Integrable matrix theory: Level statistics

Jasen A. Scaramazza, B. Sriram Shastry, and Emil A. Yuzbashyan

Phys. Rev. E **94**, 032106 — Published 2 September 2016

DOI: 10.1103/PhysRevE.94.032106

Integrable matrix theory: level statistics

Jasen A. Scaramazza[1], B. Sriram Shastry[2], Emil A. Yuzbashyan[1]
[1] *Center for Materials Theory, Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08854, USA*
[2] *Physics Department, University of California, Santa Cruz, CA 95064, USA*
(Dated: August 14, 2016)

We study level statistics in ensembles of integrable $N \times N$ matrices linear in a real parameter x. The matrix $H(x)$ is considered integrable if it has a prescribed number $n > 1$ of linearly independent commuting partners $H^i(x)$ (integrals of motion) $\big[H(x), H^i(x)\big] = 0$, $\big[H^i(x), H^j(x)\big] = 0$, for all x. In a recent work, we developed a basis-independent construction of $H(x)$ for any n from which we derived the probability density function, thereby determining how to choose a typical integrable matrix from the ensemble. Here, we find that typical integrable matrices have Poisson statistics in the $N \to \infty$ limit provided n scales at least as $\log N$; otherwise, they exhibit level repulsion. Exceptions to the Poisson case occur at isolated coupling values $x = x_0$ or when correlations are introduced between typically independent matrix parameters. However, level statistics cross over to Poisson at $\mathcal{O}(N^{-0.5})$ deviations from these exceptions, indicating that non-Poissonian statistics characterize only subsets of measure zero in the parameter space. Furthermore, we present strong numerical evidence that ensembles of integrable matrices are stationary and ergodic with respect to nearest neighbor level statistics.

I. INTRODUCTION

It is generally believed that the energy levels of integrable systems[1] follow a Poisson distribution[2–8]. For example, the probability that a normalized spacing between adjacent levels lies between s and $s + ds$ is expected to be $P(s)ds = e^{-s}ds$. In contrast, chaotic systems exhibit Wigner-Dyson statistics, with level repulsion $P(s) \propto s^2$ or s at small s. Moreover, level statistics are often used as a litmus test for quantum integrability even though there are integrable models that fail this test, e.g. the reduced BCS model[5] (which is a particular linear combination of commuting Gaudin Hamiltonians). In the present work, we quantify when and why Poisson statistics occur in quantum integrable models, while also characterizing exceptional (non-Poisson) behavior.

Poisson statistics have been numerically verified on a case-by-case basis for some quantum integrable systems, including the Hubbard[2] and Heisenberg[2, 3] models. On the other hand, general or analytic results on the spectra of quantum integrable models are lacking, in part due to the absence of a generally accepted unambiguous notion of quantum integrability,[9, 10] and in part because existing results usually apply to isolated models instead of members of statistical ensembles like random matrices[11]. Notably, Berry and Tabor showed [4] that level statistics in semiclassical integrable models are always Poissonian as long as the energy $E(n_1, n_2, \dots)$ is not a linear function of the quantum numbers $n_1, n_2, \dots$, i.e. the system cannot be represented as a collection of decoupled harmonic oscillators. As integrability is destroyed by perturbing the Hamiltonian, the statistics are expected to cross over from Poisson to Wigner-Dyson at perturbation strengths as small as the inverse system size[3].

Random matrix theory (RMT)[11, 12] captures level repulsion and other universal features of eigenvalue statistics in generic (non-integrable) Hamiltonians, see e.g. Fig. 2. We recently proposed an integrable matrix theory[13] (IMT) to describe eigenvalue statistics of integrable models. This theory is based on a rigorous notion of quantum integrability and provides ensembles of integrable matrix Hamiltonians with any given number of integrals of motion (see below). It is similar to RMT in that both are ensemble theories equipped with rotationally invariant probability density functions. An important difference is that random matrices do not represent realistic many-body models, while integrable ones correspond to actual integrable Hamiltonians. We therefore have access not only to typical features, but also to exceptional cases and are in a position to make definitive statements about the statistics of quantum integrable models. Here we study the nearest-neighbor level spacing distributions of the IMT ensembles.

The approach of Refs. 10, 13–18 to quantum integrability operates with $N \times N$ Hermitian matrices linear in a real parameter x. A matrix $H(x) = xT + V$ is called *integrable*[16, 17, 19] if it has a commuting partner $\widetilde{H}(x) = x\widetilde{T} + \widetilde{V}$ other than a linear combination of itself and the identity matrix and if $H(x)$ and $\widetilde{H}(x)$ have no common x-independent symmetry, i.e. no $\Omega \neq c\mathbb{1}$ such that $[\Omega, H(x)] = [\Omega, \widetilde{H}(x)] = 0$. Fixing the parameter-dependence makes the existence of commuting partners a nontrivial condition, so that only a subset of measure zero among all Hermitian matrices of the form $xT + V$ are integrable[17].

Further, integrable matrices fall into different classes (types) according to the number of independent integrals of motion. We say that $H(x)$ is a type-M integrable matrix if there are precisely $n = N - M > 1$ linearly independent $N \times N$ Hermitian matrices[20] $H^i(x) = xT^i + V^i$ with no common x-independent symmetry such that

$$\big[H(x), H^i(x)\big] = 0, \quad [H^i(x), H^j(x)] = 0, \qquad (1)$$

for all x and $i, j = 1, \ldots, n$. A type-M family of integrable matrices (*integrable family*) is an n-dimensional vector space[20], where $H^i(x)$ provide a basis. The general member of the family is

$$H(x) = \sum_{i=1}^{n} d_i H^i(x), \tag{2}$$

where d_i are real numbers. The maximum possible value of n is $n = N - 1$, corresponding to type-1 or maximally commuting Hamiltonians.

Examples of well-known many-body Hamiltonians that fit into this definition of integrability are the Gaudin, 1D Hubbard and XXZ models, where x corresponds to the external magnetic field, Hubbard U and the anisotropy, respectively. Note however that these models have various x-independent symmetries, such as the z-component of the total spin, total momentum etc. Taken at a given number of spins or sites, they break down into sectors (matrix blocks) characterized by certain parameter-independent symmetry quantum numbers. Such blocks are integrable matrices according to our definition. For instance, the 1D Hubbard model on six sites with three spin up and three spin down electrons is a direct sum of integrable matrices of various types[17]. Sectors of Gaudin magnets, where the z-component of the total spin differs by one from its maximum or minimum value (one spin-flip), or, equivalently, the one Cooper pair sector of the BCS model are type-1 [16], while other sectors are integrable matrices of higher types.

Prior work[10, 15–18] constructed all type-1, 2, 3 integrable matrices and a certain subclass of arbitrary type-M, determined exact eigenvalues and eigenfunctions of these matrices, investigated the number of level crossings as a function of size and type, and showed that type-1 integrable families satisfy the Yang-Baxter equation. The present work is a continuation of Ref. 13 where we formulated a rotationally invariant parametrization of integrable matrices and derived an appropriate probability density function (PDF) for the parameters, i.e. for ensembles of integrable matrices of any given type. The derivation is similar to that in the RMT and is based on either maximizing the entropy of the PDF or, equivalently, postulating statistical independence of independent parameters and rotational invariance of the PDF. Here we use the results of Ref. 13 to generate and study numerically and analytically level spacing distributions in ensembles of integrable matrices of various types as well as in individual matrices.

Our main results are as follows. For a generic choice of parameters, the level statistics of integrable matrices $H(x)$ are Poissonian in the limit of the Hilbert space size $N \to \infty$ if the number of conservation laws n scales at least as $\log N$, see Fig. 1 for an example. Exceptions to Poisson statistics fall into two categories. First, it is always possible to construct an integrable matrix that has *any* desired level spacing distribution at a given isolated value, $x = x_0$, of the coupling (or external field) parameter. For a typical type-1 matrix there is always a single

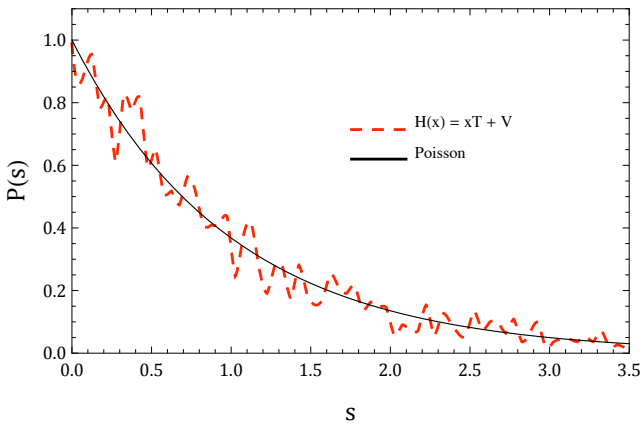

FIG. 1: (color online) The level spacing distribution for a 4000×4000 real symmetric integrable matrix $H(x) = xT + V$ at $x = 1$. This particular matrix is a sum of 200 linearly independent matrices that commute for all values of the real parameter x. Note that the spacing distribution is maximized at $s = 0$, a feature known as level clustering. The smooth curve is a Poisson distribution, which is theorized to be typical of integrable matrices. Compare to the generic real symmetric matrix case in Fig. 2.

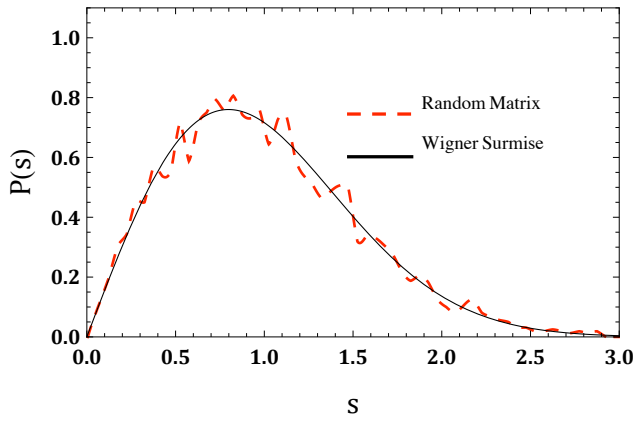

FIG. 2: (color online) The level spacing distribution of a 4000×4000 random real symmetric matrix with entries chosen as independent random numbers from a normal distribution of mean 0 and off-diagonal variance 1/2 (diagonal variance of 1). Such a matrix belongs to the Gaussian orthogonal ensemble (GOE) of real symmetric matrices, studied in random matrix theory (RMT). The main feature of the spacing distribution here is its vanishing for small spacings, also known as level repulsion. The smooth curve is the Wigner surmise $P(s) = \frac{\pi}{2} s e^{-\frac{\pi}{4} s^2}$. See the integrable matrix case in Fig. 1.

value of x where the statistics are Wigner-Dyson. The distribution quickly crosses back over, however, to Poisson at deviations from x_0 of size $\delta x \sim N^{-0.5}$, with the crossover centered at $\delta x \sim N^{-1}$. Second, one obtains non-Poissonian distributions by introducing correlations among the ordinarily independent parameters character-

izing an integrable matrix $H(x)$; the reduced BCS model falls into this category. The statistics again revert to Poisson at $\mathcal{O}(N^{-0.5})$ deviations from such correlations. We also show numerically that as $N \to \infty$, integrable matrix ensembles satisfy two distinct definitions of ergodicity with respect to the nearest-neighbor spacing distribution $P(s)$. Not only are the statistics of a single matrix representative of the entire ensemble, but the statistics of the j-th bulk spacing across the ensemble are independent of j.

In Sect. II, we present numerical results on the level statistics of type-1 matrices, defined to be integrable matrices $H(x)$ with the maximum number $n_{\max} = N - 1$ of linearly independent commuting partners. Section III contains numerical results for integrable matrices with $n \leq n_{\max}$. We present our analytical justification of numerical results using perturbation theory in Sect. IV. Finally, we give numerical results on ergodicity in Sect. V.

II. LEVEL STATISTICS OF TYPE-1 INTEGRABLE MATRICES

A. Type-1 families, primary parametrization

Although our definition of integrable matrices encompasses the general Hermitian case, we restrict our focus in this work to real symmetric matrices. We begin with type-1 integrable $N \times N$ families which contain $N - 1$ nontrivial commuting partners in addition to the scaled identity $(c_1 x + c_2)\mathbb{1}$. Such matrices are the simplest to construct, for the parametrization of type-M integrable families increases in complexity with M. Results on these higher types are deferred to Sect. III.

We first summarize the essential points of the basis-independent type-1 construction of Ref. 13 in order to arrive at the parametrization of Eq. (4) useful for numerical calculations. By considering linear combinations of the $N - 1$ basis matrices, defined in Eq. (2), and the identity, one can prove that every type-1 family contains a particular integrable matrix $\Lambda(x)$ with rank-1 T-part

$$\Lambda(x) = x\,|\gamma\rangle\,\langle\gamma| + E, \tag{3}$$

i.e. $[H(x), \Lambda(x)] = 0$ for all x and any $H(x) = xT + V$ in the family. There is an additional restriction $[V, E] = 0$, which follows from $\mathcal{O}(x^0)$ term in the commutator. It can be shown that the matrices E and V and the vector $|\gamma\rangle$ completely determine a given type-1 matrix $H(x) = xT + V$ modulo an additive constant proportional to the scaled identity.

If we consider any type-1 $H(x)$ in the shared eigenbasis of E and V, we find that the matrix elements of $H(x)$ can be parametrized in terms of the N eigenvalues ε_i of E, the N eigenvalues d_i of V, and the N vector components γ_i of $|\gamma\rangle$. Statistical arguments borrowed from RMT in Ref. 13 identify the ε_i and d_i as two independent sets of eigenvalues drawn from the Gaussian orthogonal ensemble. The γ_i are drawn from a $\delta(1 - |\gamma|^2)$ distribution. With these parameters, *any* $N \times N$ type-1 integrable matrix $H(x) = xT + V$ can be constructed in the following way:

$$\begin{aligned} [H(x)]_{ij} &= x\gamma_i\gamma_j \frac{d_i - d_j}{\varepsilon_i - \varepsilon_j}, \quad i \neq j, \\ [H(x)]_{jj} &= d_j - x\sum_{k\neq j} \gamma_k^2 \frac{d_j - d_k}{\varepsilon_j - \varepsilon_k}. \end{aligned} \tag{4}$$

We call Eq. (4) the "primary" parametrization, which is given specifically in the basis where V is diagonal and can be transformed into any other basis by an orthogonal transformation. Note that the quantities d_j act as coefficients of linear combination of *basis matrices* $H^i(x)$ defined by setting $d_j = \delta_{ij}$ in Eq. (4). Explicitly, nonzero matrix elements of $H^i(x)$ are

$$\begin{aligned} [H^i(x)]_{ij} &= [H^i(x)]_{ji} = x\frac{\gamma_i\gamma_j}{\varepsilon_i - \varepsilon_j}, \quad j \neq i, \\ [H^i(x)]_{jj} &= -x\frac{\gamma_i^2}{\varepsilon_i - \varepsilon_j}, \quad j \neq i, \\ [H^i(x)]_{ii} &= 1 - x\sum_{k\neq i} \frac{\gamma_k^2}{\varepsilon_i - \varepsilon_k}. \end{aligned} \tag{5}$$

and

$$H(x) = \sum_{i=1}^{N} d_i H^i(x). \tag{6}$$

From Eq. (6) we see that the ε_i and γ_i uniquely identify a type-1 commuting family whereas the choice of d_i produces a given member of the family.

To describe the spectrum of $H(x)$, we introduce an additional N parameters $\lambda_j = \lambda_j(x)$ determined by the following equation[16]:

$$\frac{1}{x} = \sum_{k=1}^{N} \frac{\gamma_k^2}{\lambda_j - \varepsilon_k}. \tag{7}$$

One can graphically verify that for any non-degenerate choice of ε_k there are N real solutions λ_j to Eq.(7) that interlace the ε_k. The N eigenvectors $v(x)$ and eigenvalues $\eta(x)$ of $H(x)$ are labelled by λ_j and take the form

$$[v_{\lambda_j}(x)]_k = \frac{\gamma_k}{\lambda_j - \varepsilon_k}, \quad \eta_{\lambda_j}(x) = x\sum_{i=1}^{N} \frac{d_i\gamma_i^2}{\lambda_j - \varepsilon_i}. \tag{8}$$

The components of the (unnormalized) eigenvectors $v_{\lambda_j}(x)$ are independent of the choice of d_i in Eq. (6), and are thus common to any member of the family defined by ε_k and γ_k.

B. Universality of Poisson statistics

Equipped with parametrizations of integrable matrix ensembles based on the number of commuting partners

in a family, we can quantitatively outline both the origin and the robustness of Poisson statistics in these ensembles. We first explore the latter with numerical tests of the statistics of integrable matrices in Sects. II C - III C. For clarity of exposition, the numerical results of Sects. II C, II D and II E are demonstrated strictly for type-1 matrices. In Sect. III, we show that the same results apply generally to a construction of higher type integrable matrix families that by definition contain fewer than the maximum number of conservation laws. We present analytical considerations of numerical results in Sect. IV.

We emphasize that regardless of the choice of parameters we find Poisson level statistics in the overwhelming majority of cases, even near isolated points in parameter space with non-Poissonian statistics. For example, the least biased choice for d_i in Sect. II A enforces GOE statistics at $x = 0$ since $H(0) = V$; by effecting a shift $x \to x + x_0$, the equivalent invariant statement is that each type-1 matrix has a parameter value x_0 such that $H(x_0)$ has Wigner-Dyson statistics. Another exception to Poisson statistics is when d_i and ε_i are correlated so that $d_i = f(\varepsilon_i)$, a smooth function at least over almost the entire range of ε_i. Nonetheless, as soon as we deviate from x_0 or $f(\varepsilon_i)$, the results of Sects. II C and II D show that statistics quickly revert to Poisson at deviations scaling as $\delta \sim N^{-0.5}$ in the limit $N \to \infty$.

Generally, we find that random linear superpositions of basis matrices within a given integrable family are crucial for obtaining Poisson level statistics. Basis matrices themselves, defined in Eq. (6) for the primary type-1 construction and in Eq. (22) for more general integrable matrices, show non-Poissonian statistics with strong level repulsion. Such repulsion washes away, however, for $H(x)$ that are random linear combinations of sufficiently many basis matrices. We see this behavior in Sect. II E for all type-1 matrices, i.e. independent of the number m of basis matrices (conservation laws) in linear combination as long as $m > \mathcal{O}(\log N)$.

We fit all spacing distributions $P(s)$ to the Brody function[21] $P(s, \omega)$, where ω is the Brody parameter

$$P(s, \omega) = a(\omega) s^{\omega} e^{-b(\omega) s^{\omega+1}}. \tag{9}$$

The distribution in Eq. (9) has unit mean and norm with appropriate choices of constants $a(\omega)$ and $b(\omega)$. It interpolates between a Poisson distribution $P(s) = e^{-s}$ at $\omega = 0$ and the Wigner surmise $P(s) = \frac{\pi}{2} s e^{-\frac{\pi}{4} s^2}$ at $\omega = 1$, and hence is a convenient fitting function. The Brody parameter ω can take all values $\omega > -1$, which means it also can detect enhanced level clustering or repulsion.

Note, however, that the Wigner surmise is not the exact nearest neighbor spacing distribution of GOE matrices. One may therefore expect our numerics to produce an $\omega \neq 1$ for GOE matrices. This is indeed the case, see for example Fig. 4 where $\omega \approx .956$ for GOE matrices. The exact distribution $P(s)$ can be found in Ref. 11 and was originally derived by Gaudin in terms of a Fredholm determinant[22]. Using Ref. 23 and a few lines of Mathematica code, we find that the same fitting procedure used for numerically generated matrices produces $\omega \approx 0.957$. Note that it is important to exclude $P(0) = 0$ in the fitting procedure for numerically generated finite-sized matrices.

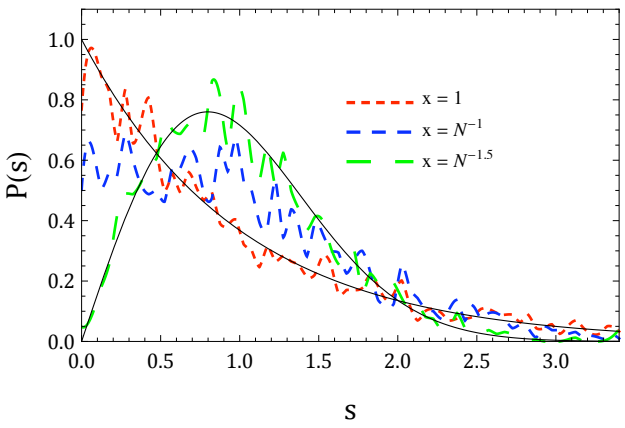

FIG. 3: (color online) Crossover in coupling x of the level statistics of type-1 integrable $N \times N$ matrices $H(x) = xT + V$, $N = 4000$. See Sect. II A for their parametrization. V is a random matrix so that $H(x = 0)$ has level repulsion. Each distribution contains the levels statistics of a single matrix $H(x)$ at a given value of x. Note that some level repulsion has set in by $x = N^{-1}$. Each numerical distribution is fit to the Brody function $P(s, \omega)$ from Eq. (9); for couplings $x = \left(1, N^{-1}, N^{-1.5}\right)$ the fits give $\omega = (0.01, 0.30, 0.94)$, respectively. The solid lines are reference plots of a Poisson distribution $P(s) = e^{-s}$ and the Wigner Surmise $P(s) = \frac{\pi}{2} s e^{-\frac{\pi}{4} s^2}$. See Fig. 4 for more on this crossover.

C. Crossover in coupling parameter x

Here we show that even if the statistics are non-Poissonian at a given coupling value $x = x_0$ (we set $x_0 = 0$), level clustering is restored at small deviations from x_0. For any N, the matrices T and V each have eigenvalues that mostly lie on an $\mathcal{O}(1)$ interval centered about zero. We consider the primary type-1 construction encountered in Eq. (4) and explore the level statistics of large matrices. In Fig. 3, we see qualitatively how the statistics change with x when $N = 4000$. We find Poisson statistics at $x \sim 1$ until a crossover to level repulsion begins near $x = N^{-0.5}$ and ends near $x = N^{-1.5}$.

To verify that the crossover scaling inferred from Fig. 3 is correct for all $N \gg 1$, in Fig. 4 we plot how the Brody parameter ω (see Eq. (9)) evolves with x for various choices of N. It turns out that $\omega(x, N)$ can be fit to a relatively simple function, for any $N \gg 1$

$$\omega(x, N) = \alpha - \beta \tanh\left(\frac{\log_N x - X_0}{Z}\right). \tag{10}$$

The numbers (α, β, X_0, Z) are fit parameters and take

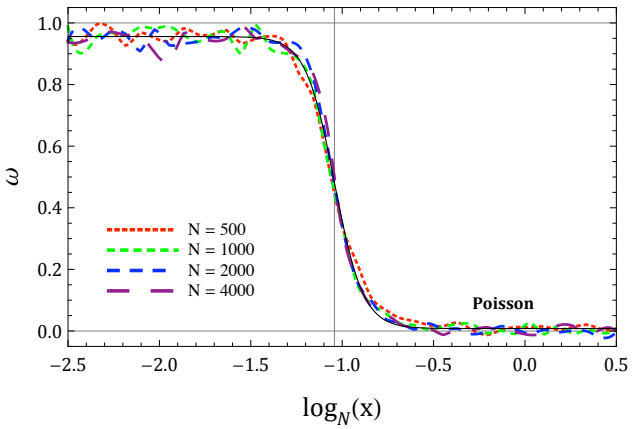

FIG. 4: (color online) Crossover in level statistics with variation of coupling parameter x in type-1 integrable $N \times N$ matrices $H(x) = xT + V$, quantified by the Brody parameter $\omega(x, N)$ from Eq. (9). The two important limits are $\omega = 0$ for Poisson statistics and $\omega = 1$ for random matrix (Wigner-Dyson) statistics. Each plotted value $\omega(x, N)$ is computed for the combined level spacing distribution of several matrices from the ensemble. We extract the crossover scale by fitting $\omega(x, N)$ to Eq. (10) (solid curve) to all curves simultaneously, where most notably $X_0 \sim -1$ for all N investigated, indicating that crossovers to Poisson statistics are centered at that value for integrable matrices $H(x)$ when $H(x = 0)$ has level repulsion. The middle of the crossover is indicated by a vertical line.

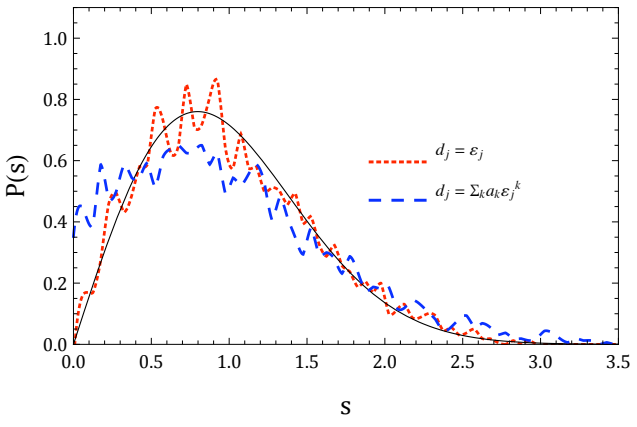

FIG. 5: (color online) Level statistics of two $N \times N$ type-1 integrable matrices $H(x) = xT + V$, $x = 1$ and $N = 4000$, when correlations are introduced between d_j and ε_j (see Eqs. (4, 8), and then Eq. (11) for an example). Note that in contrast to Fig. 3, these integrable matrices exhibit level repulsion even for $x = 1$. Each of the two curves is generated from a single matrix. One numerical curve corresponds to the case when $d_i = \varepsilon_i$ and the other is when $d_i = \sum_{k=1}^{4} A_k h_k(\varepsilon_i)$, where $h_k(z)$ is the k-th order Hermite polynomial and $(A_1, A_2, A_3, A_4) = (2.3, 2.16, -1.46, 0.51)$, chosen randomly. Note that the polynomial dependence weakens the level repulsion as compared to the linear case. If higher order polynomials are included, the level repulsion eventually gives way to Poisson statistics. The solid curve is the Wigner surmise $P(s) = \frac{\pi}{2} s e^{-\frac{\pi}{4} s^2}$. See Fig. 7 for more on this behavior.

the values $(0.482, 0.474, -1.04, 0.157)$ in Fig. 4. Most important is that for any $N \gg 1$ we find $X_0 \sim -1$, which solidifies our claim that the crossover occurs between $x \sim N^{-1.5}$ and $x \sim N^{-0.5}$. Analytical arguments explaining this scaling are given in Sect. IV.

D. Correlations between matrix parameters

In the eigenbasis of V, our parametrization of integrable $N \times N$ matrices is given in terms of about $3N$ independent parameters (up to a change of basis). Through an explicit construction of the probability density function of integrable matrices obtained through basis-independent considerations, Ref. 13 shows that for a typical integrable matrix, d_i and ε_i are indeed uncorrelated. We see in this section that if correlations are introduced between ε_i and d_i, the statistics become non-Poissonian. Small perturbations about these correlations, however, bring the statistics immediately back to Poisson. In this section $x = 1$ for all matrices considered.

Continuing with type-1 matrices in the primary parametrization, Eq. (4), we recall that the eigenvalues η_{λ_j} of such a matrix $H(x) = xT + V$ are given by Eq. (8),where the $\lambda_j = \lambda_j(x)$ are obtained from Eq. (7). As we saw in Sect. II C, a typical choice of parameters will produce Poisson statistics, but this changes if we let d_i be some smooth function of ε_i. The simplest case is shown in Fig. 5 for which $d_i = \varepsilon_i$. As discussed in Refs. 13, 17, $H(x)$ for this choice of parameters describes a sector of the reduced BCS model and, independently, a short range impurity in a weakly chaotic metallic quantum dot studied in Refs. 24, 25.

The level repulsion for this case can be understood by a simple manipulation of Eq. (8) when $d_i = \varepsilon_i$:

$$\begin{aligned} \eta_{\lambda_j} &= x \sum_{i=1}^{N} \frac{\varepsilon_i \gamma_i^2 + \lambda_j \gamma_i^2 - \lambda_j \gamma_i^2}{\lambda_j - \varepsilon_i} \\ &= -x \sum_{i=1}^{N} \gamma_i^2 + \lambda_j, \end{aligned} \tag{11}$$

where we used Eq. (7). Then when $d_i = \varepsilon_i$ the eigenvalues of $H(x)$ are just the λ_j up to an additive constant. For the case when ε_i are random matrix eigenvalues, Ref. 24 derives the joint probability density of the set $\{\varepsilon_i, \lambda_j\}$ and Ref. 25 demonstrates that the λ_j are subject to the same level repulsion as the ε_i. Note also that Eq. (7) implies λ_j lie between consecutive ε_i and therefore the eigenvalues in Eq. (11) can have no crossings at any finite x. Numerically, we have found that λ_j exhibit level repulsion for *any* choice of ε_i (see Fig. 6). Fig. 5 also shows the level repulsion induced when $d_i = \sum_{k=1}^{4} A_k h_k(\varepsilon_i)$, where $h_k(\varepsilon_i)$ is the k-th order Hermite polynomial and A_k are independent random

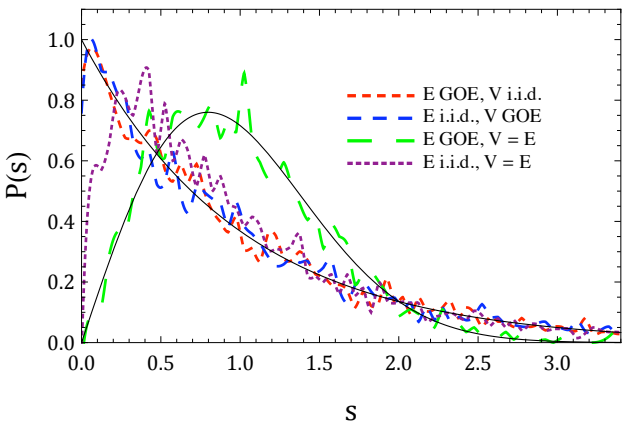

FIG. 6: (color online) Illustrating that conclusions drawn about correlations between d_i (eigenvalues of V) and ε_i (eigenvalues of E) are independent of the particular choice of ε_i. Pictured are four numerically generated nearest-neighbor spacing distributions $P(s)$ for 5000×5000 type-1 matrices, $x = 1$, when the d_i and ε_i are either from a random matrix (GOE) or are independently and identically distributed numbers (i.i.d.) from a normal distribution. Each curve represents the level statistics of a single matrix chosen from the type-1 ensemble. Level repulsion survives in the two cases where d_i and ε_i are correlated ($V = E$), even though the overall shape of $P(s)$ depends on whether E's eigenvalues are GOE or i.i.d. numbers. The solid curves are the usual Poisson distribution $P(s) = e^{-s}$ and the Wigner surmise $P(s) = \frac{\pi}{2} s e^{-\frac{\pi}{4}s^2}$. We do not include plots for different choices of γ_i, which do not affect the general character of the results.

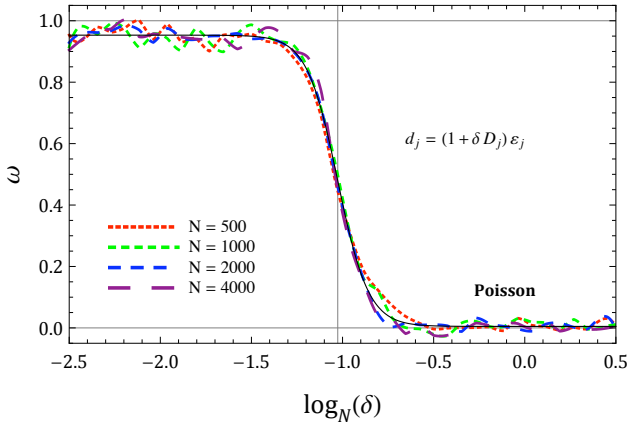

FIG. 7: (color online) Variation in the Brody parameter $\omega(\delta, N)$ when $d_i = \varepsilon_i(1 + \delta D_i)$ in the level statistics of $N \times N$ type-1 integrable matrices $H(x)$ for various N, $x = 1$. The number δ is a parameter controlling the size of the perturbation from correlation, and D_i is an $\mathcal{O}(1)$ random number from a normal distribution. Note that the crossover in δ is very similar to that in x shown in Fig. 4. The numerical curves are fit to the function $\omega(\delta, N)$ given in Eq. (12) (solid curve), with a crossover centered at $X_0 \sim -1$, indicating that crossovers to Poisson statistics are centered at that value. Each plotted value $\omega(\delta, N)$ is computed for the combined level spacing distribution of several matrices from the ensemble. A vertical line indicates the center of the crossover on the plot. For a similar plot for nonlinear functions $d_i(\varepsilon_i)$ see Fig. 8.

numbers drawn from a normal distribution. In this case, the level repulsion is mitigated relative to the case of linear correlation. Sums to higher orders of $h_k(\varepsilon_i)$ (or any higher order polynomial) will eventually bring the statistics back to Poisson.

We now investigate the stability of induced level repulsion in $H(x)$ when correlations between d_i and ε_i are broken. In Fig. 7, we let $d_i = \varepsilon_i(1 + \delta D_i)$ where D_i is an $\mathcal{O}(1)$ random number from a normal distribution and δ is a number controlling the size of the perturbation. The crossover to Poisson statistics as δ increases is very similar to that in Fig. 4, which shows the crossover with x. In fact, we can fit the Brody parameter $\omega(\delta, N)$ to

$$\omega(\delta, N) = \alpha - \beta \tanh\left(\frac{\log_N \delta - X_0}{Z}\right). \qquad (12)$$

Note that Eq. (12) is just Eq. (10) with the substitution $x \to \delta$. We find that the crossover occurs over the range $N^{-1.5} \lesssim \delta \lesssim N^{-0.5}$, indicating that any perturbation to correlations will immediately destroy level repulsion as $N \to \infty$. In particular, Fig. 7 gives $(\alpha, \beta, X_0, Z) = (0.479, 0.474, -1.03, 0.169)$ for linear correlations. This scaling is not restricted to the case $d_i = \varepsilon_i$, as seen in Fig. 8 where we again consider $d_i = \sum_{k=1}^{4} A_k h_k(\varepsilon_i)$ and find a similar crossover with $(\alpha, \beta, X_0, Z) = (0.237, 0.233, -0.914, 0.206)$.

E. Basis matrices: how many conservation laws?

Here we demonstrate that in order to obtain Poisson statistics, the number m of linearly independent conservation laws contained in an $N \times N$ integrable type-1 matrix can be much less than N. Consider a combination of m basis matrices $H^i(x)$ defined in Sect. II A

$$H(x) = \sum_{i=1}^{m} d_i H^i(x), \quad m \leq N - 1. \qquad (13)$$

From the sum in Eq. (13), we can determine the number m needed to obtain Poisson statistics. Individual basis matrices $H^i(x)$ will exhibit level repulsion, and it is only when an integrable matrix is formed from an *uncorrelated* (w.r.t. ε_i, see Sect. II D) linear combination of sufficiently many of them will we observe Poisson statistics. Level repulsion in this case can be qualitatively understood by reasoning that a basis matrix only "contains" one nontrivial conservation law, itself. More concretely, we see from Eq. (8) that the eigenvalues of $H^i(x)$ are $x\gamma_i^2(\lambda_j - \varepsilon_i)^{-1}$, i.e. they are simple, mostly smooth functions of λ_j, which exhibit level repulsion.

Fig. 9 quantifies how many basis matrices m (i.e. conservation laws) are needed for Poisson statistics as a function of N, the matrix size. We find numerically that the plots of the Brody parameter ω (see Eq. (9)) vs. the number m of basis matrices in linear combination can be

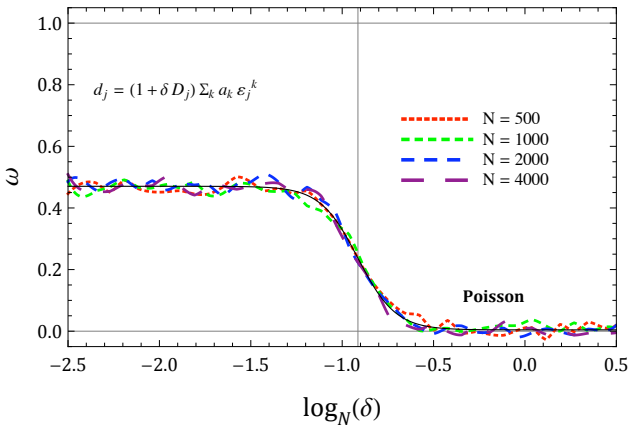

FIG. 8: (color online) Variation in the Brody parameter $\omega(\delta, N)$ when $d_i = \sum_{k=1}^{4} A_k h_k(\varepsilon_i)(1+\delta D_i)$ in the level statistics of $N \times N$ type-1 integrable matrices $H(x=1)$ for various N. Here δ quantifies the deviation from the point $\delta = 0$ where the parameters d_i and ε_i defining the matrices are correlated, D_i and A_k are $\mathcal{O}(1)$ random numbers from a normal distribution, $h_k(z)$ is the k-th order Hermite polynomial. Each $\omega(\delta, N)$ is computed for the combined level spacing distribution of several matrices from the ensemble. The crossover in δ is very similar to that in x in Fig. 4 and in δ for linear correlations in Fig. 7. Because the correlations are nonlinear, the level repulsion is diminished in comparison to previous cases. Despite this, the crossover still demonstrates the same scaling – fitting the data to $\omega(\delta, N)$ given in Eq. (12) (solid curve), with a crossover centered at $X_0 \sim -1$ (vertical line), shows that $\delta \propto N^{-0.5}$ is enough for statistics to revert to Poisson.

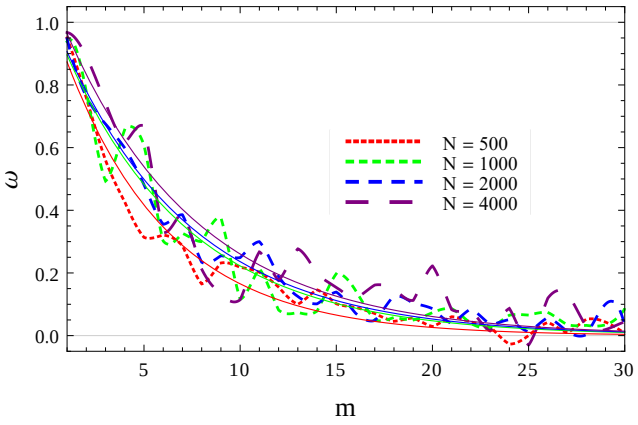

FIG. 9: (color online) The Brody parameter $\omega(m, N)$ (see Eq. (9)) vs. number m of type-1 basis matrices $H^i(x)$ in linear combination $H(x) = \sum_{i=1}^{m} d_i H^i(x)$ for various N, $x = 1$. The fits presume exponential decay and are expressed in terms of two parameters (a, b) from Eq. (14). For $N = (500, 1000, 2000, 4000)$ we find the decay constant $b = (1.15, 1.07, 1.14, 1.21)$, indicating that we only need $m_{\min} \approx \log N$ conservation laws for Poisson statistics to emerge. Figs. 15 and 16 show similar plots for higher types. Each plotted $\omega(m, N)$ is computed for the combined level spacing distribution of several matrices from the ensemble.

fit to a simple function

$$\omega(m, N) = a \exp\left[-\frac{b}{\log N}\, m\right], \tag{14}$$

where a and b are real constants. The fact that for different values of N we find that $b \sim 1$ supports the notion that we need only about $\log N$ conservation laws in order to induce Poisson statistics. We make this claim with caution because we only have data for $500 \leq N \leq 4000$, a range over which $\log N$ does not vary significantly. More precisely, Fig. 9 shows that having $m = \mathcal{O}(1)$ conservation laws is insufficient for inducing Poisson statistics, and that a useful upper bound on the lowest m necessary for Poisson statistics is $m_{\min} < \mathcal{O}(N^{\alpha})$ where $0 < \alpha < 0.20$. We obtain the factor of 0.20 by rewriting Eq. (14) assuming the decay constant has power law dependence on N instead of logarithmic dependence

$$\omega(m, N) = a \exp\left[-\frac{c}{N^{\alpha}}\, m\right]. \tag{15}$$

Numerically we found that the parameter b in Eq. (14) satisifies $1.07 \leq b \leq 1.21$ when $500 \leq N \leq 4000$. By matching exponents between Eq. (15) and Eq. (14) for $(b_1, N_1) = (1.21, 500)$ and $(b_2, N_2) = (1.07, 4000)$, we find a maximum exponent $\alpha = 0.198$.

The basis matrices $H^i(x)$ contained in any integrable $H(x)$ are linearly independent conservation laws. The observed dependance of $P(s)$ on the number m of basis matrices in linear combination is reminiscent of the early work of Rosenzweig and Porter[26] (RP) on the nearest neighbor spacing distribution of superpositions of independent spectra. Although the spectra of basis matrices $H^i(x)$ are not strictly independent and are added together instead of superposed ("superposed" here means "combined into a single list"), we see the same qualitative behavior as described by RP: a single basis matrix has level repulsion, but a sufficiently large number combined have Poisson statistics. In the case of m independent, superposed spectra with vanishing $P(0)$ that contribute equally to the mean level density, the value $P_m(0)$ of the superposed spectrum is given by the RP result

$$P_m(0) = 1 - \frac{1}{m}. \tag{16}$$

We see in Fig. (10) that $P_m(0)$ for m basis matrices in linearly combination differs from the RP result for small m, as expected, but asymptotically approaches Eq. (16) for large m and large N. Thus it seems reasonable to conceptually understand the emergence of Poisson level statistics in integrable matrices $H(x)$ as arising from the existence of conservation laws, whose spectra are statistically independent for large m and N.

Integrable matrix spectra are similar in structure to those of semiclassically integrable models studied by Berry and Tabor[4]. Such spectra are also sums (or simple functions of) rigid spectra, and they have Poisson nearest neighbor level statistics in the semiclassical limit.

Berry's work[27] on semiclassical models shows that

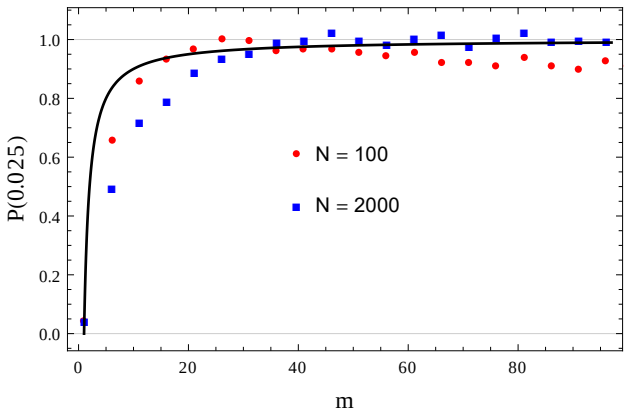

FIG. 10: (color online) Plot of numerically generated $P_m(0.025)$ for *linear combinations* of m type-1 basis matrices, Eq. (13), for $N = 100$ and $N = 2000$ at $x = 1$. The solid curve gives the Rosenzweig-Porter prediction of $P_m(0) = 1-1/m$ for *superpositions* of m independent random matrix spectra. Physically, the RP curve represents $P_m(0)$ for the combined spectra of m blocks of different (parameter-independent) quantum numbers of a Hamiltonian. We note that although different mechanisms are involved in the RP and integrable matrix approach to Poisson statistics, the behavior of $P(0)$ is similar. This gives heurstic justification to why the existence of parameter-dependent conservation laws in $H(x)$ implies Poisson statistics. The sub-Poisson behavior for $N = 100$ is a finite-size effect.

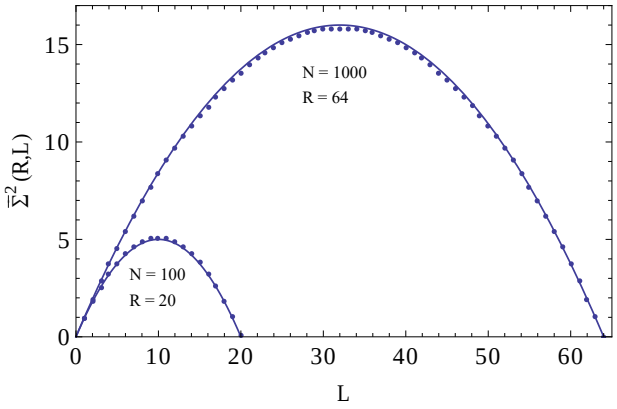

FIG. 11: (color online) Ensemble averaged number variance $\overline{\Sigma}^2(R, L)$ in $N \times N$ type-1 matrices $H(x) = xT + V$ at $x = 1$ for $N = 100$ and $N = 1000$. In order to achieve a constant mean level spacing normalized to unity, we selected the middle $R = 2\sqrt{N}$ eigenvalues from each matrix and used periodic boundary conditions on the list of eigenvalues. The results are in excellent agreement with the Poisson ensemble predictions (solid curves), given by Eq. (17). There is no saturation on the ensemble average. We averaged over 10^4 matrices for $N = 100$ and 500 matrices for $N = 1000$.

longer range spectral statistics of integrable and chaotic models deviate from the predictions of the Poisson ensemble[28] and Gaussian random matrix theory, respectively. Similar behavior occurs in purely quantum systems[29]. An example of such a long range statistic is $\Sigma^2(L)$, the spectral variance of the average number of eigenvalues contained in an interval of length L. For independent random numbers with unit mean spacing in an infinitely large spectrum, $\Sigma^2(L) = L$. For a given Hamiltonian, $\Sigma^2(L)$ will eventually saturate[30] at some $L_{\max}$, which depends on the system's classical periodic orbits and the energy scale.

We find no evidence of saturation of $\Sigma^2(L)$ in type-1 matrices on the ensemble average. Because we work with finite size spectra, we compare numerically generated $\Sigma^2(L)$ to the corresponding Poisson ensemble averaged result for lists of R independent numbers with unit mean spacing and periodic boundary conditions

$$\overline{\Sigma}^2(R, L) = L\left(1 - \frac{L}{R}\right). \qquad (17)$$

The overline indicates an average over the Poisson ensemble. Because numerical unfolding (see Appendix A) introduces spurious effects in long range spectral observables, we instead average over small regions containing $R = 2\sqrt{N}$ eigenvalues in the centers of $N \times N$ matrices where the level density is approximately constant. As seen in Fig. 11, the spectral variance of type-1 matrices satisfies Eq. (17), even at relatively small N.

While there is no saturation on the ensemble average, $\Sigma^2(R, L)$ in the Poisson ensemble has large fluctuations for $L \sim R/2$. Figs. 12 and 13 shows how individual members of the Poisson ensemble and individual type-1 matrices can both exhibit saturation to values of $\Sigma^2(R, L)$ much smaller than Eq. (17) and have a spectral variance greatly exceeding Eq. (17). Type-M matrices, whose construction is detailed in the next section, exhibit similar behavior in $\Sigma^2(R, L)$ for small M, but we have not quantified how precisely $\Sigma^2(R, L)$ changes with increasing M.

Recent work by Prakash and Pandey[31] shows that a two particle non-interacting embedded matrix ensemble[32] exhibits saturation of $\Sigma^2(L)$ on the ensemble average. Embedded matrix ensembles model the structure of many body systems by constructing eigenenergies out of random k-body interactions between m particles, $k < m$. Ref. 31 contains an extended discussion of saturation and helpful references. We do not pursue spectral variance further in this work.

III. STATISTICS OF INTEGRABLE MATRICES OF HIGHER TYPES

A. Ansatz Type-M families

We do not yet have a method for directly generalizing the type-1 primary parametrization from Sect. II A to higher type matrices that by definition have fewer commuting partners. Instead, we present another parametrization that produces a subset of integrable families of any type $M \geq 1$. The construction is in terms of

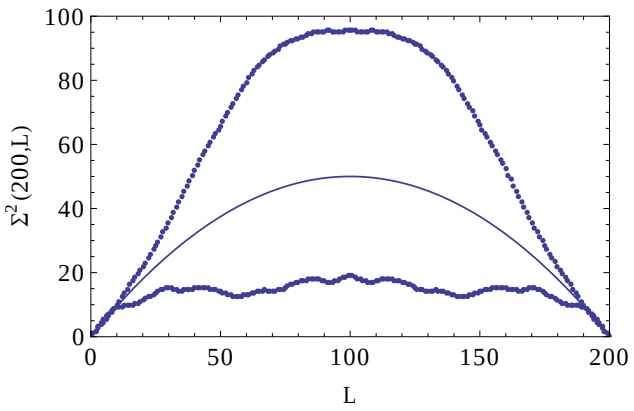

FIG. 12: (color online) Deviations from the Poisson ensemble average Eq. (17) (solid curve) of number variance $\overline{\Sigma}^2(200, L)$ from of two members of the Poisson ensemble. Shown are the number variances of two different lists of 200 independent numbers from a flat distribution in order to illustrate the large fluctuations of long-range spectral observables in the Poisson ensemble. See Fig. 13 for similar behavior in type-1 spectra.

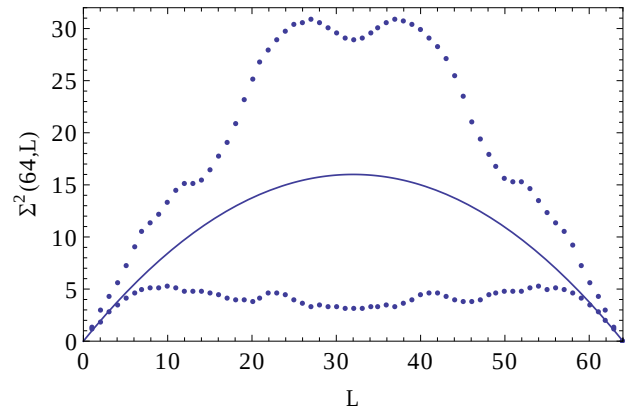

FIG. 13: (color online) Deviations from the Poisson ensemble average Eq. (17) (solid curve) of number variance $\overline{\Sigma}^2(64, L)$ from of two members of the $N = 1000$ type-1 ensemble. Shown are the number variances of two matrices used in the ensemble average of Fig. 11. The saturation observed in the more rigid of the two spectra is remniscent of that seen in members of the Poisson ensemble, see Fig. 12.

$3N+1$ real parameters so that in choosing values for them one obtains a matrix $H(x) = xT+V$ with a desired number n of nontrivial commuting partners ($n = N-M$) and no parameter-independent symmetries. As in the type-1 primary parametrization, the parameters can be traced back to eigenvalues of two commuting constant random matrices and a random vector.

Here we present the results; more details can be found in Ref. 17 while the rotationally invariant construction is given in Ref. 13. Again in the diagonal basis of V, the most general member of an ansatz type-M commuting family is

$$
\begin{aligned}
&[H(x)]_{ij} = x\gamma_i\gamma_j \left(\frac{d_i - d_j}{\varepsilon_i - \varepsilon_j}\right) \frac{\Gamma_i + \Gamma_j}{2}, \quad i \neq j, \\
&[H(x)]_{ii} = \\
&\quad = d_i - x \sum_{j\neq i} \gamma_j^2 \left(\frac{d_i - d_j}{\varepsilon_i - \varepsilon_j}\right) \frac{1}{2} \frac{(\Gamma_i + \Gamma_j)(\Gamma_j + 1)}{\Gamma_i + 1},
\end{aligned}
\tag{18}
$$

where

$$
\begin{aligned}
&\langle i|i\rangle \equiv \sum_{j=1}^{N} \frac{\gamma_j^2}{(\lambda_i - \varepsilon_j)^2}, \\
&d_i = \frac{1}{x_0} \sum_{j=1}^{N-M} \frac{g_j}{\langle j|j\rangle} \frac{1}{\lambda_j - \varepsilon_i}, \\
&\Gamma_i = \pm\sqrt{1 + \frac{1}{x_0} \sum_{j=N-M+1}^{N} \frac{P_j}{\langle j|j\rangle} \frac{1}{\lambda_j - \varepsilon_i}}.
\end{aligned}
\tag{19}
$$

This parametrization gives all type-1, 2, and 3 integrable matrices and only a subset of such for higher types. We call matrices obtained by this construction *ansatz type-M* as opposed to all type-M, these two notions being equivalent for $M = 1, 2, 3$.

Basis-independent considerations from Ref. 13 identify λ_i as eigenvalues of a matrix Λ selected from the GOE and γ_i as selected from a $\delta(1 - |\gamma|^2)$ distribution, as was the case for the primary parametrization of type-1 matrices in Sect. II A. One may alternatively select the ε_i as eigenvalues of a GOE matrix E and from them derive the λ_i. We find that this choice has no effect on the statistics. Unique to the ansatz parametrization are the $(N - M)$ parameters g_i and M parameters P_i. Ref. 13 identifies these parameters as eigenvalues selected from an $N \times N$ GOE matrix G[33] satisfying $[G, \Lambda] = 0$. The sign of Γ_i can be chosen arbitrarily for each i and each set of sign choices corresponds to a different commuting family. The λ_i by construction are solutions of the following equation with arbitrary (but fixed) real $x_0 \neq 0$:

$$
\begin{aligned}
f(\lambda_i) &\equiv \sum_{j=1}^{N} \frac{\gamma_j^2}{\lambda_i - \varepsilon_j} - \frac{1}{x_0} = 0, \\
F(\varepsilon_i) &\equiv \sum_{j=1}^{N} \frac{1}{\langle j|j\rangle} \frac{1}{\lambda_j - \varepsilon_i} - x_0 = 0.
\end{aligned}
\tag{20}
$$

The second line of Eq. (20) follows from the first by writing both the partial fraction decomposition and factorized form of $F(z) = 1/f(z)$ and matching residues. Eqs. (19) and (20) mean that ansatz type-M matrices are written in terms of an auxiliary primary type-1 problem with parameter x_0 and (unnormalized) eigenstates $|i\rangle$, see Eq. (7) and Ref. 13. Note the important distinction between x and x_0 – namely that x is free but x_0 is fixed for a given family of commuting matrices.

Due to the square root in the expression for Γ_i, Eq. (19), a given set of P_i will typically result in a complex set of Γ_i. The matrix $H(x)$ will subsequently be complex symmetric, rather than real, although it will still satisfy all requirements of integrability. Because in this work we study the eigenvalues of real symmetric integrable matrices, we elect to reparametrize Γ_i in a way that guarantees they be real without awkwardly scaling each set of P_i

$$\Gamma_i = \pm\sqrt{\frac{\prod_{j=1}^{N}(\phi_j - \varepsilon_i)}{\prod_{k=1}^{N}(\lambda_k - \varepsilon_i)}}, \tag{21}$$

where the M ϕ_j are real parameters such that (upon ordering ε_j and λ_j for argument's sake) $\varepsilon_j < \phi_j < \lambda_j$ if $x_0 > 0$ and $\lambda_j < \phi_j < \varepsilon_j$ if $x_0 < 0$. The resulting Γ_i are real-valued. As there is no existing basis-independent interpretation for ϕ_j, we simply choose them from a uniform distribution on their allowed intervals. We find that the choice of ϕ_i or P_i to generate the Γ_j has a numerically undetectable effect on the eigenvalue statistics.

Varying parameters g_j produces different matrices within the same commuting family, while varying the remaining parameters $\gamma_i, \lambda_i, \phi_i, x_0$ generates sets of matrices from different families. A natural way to choose a basis for the ansatz type-M commuting family is to define the $n = N - M$ nontrivial $H^k(x)$ such that $g_j = \delta_{kj}$ in Eq. (19) for $1 \leq j \leq N - M$. In other words,

$$\begin{aligned} &H^k(x) = xT^k + V^k \text{ is given by Eq. (18) with} \\ &\qquad d_i \to d_i^k = \frac{1}{x_0}\frac{1}{\langle k|k\rangle}\frac{1}{\lambda_k - \varepsilon_i}. \end{aligned} \tag{22}$$

for $k = 1, \ldots, N - M$. In particular,

$$V^k = \mathrm{Diag}(d_1^k, d_2^k, \ldots, d_N^k) \tag{23}$$

A general member of the commuting family is

$$H(x) = \sum_{k=1}^{N-M} g_k H^k(x). \tag{24}$$

up to a multiple of the identity trivial to the study of level spacing statistics.

Ansatz type-M families have an exact solution in terms of a single equation similar to Eq. (7) given in Ref. 17, which has slight differences in notation as compared to here. To study level statistics of ansatz matrices, we numerically diagonalize them rather than use the computationally cumbersome exact solution.

A fundamental difference between ansatz type-M matrices and the primary type-1 parametrization is that the eigenvalues of the matrix V in the former are heavily constrained by Eq. (19), while in the latter they are free parameters. In particular, as explained in Ref. 13 the primary type-1 V is selected from the GOE, while the ansatz V is a certain primary type-1 matrix evaluated at $x = -x_0$, i.e.

$$V(x_0) = -x_0 T_{H_1} + H_1, \tag{25}$$

where H_1 has $N - M$ arbitrary eigenvalues g_i and M eigenvalues equal to zero. By the results of Sect. II, ansatz $V = V(x_0)$ will typically have Poisson statistics. The resolution to this apparent disconnect between the two parametrizations is that for $|x_0| \ll 1$, $V(x_0)$ will have the eigenvalue statistics of H_1. We argue in Ref. 13 that the $N - M$ g_i are a subset of eigenvalues of an $N \times N$ matrix from the GOE, so that for M not too large and $x_0 \ll 1$ we obtain Wigner-Dyson statistics in ansatz V.

We then forgo studying crossovers in the coupling x of level statistics of ansatz type-M matrices $H(x) = xT + V$ because ansatz V have Poisson statistics for typical parameter choices. Instead, we focus on the behavior of the statistics with respect to parameter correlations, the number M and the number of basis matrices. In all numerical work on ansatz matrices we set $x_0 = 1$, as this is a typical coupling value for the auxiliary type-1 problem.

B. Correlations in ansatz parameters

Building on the results of Sect. II D, here we explore effects of parameter correlations on the statistics in general type-M ansatz matrices. Introducing correlations between d_i and ε_i in this case is more complicated than in Sect. II D because the d_i here are not all independent. Fortunately, Eq. (19) admits a simple way to produce such correlations. As an example, consider the case when $g_j = \lambda_j$

$$\begin{aligned} d_i &= \frac{1}{x_0}\sum_{j=1}^{N-M}\frac{\lambda_j - \varepsilon_i + \varepsilon_i}{\langle j|j\rangle}\frac{1}{\lambda_j - \varepsilon_i} \\ &= \frac{1}{x_0}\,\varepsilon_i\left(\sum_{j=1}^{N-M}\frac{1}{\langle j|j\rangle}\frac{1}{\lambda_j - \varepsilon_i}\right) + (\text{const}) \\ &= \varepsilon_i\left(1 - \frac{1}{x_0}\sum_{j=1}^{M}\frac{1}{\langle j|j\rangle}\frac{1}{\lambda_j - \varepsilon_i}\right) + (\text{const}), \end{aligned} \tag{26}$$

where the second part of Eq. (20) was used. The sums in the third line of Eq. (26) introduce a randomizing factor that has a weak effect for small M but that destroys the correlation between d_i and ε_i at intermediate values of M. Fig. 14 shows the now familiar level statistics crossover in δ for ansatz matrices of different size and type with $g_k = \lambda_k(1 + \delta G_k)$, where G_k is an $\mathcal{O}(1)$ random number chosen from a normal distribution and δ a parameter controlling the size of the perturbation. Just as in Sect. II D, the crossover to Poisson statistics is centered about $\delta \sim N^{-1}$. More generally, we can induce level repulsion in ansatz type-M matrices if $M \ll N$ when $g_k = f(\lambda_k)$, a smooth function of λ_k.

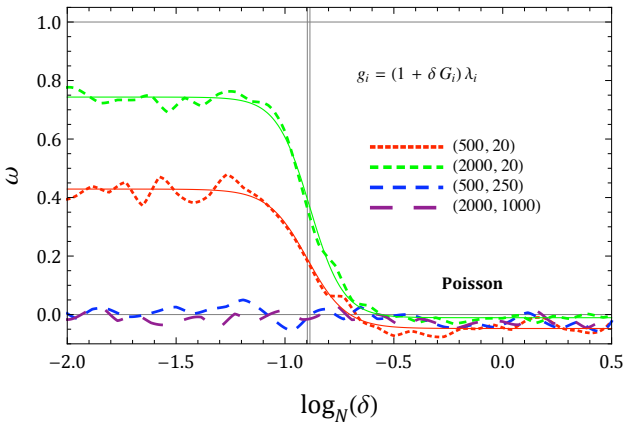

FIG. 14: (color online) Variation in the Brody parameter $\omega(\delta, N)$ when $g_i = \lambda_i(1+\delta G_i)$ in the level statistics of $N \times N$ ansatz type-M integrable matrices $H(x = 1)$, Eq. (18), for various N and M. Ordered pairs in the legend indicate size and type (N, M) of the matrices, δ controls the strength of the perturbation from the point $\delta = 0$ where the parameters g_i and λ_i defining these integrable matrices are correlated, and G_i is an $\mathcal{O}(1)$ random number from a normal distribution. The crossover in δ for small M is similar to the primary type-1 crossovers in δ and x seen in Figs. 4,7 and 8. For larger M, correlations cannot be introduced by this method, see Eq. (26). Despite type-M matrices having fewer than the maximum number of conservation laws, the crossover still demonstrates the scaling given in Eq. (12) (solid curves) with a crossover centered around $X_0 \sim -1$ (vertical line). As before, deviations from correlation of size $\delta \propto N^{-0.5}$ are enough for the statistics to become Poisson. Each plotted value $\omega(\delta, N)$ is computed for the combined level spacing distribution of several matrices from the ensemble. For the case of correlations in ansatz matrices, we choose all $\Gamma_k > 0$ in order to avoid pathological statistics in $H(x)$.

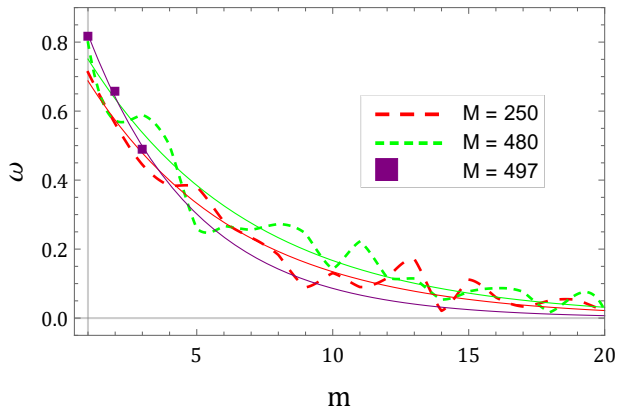

FIG. 15: (color online) Graph of the Brody parameter $\omega(m, N)$ given by Eq. (9) vs. number m of ansatz type-M basis matrices $H^k(x)$, see Eq. (22), contained in linear combination $H(x) = \sum_{k=1}^{m} g_k H^k(x)$ for $N = 500$, $x = 1$. The fits presume exponential decay and are expressed in terms of two parameters (a, b) from Eq. (14). For $M = (250, 480)$ we find the decay constant $b = (1.13, 1.04)$, indicating that we only need $m_{\min} \approx \log N$ conservation laws for Poisson statistics to emerge, independent of type. We do not observe Poisson statistics for $M = 497$ because the maximum number of nontrivial basis matrices is 3 in this case, and we see that we need at least ~ 15 conservation laws for Poisson statistics to start emerging for $N = 500$. See Fig. 16 for a similar plot for $N = 2000$ and Fig. 9 for the same concept in type-1 matrices. Each plotted value $\omega(m, N)$ is computed for the combined level spacing distribution of several matrices from the ensemble.

C. Basis matrices: ansatz higher types

We now generalize the type-1 results of Sect. II E to apply to all ansatz type-M matrices. Recall that a general ansatz type-M matrix $H(x) = xT + V$ can be written as a linear combination of basis matrices $H^k(x)$ for which $g_i = \delta_{ik}$ (see Eq. (24)).

We see again in Figs. 15 and 16 that Poisson statistics emerge for relatively small linear combinations of basis matrices. Denoting m as the number of conservation laws contained in a linear combination, i.e.

$$H(x) = \sum_{i=1}^{m} g_k H^k(x), \quad m \leq N - M, \tag{27}$$

we investigate the Brody parameter $\omega(m, N)$ from Eq. (14). In Fig. 15, $N = 500$, $\omega(m, N)$ decays to zero as a function of m in nearly the same way for $M = 470$ as for $M = 20$. It is only for very large M, such as $M = 497$, that level clustering is forbidden, and this only because we can use a maximum of 3 nontrivial basis matrices. Similar behavior emerges for $N = 2000$ in Fig. 16. For all N and M tested we find $b \sim 1$ (with precise values given in the captions). Therefore, we can estimate a similar bound as in Sect. II E for the minimum number of conservation laws needed for Poisson level statistics, namely $m_{\min} < \mathcal{O}(N^{\alpha})$ where $0 < \alpha < 0.24$, obtained from the $M = N/2$ cases. Since m cannot exceed the total number of conservation laws $n = N - M$ for type-M matrices, this provides a lower bound $n_{\min} = m_{\min} < \mathcal{O}(N^{\alpha})$ consistent with $m_{\min} \approx \log N$.

IV. ANALYTICAL RESULTS: PERTURBATION THEORY

Some of the numerical observations found in Sects. II and III can be understood using perturbation theory in the parameter x. We restrict our analysis to the primary type-1 parametrization because our arguments for this case are much more transparent than for the ansatz construction. The analysis for ansatz matrices is similar.

The eigenvalues $\eta_m(x)$ of $H(x)$ to first order in x are given by the second equation in Eq. (4), where we set constant $|\gamma_j|^2 = N^{-1}$ for clarity and to achieve proper

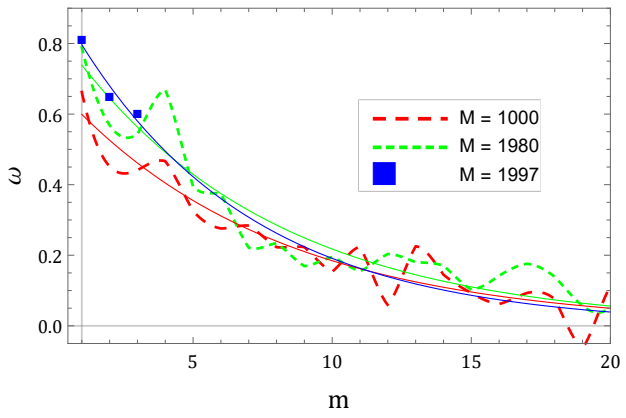

FIG. 16: (color online) Brody parameter $\omega(m,N)$ (see Eq. (9)) vs. number m of ansatz type-M basis matrices $H^k(x)$, see Eq. (22), contained in linear combination $H(x) = \sum_{k=1}^{m} g_k H^k(x)$ for $N = 2000$, $x = 1$. The fits presume exponential decay and are expressed in terms of two parameters (a,b) from Eq. (14). For $M = (1000, 1980)$ we find the decay constant $b = (0.99, 1.03)$, indicating that we only need $m_{\rm min} \approx \log N$ conservation laws for Poisson statistics to emerge, independent of type. We do not observe Poisson statistics for $M = 1997$ because the maximum number of nontrivial basis matrices is 3 in this case, and we see that we need at least ~ 20 conservation laws for Poisson statistics to start emerging for $N = 2000$. See Fig. 15 for a similar plot for $N = 500$ and Fig. 9 for the same concept in type-1 matrices. Each plotted value $\omega(m,N)$ is computed for the combined level spacing distribution of several matrices from the ensemble.

scaling for large N

$$\eta_m(x) \approx d_m - \frac{x}{N} \sum_{j \neq m} \left(\frac{d_m - d_j}{\varepsilon_m - \varepsilon_j} \right). \qquad (28)$$

The first term comes from V, which has a Wigner-Dyson $P(s)$, and the second term from T, which is determined by the integrability condition and whose level statistics we do not control. Let us estimate the x at which the two terms in Eq. (28) become comparable. Note that d_k and ε_k both lie on $\mathcal{O}(1)$ inteverals so that T and V scale in the same way for large N. Suppose ε_k are ordered as $\varepsilon_1 < \varepsilon_2 < \cdots < \varepsilon_N$. When d_k and ε_k are uncorrelated $d_m - d_j$ is $\mathcal{O}(1)$ when j is close to m, i.e. when $(\varepsilon_m - \varepsilon_j) = \mathcal{O}(N^{-1})$. The second term in Eq. (28) is then $x c_m \ln N$, where $c_m = \mathcal{O}(1)$ is a random number only weakly correlated with d_m. We performed simple numerical tests that confirm this scaling argument.

If we now order d_m, c_m in general will not be ordered, i.e. if $d_{m+1} > d_m$ is the closest level to d_m and therefore $(d_{m+1} - d_m) = \mathcal{O}(N^{-1})$, the corresponding difference $(c_{m+1} - c_m) = \mathcal{O}(1)$. The contributions to level-spacings from the two terms in Eq. (28) become comparable for $x = x_c \approx 1/(N \ln N)$. It makes sense that the second term introduces a trend towards a Poisson distribution because it is a (nonlinear) superposition of ε_k and d_k – eigenvalues of two uncorrelated random matrices. Thus, we expect a crossover from Wigner-Dyson to Poisson statistics near $x = x_c$. In our numerics we observe a crossover over the range $N^{-1.5} \lesssim x \lesssim N^{-0.5}$ centered about $x_c \sim N^{-1}$ likely because we do not reach large enough N to detect the log component of the crossover.

This argument breaks down when $d_k = f(\varepsilon_k)$, since in this case $(d_m - d_j) = \mathcal{O}(N^{-1})$ when $(\varepsilon_m - \varepsilon_j) = \mathcal{O}(N^{-1})$. The two terms in Eq. (28) become comparable only at $x = \mathcal{O}(1)$; moreover, the second term no longer trends towards Poisson statistics. Relaxing the correlation between d_k and ε_k with $d_k = f(\varepsilon_k)(1 + \delta D_k)$, $D_k = \mathcal{O}(1)$, and going through the same argument, one expects a crossover to Poisson statistics at $\delta = \mathcal{O}(1/N \ln N)$ when $x = \mathcal{O}(1)$.

The level repulsion observed in basis matrices is a consequence of the level repulsion implicit in the parameters λ_i, independent of the choice of ε_i, see the text below Eq. (11) and Fig. 6. Indeed, basis matrices $H^i(x)$ in the primary type-1 parametrization, Eq. (4), have eigenvalues $\eta_j^i(x) = x\gamma_i^2(\lambda_j - \varepsilon_i)^{-1}$, which is a smooth function of λ_j except near ε_i. The $\eta_j^i(x)$ therefore inherit the level repulsion of the λ_j. Analogous reasoning applies to ansatz basis matrices.

V. ERGODICITY IN INTEGRABLE MATRIX ENSEMBLES

The discussion and figures in this section make frequent reference to the "primary" construction of type-1 integrable matrices and the "ansatz" construction of type-M integrable matrices. These parametrizations are introduced in Sect. II A and Sect. III A, respectively. Ensemble averages are taken with respect to the probability distributions for integrable matrices introduced in Ref. 13.

One of the goals of this work is to determine the extent to which ensembles of integrable matrices are "ergodic". Intuitively, an ensemble is called ergodic if a single randomly selected member has properties that are typical of the entire ensemble. Bohigas and Gianonni[34] expound the subject in generality for random matrices, and here we focus numerically on the meaning of ergodicity with regards to the nearest-neighbor level spacing distribution of integrable matrices. Rigorous results on ergodicty for Gaussian ensembles and the Poisson ensemble were derived by Pandey[35].

We distinguish between three separate ways of generating nearest-neighbor eigenvalue spacing distributions for $N \times N$ integrable matrix ensembles. We call $P_{i,N,R}(s)$ the level spacing distribution, normalized to unity, of the i-th member of the ensemble obtained from a spectral region R containing many eigenvalues (infinitely many as $N \to \infty$). The normalized distribution of spacings in R from all matrices in the ensemble is called $\mathcal{P}_{N,R}(s)$. A third way to characterize spacing statistics is through the normalized distribution of the j-th eigenvalue spacing of

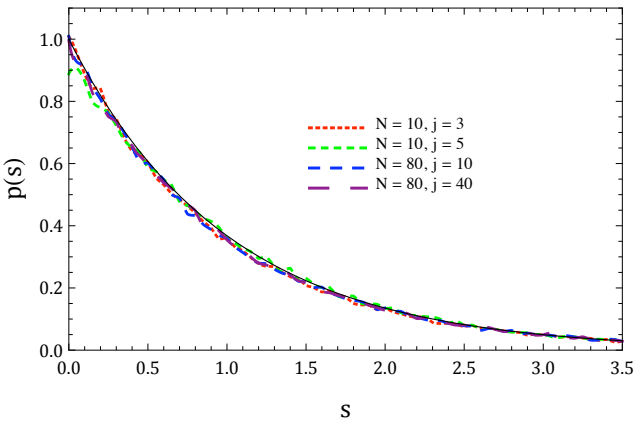

FIG. 17: (color online) Demonstrating the stationary property Eq. (29) in type-1 $N \times N$ matrices $H(x)$, $x = 1$ in the primary parametrization. The four numerical curves show the statistics $p_{N,j}(s)$ for $(N, j) = (10, 3)$, $(10, 5)$, $(80, 10)$ and $(80, 40)$, each containing 10^5 eigenvalue spacings. The statistics are nearly independent of j for $N = 10$, and for $N = 80$ there is no perceptible difference between $j = 10$ and $j = 40$. The solid line is a Poisson distribution $p(s) = e^{-s}$. Stationarity is shown to hold also for type-M ansatz matrices in Fig. 18.

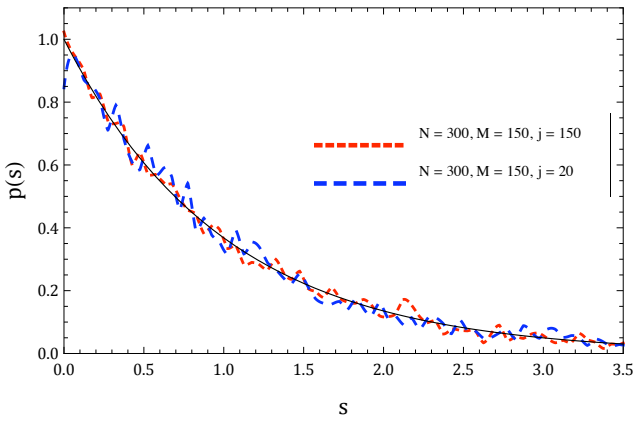

FIG. 18: (color online) Demonstrating the stationarity property Eq. (29) in ansatz type-150 $N \times N$ matrices $H(x)$, $x = 1$ and $N = 300$. The two numerical curves show the statistics $p_{N,j}(s)$ for $(N, j) = (300, 150)$ and $(300, 20)$, each containing $\sim 10^4$ eigenvalue spacings. The statistics are nearly independent of j, although higher N would be needed in order for the differences to disappear. The solid line is a Poisson distribution $p(s) = e^{-s}$.

all matrices in the ensemble, which we call $p_{N,j}(s)$. Both the regions R and the numbers j are stipulated to be far from the edges of the spectrum. In general, $P_{i,N,R}(s)$, $\mathcal{P}_{N,R}(s)$ and $p_{N,j}(s)$ are distinct distributions. Conceptually, $\mathcal{P}_{N,R}(s)$ and $p_{N,j}(s)$ are ensemble properties while $P_{i,N,R}(s)$ characterizes the spectrum of an individual matrix. In the following definitions, we assume that the spacing distributions converge to a well-defined limit as $N \to \infty$, unlike known pathological examples such as the semiclassical spacing distribution of a harmonic chain[4]. This assumption is supported numerically.

We now describe a precise notion[35] of ergodicity that characterizes the limiting behavior of $P_{i,N,R}(s)$, $\mathcal{P}_{N,R}(s)$ and $p_{N,j}(s)$ as $N \to \infty$. First, we must determine whether $p_{N,j}(s)$ is asymptotically stationary, i.e. independent of j

$$\lim_{N\to\infty} p_{N,j}(s) = p(s). \tag{29}$$

In the case of type-1 matrices in the primary parametrization, we see in Fig. 17 that the graphs of two different $p_{10,j}(s)$ closely resemble those of two different $p_{80,j}(s)$, the latter of which are clearly Poisson. The same is true for ansatz matrices of any type, but the convergence to a Poisson distribution does not become apparent until $N = 300$ as in Fig. 18. We conclude that Eq. (29) is true for integrable matrices.

We now turn to the notion of spectral averaging, i.e. the function $P_{i,N,R}(s)$. If Eq. (29) holds, the ensemble averaged $P_{i,N,R}(s)$, called $\mathcal{P}_{N,R}(s)$, satisfies

$$\lim_{N\to\infty} \mathcal{P}_{N,R}(s) = p(s), \tag{30}$$

independent of the region R. In practice, we numerically unfold the spectrum (see Appendix A) in order to take into account any effects a non-stationary mean level spacing can have on $P_{i,N,R}(s)$, which characterizes fluctuations about the mean level spacing. In this work, we say integrable matrices are spectrally stationary if

$$\lim_{N\to\infty} P_{i,N,R}(s) = P_i(s), \tag{31}$$

and ergodic with respect to nearest neighbor level statistics if

$$P_i(s) = p(s). \tag{32}$$

Two points are to be made about Eq. (31) and Eq. (32). First, Eq. (31) is similar in spirit to, but not implied by, Eq. (29). Figs. 19-22 show for various integrable matrices, basis matrices included, that the level statistics from a single large matrix, $P_{i,N,R}(s)$, do not depend on the spectral region R used.

Second, the limiting distribution is independent of the index i, which means that a single matrix's spacing distribution is typical of the ensemble. In rigorous work on Gaussian ensembles[35], this is proved by showing the ensemble averaged variance of $P_{i,N,R}(s)$ vanishes as $N \to \infty$. In this work, we compare numerically generated graphs of spectral spacing distributions to ensemble averaged ones for large N. By comparing Figs. 23, 24 to Figs. 19, 20, we see that for large N, $P_{i,N,R}(s) \to p(s)$.

The properties of stationarity and ergodicity are useful if they set in quickly for small N, because smaller matrices are more accessible both analytically and computationally. A classic example in Gaussian random matrix theory is the Wigner surmise, derived from 2×2 matrices

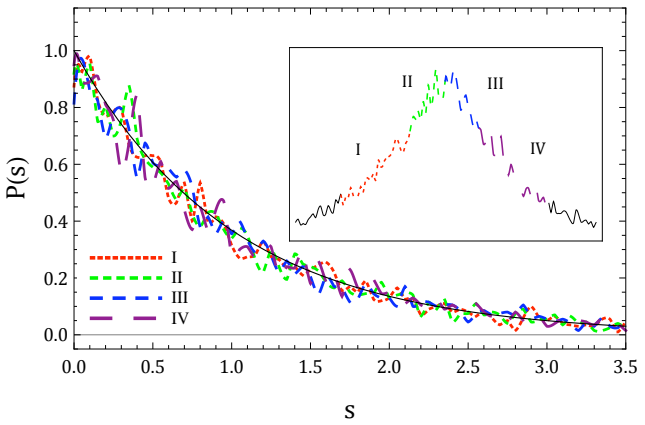

FIG. 19: (color online) Demonstrating spectral stationarity Eq. (31) in type-1 matrices. Shown are the level statistics $P_{i,N,R}(s)$ of a single (i-th member of the ensemble) type-1 integrable matrix $H(x)$, $x = 1$ and $N = 20000$, for different regions R of its spectrum containing 4000 eigenvalues each. The inset shows the density of states of this matrix and indicates which numerical curve corresponds to which region R. The distributions $P_{i,N,R}(s)$ shown are independent of R, indicating that type-1 matrix spectra are stationary with respect to nearest neighbor level statistics. Noting that these distributions are Poisson, $P_{i,N,R}(s) \approx e^{-s}$ (solid curve) and comparing to Fig. 23 which gives $\mathcal{P}_{N',R}(s) \approx e^{-s}$ for $N' = 2000$, we see that ergodicity, Eq. (32), is satisfied for type-1 integrable matrices.

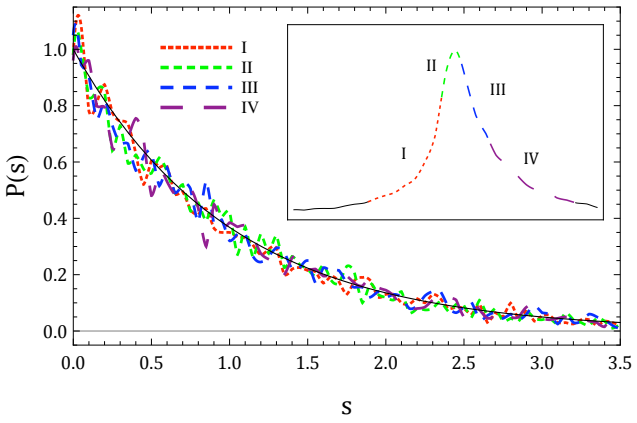

FIG. 20: (color online) Level statistics $P_{i,N,R}(s)$ of a single integrable matrix $H(x)$, $x = 1$, $N = 20000$ and $M = 10000$, for different regions R of its spectrum (the subscript i indicates $H(x)$ is the i-th matrix in the ensemble) containing 4000 eigenvalues each. Inset: the density of states of $H(x)$ showing the correspondence between the distributions and regions R. The distributions $P_{i,N,R}(s)$ are independent of R, indicating that type-M matrix spectra are stationary with respect to nearest neighbor level statistics, i.e., Eq. (31) holds. Noting that these distributions are Poisson, $P_{i,N,R}(s) \approx e^{-s}$ (solid curve) and comparing to Fig. 24 which gives $\mathcal{P}_{N',R}(s) \approx e^{-s}$ for $N' = 2000$ $M' = 1000$, we verify the ergodic property, Eq. (32).

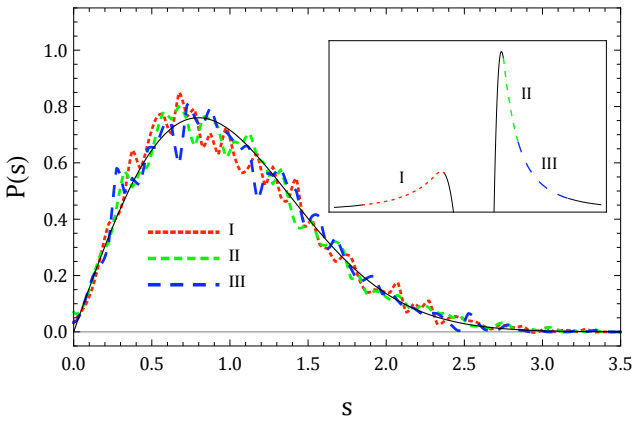

FIG. 21: (color online) Demonstrating spectral stationarity, Eq. (31), in level statistics of primary type-1 basis matrices (defined in Sect. II A). Shown are the level statistics $P_{i,N,R}(s)$ of a single type-1 integrable basis matrix, $x = 1$ and $N = 20000$, for different regions R of its spectrum (the subscript i indicates $H(x)$ is the i-th matrix in the ensemble). Each spectral region R contains 4000 eigenvalues. The inset shows the density of states of this matrix and indicates which numerical curve corresponds to which region R. The distributions $P_{i,N,R}(s)$ shown are independent of R, indicating that type-1 basis matrix spectra are stationary with respect to level statistics. Even though there is a band gap, the level statistics on either side of the gap are the same. The solid curve is the Wigner surmise $P(s) = \frac{\pi}{2} s e^{-\frac{\pi}{4} s^2}$.

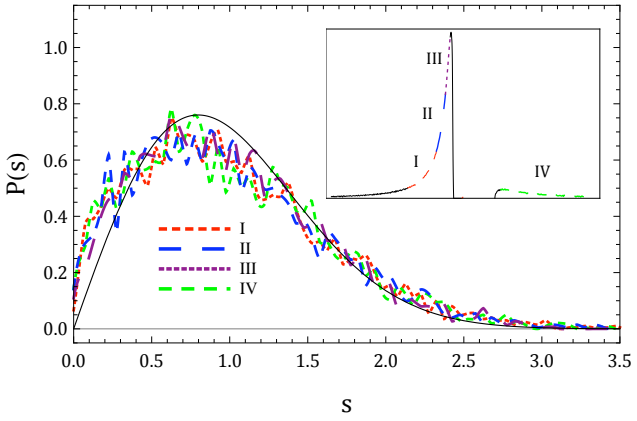

FIG. 22: (color online) Demonstrating spectral stationarity, Eq. (31) in level statistics of ansatz basis matrices (defined in Sect. III A). Shown are the level statistics $P_{i,N,R}(s)$ of a single type-10000 integrable ansatz basis matrix, $x = 1$ and $N = 20000$, for different regions R of its spectrum (the subscript i indicates $H(x)$ is the i-th matrix in the ensemble). The inset shows the density of states of this matrix and indicates which numerical curve corresponds to which region R. The distributions $P_{i,N,R}(s)$ shown are independent of R, indicating that type-M basis matrix spectra are stationary with respect to level statistics. Even though there is a band gap, the level statistics on either side of the gap are the same. The solid curve is the Wigner surmise $P(s) = \frac{\pi}{2} s e^{-\frac{\pi}{4} s^2}$. Regions I - III use 4000 eigenvalues apiece, while region IV uses only 3000 and gets to within 1% of the spectrums edge.

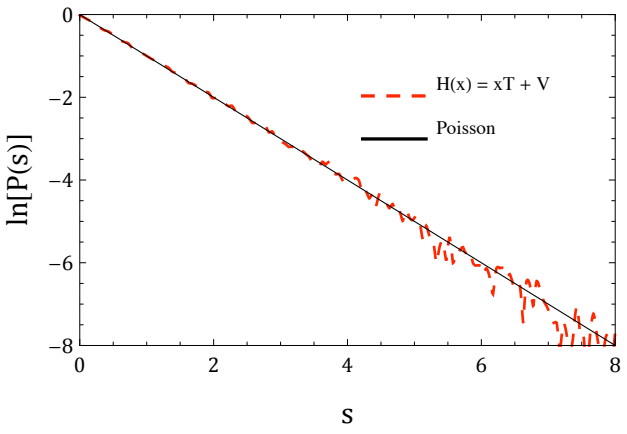

FIG. 23: (color online) Demonstrating ergodicity Eq. (32) in type-1 matrices (continuing from Fig. 19). A plot of $\ln \mathcal{P}_{N,R}(s)$, $N = 2000$ for 100 type-1 matrices in the primary parametrization. We do not specify the spectral region R because Fig. 19 shows that the statistics are independent of R. The solid line is $f(s) = -s$, indicating that $\mathcal{P}_{N,R}(s)$ is indeed Poisson for $N = 2000$ type-1 matrices.

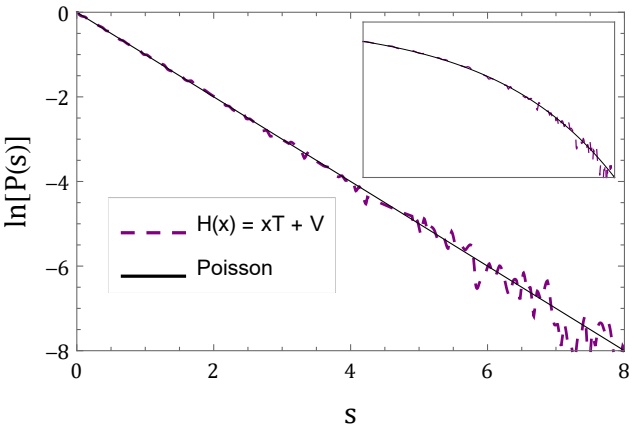

FIG. 24: (color online) Demonstrating ergodicity Eq. (32) in type-M ansatz matrices (continuing from Fig. 20). A plot of $\ln \mathcal{P}_{N,R}(s)$, $N = 2000$ for 100 type $M = 1000$ matrices in the ansatz parametrization. We do not specify the spectral region R because Fig. 20 shows that the statistics are independent of R. The solid line is $f(s) = -s$, indicating that $\mathcal{P}_{N,R}(s)$ is indeed Poisson for $N = 2000$ type-1000 matrices. Inset: Log-log plot of the same data.

(see Fig. 2), which is extremely useful for characterizing $p(s)$ in the GOE.

We have seen that the nearest neighbor level statistics of integrable matrices $H(x)$ are generally stationary and ergodic, but the property does not set in for small N as quickly as it does for Gaussian random matrices. As an example, Figs. 25, 26 show $p_{3,2}(s)$, the distribution of the 2nd eigenvalue spacing for $N = 3$, $M = 1$. This distribution differs markedly from a Poisson distribution, especially in the small s and large s regions. For small s

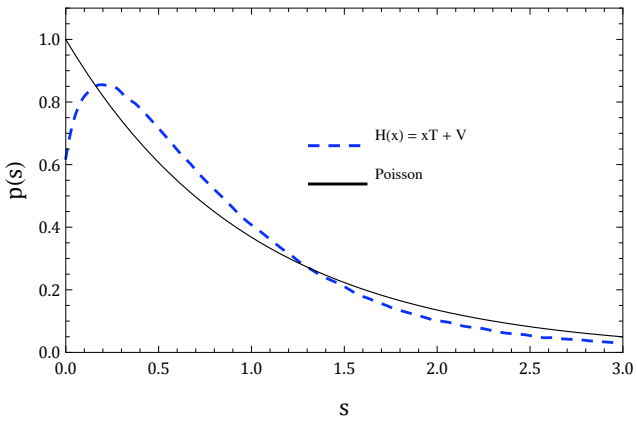

FIG. 25: (color online) Plot of the statistics $p_{3,2}(s)$, the second spacing of 10^6 type-1 integrable matrices $H(x)$ of size $N = 3$ with $x = 1$. The distribution is not Poisson (solid line) and actually has a power law tail (see Fig. 26 for more on the tail). In order to observe the limit $p(s)$ of type-1 integrable matrices, defined in Eq. (29), we need to go to larger N as in Fig. 17.

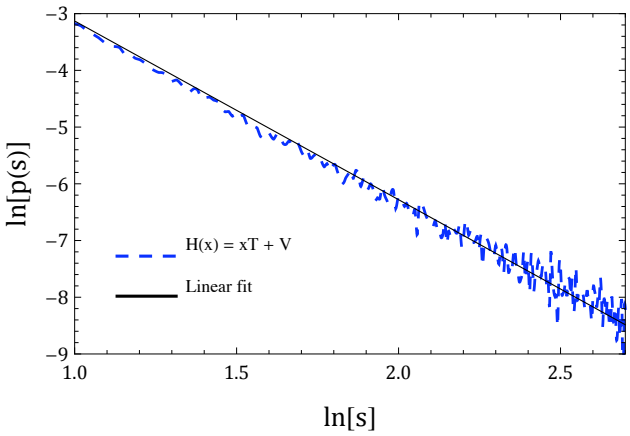

FIG. 26: (color online) Log-log plot of the tail of the distribution $p_{3,2}(s)$ shown in Fig. 25, the statistics of the second spacing of 10^6 primary type-1 integrable matrices $H(x)$ of size $N = 3$ with $x = 1$. The linear fit $f(s) = -3.15s + 0.02$ shows that this portion of the tail of the distribution $p_{3,2}(s)$ follows a power law $s^{-\alpha}$ with exponent $\alpha \approx 3.15$. Because the distribution $p_{N,j}(s)$ transitions to Poisson for large N, as evidenced by Fig. 17 for type-1 primary matrices and Fig. 18 for type-M ansatz matrices, we conclude that exponential behavior in the far tail of $p_{N,j}(s)$ likely emerges only in the limit $N \to \infty$.

there is slight level repulsion and for large s Fig. 26 shows that the decay of $p_{3,2}(s)$ is a power law. Numerical data generated in Sects. II and III used both $P_{i,N,R}(s)$ and $\mathcal{P}_{N,R}(s)$ to represent level statistics of integrable matrices. The results of this section show that for large N, it is valid to treat these two distinct distributions as equal.

VI. CONCLUSION

Just as ensemble averages in ordinary RMT are used to predict the average behavior of generic quantum systems, there now exists an analogous ensemble theory-integrable matrix theory - which we have used to firmly establish the source of Poisson statistics and exceptions in quantum integrable models.

The goal of this work was to demonstrate two properties of ensembles of type-M integrable $N \times N$ matrices linear in a coupling parameter $H(x) = xT+V$ as $N \to \infty$:

1) The nearest neighbor spacing distribution $P(s)$ is Poisson, $P(s) = e^{-s}$, for generic choices of parameters for almost all M. There are cases of level repulsion, but they correspond to sets of measure zero in parameter space.

2) Integrable matrix ensembles are both stationary and ergodic with respect to nearest neighbor level statistics as defined in Sect. V. It remains to show whether this ergodicity extends to longer range spectral statistics, such as the number variance $\Sigma^2(L)$.

We find that integrable $N \times N$ matrices $H(x)$ have Poisson statistics as long as the number of conservation laws exceeds $n_{\min} \approx \log N$ (or at most $n_{\min} < N^{0.24}$). Basis-independent considerations require (for type-1) GOE statistics at a fixed x_0, but we find that Poisson statistics return at deviations $\delta x \sim N^{-0.5}$. Correlations between otherwise independent parameters also induce level repulsion, but Poisson statistics again return at $\mathcal{O}(N^{-0.5})$ deviations from such correlations. In both cases the crossover occurs roughly over the range $N^{-1.5} \lesssim \delta \lesssim N^{-0.5}$.

Some parameter choices produce matrices that correspond to sectors of certain known quantum integrable models, although general parameter choices do not map to known models. Most important is that, in addition to the linearity in x condition, the ensembles of matrices studied in this work are only constrained by symmetry requirements just like the Gaussian random matrix ensembles. The only difference here is that in the integrable case there are many more symmetries, and they are parameter dependent. We therefore expect our results to apply generally to quantum integrable models with coupling parameters.

Although we justified the numerical results to a certain degree using perturbation theory, an analytic justification for Poisson statistics for integrable matrices is still lacking. Given the relatively simple construction of integrable matrices through basis-independent relations (i.e. matrix equations) involving familiar RMT quantities such as GOE matrices and random vectors[13], we surmise that an analytic demonstration of our numerical results might be feasible – especially in the type-1 case, see, e.g., the discussion below Eq. (11) and Refs. 24, 25.

It is interesting to note that many-body localized[36] (MBL) systems are also expected to display Poisson level statistics[37, 38], and there exist random matrix ensembles which model localization and its statistical signatures[39, 40]. Such ensembles are basis-dependent, which is natural because localization is a basis-dependent property. The commutation requirements of integrable systems, however, are basis-independent, and therefore so is the accompanying integrable matrix theory. A priori, many-body localization and integrability are two independent concepts[41]. Despite this fact, integrable matrices do exhibit a parameter-dependent localization property[43] in which almost all eigenstates of the matrix $H(x) = xT + V$ are localized in the basis of V for all values of x. The stability of this property when a random matrix perturbation is added to $H(x)$, including the possibility of a multifractal phase accompanying the localized and delocalized regimes[40], is the subject of future study.

Acknowledgments

This work was supported in part by the David and Lucille Packard Foundation and by the National Science Foundation under Grant No. NSF PHY11-25915. E. A. Y. also thanks the University of California at Santa Cruz and Kavli Institute for Theoretical Physics, where a significant part of this research was conducted, for hospitality. The work at UCSC was supported by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES) under Award # FG02-06ER46319. We acknowledge Daniel Hansen's contribution at an early stage of this work, especially to the unfolding technique in Appendix A. We also thank Joel Lebowitz for helpful discussions. Finally, we thank the PRE referees for suggestions and questions that led to a considerably improved draft of this work.

Appendix A: Unfolding spectra

The eigenvalue spacing distributions $P(s)$, $\mathcal{P}(s)$ and $p(s)$ (see Sect. V for definitions) considered in the level statistics data in this work characterize the fluctuations of spacings about their local means. Unfortunately, a non-constant density of states renders the actual spacings inadequate for measuring these fluctuations. Unfolding the spectrum of a matrix refers to the replacement of the actual eigenvalues η_j with a new set of numbers with unit mean spacing, but that preserve the character of local fluctuations.

We employ a simple method that essentially approximates the inverse density of states (i.e. mean level spacing) of a given matrix through linear interpolations. First, we write the eigenvalues η_j in increasing order, and express the j-th eigenvalue η_j in terms of the actual spacings S_k

$$\eta_j = \eta_1 + \sum_{k=0}^{j-1} S_k. \tag{A1}$$

No unfolding has taken place as of yet – i.e. this is an exact expression. Now we postulate that we can write the k^{th} spacing S_k as the product of a smoothly varying local mean spacing s_k and an $\mathcal{O}(1)$ fluctuating number $\rho_k = 1 + \delta_k$

$$\begin{aligned}\eta_j &= \eta_1 + \sum_{k=0}^{j-1} s_k(1+\delta_k) \\ &= \eta_1 + \sum_{k=0}^{j-1} s_k \rho_k.\end{aligned} \tag{A2}$$

Note that the ρ_k have the form of fluctuating numbers with unit mean - they will therefore serve as our unfolded spacings. By their definition we can write them as

$$\rho_k = \frac{\eta_{k+1} - \eta_k}{s_k}. \tag{A3}$$

Therefore if we calculate the smoothly varying mean level spacings s_k from the given data, we can easily find the unfolded spacings. The ambiguity in our particular unfolding procedure lies in the calculation of s_k because its definition involves choosing how many spacings over which to average, a quantity we call $2r$

$$s_k = \frac{\eta_{k+r} - \eta_{k-r}}{2r}. \tag{A4}$$

It is important to realize that s_k is just the inverse of the density of states. The parameter r is arbitrary except that it must satisfy two conditions:

1) r must be large enough to contain many eigenvalues, which is necessary in order for s_k to be a smooth function of k.

2) r cannot be too large or else s_k will actually smooth over features in the true inverse density of states.

In practice we have chosen r to be the floor function of the square root of the maximum number of eigenvalue spacings ν taken from each matrix. To avoid edge effects we have selected $\nu = 0.8N$. Then $r = \left\lfloor \sqrt{0.8N} \right\rfloor$. Here are some typical values of r used in this paper

$$\begin{aligned} N = 500, &\quad r = 22, \\ N = 1000, &\quad r = 31, \\ N = 2000, &\quad r = 44. \end{aligned} \tag{A5}$$

Such a choice of r grows with N but also is small compared with N. In other words, we satisfy the requirement $1 \ll r \ll N$ as $N \to \infty$.

For even the best choices of r, our unfolding method can still fail if the inverse density of states varies too quickly or has singularities. Such a situation arises for example in small linear combinations of basis matrices (defined in Eq. (6) and Eq. (24)) for any type. Consider, for example, the insets of Figs. 21 and 22 that show the densities of state of integrable basis matrices. The large portions of the spectra where the density of states $D(\eta)$ drops to zero is typical of small linear combinations of basis matrices. This behavior is generic for basis matrices of all types, and it can be understood by first considering the expression for the eigenvalues of a type-1 basis matrix (in the primary parametrization) where $d_k = \delta_{k,q}$

$$\eta_j = \frac{\gamma_q^2}{\lambda_j - \varepsilon_q}. \tag{A6}$$

As both the λ_j and ε_j have finite support, η_j in this case will only approach within a finite distance of zero.

An analogous argument exists for basis matrices in the ansatz parametrization for any type. For linear combinations of a small number of basis matrices, such gaps may overlap, but a mean level spacing s_k will still be ill-defined in many parts of the spectrum. As the number of basis matrices in the linear combination increases, the gaps smooth out until s_k is well-defined everywhere.

Given such gaps in spectra, no choice of r will give the consistent level statistics. This can be seen numerically by varying r and observing that $P(s)$ is strongly dependent on r. We must then avoid regions of the spectrum where $1/D(\eta)$ is poorly behaved.

The difficulty in this task is to automate it so that we can unfold many matrices in succession without having to examine each one by hand. Our solution is to notice that if there are a small number of spacings in the spectrum that are many times the local mean spacing, the standard deviation of the set of actual spacings will be large. If the standard deviation (normalized by the mean) of the actual spacings is near unity, we can be sure that there are no huge jumps such as the ones in Figs. 21 and 22.

With these considerations, here is our unfolding algorithm:

(1) Calculate $SD = \dfrac{\text{Standard Deviation}}{\text{Mean}}$ of the middle $(80\% + 2r)$ of the spectrum's actual spacings. If $SD < 1.5$, unfold this region of the spectrum with $r = \left\lfloor \sqrt{0.8N} \right\rfloor$ and continue to next matrix. If not, continue to step (2).

(2) Shift the region of the spectrum in question to the right by 10 eigenvalues.

(3) If $\eta_{\max} > \eta_{0.9N}$, use 10 fewer spacings AND restart $\eta_{\min} = \eta_{0.1N}$

(4) Calculate SD. If $SD > 1.2$, repeat back to step (2). If $SD < 1.5$ unfold this region of the spectrum with $r = \left\lfloor \sqrt{0.8N} \right\rfloor$ and continue to next matrix.

This procedure allows for fewer than $0.8N$ of the spacings to be used, but we are guaranteed to only investigate regions of the spectrum where the mean level spacing accurately represents the size of a typical spacing. Once a sufficiently large number of basis matrices are used in linear combination, the entire middle 80% of the spectrum behaves smoothly and the procedure given above terminates at step (1) for each matrix. The choice of a maximum SD of 1.2 is somewhat arbitrary, and in some parts of this work we used 1.5 in order to increase computation speed (i.e. keep more eigenvalue spacings per

matrix). Apart from slight differences in distributions, our results are unaffected by this arbitrariness.

The unfolding procedure used in this paper assumes that the level statistics are the same in all regions of the spectrum, excluding the edges. Although in principle a Hermitian matrix can have different spectral statistics in different parts of its spectrum, we numerically showed in Sect. V that the statistics are the same in all parts of the spectrum of integrable matrices $H(u)$, i.e. they are translationally invariant.

[1] In this introduction, by quantum integrable systems, models, or Hamiltonians we mean exactly solvable many-body models also colloquially called integrable. On the other hand, an 'integrable matrix' is a precise concept defined in the 4-th and 5-th paragraphs.

[2] D. Poilblanc, T. Ziman, J. Bellissard, F. Mila, G. Montambaux, Europhys. Lett. **22**, 537 (1993).

[3] D. A. Rabson, B. N. Narozhny, and A. J. Millis, Phys. Rev. B **69**, 054403 (2004).

[4] M. V. Berry and M. Tabor, Proc. R. Soc. Lond. A **356**, 375 (1977).

[5] A. Relaño, J. Dukelsky, J.M.G. Gómez, and J. Retamosa, Phys. Rev. E **70**, 026208 (2004).

[6] H.-J. Stöckmann and J. Stein, Phys. Rev. Lett. **64**, 2215 (1990).

[7] C. Ellegaard, T. Guhr, K. Lindemann, H. Q. Lorensen, J. Nygard, and M. Oxborrow, Phys. Rev. Lett. **75**, 1546 (1995).

[8] R. Püttner, B. Grémaud, D. Delande, M. Domke, M. Martins, A. S. Schlachter, G. Kaindl, Phys. Rev. Lett. **86**, 3747 (2000).

[9] J.-S. Caux, J. Mossel, J. Stat. Mech. **2011**, 02023 (2011).

[10] E. A. Yuzbashyan and B. S. Shastry, J. Stat. Phys. **150**, 704 (2013).

[11] M. L. Mehta, *Random Matrices*. (Academic Press, San Diego, 1991).

[12] T. Guhr, A. Müller-Groeling, and H.A. Weidenmüller, Phys. Rep. **299**, 189 (1998).

[13] E. A. Yuzbashyan, B. S. Shastry, and J. A. Scaramazza, Phys. Rev. E **93**, 052114 (2016).

[14] E. A. Yuzbashyan, B. L. Altshuler and B. S. Shastry, J. of Phys. A **35**, 7525 (2002).

[15] B. S. Shastry, J. Phys. A **38**, L431 (2005).

[16] H. K. Owusu, K. Wagh, and E. A. Yuzbashyan, J. of Phys. A **42**, 035206 (2009).

[17] H. K. Owusu and E. A. Yuzbashyan, J. of Phys. A **44**, 395302 (2011).

[18] B. S. Shastry, J. Phys. A **44**, 052001 (2011).

[19] One can equivalently attach the parameter to the V-part, i.e. write $H(u) = T+uV$ as in some of the previous work, or introduce two parameters, $H(x,u) = xT + uV$, see Ref. 13 for more detail.

[20] Two matrices are considered equal if they differ by a multiple of the identity $(c_1x+c_2)\mathbb{1}$, which is not included in $H^i(x)$. For example, linear independence means that $\sum_i a_i H^i(x) = (c_1x + c_2)\mathbb{1}$ iff $a_i = 0$.

[21] T. A. Brody. Lett. N. Cimento **7**, 482 (1973).

[22] M. Gaudin, Nuc. Phys. **25**, 447 (1961).

[23] F. Bornemann, Math. Comp. **79**, 871 (2009).

[24] I. L. Aleiner and K. A. Matveev, Phys. Rev. Lett. **80**, 814 (1998).

[25] E. Bogomolny, P. Leboeuf, and C. Schmit, Phys. Rev. Lett. **85**, 2486 (2000).

[26] N. Rosenzweig and C. E. Porter, Phys. Rev. **120**, 1698 (1960).

[27] M. V. Berry, Proc. R. Soc. Lond. A **400**, 229 (1985).

[28] Any ensemble composed of lists of independent random numbers is referred to as the "Poisson" ensemble.

[29] T. H. Seligman, J. J. M. Verbaarschot, and M. R. Zirnbauer, J. Phys. A **18**, 2751 (1985).

[30] Berry works with the spectral ridigity, which is a statistic closely related to the number variance.

[31] R. Prakash and A. Pandey, Phys. Rev. E **93**, 052225 (2016).

[32] T. A. Brody, J. Flores, J. B. French, P. A. Mello, A. Pandey, and S. S. M. Wong, Rev. Mod. Phys. **53**, 385 (1981).

[33] In this work, the g_j indeed are eigenvalues from an $N \times N$ GOE matrix, but are selected as a set of $N-M$ consecutive eigenvalues instead of $N-M$ randomly chosen ones. This choice does not affect the level statistics.

[34] O. Bohigas and M. J. Giannoni, Ann. Phys. **89**, 393 (1975).

[35] A. Pandey, Ann. Phys. **119**, 170 (1979).

[36] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, Ann. Phys. **321**, 1126 (2006).

[37] V. Oganesyan and D. Huse, Phys. Rev. B **75**, 155111 (2007).

[38] A. Pal and D. Huse, Phys. Rev. B **82**, 174411 (2010).

[39] P. Shukla, Phys. Rev. E **71**, 026226 (2005).

[40] V. E. Kravtsov, I. M. Khaymovich, E. Cuevas, and M. Amini, N. J. Phys. **17**, 122002 (2015).

[41] MBL seems to be a broader class. Even though certain integrals of motion have been identified in MBL systems (see Ref. 42 for example), they have not been shown to be exactly solvable or integrable as defined in this paper.

[42] M. Serbyn, Z. Papić, and D. A. Albanin, Phys. Rev. Lett. **111**, 127201 (2013).

[43] R. Modak, S. Mukerjee, E. A. Yuzbashyan, and B. S. Shastry, N. J. Phys. **18**, 033010 (2016).