Typical and large-deviation properties of minimum-energy paths on disordered hierarchical lattices

O. Melchert\textsuperscript{a} and A.K. Hartmann\textsuperscript{b}

Institut für Physik, Universität Oldenburg, Carl-von-Ossietzky Strasse, 26111 Oldenburg, Germany

Received 20 March 2013 / Received in final form 3rd June 2013
Published online 15 July 2013 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2013

Abstract. We perform numerical simulations to study the optimal path problem on disordered hierarchical graphs with effective dimension $d_{\text{eff}} \approx 2.32$. Therein, edge energies are drawn from a disorder distribution that allows for positive and negative energies. This induces a behavior which is fundamentally different from the case where all energies are positive, only. Upon changing the subtleties of the distribution, the scaling of the minimum energy path length exhibits a transition from self-affine to self-similar. We analyze the precise scaling of the path length and the associated ground-state energy fluctuations in the vicinity of the disorder critical point, using a decimation procedure for huge graphs. Further, using an importance sampling procedure in the disorder we compute the negative-energy tails of the ground-state energy distribution up to 12 standard deviations away from its mean. We find that the asymptotic behavior of the negative-energy tail is in agreement with a Tracy-Widom distribution. Further, the characteristic scaling of the tail can be related to the ground-state energy fluctuations, similar as for the directed polymer in a random medium.

1 Introduction

Many problems in physics and computer science can conveniently be modeled using graphs. Thereby it is often inevitable to assign attributes to the edges that assist in specifying the problem under consideration. E.g., weighted graphs, where a weight is associated with each edge, might be used to model disordered environments. For a given weighted graph the minimum-energy path (MWP) problem refers to the paradigmatic optimization problem of finding a simple (i.e. loopless) path, connecting two distinguished nodes of the graph, along which the sum of the edge weights is minimal. The MWP problem quite naturally lends itself to study a multitude of lattice-path models in the context of disordered systems. In this regard, it has proven to be useful in order to characterize, e.g., linear polymers in random media \cite{1–5}, domain wall excitations in disordered environments such as spin glasses \cite{6–8} and the solid-on-solid model \cite{9}. Due to this relation to physical problems, the weight will be denoted as energy in the following. If the disorder is drawn from a distribution that allows for nonnegative energy only, as for the canonical “directed polymer in a random medium” (DPRM), the groundstate configuration of the polymer can be computed efficiently using Dijkstra's algorithm \cite{10,11}. However, if the disorder distribution allows for edge-energies of either sign, as for the problem of finding a minimum energy domain wall in 2D Ising spin glasses \cite{7,8,12} (given that there is no closed path with a negative energy) or more generally for the negative-weight percolation (NWP) problem \cite{13–18}, the solution of the MWP problem requires a nontrivial transformation to an auxiliary minimum-weight perfect matching problem \cite{19}. Furthermore, the properties of MWPs with negative edges are fundamentally different from the case where all edge energies are non-negative \cite{7,8,12}.

To specify the NWP more precisely, one considers, say, a regular $d = 2$ square lattice graph with side length $L$ and free boundaries in one direction, periodic boundaries in the remaining direction, and energies drawn from a distribution that allows for edge energies of either sign. The details of the energy distribution are controlled by a tunable disorder parameter. For a given realization of the disorder one might be interested in, say, an agent “harvesting” the negative energies (seeing as a negative cost, i.e., a resource, e.g., an energy) along a freely adjustable path between two given points. This means the walker might have to use edges with positive energies as well, i.e., spend some amount of the resource. In addition some resources might be harvested, in parallel or in competition to the walker, by other walkers which are not restricted to walk between two given endpoints. These other walkers are only present for walks where the amount of the harvested resource is larger than the amount of the resource spent. Mathematically this means we consider configurations consisting of a single path and a set of loops, i.e., closed paths, such that the total sum of the energies assigned to the edges that build up the path and the loops attains a minimum. As
an additional optimization constraint the path might be forced to span the lattice along the direction with the free boundaries. This means, the walker covers a large fraction of the lattice which allows him to maximize the amount of the harvested resource. Further the path and the loops are simple and are not allowed to intersect each other. Therefore, they exhibit an “excluded volume” quite similar to usual self avoiding walks (SAWs) [20]. A pivotal observation is that the NWP model features a disorder driven, geometric phase transition, signaled by the emergence of paths that span the lattice along the direction with the periodic boundary conditions. In this regard, depending on the disorder parameter, one can identify two distinct scaling regimes: (i) a phase where the paths tend to be short in length, displaying a self-affine scaling with system size (see Fig. 1a), and, (ii) a phase where the paths tend to be long and exhibit a self-similar scaling (see Figs. 1b and 1c). From the previous analyses for \( d = 2 \), we found that right at the critical point, the paths are self-similar with a fractal dimension \( d_f = 1.268(1) \) (see Refs. [13,17]).

Considering real-space renormalization group schemes in the context of spin models it was realized more than three decades ago that, albeit these renormalization schemes are approximate on regular hypercubic lattices they turn out to be exact if the model is studied on certain hierarchical graphs [21–25]. Here, we study a particular MWP problem (which closely resembles the NWP problem) in a Migdal-Kadanoff-like renormalization group scheme on hierarchical lattice graphs (constructed using a “Wheatstone bridge” elementary cell (see Fig. 2) with an effective dimension \( d_{\text{eff}} \approx 2.32 \) [26–28], where an exact decimation procedure can be used to analyze huge graphs [29,30]. We address the critical behavior of the MWP in this setup by monitoring observables related to the path energy and path length. The subtleties of the construction procedure that leads to hierarchical lattices with effective dimension \( d_{\text{eff}} \approx 2.32 \) even allows to prove the transition from the self-affine to self-similar scaling of the path length, as observed for the NWP problem on hypercubic lattice graphs [13,15].

Similar to previous studies of minimum-energy path problems on hierarchical graphs [3,6,26,31–36] and regular lattices [1,2,5,10,37,38], we here consider the finite-size scaling of the length and energy of the paths, as well as the associated energy fluctuations. We further complement the simple sampling (SiSa) estimates of the path-energy distribution by an importance sampling (ImSa) procedure in the disorder [39–43], allowing to resolve the respective distributions up to 12 standard deviations away from its mean. In this regard it is found that the asymptotic behavior of the negative-energy tail is in agreement with a Tracy-Widom distribution. Further, the characteristic scaling of the tail can be related to the path energy fluctuations, similar as for the directed polymer in a random medium [32]. Note that, apart from the disorder distribution, the MWP problem considered here is similar to the optimal path problem on hierarchical lattices as studied in reference [26] (therein, the authors considered a uniform distribution of nonegative edge-energies, only).

Above we pointed out that the MWP problem studied here closely resembles the NWP problem. At this point we would like to point out the major similarities and differences of both models: similar to the NWP problem, in the MWP problem the sum of energies of edges that build up a path is object to minimization. As a major difference note that the path found in the context of the NWP problem is not necessarily the (absolute) minimum-energy path. The reason is that in the NWP problem, a global minimum of the energy for a single path plus a (possibly empty) set of loops (all with negative energy) is searched for (see Fig. 1). This is in contrast to the MWP problem on hierarchical lattices, where a particular decimation scheme (see Sect. 2), allows to obtain a truly minimum-energy path in a framework where no loops are considered. However, also note that at the critical point of the NWP model (see Fig. 1b), the “additional” loops are small and resemble a rather dilute “gas” of loops which are unlikely to affect the statistics of the path that spans the system in between the free boundaries. Hence, in the vicinity of the critical point of the model we expect the MWP problem studied here to provide a reasonable approximation to the path in the NWP problem.

The remainder of the presented article is organized as follows. In Section 2, we explain the construction procedure to obtain the hierarchical lattice graphs and we outline the pool method used compute the properties of the paths for huge graphs. In Section 3, we list the results of our numerical simulations in terms of which we locate the self-affine to self-similar transition of the path length and where we put under scrutiny the path-energy distribution. In Section 4 we conclude with a summary.

## 2 Model and algorithm

The hierarchical lattices considered in the remainder of the presented article can be constructed using a simple deterministic rule. This rule specifies how the individual edges in a graph at a given iteration step \( I \) need to be transformed in order to obtain a graph at iteration step \( I + 1 \). Let \( G_I = (V_I, E_I) \) denote a hierarchical graph that consists of a set of

---

**Fig. 1.** Samples of minimum-energy configurations consisting of one path (forced to span the lattice along the direction with the free boundaries) and a set of loops for a 2D square lattice with side length \( L = 64 \) and periodic boundaries in the horizontal direction. The snapshots relate to different values of the disorder parameter \( \rho \), where (a) \( \rho < \rho_c \), (b) \( \rho = \rho_c \), and, (c) \( \rho > \rho_c \). In the limit of large system sizes and above the critical point \( \rho_c \), paths might span the lattice along the direction with the periodic boundaries.
nodes \( i \in \mathcal{V}_I \) and a set \( \mathcal{E}_I \subset \mathcal{V}_I^2 \) of undirected edges \( e = \{i,j\} \in \mathcal{E}_I \). The number of edges is given by \( M_I = |\mathcal{E}_I| \). The transformation in order to proceed from \( G_I \) to \( G_{I+1} \) reads as follows: each edge \( e = \{i,j\} \in \mathcal{E}_I \) is replaced by a subgraph \( G' \) consisting of four nodes \( \{i,k_0^{(ij)},k_1^{(ij)},j\} \) (therefore, the set of nodes needs to be amended by two nodes \( k_0^{(ij)} \) and \( k_1^{(ij)} \)) and five edges \( \{i,k_0^{(ij)}\}, \{i,k_1^{(ij)}\}, \{k_0^{(ij)},k_1^{(ij)}\}, \{k_0^{(ij)},j\}, \{k_1^{(ij)},j\} \). The nodes \( i \) and \( j \) are referred to as the terminal nodes of the subgraph. After the transformation is completed the number of edges increased to \( M_{I+1} = 5 \times M_I \), and the linear extension of the graph has doubled, i.e. \( L_{I+1} = 2 \times L_I \). At \( I = 0 \), the construction procedure is started with a single edge, meaning that \( M_0 = 1 \) and \( L_0 = 1 \). Hence, \( M_{I} = 5^I \) and \( L_I = 2^I \). From the increase of the number of edges \( M_{I} \) as a function of the linear extension \( L_I \) of the graphs according to \( M_{I} = 2^I \log_5(5) = L_I^{\text{eff}} \) it is possible to obtain the effective (fractal) dimension of the hierarchical lattices as \( d_{\text{eff}} = \log_5(5) \approx 2.32 \). The construction procedure is illustrated in Figure 2, where, starting with a single edge at \( I = 0 \), the two steps \( G_0 \rightarrow G_1 \rightarrow G_2 \) are shown explicitly. Finally, a path is represented by an ordered set of edges. E.g., regarding the subgraph \( G' \), a possible path that connects its terminal nodes \( i \) and \( j \) reads \( p = \{(i,k_0^{(ij)}),(k_0^{(ij)},k_1^{(ij)}),(k_0^{(ij)},j),(k_1^{(ij)},j)\} \).

The minimum-energy path problem we address here reads as follows. Let \( s, t \) denote the endnodes of the single edge at \( I = 0 \). Perform a number of \( I_{\text{max}} \) iteration steps to yield a hierarchical graph \( G_{I_{\text{max}}} \) and assign a random energy to each edge, drawn from a given disorder distribution. Finally, compute a minimum energy \( s-t \) path for the graph \( G_{I_{\text{max}}} \). The precise topology of the resulting path depends on the particular realization of the disorder and has length \( \ell \in [2^{I_{\text{max}}}, 3^{I_{\text{max}}}]. \) Bear in mind that in order to compute one such path, a graph with \( M_{I_{\text{max}}} = 5^{I_{\text{max}}} \) edges needs to be constructed. Consequently, a number of \( M_{I_{\text{max}}} \) random deviates need to be drawn from the disorder distribution. A more efficient way to sample minimum energy \( s-t \) paths for the case of hierarchical graphs at large values of \( I \) is provided by the pool method [29,30]. Therein one maintains a set of \( I_{\text{max}} \) pools \( P_I \) of (effective) edges, with \( I = 0, \ldots, I_{\text{max}} - 1 \). The number of edges in each pool is the same and is denoted by \( N \). An individual edge carries two attributes \( e = (E, \ell) \), where \( E \) denotes the energy and \( \ell \) the length of a path associated with the edge. At \( I = 0 \) the edges are initialized with \( \ell = 1 \) and the energies \( E \) are drawn from a specified disorder distribution \( P_0(E) \), signifying the “single edge level”. In order to proceed from pool \( P_I \) to \( P_{I+1} \), the following three-step decimation procedure, sketched in Figure 3, has to be repeated \( N \) times:

(i) pick five edges \( e_1, \ldots, e_5 \) at random from pool \( P_I \). Combine these to form a subgraph \( G' \) as explained earlier;

(ii) from the four distinct paths that connect the terminal nodes of \( G' \), determine the minimal energy path \( p^* \), i.e. the path \( p^* \in \{(e_1,e_2),(e_3,e_4),(e_1,e_5,e_4),(e_3,e_5,e_2)\} \) for which \( E^* = \sum_{e \in p^*} E(e) \equiv \min \). Correspondingly, the length of the path reads \( \ell^* = \sum_{e \in p^*} \ell(e) \);

(iii) set up a new edge having attributes \( e = (E^*, \ell^*) \) and add it to pool \( P_{I+1} \).

After \( P_0 \) has been initialized, all pools up to \( I = I_{\text{max}} \) might be filled in this manner. Note that an edge \( e \in P_I \) effectively corresponds to a hierarchical graph \( G_I \), i.e. it has a “hidden” substructure that allows to represent a minimum-energy path with length \( \ell \in [2^I, 3^I] \). The attributes of the edge encode the characteristics of the respective path, i.e. its energy \( E \) and length \( \ell \). A pool \( P_I \), thus, consists of \( N \) instances of minimum-energy paths for hierarchical graphs at iteration step \( I \). Thereby, the computational resources needed to fill a pool stay constant as \( I \) increases.

Further, each pool specifies its own distributions \( P_I(E) \) and \( P_I(\ell) \) of path energies and path lengths, respectively. These allow to quantify the scaling behavior of the average path length with system size as \( \langle \ell \rangle \sim L_I^{\nu(E)} \), defining the fractal dimension \( d_f \) of the paths, the average path energy \( \langle E \rangle \sim L_I^{\nu(E)} \), and the fluctuation of the path energies as \( \var(E) = \langle E^2 \rangle - \langle E \rangle^2 \sim L_I^{2\nu(\var)} \). Note that these energy fluctuations are measured with respect to the linear extension of the considered lattice graphs. This can be compared the ground-state energy fluctuations of the DPRM, which are commonly measured for polymers of a fixed length \( L \), giving rise to the fluctuation exponent \( \omega \) defined as \( \var(E_{GS}) \sim L^{2\omega} \) (see Ref. [3]). Hence, to compare our results with the DPRM case, one should rewrite the energy fluctuations as a function of the average path length \( \ell \). This leads to a corresponding estimate of \( \omega \) via using the relation \( \omega = \Omega/d_f \).
In the following section, we will use the algorithmic procedure outlined above in order to study the minimum-energy path problem for an increasing number of iteration steps $\ell$ and for a large range of values of the disorder parameter $\rho$.

3 Results

In the presented study, the initial disorder distribution $P_0(E)$, used to set the edge-energies in the initial pool $P_0$, is a Gaussian with mean $\mu$ and width $\sigma^2 = 1$, i.e.,

$$P_0(E) = \exp \left[-(E-\mu)^2/2\right]/\sqrt{2\pi}.$$  

(1)

A tunable disorder parameter is defined as $\rho = 1/\mu$, so that the standard normal distribution is recovered in the limit $\rho \to \infty$. Typical values for the pool-size and iteration steps are $N = 10^6$ and $I_{\text{max}} = 30$, respectively.

To facilitate intuition, note that as $\mu \to \infty$, typical minimum-energy paths will exhibit positive energies. Thus, an increasing path length will lead to an increasing path energy. Hence, a minimum-energy path will tend to be short in length. Considering hierarchical lattice graphs $G_I$, one might consequently expect a scaling behavior $\langle \ell \rangle \propto 2^I$, implying a scaling exponent $d_f = 1$. On the other hand, as $\mu \to -\infty$, typical minimum-energy paths will exhibit negative energies. Therefore, an increasing path length results in a decreasing path energy, leading to expect $\langle \ell \rangle \propto 3^I$, and therefore $d_f = \log_3(3) \approx 1.585$. In between these two extremal “trivial” cases, there exists a particular value $\rho_c = 1/\mu_c$ of the disorder parameter that signifies the onset of “proliferation”, where $\langle E \rangle = 0$ as $\rho \to \infty$. Here, the average path length exhibits a non-trivial scaling behavior displayed by a scaling exponent $1 < d_f < 1.585$.

At first we attempt to estimate the value of $\rho_c$ by means of the secant method [44], considering different initial pools at a given value of $I$. Next, we consider one particular initial pool to quantify the scaling behavior of the average minimum-energy path length and energy. Finally, we put under scrutiny the probability density of minimum-energy path energies.

3.1 Location of the critical point where the minimum-energy path energy vanishes

In order to approximate the critical point $\rho_c$ we considered pools of size $N = 10^6$ and $I \leq 22$. So as to arrive at an estimate of $\rho_c$ at iteration step $\ell$ we proceeded as follows: using the secant method we prepared a number of $M = 32$ independent estimates of effective critical values $\rho_{\text{eff}}^{(i)}$, where $\langle E \rangle \approx 0$ at the considered value of $I$. The distribution of these effective critical values (see inset of Fig. 4) is characterized by the average $\rho_c(\ell) = (1/M) \sum_{i=1}^{M} \rho_{\text{eff}}^{(i)}$. E.g., at $I = 22$ we yield $\rho_c(\ell) = 0.8435(7)$, wherein the standard deviation among the $M$ independent estimates reads $\sigma_{\rho_c} = 0.004$. The averages exhibit the scaling behavior $\rho_c(\ell) = \rho_c + a 2^{-\ell b}$, where a fit [45] to the interval $I \in [5, 22]$ yields $\rho_c = 0.8436(2)$, $a = O(1)$, and $b = 0.866(4)$ (see Fig. 4). The results did not depend much on the pool size, e.g. considering $N = 10^6$ and proceeding as above we find $\rho_c = 0.8435(1)$ and $b = 0.868(5)$.

3.2 Trivial to non-trivial transition of the average minimum-energy path length

As it appears, for large values of $I$ and $\rho_c \approx 0.84$ one should observe a vanishing average path energy. In the presented subsection we consider a single pool of size $N = 10^6$ and $I \leq 30$ in order to assess the scaling behavior of the minimum-energy path length and energy. For that pool we find that the average path energy changes its sign at $\rho = 0.8367$ (see inset of Fig. 5a). Further, at that approximate value the scaling behavior of the fluctuations related to the path energy and length change significantly (see Figs. 5a and 5b). Note that the shape of the distributions $P(E)$ and $P(\ell)$ of the path energies and path lengths, respectively, change during the iteration steps. E.g., the initial distribution of path energies $P(E)$ is a Gaussian with mean $1/\rho$ and unit width. At large values of $I$ it attains a limiting shape consistent with a Tracy-Widom scaling form (as shown in Sect. 3.3) (see Fig. 6a). Initially, the distribution of path lengths is a delta-distribution centered at $\ell = 1$. After one iteration, the probability mass function of path lengths in pool $P_1$ has two nonzero contributions at $\ell = 2$ and $\ell = 3$ (see Fig. 6b) with statistical weight $p(\ell = 2) = 0.8224$ and $p(\ell = 3) = 0.1776$ and $\langle \ell \rangle = 2.1776(4)$. During further iterations, the shape of the distribution of path lengths changes: while for iteration steps $I = 2, \ldots, 15$ it exhibits an exponential decay in...
the large-ℓ tail, the rescaled distribution attains the limiting shape of a standard normal distribution in the limit of large I (see Fig. 6b). Also note that the summary measures related to the distributions depend on I and ρ. E.g., considering $P_I(\ell)$, the scaling behavior of the mean and width are described by the exponents $d_E$ and $\Omega$ through the scaling relations $\langle \ell \rangle \propto 2^{d_E I}$ and $\text{var}(\ell) \propto 2^{d_E I \Omega}$. The precise values of the scaling exponents then depend on the model parameter $\rho$ at which the simulations are performed (see discussion below).

The precise location of the peak position related to var(ℓ) (see Fig. 5b) can be used to define an iteration-step dependent effective critical point $\rho_{\text{eff}}(I)$. Seen as a function of $I$, these effective critical points can be used to pinpoint the precise location where the proliferation transition of the path length occurs in the limit $I \rightarrow \infty$. The effective critical values exhibit a scaling of the form $\rho_{\text{eff}}(I) = \rho_c + a 2^{-b I}$, where a fit to the interval $I \in [10, 20]$ yields the estimates $\rho_c = 0.836688(1)$, $a = O(1)$, and $b = 0.76(2)$ (see inset of Fig. 5b). As pointed out above, for $\rho < \rho_c$ an increase in path length most likely results in an increasing path energy. Hence, one can expect that for $\rho < \rho_c$ the minimum-energy-path problem investigated here effectively corresponds to the optimal path problem on hierarchical lattices in the limit of “weak” disorder, studied in reference [26] via numerical simulations (in Ref. [26], “weak” disorder is governed by the objective-function intrinsic parameter $\alpha = 1$). In that study, a nonequilibrium uniform disorder distribution was considered and the scaling behavior of the path energy and length at different values of $\alpha = 1, \ldots, \infty$ was analyzed (the extremal choice $\alpha = \infty$ signifies the so-called limit of “strong” disorder which is also well studied on regular lattice graphs, see, e.g., Refs. [38,46]). In particular, at $\alpha = 1$ it was found that $\Omega = 0.30(1)$ and $d_f = 1$ (note that in Ref. [26] the exponents are called $\omega$ and $D_f$, respectively). From this it is immediate to expect $d_f = 1$, $d_E = 1$ and $\Omega = 0.3$ for $\rho < \rho_c$. From a direct fit to the scaling
forms $\langle \ell \rangle \propto 2^{d_I}$, $\langle E \rangle \propto 2^{d_E}$ and $\text{var}(E) \propto 2^{2d_I}$ we obtained the numerical values for the scaling exponents $d_f$, $d_E$ and $\Omega$ (and consequently the “corrected” exponent $\omega$) as shown in Figure 7 and listed in Table 1. For values of $\rho$ below and above the critical point, the fits were restricted to large iteration steps, i.e. $I \in [20, 30]$. However, note that it is difficult to prepare a system right at $\rho_c$ as $I$ increases, fluctuations will eventually cause the system to assume the asymptotic scaling behavior characteristic for $\rho < \rho_c$ or $\rho > \rho_c$. Hence, in order to obtain the scaling exponents for $\rho \approx \rho_c$, the fitting procedure was restricted to intermediate iteration steps $I \in [5, 10]$ ($I \in [5, 15]$ in case of $\langle E \rangle$), only. Anyway, for the case of the energy, we actually require $\langle E \rangle \approx 0$ at $\rho = \rho_c$, hence the value of $d_E$ right at the critical point is of limited relevance and a pure numerical artifact. Further, note that in the vicinity of $\rho_c$, the correlations that build up in pools of finite size might be rather strong and biased. As a remedy, one might e.g. consider a “pool refreshing” approach, similar to the one presented in reference [30], where, in order to obtain subgraphs at iteration step $I$, the subgraphs at previous iteration steps are also updated systematically.

Another means to quantify the scaling behavior of the above observables is given by the local scaling exponents. E.g., denoting the average minimum-energy path length at a given value of $\rho$ and iteration step $I$ as $\langle \ell(\rho) \rangle_I$, one can obtain the local analog to the fractal dimension as $d_{loc}^{\ell}(\rho) = \log_2(\langle \ell(\rho) \rangle_{I+1}/\langle \ell(\rho) \rangle_{I})$. The respective error can be obtained via error propagation as $\delta d_{loc}^{\ell} = \delta(\langle \ell(\rho) \rangle_{I+1}/\langle \ell(\rho) \rangle_{I}) + \delta(\langle \ell(\rho) \rangle_{I+1}/\langle \ell(\rho) \rangle_{I+1})$. The resulting local scaling exponents are shown in Figure 8. The local equivalent $f_{loc}^{\ell}$ of the energy fluctuation exponent can be computed in similar manner (see inset of Fig. 8).

For a given value of $\rho$ and for increasing $I$, the local exponents become independent of $I$ and tend to a limiting value that is within error bars in agreement with the numerical values of $d_f$ and $\Omega$ listed in Table 1.

### Table 1. Critical exponents that characterize the self-affine to self-similar transition of the minimum-energy path length on hierarchical lattice graphs.

| $\rho$ | $d_f$ | $d_E$ | $\Omega$ | $\omega$ |
|-------|-------|-------|---------|---------|
| $\rho < \rho_c$ | $1$ | $0.300(4)$ | $0.300(4)$ | |
| $\rho \approx \rho_c$ | $1.158(1)$ | $0.3496(2)$ | $0.347(5)$ | $0.300(5)$ |
| $\rho > \rho_c$ | $1.5849(31)$ | $1.5840(71)$ | $0.401(5)$ | |

### 3.3 Importance sampling results for the ground-state energy distribution

As stressed above, for a disorder parameter $\rho < \rho_c$ the minimum energy path problem considered here effectively corresponds to the optimal path problem studied in reference [26]. This is further highlighted by the probability distribution function (pdf) of minimum-energy path energies. In this regard, Figure 9 shows the simple-sampling estimate of the ground-state energy distribution for the minimum-energy path at $\rho = 0.830$ for different iteration steps $I$, obtained from pools of size $N = 10^6$. As evident...
from Figure 9, data curves corresponding to different iteration steps $I$ scale according to

$$P_I(E) = \sigma_E^{-1} P((E - \langle E \rangle)/\sigma_E), \tag{2}$$

where $\epsilon \equiv (E - \langle E \rangle)/\sigma_E$ defines a reduced energy with $\langle E \rangle$ and $\sigma_E$ describing the average and standard deviation of the distribution $P_I(E)$, respectively. This means that the scaling function $P(\epsilon)$ does not depend on the value of $I$, only. However, note that this is true for large values of $I$, only. At small values of $I$ the shape of the pdf is very well subject to changes (see discussion in Sect. 3.2 and Fig. 6a). Further, the rescaled distribution of the optimal path energy for the case were edge-energies are drawn uniformly from the interval $[0,1]$ (i.e. the case considered in Ref. [26]) is also shown (in the key, the respective symbol is marked as (u)). The dashed line indicates the guiding function fitted to the negative tail of the ground-state energy distribution.

Subsequently we address the question whether the distribution of ground-state energies in the minimum-energy path problem for $\rho < \rho_c$, $\rho \approx \rho_c$, and $\rho > \rho_c$ is consistent with a Tracy-Widom scaling form and we attempt to obtain a numerically precise estimate of the negative tail exponent $\eta$ for the above three cases. To this end, we consider an importance sampling procedure in the disorder [39,40], where the sampling process is controlled by a guiding function [41]. This allows to compute the negative-energy tails of the ground-state energy distribution up to 12 standard deviations away from its mean. Similar to reference [42], we consider the negative tails of the distribution for the reduced energy $\epsilon \leq 1$ only. We further use a guiding function

$$G(\epsilon) = \exp((a - b|\epsilon| + c|\epsilon|^3) \tag{3}$$

in order to estimate the parameters that characterize best the simple sampling distributions at the three values $\rho = 0.83$, $\rho = \rho_c$, $\rho = 0.86$. The respective estimates are listed in Table 2 and the guiding function for $\rho = 0.83$ is indicated as dashed line in Figure 9.

Let $P(\epsilon)$ describe the true probability density function of observing a minimum-energy path with reduced energy $\epsilon$ for the model under consideration (bear in mind that it holds that $P(\epsilon) = \sigma_E P_I(E)$, where $\epsilon \equiv (E - \langle E \rangle)/\sigma_E$). In order to arrive at an ImSa estimate $P^{IS}(\epsilon)$ that approximates $P(\epsilon)$, we divide the generation of the $I$ iterations into two parts, consisting of $I - \Delta I$ and $\Delta I$ iterations, where $\Delta I$ is small, we consider $\Delta I = 2, 3$ or 4. The generation of the first $I - \Delta I$ iterations, is performed in the usual simple-sampling way, leading to a large ($N = 10^6$) pool $P_{I-\Delta I}$. The final $\Delta I$ iterations should be done in a way that also the interesting tails of the distribution are sampled. For this purpose, we generate a Markov chain $P_0^{IS} \rightarrow P_1^{IS} \rightarrow P_2^{IS} \rightarrow \ldots$ of pools, which all are subsets of $P_{I-\Delta I}$, but the sampling is done in a way that also the tails of the distribution of reduced energy values are sampled.

The initial pool $P_0^{IS}$ is created by picking a (uniformly sampled) random subset of $5^{\Delta I}$ edges from $P_{I-\Delta I}$. For this pool now the final $\Delta I$ levels of the hierarchy are performed within one graph: the $5^{\Delta I}$ edges comprising the sampling pool can be arranged into one particular realization of a hierarchical graph $G_{\Delta I}$. Upon stepwise

| $\rho$ | $I$ | $a$ | $b$ | $c$ | $\eta$ | $\chi^2$/dof |
|-------|-----|-----|-----|-----|------|----------------|
| 0.83  | 20  | -0.98(4) | 0.72(5) | 0.14(8) | 1.55(3) | 1.01             |
| $\rho_c$ | 15  | -1.00(3) | 0.81(6) | 0.24(7) | 1.49(4) | 0.81             |
| 0.86  | 20  | -0.94(2) | 0.56(4) | 0.03(7) | 1.84(4) | 0.78             |
decimation $G_{\Delta I} \rightarrow G_0$ this yields one particular edge with an edge energy $E_0$ and a corresponding reduced energy $\epsilon_0 \equiv (E_0^{IS} - \langle E \rangle)/\sigma_E$, wherein $\langle E \rangle$ and $\sigma_E$ describe the simple sampling estimate of $P_I(\epsilon)$. Note that the probability in the tails is very small, hence ImSa will not change $\langle E \rangle$ and $\sigma_E$ considerably.

The Markov chain Monte Carlo step reads as follows: from a given sampling pool $P_i^{IS}$, we construct the next sampling pool $P_{i+1}^{IS}$ using the following 2-step procedure:

1. randomly choose a fraction $p$ of edges contained in the sampling pool $P_i^{IS}$ and replace those edges by new edges chosen from the large pool $P_{I-\Delta I}$. This then specifies a candidate $P'$ for the next sampling pool, characterized by the reduced energy $\epsilon'$;
2. set $P_{i+1}^{IS} = P'$ with probability

$$ P_{\text{accept}} = \min \left[ \frac{G(\epsilon)}{G(\epsilon')} \right]. \tag{4} $$

Set $P_{i+1}^{IS} = P_i^{IS}$ otherwise.

To complete the importance sampling simulation, the evolution of the initial sampling pool $P_0^{IS}$ is followed a number of $M$ steps. The resulting $M$ reduced energy values $\epsilon_0, \ldots, \epsilon_M$ comprise an auxiliary distribution $P^{IS}(\epsilon)$ that describes the probability by means of which a reduced path energy $\epsilon$ is visited within the ImSa procedure. Since the importance sampling is designed such that a sampling pool having reduced energy $\epsilon$ is encountered with probability $\propto 1/G(\epsilon)$, the distribution $P^{IS}(\epsilon)$ is further given by the ratio $P^{IS}(\epsilon) = P(\epsilon)/G(\epsilon)$ (where $P(\epsilon)$ describes the true pdf of observing a minimum-energy path with reduced energy for the considered model system). As long as the guiding function $G(\epsilon)$ provides a reasonable approximation to the true distribution of path energies, the auxiliary distribution $P^{IS}(\epsilon)$ obtained using the IS procedure is rather “flat”. Thus, regarding the target distribution $P(\epsilon)$ one might hope to improve on the negativetail statistics provided by a simple sampling approach. As discussed in reference [41], successive configurations (i.e. sampling pools) encountered during an ImSa simulation are not independent. As a remedy one might consider the autocorrelation function

$$ \chi(\Delta I) = \frac{\langle E_i E_{i+\Delta I} \rangle - \langle E_i \rangle \langle E_{i+\Delta I} \rangle}{\langle E_i^2 \rangle - \langle E_i \rangle^2} \tag{5} $$

associated to the sequence of energy values $E_i$ obtained from the importance sampling procedure. The number of Monte Carlo steps that have to elapse until the autocorrelation function decays to $1/e$ gives the respective autocorrelation time $\tau_E$. Sampling pools that are separated by $\approx \tau_E$ Monte Carlo steps can be considered effectively uncorrelated. Finally, the truncated sequence of effectively uncorrelated energy values can be analyzed similar to the simple-sampling data.

The results discussed below were obtained for the choice $p = 0.2$, where we restricted the IS procedure to $-12 \leq \epsilon \leq -1$. We further performed a number of $M = 10^7$ Monte Carlo steps to estimate the distribution of path energies in a target pool corresponding to $I = 20$. In order to assess the autocorrelation time we considered IS simulations at a disorder parameter $\rho = 0.83$ and for $\Delta I = 2, 3, 4$. The solid lines indicate a best fit of the respective data to a stretched exponential function $\chi(\Delta I) \propto \exp\left(-\Delta I/\tau_E\right)^{\beta}$. The resulting fit-parameters are listed in the text. The inset indicates the effective autocorrelation times that result from an analysis of slices of $\Delta M$ successive energy values obtained during the ImSa simulation at $\Delta I = 3$. For $\Delta M < 10^7$ the autocorrelation time appears to increase $\propto \Delta M^{0.87(3)}$ (the corresponding fit is indicated as a dashed line).

![Fig. 10.](image) The main plot shows the autocorrelation functions $\chi(\Delta I)$ for importance sampling simulations at $\rho = 0.83$ and $\Delta I = 2, 3, 4$. The solid lines indicate a best fit of the respective data to a stretched exponential function $\chi(\Delta I) \propto \exp\left(-\Delta I/\tau_E\right)^{\beta}$. The resulting fit-parameters are listed in the text. The inset indicates the effective autocorrelation times that result from an analysis of slices of $\Delta M$ successive energy values obtained during the ImSa simulation at $\Delta I = 3$. For $\Delta M < 10^7$ the autocorrelation time appears to increase $\propto \Delta M^{0.87(3)}$ (the corresponding fit is indicated as a dashed line).
Table 3. Exponent $\eta$, describing the scaling of the negative tail of the pdf $P(\epsilon)$, obtained by fitting the function $G(\epsilon)$ (see text) to the data resulting from the ImSa procedure. From left to right: value $\rho$ of the disorder parameter, interval over which the fit was performed, exponent $\eta$ (as well as $1 - 1/\eta$) and reduced chi-square $\chi^2$/dof.

| $\rho$     | $[\epsilon_{min}, \epsilon_{max}]$ | $\eta$ | $1 - 1/\eta$ | $\chi^2$/dof |
|------------|--------------------------------------|--------|---------------|--------------|
| 0.83       | $[-11, -1]$                          | 1.42(3)| 0.30(1)       | 0.88         |
| $\rho_c$   | $[-10, -1]$                          | 1.42(2)| 0.30(1)       | 1.21         |
| 0.86       | $[-12, -1]$                          | 1.65(6)| 0.39(2)       | 1.10         |

one should keep only every $\tau$th energy value. However, in reference [42], the authors concluded that if during the $M$ Monte Carlo steps the interval $[\epsilon_{min}, \epsilon_{max}]$ is crossed sufficiently often, it is not necessary to discard any energy values obtained during the IS simulation. Here, for $\Delta I = 3$, considering the interval $[-12, -1]$ and performing $M = 10^7$ Monte Carlo steps at $\rho = 0.83$ we found a number of $n_{cross} = 366$ interval crossings. In agreement with reference [42] we observed that it makes no difference whether the sequence of energy values collected during the IS procedure was truncated or not, the resulting pdf $P(\epsilon)$ remained almost unchanged (apart from effects due to the different sample sizes used to construct the pdfs).

Below we present the results obtained for ImSa simulations at $\rho = 0.83, \rho_c, 0.86$ considering $\Delta I = 3$ and a target distribution at $I = 20$ (only the simulation at $\rho_c$ was carried out for a target pool at $I = 15$). Once the distribution $P_{IS}(\epsilon)$ is obtained from the ImSa procedure, it can immediately be transformed to the desired pdf $P(\epsilon)$. A comparison of the SiSa and ImSa pdf shows that the absolute probabilities in the overlapping region $\epsilon \in [-6, -1]$ do not coincide. This is due to the restriction of the relative energies to the interval $\epsilon \in [-12, -1]$ during the ImSa simulation. One can easily account for this discrepancy by requiring that the SiSa and ImSa estimates coincide in the overlapping region, and by rescaling the ImSa estimate accordingly. In Figures 11a–11c, the resulting pdfs of observing a minimum-energy path with reduced energy $\epsilon$ for $\rho = 0.83, \rho_c, 0.86$ are shown, respectively. Note that the pdfs are represented using histograms that consist of 64 bins, each. In either case, the ImSa (SiSa) estimate is depicted for relative energies $\epsilon < -2$ ($\geq -2$). As evident from the figures, using the ImSa procedure probabilities as small as $P(\epsilon) \propto 10^{-20}$ can be reached, in contrast to $\propto 10^{-6}$ for a SiSa approach (cf. Fig. 9). The insets to Figures 11a–11c indicate the normalized deviation of a best fit of the function $G(\epsilon) = \exp \{a - b[1 - c^{|\epsilon|}]\}$ to the negative tail of $P(\epsilon)$. Once the fit is performed the deviation is obtained as $\Delta_{fit}(\epsilon) = (P(\epsilon) - G(\epsilon))/\Delta P(\epsilon)$, where $\Delta P(\epsilon)$ indicates the measurement error on $P(\epsilon)$ as obtained by bootstrap resampling [45]. The observation that $\Delta_{fit}$ (considering a fit of $G(\epsilon)$ to the ImSa data) is of order one and changes sign in an irregular fashion indicates that there are no systematic deviations and that the fit-function $G(\epsilon)$ represents a proper approximation to the negative tail of the observed pdf. In particular, Figure 11a highlights that a fit to the SiSa pdf might be misleading if one is interested
in the true scaling behavior of $P(\epsilon)$ as $\epsilon \to -\infty$. For that purpose, the above fit-function with fitting parameters obtained for the SiSa data (dashed curve in the main plot) was used to compute the normalized deviation to the ImSa data (dashed line in the inset). Referring to this one finds rather strong systematic deviations where, e.g., $\Delta_{\text{fit}}(-11) \approx 10$. The parameters $\eta$ that correspond to a best fit to the ImSa data are listed in Table 3.

As pointed out above, previous studies suggested that the exponent $\eta$ is related to the energy fluctuation exponent $\omega$ (listed in Tab. 1) by means of the expression $\omega = 1 - 1/\eta$. Here, for the minimum-energy path problem on hierarchical lattice graphs we find that this expression holds for all values of $\rho$ thus considered. To facilitate comparison, the values $1 - 1/\eta$ are listed in Table 3.

4 Conclusions

In the presented article we have investigated a particular MWP problem on MK hierarchical graphs. It is quite similar to earlier polymer $[2,3,31–34,36]$ and optimal path problems $[6,26,37,38]$. The important difference is that for a considerable fraction of negative edge energies, the total energy of a path may be reduced by taking longer paths (which leads to a different universality class). In the same fashion as the optimal path problem on hierarchical graphs, studied in reference $[26]$, corresponds to the (generic) non-directed optimal path problem $[10]$, the minimum-energy path problem studied here corresponds to the negative-energy percolation problem $[13]$ in which there is a path forced onto the system and where the disorder is “weak” enough to render the appearance of loops irrelevant for the scaling behavior of the path.

Here, the scaling properties of the MWP obtained after the decimation of huge hierarchical graphs change with increasing edge-disorder, leading from a phase where the path displays a self-affine scaling behavior to a phase where the path displays a (statistically) self-similar scaling behavior. We characterized the respective phase transition by monitoring the length and energy of the MWPs as function of the disorder and quantified the scaling behavior of the observables (and their fluctuations) by means of proper critical exponents. While the scaling of the observables off criticality can be explained intuitively, the scaling behavior found at the critical point of the model is non-trivial and compares well to the scaling observed for the optimal path problem on the same hierarchical lattice in the limit of “strong” disorder $[26]$. However, note that the precise optimization criteria of both models are slightly different: while in the optimal path problem $[10,26]$ one aims to minimize the largest energy along a sub-path, here one strives after minimizing the sum of energies along a sub-path. Note that this was already realized for the respective models on regular lattice graphs, where quite similar scaling exponents for the average path length were found in dimensions $d = 2$ through 6 (see Ref. $[15]$).

Further, we performed an importance sampling simulation for the ground-state energy distribution of the paths and confirmed that it is consistent with a Tracy-Widom scaling form, similar to the directed polymer in a random medium $[42]$. Using the importance sampling procedure allowed for an analysis of the ground-state energy distribution down to probabilities as small as $\omega \times 10^{-6}$ (in contrast, a convenient simple sampling approach only allows to reach $\omega \times 10^{-6}$), i.e. up to 12 standard deviations away from the mean of the respective distribution. This leads to a precise estimate of the scaling behavior of the negative tail of the ground-state energy distribution. For all values of the disorder parameter considered here, the respective exponents $\eta$ could be related to the energy-fluctuation exponents $\omega$ via the relation $\omega = 1 - 1/\eta$.

As presented here, the negative-weight percolation (NWP) model from which the minimum energy path problem is derived, is a generic statistical physics problem with quenched disorder. It is mainly of academic interest and has no direct applications. However, in 2D, the underlying optimization problem is related to the problem of finding groundstates for the 2D Ising spin glass (especially on the honeycomb lattice there is a one-to-one correspondence between the optimal paths studied in terms of the NWP problem and domain wall excitation in Ising spin glasses (see Ref. $[8]$), and from the critical exponents found for the 3D NWP problem a relation to the 3D strongly screened vortex glass model was suggested $[18]$.

One of the challenges of the presented study was to combine an exact decimation procedure for hierarchical lattices with an importance sampling Monte Carlo simulation in the quenched disorder. Using such a numerical approach we were able to shed light on the large-deviation properties of observables in those and other disordered model systems on hierarchical lattices.

During the last decades various model systems with and without disorder have been studied on hierarchical lattice graphs: e.g. the directed polymer problem $[3,31–36]$, the Ising spin glass $[29,30]$, and the random $q$-state Potts model $[50]$. However, using the conventional numerical approach only typical properties of the observables of interest could be analyzed. The approach presented here might be used in order to clarify the large-deviation properties of observables in those and other disordered model systems on hierarchical lattices that are not accessible via simple sampling Monte Carlo simulations.

OM acknowledges financial support from the DFG (Deutsche Forschungsgemeinschaft) under Grant HA3169/3-1. The simulations were performed at the HPC Cluster HERO, located at the University of Oldenburg (Germany) and funded by the DFG through its Major Instrumentation Programme (INST 184/108-1 FUGG) and the Ministry of Science and Culture (MWK) of the Lower Saxony State.

References

1. K. Kremer, Z. Phys. B 45, 149 (1981)
2. M. Kardar, Y.C. Zhang, Phys. Rev. Lett. 58, 2087 (1987)
3. B. Derrida, Physica A 163, 71 (1990)
4.  P. Grassberger, J. Phys. A 26, 1023 (1993)  
5.  R. Parshani, I.A. Braunstein, S. Havlin, Phys. Rev. E 79, 050102 (2009)  
6.  M. Cieplak, A. Maritan, J.R. Banavar, Phys. Rev. Lett. 72, 2320 (1994)  
7.  O. Melchert, A.K. Hartmann, Phys. Rev. B 76, 174411 (2007)  
8.  O. Melchert, A.K. Hartmann, Comp. Phys. Commun. 182, 1828 (2011)  
9.  K. Schwarz, A. Karrenbauer, G. Schehr, H. Rieger, J. Stat. Mech. 2009, P08022 (2009)  
10. M. Cieplak, A. Maritan, M.R. Swift, A. Ahattacharya, A.L. Stella, J.R. Banavar, J. Phys. A 28, 5693 (1995)  
11. O. Melchert, A.K. Hartmann, Phys. Rev. E 73, 051108 (2006)  
12. O. Melchert, A.K. Hartmann, New. J. Phys. 10, 043039 (2008)  
13. L. Apolo, O. Melchert, A.K. Hartmann, Phys. Rev. E 79, 031103 (2009)  
14. O. Melchert, L. Apolo, A.K. Hartmann, Phys. Rev. E 81, 051108 (2010)  
15. O. Melchert, A.K. Hartmann, M. Mézard, Phys. Rev. E 84, 041106 (2011)  
16. C. Norrenbrock, O. Melchert, A.K. Hartmann, arXiv:1205.1412 (2012)  
17. G. Clausen, L. Apolo, O. Melchert, A.K. Hartmann, Phys. Rev. E 86, 056708 (2012)  
18. R.K. Ahuja, T.L. Magnanti, J.B. Orlin, Network Flows: Theory, Algorithms, and Applications (Prentice Hall, New Jersey, 1993)  
19. D. Stauffer, A. Aharony, Introduction to Percolation Theory (Taylor and Francis, London, 1994)  
20. B.W. Southern, A.P. Young, J. Phys. C 10, 2179 (1977)  
21. A.N. Berker, S. Ostlund, J. Phys. C 12, 4961 (1979)  
22. M. Kaufman, R.B. Griffiths, Phys. Rev. B 24, 496 (1981)  
23. R.B. Griffiths, M. Kaufman, Phys. Rev. B 26, 5022 (1982)  
24. M. Kaufman, R.B. Griffiths, Phys. Rev. B 30, 244 (1984)  
25. M. Cieplak, A. Maritan, M.R. Swift, A. Ahattacharya, A.L. Stella, J.R. Banavar, J. Phys. A 28, 5693 (1995)  
26. O.R. Salmon, B.T. Agostini, F.D. Nobre, Phys. Lett. A 374, 1631 (2010)  
27. R. Teodoro, C.G. Bezerra, A.M. Mariz, F.A. da Costa, arXiv:1005.3863v1 (2010)  
28. A.J. Bray, S. Feng, Phys. Rev. B 36, 8456 (1987)  
29. S. Boettcher, Eur. Phys. J. B 33, 439 (2003)  
30. B. Derrida, R.B. Griffiths, Europhys. Lett. 8, 111 (1989)  
31. P. Doussal, J. Machta, J. Stat. Phys. 64, 541 (1991)  
32. P. Devillard, Phys. Rev. Lett. 70, 1124 (1993)  
33. M.S. Cao, J. Stat. Phys. 71, 51 (1993)  
34. Y. Shussman, A. Aharony, J. Stat. Phys. 80, 147 (1995)  
35. C. Monthus, T. Garel, Phys. Rev. E 77, 021132 (2008)  
36. A. Hansen, J. Kertész, Phys. Rev. Lett. 93, 040601 (2004)  
37. S.V. Buldyrev, S. Havlin, H.E. Stanley, Phys. Rev. E 73, 036128 (2006)  
38. A.K. Hartmann, Phys. Rev. E 65, 056102 (2002)  
39. A. Engel, R. Monasson, A.K. Hartmann, J. Stat. Phys. 117, 387 (2004)  
40. M. Körner, H.G. Katzgraber, A.K. Hartmann, J. Stat. Mech. 2006, P04005 (2006)  
41. C. Monthus, T. Garel, Phys. Rev. E 74, 051109 (2006)  
42. S. Wollsheimer, B. Burghardt, A.K. Hartmann, Algorithms for Molecular Biology 2, 9 (2007)  
43. W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C (Cambridge University Press, Cambridge, 1992)  
44. A.K. Hartmann, Practical Guide to Computer Simulations (World Scientific, Singapore, 2009)  
45. S.V. Buldyrev, S. Havlin, E. López, H.E. Stanley, Phys. Rev. E 70, 035102(R) (2004)  
46. T. Halpin-Healy, Y.-C. Zhang, Phys. Rep. 254, 215 (1995)  
47. D.C. Johnston, Phys. Rev. B 74, 184430 (2006)  
48. G. Franzese, A. Fierro, A. De Candia, A. Coniglio, Physica A 257, 376 (1998)  
49. F. Igloi, L. Turban, Phys. Rev. B 80, 134201 (2009)