Localization of the Maximal Entropy Random Walk

Z. Burda,[1,*] J. Duda,[1] J. M. Luck,[2] and B. Waclaw[3]
[1]Marian Smoluchowski Institute of Physics, Jagellonian University, Reymonta 4, 30-059 Kraków, Poland
[2]Institut de Physique Théorique, CEA IPhT and CNRS URA 2306, CEA Saclay, 91191 Gif-sur-Yvette cedex, France
[3]Institut für Theoretische Physik, Universität Leipzig, Postfach 100 920, 04009 Leipzig, Germany
(Received 6 November 2008; published 23 April 2009)

We define a new class of random walk processes which maximize entropy. This maximal entropy random walk is equivalent to generic random walk if it takes place on a regular lattice, but it is not if the underlying lattice is irregular. In particular, we consider a lattice with weak dilution. We show that the stationary probability of finding a particle performing maximal entropy random walk localizes in the largest nearly spherical region of the lattice which is free of defects. This localization phenomenon, which is purely classical in nature, is explained in terms of the Lifshitz states of a certain random operator.

DOI: 10.1103/PhysRevLett.102.160602 PACS numbers: 05.40.Fb, 72.15.Rn, 89.70.Cf

Since the seminal papers by Einstein [1] and Smoluchowski [2] which formulated the theory of Brownian motion and diffusive processes, random walk (RW), the paradigmatic discrete-time realization of these processes, has continuously attracted attention. RW has been discussed in thousands of papers and textbooks in statistical physics, economics, biophysics, engineering, particle physics, etc., and is still an active research area (see, e.g., [3]). Mathematically speaking, RW is a Markov chain which describes the trajectory of a particle taking successive random steps. In the simplest case of generic random walk (GRW) on a lattice, at each time step the particle chooses at random one of the adjacent nodes and jumps to it. In the continuum limit, the probability density of finding the particle at a given position obeys the diffusion equation. When the lattice is regular (i.e., all nodes have the same degree), it is easy to show that all trajectories (sequences of nodes visited by the particle) of a given length between two given points of the lattice are equiprobable, and thus have maximal entropy. It will, however, be shown below that, as soon as the lattice is not regular, GRW trajectories are no longer equiprobable. Equiprobable trajectories naturally enter the path-integral formalism [4] of quantum mechanics, where trajectories are only weighted by their length, playing the role of the action in the absence of a potential energy. The question therefore arises whether random trajectories in curved spaces (discretized as irregular lattices) should be constructed by GRW or rather by another kind of RW that leads to equiprobable paths. Finally, another example of interest is provided by the path-integral Monte Carlo methods [5], where the key issue is precisely to generate the ‘‘right’’ path statistics.

In this Letter we examine physical properties of a RW defined by the requirement that all trajectories between two given points are equiprobable, even on an irregular lattice. We shall see that this new definition leads to a dramatic change in the behavior of RW on a irregular lattice. Let us summarize our main results. First, we define the maximal entropy random walk (MERW) and show that it indeed maximizes the entropy of trajectories, in contrast to generic random walk (GRW), which has smaller entropy. Second, we discuss a surprising effect of localization of MERW trajectories in the presence of weak disorder. This is a purely classical example of the Lifshitz phenomenon [6]. Some kind of localization has been observed before in RW on networks with a broad distribution of nodes degrees [7], but for MERW the effect is completely different in nature, since it can be triggered by any small amount of inhomogeneity.

To begin, let us consider quite generally a particle hopping from node to node on a given finite, connected graph. The graph is defined by a symmetric adjacency matrix A, with elements $A_{ij} = 1$ if i and j are neighboring nodes and $A_{ij} = 0$ otherwise. The hopping is a discrete-time local Markov process: the particle sitting at some moment at node i will hop to a neighboring node j with probability P_{ij}, independently of the past history. The elements of the transition matrix are $P_{ij} = 0$ if $A_{ij} = 0$, that is if nodes i, j are not linked, and for each node i one has $\sum_j P_{ij} = 1$.

The main quantity of interest is the probability, $\pi_i(t)$, of finding the particle at node i at time t. One can calculate it recursively, as $\pi_i(t+1) = \sum_j \pi_j(t) P_{ji}$. Using spectral properties of the matrix P_{ij}, one can show that $\pi_i(t)$ generically reaches a unique stationary state π_i^* obeying the following equation:

$$\pi_i^* = \sum_j \pi_j^* P_{ji}. \tag{1}$$

For GRW, $P_{ij} = A_{ij}/k_i$, where $k_i = \sum_j A_{ij}$ is the degree of node i. This means that the particle hops to an adjacent node with the same probability for all neighbors. The stationary distribution of GRW reads

$$\pi_i^* = \frac{k_i}{\sum_j k_j}. \tag{2}$$

Another quantity of interest, especially important from the point of view of entropy, is the probability $P(\gamma_{i_0 i_t}^{(t)})$ of generating a trajectory $\gamma_{i_0 i_t}^{(t)}$ of length t passing through the nodes $(i_0, i_1, \ldots, i_{t-1}, i_t)$:

$$P(\gamma_{i_0 i_t}^{(t)}) = P_{i_0 i_1} P_{i_1 i_2} \cdots P_{i_{t-1} i_t}. \quad (3)$$

In general $P(\gamma_{i_0 i_t}^{(t)})$ depends on all nodes on the trajectory. For GRW we have $P(\gamma_{i_0 i_t}^{(t)}) = 1/(k_{i_0} k_{i_1} \ldots k_{i_{t-1}})$, so that the trajectories are manifestly not equiprobable. An exception is GRW on a k-regular graph, whose nodes have identical degrees, as, for instance, on a regular lattice. In general, however, trajectories produced by GRW are not maximally random. As we will see below, there exists, though, a natural choice of P_{ij} such that all trajectories of given length t and given endpoints are equiprobable. This choice defines MERW.

Let us now present the explicit construction of MERW. Let ψ_i be the normalized eigenvector ($\sum_i \psi_i^2 = 1$) corresponding to the largest eigenvalue λ of the adjacency matrix A:

$$\sum_j A_{ij} \psi_j = \lambda \psi_i. \quad (4)$$

The eigenvalue λ is clearly in the range $k_{\min} \le \lambda \le k_{\max}$, where $k_{\min}$ and $k_{\max}$ are the maximal and minimal node degrees of the graph, respectively. The Frobenius-Perron theorem tells us that all the ψ_i are of the same sign, so that one can choose $\psi_i > 0$. Let us use this eigenvector to define the following transition matrix:

$$P_{ij} = \frac{A_{ij}}{\lambda} \frac{\psi_j}{\psi_i}. \quad (5)$$

By construction, the entries P_{ij} are positive if i and j are neighboring nodes. They are also properly normalized: $\sum_j P_{ij} = 1$. A similar construction has been recently proposed in the context of optimal information coding [8]. The weight (3) is now independent of intermediate nodes:

$$P(\gamma_{i_0 i_t}^{(t)}) = \frac{1}{\lambda^t} \frac{\psi_{i_t}}{\psi_{i_0}}, \quad (6)$$

and thus all trajectories having length t and given endpoints i_0 and i_t are equiprobable. For a closed trajectory, the probability (6) only depends on its length t. The stationary distribution of MERW is

$$\pi_i^* = \psi_i^2, \quad (7)$$

which is easy to check by combining Eqs. (5) and (1). It is a normalized probability: $\sum_i \pi_i^* = 1$, and the detailed balance condition is fulfilled: $\pi_i^* P_{ij} = \pi_j^* P_{ji}$.

We intuitively see that random trajectories generated by MERW are more random than those generated by GRW, since the probability of a given random path (6) is independent of intermediate nodes. This statement can be quantified by comparing the entropy production rates of both Markov processes. Let $P(i_0, i_1, \ldots, i_t)$ be the probability of a sequence $(i_0, i_1, \ldots, i_t)$ in the set of all sequences of length t generated by the Markov chain. The Shannon entropy in this set of sequences is

$$S_t = -\sum_{i_0, i_1 \ldots i_t} P(i_0, \ldots, i_t) \ln P(i_0, \ldots, i_t). \quad (8)$$

One can show [9] that the entropy S_t asymptotically grows at a fixed rate

$$s \equiv \lim_{t\to\infty} \frac{S_t}{t} = -\sum_i \pi_i^* \sum_j P_{ij} \ln P_{ij}. \quad (9)$$

For GRW, with $P_{ij} = A_{ij}/k_i$ and π_i^* from Eq. (2), we obtain the entropy production rate

$$s_{\mathrm{GRW}} = \frac{\sum_i k_i \ln k_i}{\sum_i k_i}, \quad (10)$$

whereas MERW, with transition rates (5) and stationary distribution (7), yields $s_{\mathrm{MERW}} = \ln\lambda$. We now show that s_{MERW} is indeed the maximal entropy rate which can be obtained for any stochastic process generating trajectories on the graph. The number of trajectories of length t on the graph is $N_t = \sum_{i,j} (A^t)_{ij}$, where A^t is the tth power of the adjacency matrix. We therefore obtain the asymptotic value

$$s_{\mathrm{MAX}} = \lim_{t\to\infty} \frac{\ln N_t}{t} = \ln\lambda, \quad (11)$$

which sets the upper limit for the entropy production rate of such processes. We see that $s_{\mathrm{MERW}} = s_{\mathrm{MAX}}$, so that MERW indeed maximizes entropy. The inequality $s_{\mathrm{GRW}} \le s_{\mathrm{MAX}}$ is equivalent to $\sum_i k_i \ln k_i / \sum_j k_j \le \ln\lambda$. For a k-regular graph, $s_{\mathrm{GRW}} = s_{\mathrm{MERW}} = \ln k$. Similarly, for a bipartite graph which has nodes of degree k in one partition and of degree k' in the other one, $s_{\mathrm{GRW}} = s_{\mathrm{MERW}} = \frac{1}{2} \times \ln(kk')$.

As already mentioned, GRW and MERW are identical on a k-regular graph. The question then arises how much the two types of RW differ on a graph or lattice with some irregularities. For definiteness, imagine that we remove at random a small fraction $q \ll 1$ of nonadjacent links from an $L \times L$ square lattice with periodic boundary conditions. In this way we obtain a lattice with a weak disorder (dilution), where most of the nodes are of degree $k = 4$ and some of degree $k = 3$. The stationary distribution π_i^* for GRW is given by Eq. (2), so that the probability of finding the particle after a long time at a defective node is equal to $3/4$ of the probability at an intact one. The situation looks completely different for MERW, as shown in Fig. 1, presenting density plots of π_i^* for different densities of defects, obtained by diagonalizing A numerically and using Eq. (7). For a very low density q of defects, the probability π_i^* is smaller in the neighborhood of de-

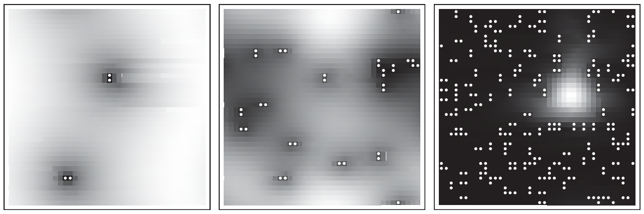

FIG. 1. Density plots of π_i^* for a periodic square lattice of size $L \times L$, for $L = 40$ and the fractions $q = 0.001$, 0.01, 0.1 of removed links. The nodes incident with removed links are marked with circles. Data are obtained numerically by applying the Lanczos algorithm to the adjacency matrix. For a sparse matrix used here, the algorithm has a complexity $O(N^2)$, where $N = L^2$ in two dimensions. Plots for other values L, q can be produced using the demonstration [16].

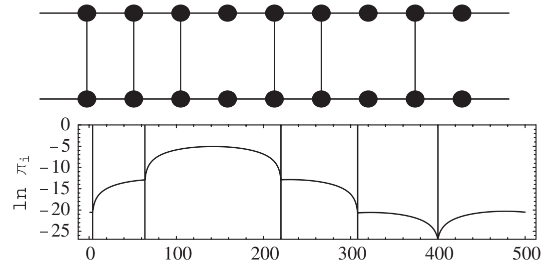

FIG. 2. Top: a ladder with randomly removed rungs. Bottom: stationary distribution π_i^* on the ladder for $L = 500$ and five randomly positioned defects marked with vertical lines.

fects, like in the GRW case. However, if the number of defects increases, the stationary distribution π_i^* becomes localized in a nearly circular region. We will indeed show, using the Lifshitz argument [6], that this localization phenomenon takes place for any nonzero fraction of defects, provided the linear size L of the system is large enough, and that the radius of the localization region grows as $(\ln L)^{1/2}$.

Let us start with a $1d$ example, in order to build up some intuition. As a model example we shall consider a ladder graph with periodic boundary conditions, with a fraction q of randomly removed rungs, as shown in Fig. 2. In order to define the transition probabilities (5) we have to solve the eigenvalue problem for the adjacency matrix A. Let L be length of the ladder. Taking into account the symmetry between both legs, we have

$$\psi_{i+1} + \psi_{i-1} + r_i \psi_i = \lambda \psi_i, \tag{12}$$

where i runs over the L nodes in the lower leg of the ladder, say, and $r_i = 1$ if there is a rung at the position i, and $r_i = 0$ otherwise. Introducing the discrete Laplacian $\Delta_{ij} = \delta_{i,j+1} + \delta_{i,j-1} - 2\delta_{ij}$, Eq. (12) can be recast as

$$-(\Delta \psi)_i + v_i \psi_i = E \psi_i, \tag{13}$$

where $E = 3 - \lambda$, whereas $v_i = 1 - r_i$ form a random binary sequence with a frequency of unities or defects ($v_i = 1$) equal to q and a frequency of zeros ($v_i = 0$) equal to $p = 1 - q$. Each sequence of nodes without defects ($v_i = 0$) is said to form a well. Equation (13) is formally identical to the eigenvalue equation of the following trapping problem. A particle performs a RW in continuous time on the $1d$ lattice. Defects act as static traps: whenever the particle sits at node i, it is annihilated at rate v_i per unit time. Trapping problems of this kind have been studied extensively [10]. The asymptotic longtime falloff of the survival probability is known to be related to the so-called Lifshitz tail in the density of states of Eq. (13) as $E \to 0$. In the present context, the Lifshitz argument [6] predicts that the ground state of Eq. (13) is well approximated by that of the longest well, i.e., $-(\Delta\psi)_i = E_0 \psi_i$ ($i = 1, \ldots, w$), with Dirichlet boundary conditions $\psi_0 = \psi_{w+1} = 0$, where w is the length of that well. We obtain $\psi_i \sim \sin[i\pi/(w+1)]$ and $E_0 = 2[1 - \cos\pi/(w+1)] \approx \pi^2/w^2$. In the $1d$ situation [11], this argument is known to essentially give an exact description of the ground state.

In the case of MERW, we therefore predict that the whole stationary probability is asymptotically localized on the longest well, i.e., the longest sequence without defects. The Lifshitz picture is well illustrated by Fig. 2, showing plots of the stationary density π_i^*, obtained by numerical diagonalization of A. The asymptotic growth of the length w of the longest well can be estimated as follows. The mean number of unities in the sequence grows as Lq. The mean number of those followed by one zero is Lqp, by two zeros is Lqp^2, and so on, so that there are Lqp^n wells of length n, i.e., consisting of n zeros. The length w of the longest well, given by the estimate $Lqp^w \sim 1$, grows logarithmically with the system size, as $w \approx \ln L/|\ln p|$, so that $E_0 \approx (\pi|\ln p|/\ln L)^2$. In Fig. 3 we show that the ground-state energy E_0 obtained by numerically solving Eq. (13), averaged over binary disorder for $q = 0.1$, agrees with the above estimate for L large enough.

The Lifshitz argument can be generalized to higher-dimensional situations [12]. The ground state of the discretized Schrödinger equation (13) is localized in the largest *Lifshitz sphere*, defined as the largest nearly spherical

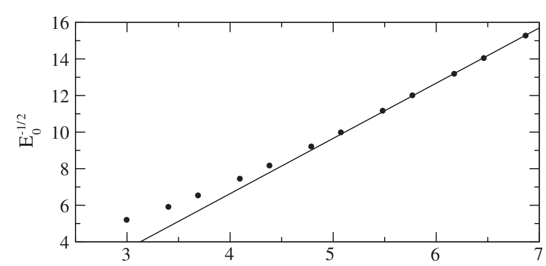

FIG. 3. Ground-state energy E_0 of Eq. (13) on ladders versus $\ln L$ for $L = 20, \ldots, 960$ and $q = 0.1$. The solid line shows the estimate $E_0^{-1/2} = \ln L/(\pi|\ln p|) + B$, with B fitted to the rightmost data point.

region of the lattice which is free of defects. Taking again for definiteness the example of the square lattice, the radius $R_{\max}$ of the largest Lifshitz disk can be estimated to be $R_{\max} \approx (\ln L/(\pi|\ln p|))^{1/2}$, since the number of circular regions of radius R with no defects is of order $L^2 p^{2\pi R^2}$, as there are two links per node. In higher dimension d, the above estimate reads $R_{\max} \sim (\ln L/|\ln p|)^{1/d}$, skipping constants. Hence the stationary probability of MERW on a d-dimensional lattice in the presence of any amount of disorder is localized in the largest Lifshitz sphere, whose *volume* grows as $\ln L$. Inside the sphere, the MERW essentially has a diffusive character.

In conclusion, the MERW introduced in this work is a local process, in the sense that the particle jumps to neighboring nodes. The local transition rates, however, depend on the global structure of the graph through its Perron-Frobenius vector. MERW is indeed the result of globally maximizing the entropy of trajectories, in contrast to GRW, which can be viewed as the result of a greedy entropy maximization at each step. On the other hand, although MERW requires the knowledge of the entire system, this knowledge is not always advantageous for the random walker. It indeed turns out that generating all paths with equal probability may prevent the walker from exploring the entire space. This occurs, e.g., on lattices with weak dilution, where the particle gets localized by a Lifshitz phenomenon. Interestingly, this effect can be put in perspective with ergodicity breaking resulting from entropy barriers [13]. It is therefore tempting to consider MERW as a simple model of evolution in a flat fitness landscape but with entropic traps. In the course of evolution, the system goes through consecutive metastable states, i.e., larger and larger local Lifshitz spheres, until it finally reaches the true ground state, i.e., the largest Lifshitz sphere.

Let us close with a comment on the connection between MERW and path integrals in curved space. The quantum-mechanical amplitude for a free particle propagating in a curved, discretized space-time is usually calculated as a sum over GRW trajectories. In quantum gravity [14] the d-dimensional space is discretized into simplices. If a free particle propagates on the dual graph formed by the centers of simplices, the underlying graph is $(d + 1)$ regular, so that MERW and GRW are equivalent. On the other hand, if it propagates on the original graph, which is irregular, then the two types of RW are different. It would be interesting to compare the MERW and GRW quantum particle propagation in Lorentzian quantum gravity [15]. It may happen that the free particle will tend to localize in tubelike regions extended in the temporal direction due to spatial "defects" that arise from quantum fluctuations of Lorentzian geometry.

We would like to thank P. Bialas and F. David for discussions. This work was supported in part by the Alexander von Humboldt Foundation's Grant No. 3.4-Fokoop-DEU/1117877, the Marie Curie ToK project "COCOS", No. MTKD-CT-2004-517186, the EC-RTN Network "ENRAGE", No. MRTN-CT-2004-005616, and the Polish Ministry of Science Grant No. 1P03B-04029 (2005–2008).

*zdzislaw.burda@uj.edu.pl

[1] A. Einstein, Ann. Phys. (Leipzig) **17**, 549 (1905); **19**, 371 (1906).
[2] M. Smoluchowski, Ann. Phys. (Leipzig) **21**, 756 (1906).
[3] C. Anteneodo and W. A. M. Morgado, Phys. Rev. Lett. **99**, 180602 (2007); S. N. Majumdar and R. M. Ziff, Phys. Rev. Lett. **101**, 050601 (2008).
[4] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
[5] J. H. Hetherington, Phys. Rev. A **30**, 2713 (1984).
[6] I. M. Lifshitz, Adv. Phys. **13**, 483 (1964); Sov. Phys. Usp. **7**, 549 (1965).
[7] J. D. Noh and H. Rieger, Phys. Rev. Lett. **92**, 118701 (2004).
[8] J. Duda, arXiv:0710.3861.
[9] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948); **27**, 623 (1948); B. McMillan, Ann. Math. Stat. **24**, 196 (1953).
[10] B. Y. Balagurov and V. G. Vaks, Zh. Eksp. Teor. Fiz. **65**, 1939 (1973) [Sov. Phys. JETP **38**, 968 (1974)]; M. D. Donsker and S. R. S. Varadhan, Commun. Pure Appl. Math. **28**, 525 (1975); **32**, 721 (1979); P. Grassberger and I. Procaccia, J. Chem. Phys. **77**, 6281 (1982); J. W. Haus and K. W. Kehr, Phys. Rep. **150**, 263 (1987).
[11] Th. M. Nieuwenhuizen, Physica (Amsterdam) **167A**, 43 (1990); J. M. Luck, *Systèmes Désordonnés Unidimensionnels* (Collection Aléa-Saclay, Saclay, 1992), and the references therein.
[12] L. A. Pastur, Russ. Math. Surv. **28**, 1 (1973); R. Friedberg and J. M. Luttinger, Phys. Rev. B **12**, 4460 (1975); W. Kirsch and F. Martinelli, Commun. Math. Phys. **89**, 27 (1983).
[13] F. Ritort, Phys. Rev. Lett. **75**, 1190 (1995); F. Ritort and P. Sollich, Adv. Phys. **52**, 219 (2003).
[14] J. Ambjørn, Z. Burda, J. Jurkiewicz, and C. F. Kristjansen, Acta Phys. Polon. **23**, 991 (1992).
[15] J. Ambjørn, J. Jurkiewicz, and R. Loll, Phys. Rev. Lett. **95**, 171301 (2005).
[16] B. Waclaw, http://demonstrations.wolfram.com/GenericRandomWalkAndMaximalEntropyRandomWalk/.