Critical behavior of the Anderson model on the Bethe lattice via a large-deviation approach

Giulio Biroli,[1] Alexander K. Hartmann,[2] and Marco Tarzia[3,4]
[1]*Laboratoire de Physique Statistique, Ecole Normale Supérieure, PSL Research University, 24 rue Lhomond, 75005 Paris, France*
[2]*Institut für Physik, Universität Oldenburg, 26111 Oldenburg, Germany*
[3]*LPTMC, CNRS-UMR 7600, Sorbonne Université, 4 Pl. Jussieu, F-75005 Paris, France*
[4]*Institut Universitaire de France, 1 rue Descartes, 75005 Paris, France*

(Received 15 October 2021; revised 5 February 2022; accepted 24 February 2022; published 11 March 2022)

We present a new, large-deviation approach to investigate the critical properties of the Anderson model on the Bethe lattice close to the localization transition in the thermodynamic limit. Our method allows us to study accurately the distribution of the local density of states (LDoS) down to very small probability tails as small as 10^{-50}, which are completely out of reach for standard numerical techniques. We perform a thorough analysis of the functional form of the tails of the probability distributions of the LDoS, which turn out to be very well described by the functional form predicted by the supersymmetric formalism and yields a direct and transparent estimation of the correlation volume very close to the Anderson transition. Such a correlation volume is found to diverge exponentially when the localization is approached from the delocalized regime, in a singular way that is compatible with the analytic predictions of the supersymmetric treatment.

DOI: 10.1103/PhysRevB.105.094202

I. INTRODUCTION

After more than a half century, the subject of Anderson localization [1] is still very much alive [2–4], as proved by the recent observations of Anderson localization of of cold atomic gases in 1D [5–7] and 3D [8–10], of kicked rotors in 3D [11], and of classical sound elastic waves in 3D [12]. On the theoretical side, several questions remain open: Although there is by now a good understanding of the localization transition in low-dimensional systems, culminating in a functional renormalization group analysis by a $2+\epsilon$ expansion [13], the behavior in high dimensions [14], in particular, the existence of an upper critical dimension and the relationship with Bethe lattice analysis [15], is still an issue. Recently, there has been a renewal of interest on this problem because of its relationship with many-body localization (MBL) [16]. This is a fascinating new kind of phase transition between a low-temperature nonergodic phase—a purely quantum glass—and a high-temperature ergodic phase [17–22]. This phenomenon has been argued to take place for several disordered isolated interacting quantum systems, and can be thought of as localization in the Fock space of Slater determinants, which play the role of lattice sites in a disordered Anderson tight-binding model. A paradigmatic representation of this transition [16,23–28] is indeed (single-particle) Anderson localization on a very high-dimensional hierarchical lattice, which for spinless electrons consists in an N-dimensional hypercube (where $N \gg 1$ is the number of sites of the lattice system). Although the analogy between MBL and Anderson localization on the Bethe lattice involves several drastic simplifications (e.g., the correlation between random energies is neglected as well as the specific structure of the Hilbert space), it is very useful to obtain a qualitative understanding of the problem [28–31].

Localization had an impact on several fields, in particular random matrices and quantum chaos. As a matter of fact, in the delocalized phase the level statistics is described by random matrix theory and generally corresponds to the Gaussian orthogonal ensemble (GOE), whereas instead in the localized phase is determined by Poissonian statistics because wave functions close in energy are exponentially localized on very distant sites and hence do not overlap; thus, contrary to the GOE case, there is no level repulsion and eigenenergies are distributed similarly to random points thrown on a line.

The relationship with quantum chaos goes back to the Bohigas-Giannoni-Schmidt conjecture, which states that the level statistics of chaotic (or ergodic) systems is given by random matrix theory, whereas integrable systems instead are characterized by Poissonian statistics [32]. This result can be fully worked out and understood in the semi-classical limit [33,34]: for a quantum chaotic system, in the $\hbar \to 0$ limit, wave functions at a given energy become uniformly spread over the microcanonical hypersurface of the configuration space. They are fully delocalized as expected for an ergodic classical system that covers regions with the same energy uniformly. Instead, quantum nonergodic models, such as integrable systems, are characterized by Poissonian statistics and localized wave functions. All those results support a general relationship between delocalization–GOE statistics–ergodicity (similarly between localization–Poisson statistics–lack of ergodicity).

However, in the last decade several numerical studies [35–42] were performed for the Anderson model on the *Bethe lattice*, in fact, on *random-regular Graphs* (RRG), with N nodes and a parameter W controlling the strength of the local disorder. This is a class of random lattices that have locally a

tree-like structure but do not have boundaries [43], see below for a precise definition. The results suggested the possibility of the existence of an intermediate delocalized but nonergodic phase characterized by multifractal eigenfunctions in a broad disorder range preceding the localization transition, as first suggested in Ref. [24]. The arguments in favor of this scenario rely mostly on numerical extrapolations of results obtained from exact diagonalization (ED) of large but finite samples, and the existence of such a phase in the thermodynamic limit has been strongly questioned during recent years [28,44–49].

Although the possibility of such multifractal delocalized phase is clearly very intriguing, especially due to its relationship with MBL [24], it appears to be in explicit conflict with the analytical predictions based on the supersymmetric approach for the Anderson model on sparse random graphs [50–55]. Moreover, recent numerical investigations based on the finite-size scaling of the spectral and the wave functions's statistics on the delocalized side of the Anderson model on RRG [44,47,48] and similar sparse random lattices [45,46] provided strong indications against the existence of a truly intermediate nonergodic extended phase. These investigations highlighted a nonmonotonous behavior of the observables as a function of the system size on the delocalized side of the transition, which can be explained in terms of (i) the presence of a characteristic scale which diverges exponentially fast approaching the transition and is already very large far from it [44–48] and (ii) the localized nature of the critical point in the limit of an infinite dimension [14,51,53–55]. The combination of these two elements produce *dramatic and highly nontrivial finite-size effects* even very far from the critical point and give rise to a strong nonergodic behavior in a crossover region where the *correlation volume* $N_c(W)$ is larger than the accessible system sizes. (On the contrary, there is by now a general consensus on the fact that the delocalized phase of the Anderson model on the loop-less Cayley is genuinely multifractal [30,56,57]).

Note that the thorough characterization of such crossover regime has not only an academic interest, but has also some important practical implications. In fact, the crossover scale turns out to be so large even far below the localizaiton transition that the multifractal exponents associated to the spectral statistics appear to be independent on the system size N in a broad range of sizes smaller than N_c, producing an effective nonergodic behavior on several decades of length and timescales [29,30,39,40]. Yet, a precise characterization of the correlation volume, in particular from the numerical point of view, remains elusive. Direct numerical simulations would need to focus on intractably large system sizes. The Anderson transition on tree-like lattices offers, however, an alternative route since it allows for an exact solution [15,47–55,58–62]. This can be obtained in terms of the exact self-consistent equations for the Green's functions (in the thermodynamic limit), which allow to establish the transition point and the corresponding critical behavior. However, even this approach suffers from the dramatic increase of the correlation volume, which controls the cutoff of the probability distribution of the imaginary part of the Green's function [i.e., the local density of states (LDoS)] [47,50–52,61]. Since $N_c(W)$ is so large even far away from the transition, the cutoff occurs in the far tails of the distribution, which cannot be properly sampled with standard numerical techniques such as the population dynamics algorithm even using huge populations [61]. Here, we solve this problem by putting forward a novel large-deviation technique which allows one to sample very accurately the tails of the probability distribution of the LDoS down to extremely small probabilities, and highlight with great accuracy the crossover scale and its critical behavior.

The outcomes of this analysis are in agreement with the predictions of the supersymmetric approach [47,50–53] and are compatible with a correlation volume which diverges exponentially fast as the Anderson localization is approached, as $N_c(W) \approx A\, e^{c/(W_L - W)^{\nu}}$, with $\nu = 1/2$ and W_L being the critical disorder strength.

The paper is organized as follows. In the next section we introduce the model and briefly review previous results and studies. In Sec. III we present some recent numerical results of the spectral statistics obtained from ED of the Anderson model on the RRG of finite size. In Sec. IV we describe the new large deviation approach to sample efficiently the tails of the distributions of the Green's functions and directly estimate the correlation volume close to W_L. Finally, in Sec. VI we discuss the physical implications of our results, providing some concluding remarks and perspectives for future work.

II. MODEL AND STATE OF THE ART

The model we focus on consists in noninteracting spinless electrons in a disordered potential

$$\mathcal{H} = -t \sum_{\langle i,j \rangle} (c_i^\dagger c_j + c_j^\dagger c_i) - \sum_{i=1}^{N} \epsilon_i\, c_i^\dagger c_i, \tag{1}$$

where the first sum runs over all the nearest-neighbors sites of the lattice, the second sum runs over all N sites; $c_i^\dagger$, c_i are fermionic creation and annihilation operators, and t is the hopping kinetic energy scale, which we take equal to 1. The on-site energies ϵ_i are i.i.d. random variables uniformly distributed in the interval $[-W/2, W/2]$:

$$p(\epsilon) = U\left(-\frac{W}{2}, \frac{W}{2}\right) \equiv \frac{1}{W}\,\theta\left(\frac{W}{2} - |\epsilon|\right). \tag{2}$$

As anticipated in the Introduction, the lattice that we consider is a $(k+1)$-RRG, i.e., a lattice chosen uniformly at random among all graphs of N sites where each of the sites has connectivity $k+1$. The properties of such random graphs were extensively studied (see Ref. [43] for a review). A RRG can be essentially viewed as a finite portion of a tree wrapped onto itself. It is known in particular that for large number of sites any finite portion of such a graph is a tree with a probability going to 1 as $N \to \infty$, and that the RRG has large loops of typical length of order $\ln N$ [43].

The model (1) is then a sum of two random matrices, $\mathcal{H} = \mathcal{C} + \mathcal{D}$: $\mathcal{C}$ is the connectivity matrix of the RRG, $\mathcal{C}_{ij} = -t$ if sites i and j are connected and zero otherwise. $\mathcal{D}$ is the diagonal matrix corresponding to the on-site random energies $\mathcal{D}_{ij} = \epsilon_i \delta_{ij}$. It is known from previous studies that the first ensemble of sparse random matrices belongs to the GOE universality class (with fully delocalized eigenvectors) [63,64], while the second is described by definition by Poissonian statistics (with fully localized eigenvectors).

Localization on the RRG was first studied by Abou-Chacra, Anderson, and Thouless [15] and then later by many others, see Refs. [35–42,44,46–55,58–62] and references therein. Many similarities, but also few important differences, with the three dimensional behavior have been found. The differences mainly concern the critical properties. Contrary to the finite-dimensional case, the critical behavior is not power-law-like but instead exponential, i.e., one finds essential singularities approaching the localization transition from the delocalized regime [47,50–55]. Moreover, the inverse participation ratio (IPR), defined as $\langle \sum_{i=1}^{N} |\psi_\alpha(i)|^4 \rangle$, is found to have a discontinuous jump at the transition from a $O(1)$ toward a $1/N$ scaling [50], instead of being continuous at the transition. Arguments based on supersymmetric field theory indicate that the level statistics should display a transition from GOE to Poissonian statistics concomitant with the localization transition [50,51] (see also Ref. [49]). However, the first numerical studies did not fully support this claim [35,65]. Moreover, the arguments of the authors of Ref. [24] indicated that the two transitions might actually not coincide. As discussed above, the possibility of the existence of an intermediate phase, which is delocalized and yet still not ergodic, were first suggested in Ref. [35]. These findings triggered a lot of activity. In Ref. [36], based on the numerical extrapolation of the spectrum of fractal dimensions of finite-size systems, it was conjectured that the eigenstates were multifractal in the entire delocalized phase. More recently, the authors of Refs. [37,38,42] combined EDs and semi-analytical calculations to claim the existence of the intermediate nonergodic but delocalized phase in a broad disorder strength $W_E < W < W_L$. These claims have been questioned by the numerical investigations of Refs. [44–49], which analyzed the level and eigenfunction statistics on the delocalized side of the Anderson transition on the RRG and similar sparse random lattices, and unveiled the existence of very strong finite-size effects with a characteristic crossover scale $N_c(W)$ associated to a pronounced nonmonotonous behavior of the observables as a function of N, and which diverges exponentially fast as the localization transition is approached. The origin of the nonmonotonicity has been traced back to the localized nature of the Anderson critical point in the limit of infinite dimensions [14,51,53–55]: For $N \ll N_c$ the system flows towards the Anderson transition fixed point whose properties on the RRG are analogous to the localized phase, whereas for $N \gg N_c$ the system approaches the $N \to \infty$ ergodic behavior. The conclusions of these investigations are thus that the system is ergodic in the entire delocalized phase, but is characterized by dramatic and nontrivial finite-size effects even very far from the critical point, giving rise to an apparent nonergodic behavior in a crossover region where the correlation volume is larger than the accessible system sizes. Nonetheless, as explained in the Introduction, a precise characterization of the correlation volume N_c is still missing.

In the following, without loss of generality, we focus on the $k = 2$ case (i.e., total connectivity $k + 1 = 3$) and on the middle of the spectrum $E = 0$. Previous studies of the transmission properties and dissipation propagation attempted to determine the critical value of the disorder at which the localization transition takes place [15,57,58,61,62]. The most accurate and precise estimation of W_L in the thermodynamic limit has been performed in Ref. [61], based on a direct high-precision numerical diagonalization of the integral operator governing the linear stability of the recursive cavity equations for the Green's functions, yielding $W_L \simeq 18.17 \pm 0.01$. Previous analysis of the spectral properties instead suggested the presence of the nonergodic delocalized phase in the range $10 \approx W_E < W < W_L$ [35,37,38].

III. EXACT DIAGONALIZATION ON THE RRG

The purpose of this section is to set the stage and show the known results from the recent literature that support the presence of the correlation volume $N_c(W)$ and its very fast increase [44–48]. These results are obtained from EDs of the Hamiltonian (1) on the RRG for several system sizes $N = 2^n$, from $n = 6$ to $n = 15$, and for several values of the disorder strength W on the delocalized side of the Anderson transition in the disorder range where previous studies suggested the possibility of the existence of a multifractal delocalized phase [35,37,38,42] $W_E < W < W_L$. For each value of N and W, we average over both the on-site quenched disorder and on RRG realizations, taking (at least) 2^{22-n} different samples. Since we are interested in $E = 0$, we also average over $1/8$ of the eigenstates centered around the middle of the band (we checked that taking $1/16$ or $1/32$ of the states does not alter the results, but yields a poorer statistics).

We start by focusing on numerical results for the level statistics of the Anderson model on the RRG which unveil the nonmonotonic behavior of the relevant observables [44,45,48]. We study the statistics of level spacing of neighboring eigenvalues: $s_\alpha = E_{\alpha+1} - E_\alpha \geqslant 0$, where E_α is the energy of the αth eigenstate in the sample. In the delocalized regime level crossings are forbidden. Hence the eigenvalues are strongly correlated and the level statistics is expected to be described by random matrix theory (more precisely, several results support a general relationship between delocalization and the Wigner's surmise of the GOE). Conversely, in the localized phase wave functions close in energy are exponentially localized on very distant sites and do not overlap. Thus there is no level repulsion and eigenvalues should be distributed similarly to random points thrown on a line (Poissonian statistics). To avoid the difficulties related to the unfolding of the spectrum, we follow the authors of Ref. [66] and measure the ratio of adjacent gaps

$$r_\alpha = \frac{\min\{s_\alpha, s_{\alpha+1}\}}{\max\{s_\alpha, s_{\alpha+1}\}},$$

and obtain the probability distribution which displays a universal form depending on the level statistics [66]. In particular, $\langle r \rangle$ is expected to converge to its GOE and Poissonian counterpart in the extended and localized regime [67], allowing to discriminate between the two phases as $\langle r \rangle$ changes from $\langle r \rangle_{\rm GOE} \simeq 0.53$ to $\langle r \rangle_P \simeq 0.39$, respectively.

The GOE-Poisson transition can also be captured by correlations between nearby eigenstates such as the mutual overlap between two subsequent eigenvectors, defined as

$$q_m = \sum_{i=1}^{N} |\psi_\alpha(i)||\psi_{\alpha+1}(i)|.$$

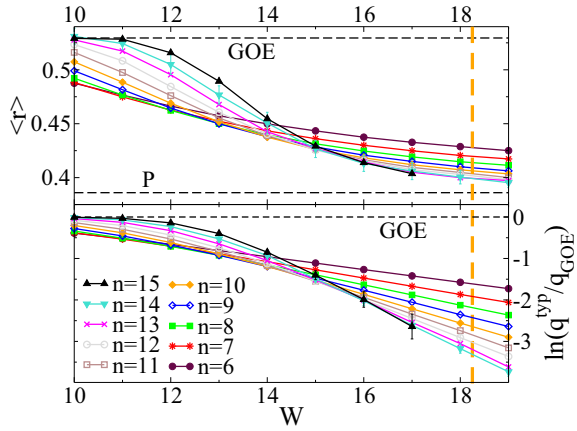

FIG. 1. $\langle r\rangle$ (upper panel) and $\ln(q^{\rm typ}/q_{\rm GOE})$ (lower panel) as a function of the disorder W for several system sizes $N = 2^n$ with n from 6 to 15. The horizontal dashed lines correspond to the reference GOE and Poissonian asymptotic values. The vertical orange dashed line spots the position of the Anderson localization transition, $W_L \simeq 18.17$ [61].

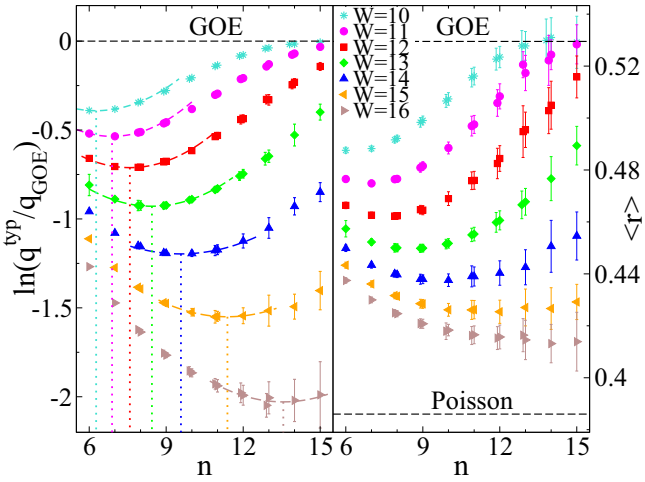

FIG. 2. $\ln(q^{\rm typ}/q_{\rm GOE})$ (left panel) and $\langle r\rangle$ (right panel) as a function of $n = \log_2 N$ for $W = 10, 11, \ldots, 16$. The data show the nonmonotonic behavior of $q^{\rm typ}$ and $\langle r\rangle$. The position of the minimum $n_c(W)$ extracted from $q^{\rm typ}(W)$, corresponding to the volume $N_c(W) = 2^{n_c(W)}$, is represented by the vertical dotted lines.

In the GOE regime the wave functions's amplitudes are i.i.d. Gaussian random variables of zero mean and variance $1/N$ [68], hence $\langle q\rangle$ converges to $\langle q\rangle_{\rm GOE} = 2/\pi$. Conversely, in the localized phase two successive eigenvector are typically peaked around very distant sites and do not overlap, and therefore $\langle q\rangle_P \to 0$ for $N \to \infty$. At first sight this quantity seems to be related to the statistics of wave functions's coefficients rather than to energy gaps. Nonetheless, in all the random matrix models that have been considered in the literature so far, one empirically finds that $\langle q\rangle$ is directly associated to the statistics of gaps between neighboring energy levels [69].

In Fig. 1 we show the behavior of the average value of the ratio of adjacent gaps $\langle r\rangle$ and of (the logarithm of) the typical value of the mutual overlap between subsequent eigenvectors $q^{\rm typ} = e^{\langle \ln q\rangle}$ as a function of the disorder W, for several system sizes $N = 2^n$, with n from 6 to 15. As expected, for small (or large) enough disorder we recover the universal values $\langle r\rangle_{\rm GOE} \simeq 0.53$ and $q^{\rm typ}_{\rm GOE} = 2/\pi$ (or $\langle r\rangle_P \simeq 0.39$ and $q^{\rm typ}_P \to 0$) corresponding to GOE (or Poissonian) statistics. However, as pointed out in Ref. [35] the different curves corresponding to different values of N cross much before the localization transition, occurring at $W_L \simeq 18.17$ (indicated by the vertical dashed line in the plot). This behavior was interpreted in terms of an intermediate delocalized but nonergodic phase [35]. Nevertheless, by carefully analyzing the data, we realized that the crossing point is, in fact, slowly but systematically drifting towards larger values of W as N is increased (see inset of Fig. 3), as also observed [44,45,48].

This is clearly unveiled by Fig. 2, where we plot the behavior of $q^{\rm typ}$ and $\langle r\rangle$ as a function of $n = \log_2 N$, for several values of the disorder belonging to the range where the curves of $\langle r\rangle$ and $q^{\rm typ}$ for different n cross, i.e., $10 \lesssim W \lesssim 16$. One indeed observes that, in this region, $q^{\rm typ}$ and $\langle r\rangle$ become nonmonotonic functions of n. The position of the minimum of $q^{\rm typ}$ (highlighted by dashed vertical lines in the left panel of Fig. 2) naturally defines a characteristic system size $N_c(W) = 2^{n_c(W)}$ governing the crossover from Poissonian to GOE statistics (on the scale of the mean level spacing): For $N < N_c(W)$ one has indeed that $q^{\rm typ}$ decreases as the system size is increased, as expected for localized wave functions, whereas for $N > N_c(W)$ it is an increasing function of n and eventually converges to the GOE universal value. The same nonmonotonic behavior as a function of the system size is found for $\langle r\rangle$ (right panel of Fig. 2).

The finite-size scaling behavior of the statistics of the wave functions's amplitudes can be analyzed in a similar way, as was recently shown in Refs. [44,45,48]. This can be done by computing the moments

$$\Upsilon_q(N) = \left\langle \log\left(\sum_{i=1}^{N} |\psi_\alpha(i)|^{2q}\right)\right\rangle,$$
$$\Upsilon_1(N) = -\left\langle \sum_{i=1}^{N} |\psi_\alpha(i)|^2 \log|\psi_\alpha(i)|^2\right\rangle, \qquad (3)$$

from which the flowing fractal dimensions D_q are obtained as logarithmic derivatives with respect to $\log N$ [70]:

$$D_q(N) = \frac{1}{1-q}\frac{\partial \Upsilon_q(N)}{\partial \log N}, \quad D_1(N) = \frac{\partial \Upsilon_1(N)}{\partial \log N}. \qquad (4)$$

As discussed at length in Refs. [44,45,48], the anomalous dimensions D_1 and D_2 exhibit a behavior very much analogous to that observed for the level statistics in Fig. 1. For moderate disorder $W \lesssim 10$, they approach their ergodic value 1. For stronger disorder $10 \lesssim W \lesssim 16$, one observes a nonmonotonous behavior as a function of the system size: They first flow toward 0 for small enough N (corresponding to a localized behavior) and then start to increase at larger N. The characteristic size governing the crossover from the localized behavior to the extended behavior turns out to be very similar (within the numerical accuracy) to the one governing the crossover for the level statistics.

From the moments of the wave functions's amplitudes Υ_q one can also compute the entire spectrum of fractal

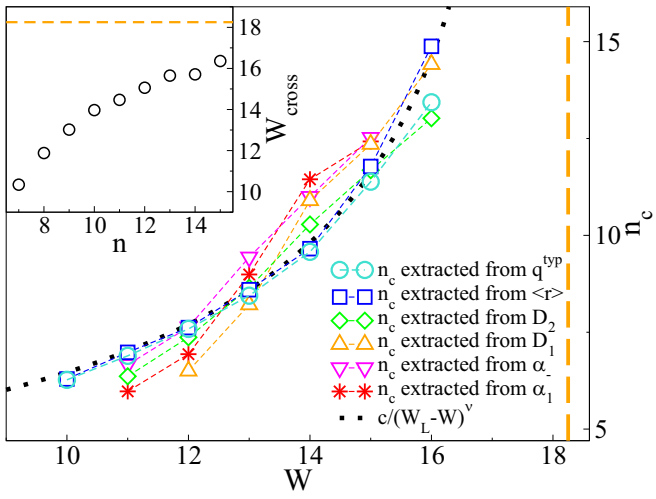

FIG. 3. Main panel: Characteristic crossover scales $n_c(W) = \log_2 N_c(W)$ extracted from different observables related to the level statistics ($\langle r \rangle$, squares and q^{typ}, circles) and to the statistics of the wave functions's amplitude, i.e., the fractal dimensions D_2 (diamonds) and D_1 (up triangles), the lower edge α_- of the support of the multifractal spectrum $f(\alpha)$ (down triangles), and the point α_1 where $f(\alpha_1) = \alpha_1$ and $f'(\alpha_1) = 1$ (stars). See Refs. [44,48] for more details. The black dotted curve is a fit of the form $n_c(W) \propto c/(W_L - W)^{\nu}$ with $c \approx 20$ and $\nu \approx 0.6$, where we used the value $W_L = 18.17$ [61]. Inset: Evolution with n of the crossing point of the curves of $q^{\text{typ}}(W)$ of Fig. 1 for two subsequent system sizes.

dimensions $f(\alpha)$, i.e., its Legendre transform (see, e.g., Refs. [36,48] for more details). The shape of $f(\alpha)$ can be characterized, for instance, by the left edge of its support α_- and by the point where $f(\alpha_1) = \alpha_1$ and $f'(\alpha_1) = 1$. As was shown in Ref. [48], the evolution of $f(\alpha)$ with the system size is also nonmonotonic: For small enough sizes the support of $f(\alpha)$ gets broader (corresponding to more localized states) and α_- and α_1 decrease upon increasing N. Conversely, for larger sizes the support of $f(\alpha)$ shrinks back (corresponding to more extended states), and α_- and α_1 increase with N. This kind of behavior is observed in the entire region $10 \lesssim W \lesssim 16$. The crossover scale governing the nonmonotonic evolution of the multifractal spectrum is once again very close (within the numerical accuracy) with the one characterizing the nonmonotonicity of the level statistics and of the anomalous dimensions.

This is clearly illustrated by the main panel of Fig. 3, where we plot the characteristic crossover scales $n_c(W)$ extracted from the different probes described above showing that, within the numerical accuracy, they all yield a very similar dependence on the disorder strength W (see Refs. [44,45,48] for more details). These results support the existence of a unique characteristic scale which controls the finite-size behavior of the statistics of the energy levels and of the wave functions's amplitudes, and is in agreement with the supersymmetric approach [47,50–55], which predicts a unique transition from a phase characterized by Poissonian statistics, localization, and lack-of-ergodicity to one displaying GOE statistics, delocalization, and full ergodicity.

As anticipated above, the nonmonotonic behavior can be explained in terms of the nature of the Anderson critical point on the RRG [44], which has properties similar to that of the localized phase [14,51,53,54,54,55], with critical level statistics of Poissonian form and strongly localized critical wave functions. The observables of systems of size $N \ll N_c(W)$ would then first flow upon increasing N towards the critical values, which tend, for $d \to \infty$, to the ones of the localized phase. Then, when N becomes larger than the correlation volume N_c, the observables flow towards their standard values in the delocalized, fully ergodic phase.

The black dotted curve of Fig. 3 shows a fit of the data of the form $n_c \propto c/(W_L - W)^{\nu}$, implying an exponential divergence of the correlation volume at the transition point. However, the numerical data are clearly too far from W_L (and perhaps too noisy) to obtain an accurate estimation of ν. Yet the value of the exponent is not too far from the one predicted by the supersymmetric analysis $\nu = 1/2$ [50–55]. In the next section we put forward a new large-deviation approach which allows one to access the crossover scale from the solution of the self-consistent equations for the Green's functions in the thermodynamic limit, providing a much more stringent test of the analytic predictions.

IV. SELF-CONSISTENT ITERATION EQUATIONS FOR THE GREEN'S FUNCTIONS AND LARGE DEVIATION METHOD

As discussed in the Introduction, the Anderson model on tree-like structures allows for an exact solution in the limit of infinite lattices [15,47,49–55,58–61], which yield the probability distribution function of the diagonal elements of the resolvent matrix, defined as $\mathcal{G}(z) = (\mathcal{H} - z\mathcal{I})^{-1}$.

To obtain the recursive equations, the key objects are the so-called *cavity* Green's functions, $G_{i\to j}(z) = [(\mathcal{H}_{i\leftrightarrow j} - z\mathcal{I})^{-1}]_{ii}$, i.e., the diagonal elements on site i of the resolvent matrix of the modified Hamiltonian $\mathcal{H}_{i\leftrightarrow j}$ where the edge between the site i and one of its neighbors j has been removed.

Take a given site i and its neighbors $\{l_1, \ldots, l_{k+1}\}$ living on an infinite tree. If one removes the site i from the graph, then the sites $\{l_1, \ldots, l_{k+1}\}$ are uncorrelated since the lattice would break in $k+1$ semi-infinite disconnected branches. One then obtains (e.g., by direct Gaussian integration or using the block matrix inversion formula) the following iteration relations for the cavity Green's functions [15]:

$$G^{-1}_{i\to l_m}(z) = -\epsilon_i - z - t^2 \sum_{l_j \in \partial i/l_m} G_{l_j \to i}(z), \tag{5}$$

where l_m with $m = 1, \ldots, k+1$ denote the excluded neighbor of i, $z = E + i\eta$, η is an infinitesimal imaginary regulator which smoothens out the pole-like singularities in the right-hand sides, ϵ_i is the on-site random energy taken from the distribution (2), and $\partial i/l$ denotes the set of all $k+1$ neighbors of i except l. (Note that for each site with $k+1$ neighbors one can define $k+1$ cavity Green's functions and $k+1$ recursion relations of this kind.) After that the solution of Eq. (5) has been found, one can finally obtain the diagonal elements of the resolvent matrix of the original problem on a given site i as a function of the cavity Green's functions for all the neighboring

sites [58]

$$\mathcal{G}_i^{-1}(z) = -\epsilon_i - z - t^2 \sum_{l_j \in \partial i} G_{l_j \to i}(z). \tag{6}$$

In the following we will focus on the middle of the spectrum ($E = 0$) and set $t = 1$.

The statistics of the diagonal elements of the resolvent gives—in the $\eta \to 0^+$ limit—the spectral properties of $\mathcal{H}$. In particular, the probability distribution of the LDoS at energy E is given by

$$\rho_i = \sum_\alpha |\psi_\alpha(i)|^2 \, \delta(E - E_\alpha) = \lim_{\eta \to 0^+} \frac{1}{\pi} \operatorname{Im} \mathcal{G}_i(z), \tag{7}$$

from which the average density of states (DoS) is simply given by $\rho = (1/N) \sum_i \rho_i = 1/(N\pi) \operatorname{Tr} \operatorname{Im} \mathcal{G}$.

Note that, however, on finite RRGs when site i is removed from the graph, the neighbors $\{l_1, \ldots, l_{k+1}\}$ are not truly decoupled since they are still connected by some (typically large) loop present somewhere in the system. Since the average size of the loops scales as $\ln N$ [43], it is reasonable to expect that Eqs. (5) and (6) become asymptotically exact in the thermodynamic limit as the cavity Green's functions on sites $\{l_1, \ldots, l_{k+1}\}$ become uncorrelated in absence of site i if the typical length of the loops, which connect them is larger than the correlation length. This has been, in fact, proven rigorously in Ref. [71] using the local convergence of RRGs to Cayley trees.

Since the Green's functions $G_{i \to j}$ and $\mathcal{G}_i$ are random variables, Eqs. (5) and (6) naturally lead to functional equations on their probability distribution $Q(G)$ and $P(\mathcal{G})$. From Eq. (5) one first gets the self-consistent functional equation for the probability distributions of the cavity Green's functions in the $N \to \infty$ limit (averaged over the on-site disorder and on different realizations of the random lattice)

$$Q(G) = \int dp(\epsilon) \prod_{l=1}^{k} \mathrm{d}Q(G_l) \, \delta\left(G^{-1} + \epsilon + z + \sum_{i=1}^{k} G_l \right), \tag{8}$$

where $p(\epsilon)$ is the probability distribution of the on-site random energy, Eq. (2). Once the fixed point of Eq. (8) is obtained, using Eq. (6) one can compute the probability distribution of the diagonal elements of the resolvent

$$P(\mathcal{G}) = \int dp(\epsilon) \prod_{l=1}^{k+1} \mathrm{d}Q(G_l) \, \delta\left(\mathcal{G}^{-1} + \epsilon + z + \sum_{l=1}^{k+1} G_l \right). \tag{9}$$

This set of functional equations can be solved numerically with an arbitrary degree of precision using a population dynamics algorithm [15,37,38,46,58,61,72].

For a population with $\mathcal{M}$ elements, the statistics of histograms obtained from the population yields directly an accuracy $O(\mathcal{M}^{-1})$. Nevertheless, for any such a population-dynamics approach, one not only has the population itself available, but still the equations used to iterate the population, here Eq. (8). This can be exploited to yield information about the desired distribution far beyond the straightforward accuracy $O(\mathcal{M}^{-1})$. This we will show below by introducing a new type of advanced large-deviation algorithm which allows us to sample the distribution $Q(G)$ of cavity Green's function with a very high precision in the tails.

To state the large-deviation algorithm, we need to specify explicitly the population dynamics approach [61,72]: We store a population $\{G_l\}$ of $\mathcal{M}$ complex-valued elements $G_l = a_l + ib_l$ ($l = 1, \ldots, \mathcal{M}$), i.e., $a_l = \operatorname{Re}(G_l)$ and $b_l = \operatorname{Im}(G_l)$. For each iteration step, we pick k randomly chosen elements $G_{l_j} = a_{l_j} + ib_{l_j}$ from the population and draw a uniformly distributed random number ϵ for the local energy according to Eq. (2). This allows us to calculate a new element from Eq. (5). Since below we will access the imaginary part of G separately, we use Eq. (5) in the following explicit form:

$$\begin{aligned} a + ib &= \frac{\left(-\epsilon - E - \sum_{j=1}^{k} a_{l_j}\right) + i\left(\sum_{j=1}^{k} b_{l_j} + \eta\right)}{\left(-\epsilon - E - \sum_{j=1}^{k} a_{l_j}\right)^2 + \left(\sum_{j=1}^{k} b_{l_j} + \eta\right)^2} \\ &\equiv f_{E + \sum_{j=1}^{k} a_{l_j}, \sum_{j=1}^{k} b_{l_j} + \eta}(\epsilon), \end{aligned} \tag{10}$$

which implies the definition of $f_{A,B}(\epsilon)$,

$$f_{A,B}(\epsilon) = \frac{(-\epsilon - A) + iB}{(-\epsilon - A)^2 + B^2} \tag{11}$$

for convenience. The iteration step is completed by replacing one randomly chosen element by the new one. This iteration is always performed until approximate convergence of the population, as established by monitoring the mean, variance, and few very small quantiles as well as the full shape of the distribution. Naturally, the resolution of the approximated distribution, represented by the population, is determined by the number of elements $\mathcal{M}$ in the population, as was deeply investigated in Ref. [61]. $\mathcal{M}$ should not be confused with the system size. The population dynamics approach implicitly assumes an infinite system size. Still, the finite population size plays a somewhat similar role as far as the critical properties are concerned. In particular, one observes that the critical point W_L systematically drifts to larger values as the pool size is increased, similarly to what observed in the inset of Fig. 3 increasing N.

Previous studies [15,58,61] showed that, in the localized phase, $W > W_L$ the iteration equations are unstable with respect to the imaginary regulator η: $Q(G)$ and $P(\mathcal{G})$ are singular and the average DoS vanishes in the $\eta \to 0^+$ limit. Conversely, in the metallic phase the probability distributions converge to stable nonsingular η-independent distribution functions, provided that η is sufficiently small.

For the distribution $Q(b)$ of the imaginary part $b \equiv \operatorname{Im} G$ we aim at obtaining the distribution to a high precision, i.e., deep in the tails. For this purpose, we implemented a large-deviation approach, which is explained next. Standard large-deviation algorithms rely on the sampling of biased distributions and unbiasing the obtained data in the end. Such approaches have been widely used, e.g., to study the large-deviation properties of random graphs [73,74], biological sequence alignments [75], protein folding [76], random walks [77,78], models of transport [79,80], the Kardar-Parisi-Zhang equation [81], nonequilibrium work processes [82], and many more. We tried such an approach based on a bias here, but were not able to see convergence of the used Markov chains deep enough in the tails. For this reason, we developed a very different approach here.

To convey the main idea, we notice that for any given set of randomly selected elements $\{G_{l_j}\}$, the next (and only) step is to sample random energy values according to the uniform distribution to obtain the probability of the imaginary part b conditioned to this set. This means, for the given set and given values of E and η, corresponding to $A = E + \sum_{j=1}^{k} a_{l_j}$ and $B = \sum_{j=1}^{k} b_{l_j} + \eta$, we have, by using a standard property of the delta function $\delta(x)$ and by using that the probability density for the local energies is simply $1/W$,

$$Q_{A,B}(b) = \int_{-W/2}^{W/2} \delta(b - \tilde{f}_{A,B}(\tilde{\epsilon}))\frac{1}{W} d\tilde{\epsilon} = \frac{1}{W}\int_{-W/2}^{W/2} \sum_{l:\tilde{f}_{A,B}(\tilde{\epsilon}_l)=b} \frac{1}{|\tilde{f}'_{A,B}(\tilde{\epsilon}_l)|}\delta(\tilde{\epsilon} - \tilde{\epsilon}_l) d\tilde{\epsilon}, \quad (12)$$

where $\tilde{f}_{A,B} = \mathrm{Im} f_{A,B}$ and $\tilde{\epsilon}_l$ are those real-valued zeros of $b - \tilde{f}_{A,B}(\tilde{\epsilon})$ which are located in the interval $[-W/2, W/2]$, and $\tilde{f}'_{A,B}(\epsilon)$ is the derivative of $\tilde{f}_{A,B}$ with respect to ϵ. The zeros of the function $b - \tilde{f}_{A,B}(\tilde{\epsilon})$ are easy to find because we only have to solve a quadratic equation, leading to $\tilde{\epsilon}_l = A \pm \sqrt{B/b - B^2}$.

Let us now assume that a arbitrary value b is given (fixed), where we want to evaluate $Q(b)$. The requirement that we only have to consider real-valued roots leads immediately to $b \leqslant 1/B$, i.e., $Q_{A,B}(b) = 0$ for $b > 1/B$.[1] This, on the other hand, means that to evaluate $Q(b)$, we could sample from the population such that only values are considered which follow this condition, i.e., where $B \leqslant 1/b$, i.e., $\sum_{j=1}^{k} b_{l_j} + \eta \leqslant 1/b$ holds. A simple way to achieve this restricted sampling is to sample values with $b_{l_j} \leqslant 1/b - \eta > 0$ since larger values will immediately lead to $Q(b) = 0$. Still, because a sum $B = \sum_{j=1}^{k} b_{l_j} + \eta$ is calculated, sometimes the combined sample values will not meet the condition $B \leqslant 1/b$, hence this gives rise no contribution to $Q(b)$ as well. But this rejection happens much less frequently compared to the sampling from the full distribution.[2] Thus, we restrict the sampling of all k-tuples to the region $\leqslant 1/b - \eta$ and include a bias $[\int_0^{1/b-\eta} \hat{Q}(\tilde{b})d\tilde{b}]^k$ ($\hat{Q}$ is the approximation of the true probability as given by the finite population) to all values of $Q_{A,B}(b)$ as calculated from Eq. (12). We technically achieved the restricted sampling by once sorting the population obtained in the standard population dynamics according to the value of the imaginary part b_l and subsequently drawing uniformly inside the desired range. Note that if the $l_{\max}$th element of the sorted population is the largest element which is inside the desired range, the bias is simply $(l_{\max}/\mathcal{M})^k$. For each value of b we were interested in, we performed N_{est} times this step of estimating $Q(b)$ and averaged over these estimates. In Fig. 4 the algorithm is summarized.

algorithm sampling $Q(b)$
begin
  Initialize population of $\mathcal{M}$ members.
  Iterate population using Eq. (10) until convergence
  for b in desired range
  begin
    $s = 0$
    for t=1 to N_{est}
    begin
      sample k elements $\{g_{l_i} = a_{l_i} + \mathrm{i}b_{l_i}\}$ with $b_{l_i} \leq 1/b - \eta$
      $A = E + \sum_{j=1}^{k} a_{l_j}$, $B = \sum_{j=1}^{k} b_{l_j} + \eta$
      calculate $Q_{A,B}(b)$ according to Eq. (12)
      $s = s + Q_{A,B}(b) \times [\int_0^{1/b-\eta} \hat{Q}(\tilde{b})d\tilde{b}]^k$
    end
    print b, s/N_{est}
  end
end

FIG. 4. Summary of the large-deviation sampling algorithm for the distribution of the imaginary part of the cavity Green's function (see text).

V. RESULTS

We applied the large-deviation approach described above within computer simulations [83] to obtain the distribution of the cavity Green's function for the Anderson model for the Bethe lattice with degree $k + 1 = 3$ with $E = 0$ and $\eta = 0$ for values of the disorder parameter $W \in [13, 17.3]$. For the population dynamics approach, we used a population size $\mathcal{M} = 10^7$. As discussed in Ref. [61] in great detail and mentioned above, the transition point was expected to be shifted to a slightly smaller value of the disorder compared to the exact asymptotic estimation obtained by diagonalizing explicitly the integral operator that governed the linear stability of the cavity equations in the localized phase in the $\mathcal{M} \to \infty$ limit, $W_L \simeq 18.17 \pm 0.01$. Below we indeed found a critical disorder of about $W_L \approx 17.77$.

To speed up convergence, since the imaginary parts b_l of the elements G_l are typically small with increasing value of G, we initialized the elements with random values uniformly distributed for the real parts as $a_l \sim U(-1, 1)$ and for the imaginary parts as $b_l \sim 10^{-\delta} U(0.5, 1.5)$. We used $\delta = 0$ (no special scaling) for $W \leqslant 16$ and $\delta = 9$ for $16 < W \leqslant 17.4$. For all values of W, we observed convergence when iterating the population 10^4 times [i.e., $10^4 \times \mathcal{M}$ times Eq. (10) is evaluated]. For the final estimate of $Q(b)$ we used $N_{\mathrm{est}} = 10^4$ and considered logarithmically spaced values of $b \geqslant 1$.

The resulting distributions $Q(b)$ for the imaginary part $b = \mathrm{Im}G$ is shown in Fig. 5. Note that using the large-deviation approach, probability densities as small as 10^{-50} can be accessed with a very high precision, well below any probability reached by a standard population dynamics approach. To extract the correlation volume, we assumed that the distribution follows the heuristic shape

$$f(b) = f_0 b^{-\lambda} \exp[-(b/N_c)^{\alpha}], \quad (13)$$

where the behavior for small values of $\mathrm{Im}G$ is governed by a power decay with exponent λ and the tail behavior by a stretched exponent with exponent α and scale N_c.

[1] This also follows directly from Eq. (10) because the imaginary part can be bounded from above by the value obtained for $(\epsilon - A)^2 = 0$.

[2] This could be improved even more by sampling the first element such that $b_{l_1} \leqslant 1/b - \eta$, then by sampling the second one such that $b_{l_2} \leqslant 1/b - \eta - b_{l_1}$ and so on, but this would increase the efficiency only by an factor of, at most, k (here $k = 2$), which we neglected because the final sampling is very fast anyway, the order of few seconds on a standard PC, as compared to the equilibration of the population, which takes more than one day.

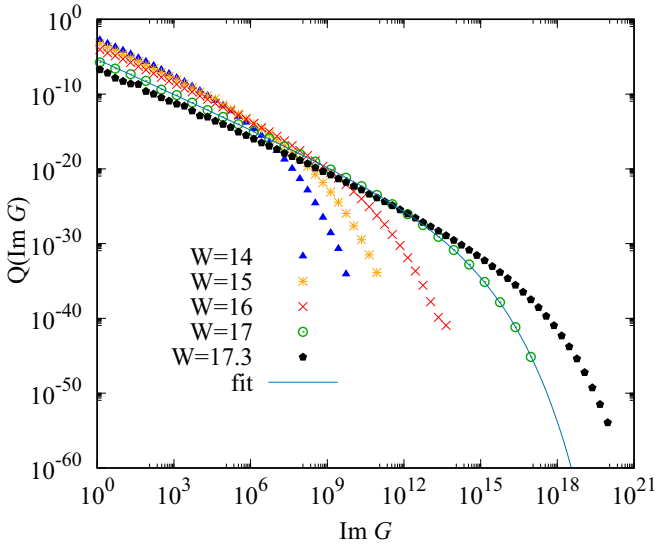

FIG. 5. Distribution $Q(\mathrm{Im}G)$ of the ImG of cavity Green's function for some values $W \in [13, 17.3]$. The line shows the result of a fit according to Eq. (13) to determine the correlation volume, see text.

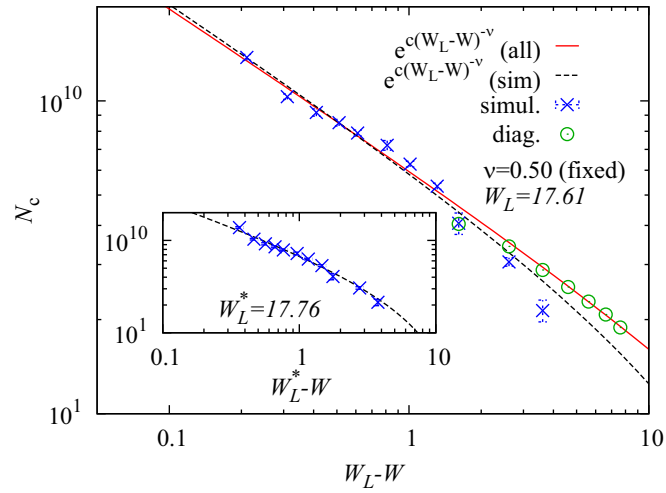

FIG. 6. Log of the correlation volume N_c as a function of the distance of the disorder parameter W from the critical point W_L. N_c is obtained from the cutoff of the tails of $Q(\mathrm{Im}G)$ using the large-deviation approach (crosses) and from the nonmonotonic behavior of q^{typ} using EDs (circles, corresponding to the circles of Fig. 3). The lines show the result of fits which model the divergence of the scale at W_L according to Eq. (14). The lower line is when fitting the large-deviation data only, while the upper line is for all data combined. Since we show all data and the fits as a function of $W_L - W$, and since the data points for the large deviation data are used for two different fits, we have to use for the plot a single value of W_L. Since the match between fit and data was better, we took the value of W_L as obtained from the fit for the combined data, which resulted in $W_L = 17.61(3)$. Thus, for the other shown fit using just the large deviation data, we fixed W_L to this value. Note that when, for testing purposes, W_L is allowed to adjust as well when fitting just the large-deviation data, a similar value $W_L = 17.76(8)$ results, but with better agreement, see inset.

Note that we also tried the fitting form given in Eq. (57) of Ref. [52], corresponding of the analytic prediction of the supersymmetric treatment for the asymptotic behavior of the tails of the probability distribution close to W_L. Such a function fits equally well the data of Fig. 5. However, it contains trade-off parameters for the correlation volume, i.e., it is possible to obtain good fits to the tail of the distributions over broader ranges of N_c for suitably chosen combinations of the values of the other parameters. Therefore, to obtain a more informative estimation of the correlation volume, we finally only considered Eq. (13). In any case, a key result of this analysis is that the functional form predicted by Ref. [52] fits remarkably well the tails of the distributions of the imaginary part of the Green's function over dozens of decades down to very small probability density, thereby providing a very stringent test for the validity of the supersymmetric approach.

By fitting the (log of the) distributions using the heuristic function (13) for the different values of W, we obtained the cutoff scale as a function of disorder strength W. Note that for λ we obtained values near 1.5, compatible with the prediction of the authors or Refs. [47,52]. We thus fixed $\lambda = 1.5$ for all values of W, resulting in less noisy data for N_c for the final fits. Note that the error bars obtained for W, also used practically below for fits, are purely statistical, i.e., rather small. For the exponent α, we obtained values in the range $\alpha \in [0.163(2), 0.208(2)]$ with a decreasing trend for growing values of λ. The results for N_c are shown in Fig. 6. We also show on the same plot the estimation of the correlation volume extracted from the nonmonotonic behavior of q^{typ} obtained via EDs (circles of Fig. 3, see also Fig. 2). This comparison is very insightful for two reasons. (i) The largest correlation volume obtained using the large-deviation approach for $W = 17.3$ is about 5.8×10^{13} which is almost 2^{46}. Thus, to observe such correlation volumes directly using EDs, one would have to treat RRGs of at least this size, which is, compared to the results shown in Sec. III, clearly not feasible with the current methods. (ii) The estimations of N_c obtained from the nonmonotonic behavior of the spectral statistics and from the cutoff of the tails of the probability distribution of the LDoS can have a different prefactor A appearing in Eq. (14), and asymptotically coincide only close enough to the Anderson transition. Far from the transition the two estimations can lead to quite different results. Yet, Fig. 6 shows that the two estimations of N_c are in surprisingly good agreement, even far below W_L.

We fitted the resulting scale values to the function

$$N_c(W) = A\, e^{c/(W_L - W)^{\nu}} \tag{14}$$

[actually by fitting $\log N_c(W) = \log A + c(W - W_L)^{\nu}$ to log of the measured scale]. We set the exponent ν to the value predicted by the supersymmetric formalism $\nu = 0.5$ [47,50–55,61]. When allowing W_L to adjust freely, we obtained estimates $W_L = 17.77(8)$ and $c = 21(3)$ (just statistical error bars), with a χ^2-score per degree of freedom equal to 5.1. This is of medium quality, but can be explained by the rather small purely statistical error bars for $N_c(W)$. The actual fit to the data looks rather good, as is visible in the inset of Fig. 6.

Although, as explained above, the correlation volume estimated from the cutoff of the tails using the large-deviation

method and from the nonmonotonicity of the observables related to the spectral statistics obtained from EDs might have different prefactors, for the sake of completeness it might be instructive to try to mix these data and fit them with the same function. More precisely, we combined the large deviation data and the data for N_c extracted from the nonmonotonic behavior of $q^{\rm typ}$ when using ED shown in Sec. III. Setting again $\nu = 0.5$ and letting W_L free to vary, we got $W_L = 17.61(3)$ and $c = 15(1)$. Since the data for N_c does not carry error bars, we cannot state a χ^2-score here.

We also tried a fit to the combined data with ν allowing to adjust freely. Here we obtained a value of $\nu = 0.26(15)$ which is compatible within 2 sigma with the predicted value $\nu = 0.5$. Since the error bar of the fit value is rather large, we prefer to use the fixed predicted value. Furthermore, the fact that one finds a value of ν smaller than $1/2$ is not surprising: As was recently discussed thoroughly in Ref. [61], obtaining the theoretical value $\nu = 0.5$ is indeed extremely difficult numerically. In this paper the authors found that the estimated value of the exponent obtained by fitting the divergence of the typical DoS using the standard pool method was affected by strong corrections due to the finite size of the pool and was significantly smaller than $1/2$ even when using huge populations. Upon increasing $\mathcal{M}$ the value of ν systematically increases and gets closer and closer to $1/2$. We argue that a similar effect is likely to be at play also within our large-deviation approach and that to get the correct value of the exponent one should carefully repeat the analysis increasing the size of the pool used to describe the bulk part of the probability distribution.

Finally, for another comparison, we also performed a fit of Eq. (14) to just the large-deviation-based data obtained here, but with the value $W_L = 17.61$ fixed, resulting in $c = 15.4(8)$ and a lower χ^2-score per degree of freedom of 6.1. As is visible in the main plot of Fig. 6, the match between the fitted function and data is less good here than for the other two fits.

VI. CONCLUSION AND PERSPECTIVES

In this paper we introduced a new large-deviation approach to investigate the critical behavior of the Anderson model on the RRG. This approach allows us to study the distribution of the imaginary part of the cavity Green's function down to very small probability tails, which is completely out of reach for standard numerical techniques.

In fact, as shown in Sec. III and previously discussed in Refs. [44,45,48], EDs clearly indicate the existence of a characteristic crossover scale $N_c(W)$ governing the finite-size effects of several observables and probes associated to the statistics of the gaps and of the eigenfunctions' amplitudes: For small sizes $N \ll N_c$ these observables seem first to flow towards towards the critical value upon increasing N (which on the RRG correspond to the ones of the localized phase [14,51,53–55]), and then for $N \gg N_c$ eventually approach the values corresponding to a standard delocalized, fully ergodic, phase. Although the ED estimation of $N_c(W)$ is compatible with an exponential divergence of the correlation volume upon approaching the Anderson transition, the numerical data are limited to relatively small sizes, $N \leqslant 2^{15}$, and thus can only access a disorder range too far from the transition to allow for an accurate determination of its critical behavior.

On the contrary, the large-deviation extension of the population dynamics approach allowed us to obtain accurately the distribution of the imaginary part of the cavity Green's function to very small probability densities as 10^{-50} (to obtain them by ED one would need a system size at least as large as $N = 2^{46}$ sites). The main idea is to first perform a standard population dynamics till convergence. In a second step, a biased sampling of the such-obtained histogram is made. This works out, because for given values of ImG, only a restricted range of the histograms contributes, and the magnitude of this range determines the bias used.

The functional form predicted by the supersymmetric formalism [52] fits remarkably well the distributions over dozens of decades and down to very low values of the probability density, thereby providing a very stringent test of the theoretical approach. These fits also yield a direct numerical estimation of the correlation volume N_c close to the critical disorder. Our data are compatible with a divergence of the logarithm of the correlation volume with a power of $\nu = 0.5$ [46,47,50–55,61]. The corresponding transition value that we find is $W_L \approx 17.77$.

All in all these results provide another transparent and coherent argument supporting the idea that the Anderson model on the RRG becomes fully ergodic in the entire delocalized phase, in agreement with the recent results of the authors of Refs. [44–46,48] and with the predictions of the authors of Refs. [47,50–55] based on supersymmetric field theory. Nonetheless, ergodicity establishes on a system size which becomes exponentially large as the localization transition is approached, and exceeds the system sizes accessible via ED well before the localization transition, resulting in a very wide crossover region in which the system looks as if it were in a mixed (delocalized but nonergodic) phase for all practical purposes, i.e., on finite but large length and timescales [volumes smaller than $N_c(W)$ and times smaller than $\hbar/N_c^{-1}(W)$].

In light of the analogy between Anderson localization on Bethe lattices and many-body localization [16,23–29], the results presented above might help us understand the highly nontrivial properties of the delocalized phase of many-body interacting disordered systems exhibiting MBL.

On the methodological side, our approach might present a new type of large-deviation approach. It could be helpful also for other models, where similar self-consistent equations, like for field distributions, are obtained. The key property is Eq. (12), which gives the contribution to an arbitrary location b of the desired distribution Q as a function of any given sample of the previously obtained population and as a function of the underlying disorder distribution. In the case where the sampling of the population can be effectively restricted to the relevant values, depending on b, and if the zeros of the delta function can be obtained efficiently, our proposed approach should be useful.

ACKNOWLEDGMENTS

This research was partially supported by a grant from the Simons Foundation (Grant No. 454935). We thank the Centre National de la Recherche Scientifique for supporting AKH during a guest professorship at Sorbonne Université. The

simulations were performed at the HPC Cluster CARL, located at the University of Oldenburg (Germany) and funded by the Deutsche Frschungsmemeinschaft (DFG) through its Major Research Instrumentation Program (INST 184/157-1 FUGG) and the Ministry of Science and Culture (MWK) of the Lower Saxony State.

[1] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).
[2] P. A. Lee and T. V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985).
[3] F. Evers and A. D. Mirlin, Rev. Mod. Phys. **80**, 1355 (2008).
[4] A. Lagendijk, B. V. Tiggelen, and D. S. Wiersma, Phys. Today **62**(8), 24 (2009).
[5] A. Aspect and M. Inguscio, Phys. Today **62**(8), 30 (2009).
[6] G. Roati, C. D'Errico, L. Fallani, M. Fattori, C. Fort, M. Zaccanti, G. Modugno, M. Modugno, and M. Inguscio, Nature (London) **453**, 895 (2008).
[7] J. Billy, V. Josse, Z. Zuo, A. Bernard, B. Hambrecht, P. Lugan, D. Clément, L. Sanchez-Palencia, P. Bouyer, and A. Aspect, Nature (London) **453**, 891 (2008).
[8] S. S. Kondov, W. R. McGehee, J. J. Zirbel, and B. DeMarco, Science **334**, 66 (2011).
[9] F. Jendrzejewski, A. Bernard, K. Müller, P. Cheinet, V. Josse, M. Piraud, L. Pezzé, L. Sanchez-Palencia, A. Aspect, and P. Bouyer, Nat. Phys. **8**, 398 (2012).
[10] G. Semeghini, M. Landini, P. Castilho, S. Roy, G. Spagnolli, A. Trenkwalder, M. Fattori, M. Inguscio, and G. Modugno, Nat. Phys. **11**, 554 (2015).
[11] J. Chabé, G. Lemarié, B. Grémaud, D. Delande, P. Szriftgiser, and J. C. Garreau, Phys. Rev. Lett. **101**, 255702 (2008).
[12] H. Hu, A. Strybulevych, J. H. Page, S. E. Skipetrov, and B. A. van Tiggelen, Nat. Phys. **4**, 945 (2008).
[13] M. S. Foster, S. Ryu, and A. W. W. Ludwig, Phys. Rev. B **80**, 075101 (2009).
[14] E. Tarquini, G. Biroli, and M. Tarzia, Phys. Rev. B **95**, 094204 (2017).
[15] R. Abou-Chacra, P. W. Anderson, and D. J. Thouless, J. Phys. C **6**, 1734 (1973).
[16] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, Ann. Phys. (NY) **321**, 1126 (2006).
[17] I. V. Gornyi, A. D. Mirlin, and D. G. Polyakov, Phys. Rev. Lett. **95**, 206603 (2005).
[18] E. Altman and R. Vosk, Annu. Rev. Condens. Matter Phys. **6**, 383 (2015).
[19] Nandkishore and D. A. Huse, Annu. Rev. Condens. Matter Phys. **6**, 15 (2015).
[20] D. A. Abanin and Z. Papić, Ann. Phys. (Leipzig) **529**, 1700169 (2017).
[21] N. Alet and N. Laflorencie, C. R. Phys. **19**, 498 (2018).
[22] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Rev. Mod. Phys. **91**, 021001 (2019).
[23] D. E. Logan and P. G. Wolynes, Phys. Rev. B **36**, 4135 (1987); J. Chem. Phys. **93**, 4994 (1990); R. Bigwood, M. Gruebele, D. M. Leitner, and P. G. Wolynes, Proc. Natl. Acad. Sci. **95**, 5960 (1998).
[24] B. L. Altshuler, Y. Gefen, A. Kamenev, and L. S. Levitov, Phys. Rev. Lett. **78**, 2803 (1997).
[25] Ph. Jacquod and D. L. Shepelyansky, Phys. Rev. Lett. **79**, 1837 (1997).
[26] A. De Luca and A. Scardicchio, Europhys. Lett. **101**, 37003 (2013).
[27] S. Roy and D. E. Logan, Phys. Rev. B **101**, 134202 (2020).
[28] K. S. Tikhonov and A. D. Mirlin, Ann. Phys. (NY) **435**, 168525 (2021).
[29] G. Biroli and M. Tarzia, Phys. Rev. B **96**, 201114(R) (2017).
[30] G. Biroli and M. Tarzia, Phys. Rev. B **102**, 064211 (2020).
[31] K. S. Tikhonov and A. D. Mirlin, Phys. Rev. B **103**, 064204 (2021).
[32] O. Bohigas, M. J. Giannoni, and C. Schmit, Phys. Rev. Lett. **52**, 1 (1984).
[33] M. V. Berry, Proc. R. Soc. London A **400**, 229 (1985).
[34] A. V. Andreev, O. Agam, B. D. Simons, and B. L. Altshuler, Phys. Rev. Lett. **76**, 3947 (1996).
[35] G. Biroli, A. C. Ribeiro-Teixeira, and M. Tarzia, arXiv:1211.7334.
[36] A. De Luca, B. L. Altshuler, V. E. Kravtsov, and A. Scardicchio, Phys. Rev. Lett. **113**, 046806 (2014); A. De Luca, A. Scardicchio, V. E. Kravtsov, and B. L. Altshuler, arXiv:1401.0019.
[37] B. L. Altshuler, E. Cuevas, L. B. Ioffe, and V. E. Kravtsov, Phys. Rev. Lett. **117**, 156601 (2016); B. L. Altshuler, L. B. Ioffe, and V. E. Kravtsov, arXiv:1610.00758.
[38] V. E. Kravtsov, B. L. Altshuler, and L. B. Ioffe, Ann. Phys. (NY) **389**, 148 (2018).
[39] S. Bera, G. De Tomasi, I. M. Khaymovich, and A. Scardicchio, Phys. Rev. B **98**, 134205 (2018).
[40] G. De Tomasi, S. Bera, A. Scardicchio, and I. M. Khaymovich, Phys. Rev. B **101**, 100201(R) (2020).
[41] S. Savitz, C. Peng, and G. Refael, Phys. Rev. B **100**, 094201 (2019).
[42] M. Pino, Phys. Rev. Research **2**, 042031(R) (2020).
[43] N. C. Wormald, *Models of Random-Regular Graphs, in Surveys in Combinatorics*, London Mathematical Society Lecture Note Series, edited by J. D. Lamb and D. A. Preece (Cambridge University Press, Cambridge, England, 1999), pp. 276, 239.
[44] K. S. Tikhonov, A. D. Mirlin, and M. A. Skvortsov, Phys. Rev. B **94**, 220203(R) (2016).
[45] E. Tarquini, G. Biroli, and M. Tarzia, Phys. Rev. Lett. **116**, 010601 (2016).
[46] I. Garcia-Mata, O. Giraud, B. Georgeot, J. Martin, R. Dubertrand, and G. Lemarié, Phys. Rev. Lett. **118**, 166801 (2017).
[47] K. S. Tikhonov and A. D. Mirlin, Phys. Rev. B **99**, 024202 (2019).
[48] G. Biroli and M. Tarzia, arXiv:1810.07545.
[49] F. L. Metz and I. P. Castillo, Phys. Rev. B **96**, 064202 (2017).
[50] Y. V. Fyodorov and A. D. Mirlin, J. Phys. A: Math. Gen. **24**, 2219 (1991); Phys. Rev. Lett. **67**, 2049 (1991); Y. V. Fyodorov, A. D. Mirlin, and H.-J. Sommers, J. Phys. I France **2**, 1571 (1992).

[51] A. D. Mirlin and Y. V. Fyodorov, Nucl. Phys. B **366**, 507 (1991); Phys. Rev. B **56**, 13393 (1997).
[52] A. Mirlin and Y. Fyodorov, J. Phys. I France **4**, 655 (1994).
[53] M. R. Zirnbauer, Phys. Rev. B **34**, 6394 (1986); Nucl. Phys. B **265**, 375 (1986).
[54] K. B. Efetov, Adv. Phys. **32**, 53 (1983); Zh. Eksp. Teor. Fiz. **88**, 1032 (1985); **92**, 638 (1987); **93**, 1125 (1987).
[55] J. J. M. Verbaarshot, Nucl. Phys. B **300**, 263 (1988).
[56] K. S. Tikhonov and A. D. Mirlin, Phys. Rev. B **94**, 184203 (2016); M. Sonner, K. S. Tikhonov, and A. D. Mirlin, *ibid.* **96**, 214204 (2017).
[57] C. Monthus and T. Garel, J. Phys. A: Math. Theor. **44**, 145001 (2011).
[58] G. Biroli, G. Semerjian, and M. Tarzia, Prog. Theor. Phys. Suppl. **184**, 187 (2010).
[59] M. Aizenman and S. Warzel, J. Math. Phys. **53**, 095205 (2012); Phys. Rev. Lett. **106**, 136804 (2011).
[60] V. Bapst and G. Semerjian, J. Stat. Phys. **145**, 51 (2011); V. Bapst, J. Math. Phys. **55**, 092101 (2014).
[61] K. S. Tikhonov and A. D. Mirlin, Phys. Rev. B **99**, 214202 (2019).
[62] G. Parisi, S. Pascazio, F. Pietracaprina, V. Ros, and A. Scardicchio, J. Phys. A: Math. Theor. **53**, 014003 (2020).
[63] I. Oren, A. Godel, and U. Smilansky, J. Phys. A: Math. Theor. **42**, 415101 (2009); I. Oren and U. Smilansky, *ibid.* **43**, 225205 (2010).
[64] R. Bauerschmidt, J. Huang, A. Knowles, and H.-T. Yau, Ann. Probab. **45**, 3626 (2017); R. Bauerschmidt, A. Knowles, and H. T. Yau, Comm. Pure Appl. Math. **70**, 1898 (2017).
[65] M. Sade and R. Berkovits, Phys. Rev. B **68**, 193102 (2003).
[66] V. Oganesyan and D. A. Huse, Phys. Rev. B **75**, 155111 (2007).
[67] Y. Y. Atas, E. Bogomolny, O. Giraud, and P. Vivo, J. Phys. A: Math. Theor. **46**, 355204 (2013).
[68] C. Porter and R. Thomas, Phys. Rev. **104**, 483 (1956).
[69] Perhaps the best example of that is provided by the generalization of the Rosenzweig-Porter random matrix model of V. E. Kravtsov, I. M. Khaymovich, E. Cuevas, and M. Amini, New J. Phys. **17**, 122002 (2015), where there is an entire region of the parameter space where wave functions are delocalized, but multifractal and strongly correlated, while the statistics of the neighboring gaps is still described by the GOE ensemble. In this case one numerically finds that $\langle q \rangle$ converges to its GOE universal value $2/\pi$ irrespective of the fact that the wave functions's amplitudes are not uncorrelated Gaussian random variables of variance $1/N$.
[70] The logarithmic derivatives are computed as discrete derivatives involving the five available values of the system size closest to N, but we checked that using three or seven points does not significantly modify the results.
[71] C. Bordenave and M. Lelarge, Random Struct. Alg. **37**, 332 (2010).
[72] M. Mézard and G. Parisi, Eur. Phys. J. B **20**, 217 (2001).
[73] A. K. Hartmann, Eur. Phys. J. B **84**, 627 (2011).
[74] A. K. Hartmann and M. Mézard, Phys. Rev. E **97**, 032128 (2018).
[75] A. K. Hartmann, Phys. Rev. E **65**, 056102 (2002).
[76] D. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, J. Chem. Phys. **108**, 1964 (1998).
[77] A. K. Hartmann, S. N. Majumdar, and A. Rosso, Phys. Rev. E **88**, 022119 (2013).
[78] G. Claussen, A. K. Hartmann, and S. N. Majumdar, Phys. Rev. E **91**, 052104 (2015).
[79] C. Giardina, J. Kurchan, and L. Peliti, Phys. Rev. Lett. **96**, 120603 (2006).
[80] W. Staffeldt and A. K. Hartmann, Phys. Rev. E **100**, 062301 (2019).
[81] A. K. Hartmann, P. Le Doussal, S. N. Majumdar, A. Rosso, and G. Schehr, Europhys. Lett. **121**, 67004 (2018).
[82] A. K. Hartmann, Phys. Rev. E **89**, 052103 (2014).
[83] A. K. Hartmann, *Big Practical Guide to Computer Simulations* (World Scientific, Singapore, 2015).