Error Analysis of Free Probability Approximations to the Density of States of Disordered Systems

Jiahao Chen ((陳家豪)),* Eric Hontz,† Jeremy Moix,‡ Matthew Welborn,§ and Troy Van Voorhis‖
Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

Alberto Suárez¶
Departamento de Ingeniería Informática, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Ciudad Universitaria de Cantoblanco, Calle Francisco Tomás y Valiente, 11 E-28049 Madrid, Spain

Ramis Movassagh** and Alan Edelman††
Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA
(Received 29 February 2012; published 17 July 2012)

Theoretical studies of localization, anomalous diffusion and ergodicity breaking require solving the electronic structure of disordered systems. We use free probability to approximate the ensemble-averaged density of states without exact diagonalization. We present an error analysis that quantifies the accuracy using a generalized moment expansion, allowing us to distinguish between different approximations. We identify an approximation that is accurate to the eighth moment across all noise strengths, and contrast this with perturbation theory and isotropic entanglement theory.

DOI: 10.1103/PhysRevLett.109.036403 PACS numbers: 71.23.An

Disordered materials have long been of interest for their unique physics such as localization [1,2], anomalous diffusion [3,4] and ergodicity breaking [5]. Their properties have been exploited for applications as diverse as quantum dots [6,7], magnetic nanostructures [8], disordered metals [9,10], and bulk heterojunction photovoltaics [11–13]. However, conventional electronic structure theories require diagonalization of many explicit sampled Hamiltonians, making such calculations expensive. Alternatively, free probability theory allows a powerful nonperturbative method for computing of eigenvalues of sums of certain matrices without rediagonalizing the matrix sums [14]. This has been proposed as an approximation for general random matrices [15]; however, we are not aware of any rigorous study of its accuracy. This motivates us to describe herein a general framework for quantifying the error in terms of discrepancies in the moments of the probability distribution functions (PDFs).

Comparing two PDFs.—We propose to quantify the deviation between two PDFs using moment expansions. [16] These are widely used to describe deviations from normality in the form of Gram–Charlier and Edgeworth series [17,18]. The general case applies also to non-Gaussian reference PDFs. For two PDFs $w(\xi)$ and $\tilde{w}(\xi)$ with finite cumulants $\kappa_1, \kappa_2, \ldots$ and $\tilde{\kappa}_1, \tilde{\kappa}_2, \ldots$, and moments $\mu_1, \mu_2, \ldots$ and $\tilde{\mu}_1, \tilde{\mu}_2, \ldots$ respectively, we can define a formal differential operator which transforms $\tilde{w}$ into w [17,19]:

$$w(\xi) = \exp\left[\sum_{n=1}^{\infty} \frac{\kappa_n - \tilde{\kappa}_n}{n!}\left(-\frac{d}{d\xi}\right)^n\right]\tilde{w}(\xi). \quad (1)$$

This operator is parameterized completely by the cumulants of both distributions. The resulting Edgeworth series is asymptotic and only conditionally convergent [20].

The first k for which the cumulants κ_k and $\tilde{\kappa}_k$ differ then allows us to define a degree to which the approximation $w \approx \tilde{w}$ is valid. Expanding the exponential and using the well-known relationships between cumulants and moments allows us to state that if the first $k-1$ cumulants agree, but the kth cumulants differ, then

$$w(\xi) = \tilde{w}(\xi) + \frac{\mu_k - \tilde{\mu}_k}{k!}(-1)^k \tilde{w}^{(k)}(\xi) + O(\tilde{w}^{(k+1)}). \quad (2)$$

This series inherits the same asymptotic convergence properties as the original Edgeworth series [20,21]. Nevertheless, it is sufficient to use the leading order correction solely to quantify the error incurred by approximating one PDF by another.

The free convolution.—We now take the PDFs to be densities of states (DOSs) of random matrices. The DOS of a random matrix X is defined using the eigenvalues $\{\lambda_n^{(m)}\}$ of the M samples $X_1, \ldots, X_m, \ldots, X_M$ by

$$\rho^{(X)}(\xi) = \lim_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M}\frac{1}{N}\sum_{n=1}^{N}\delta(\xi - \lambda_n^{(m)}). \quad (3)$$

To approximate DOSs with free probability, we split the Hamiltonian

$$H = A + B \quad (4)$$

into two matrices A and B whose DOSs, $\rho^{(A)}$ and $\rho^{(B)}$ respectively, can be determined easily. In general, it is not possible to calculate the eigenvalues of H by adding the eigenvalues of A and B together; the general problem is complicated by A and B not commuting [22]. In contrast, free probability tells us that for certain noncommuting matrices A and B, the exact DOS becomes the free convolution $A \boxplus B$, i.e. $\rho^{(H)} \approx \rho^{(A\boxplus B)}$, a "sum" which can be

calculated without exact diagonalization of H [23]. We calculate the free convolution numerically by diagonalizing the free approximant [24]

$$Z = A + Q^{-1}BQ, \tag{5}$$

where Q is a $N \times N$ random matrix of Haar measure. For real symmetric matrices A and B it is sufficient to consider orthogonal matrices Q, which can be generated from the QR decomposition [25] of a Gaussian orthogonal matrix [24]. (This can be generalized readily to unitary and symplectic matrices for complex and quaternionic Hamiltonians, respectively.) The similarity transformation $Q^{-1} \cdot Q$ applies a random rotation to the basis of B with respect to A. In the $N \to \infty$ limit, the DOS $\rho^{(Z)}$ converges to the free convolution $A \boxplus B$ [14,26].

The moment expansion above provides an error analysis via discrepancies between the kth moment of the exact DOS, $\mu_k^{(H)}$, and the free approximant, $\mu_k^{(A\boxplus B)}$. By definition, the exact moments are [27]

$$\mu_k^{(H)} = \mu_k^{(A+B)} = \langle (A+B)^k \rangle, \tag{6}$$

where $\langle Z \rangle = \mathbb{E}\,\mathrm{Tr}(Z)/N$ denotes the normalized expected trace (NET) of the $N \times N$ matrix Z. Expanding the (noncommutative) binomial produces a sum of joint moments $\langle A^{n_1}B^{m_1}\cdots A^{n_r}B^{m_r}\rangle$ with the positive integer exponents n_s, m_s summing to $\sum_{s=1}^{r}(n_s + m_s) = k$. The approximation of freeness implies that the joint moments must obey, by definition [28], relations of the form

$$0 = \langle \Pi_{s=1}^{r}(A^{n_s} - \langle A^{n_s}\rangle)(B^{m_s} - \langle B^{m_s}\rangle)\rangle \tag{7a}$$

$$= \langle \Pi_{s=1}^{r} A^{n_s}B^{m_s}\rangle + \text{lower order terms}, \tag{7b}$$

where the second equality results from the linearity of the NET. Testing for $\mu_k^{(A+B)} \neq \mu_k^{(A\boxplus B)}$ then reduces to testing whether each centered joint moment of the form in (7a) is statistically nonzero. Enumerating all unique joint moments of degree k is equivalent to the combinatorics of binary necklaces, which can be generated efficiently [29].

The procedure we have described ascribes a degree k to the approximation $\rho^{(H)} \approx \rho^{(A\boxplus B)}$ given the splitting $H = A + B$. For each positive integer n, we generate all unique centered joint moments of degree n, and test if they are statistically nonzero. The lowest n for which there is at least one such term is the degree of approximation k. This is our main result.

Decomposition of the Anderson Hamiltonian.—As a concrete example, we focus on the Anderson Hamiltonian [30]

$$H = \begin{pmatrix} h_1 & J & & \\ J & h_2 & \ddots & \\ & \ddots & \ddots & J \\ & & J & h_N \end{pmatrix}, \tag{8}$$

where J is constant and the diagonal elements h_i are identically and independently distributed (iid) random variables with PDF $p_h(\xi)$. This is a real, symmetric tridiagonal matrix with circulant (periodic) boundary conditions on a one-dimensional chain. Unless otherwise stated, we assume that h_i are normally distributed with mean 0 and variance σ^2. We note that σ/J gives us a dimensionless order parameter to quantify the strength of disorder.

So far, we have only required of the decomposition scheme $H = A + B$ that $\rho^{(A)}$ and $\rho^{(B)}$ be easily computable. Are certain choices intrinsically superior to others? For the Anderson Hamiltonian, we consider two reasonable partitioning schemes:

$$H = A_1 + B_1 = \begin{pmatrix} h_1 & & & \\ & h_2 & & \\ & & h_3 & \\ & & & \ddots \end{pmatrix} + \begin{pmatrix} 0 & J & & \\ J & 0 & J & \\ & J & 0 & \ddots \\ & & \ddots & \ddots \end{pmatrix} \tag{9a}$$

$$H = A_2 + B_2 = \begin{pmatrix} h_1 & J & & & \\ J & 0 & & & \\ & & h_3 & J & \\ & & J & 0 & \\ & & & & \ddots \end{pmatrix} + \begin{pmatrix} 0 & & & & \\ & h_2 & J & & \\ & J & 0 & & \\ & & & h_4 & \cdots \\ & & & \vdots & \ddots \end{pmatrix}. \tag{9b}$$

We refer to these as scheme I and II, respectively. In scheme I, we have $\rho_{A_1} = p_h$ since A_1 is diagonal with each nonzero matrix element being iid. B_1 is simply J multiplied by the adjacency matrix of a one-dimensional chain, and therefore has eigenvalues $\lambda_n = 2J\cos(2n\pi/N)$ [31]. The DOS of B_1 is $\rho_{B_1}(\xi) = \sum_{n=1}^{N}\delta(\xi - \lambda_n)$ which converges as $N \to \infty$ to the arcsine distribution with PDF $p_{AS}(\xi) = 1/(\pi\sqrt{4J^2 - \xi^2})$ on the interval $[-2|J|, 2|J|]$. In scheme II, we have that $\rho_{A_2} = \rho_{B_2} = \rho_X$ where ρ_X is the DOS of

$$X = \begin{pmatrix} h_1 & J \\ J & 0 \end{pmatrix}.$$

The matrix X has eigenvalues $\epsilon_\pm(\xi) = h_1(\xi)/2 \pm \sqrt{h_1^2(\xi)/4 + J^2}$ and so

$$\rho_X(\xi) = \left(1 + \frac{J^2}{\xi^2}\right) p_h\left(\xi - \frac{J^2}{\xi}\right). \tag{10}$$

Numerical free convolution.—We now calculate the free convolution $A \boxplus B$ numerically by sampling the distributions of A and B and diagonalizing the free approximant (5). The exact DOS $\rho^{(A+B)}$ and free approximant $\rho^{(A\boxplus B)}$ are plotted in Figs. 1(a)–1(c) for both schemes for low, moderate and high noise regimes ($\sigma/J = 0.1$, 1, 10, respectively). For scheme I, we observe excellent agreement between $\rho^{(H)}$ and $\rho^{(A_1\boxplus B_1)}$ across all values of σ/J, which is evident from visual inspection; in contrast, scheme II shows variable quality of fit. We can understand this difference using the procedure outlined above to analyze the accuracy of the approximations $\rho^{(H)} \approx \rho^{(A_1\boxplus B_1)}$ and $\rho^{(H)} \approx \rho^{(A_2\boxplus B_2)}$. For scheme I, the approximation (2) is of degree $k = 8$; the discrepancy lies solely in the term $\langle (A_1B_1)^4 \rangle$ [32]. Free probability expects this term to vanish, but its true value is nonzero. The matrix A_1 weights each path by a factor of h, while B_1 weights each path by J and requires a hop to an adjacent site. The explicit products of matrix elements can then be expressed diagrammatically with closed paths as shown in Fig. 2. Consequently, we can write explicitly

$$\begin{aligned} \langle (A_1B_1)^4 \rangle &= \langle h_i J h_{i-1} J h_i J h_{i+1} J \rangle + \langle h_i J h_{i+1} J h_i J h_{i-1} J \rangle \\ &\quad + \langle h_i J h_{i-1} J h_i J h_{i-1} J \rangle + \langle h_i J h_{i+1} J h_i J h_{i+1} J \rangle \\ &= 2J^4 \mathbb{E}(h_i)^2 \mathbb{E}(h_i^2) + 2J^4 \mathbb{E}(h_i^2)^2 = 0 + 2J^4\sigma^4, \end{aligned} \tag{11}$$

where the second equality follows from the independence of the h_i's. This explains why the agreement between the free and exact PDFs is so good, as the leading order correction is in the eighth derivative of $\rho^{(A_1\boxplus B_1)}$ with coefficient $2\sigma^4J^4/8! = (\sigma J)^4/20160$. In contrast, scheme II is correct only to degree $k = 4$, where the discrepancy lies in $\langle A_2^2 B_2^2 \rangle$. Free probability expects this to be equal to $\langle A_2^2 B_2^2 \rangle = \langle A_2^2 \rangle \langle B_2^2 \rangle = \langle X^2 \rangle^2 = (J^2 + \sigma^2/2)^2$, whereas the exact value of this term is $J^2(J^2 + \sigma^2)$. Therefore, the error is in the fourth derivative of $\rho^{(A\boxplus B)}$ with coefficient $(-\sigma^4/4)/4! = -\sigma^4/96$.

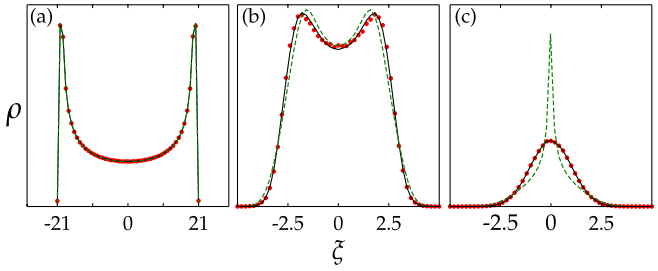

FIG. 1 (color online). Calculation of the DOS, $\rho(\xi)$, of the Hamiltonian H of (8) with $M = 5000$ samples of 2000×2000 matrices for (a) low, (b) moderate and (c) high noise ($\sigma/J = 0.1$, 1 and 10, respectively, with $\sigma = 1$). For each figure we show the results of free convolution defined in scheme I ($\rho^{(A_1\boxplus B_1)}$; black solid line), scheme II ($\rho^{(A_2\boxplus B_2)}$; green dashed line) and exact diagonalization ($\rho^{(H)}$; red dotted line).

Analytic free convolution.—Free probability allows us also to calculate the limiting distribution of $\rho^{(A\boxplus B)}$ in the macroscopic limit $N \to \infty$ and $M \to \infty$, allowing the cost of numerical sampling and matrix diagonalization to be sidestepped entirely. The key tool is the R-transform $R(w) = g^{-1}(w) - w^{-1}$ [23], where g^{-1} is defined implicitly via the Cauchy transform (i.e., its retarded Green function)

$$w = \lim_{\epsilon \downarrow 0} \int_{\mathbb{R}} \frac{\rho(\xi)}{g^{-1}(w) - (\xi + i\epsilon)} d\xi. \tag{12}$$

For freely independent A and B, the R transforms linearize the free convolution, i.e. $R^{(A\boxplus B)}(w) = R^{(A)}(w) + R^{(B)}(w)$, and the PDF can be recovered from the Plemelj–Sokhotsky inversion formula by

$$\rho^{(A\boxplus B)}(\xi) = \frac{1}{\pi} \operatorname{Im}[(g^{(A\boxplus B)})^{-1}(\xi)] \tag{13a}$$

$$g^{(A\boxplus B)}(w) = R^{(A\boxplus B)}(w) + w^{-1}. \tag{13b}$$

We apply this to scheme I with each iid h_i following a Wigner semicircle distribution with PDF $p_W(\xi) = \sqrt{4 - \xi^2}/4\pi$ on the interval $[-2, 2]$. (The analytic calculation is considerably easier than for Gaussian noise.) First, calculate the Green function $G^{(A_1)}(z) = (z - \sqrt{z^2 - 4})/2$. Next, take the functional inverse $g^{(A_1)}(w) = (G^{(A_1)})^{-1}(w) = w + 1/w$. Subtracting $1/w$ finally yields the R-transform $R^{(A)}(w) = w$. Similarly with $\rho^{(B_1)} = p_{AS}$, we find its Cauchy transform $G^{(B_1)}(z) = 1/\sqrt{z^2 - 4J^2}$, its functional inverse $g^{(B_1)}(w) = (\sqrt{1 + 4J^2w^2})/w$, and the R transform $R^{(B_1)}(w) = (-1 + \sqrt{1 + 4J^2w^2})/w$.

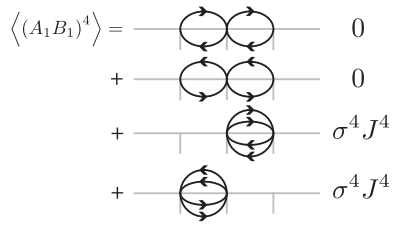

FIG. 2. Diagrammatic expansion of the term $\langle A_1B_1A_1B_1A_1B_1A_1B_1 \rangle$ in terms of allowed paths dictated by the matrix elements of A_1 and B_1 of scheme I in (9a).

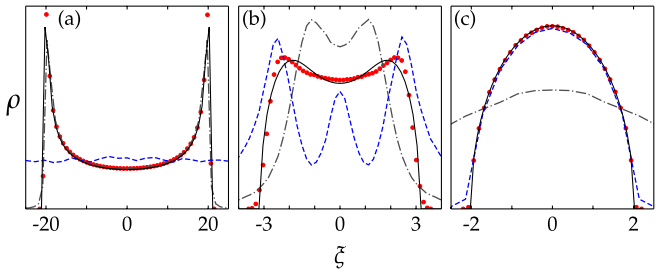

FIG. 3 (color online). DOS, $\rho(\xi)$, of the Hamiltonian (8) with $M = 5000$ samples of 2000×2000 matrices with (a) low, (b) moderate and (c) high semicircular on-site noise ($\sigma/J = 0.1$, 1 and 10, respectively, with $\sigma = 1$), as calculated with exact diagonalization (red dotted line), free convolution (black solid line), and perturbation theory with A_1 as reference (blue dashed line) and B_1 as reference (gray dash-dotted line). The partitioning scheme is scheme I of (9a).

To perform the free convolution analytically, we add the R-transforms to get $R^{(A_1 \boxplus B_1)}(w) = R^{(A_1)}(w) + R^{(B_1)}(w)$, from which we obtain $g^{(A_1 \boxplus B_1)}(w) = w + (\sqrt{1 + 4J^2 w^2})/w$. The final steps are to calculate the functional inverse $(g^{(A_1 \boxplus B_1)})^{-1}$ and take its imaginary part to obtain $\rho^{(A_1 \boxplus B_1)}$. Unfortunately, $(g^{(A_1 \boxplus B_1)})^{-1}$ cannot be written in a compact closed form; nevertheless, the inversion can be calculated numerically. We present calculations of the DOS as a function of noise strength σ/J in Fig. 3, showing again that the free convolution is an excellent approximation to the exact DOS.

Comparison with other approximations.—We compare the free approximations to the results of standard second-order matrix perturbation theory [33], as shown in Fig. 3. Unsurprisingly, perturbation theory produces results that vary strongly with σ/J, and that the different series, based on whether A is considered a perturbation of B or vice versa, have different regimes of applicability. Furthermore, it is clear even from visual inspection that the second moment of the DOS calculated using second-order perturbation theory is not always correct. In contrast, the free convolution produces results with a more uniform level of accuracy across the entire range of σ/J, and that we have at least the first three moments being correct [34].

It is also natural to ask what mean field theory, another standard tool, would predict. Interestingly, the limiting behavior of scheme I as $N \to \infty$ is equivalent to both the coherent potential approximation (CPA) [35–37] in condensed matter physics, and the Blue's function formalism in quantum chromodynamics for calculating one-particle irreducible self-energies [38]. The breakdown in the CPA in the term $\langle (A_1 B_1)^4 \rangle$ is known [1,39]; however, to our knowledge, the magnitude of the deviation was not explained. Our error analysis framework provides such a quantitative explanation.

Finally, we discuss the predictions of isotropic entanglement (IE) theory, which linearly interpolates the fourth cumulant between the classical convolution $\rho^{(A*B)}(\xi) = \int_{-\infty}^{\infty} \rho^{(A)}(\xi)\rho^{(B)}(x - \xi)dx$ and the free convolution $\rho^{(A \boxplus B)}(\xi)$ [34,40]. Given the eigenvalues Λ_A, Λ_B of the matrices A and B, the classical convolution $\rho^{(A*B)}(\xi)$ can be computed from the eigenvalues of the random matrix $Z_{\mathrm{cl}} = \Lambda_A + \Pi^{-1}\Lambda_B \Pi$, where Π is a $N \times N$ random permutation matrix. This compares with the free convolution sampled from $Z' = \Lambda_A + Q^{-1}\Lambda_B Q$, which has the same eigenvalues as the free approximant (5) by orthogonal invariance of the Haar measure of Q. As discussed previously, the lowest three moments of Z and H are identical; this turns out to be true also for Z_{cl} [34]. Therefore, IE proposes to interpolate via the fourth cumulant, with interpolation parameter p defined as

$$p = \frac{\kappa_4^{(H)} - \kappa_4^{(A \boxplus B)}}{\kappa_4^{(A*B)} - \kappa_4^{(A \boxplus B)}} \tag{14}$$

For scheme I, IE always favors the free convolution limit ($p = 0$) over the classical limit ($p = 1$); this follows from our previous analysis that $\kappa_4^{(H)} = \kappa_4^{(A_1 \boxplus B_1)}$. In scheme II, however, we observe the unexpected result that p is always negative regardless of the noise strength σ/J. From our previous analysis, $\kappa_4^{(A_2 + B_2)} - \kappa_4^{(A_2 \boxplus B_2)} = -\sigma^4/4$. Additionally, $\kappa_4^{(A_2 * B_2)} \neq \kappa_4^{(A_2 \boxplus B_2)}$ where the only discrepancy lies is in the so-called departing term $\langle A_2 B_2 A_2 B_2 \rangle$ [34,40]. This term contributes 0 to $\kappa_4^{(A_2 \boxplus B_2)}$ but has value $\langle A_2^2 \rangle \langle B_2^2 \rangle = (J^2 + \sigma^2/2)^2$ in $\kappa_4^{(A_2 * B_2)}$, since for the classical convolution, $\langle \Pi_{s=1}^{r} (A_2^{n_s} B_2^{m_s}) \rangle = \langle A_2^{\sum_{s=1}^{r} n_s} \rangle \langle B_2^{\sum_{s=1}^{r} m_s} \rangle$. Thus, $p = -2(2(\frac{\sigma}{J})^{-2} + 1)^{-2}$ which is manifestly negative.

In conclusion, the accuracy of approximations using the free convolution depend crucially on the way the Hamiltonian is partitioned. Scheme I describes an unexpectedly accurate approximation for the DOS of disordered Hamiltonians for all system sizes N and noise strengths σ/J. Our error analysis explains why this approximation is correct to degree 8, and also provides a general framework for understanding the performance of other approximations. We expect our results to be generally applicable to arbitrary Hamiltonians, and pave the way toward constructing even more accurate approximations using free probability with rigorous error bars. Our results represent an optimistic beginning to the use of powerful and highly accurate nonperturbative methods for studying the electronic properties of disordered condensed matter systems regardless of the strength of noise present. Thus, we expect these methods to be especially useful for studying the unique physics enabled by noise.

J. C., E. H., M. W., T. V., R. M., and A. E. acknowledge funding from NSF SOLAR Grant No. 1035400. J. M. acknowledges support from NSF CHE Grant No. 1112825 and DARPA Grant No. N99001-10-1-4063. A. E. acknowledges additional funding from NSF DMS Grant No. 1016125. A. S. acknowledges funding from

Spain's Dirección General de Investigación, Project TIN2010-21575-C02-02. We thank Jonathan Novak, Sebastiaan Vlaming, and Raj Rao for insightful discussions.

*jiahao@mit.edu
†ehontz@mit.edu
‡jmoix@mit.edu
§welborn@mit.edu
‖tvan@mit.edu
¶alberto.suarez@uam.es
**ramis@mit.edu
††edelman@math.mit.edu

[1] D. J. Thouless, Phys. Rep. **13**, 93 (1974).
[2] F. Evers and A. Mirlin, Rev. Mod. Phys. **80**, 1355 (2008).
[3] J.-P. Bouchaud and A. Georges, Phys. Rep. **195**, 127 (1990).
[4] M. F. Shlesinger, G. M. Zaslavsky, and J. Klafter, Nature (London) **363**, 31 (1993).
[5] R. G. Palmer, Adv. Phys. **31**, 669 (1982).
[6] E. Barkai, Y. Jung, and R. Silbey, Annu. Rev. Phys. Chem. **55**, 457 (2004).
[7] F. D. Stefani, J. P. Hoogenboom, and E. Barkai, Phys. Today **62**, No. 2, 34 (2009).
[8] A. Hernando, J. Phys. Condens. Matter **11**, 9455 (1999).
[9] J. Dyre and T. Schrøder, Rev. Mod. Phys. **72**, 873 (2000).
[10] J. S. Dugdale, *The Electrical Properties of Disordered Metals*, Cambridge Solid State Science Series (Cambridge University Press, Cambridge, England, 2005).
[11] J. Peet, A. J. Heeger, and G. C. Bazan, Acc. Chem. Res. **42**, 1700 (2009).
[12] S. Difley, L.-P. Wang, S. Yeganeh, S. R. Yost, and T. Van Voorhis, Acc. Chem. Res. **43**, 995 (2010).
[13] S. R. Yost, L.-P. Wang, and T. Van Voorhis, J. Phys. Chem. C **115**, 14 431 (2011).
[14] D. Voiculescu, Inventiones Mathematicae **104**, 201 (1991).
[15] P. Biane, in *Quantum Probability Communications* (World Scientific, Singapore, 1998), Vol. 11, Chap. 3, p. 55.
[16] J. Chen and A. Edelman, arXiv:1204.2257.
[17] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics* (Edward Arnold, London, 1994).
[18] S. Blinnikov and R. Moessner, Astron. Astrophys. Suppl. Ser. **130**, 193 (1998).
[19] D. Wallace, Ann. Math. Stat. **29**, 635 (1958).
[20] H. Cramér, *Random Variables and Probability Distributions*, Cambridge Tracts in Mathematics and Mathematical Physics (Cambridge University Press, Cambridge, England, 1970), 3rd ed., p. 118.
[21] P. Hall, *The Bootstrap and Edgeworth Expansion*, Springer Series in Statistics (Springer-Verlag, New York, 1992).
[22] A. Knutson and T. Tao, Not. Amer. Math. Soc. **48**, 175 (2001), http://www.ams.org/notices/200102/.
[23] D. Voiculescu, in *Operator Algebras and their Connections with Topology and Ergodic Theory*, Lecture Notes in Mathematics Vol. 1132, edited by H. Araki, C. Moore, S.-V. Stratila, and D.-V. Voiculescu (Springer, New York, 1985), p. 556.
[24] P. Diaconis, Not. Amer. Math. Soc. **52**, 1348 (2005), http://www.ams.org/notices/200511/.
[25] G. H. Golub and C. F. van Loan, *Matrix Computations* (Johns Hopkins, Baltimore, MD, 1996), 3rd ed.
[26] D. Voiculescu, in *Proceedings of the International Congress of Mathematicians* (Birkhäuser Verlag, Zürich, Switzerland, 1994), p. 227.
[27] M. L. Mehta, *Random Matrices*, Pure and Applied Mathematics (Elsevier/Academic Press, Amsterdam; San Diego, CA, 2004), 3rd ed.
[28] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability*, London Math. Soc. Lecture Note Ser., Vol. 335 (Cambridge University Press, Cambridge, England, 2006).
[29] J. Sawada, SIAM J. Comput. **31**, 259 (2001).
[30] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).
[31] G. Strang, SIAM Rev. **41**, 135 (1999).
[32] I. Popescu, personal communication (2011).
[33] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge, UK, 1990).
[34] R. Movassagh and A. Edelman, arXiv:1012.5039.
[35] P. Neu and R. Speicher, Z. Phys. B **95**, 101 (1994).
[36] P. Neu and R. Speicher, J. Phys. A **28**, L79 (1995).
[37] P. Neu and R. Speicher, J. Stat. Phys. **80**, 1279 (1995).
[38] A. Zee, Nucl. Phys. **B474**, 726 (1996).
[39] J. Blackman, D. Esterling, and N. Berk, Phys. Rev. B **4**, 2412 (1971).
[40] R. Movassagh and A. Edelman, Phys. Rev. Lett. **107**, 097205 (2011).