPOSITION 5. *We have*

$$\mathcal{E}^{\text{COE}_{2N-1}}((-\theta, \theta); z) = \frac{(1 - z)\mathcal{E}^{O^-(2N)}((0, \theta); \bar{z}) + \mathcal{E}^{O^+(2N)}((0, \theta); \bar{z})}{2 - z}. \tag{62}$$

Proof. We read off from [36, theorem 7.1] that

$$\text{even } |\text{COE}_{2N-1}| = O^-(2N), \quad \text{odd } |\text{COE}_{2N-1}| = O^+(2N). \tag{63}$$

Substituting in (56) and manipulating as in the derivation of (61) gives (62). ■

With $v = (-)^N$, the cases (61) and (62) can be combined into a single formula that holds for both parities of N :

$$\mathcal{E}^{\text{COE}_N}((-\theta, \theta); z) = \frac{(1 - z)\mathcal{E}^{\text{O}^v(N+1)}((0, \theta); \bar{z}) + \mathcal{E}^{\text{O}^{-v}(N+1)}((0, \theta); \bar{z})}{2 - z}. \tag{64}$$

The eigenvalues for $\text{O}^\pm(N + 1)$ in $(0, \pi)$ form a determinantal point process with kernel (see, e.g., [6, proposition 5.5.3])

$$\frac{1}{2\pi} \left(\frac{\sin N(x - y)/2}{\sin(x - y)/2} \mp v \frac{\sin N(x + y)/2}{\sin(x + y)/2} \right) = \frac{N}{2\pi} K^{N, \mp v}(xN/2\pi, yN/2\pi), \tag{65}$$

where

$$K^{N, \pm}(x, y) = K^N(x, y) \pm K^N(x, -y). \tag{66}$$

Scaling the eigenangles of COE_N to have unit mean spacing by $\theta = 2\pi s/N$ thus yields the following result:

COROLLARY 1. *Let $\mathbb{K}_s^{N, \pm}$ denote the integral operator on $(0, s)$ with kernel (66) and denote $\bar{z} = 2z - z^2$. We have*

$$E^{\text{COE}_N}((-s, s); z) = \frac{(1 - z) \det(I - \bar{z}\mathbb{K}_s^{N, -}) + \det(I - \bar{z}\mathbb{K}_s^{N, +})}{2 - z}. \tag{67}$$

As in (31), thinning is expressed by

$$E_\xi^{\text{COE}_N}((-s, s); z) = E^{\text{COE}_N}((-s, s); \xi z).$$

We will now turn our attention to spacing probabilities for the CSE. These follow from knowledge of the spacing distributions for the COE. Thus, one has the interrelation [37]

$$\text{alt COE}_{2N} = \text{CSE}_N, \tag{68}$$

and it follows from this that

$$\begin{aligned} \mathcal{E}^{\text{CSE}_N}(n; (-\theta, \theta)) &= \mathcal{E}^{\text{COE}_{2N}}(2n; (-\theta, \theta)) \\ &+ \frac{1}{2} (\mathcal{E}^{\text{COE}_{2N}}(2n - 1; (-\theta, \theta)) + \mathcal{E}^{\text{COE}_{2N}}(2n + 1; (-\theta, \theta))). \end{aligned} \tag{69}$$

As noted in [6, eq. (8.158)], recalling the definition (56), and making use too of (59), this gives

$$\mathcal{E}^{\text{CSE}_N}((-\theta, \theta); z) = \frac{1}{2} \left(\mathcal{E}^{\text{O}^+(2N+1)}((0, \theta); z) + \mathcal{E}^{\text{O}^-(2N+1)}((0, \theta); z) \right). \tag{70}$$

As discussed before (65), the kernel of the determinantal point process formed by the eigenvalues of $\text{O}^\pm(2N + 1)$ is given by

$$\frac{N}{\pi} K^{2N, \mp}(xN/\pi, yN/\pi).$$

Scaling the eigenangles of CSE_N to have unit mean spacing by $\theta = 2\pi s/N$, thus yields the following result, where we use (31) to express the presence of thinning:

PROPOSITION 6. Let $\mathbb{K}_s^{N,\pm}$ denote the integral operator on $(0, s)$ with kernel (66). We have

$$E_\xi^{CSE_N}((-s, s); z) = \frac{1}{2} \left(\det \left(I - \xi z \mathbb{K}_{2s}^{2N,-} \right) + \det \left(I - \xi z \mathbb{K}_{2s}^{2N,+} \right) \right). \quad (71)$$

5.2. Expansion of spacing distributions

Making use of (33), we get the spacing distribution

$$p_\xi^{COE}(0; s) = \frac{1}{\xi} \frac{d^2}{ds^2} E_\xi^{COE_N}(0; (-s/2, s/2)), \quad (72)$$

from which knowledge of the large N expansion of the determinants in Corollary 1 will allow us to determine the terms in the corresponding expansion of the spacing.

PROPOSITION 7. Let \mathbb{K}_s^\pm and \mathbb{L}_s^\pm denote the integral operators on $(0, s)$ with kernels $K(x, y) \pm K(x, -y)$ and $L(x, y) \pm L(x, -y)$, respectively. We have

$$p_\xi^{COE}(0; s) = p_{1,\xi}(0; s) + \frac{1}{N^2} r_{1,\xi}(0; s) + O\left(\frac{1}{N^4}\right), \quad (73)$$

where, denoting $\bar{\xi} = 2\xi - \xi^2$,

$$p_{1,\xi}(0; s) = \frac{1}{\bar{\xi}} \frac{d^2}{ds^2} \left((1 - \xi) \det(\mathbb{I} - \bar{\xi} \mathbb{K}_{s/2}^-) + \det(\mathbb{I} - \bar{\xi} \mathbb{K}_{s/2}^+) \right), \quad (74)$$

and

$$r_{1,\xi}(0; s) = \frac{d^2}{ds^2} \left((1 - \xi) \Omega(\bar{\xi} \mathbb{K}_{s/2}^-) : \mathbb{L}_{s/2}^- + \Omega(\bar{\xi} \mathbb{K}_{s/2}^+) : \mathbb{L}_{s/2}^+ \right). \quad (75)$$

Remark 2. The case $\xi = 1$ of (74) is due to Gaudin [18].

As an illustration and test of the above results, we took 10^8 samples of COE_N ($N = 20$, using the $\beta = 1$ CMV sparse matrix model [38]), and from this made an empirical determination of the spacing distribution scaled to unit mean spacing. This was then subtracted from the large N limit $p_{1,\xi}(0; s)$ and the difference compared against $r_{1,\xi}(0; s)$. Both $\xi = 1$ (no thinning) and $\xi = 0.6$ were considered; see Figs. 9 and 10.

The analogue of (72),

$$p_\xi^{CSE}(0; s) = \frac{1}{\xi} \frac{d^2}{ds^2} E_\xi^{CSE_N}(0; (-s/2, s/2)), \quad (76)$$

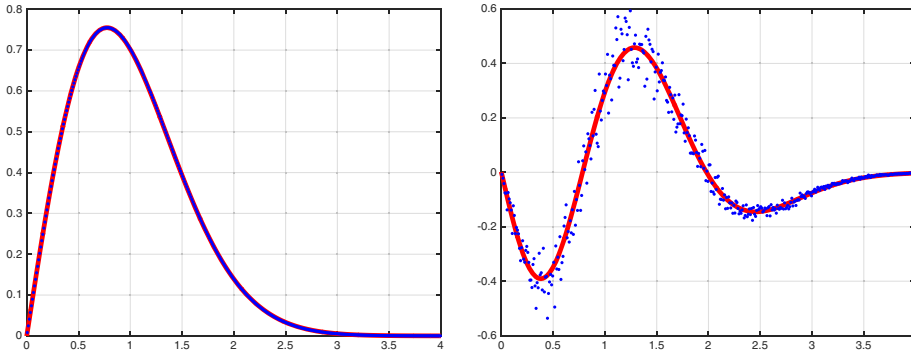


Figure 9. Zeroth next neighbor spacing: simulation versus formulas from theory for finite size COE, no thinning ($\xi = 1$). Left panel: a histogram of empirical data from COE_N with $N = 20$ scaled to unit mean spacing, computed using a bin size of 0.01 and 10^8 samples (blue dots); the large N limit $p_{1,\xi}(0; s)$ (red solid line). Right panel: the simulation data minus $p_{1,\xi}(0; s)$ scaled by N^2 (blue dots); the leading correction term $r_{1,\xi}(0, s)$ (red solid line).

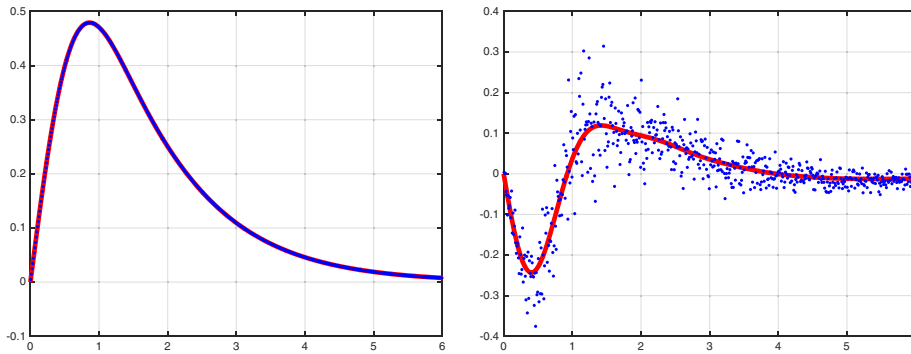


Figure 10. As in Fig. 9 but with thinning $\xi = 0.6$.

allows us to determine the first terms of the large N expansion by expanding correspondingly the determinants in Proposition 6.

PROPOSITION 8. With \mathbb{K}_s^\pm and \mathbb{L}_s^\pm as in Proposition 7 we have

$$p_\xi^{\text{CSE}}(0; s) = p_{4,\xi}(0; s) + \frac{1}{N^2} r_{4,\xi}(0; s) + O\left(\frac{1}{N^4}\right), \tag{77}$$

where

$$p_{4,\xi}(0; s) = \frac{1}{2\xi} \frac{d^2}{ds^2} (\det(\mathbb{I} - \xi \mathbb{K}_s^-) + \det(\mathbb{I} - \xi \mathbb{K}_s^+)) \tag{78}$$

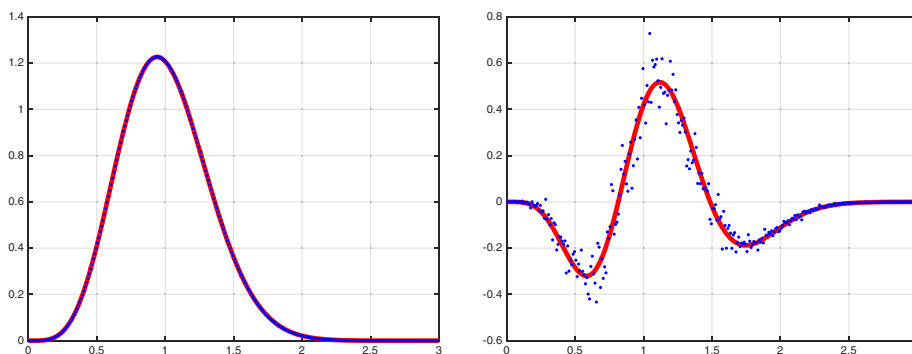


Figure 11. Zeroth next neighbor spacing: simulation versus formulas from theory for finite size CSE, no thinning ($\xi = 1$). Left panel: a histogram of empirical data from CSE_N with $N = 20$ scaled to unit mean spacing, computed using a bin size of 0.01 and 10^8 samples (blue dots); the large N limit $p_{4,\xi}(0; s)$ (red solid line). Right panel: the simulation data minus $p_{4,\xi}(0; s)$ scaled by N^2 (blue dots); the leading correction term $r_{4,\xi}(0, s)$ (red solid line).

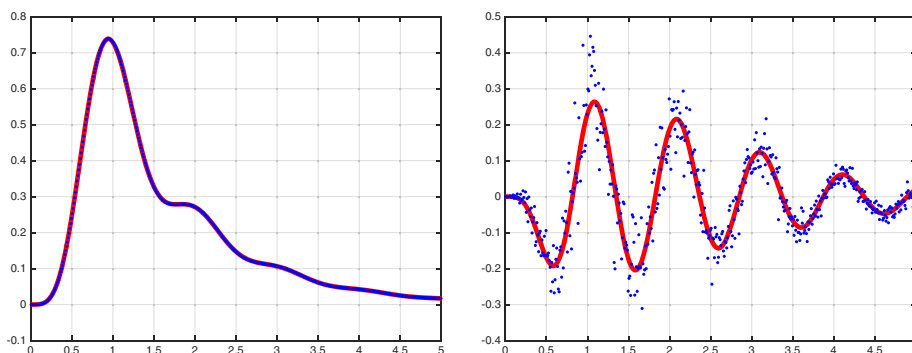


Figure 12. As in Fig. 11 but with thinning $\xi = 0.6$.

and

$$r_{4,\xi}(0; s) = \frac{1}{8} \frac{d^2}{ds^2} (\Omega(\xi \mathbb{K}_s^-) : \mathbb{L}_s^- + \Omega(\xi \mathbb{K}_s^+) : \mathbb{L}_s^+). \tag{79}$$

Remark 3. The case $\xi = 1$ of (78) is due to Mehta and Dyson [37].

As was done for the COE, to illustrate and test the above results, we took 10^8 samples of CSE_N ($N = 20$, again using the $\beta = 4$ CMV sparse matrix model [38]), and from this made an empirical determination of the spacing distribution scaled to unit mean spacing. This was then subtracted from the large N limit $p_{4,\xi}(0; s)$ and the difference compared against $r_{4,\xi}(0; s)$. Both $\xi = 1$ (no thinning) and $\xi = 0.6$ were considered; see Figs. 11 and 12.

5.3. Painlevé transcendent characterization

In addition to the integral operator characterization of the expansion terms in Propositions 7 and 8, there is one in terms of Painlevé transcendents. Such an expression in the case of the CUE has been given in [20] and restated in (14)–(17) above. To simplify we restrict ourselves to N being even, writing $N \rightarrow 2N$ for definiteness. We begin by noting that the joint eigenvalue PDF for the eigenangles $\{\theta_j\}$ of $O^\pm(2N + 1)$ matrices in the interval $(0, \pi)$, after changing variables $x_j = \sin^2(\theta_j/2)$, is proportional to

$$\prod_{l=1}^N x_l^a (1 - x_l)^b \prod_{1 \leq j < k \leq N} (x_k - x_j)^2, \quad 0 < x_l < 1,$$

with $a = \pm 1/2$, $b = -a$, which is an example of the Jacobi unitary ensemble (JUE); see, for example, [6, §3.7.1]. Hence the corresponding generating functions for the gap probabilities are related by

$$\mathcal{E}^{O^\pm(2N+1)}((0, \theta); z) = \mathcal{E}^{\text{JUE}_N} \left(\left(0, \sin^2 \frac{\theta}{2} \right); z \right) \Big|_{a=\pm 1/2, b=-a}. \tag{80}$$

The significance of this result for present purposes is that the RHS can be expressed as a σ PVI transcendent [39, 40]. Specifically, reading off from [6, eqs. (8.71), (8.75), (8.76)] we have

$$\mathcal{E}^{\text{JUE}_N}((0, \sin^2 \theta/2); z) = \exp \int_0^{\sin^2 \theta/2} \frac{f^\pm(t; z)}{t(t-1)} dt, \tag{81}$$

where $f = f^\pm$ satisfies the particular σ PVI equation

$$(t(1-t)f'')^2 + (f' - N^2)(2f + (1-2t)f')^2 - (f')^2 \left(f' - N^2 + \frac{1}{4} \right) = 0, \tag{82}$$

subject to the boundary condition

$$f^\pm(t; z) \underset{t \rightarrow 0^+}{\sim} \begin{cases} \frac{8N(N^2 - 1/4)}{3\pi} z t^{3/2}, & a = 1/2, \\ \frac{2N}{\pi} z t^{1/2}, & a = -1/2. \end{cases} \tag{83}$$

We, therefore, have

$$\mathcal{E}^{O^\pm(2N+1)}((0, \pi s/N); z) = \exp \left(- \int_0^{(\pi s)^2} f^\pm(\sin^2 \sqrt{u}/2N; z) \frac{du}{N \sqrt{u} \sin \sqrt{u}/N} \right). \tag{84}$$

To obtain an expansion consistent with (73) we make the ansatz

$$f^\pm(\sin^2(\sqrt{w}/2N); z) = f_0^\pm + \frac{f_1^\pm}{N^2} + O(N^{-4}). \tag{85}$$

Changing variables in (82) $t = \sin^2(\sqrt{w}/2N)$, substituting (85), and equating terms to leading order N^4 and to next leading order N^2 , we obtain characterizations in terms of differential equations of f_0^\pm and f_1^\pm .

PROPOSITION 9. *The leading function f_0^\pm in (85) satisfies the particular σ PIII' equation (with $v_1 = v_2 = 1/2$ in the notation of [6, eq. (8.15)])*

$$w^2(f_0'')^2 - \frac{(f_0')^2}{4} + 4f_0(f_0')^2 + w(f_0')^2 - 4w(f_0')^3 - f_0f_0' = 0, \tag{86}$$

subject to the boundary condition

$$f_0^v(w; z) \sim \begin{cases} \frac{z}{3\pi} w^{3/2}, & v = +, \\ \frac{z}{\pi} w^{1/2}, & v = -. \end{cases} \tag{87}$$

The function f_1^\pm in (85) satisfies the second-order linear differential equation

$$A_1(w; z)f_1'' + B_1(w; z)f_1' + C_1(w; z)f_1 + D_1(w; z) = 0, \tag{88}$$

where the coefficients are given in terms of $f_0 = f_0^\pm$ according to

$$A_1(w; z) = 2w^2 f_0'',$$

$$B_1(w; z) = 8f_0 f_0' + 2wf_0' - 12wf_0^2 - f_0 - \frac{f_0'}{2},$$

$$C_1(w; z) = f_0'(4f_0' - 1),$$

$$D_1(w; z) = \frac{f_0'}{3} \left(3f_0^2 + w^2 f_0'' - 2w^2(f_0')^2 + wf_0 - \frac{wf_0'}{4} - 2wf_0 f_0' \right) - \frac{f_0^2}{4},$$

subject to the boundary condition

$$f_1^\pm(w; z) \sim -\frac{zw^{3/2}}{12\pi}, \tag{89}$$

which is thus the same in both cases. In terms of f_0^\pm and f_1^\pm we have

$$\mathcal{E}^{0^\pm(2N+1)}((0, \pi s/N); z) = \exp \left(- \int_0^{(\pi s)^2} \frac{f_0^\pm(w; z)}{w} dw \right)$$

$$\times \left(1 - \frac{1}{N^2} \int_0^{(\pi s)^2} \frac{w f_0^\pm(w; z) + 6 f_1^\pm(w; z)}{6w} dw + O\left(\frac{1}{N^4}\right) \right). \tag{90}$$

Remark 4. A problem for future study is to describe how the solutions of (15) and (88) relate to the broader Painlevé theory.⁷

Scaling the eigenangles of COE_{2N} to have unit mean spacing by setting $\theta = \pi s/N$, we can now substitute (90) into (61); then—after using (31) to add the presence of thinning—substitute the result in (72) and compare with (73) to obtain characterizations of the expansion terms $p_{1,\xi}(0; s)$ and $r_{1,\xi}(0; s)$ in terms of Painlevé transcendents.

PROPOSITION 10. *We have*

$$p_{1,\xi}(0; s) = \frac{1}{\xi(2 - \xi)} \frac{d^2}{ds^2} \left((1 - \xi) \exp \left(- \int_0^{(\pi s/2)^2} \frac{f_0^+(w; 2\xi - \xi^2)}{w} dw \right) + \exp \left(- \int_0^{(\pi s/2)^2} \frac{f_0^-(w; 2\xi - \xi^2)}{w} dw \right) \right) \tag{91}$$

and

$$r_{1,\xi}(0; s) = - \frac{4}{\xi(2 - \xi)} \frac{d^2}{ds^2} \left((1 - \xi) \int_0^{(\pi s/2)^2} \frac{w f_0^+(w; 2\xi - \xi^2) + 6 f_1^+(w; 2\xi - \xi^2)}{6w} dw \times \exp \left(- \int_0^{(\pi s/2)^2} \frac{f_0^+(w; 2\xi - \xi^2)}{w} dw \right) + \int_0^{(\pi s/2)^2} \frac{w f_0^-(w; 2\xi - \xi^2) + 6 f_1^-(w; 2\xi - \xi^2)}{6w} dw \times \exp \left(- \int_0^{(\pi s/2)^2} \frac{f_0^-(w; 2\xi - \xi^2)}{w} dw \right) \right). \tag{92}$$

Remark 5. The case $\xi = 1$ of (91) agrees with the $\sigma\text{PIII}'$ formula for $p_1(0; s)$ (no thinning) reported in [41] and [32].

The differential equations (86) and (88) can be used to successively generate terms in the series expansions of f_0^\pm and f_1^\pm about the origin, extending the boundary conditions (87) and (89). Substituting these into

⁷This was raised by one of the referees.

(91) and (92) then provides us with the small s expansion of $p_{1,\xi}(0; s)$ and $r_{1,\xi}(0; s)$.

COROLLARY 2. *We have*

$$\begin{aligned} p_{1,\xi}(0; s) = & \frac{1}{6}\pi^2\xi s - \frac{1}{60}\pi^4\xi s^3 - \frac{1}{270}\pi^4(\xi - 2)\xi s^4 \\ & + \frac{\pi^6\xi s^5}{1680} + \frac{\pi^6(\xi - 2)\xi s^6}{4725} - \frac{\pi^8\xi s^7}{90720} \\ & + \frac{\pi^8(\xi - 2)(3\xi - 32)\xi s^8}{5292000} + \frac{\pi^{10}\xi s^9}{7983360} + O(s^{10}) \end{aligned} \quad (93)$$

and

$$\begin{aligned} r_{1,\xi}(0; s) = & -\frac{1}{6}\pi^2\xi s + \frac{1}{18}\pi^4\xi s^3 + \frac{1}{54}\pi^4(\xi - 2)\xi s^4 \\ & - \frac{1}{240}\pi^6\xi s^5 - \frac{4\pi^6(\xi - 2)\xi s^6}{2025} + \frac{\pi^8\xi s^7}{7560} \\ & - \frac{\pi^8(\xi - 2)(3\xi - 32)\xi s^8}{352800} - \frac{\pi^{10}\xi s^9}{435456} + O(s^{10}). \end{aligned} \quad (94)$$

We remark that setting $\xi = 1$ in (93) extends the expansion of $p_{1,\xi}(0; s)|_{\xi=1}$ given in [6, eq. (8.141)].

For the analogous results in the case of the CSE_N , scaling the eigenangles to have unit mean spacing by setting $\theta = 2\pi s/N$, we substitute (90) into (70); then—after using (31) to add the presence of thinning—substitute the result in (76) and compare with (77).

PROPOSITION 11. *We have*

$$\begin{aligned} p_{4,\xi}(0; s) = & \frac{1}{2\xi} \frac{d^2}{ds^2} \left(\exp \left(- \int_0^{(\pi s)^2} \frac{f_0^+(w; \xi)}{w} dw \right) \right. \\ & \left. + \exp \left(- \int_0^{(\pi s)^2} \frac{f_0^-(w; \xi)}{w} dw \right) \right) \end{aligned} \quad (95)$$

and

$$\begin{aligned} r_{4,\xi}(0; s) = & -\frac{1}{2\xi} \frac{d^2}{ds^2} \left(\int_0^{(\pi s)^2} \frac{w f_0^+(w; \xi) + 6 f_1^+(w; \xi)}{6w} dw \right. \\ & \left. \times \exp \left(- \int_0^{(\pi s)^2} \frac{f_0^+(w; \xi)}{w} dw \right) \right) \end{aligned}$$

$$\begin{aligned}
& + \int_0^{(\pi s)^2} \frac{w f_0^-(w; \xi) + 6 f_1^-(w; \xi)}{6w} dw \\
& \times \exp \left(- \int_0^{(\pi s)^2} \frac{f_0^-(w; \xi)}{w} dw \right). \tag{96}
\end{aligned}$$

Remark 6. As in Remark 5 the case $\xi = 1$ of (95) agrees with the σ PIII' formula for $p_4(0; s)$ (no thinning) reported in [41] and [32].

As with Corollary 2 we can use these characterizations to generate power series expansions about the origin.

COROLLARY 3. *We have*

$$\begin{aligned}
p_{4,\xi}(0; s) &= \frac{16}{135} \pi^4 \xi s^4 - \frac{128\pi^6 \xi s^6}{4725} + \frac{512\pi^8 \xi s^8}{165375} - \frac{34816\pi^{10} \xi s^{10}}{147349125} + O(s^{12}), \\
r_{4,\xi}(0; s) &= -\frac{4}{27} \pi^4 \xi s^4 + \frac{128\pi^6 \xi s^6}{2025} - \frac{128\pi^8 \xi s^8}{11025} + \frac{17408\pi^{10} \xi s^{10}}{13395375} + O(s^{12}).
\end{aligned}$$

Acknowledgments

We are grateful to A. Odlyzko for providing us with the Riemann zero data set. We would also like to thank the referees for their thoughtful comments. The research of F.B. was supported by the DFG-Collaborative Research Center, TRR 109, ‘‘Discretization in Geometry and Dynamics.’’ The research of P.J.F. and A.M. is part of the program of study supported by the ARC Centre of Excellence for Mathematical & Statistical Frontiers.

Appendix: Implementation

The method of Section 2.2 for numerically evaluating general terms of integral operators is most easily added to the Matlab toolbox described in [26]; the basic code needed to run it is as follows:

```

1 function [val, err, n] = OperatorTerm (term, varargin)
2
3 %OPERATORTERM evaluates terms of integral operators that have a scalar value.
4 %
5 % OPERATORTERM(term,K1,...,Km) returns the value of the
6 % expression 'term(K1,...,Km)' for m discrete integral
7 % operators K1,...,Km (that is, for Nyström matrices of
8 % a variable dimension n, which is chosen adaptively).
9 % The value to be returned must be scalar.
10
11 tol = 5e-15;
12 n = 7; n_max = 1000;
13

```

```

14 operators = cell(size(varargin));
15
16 val0 = inf;
17 while n < n_max
18     n = floor(1.41*n);
19     for k = 1:length(varargin)
20         operators{k} = varargin{k}(n);
21     end
22     val = term(operators{:});
23     err = abs(val-val0);
24     if err < tol; break; end
25     val0 = val;
26 end

```

From the user perspective, this command is called with a considerable symbolic look and feel. For example, the leading order term in (11), that is,

$$\det(\mathbb{I} - \mathbb{K}_s),$$

evaluates for $s = 1$ to the following value:

```

1 s = 1;
2 K = op(@(x,y) sinc(pi*(x-y)), [0,s]);
3 I = @(K) eye(size(K));
4 lead = @(K) det(I(K) - K);
5 [val, err, n] = OperatorTerm(lead, K);
6 PrintCorrectDigits(val, err);
7
8     0.170217421379185

```

and the leading correction term in (11), that is,

$$\Omega(\mathbb{K}_s) : \mathbb{L}_s = -\det(\mathbb{I} - \mathbb{K}_s) \text{tr}((\mathbb{I} - \mathbb{K}_s)^{-1} \mathbb{L}_s),$$

evaluates for $s = 1$ to the following value:

```

9 L = op(@(x,y) pi*(x-y) .* sin(pi*(x-y))/6, [0,s]);
10 corr = @(K,L) -det(I(K)-K) * trace((I(K)-K)\L);
11 [val, err, n] = OperatorTerm(corr, K, L);
12 PrintCorrectDigits(val, err);
13
14     -0.075241982465122

```

Because the estimated error has been taken into account when printing these values, they are good to about 15 digits; in both cases the adaptively chosen number of Gauss–Legendre quadrature points (i.e., the dimension of the Nyström matrices representing the integral operators) was as small as $n = 15$, which is a clear sign of the exponential convergence of the method. CPU time is about a millisecond.

References

1. F. J. DYSON, The three fold way. Algebraic structure of symmetry groups and ensembles in quantum mechanics, *J. Math. Phys.* 3:1199–1215 (1962).
2. A. ALTLAND and M. R. ZIRNBAUER, Nonstandard symmetry classes in mesoscopic normal-superconducting hybrid compounds, *Phys. Rev. B* 55:1142–1161 (1997).

3. N. M. KATZ and P. SARNAK, Zeroes of zeta functions and symmetry, *Bull. Amer. Math. Soc.* 36:1–26 (1999).
4. E. P. WIGNER, Proc. conf. on neutron physics Gatlinburg, *Oak Ridge Natl. Lab. Rep. ORNL 2309:59* (1957).
5. O. BOHIGAS, M. J. GIANNONI, and C. SCHMIT, Characterization of chaotic quantum spectra and universality of level fluctuation laws, *Phys. Rev. Lett.* 52:1–4 (1984).
6. P. J. FORRESTER, *Log-Gases and Random Matrices* (Princeton University Press, Princeton, 2010).
7. F. W. K. FIRK and S. J. MILLER, Nuclei, primes and the random matrix connection, *Symmetry* 1:64–105 (2009).
8. H.-J. STÖCKMANN, Microwave billiards and quantum chaos, *Scholarpedia*, 5(10):10243 (2010).
9. Wikipedia, Hilbert–Pólya conjecture, en.wikipedia.org/wiki/Hilbert-Pólya_conjecture (2016).
10. M. V. BERRY and J. P. KEATING, The Riemann zeros and eigenvalue asymptotics, *SIAM Rev.* 41:236–266 (1999).
11. G. SIERRA and J. RODRIGUEZ-LAGUNA, $H = xp$ model revisited and the Riemann zeros, *Phys. Rev. Lett.* 106:200201, (2011).
12. A. M. ODLYZKO, On the distribution of spacings between zeros of the zeta function, *Math. Comput.* 48:273–308 (1987).
13. J. P. KEATING and N. C. SNAITH, Random matrix theory and $\zeta(1/2 + it)$, *Commun. Math. Phys.* 214:57–89 (2001).
14. F. J. DYSON, Statistical theory of energy levels of complex systems II, *J. Math. Phys.* 3:157–165 (1962).
15. A. M. ODLYZKO, The 10^{22} -nd zero of the Riemann zeta function, in *Dynamical, Spectral, and Arithmetic Zeta Functions* (M. L. Lapidus, M. van Frankenhuysen, Eds.), Contemporary Math., Vol. 290, pp. 139–144, Amer. Math. Soc, Providence (2001).
16. E. BOGOMOLNY, O. BOHIGAS, P. LEBOEUF, and A. C. MONASTRA, On the spacing distribution of the Riemann zeros: corrections to the asymptotic result, *J. Phys. A* 39:10743–10754 (2006).
17. M. L. MEHTA, On the statistical properties of the level-spacings in nuclear spectra, *Nucl. Phys. B* 18:395–419 (1960).
18. M. GAUDIN, Sur la loi limite de l’espacement des valeurs propres d’une matrice aléatoire, *Nucl. Phys.* 25:447–458 (1961).
19. M. JIMBO, T. MIWA, Y. MÔRI, and M. SATO, Density matrix of an impenetrable Bose gas and the fifth Painlevé transcendent, *Phys. D* 1:80–158 (1980).
20. P. J. FORRESTER and A. MAYS, Finite-size corrections in random matrix theory and Odlyzko’s dataset for the Riemann zeros, *Proc. Roy. Soc. A* 471:20150436, (2015).
21. J. ILLIAN, A. PENTTINEN, H. STOYAN, and D. STOYAN, *Statistical Analysis and Modelling of Spatial Point Patterns* (Wiley, Chichester, 2008).
22. O. BOHIGAS and M. P. PATO, Missing levels in correlated spectra, *Phys. Lett. B* 595:171–176 (2004).
23. F. BORNEMANN, On the numerical evaluation of Fredholm determinants, *Math. Comp.* 79:871–915 (2010).
24. B. SIMON, *Trace Ideals and Their Applications*, 2nd ed. (Amer. Math. Soc., Providence, 2005).
25. I. BOGAERT, Iteration-free computation of Gauss-Legendre quadrature nodes and weights, *SIAM J. Sci. Comput.* 36:A1008–A1026 (2014).

26. F. BORNEMANN, On the numerical evaluation of distributions in random matrix theory: a review, *Markov Processes Relat. Fields* 16:803–866 (2010).
27. P. J. FORRESTER and A. M. ODLYZKO, Gaussian unitary ensemble eigenvalues and Riemann ζ function zeros: a non-linear equation for a new statistic, *Phys. Rev. E* 54:R4493–R4495 (1996).
28. E. BOGOMOLNY and J. P. KEATING, Gutzwiller’s trace formula and spectral statistics: beyond the diagonal approximation, *Phys. Rev. Lett.* 77:1472–1475 (1996).
29. R. J. MATHAR, Series of reciprocal powers of k -almost primes, *arXiv:0803.0900v3* (2009).
30. E. BOGOMOLNY and J. P. KEATING, A method for calculating spectral statistics based on random-matrix universality with an application to the three-point correlations of the Riemann zeros, *J. Phys. A* 46:305203, (2013).
31. J. BAIK and E. M. RAINS, Algebraic aspects of increasing subsequences, *Duke Math. J.* 109:1–65 (2001).
32. P. J. FORRESTER, Evenness symmetry and inter-relationships between gap probabilities in random matrix theory, *Forum Math.* 18:711–743 (2006).
33. F. J. DYSON, Statistical theory of energy levels of complex systems III, *J. Math. Phys.* 3:166–175 (1962).
34. J. GUNSON, Proof of a conjecture of Dyson in the statistical theory of energy levels, *J. Math. Phys.* 4:752–753 (1962).
35. M. L. MEHTA, Power series for level spacing functions of random matrix ensembles, *Z. Phys. B* 86:285–290 (1992).
36. F. BORNEMANN and P. J. FORRESTER, Singular values and evenness symmetry in random matrix theory, *Forum Math.* 28:873–891 (2016).
37. M. L. MEHTA and F. J. DYSON, Statistical theory of the energy levels of complex systems. V, *J. Math. Phys.* 4:713–719 (1963).
38. R. KILLIP and I. NENCIU, Matrix models for circular ensembles, *Int. Math. Res. Not.* 2004:2665–2701 (2004).
39. L. HAINE and J.-P. SEMENGUE, The Jacobi polynomial ensemble and the Painlevé VI equation, *J. Math. Phys.* 40:2117–2134 (1999).
40. P. J. FORRESTER and N. S. WITTE, Application of the τ -function theory of Painlevé equations to random matrices: PVI, the JUE, CyUE, cJUE and scaled limits, *Nagoya Math. J.* 174:29–114 (2004).
41. P. J. FORRESTER and N. S. WITTE, Exact Wigner surmise type evaluation of the spacing distribution in the bulk of the scaled random matrix ensembles, *Lett. Math. Phys.* 53:195–200 (2000).

TECHNICAL UNIVERSITY OF MUNICH
THE UNIVERSITY OF MELBOURNE
THE UNIVERSITY OF MELBOURNE

(Received August 21, 2016)