

Forward approximation as a mean-field approximation for the Anderson and many-body localization transitions

Francesca Pietracaprina and Valentina Ros

*SISSA - International School for Advanced Studies, Via Bonomea 265, 34136 Trieste, Italy
and INFN Sezione di Trieste, Via Valerio 2, 34127 Trieste, Italy*

Antonello Scardicchio

*Abdus Salam ICTP, Strada Costiera 11, 34151 Trieste, Italy;
INFN Sezione di Trieste, Via Valerio 2, 34127 Trieste, Italy;*

and Dipartimento di Fisica, Università degli Studi di Bari "Aldo Moro", I-70126 Bari, Italy

(Received 25 September 2015; revised manuscript received 7 January 2016; published 1 February 2016)

In this paper we analyze the predictions of the forward approximation in some models which exhibit an Anderson (single-body) or many-body localized phase. This approximation, which consists of summing over the amplitudes of only the shortest paths in the locator expansion, is known to overestimate the critical value of the disorder which determines the onset of the localized phase. Nevertheless, the results provided by the approximation become more and more accurate as the local coordination (dimensionality) of the graph, defined by the hopping matrix, is made larger. In this sense, the forward approximation can be regarded as a mean-field theory for the Anderson transition in infinite dimensions. The sum can be efficiently computed using transfer matrix techniques, and the results are compared with the most precise exact diagonalization results available. For the Anderson problem, we find a critical value of the disorder which is 0.9% off the most precise available numerical value already in 5 spatial dimensions, while for the many-body localized phase of the Heisenberg model with random fields the critical disorder $h_c = 4.0 \pm 0.3$ is strikingly close to the most recent results obtained by exact diagonalization. In both cases we obtain a critical exponent $\nu = 1$. In the Anderson case, the latter does not show dependence on the dimensionality, as it is common within mean-field approximations. We discuss the relevance of the correlations between the shortest paths for both the single- and many-body problems, and comment on the connections of our results with the problem of directed polymers in random medium.

DOI: [10.1103/PhysRevB.93.054201](https://doi.org/10.1103/PhysRevB.93.054201)

I. INTRODUCTION

The propagation of waves and quantum particles in a disordered medium is a fascinating and challenging problem in statistical mechanics, with plenty of relevance for experiments [1–3]. Among the phenomena that occur in such a setting, Anderson localization is probably the most striking [1,4]. “Anderson’s theorem” in Ref. [4] states that for sufficiently strong disorder, diffusive transport is completely suppressed in single-particle problems on a lattice. The study of the transition that separates the two phases (localized and delocalized) has resisted an exact solution for about 60 years, and numerical methods are still at the core of the advances in this topic [5].

Recently, interest in Anderson localization has surged due to the work by Basko, Aleiner, and Altshuler [6], henceforth denoted by BAA. There, the phenomenon of many-body localization (MBL), i.e., the stability of the Anderson insulator to the addition of interactions, is investigated. BAA’s work has been extended and re-interpreted in several other works [7–17], and MBL appears now to be the most robust mechanism to break the ergodicity that is typical of generic interacting systems.

The core of BAA’s analysis relies on perturbatively accounting for the interactions, at finite temperature and particle density. Technically, they consider the perturbation theory for the imaginary part of the propagator of an excitation on top of an eigenstate by means of the Keldysh formalism. For sufficiently weak interactions, the perturbative series is shown to converge with probability equal to 1. In the spirit of Ref. [4]

this implies the localization of the excitations themselves, and the absence of transport.

As an alternative route, the MBL problem can be interpreted as a single-particle tight-binding problem in the space of many-body configurations [18], with the interactions playing the role of an effective hopping. However, several issues arise in this formulation. First, the on-site energies in the resulting effective lattice are not independent variables drawn from the same distribution, but they are strongly correlated. Second, the connectivity of a configuration in the many-body problem scales with (a power of) the system size: it diverges in the thermodynamic limit, and thus it is impossible to define a limiting graph. As a consequence, one needs to define an effective connectivity which stays of $O(1)$ in the thermodynamic limit (a similar phenomenon is observed in Ref. [19]). Finally, the number of paths of a given length connecting two many-body configurations grows factorially with the distance between them, the distance being defined as the minimum number of actions of the interaction operator needed to connect the initial to the final configuration. Since the distance between two configurations can be of the order of the system size, the number of paths can grow factorially in the system size. When the mapping to a single-particle problem is applied to a system of N interacting spins, it results in a correlated-disorder problem on a (section of) an N -dimensional hypercube.

Despite these complications, in some recent works [14,16,19] this approach has been used successfully to

estimate, among other things, the boundaries of the MBL region. The analytical calculations in Refs. [14,19] were shown to be in very good agreement with the numerical results in the same papers, obtained with exact diagonalization. The analytic results are derived revisiting an approximation already used in Refs. [4,18,20,21], which consists in calculating the Green's functions by retaining only the lowest order in the hopping. Recently, this approximation has been used in Ref. [22] to derive the power-law tail of the distribution of the wave function amplitudes on a Bethe lattice in the localized phase.

In this paper we discuss in detail this approximation, dubbed "forward approximation" (FA), by illustrating its virtues and limitations and its connections to other, seemingly unrelated problems of statistical physics. The paper is organized as follows: In Sec. II we derive the expression for the wave function amplitudes in forward approximation, and discuss a criterion for localization given in terms of the probability of resonances. In Sec. III, after recalling some known results on the Anderson model on the Bethe lattice, we focus on d -dimensional systems (with $d = 3-6$). We compare the analytic results we get for the critical disorder within the FA with the numerics, showing how the approximation gives better results as the dimensionality d is increased. We then discuss the application of the aforementioned technique to a many-body problem of interacting spins in a disordered environment. In Sec. IV we discuss the relevance of the correlations and interference between different paths connecting two configurations, both for the single-particle and for the many-body problem. In the Conclusion we comment on the various possible directions in which this work can be extended, focusing in particular on the application of the forward approximation to the study of the MBL phase and of the many-body localization-delocalization transition.

II. DERIVATION

A. The forward approximation for the eigenfunctions

To begin with, we derive the expression for the wave function amplitudes in FA for a single-particle hopping on a finite lattice with on-site disorder. We consider the Hamiltonian

$$H = \sum_{i=1}^N \epsilon_i c_i^\dagger c_i + t \sum_{\langle i,j \rangle} (c_i^\dagger c_j + \text{H.c.}). \quad (1)$$

In the Anderson model, the ϵ_i are independent random variables uniformly distributed in $[-W/2, W/2]$, one for each of the N sites in the lattice. The edges $\langle i,j \rangle$ define the lattice geometry. The lattice constant is set to 1, and we denote with L the length scale characterizing the size of the lattice (for a cubic lattice in dimension d the diameter is $L = N^{1/d}$, and for a Bethe lattice or regular random graph it is $L = \ln[(N-1)(K-1)/(K+1)+1]/\ln K \approx \ln N/\ln K$, where $K+1$ is the connectivity of the lattice). The distance $d(a,b)$ between two arbitrary sites a,b of the lattice is the number of edges in a shortest path connecting them. We refer to it as the lattice distance in the following.

We consider the matrix elements of the resolvent between two states associated to the sites a,b in the lattice:

$$G(b,a,E) = \langle b | \frac{1}{E-H} | a \rangle, \quad (2)$$

at energy E . They have the following expansion in series of t :

$$G(b,a,E) = \frac{1}{E-\epsilon_a} \sum_{p \in \text{paths}(a,b)} \prod_{i \in p} \frac{t}{E-\epsilon_i}, \quad (3)$$

where the sum runs over all the paths p in the lattice connecting the sites a and b , and i labels the sites visited by the path p (site a excluded). Given two sites a,b at lattice distance n , the lowest orders in the expansion are

$$G(b,a,E) = \frac{1}{E-\epsilon_a} \frac{t}{E-\epsilon_1} \frac{t}{E-\epsilon_2} \cdots \frac{t}{E-\epsilon_{n-1}} \frac{t}{E-\epsilon_b} + \text{other paths of length } n + \cdots, \quad (4)$$

where the sites $(a,1,2,\dots,n-1,b)$ belong to one of the shortest paths connecting a and b .

On the other hand, from the spectral decomposition it follows

$$G(b,a,E) = \sum_{\alpha} \frac{\psi_{\alpha}(b)\psi_{\alpha}^*(a)}{E-E_{\alpha}}, \quad (5)$$

and thus the residue at $E = E_{\alpha}$ gives

$$\lim_{E \rightarrow E_{\alpha}} (E - E_{\alpha})G(b,a,E) = \psi_{\alpha}(b)\psi_{\alpha}^*(a), \quad (6)$$

assuming no degeneracy of the eigenvalues.

The path representation (3) does not have a pole at E_{α} , to no order in t . To get the exact poles it is necessary to resum the closed paths in the series expansion. Once this is done, the full series is recast into a sum over the *nonrepeating paths* $\text{paths}^*(a,b)$ connecting the sites a and b :

$$G(b,a,E) = \frac{1}{E-\epsilon_a - \Sigma_a(E)} \times \sum_{p \in \text{paths}^*(a,b)} \prod_{i \in p} \frac{t}{E-\epsilon_i - \Sigma_i^{(p)}(E)}. \quad (7)$$

Here $\Sigma_a(E)$ is the local self-energy at the site a , defined through the identity

$$G(a,a,E) \equiv \frac{1}{E-\epsilon_a - \Sigma_a(E)}. \quad (8)$$

It is equal to the sum of the amplitudes of all the closed paths in which site a appears only as starting and ending point; to lowest order in t

$$\Sigma_a(E) = \sum_{j \in \partial a} \frac{t^2}{E-\epsilon_j} + O(t^3), \quad (9)$$

where ∂a is the set of nearest-neighboring sites of a . The path-dependent term $\Sigma_i^{(p)}(E)$ is a modified self-energy, which resums the loops around site i , never crossing site i again, or any of the sites $(a,1,\dots,i-1)$ already visited by the nonrepeating path p .

The expansion (7) in nonrepeating paths has several advantages. First, while $\text{paths}(a,b)$ is an infinite set, even for a finite lattice, $\text{paths}^*(a,b)$ is finite for a finite lattice. Moreover, (7) is free of the divergences affecting (3) that are due to *local resonances*, i.e., to sites i,j at bounded distance satisfying $|\epsilon_i - \epsilon_j| \sim \Delta$, with Δ small (we clarify the exact meaning of *small* in Sec. II B). Local resonances necessarily occur also in the localized phase, and produce arbitrarily large factors

in Eq. (3), corresponding to the paths repeatedly hitting the resonant sites an arbitrary number of times. These repetitions of large factors lead to the divergence of the naive perturbation series in t , but are resummed into self-energy corrections in Eq. (7). The “renormalized” expansion in nonrepeating paths is found to converge in the localized phase [4], when resonances do not proliferate at asymptotically large distances in space, and the hopping hybridizes only the degrees of freedom in a finite, albeit possibly big, region of space. An analogous resummation procedure is discussed in Ref. [16], where the perturbation theory for quasilocal conserved operators is shown to converge in the MBL phase.

The expression for the eigenfunction is obtained from the resolvent as follows: the eigenenergy E_α satisfies $E_\alpha = \epsilon_a + \Sigma_\alpha(E_\alpha)$; thus the first factor of (7) has a pole at E_α with residue $|\psi_\alpha(a)|^2$, as follows from (8) and (6). Then,

$$\begin{aligned} & \lim_{E \rightarrow E_\alpha} (E - E_\alpha) G(b, a, E) \\ &= |\psi_\alpha(a)|^2 \lim_{E \rightarrow E_\alpha} \sum_{p \in \text{paths}^*(a, b)} \prod_{i \in p} \frac{t}{E - \epsilon_i - \Sigma_i^{(p)}(E)}, \end{aligned} \quad (10)$$

which gives

$$\psi_\alpha(b) = \psi_\alpha(a) \sum_{p \in \text{paths}^*(a, b)} \prod_{i \in p} \frac{t}{E_\alpha - \epsilon_i - \Sigma_i^{(p)}(E_\alpha)}, \quad (11)$$

with $\psi_\alpha(a)$ obtained from $\Sigma_\alpha(E_\alpha)$ using (6) [23].

From (11) we can read the expression of the wave function amplitudes to lowest order in t . Assume that α labels an eigenstate localized at site a for $t \rightarrow 0$. Since $\Sigma_\alpha = O(t^2)$, we have to lowest order $E_\alpha \rightarrow \epsilon_a$, $\psi_\alpha(a) \rightarrow 1$, $\Sigma_i \rightarrow 0$, giving

$$\psi_\alpha(b) = \sum_{p \in \text{spaths}(a, b)} \prod_{i \in p} \frac{t}{\epsilon_a - \epsilon_i}, \quad (12)$$

where the set $\text{spaths}(a, b) \subset \text{paths}^*(a, b)$ contains the shortest paths from a to b . Note that this derivation does not rely on the particular structure of the lattice or on the independence of the on-site energies, and thus it can be straightforwardly generalized to hopping problems on graphs with different geometry or more general distribution of the random energies, such as the ones in Sec. III C.

The effect of the modified self-energy corrections $\Sigma_i^{(p)}(\epsilon_a)$ is to weaken the role of resonances. Indeed, let us assume that in the path p there is a site i which is resonant with a , $|\epsilon_a - \epsilon_i| \sim \Delta$. In this case the forward approximation (12) will contain the very large term t/Δ . However, the correction $\Sigma_{i-1}^{(p)}(\epsilon_a)$ in the previous site also contains such large term, leading to a compensation. As we shall see in the following, neglecting this effect leads to an overestimate of the minimum disorder strength needed to localize the system. On the other hand, the loops contributing to the self-energy corrections become less relevant when the dimensionality (or connectivity) of the lattice is increased. Thus, the FA is expected to give faithful results in higher dimension (see Sec. III B).

B. Probability of resonances and criterion for localization

For a single-particle problem on a finite-dimensional lattice, we define an eigenstate ψ_α of a system of size L localized if (with probability 1 over the disorder realizations) the

probability of finding a particle at a distance $O(L)$ from the localization center of the state tends to zero in the limit of large L . More precisely, let a denote the localization center of ψ_α , and

$$\psi_r \equiv \max_{b: d(b, a) = r} |\psi_\alpha(b)|. \quad (13)$$

We define the state ψ_α localized if there exists a finite $\xi > 0$ such that

$$P\left(\frac{\ln|\psi_r|^2}{2r} \leq -\frac{1}{2\xi}\right) \rightarrow 1 \quad \text{for } r \rightarrow \infty. \quad (14)$$

Namely, we require that the random numbers $|\psi_\alpha(b)|^2$ can be enclosed in an exponential envelope for all b sufficiently far from the localization center of the state. By means of the Kubo formula, it is possible to show that this condition on the eigenstates implies the vanishing of the dc conductivity. We identify the localization length of ψ_α with the minimum value of ξ for which Eq. (14) is true. It does, in general, depend on the state; however, it is supposed to depend smoothly on the energy E_α in the thermodynamic limit. A mobility edge exists whenever there is band of energies for which such minimum is not finite. In particular, at fixed energy and at the corresponding critical value of disorder $W = W_c$, the localization length diverges and the asymptotic bound in Eq. (14) ceases to hold for any finite ξ . This entails that for any arbitrarily small, positive $\epsilon = \xi^{-1}$ and at arbitrarily large distances r from the localization center, there exist sites b such that the ratio $\ln|\psi_\alpha(b)|^2/2r$ exceeds the constant $-\epsilon$ with some finite probability. We expect the delocalized phase $W < W_c$ to be characterized by the stronger condition:

$$P\left(\frac{\ln|\psi_r|^2}{2r} \geq -\epsilon\right) \rightarrow 1 \quad \text{for } r \rightarrow \infty, \quad (15)$$

holding for any arbitrarily small, strictly positive value of ϵ .

From Eqs. (14) and (15) it follows that the localization-delocalization transition can be detected analyzing the statistics of the wave function amplitudes as a function of distance. In the following, we compute the wave function amplitudes in FA and determine numerically the probability in Eq. (15), choosing ϵ of the order of the numerical precision. We refer to the resulting probability as the “probability of resonances.” The terminology is motivated by the fact that within the FA, an amplitude of $O(1)$ at a site b at distance r from the localization center a corresponds to a resonance between the two sites. Indeed, the corresponding two-site problem can be considered as a two-level system with reduced Hamiltonian

$$h = \begin{pmatrix} 0 & h_r \\ h_r & \Delta \end{pmatrix}, \quad (16)$$

where $\Delta = \epsilon_a - \epsilon_b$, and

$$h_r = t \sum_{p \in \text{paths}^*(a, b)} \prod_{i \in p} \frac{t}{\epsilon_a - \epsilon_i - \Sigma_i^{(p)}(\epsilon_a)}, \quad (17)$$

where the products are taken over all sites in the path, excluding a, b . The sites are resonant when the energy difference $\epsilon_a - \epsilon_b$ is small, and precisely $|\Delta| < h_r$. Considering h_r to lowest order in t , one finds that this is equivalent to $|\psi(b)| > 1$, with $\psi(b)$ computed in the lowest order FA. Thus, with (14) and (15) one is probing the statistics of resonances within the FA,

and requiring that the probability to find at least a resonant site at *any* sufficiently large distance r from the localization center decays to zero in the localized phase. This criterion involving resonances allows us to obtain an estimate for the critical disorder within the FA also on the Bethe lattice, where the exact eigenstates satisfy Eq. (14) even in the delocalized phase, due to normalization.

III. RESULTS

A. Warming up: The Bethe lattice

The simplest setting for the application of the forward approximation is a Bethe lattice: in this case, given two sites a, b there is only one nonrepeating path connecting them, along which all the energies are i.i.d. This makes the problem amenable to analytic calculations [18,20,24]. We briefly recall some results in the following.

Let a be the root of the tree, and K the branching number (the connectivity is $K + 1$). Within the FA we get that the wave function at one particular point at distance L from the root is given by

$$\psi_L = \prod_{i=1}^L \frac{t}{\epsilon_a - \epsilon_i}. \quad (18)$$

This random variable has a power-law tail distribution for any distribution of ϵ_i having a support S such that $\epsilon_a \in S$, as one can see from the divergence of the first moment of the absolute value of ψ_L . It is convenient to consider the distribution of the logarithm of ψ_L , whose moments are all finite. Let us remind the reader that we choose $\epsilon_i \in [-W/2, W/2]$, and in this calculation for simplicity we set $\epsilon_a = 0$. Defining

$$x_L = \ln|\psi_L|^2 = -2 \sum_i \ln(|\epsilon_i|/t) \quad (19)$$

we find that

$$\langle x_L \rangle = 2L \ln(2et/W). \quad (20)$$

The ratio $\langle x_L \rangle/L$ is the typical decay of wave function amplitudes from the origin of localization

$$\xi_{\text{typ}}^{-1} = 2 \ln(2et/W). \quad (21)$$

However, this is not the localization length ξ as defined in Eq. (14). The latter is indeed a uniform bound over the full set of $\sim K^L$ points at distance L from the localization center: it is determined by the decay rate of the maximal amplitude over sites in each shell at distance L . This is a point-to-set correlation function decay, which is familiar in the study of disordered systems on the Bethe lattice (on a regular lattice the point-to-set is substituted by the point-to-point, but with exponentially many shortest paths leading to the final point).

The typical value of the maximal amplitude x_L^* among K^L samplings is the largest solution of

$$x_L^* : K^L P(x_L^*) \simeq 1. \quad (22)$$

Here $P(x)$ is the distribution of x_L , and the K^L paths are treated as independent. For large L , what matters is the tail of

the distribution $P(x)$. If we rescale

$$z = \frac{x_L}{2L} + \ln(W/2t), \quad (23)$$

we get, for large z ,

$$P(z) \simeq \exp[-L(z - 1 - \ln z)]. \quad (24)$$

The probability distribution of z can be found inverting its Laplace transform, which is computable since z is a sum of i.i.d. variables. The maximum z^* over K^L samplings of z is the solution of

$$0 = -z^* + 1 + \ln z^* + \ln K = \ln(z^* e K e^{-z^*}), \quad (25)$$

and $x_L^* \simeq -2L \ln(W/2t) + 2Lz^*$. Localization occurs as long as $x_L^*/L < 0$, so the critical condition can be written as

$$\ln(K e z^* e^{-z^*}) = 0, \quad (26)$$

$$z^* - \ln(W_c/2t) = 0. \quad (27)$$

By eliminating z^* we recover the familiar [20] equation:

$$W_c = 2teK \ln(W_c/2t). \quad (28)$$

Moreover, z^* is given in terms of W_c as

$$z^* = \ln(W_c/2t), \quad (29)$$

so we can write the localization length as

$$\xi^{-1} = -x_L^*/L = 2 \ln(W/W_c). \quad (30)$$

This gives a mean-field exponent for the divergence of the localization length at the transition:

$$\xi \simeq \frac{2W_c}{|W - W_c|}. \quad (31)$$

Note that $\xi_{\text{typ}} < \xi$ irrespective of the value of the disorder. The belief expressed in Ref. [22] is that this behavior persists even within the delocalized regime, in the form of very irregular (multifractal) eigenstates. Note additionally that the difference between ξ_{typ} and ξ is due to the *exponential* sampling of the probability distribution, and it extends both to the finite-dimensional cube and to the many-body case.

B. d -dimensional cube

Consider now the case of a d -dimensional lattice of side L , and let a be the site at one corner of the cube, which we treat as the origin. Given any other site b , the orientation of the nonrepeating, shortest paths from a to b induces a natural orientation of the edges of the cube, which is thus directed; see Fig. 1. Let r be the lattice distance of the sites with respect to the origin a : for a cube of side L , the maximum r is $r_{\text{max}} = (L - 1)d$, corresponding to the site at the opposite corner of the cube with respect to a .

In an infinite cube the number of points at distance r from the origin is $(r + d - 1)!/[(d - 1)!r!]$, and thus it grows at most polynomially in r , slower than $\sim r^d$. Naively, one would be led to think that the transition is given by the divergence of ξ_{typ} ; see Eq. (21). However, the number of minimum length paths leading from the origin to an arbitrary site at lattice distance r scales exponentially with r , as $\sim d^r$. Therefore, unlike in the Bethe lattice case, in finite dimension the wave

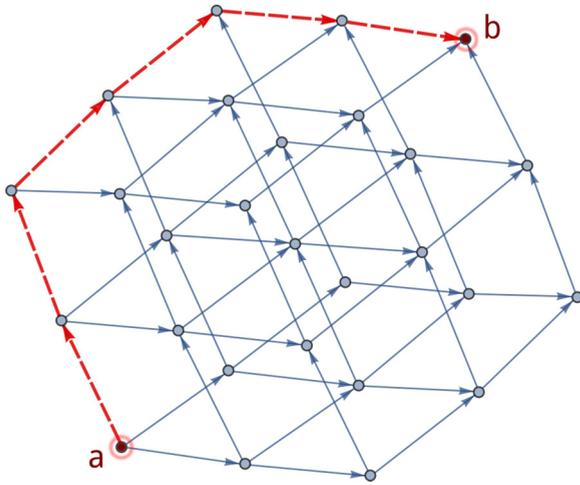


FIG. 1. Anderson model on a cube of side $L = 3$. The red, dashed edges form one of the nonrepeating paths connecting the sites a and b , of length $r = (L - 1)d = 6$. The other elements in the set $\text{paths}^*(a, b)$ are obtained following the arrows.

function amplitude at a given site b is a sum over exponentially many correlated terms

$$\psi_\alpha(b) = \sum_{p \in \text{paths}(a, b)} \prod_{i=1}^r \frac{t}{\epsilon_a - \epsilon_i} = \left(\frac{t}{W}\right)^r \psi'_\alpha(b), \quad (32)$$

where

$$\psi'_\alpha(b) = \sum_{p \in \text{paths}(a, b)} \prod_{i=1}^r \frac{1}{\epsilon'_a - \epsilon'_i} \quad (33)$$

and the random variables ϵ'_i are uniformly distributed in $[-1/2, 1/2]$. In the following, we consider the probability distribution of the random variable

$$Z_r \equiv \frac{\ln |\psi'_r|^2}{2r}, \quad (34)$$

for different values of r . Here ψ'_r denotes the maximum among all the rescaled amplitudes (33) at sites that are at lattice distance r with respect to the origin a . The probability of resonances for arbitrary values of t and W , see Eq. (15), is easily recovered from the cumulative distribution function of Z_r as

$$P\left(\frac{\ln |\psi_r|^2}{2r} \geq -\epsilon\right) = P\left(Z_r \geq \ln\left(\frac{W}{t}\right) - \epsilon\right), \quad (35)$$

with ϵ arbitrarily close to zero. According to Eqs. (14) and (15), the density of Z_r becomes asymptotically peaked at $\ln(W_c/t)$ for $r \rightarrow \infty$, with width going to zero with r . Thus, the critical value of disorder can be estimated inspecting the scaling with r of the probability density of Z_r .

The distribution of Z_r is hard to determine analytically, due to the correlation between the different shortest paths. To account for such correlations, we compute the amplitude (33) numerically by means of a transfer matrix technique, and use the resulting values to determine the probability (35) with ϵ smaller than the numerical precision. The convenience of the transfer matrix method relies on the fact that it takes only polynomial time in r , as was realized by Medina and Kardar

[21] in their treatment of the Nguyen, Spivak, and Shklovskii [25] (NSS) model [26].

The numerical computation is as follows. We fix $t = 1$ and introduce the matrix \mathcal{T} defined as

$$\mathcal{T} = \mathcal{W}A_f, \quad (36)$$

where A_f is the forward adjacency matrix of the lattice (that is, the adjacency matrix associated to the directed cube of Fig. 1), and \mathcal{W} is a diagonal matrix with components

$$\mathcal{W} = \text{diag}\left(\frac{1}{\epsilon'_a - \epsilon'_k}\right)_{k=1, \dots, L^d}. \quad (37)$$

We initialize the system in the state $|\psi^{(0)}\rangle = |a\rangle$ completely localized in the origin a , and iteratively apply the transfer matrix \mathcal{T} . A single iteration gives

$$|\psi^{(1)}\rangle \equiv \mathcal{T}|\psi^{(0)}\rangle = \frac{1}{\epsilon'_a - \epsilon'_{l_1}}|l_1\rangle + \frac{1}{\epsilon'_a - \epsilon'_{l_2}}|l_2\rangle + \dots, \quad (38)$$

where l_1, \dots, l_d are the forward neighbors of site a . The value of $\psi_\alpha(b)$ equals $\psi_\alpha(b) = \langle b | \psi^{(r)} \rangle$, where $|b\rangle$ is the state completely localized in the site b and r is the lattice distance between a and b .

We fix $\epsilon'_a = 0$ and compute the rescaled amplitude (33) for all the points b on a shell at the same lattice distance $r = r_{\max} - c$ from the origin of a hypercube of side L . Here $c \sim O(1)$ is fixed so as to have about 20 points per each size of the hypercube. We determine the maximal ψ'_r among the wave function amplitudes on those sites. We repeat the procedure for hypercubes of different sizes, with $O(10^5)$ disorder realizations for most system sizes, decreasing to $O(10^3)$ realizations only for the biggest system sizes that we consider (e.g., in $d = 3$ we take system sizes $r = 10$ through 292, with 1.5×10^5 disorder realizations up to $r = 202$ and 2.5×10^3 realizations up to $r = 292$).

As we discuss in Sec. IV, the main contribution to the transfer matrix result comes from only one of the exponentially many paths in Eq. (33), and the results obtained with the transfer matrix technique are faithfully reproduced by analyzing the statistics of the dominant path alone. The latter can be determined (see Sec. IV) with an algorithm that is computationally more efficient than the transfer matrix, allowing us to access much bigger system sizes. The results presented in this section for $d = 6, 7$, as well as for the higher values of r in $d = 3-5$, are obtained with this procedure.

1. Fluctuations of the wave function amplitudes

In Fig. 2 we plot the probability density of the variable Z_r defined in Eq. (34), for different values of r in $d = 3$. The plot shows a drift of the position of the peaks with increasing r , together with the shrinking of the width of the distribution, in agreement with the conditions (14) and (15). Plots of the r dependence of the variance $\sigma_{Z_r}^2$ of (34) are given in Fig. 3, in log-log scale for $d = 3-6$. The linear behavior indicates that the fluctuations of Z_r decay to zero as a power law in r , with a coefficient that depends on the dimensionality. The higher cumulants of the distribution exhibit a similar linear behavior in log-log scale. Moreover, the numerical computation indicates that for fixed d the probability densities

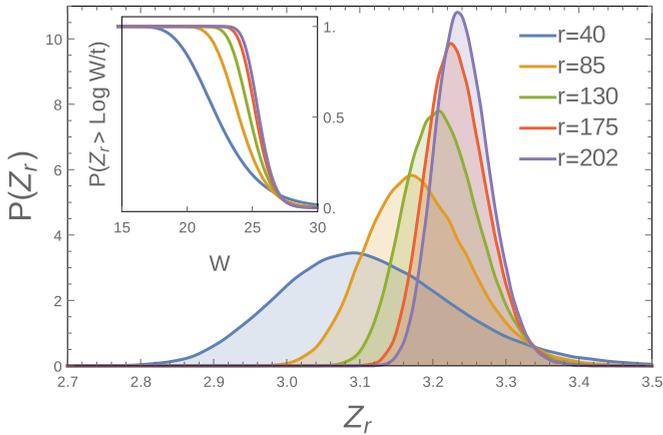


FIG. 2. Probability density of the variable Z_r defined in Eq. (34), for different r and $d = 3$. For $r \rightarrow \infty$, the curves become peaked around the critical value $\ln(W_c/t)$. Inset: Cumulative distribution function. Each curve is obtained with 1.5×10^5 disorder realizations. Very similar results are obtained for higher dimensionality.

of the variable

$$\tilde{Z}_r = \frac{Z_r - \langle Z_r \rangle}{\sigma_{Z_r}} \quad (39)$$

collapse to a limiting curve for increasing r ; see Fig. 4. As shown in the same plot, for fixed r and varying dimensionality, the distribution of \tilde{Z}_r does not change significantly, except for a weak d dependence of the tails.

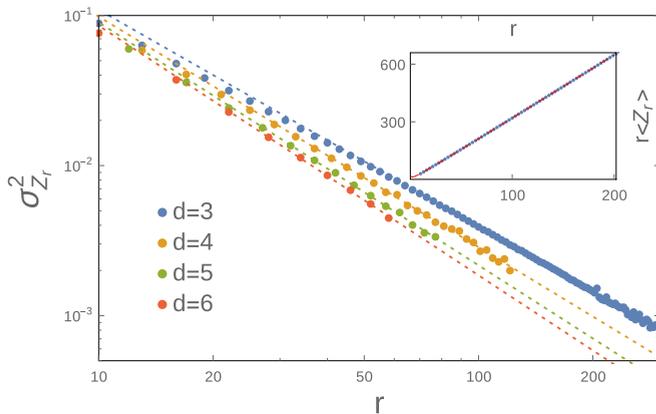


FIG. 3. Variance $\sigma_{Z_r}^2$ of the variable Z_r defined in Eq. (34). The plot is in log-log scale. The points corresponding to larger r are fitted linearly, according to the scaling form Eq. (40), and the values of the exponents $\omega_{FA}(d)$ reported in Table I are extracted from the coefficient of the linear term in the fit. The number of realizations is 1.5×10^5 for r smaller than 202, 53, 52, 40 for $d = 3, 4, 5$, and 6, respectively, and 2×10^3 for larger values of r . Inset: Mean value of the variable rZ_r defined in Eq. (34), for $d = 3$. The fit is linear with a correction $\propto r^{\omega_{FA}(3)}$, in agreement with the scaling form in Eq. (40), with the value of $\omega_{FA}(3)$ given in Table I. The results of the fit are, with reference to Eq. (42), $c_1 = -18.2 \pm 0.3$, $W_c = 27.03 \pm 0.02$, $c_2 = 29.6 \pm 0.8$. The same behavior holds for higher dimensionality and results in the estimates of the critical disorder values in Table II.

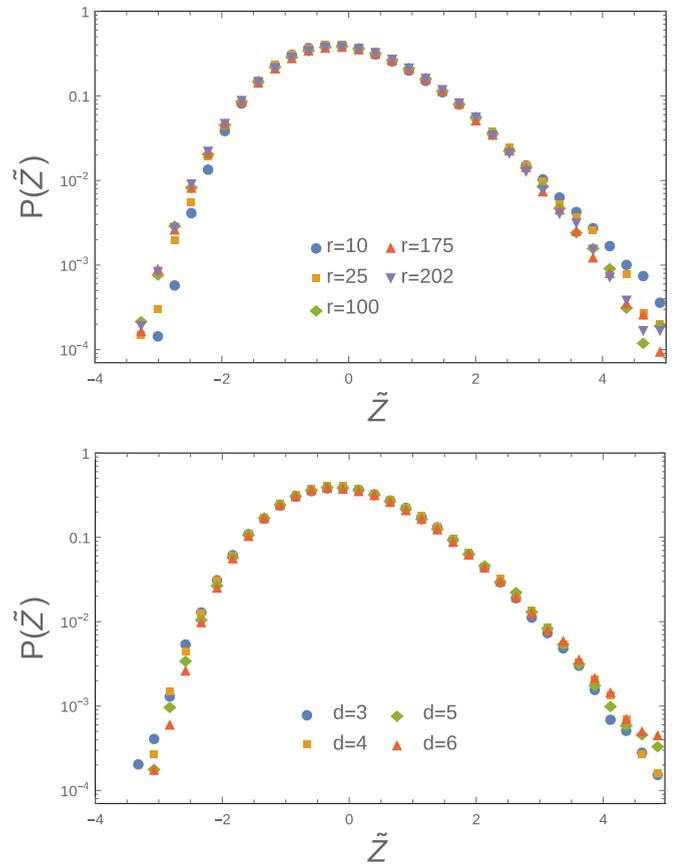


FIG. 4. Probability density $P(\tilde{Z})$ of the variable \tilde{Z}_r defined in Eq. (39). Each curve is obtained with 1.5×10^5 realizations. Top: Density of \tilde{Z}_r for different values r and $d = 3$. The curves seem to converge to a unique limiting distribution with increasing r . Bottom: Density of \tilde{Z}_r for fixed $r = 52$ and different dimensionalities.

These numerical observations are compatible with the following large r scaling form for Z_r :

$$r Z_r \underset{r \rightarrow \infty}{\sim} r \ln \left(\frac{W_c}{t} \right) + r^{\omega(d)} u, \quad (40)$$

where u is a random variable of $O(1)$ with a distribution which depends weakly on the dimensionality.

According to (40), for large r the fluctuations $\sigma_{Z_r}^2$ decay to zero with the power $r^{2[\omega(d)-1]}$. From the linear fit of $\ln(\sigma_{Z_r}^2)$ we extract the numerical estimate of the exponent in Eq. (40), which we denote with $\omega_{FA}(d)$. The results are reported in Table I.

In order to characterize the limiting distribution in Fig. 4, we compute the skewness $\text{Sk} = \kappa_3/\kappa_2^{3/2}$ and the kurtosis $\text{Kur} = \kappa_4/\kappa_2^2$ of the density of \tilde{Z}_r (here κ_i denotes the i th cumulant of the distribution). From (40) it follows that these parameters approach the ones corresponding to the variable u in the limit of large r . We restrict to $d = 3$, for which we have the largest statistics available. Plots of the r dependence of Sk and Kur are given in Fig. 5. The asymptotic values are estimated to be $\text{Sk} = 0.34 \pm 0.02$ and $\text{Kur} = 3.24 \pm 0.04$; see the caption of Fig. 5 for details.

TABLE I. Values of the exponent $\omega_{FA}(d)$ governing the decay of the fluctuations of Z_r with r ; see Eq. (40). A comparison is made with the values of the droplet exponents $\omega_{DP}(D)$ obtained numerically for the directed polymer in dimension $1 + (d - 1)$. The numerical values are taken from Appendix A in Ref. [27].

$d = D + 1$	$\omega_{FA}(d)$	$\omega_{DP}(D)$
3	0.278 ± 0.005	0.244
4	0.23 ± 0.01	0.186
5	0.191 ± 0.007	0.153
6	0.168 ± 0.006	0.130

2. Estimate of the critical disorder

To determine the critical value of disorder for $t = 1$, we extrapolate the asymptotic limit of the typical value of Z_r . Since the distribution is not fat tailed, we can equivalently consider the averages of Z_r and set

$$\langle Z_\infty \rangle \equiv \lim_{r \rightarrow \infty} \langle Z_r \rangle = \ln(W_c). \quad (41)$$

The inset in Fig. 3 shows the scaling with r of $r\langle Z_r \rangle$. The average grows linearly in r , in agreement with Eq. (40). We fit the data with the form

$$\langle r Z_r \rangle = c_1 + \ln(W_c)r + c_2 r^{\omega(d)}, \quad (42)$$

with the numerical values $\omega(d) = \omega_{FA}(d)$ reported in Table I.

The resulting estimates of the critical disorder, which we denote with W_c^{FA} , are displayed in Table II. For the smallest dimensions, a comparison is made with the critical values W_c^{num} determined in Refs. [28,29] by means of a combination of exact diagonalization and transfer matrix techniques.

The data in Table II clearly show that the FA gives an upper bound to the critical disorder, since the modification of the energy denominators provided by the (modified) self-energy corrections are neglected, and the effects of resonances are thus enhanced. However, increasing the dimensionality, the

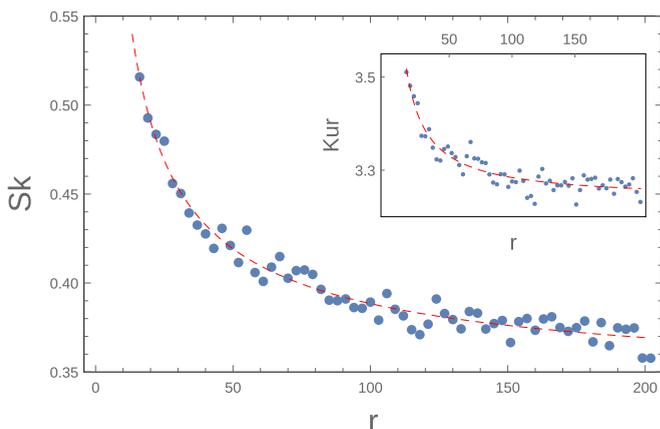


FIG. 5. Skewness Sk of the distribution of the variable \tilde{Z}_r defined in Eq. (39), for $d = 3$. The red dashed line is a fit of the form $\alpha + \beta r^\gamma$, with α, β, γ free parameters. The coefficient α is the estimate of the asymptotic value of the skewness, and it equals $Sk = 0.34 \pm 0.02$. Inset: Kurtosis Kur of the distribution of \tilde{Z}_r for $d = 3$, as a function of r . The fitting procedure is analogous to the one for the skewness, and results in $Kur = 3.24 \pm 0.04$.

TABLE II. Comparison between the critical value for localization in the Anderson model in d dimensions predicted by the forward approximation (W_c^{FA}) and the numerical results (W_c^{num}) of Ref. [28]. The relative error decreases faster than d^{-6} , presumably exponentially. For $d = 6$ the transition value $W_c^{FA} = 77.0 \pm 0.3$ can be compared with the result of Ref. [30], $W_c^{d=6} = 74.5 \pm 0.7$. This number is however an underestimation of the transition due to the choice of boundary conditions. For 7 dimensions there is no available numerics to compare with.

d	W_c^{FA}	W_c^{num}	Error
3	27.03 ± 0.03	16.536 ± 0.007	39%
4	41.4 ± 0.1	34.62 ± 0.03	16%
5	57.8 ± 0.2	57.30 ± 0.05	0.9%
6	77.0 ± 0.3		
7	93.8 ± 0.3		

discrepancy between the numerical estimates of W_c decreases; the enhanced precision of the FA result is due to the fact that the loops giving rise to the self-energy corrections become less relevant in higher dimensional lattices, and thus the FA becomes asymptotically exact in this limit.

3. Divergent length scales and critical exponents

For fixed values of W and for finite r , the probability of resonances (35) is determined by the tails of the distribution of Z_r .

For increasing r , the asymptotic limit is approached in a different way at the two sides of the transition: for $W > W_c$, the probability of resonances goes to zero exponentially with r . Below the transition, the probability converges to 1 much faster, with corrections that are only double exponential in r . We justify analytically this behavior in the Appendix by computing an approximate expression for the density of the variable Z_r . The approximation consists in considering the different paths contributing to it as independent variables.

Examples of the fits of the probability of resonances are shown in Fig. 6. To extract a W -dependent length scale $l(W)$, we perform an exponential fit of the form

$$P\left(\frac{\ln|\psi_r|^2}{2r} > 0\right) = a_1(W) \exp\left[-\frac{r}{l(W)}\right] \quad (43)$$

for $W > W_c$. For $W < W_c$ we determine $l(W)$ by means of the linear fit

$$\ln\left|\ln\left[1 - P\left(\frac{\ln|\psi_r|^2}{2r} > 0\right)\right]\right| = a_2(W) - \frac{r}{l(W)}. \quad (44)$$

The length scale $l(W)$ is plotted in Fig. 7 for $d = 3$. We expect it to diverge in the same way as the localization length/correlation length does in the localized/delocalized phase, respectively. We find that $l(W)$ diverges as a power law at a critical disorder compatible with the values of W_c^{FA} listed in Table II. A fit of the form $\ln[l(W)] = \ln c - \nu \ln|W - W_c^{FA}|$ results in an exponent that is compatible with $\nu \approx 1$ for all dimensions, consistently with the Bethe lattice picture and with the results in the Appendix. However, some deviations can be observed: a more careful analysis of the numerical data will be presented in a future publication.

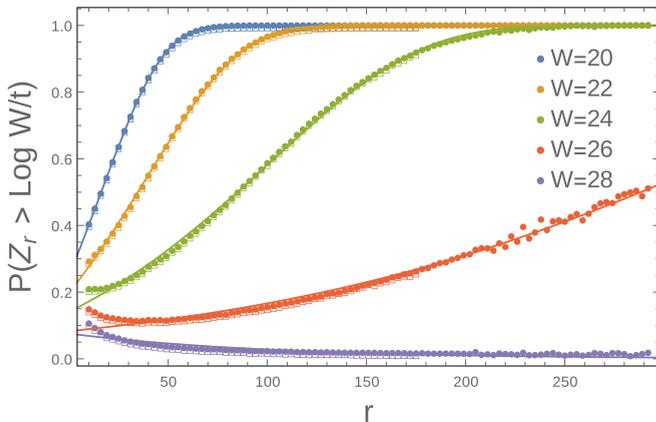


FIG. 6. Probability of resonances $P(Z_r > \ln(W/t))$ for the variable Z_r defined in Eq. (39) and $d = 3$. Asymptotically in r , the probability reaches 0 exponentially fast in the localized phase, and it reaches 1 double-exponentially fast in the delocalized phase, in agreement with the analytic computations in the Appendix. In the plot, the squares are the results of the transfer matrix calculation, the points of the dominating path (see Sec. IV), while the continuous lines are the exponential or double-exponential fits. Very similar results are obtained for higher dimensionality.

4. Connections with the problem of directed polymers in random medium

In the single-particle case, the energy denominators associated with different sites along the paths are independent variables. Thus, the expression for the wave function amplitude in FA, Eq. (12), resembles the expression for the partition function of a directed polymer (DP) in a random potential [31–33], with the thermal weights for the polymer configurations given by the amplitudes of the different paths. This analogy is not straightforward, due to the occurrence of negative

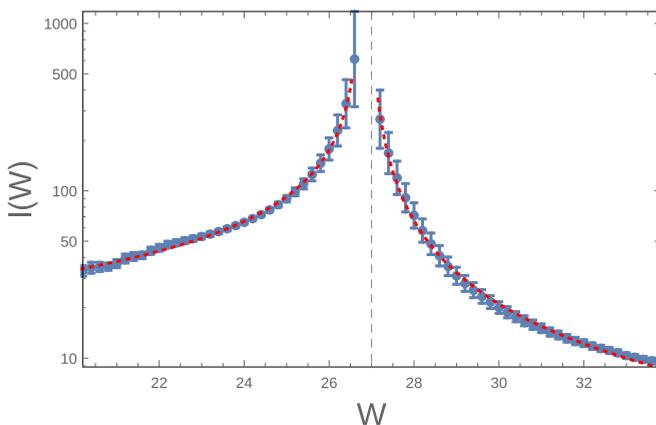


FIG. 7. Power-law divergence of the length scale $l(W)$ defined in Eqs. (43) and (44). The values of $l(W)$ for fixed W are determined from fits such as the ones in Fig. 6. The power-law fit produces a critical exponent $\nu \simeq 1$ and a critical value W_c compatible with the ones listed in Table II. The results shown here are for $d = 3$; very similar results are obtained for higher dimensionality. Notice how in the delocalized phase the distance to observe a resonance is typically larger (for the same $|W - W_c|$) than the localization length in the localized phase.

contributions in Eq. (12). Nevertheless, it has been fruitfully exploited both for the single-particle problem [34–37] and for problems of interacting spins on the Bethe lattice [38–41].

Motivated by this analogy, the authors of Ref. [35] have proposed a scaling form analogous to (40) for the logarithm $\ln g$ of the conductance of an Anderson model. There, the conductance in $d = 2$ is obtained from the Green's functions, which are computed numerically within a modified FA, the modification consisting of taking energy denominators that are not arbitrarily small but are bounded from below [42]. It is shown that the fluctuations of $\ln g$ scale with an exponent $\omega(d = 2) = 1/3$, and that the distribution of the variable u is compatible with a Tracy-Widom distribution.

These results are consistent with the conjecture [21,43] that in the strongly localized phase, where the expansion in nonrepeating paths is best controlled, the Anderson model in dimension d belongs to the same universality class of the directed polymer in dimension $1 + D$, with $D = d - 1$. In particular, the conjecture implies that in the limit of large r the distribution of $\ln g$ has the scaling form (40), with $\omega(d)$ coinciding with the droplet exponent [44] in $1 + (d - 1)$ dimensions (which is exactly known [45] to be equal to $1/3$ for $D = 1$), and u having the same distribution of the fluctuations of the free energy in the disordered phase of the polymer (distributed according to the Tracy-Widom distribution [46,47] in $D = 1$).

The values of the scaling exponents extracted from our data do not compare well with the droplet exponents $\omega(D = d - 1)$ of the DP; see Table I. Curiously, they compare within errors with $\omega(D + 1)$. We do not have an explanation for this curious behavior, and we leave its analysis for future work. Broadly speaking, the discrepancies with respect to the directed polymer results are generated by the fat tail of the distribution of the path amplitudes in Eq. (32), produced by the arbitrarily small energy denominators. It might be that the finite-size effects are more pronounced in the case of unbounded denominators. On the other hand, it is quite natural to expect that the models of nonrepeating paths with bounded amplitude considered in Ref. [35] exhibit a stronger dependence on the dimensionality, due to the fact that the domination by one single path is less pronounced in that case. We comment more on this point in Sec. IV.

C. Heisenberg model with random fields

In order to test the forward approximation on a many-body problem, we consider an XXZ spin-1/2 chain in random magnetic field,

$$H(t) = - \sum_{i=1}^L h_i s_i^z - \Delta \sum_{i=1}^L s_i^z s_{i+1}^z - t \sum_{i=1}^L (s_i^x s_{i+1}^x + s_i^y s_{i+1}^y), \quad (45)$$

where periodic boundary conditions are assumed ($s_1^\alpha = s_{L+1}^\alpha$), and the random fields h_i are uniformly distributed in $[-h, h]$. This spin Hamiltonian (45) has been studied in a large number of works [8,9,48–54], in which numerical evidence of the existence of a localization-delocalization transition is provided, mainly based on exact diagonalization results. The critical disorder is estimated [17] to be $h_c \simeq 3.72(6)$ for states

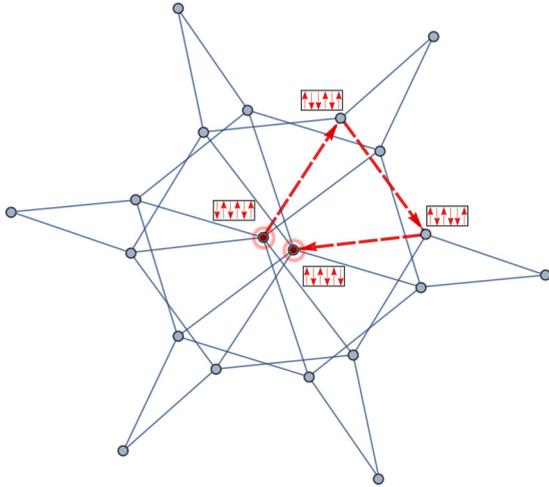


FIG. 8. Graph corresponding to the configuration space of the XXZ spin chain, see Eq. (45), of total length $L = 6$ and with periodic boundary conditions. Each site in the graph is associated with a product state in the basis of s_i^z operators. Only sites corresponding to states with zero total spin are represented. The initial Néel state $|\downarrow\uparrow\dots\rangle$ and the final, totally flipped state $|\uparrow\downarrow\dots\rangle$ are highlighted with circles. The red, dashed edges form one of the shortest paths connecting the two states, of length $L/2 = 3$.

in the middle of the energy band and parameters $t = 1$ and $\Delta = 1$.

As mentioned in the Introduction, the many-body problem can be seen as a single particle hopping problem in the “configuration space.” The latter is composed of the 2^L product states in the basis of s_i^z , which span the full Hilbert space and diagonalize $H(0)$. We denote these basis states with $|n\rangle$, and refer to them as the “computational basis.” The mapping to a hopping problem is obtained by interpreting each state $|n\rangle$ as a vertex n of a graph, with associated random energy E_n defined by $H(0)|n\rangle = E_n|n\rangle$. The third term in Eq. (45) provides the hopping between different sites, thus defining the geometry of the graph. Note that due to spin conservation, the full configuration space, and consequently the graph, are partitioned into disjoint sectors corresponding to different total spin; we restrict to the sector of total spin equal to zero, corresponding to a connected graph with $L!/[(L/2)!]^2$ vertices. See Fig. 8 for a pictorial representation of the graph for $L = 6$.

The effective hopping problem can be analyzed using the procedure set up in Sec. II A: the amplitude Ψ_α of an eigenstate of the effective single particle problem is given in forward approximation by

$$\Psi_\alpha(n_2) = \sum_{p \in \text{spaths}(n_1, n_2)} \prod_{n \in p} \frac{t}{E_{n_1} - E_n}, \quad (46)$$

where it is assumed that the eigenstate satisfies $\Psi_\alpha(n) \rightarrow \delta_{n, n_1}$ for $t \rightarrow 0$.

In the many-body language, Eq. (46) provides the expression of the coefficients of the eigenstates of (45) in the computational basis, to lowest order in the coupling t . The exponential decay of the coefficients implies localization in the configuration space, meaning that the full many-body

eigenstates are effectively a superposition of product states which differ only by configurations distant $O(1)$ from the initial configuration. The two main consequences of this structure of the wave function are that they have significantly less entanglement than ergodic states [9,50,55] and that, using Kubo’s formula for linear response [16], one can prove that they cannot support transport on macroscopic distances.

Similarly to the Anderson case, we fix an initial configuration of spins and we look at the amplitude in perturbation theory on the most distant, fully flipped configuration. In particular, we fix the localization center to be the site correspondent to the Néel state $|n_1\rangle = |\downarrow\uparrow\dots\rangle$, and consider the wave function amplitude on the site corresponding to the fully flipped Néel state $|n_2\rangle = |\uparrow\downarrow\dots\rangle$. These two sites n_1, n_2 are connected by $2(L/2)!$ paths on the graph, of length $r = L/2$ each.

By means of the transfer matrix we compute the rescaled amplitude

$$Z_r(h) \equiv \frac{\ln|\Psi_r|^2}{2r} \quad (47)$$

for different disorder strength h , with Ψ_r given by (46). We consider spin chains of size 6–20 with hopping and interaction constants respectively $t = 1$ and $\Delta = 1$, and $h = 1-6$. Note that despite that the general framework is the same as in the Anderson problem, the transfer matrix calculation is by no means identical; indeed, in the many-body case the energies associated with the different graph vertices are a linear combination of the independent random fields, and are thus correlated. Moreover, the number of paths connecting two sites proliferates with the size of the chain L , with a scaling that is faster than exponential. These paths present correlations that are much stronger with respect to the Anderson problem, as we shall discuss in more detail in Sec. IV.

1. Distribution of the wave function amplitudes and critical disorder

In Fig. 9 we show the probability density of $Z_r(h)$ for a chain of length $L = 20$ and different values of h . Since in the many-body case it is not possible to simplify the dependence

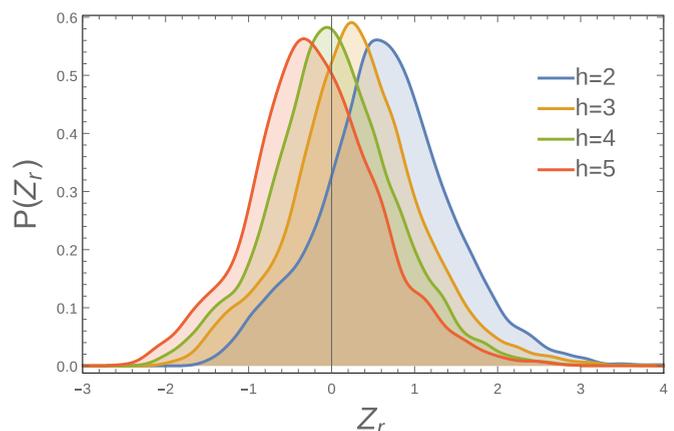


FIG. 9. Probability density of the random variable $Z_r(h)$ defined in Eq. (47), for an XXZ spin chain of length $L = 20$ (corresponding to $r = 10$) and different values of disorder h . Each curve is obtained with 3×10^3 realizations.

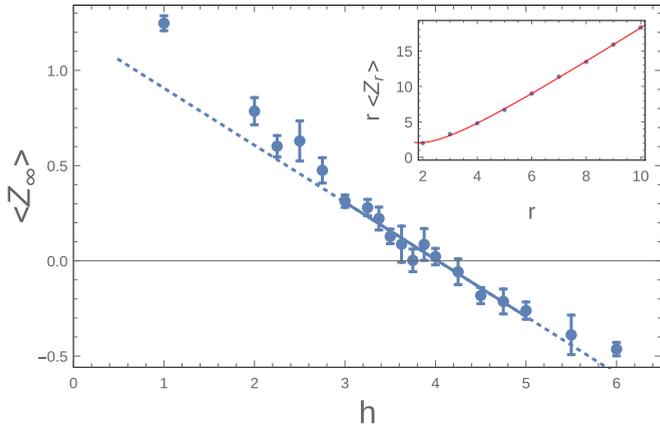


FIG. 10. Extrapolated value of the mean $\langle Z_\infty \rangle$ of the variable defined in Eq. (47). The crossing with 0 signals the many-body localization/delocalization transition for $t = 1$ [see Eq. (48)]. The error bars are obtained from the fitting procedure (see inset). The resulting transition value is $h_c = 4.0 \pm 0.3$. Inset: Finite-size scaling of $\langle Z_r \rangle$ with the distance r between the Néel states n_1 and n_2 . The plot corresponds to $h = 1$. The fit is linear with an r^{-1} correction, see Eq. (49), with parameters $c_1 = -7.2 \pm 0.4$, $\langle Z_\infty(2) \rangle = 1.23 \pm 0.02$, and $c_2 = 8.8 \pm 0.7$. The finite- r values for the mean are obtained over at least 10^4 realizations for $r < 7$ and at least 2×10^3 realizations for $r \geq 7$.

on the disorder strength h , the criterion for the transition reads

$$\langle Z_\infty(h_c) \rangle = -\ln t, \quad (48)$$

where $\langle Z_\infty(h) \rangle$ is the extrapolated value of the average of (47) for fixed h . Plots of $\langle Z_\infty(h) \rangle$ are given in Fig. 10. Here $\langle Z_\infty(h) \rangle$ is extrapolated from the finite-size values using the fitting function

$$r \langle Z_r(h) \rangle = c_1 + \langle Z_\infty(h) \rangle r + c_2 r^{-1}. \quad (49)$$

For $t = 1$, the critical point h_c is estimated from the condition $\langle Z_\infty(h_c) \rangle = 0$. The resulting value is $h_c = 4.0 \pm 0.3$, which is, as expected, larger than the result derived with exact diagonalization. Notice also that the corrections $\propto r^{-1}$ are consistent with the intuition that the $\omega_{FA} = 0$ is the correct mean-field scaling for MBL.

As we discuss in Sec. IV, in the many-body case the sum (46) is no longer dominated by a single path. Thus, the algorithm for the best path is not applicable in this context, and the limited system sizes accessible with the transfer matrix do not allow us to investigate whether a scaling form exists also for (47) in the limit of large r . For the available system sizes, the distributions of the rescaled variables $\tilde{Z}_r(h) = [Z_r(h) - \langle Z_r(h) \rangle] / \sigma_{Z_r(h)}$ do not seem to collapse to a unique curve, and the scaling of the variances $\sigma_{Z_r(h)}^2$ with r appears to be compatible with a power law, but with exponent depending on the disorder strength h . However, a more refined numerical analysis is necessary to draw a conclusion on the asymptotic behavior.

2. Divergent length scales and critical exponents

Figure 11 shows the behavior of the probability of resonances $P(Z_r(h) > -\ln t)$ as function of the distance between the Néel states. As expected, the r dependence changes

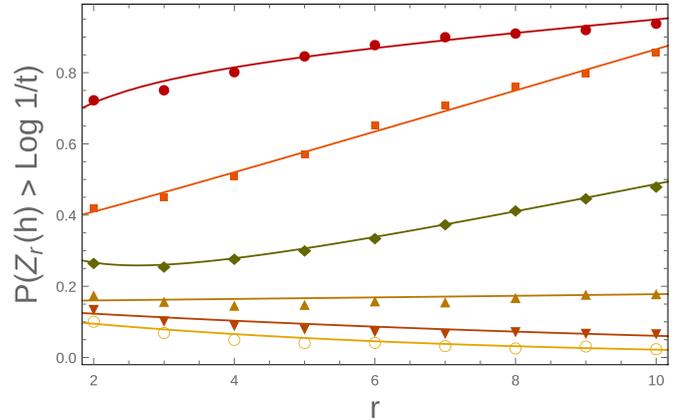


FIG. 11. Probability of resonances $P(Z_r(h) > -\ln t)$ as a function of the distance r between the two Néel states n_1 and n_2 , for $t = 1$. Asymptotically, the probability reaches zero exponentially in the localized phase and 1 in the delocalized phase. We average over 10^4 , 5×10^3 , and 3×10^3 realizations for $r \leq 8$, $r = 9$, and $r = 10$, respectively; the plotted values of the disorder h are $h = 1$ (points), $h = 2$ (squares), $h = 3$ (diamonds), $h = 4$ (upward triangle), $h = 5$ (downward triangle), and $h = 6$ (circle). Linear and exponential fits in the delocalized and localized regions respectively are plotted as continuous lines; see Eqs. (50) and (51).

with the disorder: the probability decays to zero at large h , and increases towards 1 for the smaller h . We expect the convergence to be exponential in r on both sides of the transition; however, the exponential behavior is not clearly detectable in the delocalized phase, due to the few accessible system sizes. For $h < h_c$ we extract a length scale $l(h)$ by fitting the curves in Fig. 11 with the function

$$P(Z_r(h) > 0) = a_2(h) + \frac{r}{l(h)} + \frac{b(h)}{r}. \quad (50)$$

In the localized phase we perform instead the exponential fit:

$$P(Z_r(h) > -\ln t) = a_1(h) \exp\left(-\frac{r}{l(h)}\right). \quad (51)$$

The length scales $l(h)$ extracted with this procedure are shown in Fig. 12, together with the power-law fit $l(h) = c|h - h_c|^{-\nu}$. The fit is performed separately for $h < h_c$ and $h > h_c$, resulting in an exponent close to 1 in both cases (see Fig. 12 for details). Note the asymmetry of the curve with respect to h_c , which indicates that at fixed $|h - h_c|$ the typical distance to find a resonance in the delocalized phase is larger than the localization length at the corresponding value of disorder in the localized phase. A possible consequence of this phenomenon, which occurs also in the Anderson model (see Fig. 7), could be a large “critical region” in the dynamics in the delocalized phase.

IV. THE STRUCTURE OF THE DOMINATING PATHS

As mentioned in Sec. III B, the wave function amplitudes in FA can be interpreted as the partition function for a directed polymer in random medium. However, a relevant difference is that while the weight associated with the polymer is bounded from above [32], the single factors in Eq. (33)

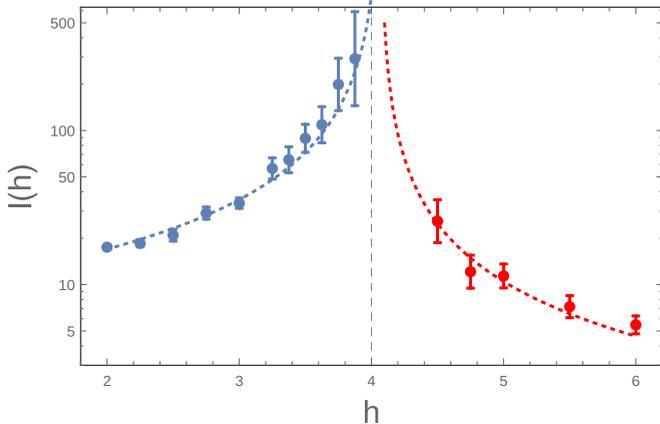


FIG. 12. Divergence of the length scales $l(h)$ extracted from the fits of the probability of resonances as a function of r . The vertical dashed line indicates the critical value h_c obtained in Fig. 10. The dotted curve is a power-law fit of the form $c|h - h_c|^{-\nu}$, resulting in a critical exponent $\nu_L = 1.12 \pm 0.06$ for $h < h_c$ and $\nu_R = 1.1 \pm 0.2$ for $h > h_c$.

are unbounded, with diverging average. As a consequence, the strongly localized phase in the Anderson model (where the FA is best controlled) always corresponds to a “frozen” phase of the directed polymer, in which most of the weight in the total sum (12) is given by one single path. An interesting question [56] is whether the freezing phenomenon persists in the delocalized phase in some form, and what this implies for the structure of the eigenstates close to the transition [22,57].

In this section, we compare the statistics of the wave function amplitudes in FA with that of the optimal path, i.e., the path with maximal amplitude, both for the single-particle problem and for the XXZ chain. We show that while in the first case the full sum is strongly dominated by the extremal path amplitude, in the many-body case most of the paths have comparable amplitude. However, the much stronger correlation between them gives rise to nonnegligible interference effects, resulting in strong cancellations.

A. The single-particle case

For the finite-dimensional case, we compute the amplitude ω_r^* of the optimal path p^* dominating the sum (12) by means of the Dijkstra algorithm [58], a graph-search algorithm that determines the path minimizing a given cost function. We consider a directed cube (such as the one in Fig. 1) and assign a positive cost χ to each directed edge $\langle i, j \rangle$:

$$\chi(i, j) \equiv \ln|\epsilon_j| - \min_k \{\ln|\epsilon_k|\}. \quad (52)$$

The total cost of a path p is the sum of the costs of the edges belonging to it, and the path p^* with maximal amplitude is the one minimizing the total cost function. In order to compare with the transfer matrix results, we compute the ratio between ω_r^* and the full sum (32) computed via the transfer matrix technique, for the same given disorder realization. The distribution of the ratios turns out to be very narrowly peaked around 1. Figure 13 displays its average as a function of the length of the paths r for $d = 3$, which is extremely close to 1, uniformly in the path length.

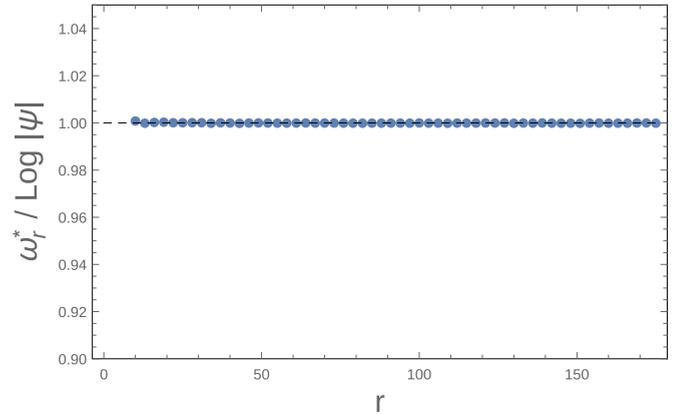


FIG. 13. Average ratio between the dominating path weight ω_r^* computed with the Dijkstra algorithm, and the sum in Eq. (32) computed using a transfer matrix technique. The ratio is taken between values corresponding to the same disorder realization. The plot corresponds to $d = 3$, and each point is averaged over 3×10^4 disorder realizations. Similar results are obtained for higher dimensionality, for those r accessible with the transfer matrix technique. The standard deviation error bars are within the point size.

As a further check of the agreement between the values computed with the two methods, we plot in Fig. 6 the r dependence of the probability (15) with $\delta = 0$, determined with the substitution $|\psi_r| \rightarrow \omega_r^*$. The data are plotted as points, which are almost indistinguishable from the transfer matrix results (squares). This indicates that the statistics of distant resonances is fully captured by the optimal path. Thus, in the single-particle case the correlation between different paths does not play a relevant role, since the sum is dominated by the extremum, as would happen for independent random variables with fat-tailed distribution. Based on this observation, the numerical analysis outlined in the previous sections can be carried out for much bigger system sizes with respect to the ones accessible with the transfer matrix technique, since the Dijkstra algorithm has lower complexity than the transfer matrix (indeed, the time complexity is $\sim O(e + v \ln v)$ and the space complexity is $\sim O(v^2)$, where v is the number of vertices and e is the number of graph edges).

We perform the same analysis also for the modified forward approximation discussed in Ref. [35], by taking the energy denominators uniformly distributed in $[-1, -W^{-1}] \cup [W^{-1}, 1]$ in $d = 3$ for two values of the cutoff, $W = 25$ and $W = 35$. We find that in this case the ratio between the maximal path and the transfer matrix result departs from one for increasing r . This suggests that more than one path dominates the transfer matrix result. It is natural to expect that in this case the number of dominating paths depends on the geometry of the system, thus introducing a stronger dependence on the dimensionality; see also the comments in Sec. III B.

For the case of unbounded denominators, we compute the inverse participation ratio (IPR) of the edge weights contributing to ω_r^* , for $\epsilon_i \in [-1, 1]$ (i.e., $W = 2$). We define

$$\text{IPR} = \frac{(\sum_i \ln|\epsilon_i|)^2}{\sum_i (\ln|\epsilon_i|)^2}, \quad (53)$$

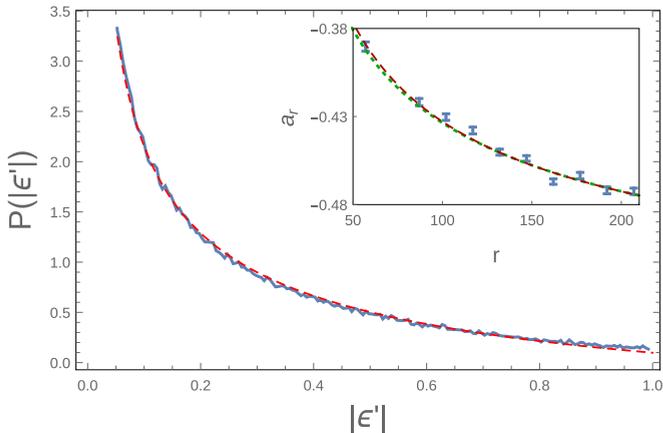


FIG. 14. Probability distribution $\rho(\epsilon)$ of the energy denominators along the optimal path; see Eq. (55). The plot corresponds to $d = 3$ and $r = 210$. The dashed red line is the fitting function of Eq. (54), with fitting parameters $c_r = -0.95 \pm 0.04$, $b_r = 1.04 \pm 0.03$, and $a_r = 0.472 \pm 0.005$. Very similar results are obtained for higher dimensionality. Inset: Plot of the exponents a_r in the fitting function of Eq. (54), as a function of r . Due to the absence of a theoretical reasoning for the finite-size scaling, we fit the curve considering logarithmic and $1/\sqrt{r}$ corrections. The green small-dashed curve is a fitting function of the form $a + c/\ln(r)$, with fit parameters $a = -0.73 \pm 0.05$ and $c = -1.4 \pm 0.3$; the red large-dashed curve is a fitting function of the form $a + c/\sqrt{r}$, with fit parameters $a = -0.57 \pm 0.02$ and $c = -1.4 \pm 0.3$. The asymptotic value a obtained with the logarithmic fitting function is compatible with the solution of Eq. (56) for $d = 3$.

where i labels the sites belonging to the optimal path p^* . We find that the disorder-averaged IPR scales linearly with the length of the path r , indicating that an extensive (in r) number of edges contributes to the total path weight, and cooperate to produce the atypically big path weights dominating (32). Figure 14 shows the distribution of the absolute value of the energies along the optimal path for $W = 2$, $d = 3$, $r = 210$, and $\epsilon_a = 0$. The fitting function has the form

$$\rho_r(\epsilon) = c_r + b_r |\epsilon|^{a_r}. \quad (54)$$

The power-law behavior is consistent with the considerations in Ref. [41]. Adapting their reasoning to the finite dimensional case, one can argue that asymptotically in r (and under the hypothesis of independent paths) the biased energy distribution along the optimal path has the form

$$\rho(\epsilon) = \frac{1 - 2x}{|\epsilon|^{2x}}, \quad (55)$$

with x solving the d -dependent equation

$$\ln\left(\frac{d}{1 - 2x}\right) - \frac{2x}{1 - 2x} = 0. \quad (56)$$

Fitting the r dependence of the coefficients c_r, b_r, a_r one finds that the asymptotic limits are in agreement with (55); for details see the inset of Fig. 14.

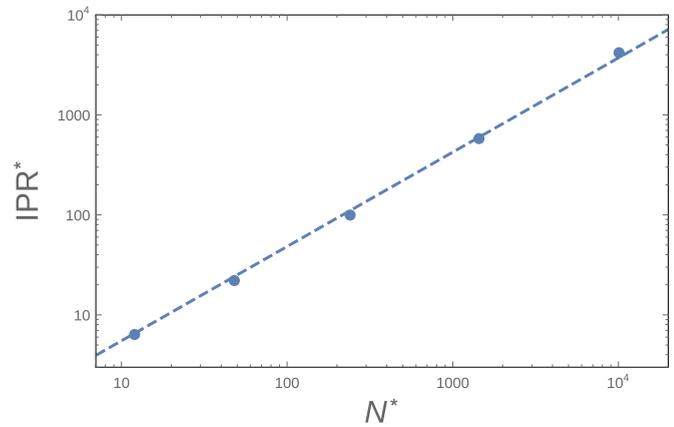


FIG. 15. Average IPR* of the paths, Eq. (57), as a function of the number of paths $N^* = 2(L/2)!$ for random spin chains of different lengths L . The IPR* is linear in the total number of paths.

B. The many-body case

When performing the same type of analysis for the Heisenberg chain, we find that the statistics of the sum (46) is not well reproduced by the optimal path alone: the distribution of the ratios between the full sum and the optimal path is very wide and peaked at values that are far from 1. Thus, despite that also in this case the amplitudes of the single paths are fat-tail distributed, there is not a single one dominating. Instead, we find that the average IPR* of the path amplitudes (which we denote with ω_p)

$$\text{IPR}^* = \frac{(\sum_p \omega_p)^2}{\sum_p \omega_p^2} \quad (57)$$

scales linearly with the total number of paths $N^* = 2(L/2)!$, indicating that there are factorially many (in the length of the chain L) paths having amplitudes that are comparable in absolute value (see Fig. 15). This is a signature of the strong correlations between the paths, which is not surprising in view of the many-body nature of the model. Following Ref. [16], one can argue that the strongest correlations are among those paths associated with processes in which the same spin flips occur, but in different order: the different orderings of the flips produce different energy denominators in Eq. (46), and thus different path amplitudes; however, the resulting terms are correlated, and one can expect that for those realizations of the random fields producing one particularly large path weight, the other ones (related to it by permutation of the order of the spin flips) will also have a large amplitude in absolute value. However, in the sum (46) the paths contribute with well defined relative signs, leading to cancellations between these factorially many terms (see Ref. [16] for an explicit calculation for a model of interacting fermions), which are fully taken into account only with the transfer matrix method.

V. CONCLUSION

In this work we have discussed the advantages and the limitations of the forward approximation applied to both single- and many-body quantum disordered systems. In particular, the FA has been used to obtain an expression for the wave

functions, that can be computed by means of a transfer matrix technique. The statistical analysis of the wave functions allows us to determine the critical values of the disorder (very accurate in large d), the critical exponents of the localization length (which are compatible with mean field), and the universal distribution of the eigenfunctions' coefficients.

For the single-particle case, the amplitudes of the wave functions in FA turned out to be very well approximated by only one path, the dominating path: this has been exploited to investigate larger system sizes with respect to the ones accessible with the transfer matrix (the algorithm that computes the best path runs in time linear in the number of edges of the underlying graph, and is not as memory-demanding as the transfer matrix of shift-invert exact diagonalization). The extremely good agreement (within statistical error) of the predicted critical values for the disorder for $d \geq 5$ suggests that the approximation should be predictive also for the properties of the wave functions at the critical point. We have not investigated in details these implications, but we foresee the wave functions to have a sparse structure which is similar to that discussed in high-coordination Bethe lattices [22]. The strong similarities with the problem of directed polymers in random medium have also been addressed; however, from the statistical analysis it emerges that the scaling exponents describing the fluctuations of the wave functions are non-mean-field, but also not equal to those of the directed polymer. Moreover, the limiting distribution of the appropriately rescaled wave functions seems to depend more weakly on the dimensionality with respect to the directed polymer case.

In the many-body problem there is no concentration of amplitude on a small number of paths, but there are strong cancellations between them: as a result, the full sum over factorially many paths is only exponentially large (or small) in the system size. The correlation between the paths has been discussed in detail in Ref. [16], and this work can be interpreted as a numerical test of the claims in that work. For the XXZ chain with random fields, the critical value predicted within the approximation ($h_c = 4.0 \pm 0.3$) is very close to the most updated result obtained with exact diagonalization; thus, this tool furnishes an alternative route to exact diagonalization, that can be applied to significantly larger system sizes. We leave to future work the question of how to incorporate higher-order corrections in the FA within the transfer matrix scheme, and how they affect the critical exponents, and the accuracy of the results.

As a concluding remark, we would like to briefly comment on the nature of the FA as a mean-field approximation for Anderson localization. For certain, the value of the critical disorder W_c for the onset of localization grows indefinitely with d , and in high d the hopping t becomes an almost negligible perturbation at the transition. The fact that the error in W_c essentially disappears around $d \simeq 6$ is a strong indication of this. This feature is quite peculiar, since in ordinary, second-order phase transitions the critical exponents above the upper critical dimension are correctly reproduced by the mean-field approximation, but the location of the transition (e.g., the critical temperature) is not. In this sense the locator expansion (that to lowest order reduces to the FA considered in this work) becomes a better suited candidate for a mean

field than the $2 + \epsilon$ expansion of the nonlinear supersymmetric sigma model (NLS σ M) [59]. It is plausible that there is a field-theoretical description of the FA which can be put in direct relation with the NLS σ M; this is an obvious direction in which to continue this work. In addition, the relation with the Bethe lattice results can be further investigated, given that the critical W_c predicted by the FA does not correspond to that of a Bethe lattice of any (integer) coordination number.

ACKNOWLEDGMENTS

A.S. would like to thank B. L. Altshuler, R. Moessner, V. Oganesyan, and S. L. Sondhi for discussions.

APPENDIX: PROBABILITY DENSITY OF Z_r

An estimate for the probability density of Z_r can be obtained making use of the fact that the sum over (33) is dominated by the single path with maximal weight:

$$Z_r \approx \max_{p \in \text{paths}} \left\{ -\frac{1}{r} \sum_{i \in p} \ln |\epsilon'_i| \right\}. \quad (\text{A1})$$

If the correlations between the different path weights are neglected, the calculation is similar to the one performed for the Bethe lattice case. In particular, one finds for the cumulative function of Z_r the following expression:

$$P(Z_r < a) = \exp \left[N_r \ln \left(1 - \frac{1}{(r-1)!} \int_{r(a-\ln 2)}^{\infty} t^{r-1} e^{-t} dt \right) \right], \quad (\text{A2})$$

where $N_r \sim d^r$ is the total number of paths on which the maximum is taken. This implies the following form for the probability density of $Z'_r = Z_r - \ln 2$:

$$p_r(z') = \frac{N_r r^r}{(r-1)!} \frac{e^{-r(z' - \ln z')}}{z'^r [1 - I_r(z')]} \exp\{N_r \ln[1 - I_r(z')]\}, \quad (\text{A3})$$

where we introduced the monotonic decreasing function

$$I_r(z') = \frac{1}{(r-1)!} \int_{z'}^{\infty} t^{r-1} e^{-t} dt. \quad (\text{A4})$$

The typical value of Z'_r , denoted z_r^* , is defined by the equation

$$N_r I_r(z_r^*) = N_r \frac{r^{r-1}}{(r-1)!} \int_{z_r^*}^{\infty} t^{r-1} e^{-t} dt = 1. \quad (\text{A5})$$

The solutions of (A5) approach a finite limit z^* for $r \rightarrow \infty$, which is related to the critical value of disorder by $z^* = \ln[W_c/(2t)]$, as previously discussed. Using that $N_r \sim d^r$ and computing the integral in Eq. (A5) with a saddle point calculation (assuming $z^* > 1$), one recovers the condition (28) for W_c , with the substitution $K \rightarrow d$.

For increasing r the probability density of Z'_r peaks at the typical value, with tails that approach zero in the limit $r \rightarrow \infty$. In particular, for $z' > z_r^*$ the decay of the tail is exponential in r . Indeed, in this regime the product $N_r I_r(z)$ is itself exponentially decreasing with r ; thus, the rightmost

exponential in Eq. (A3) rapidly converges to 1, and the distribution $p_r(z')$ approaches zero with a tail of the form

$$p_r(z') \sim e^{-r[z' - \ln z' - \ln(de)] + o(r)}. \quad (\text{A6})$$

When z' becomes smaller than the typical value z_r^* , the product $N_r I_r(z')$ increases exponentially. Since for large r the integral $I_r(z')$ is still exponentially small for all $z' > 1 + O(1/r)$, one can still set $\ln[1 - I_r(z')] \sim -I_r(z') \sim \exp[-rz' + r \ln(ez') + o(r)]$. Thus, in this regime the probability density of Z'_r decays to zero much faster, double exponentially with r :

$$p_r(z') \sim \exp(-d^r e^{-rz' + r \ln(ez')} + O(r)). \quad (\text{A7})$$

Note that (for r large enough) the interval in which $1 < z' < z_r^*$ does not shrink to zero for $d \geq 3$, given that the value z_r^* obtained from the condition (A5) is always bigger than 1. When z' approaches one, the probability in Eq. (A4) is no longer a large deviation probability; i.e., it is no longer exponentially small in r : the term $\ln[1 - I_r(z')]$ approaches a constant function of z' , and the main scaling is given by the factor d^r .

Finally, exactly at $z' = z_r^*$, using (A5) and performing the integral with an integration by parts, one finds that the probability density can be written as

$$p_r(z_r^*) = \frac{r}{1 - d^{-r}} \left[1 + \sum_{n=1}^{r-1} \frac{(r-1)!}{(r-1-n)! r^n} (z_r^*)^{-n} \right]^{-1} \times \exp\left(-1 - \frac{1}{2d^r} - \frac{1}{3d^{2r}} + \dots\right), \quad (\text{A8})$$

which diverges like r when $r \rightarrow \infty$.

Given the tails of the distribution of Z_r computed in this approximation, it is immediate to derive the asymptotic decay of the probability of resonances in the localized phase. Indeed, for $W > W_c$, the probability (35) is a large deviation for Z_r . Making use of (A6) we find

$$P\left(Z_r > \ln \frac{W}{t}\right) = \int_{\ln(\frac{W}{2t})}^{\infty} e^{-r[z' - \ln z' - \ln(de)] + o(r)} dz' = \exp\left(-\frac{r}{l(W)} + o(r)\right) \quad (\text{A9})$$

with

$$\frac{1}{l(W)} = \ln\left(\frac{W}{2tde \ln(W/2t)}\right). \quad (\text{A10})$$

Thus, within this approximation for W approaching W_c from above one finds

$$l(W) \sim \frac{W_c}{W - W_c}, \quad (\text{A11})$$

and thus the length scale diverges at the transition with a critical exponent equal to 1.

Similarly, for $2te < W < W_c$, making use of (A2) and of (A7) we find

$$P\left(Z_r < \ln\left(\frac{W}{t}\right)\right) \approx \exp\left(-\left[\frac{2tde}{W} \ln\left(\frac{W}{2t}\right)\right]^r + O(r)\right) = \exp(-e^{-r/l(W)} + O(r)). \quad (\text{A12})$$

-
- [1] E. Abrahams, *50 Years of Anderson Localization* (World Scientific, Los Angeles, 2010).
- [2] 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).
- [3] 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).
- [4] P. Anderson, *Phys. Rev.* **109**, 1492 (1958).
- [5] F. Evers and A. D. Mirlin, *Rev. Mod. Phys.* **80**, 1355 (2008).
- [6] D. Basko, I. Aleiner, and B. Altshuler, *Ann. Phys.* **321**, 1126 (2006).
- [7] V. Oganesyan and D. A. Huse, *Phys. Rev. B* **75**, 155111 (2007).
- [8] A. Pal and D. A. Huse, *Phys. Rev. B* **82**, 174411 (2010).
- [9] A. De Luca and A. Scardicchio, *Europhys. Lett.* **101**, 37003 (2013).
- [10] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. Lett.* **111**, 127201 (2013).
- [11] J. Z. Imbrie, [arXiv:1403.7837](https://arxiv.org/abs/1403.7837).
- [12] J. A. Kjäll, J. H. Bardarson, and F. Pollmann, *Phys. Rev. Lett.* **113**, 107204 (2014).
- [13] D. A. Huse, R. Nandkishore, and V. Oganesyan, *Phys. Rev. B* **90**, 174202 (2014).
- [14] C. R. Laumann, A. Pal, and A. Scardicchio, *Phys. Rev. Lett.* **113**, 200405 (2014).
- [15] A. Chandran, I. H. Kim, G. Vidal, and D. A. Abanin, *Phys. Rev. B* **91**, 085425 (2015).
- [16] V. Ros, M. Müller, and A. Scardicchio, *Nucl. Phys. B* **891**, 420 (2015).
- [17] D. J. Luitz, N. Laflorencie, and F. Alet, *Phys. Rev. B* **91**, 081103 (2015).
- [18] B. L. Altshuler, Y. Gefen, A. Kamenev, and L. S. Levitov, *Phys. Rev. Lett.* **78**, 2803 (1997).
- [19] D. A. Huse, R. Nandkishore, F. Pietracaprina, V. Ros, and A. Scardicchio, *Phys. Rev. B* **92**, 014203 (2015).
- [20] R. Abou-Chacra, D. Thouless, and P. Anderson, *J. Phys. C* **6**, 1734 (1973).
- [21] E. Medina and M. Kardar, *Phys. Rev. B* **46**, 9984 (1992).
- [22] A. De Luca, B. L. Altshuler, V. E. Kravtsov, and A. Scardicchio, *Phys. Rev. Lett.* **113**, 046806 (2014).
- [23] A similar argument gives a perturbative series in t for the conductance.
- [24] R. Abou-Chacra and D. Thouless, *J. Phys. C* **7**, 65 (1974).
- [25] V. Nguyen, B. Spivak, and B. Shklovskii, *Pis'ma Zh. Eksp. Teor. Fiz.* **41**, 35 (1985) [*JETP Lett.* **41**, 42 (1985)]; *Zh. Eksp. Teor.*

- Fiz. **89**, 11 (1985) [Sov. Phys. JETP **62**, 1021 (1985)]; Pis'ma Zh. Eksp. Teor. Fiz. **43**, 35 (1986) [JETP Lett. **43**, 44 (1986)].
- [26] Note that the present calculation does not reduce to the one for the NSS case, since a major difference between the two models is in the statistics of energy denominators: in the NSS model the binary disorder $\epsilon_i = \pm W$ (with probability p or $1 - p$) does not allow for resonances due to a single site. Rather, the resonances arise from contributions of different paths. This led to a body of work following Ref. [21], on the presence of a sign transition, where effects from different paths accumulate in order to break the sign symmetry. In the FA for the Anderson case, the energy denominators can be arbitrarily small with finite probability, generating path weights that are fat-tail distributed. The fat tail of this distribution is crucial for the considerations below.
- [27] C. Monthus and T. Garel, *J. Phys. A: Math. Theor.* **45**, 095002 (2012).
- [28] K. Slevin and T. Ohtsuki, *New J. Phys.* **16**, 015012 (2014).
- [29] Y. Ueoka and K. Slevin, *J. Phys. Soc. Jpn.* **83**, 084711 (2014).
- [30] A. M. García-García and E. Cuevas, *Phys. Rev. B* **75**, 174203 (2007).
- [31] B. Derrida, *Phys. A (Amsterdam, Neth.)* **163**, 71 (1990).
- [32] B. Derrida and H. Spohn, *J. Stat. Phys.* **51**, 817 (1988).
- [33] T. Halpin-Healy and Y.-C. Zhang, *Phys. Rep.* **254**, 215 (1995).
- [34] A. M. Somoza, M. Ortuño, and J. Prior, *Phys. Rev. Lett.* **99**, 116602 (2007).
- [35] J. Prior, A. M. Somoza, and M. Ortuño, *Eur. Phys. J. B* **70**, 513 (2009).
- [36] A. M. Somoza, P. Le Doussal, and M. Ortuño, *Phys. Rev. B* **91**, 155413 (2015).
- [37] C. Monthus and T. Garel, *Phys. Rev. B* **80**, 024203 (2009).
- [38] M. V. Feigel'man, L. B. Ioffe, and M. Mézard, *Phys. Rev. B* **82**, 184534 (2010).
- [39] L. B. Ioffe and M. Mézard, *Phys. Rev. Lett.* **105**, 037001 (2010).
- [40] M. Müller, *Europhys. Lett.* **102**, 67008 (2013).
- [41] X. Yu and M. Müller, *Ann. Phys.* **337**, 55 (2013).
- [42] Note that if this constraint is relaxed and the energy denominators are allowed to be arbitrarily small, $\log g$ is found to be proportional to the quantity $r Z_r$ that we are considering.
- [43] E. Medina, M. Kardar, Y. Shapir, and X. R. Wang, *Phys. Rev. Lett.* **62**, 941 (1989).
- [44] D. S. Fisher and D. A. Huse, *Phys. Rev. B* **43**, 10728 (1991).
- [45] D. A. Huse, C. L. Henley, and D. S. Fisher, *Phys. Rev. Lett.* **55**, 2924 (1985).
- [46] K. Johansson, *Commun. Math. Phys.* **209**, 437 (2000).
- [47] M. Prähofer and H. Spohn, *Phys. Rev. Lett.* **84**, 4882 (2000); *J. Stat. Phys.* **108**, 1071 (2002); **115**, 255 (2004).
- [48] M. Žnidarič, T. Prosen, and P. Prelovšek, *Phys. Rev. B* **77**, 064426 (2008).
- [49] J. H. Bardarson, F. Pollmann, and J. E. Moore, *Phys. Rev. Lett.* **109**, 017202 (2012).
- [50] B. Bauer and C. Nayak, *J. Stat. Mech.* (2013) P09005.
- [51] A. Nanduri, H. Kim, and D. A. Huse, *Phys. Rev. B* **90**, 064201 (2014).
- [52] Y. Bar Lev, G. Cohen, and D. R. Reichman, *Phys. Rev. Lett.* **114**, 100601 (2015).
- [53] I. Mondragon-Shem, A. Pal, T. L. Hughes, and C. R. Laumann, *Phys. Rev. B* **92**, 064203 (2015).
- [54] J. Goold, C. Gogolin, S. R. Clark, J. Eisert, A. Scardicchio, and A. Silva, *Phys. Rev. B* **92**, 180202(R) (2015).
- [55] F. Buccheri, A. De Luca, and A. Scardicchio, *Phys. Rev. B* **84**, 094203 (2011).
- [56] G. Biroli, A. Ribeiro-Teixeira, and M. Tarzia, [arXiv:1211.7334](https://arxiv.org/abs/1211.7334).
- [57] A. De Luca, A. Scardicchio, V. E. Kravtsov, and B. L. Altshuler, [arXiv:1401.0019](https://arxiv.org/abs/1401.0019).
- [58] E. Dijkstra, *Numer. Math.* **1**, 269 (1959).
- [59] K. Efetov, *Supersymmetry in Disorder and Chaos* (Cambridge University Press, Cambridge, 1999).