Diffusion in sparse networks: Linear to semilinear crossover

Yaron de Leeuw and Doron Cohen
Department of Physics, Ben Gurion University of the Negev, Beer Sheva 84105, Israel
(Received 12 June 2012; revised manuscript received 27 September 2012; published 19 November 2012)

We consider random networks whose dynamics is described by a rate equation, with transition rates w_{nm} that form a symmetric matrix. The long time evolution of the system is characterized by a diffusion coefficient D. In one dimension it is well known that D can display an abrupt percolation-like transition from diffusion ($D > 0$) to subdiffusion ($D = 0$). A question arises whether such a transition happens in higher dimensions. Numerically D can be evaluated using a resistor network calculation, or optionally it can be deduced from the spectral properties of the system. Contrary to a recent expectation that is based on a renormalization-group analysis, we deduce that D is finite, suggest an "effective-range-hopping" procedure to evaluate it, and contrast the results with the linear estimate. The same approach is useful in the analysis of networks that are described by quasi-one-dimensional sparse banded matrices.

DOI: 10.1103/PhysRevE.86.051120 PACS number(s): 05.40.−a, 02.10.Yn, 63.50.−x, 71.23.Cq

I. INTRODUCTION

The study of network systems is of interest in the diverse fields of mathematics, physics, and computer and life sciences. Commonly a network is described by a symmetric matrix that consists of real non-negative elements, e.g., the adjacency matrix or the link probability matrix, that have unique spectral properties [1,2]. Physically motivated, in this work we consider d-dimensional network systems, whose dynamics is described by a rate equation

$$\frac{dp_n}{dt} = \sum_m w_{nm} p_m. \tag{1}$$

The off-diagonal elements of $\boldsymbol{w}$ are the transition rates, while the diagonal elements are the decay rates

$$w_{nn} = -\gamma_n, \quad \gamma_n \equiv \sum_{m(\neq n)} w_{mn}. \tag{2}$$

We assume a symmetric matrix and write schematically

$$\boldsymbol{w} = \text{matrix}\{w_{nm}\}. \tag{3}$$

In some sense, one can regard $\boldsymbol{w}$ as a discrete Laplacian that is associated with the network. Clearly the physical problem is related to the study of random walk in a disordered environment [3–5].

For presentation purposes we regard the nodes of the network as *sites*, each having a location x_n. By construction, we assume that the transition rates w_{nm} are given by the expression $w_0 e^{-\epsilon_{nm}} B(x_n - x_m)$, where $B(r)$ describes the systematic dependence of the coupling on the distance between the sites, and ϵ is a random variable that might represent, say, the activation energy that is required to make a transition. Consequently the network is characterized by two functions:

$$w(r,\epsilon) \equiv w_0\, e^{-\epsilon}\, B(r), \tag{4}$$

$$\rho(r,\epsilon) \equiv \text{local density of sites}. \tag{5}$$

The latter is defined as the density of sites in (r,ϵ) space, relative to some initial site. Obviously the functional dependence of this density on r is affected by the dimensionality of the network.

Sparsity. Our interest is focused on "sparse" networks. This means that the transition rates between neighboring sites are log-wide distributed as in glassy systems. These rates span several orders of magnitudes as determined by the dispersion of r or by the dispersion of ϵ. In particular (but not exclusively) we are interested in a random site model where the rates depend exponentially on the distance between randomly distributed sites, namely, $B(r) = \exp(-r/\xi)$. In this particular case one can characterize the sparsity by the parameter

$$s = \xi/r_0, \tag{6}$$

where r_0 is the average distance between neighboring sites. We refer to such networks as "sparse" if $s \ll 1$.

Sparsity vs percolation. The problem that we consider is a variant of the percolation problem [6]. Instead of considering a bimodal distribution ("zeros" and "ones") we consider a log-wide distribution of rates [7], for which the median is much smaller than the mean value. We call such a network sparse because the large elements constitute a minority.

Sparsity vs disorder. While the standard percolation problem can be regarded as the outcome of extreme sparsity, the latter can be regarded as arising from an extreme disorder. Accordingly, the model that we are considering is a close relative of the Anderson localization problem, and therefore we shall dedicate some discussion to clarify the relation.

Physical context. The model that we address is related and motivated by various physical problems, for example, phonon propagation in disordered solids [8–10], Mott hopping conductance [7,11–15], transport in oil reservoirs [16,17], conductance of ballistic rings [18], and energy absorption by trapped atoms [19]. Optionally these models can be fabricated by combining oscillators, say, mechanical springs or electrical resistor-capacitor elements. In all these examples the issue is to understand how the *transport* is affected by the *sparsity* of a network. If the rates are induced by a driving source, this issue can be phrased as going *beyond* the familiar framework of linear response theory, as explained below.

Diffusion and subdiffusion. Our interest is focused on the diffusion coefficient D that characterizes the long time dynamics of a spreading distribution. The simplest way to define it, as in standard textbooks, is via the variance

$S(t) \equiv \langle r^2 \rangle_t$. Namely,

$$D \equiv (2d)^{-1} \lim_{t\to\infty} \frac{S(t)}{t}. \tag{7}$$

Optionally it can be defined or deduced from the decay of the survival probability $\mathcal{P}(t) \sim (Dt)^{-d/2}$. Hence it is related to the spectral properties of the transition rate matrix.

In the $d = 1$ case, it is well known [20] that D can display an abrupt percolation-like transition from diffusive ($D > 0$) to subdiffusive ($D = 0$) behavior, as the sparsity parameter drops below the critical value $s_{\rm cr} = 1$. Similar anomalies are found for fractal structures with $d < 2$, also known as "random walk on percolating clusters," see [21–25]. A question arises as to whether such a transition might happen in higher dimensions.

In [10] the spectral properties in the $d = 2$ case were investigated: on the basis of the renormalization group (RG) procedure it was deduced that $\mathcal{P}(t)$ decays in a logarithmic way, indicating anomalous (sub) diffusion. In the present work we shall introduce a different approach that implies, contrary to the simple RG treatment, that in spite of the sparsity, the long time dynamics is in fact diffusive rather than subdiffusive.

Resistor network picture. One can regard the p_n in Eq. (1) as the charge in site n; each site is assumed to have unit capacitance; hence $p_n - p_m$ is the potential difference, and $w_{nm}(p_m - p_n)$ is the current from m to n. Accordingly Eq. (1) can be regarded as the Kirchhoff equation of the circuit. While calculating D it is illuminating to exploit the implied formal analogy with a resistor network calculation [12,14,18,26]. Namely, regarding w_{nm} as connectors, it follows that D is formally like conductivity. It follows that $D[\boldsymbol{w}]$ is in general a *semilinear* function:

$$D[\lambda \boldsymbol{w}] = \lambda D[\boldsymbol{w}], \tag{8}$$

$$D[\boldsymbol{w}^a + \boldsymbol{w}^b] > D[\boldsymbol{w}^a] + D[\boldsymbol{w}^b]. \tag{9}$$

If the rates are induced by a driving source, the above super additivity implies that the analysis should go *beyond* the familiar framework of linear-response theory [27].

In this work we obtain an improved estimate for D that we call effective range hopping (ERH). Using this approach we show that in the $d = 2$ case, as s becomes small, the functional $D[\boldsymbol{w}]$ exhibits a smooth crossover from linear behavior to semilinear VRH-type dependence. Our approach is inspired by the resistor network picture of [7,12–19,27], and leads in the appropriate limit to the well-known Mott's variable range hopping (VRH) estimate for D.

Outline. We first describe some known results and some additional numerical results for the spectral properties of $d = 1$ and $d = 2$ networks and for the dependence of D on the sparsity. Then we show that an ERH procedure is useful in describing the crossover from the linear regime (no sparsity) to the semilinear regime. In the latter regime a resistor network approach is essential, and the percolation threshold manifests itself in the calculation. Finally we demonstrate that the same ERH procedure can be applied in the case of a quasi-one-dimensional network that is described by a sparse banded random matrix. The latter is of relevance to previous studies of energy absorption by a weakly chaotic system [27]. We conclude with a discussion and a short summary.

II. RANDOM SITE HOPPING MODEL

Consider a network that consists of sites that are distributed in space, locations x_n. With each bond nm we associate an activation energy $\epsilon_{nm} > 0$, and assume

$$w_{nm} = w_0\, e^{-\epsilon_{nm}}\, e^{-|x_n - x_m|/\xi}. \tag{10}$$

Accordingly we have the identification

$$B(r) = e^{-r/\xi}. \tag{11}$$

We note that in the traditional formulation of the Mott problem the activation energies are not due to some barriers, but are determined by the on-site binding energies, namely, $\epsilon_{nm} = |\varepsilon_n - \varepsilon_m|/T$, where T is the temperature. In this paper we treat the ϵ_{nm} as an uncorrelated random variable.

The density of sites relative to some initial site is characterized by a joint distribution function

$$\rho(r,\epsilon)dr\, d\epsilon = \frac{\Omega_d\, r^{d-1} dr}{r_0^d}\, f(\epsilon)d\epsilon, \quad \Omega_d = 2,2\pi,4\pi. \tag{12}$$

We distinguish between the Mott hopping model and the degenerate hopping model. Namely,

$$f(\epsilon) = 1 \quad \text{Mott hopping model,} \tag{13}$$

$$f(\epsilon) = \delta(\epsilon) \quad \text{Degenerate hopping model.} \tag{14}$$

The normalization of $f(\epsilon)$ as defined above fixes the value of the constant r_0^d, which we regard as the "unit cell." In the numerics we set the units of distance such that $r_0 = 1$.

In the traditional formulation of the Mott problem it is assumed that mean level spacing within ξ^d is Δ_ξ, such that the number of accessible sites is $(d\varepsilon/\Delta_\xi)(d^3r/\xi^d)$. By the convention of Eq. (12) this implies that the unit cell dimension is temperature dependent

$$r_0^d = \left(\frac{\Delta_\xi}{T}\right)\xi^d \quad \text{[for Mott model].} \tag{15}$$

We reemphasize that the number of sites per unit volume in the Mott problem is infinite, but effectively only $\sim T/\Delta_\xi$ sites are accessible within ξ^d per attempted transition. It is convenient to characterize a random site model by a sparsity parameter that is defined as in Eq. (6). Accordingly

$$s \equiv \frac{\xi}{r_0} = \left(\frac{T}{\Delta_\xi}\right)^{1/d} \quad \text{[for Mott model].} \tag{16}$$

We refer to a network as sparse if $s \ll 1$.

The lattice model with near-neighbor (n.n.) transitions is one of the most popular models in statistical mechanics: in particular the random walk problem on a lattice is a standard textbook example. If the rates are generated from a log-wide distribution, it can be regarded as a variant of the random site hopping model. For details see Appendix A. In particular we note that the $d = 1$ version is formally equivalent: it does not matter whether the distribution of w is due to random distances r or due to random activation energies ϵ.

Finally we note that a quasi-one-dimensional version of the random site model arises in the study of energy absorption as explained in Appendix B, and later addressed in Sec. XI.

III. CHARACTERIZATION OF TRANSPORT

The long time dynamics that takes place on the network is characterized by the spreading $S(t)$ and by the survival probability $\mathcal{P}(t)$. If the system is diffusive, these functions have the following functional form:

$$S(t) = \langle r^2 \rangle_t \sim (2d)Dt, \tag{17}$$

$$\mathcal{P}(t) \sim \frac{r_0^d}{(4\pi\, Dt)^{d/2}}. \tag{18}$$

See Appendix C for details. The diffusion coefficient D appears here in consistency with its definition in Eq. (7). We note that in the case of subdiffusion

$$S(t) \propto t^{\alpha}, \quad [\alpha < 1], \tag{19}$$

which implies by Eq. (7) that $D = 0$.

The spectrum of the matrix $\boldsymbol{w}$ consists of the trivial eigenvalue $\lambda_0 = 0$ that is associated with a uniform distribution, and a set of negative numbers $-\lambda_k$ that describe the decaying modes. The spectral function $\mathcal{N}(\lambda)$ counts the number of eigenvalues up to the value λ. We normalize it per site such that $\mathcal{N}(\infty) = 1$. The associated density of eigenvalues $g(\lambda)$ is related to $\mathcal{P}(t)$ by a Laplace transform. See Appendix C for details. It follows that in the case of a diffusive system

$$\mathcal{N}(\lambda) = \int^{\lambda} g(\lambda)d\lambda \sim \left(\frac{r_0}{2\pi}\right)^d \left[\frac{\lambda}{D}\right]^{d/2}. \tag{20}$$

In Appendix D we clarify that this expression agrees with the Debye law. Accordingly the calculation of D parallels the calculation of the speed of sound c in the Debye model.

Regarded as a transport coefficient D relates the probability current to the density gradient. This is known as Fick's law. From the discussion in the Introduction it follows that D is like the *conductivity* of a resistor network, which relates the electrical current to the voltage difference. Some further details on the practical calculation of the conductivity are presented in Appendix E. On the basis of this analogy it should be clear that $D[\boldsymbol{w}]$ is in general a *semilinear* function of the rates, see Eq. (9).

IV. EXACT AND NUMERICAL RESULTS FOR THE $d = 1$ LATTICE MODEL

In the case of a $d = 1$ lattice model with n.n. transitions it is natural to use the notation $w_n = w_{n,n-1}$. Pointing out the analogy with adding connectors in series, the expression for D is

$$D = \left(\frac{1}{N}\sum_n \frac{1}{w_n}\right)^{-1} = \frac{s-1}{s}\, w_0[s > 1]. \tag{21}$$

The calculation that leads to the last equality has been done with the distribution of Eq. (A2), where $s \equiv \xi/r_0$. Note that we have here a serial addition of resistors $R = \sum_n R_n$, where $R_n = 1/w_n$. For $s < 1$ the distribution of each R_n is dominated by the large values, hence $R = \infty$. On the other extreme for $s > 1$ the distribution of the R_n has finite first and second moments, and accordingly the result for R becomes self-averaging, as implied by the central limit theorem. This means the D is well defined only for $s > 2$. For $1 < s < 2$ the result for the average R is finite but not self-averaging.

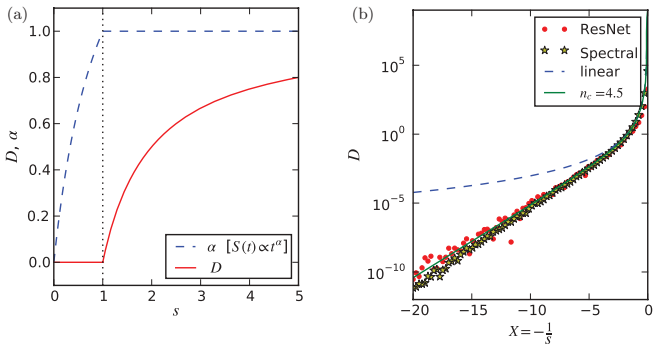

FIG. 1. (Color online) Spreading in (a) the $d = 1$ lattice model and (b) the $d = 2$ degenerate random site model. Panel (a) is based on known exact results. Its dashed blue line is the power α of the spreading, showing a subdiffusive regime for $s < 1$, and a diffusive regime for $s > 1$. Its solid red line is the diffusion coefficient D, which is zero in the subdiffusive regime. Panel (b) displays numerical results that refer to a network that consists of $N = 2000$ sites randomly scattered over a square with periodic boundary conditions. The vertical axis is the diffusion coefficient D in a logarithmic scale, while the horizontal axis is $X = -1/s$. The numerical red dots are based on a resistor network calculation (see Appendix E), while the stars are extracted from the spectral analysis (see Fig. 2). The dashed line is the linear estimate (corresponds to $n_c = 0$), while the solid line is the ERH estimate with $n_c = 4.5$. One observes that the ERH calculation describes very well the departure from the linear prediction.

The dependence of D on s is illustrated in Fig. 1(a). In the subdiffusive regime ($s < 1$), where the result for the diffusion coefficient is $D = 0$, the dynamics becomes subdiffusive. The explicit results for the survival probability and for the spreading are known [20]:

$$S(t) \sim t^{2s/(1+s)}, \tag{22}$$

$$\mathcal{P}(t) \sim t^{-s/(1+s)}, \tag{23}$$

and the associated spectral function is

$$\mathcal{N}(\lambda) \sim \lambda^{s/(1+s)}. \tag{24}$$

The numerical demonstration of the latter expectation is displayed in Fig. 2 (left upper panel). We clearly see that for $s < 1$ the asymptotic slope corresponds to subdiffusion, while for $s > 1$ it corresponds to diffusion.

V. NUMERICAL RESULTS FOR THE $d = 2$ RANDOM SITE MODEL

Results for the spectral counting function of the degenerate $d = 2$ random site model are presented in Fig. 2 (right upper panel). We also display there (in the lower panel) the participation number (PN) for each eigenstate. The PN of an eigenstate that corresponds to an eigenvalue λ_k is conventionally defined as

$$\mathrm{PN} \equiv \left[\sum_n |\langle n|\lambda_k\rangle|^4\right]^{-1}. \tag{25}$$

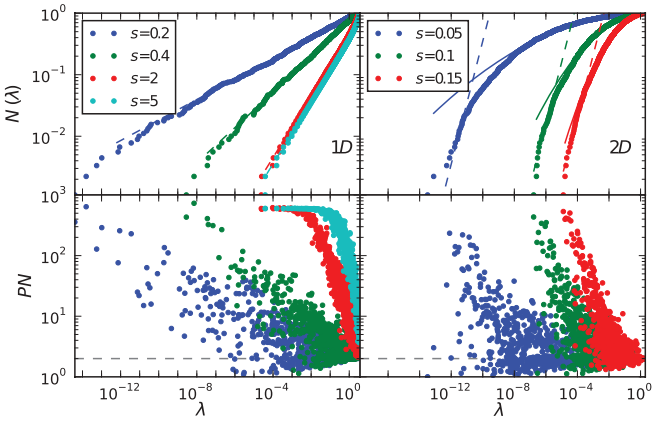

FIG. 2. (Color online) Cumulative eigenvalue distributions $\mathcal{N}(\lambda)$ for the $d = 1$ (1D) and for the $d = 2$ (2D) models of Fig. 1, and the respective PN of the eigenstates (lower panels). Several representative values of s are considered. The dots are determined via numerical diagonalization of $N \times N$ matrices, each representing a network that consists of $N = 1000$ sites randomly scattered over a square with periodic boundary conditions. There is a striking difference between the $d = 1$ and the $d = 2$ cases. For $d = 1$, the log-log slope of $\mathcal{N}(\lambda)$, see dashed lines, is less than $d/2$ for sparse networks ($s < 1$), meaning that we have subdiffusion. In the $d = 2$ case the small-λ log-log slope is always $d/2$, which corresponds to normal diffusion. The solid lines in the upper 2D plot are according to the RG analysis of [10], namely, Eq. (26). The horizontal dashed line in the lower panels indicates the special value PN= 2 that corresponds to dimer formation.

As expected from the study of localization in a disordered elastic medium [28], the PN becomes larger in the limit $\lambda \to 0$, without apparent indication for a mobility threshold.

Assuming localized modes that are conceived via dimerization of neighboring sites, $\mathcal{N}(\lambda)$ should equal the probability $\exp[-\mathsf{V}(r)/r_0^d]$ not to have any neighboring site within the volume $\mathsf{V}(r)$ of the sphere $2w_0 \exp(-r/\xi) > \lambda$. The RG analysis of [10] refines this naive expectation, adding a factor of 2 in the exponent, leading to

$$\mathcal{N}(\lambda) = \exp\left\{ -\frac{\Omega_d}{2d}\left[-s \ln\left(\frac{\lambda}{2w_0}\right)\right]^d \right\}, \tag{26}$$

where $s \equiv \xi/r_0$. This expectation is represented in Fig. 2 (right upper panel) by solid lines. We see that it fails to capture the small λ regime, where the distribution corresponds to diffusive behavior.

Extracting D via fitting to Eq. (20) we get Fig. 1(b). We see that in the $d = 2$ model *there is no abrupt crossover to subdiffusion*. We therefore would like to find a way to calculate D, and hence to have the way to determine the small λ asymptotics.

Note added. One should conclude that the RG of Ref. [10] applies only to the analysis of the high frequency response, while our interest is focused in the low frequency (direct current) analysis. The crossover between the two regimes is implied. For more details in this direction see a followup work [29] that confirms our physical picture and demonstrates numerically the implied crossover.

VI. LINEAR AND ERH ESTIMATES FOR THE DIFFUSION COEFFICIENT

The standard way to calculate diffusion in a $d = 1$ random walk problem is to inspect the transient growth of the variance $\mathrm{Var}(n) = 2Dt$. In the stochastic context, if we start at site n we have $\mathrm{Var}(n) = \sum_{n'} p_{n'}(n'-n)^2$, with $p_{n'} = w_{n'n}t$, hence

$$D_n = \frac{1}{2}\sum_{n'}(n'-n)^2\, w_{n'n}. \tag{27}$$

The generalization to more than one dimension is straightforward. Averaging the transient expression over the starting point we get the result

$$D_{\mathrm{linear}} = \frac{1}{2d}\iint w(r,\epsilon) r^2 \rho(r,\epsilon) d\epsilon\, dr. \tag{28}$$

This expression is strictly linear. It describes correctly the average transient spreading. In the absence of disorder we can trust it for arbitrary long time. But if we have a disordered or sparse network, the possibility for transport is related to the theory of percolation [7,13,14]. We are therefore motivated to introduce an approximation scheme that takes the percolation aspect into account. We shall refer to this scheme as effective range hopping (ERH) because it is a variation on the well-known VRH procedure.

Inspired by [7,13,14] we look for the threshold w_c that is required for percolation. In the ERH scheme we suggest using the following equation for its determination:

$$\iint_{w(r,\epsilon)>w_c} \rho(r,\epsilon) dr\, d\epsilon = n_c. \tag{29}$$

Here n_c is the effective coordination number that is required for getting a connected sequences of transitions. For a $d = 2$ square lattice model it is reasonable to set $n_c = 2$, reflecting the idea of forming a simple chain of transitions. Rephrased differently the requirement is to have an average of 50% connecting bonds per site. For a $d = 2$ random site model one should be familiar with the problem of percolation in a system that consists of randomly distributed discs. The effective coordination number that is required for getting percolation in such a model is $n_c = 4.5$, as found in [30], and further discussed in Sec. IV A 1 of [31].

The second step in the ERH scheme is to form an effective network whose sparse elements are suppressed to the threshold value. Then it is possible to use the linear formula Eq. (28). Hence we get

$$D_{\mathrm{ERH}} = \frac{1}{2d}\iint \min\{w(r,\epsilon), w_c\}\, r^2\, \rho(r,\epsilon)\, d\epsilon\, dr. \tag{30}$$

This expression, as required, is semilinear rather than linear. It looks like the linear estimate of Eq. (28), but it involves a network with w_{nm} that are equal or smaller to the original values. The "suppressed" connectors are those that are too sparse to form percolating trajectories.

VII. VARIABLE RANGE HOPPING ESTIMATE

The ERH is similar to the generalized VRH procedure that we have used in previous publications [18,19]. The traditional VRH is based on the idea of associating an energy cost $\varepsilon(r)$ to

a jump that has range r. Namely,

$$\varepsilon(r) \sim \left[\frac{\Omega_d}{d}\, r^d\right]^{-1} \Delta_0, \tag{31}$$

corresponding to the average level spacing of the sites within a range r. In our notation, $\epsilon(r) \equiv \varepsilon(r)/T$. For the general network models that we consider here, the relation between ϵ and r is determined through the equation

$$\int_0^{\epsilon}\int_0^{r} \rho(r',\epsilon')dr'\,d\epsilon' = n^*, \tag{32}$$

where n^* is of order unity. In fact we shall deduce later, in Sec. X, that for consistency with the ERH estimate this value should be $n^* = n_c/d$. With the substitution of Eq. (12) the tradeoff equation can be written as

$$\Omega_d \left(\frac{r}{r_0}\right)^d F(\epsilon) = n_c, \tag{33}$$

where $F(\epsilon)$ is the cumulative distribution function that corresponds to the density $f(\epsilon)$. In the Mott problem $F(\epsilon) = \epsilon$, and Eq. (31) is recovered. In words, Eq. (32) asks what is the ϵ window that is required in order to guarantee that the particle will be able to find with probability of order unity an accessible site within a range r. Larger jumps allow smaller cost. Then we estimate D as follows:

$$D_{\rm VRH} \sim w^*(r^*)^2, \tag{34}$$

where r^* is the optimal range that maximizes $w(r,\epsilon(r))$, with associated energy cost $\epsilon^* = \epsilon(r^*)$, and effective transition rate $w^* = w(r^*,\epsilon^*)$. See Fig. 3 for illustration.

The VRH estimate, unlike the ERH, does not interpolate with the linear regime. It can be used to estimate D only if the system is very sparse ($s \ll 1$). It can be regarded as an asymptotic evaluation of the ERH integral: it assumes that the hopping is dominated by the vicinity of the optimal point (r^*,ϵ^*). Accordingly, VRH-to-ERH consistency requires

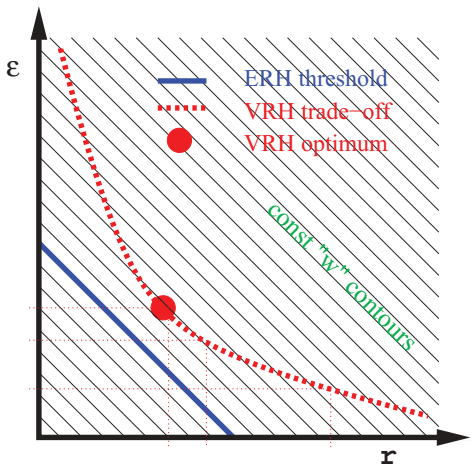

FIG. 3. (Color online) Comparing VRH with the ERH procedure. The solid blue line that corresponds to the ERH threshold w_c encloses an "area" that corresponds to n_c. The VRH tradeoff is represented by the dashed red line. The VRH optimum is represented by the thick red dot. The VRH-to-ERH consistency requirement [Eq. (46)] is to have the VRH optimum sitting on the solid blue line.

the identification $w^* = w_c$. However, using known results from percolation theory, one possibly can further refine the determination of the optimal value w^*. Namely, a somewhat smaller value than the threshold value w_c might allow a better connectivity. As s becomes very small, the effective range δw in the ERH integral, which contains the dominant contribution, becomes very small compared with $w_c - w^*$, and one should be worried about the implied (subdominant) correction. This speculative crossover is beyond the scope of the present study, and possibly very hard to detect numerically. A useful analogy here is with the crossover from "mean-field" to "critical" behavior in the theory of phase transition, as implied by the Ginzburg criterion.

VIII. ERH CALCULATION FOR THE $d = 2$ LATTICE MODEL

The $d = 2$ lattice model, as defined in Appendix A, is the simplest and most common example for studies of percolation and percolation-related problems. We substitute into Eq. (29) the effective density Eq. (A3) with the coordination number $c_L = 4$, and deduce that w_c is merely the *median* value of the n.n. transition rates. The ERH calculation using Eq. (30) with Eq. (A3) requires a simple $f(\epsilon)d\epsilon$ integration, which can be rewritten as a $\tilde{f}(w)dw$ integral. This integral is the sum of $w > w_c$ and $w < w_c$ contributions, namely,

$$D_{\rm ERH} = \left[\frac{1}{2}w_c + \frac{1}{2}\int_0^{w_c} w\,\tilde{f}(w)dw\right] r_0^2. \tag{35}$$

Note that the first term in the square brackets originates from the $w > w_c$ contribution. Note also that the result is $D = w_c r_0^2$ for a delta distribution, i.e., in the absence of disorder.

IX. ERH CALCULATION FOR THE DEGENERATE HOPPING MODEL

We now turn to the calculation of the ERH estimate for the degenerate hopping model. The ERH threshold can be written as $w_c = w_0 \exp(-r_c/\xi)$, where r_c is determined through Eq. (29), which takes the form

$$\int_0^{r_c} \frac{\Omega_d r^{d-1} dr}{r_0^2} = n_c, \tag{36}$$

leading to

$$w_c = w_0 \exp\left(-\frac{r_c}{\xi}\right), \tag{37}$$

$$r_c \equiv \left(\frac{d}{\Omega_c} n_c\right)^{1/d} r_0. \tag{38}$$

The calculation of the ERH integral of Eq. (30) is detailed in Appendix F. We note that the linear approximation of Eq. (28) is formally obtained by setting $r_c = 0$, leading to

$$D_{\rm linear} = \frac{(d+1)!\,\Omega_d}{2d}\, s^{d+2}\, w_0 r_0^2. \tag{39}$$

Then it is possible to write the result of the ERH integral as

$$D_{\rm ERH} = {\rm EXP}_{d+2}\left(\frac{1}{s_c}\right) e^{-1/s_c}\, D_{\rm linear} \tag{40}$$

where $s_c = \xi/r_c$, and

$$\mathrm{EXP}_\ell(x) = \sum_{k=0}^{\ell} \frac{1}{k!} x^k. \tag{41}$$

The linear result is formally obtained by setting $n_c = 0$ or in the $d \to \infty$ limit. In the other extreme of $s \ll 1$ we get a VRH-like dependence

$$D \sim e^{-1/s_c}, \quad \text{for} \quad s \ll 1. \tag{42}$$

Numerical verification. To obtain an ERH estimate we have to fix the parameter n_c in Eq. (36). One approach is to regard it as a free fitting parameter. But it is of course better not to use any fitting parameters. Fortunately we know from [30,31] that $n_c = 4.5$ is the average number of bonds required to get percolation. The verification of the ERH estimate for the random site model with this value is demonstrated in Fig. 1(b).

X. ERH CALCULATION FOR THE MOTT HOPPING MODEL

We turn to calculating the ERH estimate for the nondegenerate Mott hopping model, and contrast it with the linear approximation, and with the traditional VRH estimate. The ERH threshold is determined through Eq. (29), leading to

$$w_c = w_0 \exp(-\epsilon_c), \tag{43}$$

$$\epsilon_c \equiv \left(\frac{d}{\Omega_d}\frac{n_c}{s^d}\right)^{1/(d+1)}. \tag{44}$$

In the VRH procedure the optimal hopping range is found by maximizing $w(r,\epsilon)$ along the tradeoff line of Eq. (33), as illustrated in Fig. 3, leading to

$$r^* = \left(\frac{d^2}{\Omega_d} n^* s\right)^{1/(d+1)} r_0, \tag{45}$$

and the associated rate is

$$w^* = w_c \quad \text{provided} \quad n^* = n_c/d. \tag{46}$$

This identification is necessary if we want the VRH to describe correctly the asymptotic dependence of D on s.

The calculation of the ERH integral of Eq. (30) is detailed in Appendix F. Thanks to our conventions the linear result is the same as Eq. (39), and the the final result can be written as follows:

$$D_{\mathrm{ERH}} = \mathrm{EXP}_{d+3}(\epsilon_c)\, e^{-\epsilon_c} D_{\mathrm{linear}}, \tag{47}$$

where $\mathrm{EXP}(x)$ is the polynomial defined in Eq. (41). The linear result is formally obtained by setting $\epsilon_c = 0$ or in the $d \to \infty$ limit.

We see that the VRH estimate can be regarded as an asymptotic approximation that holds for $s \ll 1$. Using Eqs. (15) and (16) we deduce from Eqs. (39) and (47) that

$$D_{\mathrm{linear}} \propto T, \tag{48}$$

while for $s \ll 1$,

$$D_{\mathrm{ERH}} \sim \left(\frac{1}{T}\right)^{2/(d+1)} \exp\left[-\left(\frac{T_0}{T}\right)^{1/(d+1)}\right], \tag{49}$$

where T_0 is a constant.

XI. ERH CALCULATION FOR THE BANDED QUASI-ONE-DIMENSIONAL MODEL

We can apply the ERH calculation also to the case of the quasi-one-dimensional model that we have studied in the past [18,19]. This model is motivated by studies of energy absorption [27]. For details see Appendix B. The network is defined by a banded matrix $\boldsymbol{w}$. For simplicity we assume that the sites are equally spaced and that the reason for the sparsity is the log-wide distribution of the in-band elements.

The ERH threshold w_c is deduced from Eq. (29). For a general $B(r)$ and $f(\epsilon)$ one can integrate over $d\epsilon$, and then it takes the form

$$\int_0^\infty \frac{\Omega_d\, r^{d-1} dr}{r_0^d}\, F\left[\log\left(\frac{w_0}{w_c} B(r)\right)\right] = n_c, \tag{50}$$

where $F(\epsilon)$ is the cumulative distribution function that corresponds to the density $f(\epsilon)$. Here we are considering a $d = 1$ network. However, we are dealing with a banded matrix which in some sense is like adding an extra (but bounded) dimension to the lattice.

Specifically we assume that $B(r) = 1$ within the band, and zero for $|r| > b$. The nonzero elements have a log-box distribution, namely, ϵ is distributed uniformly over a range $[0,\sigma]$. To have large σ means "sparsity." One should notice that this sparsity is less traumatic than having $s \ll 1$ in the $d = 1$ lattice model that we considered in Sec. IV. This is because the distribution is bounded from below by finite nonzero values. Accordingly we cannot have subdiffusion here.

We now turn to estimate D using the the ERH procedure. It should be clear that the success here is not guaranteed for reasons that we further discuss in the last paragraph of this section. From Eq. (50) it follows that $w_c = w_0 \exp(-\epsilon_c)$, where ϵ_c is the solution of

$$2bF(\epsilon_c) = n_c. \tag{51}$$

For the assumed ϵ distribution the solution of this equation is trivial

$$\epsilon_c = \frac{n_c}{2b}\sigma. \tag{52}$$

While doing the ERH integral of Eq. (30) note that the integral dr should be replaced by a sum. It is convenient to define

$$\tilde{b} \equiv \sum_{r=1}^{b} r^2 = \frac{1}{6} b(b+1)(2b+1). \tag{53}$$

Then the ERH estimate takes the form

$$D_{\mathrm{ERH}} = \frac{1}{\sigma}\left[\left(1 + \frac{n_c}{2b}\sigma\right) e^{-\frac{n_c}{2b}\sigma} - e^{-2\sigma}\right] \tilde{b} w_0. \tag{54}$$

The linear estimate of Eq. (28) is formally obtained by setting $n_c = 0$, and in the absence of disorder it obviously reduced to $D = \tilde{b} w_0$. We define

$$g_s = D/D_{\mathrm{linear}}. \tag{55}$$

Numerical results are presented in Fig. 4, and they agree with the ERH estimate.

At this point one wonders whether D can be extracted from the spectral analysis, i.e., via fitting to Eq. (20). In Fig. 4(c) we plot the D that is extracted from the spectral

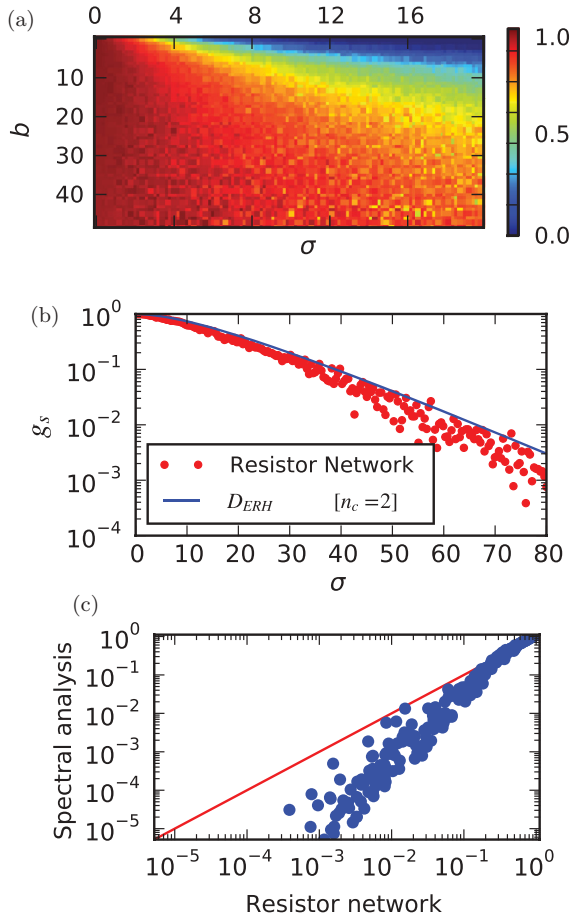

FIG. 4. (Color online) We consider a quasi $d = 1$ network that consists of $N = 1000$ sites with periodic boundary conditions. The network is described by a sparse banded matrix. The bandwidth is b, and the log-width of the rate distribution is σ. See text for details. (a) The numerical result for $g_s = D/D_{\text{linear}}$ imaged as a function of σ and b. The values of D are found via a numerical resistor network calculation, see Appendix A. (b) Subset of results that refer to the $b = 10$ matrix. The curve is the ERH prediction. (c) Scatter diagram shows the correlation between the D that was extracted from the spectral analysis and the D that was found via the resistor network calculation.

analysis versus the D that has been found via the resistor network calculation. We observe that the obtained values are much smaller. Our interpretation for that is as follows. The density of eigenvalues is related to the survival probability $\mathcal{P}(t)$ via a Laplace transform. For a quasi-one-dimensional system there is a short-time, $d = 2$ like, relatively fast transient. Consequently the $d = 1$ decay holds only asymptotically with a smaller prefactor. Accordingly we do not know whether there is a wise way to deduce D from the spectral analysis in the case of a quasi-one-dimensional network.

Concluding this section we would like to warn the reader that the use of the percolation picture in $d = 1$ is somewhat problematic: strictly speaking there is no percolation transition. Obviously for $b = 1$ we are back with the $d = 1$ lattice model for which there is subdiffusion if $s < s_{\text{cr}}$ with $s_{\text{cr}} = 1$. However, if b is reasonably large, it is not feasible to encounter such an anomaly in practice. Even if the distribution is not bounded from below, the redundancy due to $b > 1$ would lower the effective value of s_{cr}. Furthermore, in the Fermi-golden-rule picture (see next section) the occurrence of "weak links" along the band are practically not possible because the matrix elements V_{nm} are not uncorrelated random variables. We can refer to this as the *rigidity*. This rigidity is implied by semiclassical considerations.

XII. SEMILINEAR RESPONSE PERSPECTIVE

Considering models of energy absorption (see Appendix B), it is assumed that the transition rate w_{nm}, between unperturbed energy levels m and n, is determined by a driving source that has spectral content $\tilde{S}(\omega)$. The Fermi golden rule can be written as

$$w_{nm} = \tilde{S}(E_n - E_m)\,|V_{nm}|^2, \tag{56}$$

where V_{nm} is the perturbation matrix in the Hamiltonian. Accordingly we can write instead of $D = D[\boldsymbol{w}]$ an implied relation $D = D[\tilde{S}(\omega)]$. This relation is in general semilinear. This means that only the first property below, which corresponds to Eq. (8), is satisfied, not the second one.

$$D[\lambda\tilde{S}(\omega)] = \lambda\, D[\tilde{S}(\omega)], \tag{57}$$

$$D[\tilde{S}_a(\omega) + \tilde{S}_b(\omega)] = D[\tilde{S}_a(\omega)] + D[\tilde{S}_b(\omega)]. \tag{58}$$

To have a semilinear rather than linear response may serve as an experimental signature for the applicability of resistor-network modeling of energy absorption. We note, however, that if the the driving were added "on top" of a bath, the response would become linear at small intensities. Namely, if one substituted

$$\tilde{S}(\omega)_{\text{total}} = \tilde{S}_{\text{bath}}(\omega) + \tilde{S}(\omega), \tag{59}$$

it would be possible to linearize D with respect to the $\tilde{S}(\omega)$ of the driving source.

The statement that VRH is a "semilinear response" theory rather than "linear response" theory is a source for nonconstructive debates on terminology. The reason for the confusion about this point is related to the physical context. Do we calculate "current vs bias" or do we calculated "diffusion vs driving"? The response is linear in the former sense, but semilinear in the latter sense.

XIII. DISCUSSION

It should be clear that there are two major routes in developing a theory for D. Instead of deducing it from spectral properties as in [10], one can try to find ways to evaluate it directly via a resistor network calculation [7,12–15], leading in the standard Mott problem to the VRH estimate for D.

In [18,19,27] this approach was extended to handle "sparse" banded matrices whose elements have log-wide distribution, leading to a generalized VRH estimate. In this work we have pursued the same direction and obtained an improved estimate for D, the ERH estimate. Using this approach we showed that in the $d = 2$ case, as s becomes small, the functional $D[\boldsymbol{w}]$ exhibits a smooth crossover from linear behavior to semilinear VRH-type dependence.

Relation to other models. Disregarding the sparsity issue, the model that we were considering is a close relative of the Anderson localization problem. However, it is not the same

problem, and there are important differences that we would like to highlight. For the purpose of this discussion it is useful to be reminded that the hopping problem that we have addressed is essentially the same as studying the spectrum of vibrations in a disordered elastic medium. Hence $D^{1/2}$ parallels the speed of sound c of the Debye model. See Appendix D.

Mott vs Anderson. In the hopping model all the off-diagonal elements are positive numbers, while the negative diagonal elements compensate them. It follows that we cannot have "destructive interference," and therefore we do not have genuine Anderson localization. Consequently in general we might have diffusion, even in $d = 1$. In $d = 2$ we have a percolation threshold, which is again not like Anderson localization. See the discussion of fractons in [24].

Debye vs Anderson. In the standard Anderson model the eigenvalues form a band $\lambda \in [-\lambda_c,\lambda_c]$. The states at the edge of the band are always localized. The states in the middle of the band might be delocalized if $d > 2$. The spectrum that characterizes the hopping model does not have the same properties. With regard to the localization of vibrations in a disordered elastic medium [28], it has been found that the spectrum is $\lambda \in [0,\lambda_c]$. The ground state is always the $\lambda = 0$ uniform state. The localization length diverges in the limit $\lambda \to 0$. Consequently the Debye density of states is not violated: the spectrum is asymptotically the same as that of a diffusive (nondisordered) lattice. It follows that the survival probability should be like that of a diffusive system, and therefore we also expect, and get, diffusive behavior for the transport: spreading that obeys a diffusion equation.

XIV. SUMMARY

This work was originally motivated by the necessity to improve the resistor-network analysis of the diffusion in quasi-one-dimensional networks [18], and additionally from the desire to relate it to the recent RG studies [10] of the spectral properties of random site networks. The key issue that we wanted to address was the crossover from linear-like to semilinear dependence of D on the rates. This crossover show up as the "sparsity" of the system is varied.

It should be clear that unlike the RG-based expectation of [10], our analysis indicates that there is no subdiffusive behavior in $d = 2$. Accordingly, the anomalous $\log(t)$ spreading that is predicted in [10] should be regarded as a transient: for very small value of the sparsity parameter this transient might have a very long duration, but eventually normal diffusion takes over.

One can regard sparsity as an extreme type of disorder: the rates are distributed over many orders of magnitude. Still, unlike the $d = 1$ case, the implication of sparsity in $d = 2$ is not as dramatic: there is no phase transition between two different results, but a smooth crossover. It is therefore clear that our statements are consistent with those of older works that relate to the diverging localization properties of the low frequency vibrations in a disordered elastic medium [28].

The effective range hopping (ERH) procedure that we tested in this paper is a refinement of well-known studies of variable range hopping [7,11–15]. We used the insight of [7,13,14] that connects VRH with the theory of percolation.

Disregarding possible inaccuracy in the determination of the optimal rate, the ERH calculation provides a *lower* bound for D. Accordingly, by obtaining a nonzero result it is rigorously implied that D is finite. The purpose of the numerics was to demonstrate that in practice the outcome of the ERH calculation provides a very good estimate of the actual result, interpolating very well the departure from linearity.

It was important for us to clarify that a large class of networks can be treated *on an equal footing*. In particular we demonstrated that the application of the ERH estimate does not require any fitting parameters. We have verified that the same prescription can be applied to both the $d = 2$ lattice model and the $d = 2$ random site model, provided one uses the appropriate percolation threshold that is known from percolation theory.

For the traditional Mott hopping model and its degenerated version we obtained the refined expressions Eq. (47) and Eq. (40), respectively. In these expressions the *full* dependence on the dimensionality (d) is explicit, and the crossover to a linear response as a function of the sparsity (s) is transparent. Note that in the degenerate random site model the sparsity is merely a geometrical feature, while in the nondegenerate Mott model the sparsity depends on the temperature as implied by Eq. (16).

We would like to reemphasize that the original motivation for this work was the study of energy absorption by driven mesoscopic systems. In this context the implication of the semilinear crossover is the breakdown of linear response theory. The latter issue has been extensively discussed in past publications [27].

ACKNOWLEDGMENTS

We thank Amnon Aharony, Ariel Amir, Ora Entin-Wohlman, Rony Granek, and Joe Imry for illuminating discussions, comments, and references. This work has been supported by the Israel Science Foundation.

APPENDIX A: LATTICE MODEL WITH N.N. HOPPING

For $s \ll 1$ the $d = 1$ random site model is essentially equivalent to a lattice model with equally spaced sites, near-neighbor transitions, and random ϵ. From the identification $\epsilon = r/\xi$ it follows that the distribution of the "activation energy" is

$$f(\epsilon) = s\exp(-s\epsilon), \quad s \equiv \xi/r_0. \tag{A1}$$

This implies that the the distribution of the rates is

$$\tilde{f}(w)dw = \frac{s\, w^{s-1}dw}{w_0^s}[w < w_0]. \tag{A2}$$

The density of sites to which a transition can occur is

$$\rho(r,\epsilon) = c_L\delta(r - r_0)\, f(\epsilon), \tag{A3}$$

where $c_L = 2$ is the coordination number. This corresponds to the $d = 1$ case of Eq. (12).

The $d = 2$ version of the lattice model has no strict relation to the $d = 2$ random site model. A popular choice is to assume a box distribution for the activation energy within some interval $0 < \epsilon < \sigma$. The density of sites to which a transition can occur is $2\pi r f(\epsilon)$ for large r, as implied by Eq. (12). But for small r

the effective density is given by Eq. (A3) with the coordination number $c_L = 4$.

APPENDIX B: QUASI-ONE-DIMENSIONAL BANDED MATRIX MODEL

On equal footing we consider the quasi-one-dimensional banded lattice model. This model is motivated by studies of energy absorption [27]. In this context the transition rates are determined by the Fermi golden rule (FGR). Hence we write

$$w_{nm} = w_0\, e^{-\epsilon_{nm}}\, B(E_n - E_m). \tag{B1}$$

Here n and m are unperturbed energy levels of the system, but we shall keep calling them "sites" in order to avoid duplicate terminology. The density of sites relative to some initial site is characterized by the same joint distribution function as for the $d = 1$ network,

$$\rho(r,\epsilon) = 2f(\epsilon). \tag{B2}$$

Here $r = |E_n - E_m|$ is the distance between the energy levels, which is formally analogous to $r = |x_n - x_m|$ in the random site hopping model. We use here units such that the mean level spacing is unity. In the later numerical analysis we assume equally spaced levels such that the distance is simply $r = |n - m|$.

In the physical context the band profile $B(r)$ is determined by the semiclassical limit, while the distribution of the ϵ values is implied by the intensity statistics of the matrix elements. This intensity statistics is known as Porter-Thomas in the strongly chaotic case, corresponding to the Gaussian ensembles, but it becomes log-wide for systems with "weak quantum chaos" [19], reflecting the sparsity that shows up in the limiting case of integrable system [32].

In the numerical analysis we have considered simple banded matrices for which $B(r) = 1$ for $r \leqslant b$, and zero otherwise. Accordingly $1 + 2b$ is the bandwidth. The elements within the band are log-box distributed: this means that ϵ is distributed uniformly over a range $[0,\sigma]$. Note that log-box distribution is typical of glassy systems, where the tunneling rate depends exponentially on the distance between the sites.

APPENDIX C: NUMERICAL EXTRACTION OF D

In a diffusive system the coarse-grained spreading is described by the standard diffusion equation, with an evolving Gaussian distribution

$$\rho(x;t) = \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi S_x(t)}} \exp\left[-\frac{x_i^2}{2S_x(t)}\right], \tag{C1}$$

where $S_x(t) = 2Dt$. It follows from this expression that

$$S(t) = \langle r^2(t)\rangle = (2d)Dt. \tag{C2}$$

Starting with all the probability concentrated in one unit cell we get for the survival probability

$$\mathcal{P}(t) \sim \frac{r_0^d}{(4\pi\, Dt)^{d/2}}. \tag{C3}$$

The eigenvalues of the diffusion equation are

$$\lambda_k = Dq_k^2, \quad k = \text{index}, \tag{C4}$$

where the possible values of the momentum are determined by the periodic boundary conditions as $q = (2\pi/L)\vec{k}$. It follows that the cumulative number of eigenstates per site is

$$\mathcal{N}(\lambda) = \left(\frac{r_0}{2\pi}\right)^d \frac{\Omega_d}{d}\left[\frac{\lambda}{D}\right]^{d/2}. \tag{C5}$$

It is well known that the survival probability is related to the eigenvalues of $\boldsymbol{w}$ through the relation

$$\mathcal{P}(t) = \frac{1}{N}\sum_{\lambda} e^{-\lambda t} \equiv \int_0^\infty g(\lambda)d\lambda\; e^{-\lambda t}. \tag{C6}$$

For a diffusive system one can verify that the expressions above for $g(\lambda)$ and $\mathcal{P}(t)$ are indeed related by a Laplace transform. More generally, it follows that D can be deduced from the asymptotic behavior of $g(\lambda)$ in the $\lambda \to 0$ limit where the diffusive description is valid. In contrast to that for large λ, we expect $g(\lambda)$ to coincide with the distribution of the decay rates $\gamma_n = \sum_m w_{mn}$, reflecting localized modes.

APPENDIX D: RELATION TO DEBYE MODEL

Consider a system of unit masses that are connected by springs. Once can describe the system by a matrix $\boldsymbol{w}$ whose off-diagonal elements w_{nm} are the spring constants. The eigenfrequencies are determined accordingly, namely, $\omega_k = \sqrt{\lambda_k}$. Assuming that the low lying modes are like acoustic phonons with dispersion $\omega = c|q|$, where c is the so-called speed of sound, one deduces that

$$\omega_k = c|q_k|, \quad k = \text{index}, \tag{D1}$$

Consequently the associated counting function is as in the Debye model:

$$\mathcal{N}(\omega) = \left(\frac{r_0}{2\pi c}\right)^d \frac{\Omega_d}{d}\,\omega^d. \tag{D2}$$

Comparing the above expressions with Eqs. (C4) and (C5) it follows that the calculation of c^2 is formally the same as the calculation of D.

APPENDIX E: RESISTOR NETWORK CALCULATION

The diffusion coefficient D is formally like the calculation of the conductivity of the network. Therefore it can be determined via a numerical solution of a circuit equation. It is convenient to use the language of electrical engineering to explain how the resistor network calculation is carried out in practice. Accordingly we use in this appendix the notation $\boldsymbol{G}$ instead of $\boldsymbol{w}$ for the matrix that describes the resistor network, and σ instead of D for its conductivity. We define a vector $\boldsymbol{V} = \{V_n\}$, where V_n is the voltage at node n, analogous to p_n. We also define a vector $\boldsymbol{I} = \{I_n\}$ of injected currents. The Kirchhoff equation [Eq. (1)] for a steady state can be written as $\boldsymbol{GV} = 0$.

If the nodes were connected to external "reservoirs" the Kirchhoff equation would take the form $\boldsymbol{GV} = \boldsymbol{I}$. The matrix $\boldsymbol{G}$ has an eigenvalue zero which is associated with a uniform voltage eigenvector. Therefore, it has a pseudo-inverse rather than an inverse, and consequently the Kirchhoff equation has a solution if and only if the net current is $\sum_n I_n = 0$.

For the purpose of calculating the conductivity we add a source $I_1 = -1$ and a drain $I_2 = 1$. We select the location of the source (site 1) and the drain (site 2) away from the endpoints. From the solution of the Kirchhoff equation we deduce

$$\sigma[d=1] = [(V_2 - V_1)/L]^{-1}, \tag{E1}$$

where L is the distance between the contacts.

With regard to the quasi-one-dimensional model, we take the distance between the contacts to be $L' = N/2$ and look at the voltage drop along an inner segment of length $L = L' - 2b$, to avoid the transients at the contact points.

To find the conductivity in the $d = 2$ case we select contacts points that have distance $L \sim (N/2)^{1/2}$, and use the formula

$$\sigma[d=2] = [(V_2 - V_1)/\ln(L/\ell)]^{-1}, \tag{E2}$$

where $\ell \sim 1$ is the shift of the measurement point from the contact point. Here the voltage drop is divided by $\ln(L/\ell)$ instead of L, reflecting the two-dimensional geometry of the flow.

APPENDIX F: CALCULATION OF THE ERH INTEGRAL

The calculation of the ERH integral for the random site model involved the incomplete Γ function [33]

$$\Gamma(\ell+1,x) = \int_0^x r^\ell e^{-r} dr = \ell!\,\mathrm{EXP}_\ell(x)\,e^{-x}. \tag{F1}$$

We first consider the degenerate Mott model. We substitute in Eq. (30), the $w(r,\epsilon)$ of Eq. (4), and the $\rho(r,\epsilon)$ of Eq. (12) with Eq. (14). Thanks to the $\delta(\epsilon)$ we are left just with a dr integration that is split into the domains $0 < r < r_c$ and $r > r_c$. Namely,

$$\begin{aligned} D_{\mathrm{ERH}} &= \frac{w_0\Omega_d}{2d}\int_0^{r_c} e^{-r_c/\xi}\frac{r^{d+1}}{r_0^d}dr \\ &\quad + \frac{w_0\Omega_d}{2d}\int_{r_c}^{\infty} e^{-r/\xi}\frac{r^{d+1}}{r_0^d}dr \\ &= \frac{w_0\Omega_d}{2d}e^{-r_c/\xi}\frac{r_c^{d+2}}{d+2}\frac{1}{r_0^d} \\ &\quad + \frac{w_0\Omega_d}{2d}\frac{\xi^{d+2}}{r_0^d}\Gamma\left(d+2,\frac{r_c}{\xi}\right) \\ &= \frac{w_0\Omega_d\xi^{d+2}}{2d(d+2)r_0^d}\Gamma\left(d+3,\frac{r_c}{\xi}\right). \end{aligned} \tag{F2}$$

This leads directly to Eq. (40) with Eq. (39).

Turning to the nondegenerate Mott model we have to deal with a two-dimensional integral $dr\,d\epsilon$ that has, as in the previous case, two domains $w > w_c$ and $0 < w < r_c$. The two domains are separated by the line $\epsilon + (r/\xi) = \epsilon_c$. It is therefore natural to change variables:

$$x = \epsilon + (r/\xi), \tag{F3}$$

$$y = \tfrac{1}{2}[-\epsilon + (r/\xi)], \tag{F4}$$

hence

$$\begin{aligned} D_{\mathrm{ERH}} &= \frac{w_0\Omega_d}{2dr_0^d}\int_0^{\epsilon_c}\xi dx\int_{-x/2}^{x/2}dy\,e^{-\epsilon_c}\left(\xi y+\xi\frac{x}{2}\right)^{d+1} \\ &\quad + \frac{w_0\Omega_d}{2dr_0^d}\int_{\epsilon_c}^{\infty}\xi dx\int_{-x/2}^{x/2}dy\,e^{-x}\left(\xi y+\xi\frac{x}{2}\right)^{d+1} \\ &= \frac{w_0\Omega_d}{2dr_0^d}\xi^{d+2}e^{-\epsilon_c}\frac{\epsilon_c^{d+3}}{(d+2)(d+3)} \\ &\quad + \frac{w_0\Omega_d}{2dr_0^d(d+2)}\xi^{d+2}\,\Gamma(d+3,\epsilon_c) \\ &= \frac{w_0\Omega_d\xi^{d+2}}{2d(d+2)(d+3)r_0^d}\Gamma(d+4,\epsilon_c). \end{aligned} \tag{F5}$$

This leads directly to Eq. (47) with Eq. (39).

[1] S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. E **68**, 046109 (2003).

[2] S. Bradde, F. Caccioli, L. Dall'Asta, and G. Bianconi, Phys. Rev. Lett. **104**, 218701 (2010).

[3] J.-P. Bouchaud and A. Georges, Phys. Rep. **195**, 127 (1990).

[4] H. Scher and E. W. Montroll, Phys. Rev. B **12**, 2455 (1975).

[5] D. Kaya, N. L. Green, C. E. Maloney, and M. F. Islam, Science **329**, 656 (2010).

[6] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (CRC, Boca Raton, FL, 1994)

[7] B. I. Halperin, Physica D **38**, 179 (1989).

[8] S. R. Nagel, A. Rahman, and G. S. Grest, Phys. Rev. Lett. **47**, 1665 (1981).

[9] W. Schirmacher and M. Wagener, Philos. Mag. B **65**, 607 (1992).

[10] A. Amir, Y. Oreg, and Y. Imry, Phys. Rev. Lett. **105**, 070601 (2010); Phys. Rev. B **77**, 165207 (2008).

[11] N. F. Mott, Philos. Mag. **22**, 7 (1970); N. F. Mott and E. A. Davis, *Electronic Processes in Non-crystalline Materials* (Clarendon, Oxford, 1971).

[12] A. Miller and E. Abrahams, Phys. Rev. **120**, 745 (1960).

[13] V. Ambegaokar, B. Halperin, and J. S. Langer, Phys. Rev. B **4**, 2612 (1971).

[14] M. Pollak, J. Non-Cryst. Solids **11**, 1 (1972).

[15] B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Springer-Verlag, Berlin, 1984).

[16] A. Aharony, E. I. Hinrichsen, A. Hansen, J. Feder, T. Jossang, and H. H. Hardy, Physica A **177**, 260 (1991).

[17] L. Hinrichsen, A. Aharony, J. Feder, A. Hansen, T. Jossang, and H. H. Hardy, Transp. Porous Media **12**, 55 (1993).

[18] S. Bandopadhyay, Y. Etzioni, D. Cohen, Europhys. Lett. **76**, 739 (2006); D. Cohen, Phys. Rev. B **75**, 125316 (2007); A. Stotland, T. Kottos, and D. Cohen, *ibid.* **81**, 115464 (2010).

[19] A. Stotland, D. Cohen, and N. Davidson, Europhys. Lett. **86**, 10004 (2009); A. Stotland, L. M. Pecora, and D. Cohen, *ibid.* **92**, 20009 (2010); Phys. Rev. E **83**, 066216 (2011).

[20] S. Alexander, J. Bernasconi, W. R. Schneider, and R. Orbach, Rev. Mod. Phys. **53**, 175 (1981).

[21] S. Havlin, D. Movshovitz, B. Trus, and G. H. Weiss, J. Phys. A **18**, L719 (1985).

[22] R. Granek and J. Klafter, Phys. Rev. Lett. **95**, 098106 (2005).
[23] S. Reuveni, R. Granek, and J. Klafter, Proc. Natl. Acad. Sci. **107**, 13696 (2010); Phys. Rev. E **81**, 040103 (2010).
[24] D. Ben-Avraham and S. Havlin, *Diffusion and Reactions in Fractals and Disordered Systems* (Cambridge University, New York, 2000).
[25] A. Klemm, R. Metzler, and R. Kimmich, Phys. Rev. E **65**, 021112 (2002).
[26] F. Camboni and I. M. Sokolov, Phys. Rev. E **85**, 050104(R) (2012).
[27] For a review and further references, see D. Cohen, arXiv:1202.5871.
[28] K. Ishii, Prog. Theor. Phys. Suppl. **53**, 77 (1973); S. John, H. Sompolinsky, and M. J. Stephen, Phys. Rev. B **27**, 5592 (1983); W. Kantelhardt and A. Bunde, Phys. Rev. E **56**, 6693 (1997); Q. Li, C. M. Soukoulis, and G. S. Grest, Phys. Rev. B **41**, 11713 (1990).
[29] A. Amir, J. J. Krich, V. Vitelli, Y. Oreg, and Y. Imry, arXiv:1209.2169.
[30] N. W. Dalton, C. Domb, and M. F. Sykes, Proc. Phys. Soc. **83**, 496 (1964).
[31] G. E. Pike and C. H. Seager, Phys. Rev. B **10**, 1421 (1974).
[32] T. Prosen and M. Robnik, J. Phys. A **26**, L319 (1993); E. J. Austin and M. Wilkinson, Europhys. Lett. **20**, 589 (1992); Y. Alhassid and R. D. Levine, Phys. Rev. Lett. **57**, 2879 (1986); Y. V. Fyodorov, O. A. Chubykalo, F. M. Izrailev, and G. Casati, *ibid.* **76**, 1603 (1996).
[33] *NIST Handbook of Mathematical Functions*, edited by F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark (Cambridge University Press, New York, NY, 2010), Eqs. (8.4.8) and (8.4.11), http://dlmf.nist.gov.