Statistics of the number of minima in a random energy landscape

Satya N. Majumdar and Olivier C. Martin

CNRS, Université Paris-Sud, UMR8626, LPTMS, Orsay Cedex, F-91405, France

We consider random energy landscapes constructed from d-dimensional lattices or trees. The distribution of the number of local minima in such landscapes follows a large-deviation principle and we derive the associated law exactly for dimension 1. Also of interest is the probability of the maximum possible number of minima; this probability scales exponentially with the number of sites. We calculate analytically the corresponding exponent for the Cayley tree and the two-leg ladder; for two- to five-dimensional hypercubic lattices, we compute the exponent numerically and compare to the Cayley tree case.

INTRODUCTION

For many materials that are glassy [1] local minima of the energy (or of the free energy) trap the system for long times, leading to subtle equilibrium and out-of-equilibrium properties. Energy landscapes provide a simple conceptual framework for modeling these systems but in fact their use goes much beyond that. For instance, the vacua of string theories are expected to proliferate enormously, and the problem of estimating the number of local minima [2,3] of string energy landscapes is still an open problem. Other examples include rugged (random) energy landscapes in evolutionary biology [4], quantum cosmology [5], manifolds in random media [6], glassy systems [1,7], random potentials [8,9], and the associated problems in random matrix theory [3,10,11]. Typically these systems consider a particle, a configuration of particles, or even a manifold, subject to a random potential. One extreme case for its visual simplicity is that of a point particle in a random Gaussian potential; on the opposite extreme are energy landscapes associated with the configuration of a many body system. Examples in this last category include (1) the p-spin glass model [12], which in the limit of large p reduces to the far simpler random energy model of Derrida [13]; (2) atomic clusters [14] and other glassy systems with no quenched disorder [15]; (3) random manifolds in random media [6].

For any landscape, it is desirable to know the statistical properties of the minima (or more generally of the saddle points). A quantity frequently considered is the expected number of minima as a function of their energy [9]. It may also be of interest to consider how the set of minima are organized topologically, e.g., whether the barrier tree [16] is ultrametric. Our focus here is to better understand the statistics of the total number of local minima in random energy landscapes; our underlying space is a regular (Euclidean) lattice on each site of which resides a random energy. Let M denote the total number of local minima for given values of the energies on each site. Evidently M is a random variable as it varies from one realization of the landscape to another. We are interested in the statistics of M as a function of the number N of lattice sites. Similar questions were studied recently in the context of random permutations [17,18], ballistic deposition [19], and in simple models of glasses [20]. In this paper, we provide a number of analytical results on the distribution of the total number of local minima for random energy landscapes on several lattices; the moments of M are easily derived, so our focus concerns mainly the probability of *large deviations* of M from its mean value, i.e., the probabilities of atypical configurations. Furthermore, we ask what is the probability of M being at its maximum, corresponding to the limit where the minima are maximally packed on the lattice.

The paper is organized as follows. In Sec. I we specify the model and show that the statistical properties of the minima are independent of the individual distribution of energies as long as these energies are independent from site to site and are drawn from a continuous distribution. In Sec. II we cover some of the simplest properties of the statistics of M and formulate the large-deviation principle. Then we derive the closed form expression for the large-deviation function in the case of the one-dimensional lattice in Sec. III. The focus of the rest of the paper is the maximum packing problem. In Sec. IV we determine the probability of maximum packings for several solvable cases, namely, Cayley trees and a two-leg ladder. The case of d-dimensional hypercubic lattices is then treated by numerical computation in Sec. V. Finally, these different results are discussed and some closing remarks are given.

I. THE MODEL

We start with a regular lattice; on each site i exists a random energy E_i, drawn independently from site to site from a common distribution $\rho(E)$. We assume $\rho(E)$ is continuous and normalized to unity, i.e., $\int_{-\infty}^{\infty} \rho(E)dE = 1$. For any choice of the set of E_i's we obtain an energy landscape, i.e., a topological space (defined through nearest neighbors on the lattice) with an energy for each element. For a given realization of the landscape, a site i is a local minimum if $E_i < E_j$ where j denotes any nearest neighbor sites of i. Hereafter we shall denote by M the number of local minima on this landscape.

For any given realization of the landscape, one can formally express M as

$$M = \sum_{i=1}^{N} \prod_{j/\langle ij \rangle} \theta(E_j - E_i) \qquad (1)$$

where the product runs over all nearest neighbors j of site i and $\theta(x)$ is the Heaviside theta function. We are interested in

computing the probability distribution $P(M,N)$ of the total number of local minima (or equivalently that of the maxima).

The first important observation is that the distribution $P(M,N)$ is completely independent of the energy distribution $\rho(E)$. To see this, we formally express the distribution as

$$P(M,N) = \int \cdots \int \delta\left(M - \sum_{i=1}^{N} \prod_{j} \theta(E_j - E_i) \right) \prod_{k=1}^{N} \rho(E_k) dE_k \tag{2}$$

where $\delta(x)$ is the Dirac delta function. Next we make a change of variable for each k,

$$x_k = \int_{-\infty}^{E_k} \rho(E) dE. \tag{3}$$

Clearly x_k is a monotonically increasing function of E_k. Therefore, $\theta(E_j - E_i) = \theta(x_j - x_i)$. Moreover, since $\rho(E)$ is normalized to unity, the variable x varies from 0 to 1 and so Eq. (2) simply becomes

$$P(M,N) = \int_0^1 dx_1 \cdots \int_0^1 dx_N \delta\left(M - \sum_{i=1}^{N} \prod_{j} \theta(x_j - x_i) \right). \tag{4}$$

We see that the energy distribution $\rho(E)$ simply drops out and $P(M,N)$, for arbitrary $\rho(E)$, is universal and is the same as when the "new" energy variable x_i is drawn independently from a uniform distribution over $x \in [0,1]$. Thus, the model is simplified: at each lattice site exists a random number $x_i \in [0,1]$ drawn independently from a uniform distribution. We will refer to this model as the random minima model, for which we want to compute the distribution $P(M,N)$ of the total number of local minima.

Following the mapping to the uniform distribution, it follows that the random minima model is just the continuous version of the permutation generated landscape recently studied by Hivert et al. [19]. They considered a set of integers $\{1,2,3,\ldots,N\}$. Each of the $N!$ permutations of this set defines a random energy landscape and they occur with equal probability. Thus the energy at a site is now a discrete integer drawn uniformly from the set $[1,2,3,\ldots,N]$ with equal probability $1/N$. Thus the statistics of the minima in the permutation generated landscape will be identical to that of the random minima model.

II. GENERAL PROPERTIES OF THE DISTRIBUTION OF M

The mean and the variance of M are relatively straightforward to compute on an arbitrary lattice since they depend only on the local properties of the landscape. To see this, let us define the variable

$$\eta_i = \prod_{j} \theta(x_j - x_i) \tag{5}$$

where j runs over the nearest neighbors of i and the x_i's are independent random numbers in $[0,1]$ drawn from the uni-

form distribution. This η_i is an indicator function which is 1 if site i is a local minimum and 0 otherwise. Then it follows that

$$M = \sum_{i=1}^{N} \eta_i. \tag{6}$$

Taking the average in Eq. (6) and using the translational invariance of the lattice, it follows that

$$\frac{\langle M \rangle}{N} = \left\langle \prod_{j/\langle ij \rangle} \theta(x_j - x_i) \right\rangle = \int_0^1 dx_i \left[\int_{x_i}^1 dx_j \right]^n = \frac{1}{n+1} \tag{7}$$

where n is the number of nearest neighbors of any site, i.e., the coordination number of the lattice. This result also follows trivially from a combinatorial argument on the permutation landscape: consider the site i with its n neighbors. The number of ways one can arrange the integers on these $(n+1)$ sites with the restriction that i is a local minimum is $n!$ clearly. On the other hand, the total number of unrestricted configurations is $(n+1)!$ so the probability that the site i is a minimum is simply $n!/(n+1)! = 1/(n+1)$.

The calculation of the variance takes a few more steps. Squaring Eq. (6) and taking the average gives

$$\langle M^2 \rangle = \sum_{i,j} \langle \eta_i \eta_j \rangle. \tag{8}$$

Note that the η_i's are correlated random variables but only over a short range. The calculation of the correlation function $\langle \eta_i \eta_j \rangle$ can therefore be performed by hand (it only involves the calculation of simple integrals involving a maximum of $2n$ sites). For example, in one dimension (1D), one gets

$$\sigma^2 = \langle M^2 \rangle - \langle M \rangle^2 \rightarrow \frac{2}{45} N \quad \text{as } N \rightarrow \infty. \tag{9}$$

This result has been obtained in various contexts before, such as in the calculation of the number of metastable states in 1D Ising spin glass [21] and also in the context of the occupation time of a non-Markovian sequence in 1D [22]. Recently, the variance for the 2D square lattice was computed in the permutation model by Hivert et al. [19]. For large N, one gets

$$\sigma^2 = \frac{13}{225} N. \tag{10}$$

In general, on any arbitrary lattice, for large N

$$\langle M \rangle = aN,$$

$$\langle M^2 \rangle - \langle M \rangle^2 = bN, \tag{11}$$

where a and b are lattice-dependent numbers that can be computed either using integrals or by combinatorics in the permutation model.

Near the mean $\langle M \rangle$ and within a region $|M - \langle M \rangle| = O(\sqrt{N})$, the distribution $P(M,N)$ is expected to be a Gaussian [19] with mean and variance given in Eq. (11),

$$P(M,N) \sim \frac{1}{\sqrt{2\pi bN}} \exp[-(M-aN)^2/2bN]. \quad (12)$$

However, for M far from the mean, one would expect deviations in $P(M,N)$ from the Gaussian form. In this paper, we are interested in the probabilities of such large deviations, i.e., we are interested in computing the probabilities of occurrences of configurations that are far from typical. On general grounds, in the limit $M \to \infty$, $N \to \infty$ but with their ratio M/N fixed but arbitrary, one expects that $P(M,N)$ has the form

$$P(M,N) \sim \exp\left[-N\Phi\left(\frac{M}{N}\right)\right] \quad (13)$$

where $\Phi(y)$ is a large-deviation function. Now, on a given lattice, M can take values from 0 to a maximal number $M_{max} < N$. The upper limit follows from the fact that, once a site is a local minimum, none of its nearest neighbors can be a local minimum. Thus there is a nearest neighbor exclusion principle for the minima. This constraint indicates that one cannot pack arbitrarily large number of minima on the lattice and there is an upper bound on the number of minima. For example, on a bipartite lattice, one can pack at most $M_{max} = N/2$ local minima, one at every alternate site. In particular, on a square lattice the minima will be placed on a checkerboard pattern. Thus, in this case, the large-deviation function $\Phi(y)$ is defined for $0 \le y \le y_{max} = 1/2$. On the other hand, on a Cayley tree with μ the number of branches and N sites, we will see later that $M_{max} = \mu N/(\mu+1)$; thus $y_{max} = \mu/(\mu+1)$.

As mentioned above, in the vicinity of its mean, i.e., for $|M - \langle M \rangle| = O(\sqrt{N})$, the distribution $P(M,N)$ is Gaussian as in Eq. (12). This indicates that the large-deviation function $\Phi(y)$ is quadratic near $y = a$,

$$\Phi(y) \approx \frac{(y-a)^2}{2b} \quad (14)$$

such that $P(M,N) \approx \exp[-N(y-a)^2/2b] \sim \exp[-(M-aN)^2/2bN]$ has the required Gaussian form. However, far from the mean, $P(M,N)$ will have non-Gaussian tails, indicating a departure of $\Phi(y)$ from the simple quadratic form. A knowledge of the function $\Phi(y)$ would then allow one to compute the probabilities of occurrences of atypical configurations, such as the probability of a configuration with the lowest number of minima (e.g., $M = 1$) or the ones with the maximal number of minima ($M = M_{max}$).

A particular focus of this paper will be to compute the probability of a maximally packed configuration, i.e., when $M = M_{max}$. It follows from the general form in Eq. (13) that the probability of this maximal packing should decay exponentially with the system size for large N,

$$P(M_{max},N) \sim \exp[-N\Phi(y_{max})] \sim \gamma^{-N} \quad (15)$$

where $\gamma = \exp[\Phi(y_{max})]$ is a lattice-dependent constant. Of course, this expression also gives the scaling of the fraction of permutations which have M_{max} local minima in the permutation landscape. It is then natural to interpret $\ln \gamma$ as the entropy cost per site to achieve maximally packed minima.

We shall compute the constant γ exactly for a number of lattices. For the 1D chain, we will show that

$$\gamma = \frac{\pi}{2} = 1.570\,79\ldots. \quad (16)$$

Note that this result appeared before in the context of metastable states in 1D Ising spin glasses at zero temperature [21]. Furthermore, the same $(\pi/2)^{-N}$ decay also appeared as the persistence probability of a 1D non-Markovian sequence [22,23]. For the 1D case, one can also compute the full large-deviation function $\Phi(y)$ (see Sec. III). Another lattice where we can calculate γ exactly is the Cayley tree with μ branches where we show that

$$\gamma = \frac{1}{\mu+1} B\left(\frac{1}{\mu+1}, \frac{1}{\mu+1}\right) \quad (17)$$

where $B(x,y)$ is the Beta function. As expected, for $\mu = 1$, it reduces to the 1D result in Eq. (16). As μ increases, γ increases slowly, and as $\mu \to \infty$, $\gamma \to 2$. A Cayley tree with an infinite number of branches corresponds to a hypercubic lattice in d dimensions in the $d \to \infty$ limit. Thus, based on these two limiting results, one expects the constant γ on any d-dimensional lattice to satisfy the bounds

$$\pi/2 \le \gamma \le 2. \quad (18)$$

Our numerical simulations for $2 \le d \le 5$ are consistent with these bounds. We have also been able to compute γ exactly for a two-leg ladder where we show that

$$\gamma = \left[\frac{9\alpha^2}{8}\right]^{1/3} = 1.576\,57\ldots \quad (19)$$

where $\alpha = 1.866\,35\ldots$ is the smallest positive root of the Bessel function $J_{-1/3}(z) = 0$. The result in Eq. (19) is, of course, consistent with the general bounds in Eq. (18).

We now consider these successive cases in detail.

III. EXACT LARGE-DEVIATION FUNCTION IN ONE DIMENSION

In a 1D chain of size N, the distribution $P(M,N)$ of the total number of local minima was computed exactly by Derrida and Gardner [21], although they did not calculate the large-deviation function $\Phi(y)$ explicitly. However, from their result for the generating function, it is easy to derive the large-deviation function by a Legendre transform. For the sake of completeness, we provide here a brief derivation of the 1D result, albeit by a slightly different method.

It is useful to define the generating function or the partition function $Z(z,N) = \Sigma_M P(M,N)z^M$ where z is the fugacity or the weight associated with each minimum. For simplicity we consider an open chain of size N and let $x_N = x$ be the value of the random variable at the Nth site. To write a recursion relation for the partition function, it is convenient to define two restricted partition functions $Z_1(x,z,N)$ and $Z_0(x,z,N)$ denoting, respectively, the partition functions conditioned on the fact that the Nth site has value $x_N = x$ and that it is, respectively, a local minimum (i.e., $x_N < x_{N-1}$) or a local

maximum ($x_N > x_{N-1}$). Note that since we are considering an open chain, the last site (N) has only one neighbor to its left, namely, the ($N-1$)th site. Knowing the restricted partition functions, one can compute the full partition function from the relation

$$Z(z,N) = \int_0^1 dx[Z_1(x,z,N) + Z_0(x,z,N)]. \qquad (20)$$

The restricted partition functions satisfy a pair of simple recursion relations

$$Z_1(x,z,N) = z\int_x^1 Z_0(y,z,N-1)dy + \int_x^1 Z_1(y,z,N-1)dy, \qquad (21)$$

$$Z_0(x,z,N) = \int_0^x [Z_0(y,z,N-1) + Z_1(y,z,N-1)]dy. \qquad (22)$$

These recursion relations can be easily understood by considering all possibilities when one adds a new site to the chain. If the new site is a minimum, we need to attach a factor z. On the other hand, if the ($N-1$)th site was a minimum and it ceases to be a minimum after the addition of the Nth site, we have to detach a factor z.

It follows from Eqs. (21) and (22) that the restricted partition functions satisfy the boundary conditions $Z_1(x=1,z,N)=0$ and $Z_0(x=0,z,N)=0$. For large N, one expects a separation of variables between x and N of the form

$$Z_{1,0}(x,z,N) \sim \lambda^{-N} f_{1,0}(x) \qquad (23)$$

where $\lambda(z)$ is a function of z only (but independent of x) and the functions $f_1(x)$ and $f_0(x)$ satisfy the boundary conditions $f_1(1)=0$ and $f_0(0)=0$. Substituting this ansatz in Eqs. (21) and (22) and subsequently differentiating with respect to x, we get a pair of differential equations

$$\frac{df_1}{dx} = -\lambda f_1(x) - z\lambda f_0(x), \qquad (24)$$

$$\frac{df_0}{dx} = \lambda f_1(x) + \lambda f_0(x). \qquad (25)$$

Diagonalizing the [2×2] matrix, one obtains the solutions

$$f_1(x) = ae^{\lambda\sqrt{1-z}x} + be^{-\lambda\sqrt{1-z}x}, \qquad (26)$$

$$f_0(x) = -\frac{a}{(1-\sqrt{1-z})}e^{\lambda\sqrt{1-z}x} - \frac{b}{(1+\sqrt{1-z})}e^{-\lambda\sqrt{1-z}x}, \qquad (27)$$

where a and b are arbitrary constants. The two boundary conditions $f_1(1)=0$ and $f_0(0)=0$ yield two relations between a and b,

$$ae^{\lambda\sqrt{1-z}} + be^{-\lambda\sqrt{1-z}} = 0, \qquad (28)$$

$$\frac{a}{(1-\sqrt{1-z})}e^{\lambda\sqrt{1-z}} + \frac{b}{(1+\sqrt{1-z})}e^{-\lambda\sqrt{1-z}} = 0. \qquad (29)$$

Eliminating a and b between Eqs. (28) and (29) determines the eigenvalue $\lambda(z)$ exactly,

$$\lambda(z) = \begin{cases} \dfrac{1}{2\sqrt{1-z}}\ln\left(\dfrac{1+\sqrt{1-z}}{1-\sqrt{1-z}}\right) & \text{for } 0 \le z \le 1, \qquad (30) \\[4mm] \dfrac{1}{\sqrt{z-1}}\tan^{-1}(\sqrt{z-1}) & \text{for } z \ge 1. \qquad (31) \end{cases}$$

Note that the function $\lambda(z)$ is analytic at $z=1$. The form in Eq. (31) is just an analytical continuation of the form in Eq. (30) for $z \ge 1$.

Substituting the large N form of Z_1 and Z_0 in Eq. (20), one obtains the large-N behavior of the partition function,

$$Z(z,N) = \sum_M P(M,N)z^M \sim [\lambda(z)]^{-N} \qquad (32)$$

where $\lambda(z)$ is given in Eqs. (30) and (31). Substituting the anticipated form of $P(M,N) \sim \exp[-N\Phi(M,N)]$ in Eq. (32) one gets

$$Z(z,N) = \sum_M P(M,N)z^M \sim \int dy \exp[-N(\Phi(y) - y\ln z)]$$

$$\sim [\lambda(z)]^{-N}. \qquad (33)$$

Taking the $N \to \infty$ limit in Eq. (33) gives

$$\min_y[\Phi(y) - y\ln z] = \ln[\lambda(z)]. \qquad (34)$$

Inverting via the Legendre transform finally gives the large-deviation function

$$\Phi(y) = \max_z[\ln(\lambda(z)) + y\ln z], \qquad (35)$$

where $\lambda(z)$ is given in Eqs. (30) and (31). Note that determining $\Phi(y)$ from Eq. (35) requires a knowledge of $\lambda(z)$ for all $z \ge 0$; thus we need both formulas of $\lambda(z)$ in Eqs. (30) and (31).

We have obtained $\Phi(y)$ from Eq. (35) using MATHEMATICA and it is displayed in Fig. 1. Since the maximal value of M in 1D is $N/2$, the allowed range of y is $0 \le y \le 1/2$. One can analytically obtain the form of $\Phi(y)$ in the three limiting cases $y \to 0$, $y \to 1/3$, and $y \to 1/2$. First let us consider the limit $y \to 0$. To find $\Phi(y)$ in this limit we need to use the $z \to 0$ form of $\lambda(z)$ in Eq. (35). As $z \to 0$, it is easy to see from Eq. (30) that $\lambda(z) \to \ln(4/z)/2$. Substituting this behavior in Eq. (35) and subsequently maximizing the right-hand side (RHS) of Eq. (35) we find that $\Phi(y)$ diverges logarithmically, $\Phi(y) \approx \ln[\ln 2/(ey)]$, as $y \to 0$. Next we consider the $y \to 1/3$ limit. Note that the mean number of minima $\langle M \rangle = N/3$ in $d=1$ which follows from the general result in Eq. (7). Thus the $y \to 1/3$ limit corresponds to behavior of M near its mean and one expects a quadratic form for $\Phi(y)$ near $y=1/3$. Indeed this also follows from Eq. (35). The limit $y \to 1/3$ corresponds to using the $z \to 1$ behavior of $\lambda(z)$ in Eq. (35). Substituting $z=1+\epsilon$ in Eq. (31) and expanding in powers of ϵ, one gets $\ln[\lambda(z)] = -\epsilon/3 + 13\epsilon^2/90$
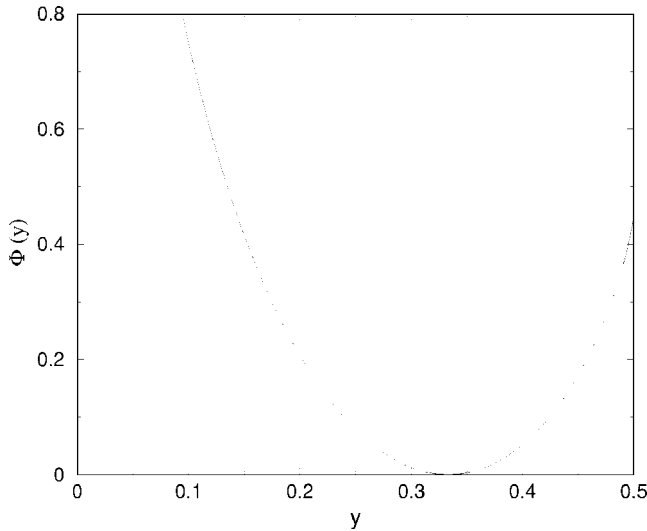
FIG. 1. The large deviation function $\Phi(y)$ for $0 \le y \le 1/2$ in 1D obtained from Eq. (35) using MATHEMATICA.

$+O(\epsilon^3)$. Substituting this result on the RHS of Eq. (35) and maximizing, one gets the expected quadratic behavior $\Phi(y) \approx 45(y-1/3)^2/4$ near $y \to 1/3$. This is thus a special case in 1D of the general behavior in Eq. (14) with $a=1/3$ and $b=2/45$. Finally, to derive the maximally packed limit $y \to 1/2$, we need to use the $z \to \infty$ behavior of $\lambda(z)$ in Eq. (35). In this limit, it follows from Eq. (31) that $\ln[\lambda(z)] \to \ln(\pi/2) - \ln(z)/2 - 2z^{-1/2}/\pi$. Using this form and maximizing the RHS of Eq. (35) we get $\Phi(y) \approx \ln(\pi/2) + 2(1/2 - y)\ln[\pi(1/2-y)/e]$ as $y \to 1/2$. Thus, summarizing the three limiting behaviors,

$$\Phi(y) \approx \begin{cases} \ln[\ln 2/(ey)] & \text{as } y \to 0, \\ \dfrac{45}{4}(y-1/3)^2 & \text{as } y \to 1/3, \\ \ln(\pi/2) + 2(1/2-y)\ln[\pi(1/2-y)/e] & \text{as } y \to 1/2. \end{cases} \tag{36}$$

Note that, as one approaches the maximally packed limit $y \to 1/2$, $\Phi(y_{max}=1/2) = \ln(\pi/2)$. Thus, it follows from Eq. (15) that in 1D, for large N,

$$P(M_{max} = N/2, N) \sim \gamma^{-N}, \quad \text{with } \gamma = \pi/2, \tag{37}$$

the result declared in Eq. (16).

IV. PROBABILITY OF THE MAXIMALLY PACKED CONFIGURATION: SOLVABLE CASES

In this section, we focus only on the maximally packed configuration (where the number of minima on the lattice is maximal). From the general large-deviation theory, we have already argued that the probability of such a configuration $P(M_{max}, N)$ is expected to decay exponentially with the system size N as in Eq. (15). The goal is to compute the nontrivial constant γ. In the previous section, we have shown
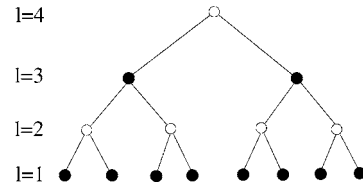

FIG. 2. The maximally packed configuration of local minima (denoted by black dots) on a Cayley tree with $\mu=2$ branches and $n=4$ generations. The layers are labeled $l=1,2\ldots$ starting with the bottom layer.

that, in a 1D chain, $\gamma=\pi/2$. In this section, we compute γ exactly in a few other solvable cases, notably for a Cayley tree with μ branches and also for a two-leg ladder in 2D.

A. Exact calculation of γ on a Cayley tree

We consider a Cayley tree with μ branches and n generations. We label the generations by $l=1,2,\ldots,n$ starting from the leaf sites at the bottom (see Fig. 2). The total number of sites on the tree is

$$N = 1 + \mu + \mu^2 + \cdots + \mu^{n-1} = \frac{\mu^n - 1}{\mu - 1}. \tag{38}$$

Note that in the limit $\mu \to 1^+$, the tree reduces to a 1D chain with $N=n$ sites.

At each site of this tree resides a random number x_i drawn independently from the uniform distribution over [0, 1]. We want to calculate the probability of the configuration with the maximum possible number of local minima. Note that, if a site is a local minimum, its neighbors cannot be local minima. To find the maximally packed configuration, we note that the number of sites at the bottom layer ($l=1$) increases exponentially with n as μ^{n-1}. Thus, to achieve maximal packing, it is necessary to fill up the bottom layer (the leaf sites) with local minima. Then the layer just above the bottom layer ($l=2$) is devoid of local minima. The next layer ($l=3$) can again be packed with local minima. Thus, the maximally packed configuration is the one where alternate layers are fully packed with minima, starting with a fully packed layer at the bottom, as shown in Fig. 2. The total number of minima M_{max} in this maximally packed configuration depends on whether the number of generations n is even or odd,

$$M_{max} = \begin{cases} \dfrac{\mu^{n+1} - 1}{\mu^2 - 1}, & \text{even } n, & (39) \\ \dfrac{\mu^{n+1} - \mu}{\mu^2 - 1}, & \text{odd } n. & (40) \end{cases}$$

For n even, the root is not a minimum, whereas for n odd, the root is a minimum. In either case, for n large, the total number of minima in the maximally packed configuration is proportional to the total number of sites N,

$$M_{\max} \approx \frac{\mu}{\mu+1} N. \tag{41}$$

Having identified the maximally packed configuration, we will next compute the probability of its occurrence and show that for large N

$$P(M_{\max}, N) \sim \gamma^{-N} \quad \text{where} \quad \gamma = \frac{1}{\mu+1} B\left(\frac{1}{\mu+1}, \frac{1}{\mu+1}\right). \tag{42}$$

For simplicity, let us assume that n is even. One can perform an identical calculation when n is odd. For n even, all the even layers (labeled by $l=2m$ with $m=1,2,\ldots,n/2$) are devoid of local minima whereas all the odd layers (labeled by $l=2m-1$ with $m=1,2,\ldots,n/2$) are fully packed with local minima. Our aim is to write a recursion relation. For this, it is convenient to define two probabilities $P_{2m}(x)$ and $Q_{2m-1}(x)$ defined, respectively, for even and odd layers. We define $P_{2m}(x)$ as the probability that a subtree of $2m$ generations (counted from the bottom of the tree) has a random variable x at its root and that x is *not* a local minimum. On the other hand, $Q_{2m-1}(x)$ is the probability that a subtree of $(2m-1)$ generations (again counted from the bottom of the tree) has a random variable x at its root and that x is a local minimum. It is then easy to see that they satisfy the recursion relations

$$P_{2m}(x) = \left[\int_0^x Q_{2m-1}(y) dy\right]^\mu, \tag{43}$$

$$Q_{2m+1}(x) = \left[\int_x^1 P_{2m}(y) dy\right]^\mu, \tag{44}$$

where $m=1,2,\ldots,n/2$. The recursions start with the initial condition $Q_1(x)=1$. Since n is even, the root of the full tree with n generations is not a local minimum and hence the probability of the full tree is just $P_n(x)$ given that the value at the root is x. The probability of the maximally packed configuration is then obtained by integrating over x at the root,

$$P(M_{\max}, N) = \int_0^1 P_n(x) dx. \tag{45}$$

Thus, we need to solve Eqs. (43) and (44) and then substitute the solution for $P_n(x)$ in Eq. (45) to calculate $P(M_{\max}, N)$.

To solve the nonlinear recursion relations, it is convenient to define $p_{2m}(x) = [P_{2m}(x)]^{1/\mu}$ and $q_{2m+1}(x) = [Q_{2m+1}(x)]^{1/\mu}$. Then the recursions in Eqs. (43) and (44) become

$$p_{2m}(x) = \int_0^x q_{2m-1}^\mu(y) dy, \tag{46}$$

$$q_{2m+1}(x) = \int_x^1 p_{2m}^\mu(y) dy, \tag{47}$$

starting with $q_1(x)=1$ for $0 \leq x \leq 1$. It follows from Eqs. (46) and (47) that they satisfy the boundary conditions $p_{2m}(0) = 0$ and $q_{2m+1}(1)=0$ for all $m \geq 1$. However, $p_{2m}(1)$ and

$q_{2m+1}(0)$ are nonzero. It is then useful to define the ratios $f_{2m}(x) = p_{2m}(x)/p_{2m}(1)$ and $g_{2m+1}(x) = q_{2m+1}(x)/q_{2m+1}(0)$ so that $f_{2m}(1)=1$, $f_{2m}(0)=0$ and $g_{2m+1}(0)=1$, $g_{2m+1}(1)=0$. In terms of the new functions, the recursions become

$$f_{2m}(x) = \frac{q_{2m-1}^\mu(0)}{p_{2m}(1)} \int_0^x g_{2m-1}^\mu(y) dy, \tag{48}$$

$$g_{2m+1}(x) = \frac{p_{2m}^\mu(1)}{q_{2m+1}(0)} \int_x^1 f_{2m}^\mu(y) dy, \tag{49}$$

starting with $g_1(x)=1$ for $0 \leq x \leq 1$. As m increases, we expect that the ratio functions $f_{2m}(x)$ and $g_{2m+1}(x)$ will approach their respective m-independent fixed point forms $f(x)$ and $g(x)$. This means that as $m \to \infty$, $q_{2m-1}^\mu(0)/p_{2m}(1) \to \lambda_1$, and $p_{2m}^\mu(1)/q_{2m+1}(0) \to \lambda_2$ and

$$f(x) = \lambda_1 \int_0^x g^\mu(y) dy, \tag{50}$$

$$g(x) = \lambda_2 \int_x^1 f^\mu(y) dy, \tag{51}$$

with the boundary conditions $f(0)=0$, $f(1)=1$ and $g(0)=1$, $g(1)=0$. These boundary conditions will determine the eigenvalues λ_1 and λ_2.

Differentiating Eqs. (50) and (51) with respect to x gives

$$\frac{df}{dx} = \lambda_1 g^\mu(x), \tag{52}$$

$$\frac{dg}{dx} = -\lambda_2 f^\mu(x). \tag{53}$$

Multiplying (52) by $\lambda_2 f^\mu(x)$ and (53) by $\lambda_1 g^\mu(x)$, adding, and then integrating, we find a conserved quantity

$$\lambda_2 f^{\mu+1}(x) + \lambda_1 g^{\mu+1}(x) = C \tag{54}$$

where C is a constant independent of x. Putting $x=0,1$ and using the respective boundary conditions gives $C=\lambda_1=\lambda_2$. Thus, $\lambda_1=\lambda_2=\lambda$ and

$$f^{\mu+1}(x) + g^{\mu+1}(x) = 1. \tag{55}$$

The common eigenvalue λ is yet to be determined. Eliminating $g(x)$ between Eqs. (52) and (55) gives

$$\frac{df}{dx} = \lambda [1 - f^{\mu+1}(x)]^{\mu/(\mu+1)} \tag{56}$$

subject to the boundary conditions $f(0)=0$ and $f(1)=1$. Integrating and using $f(0)=0$ we get

$$\int_0^{f(x)} \frac{dz}{(1 - z^{\mu+1})^{\mu/(\mu+1)}} = \lambda x. \tag{57}$$

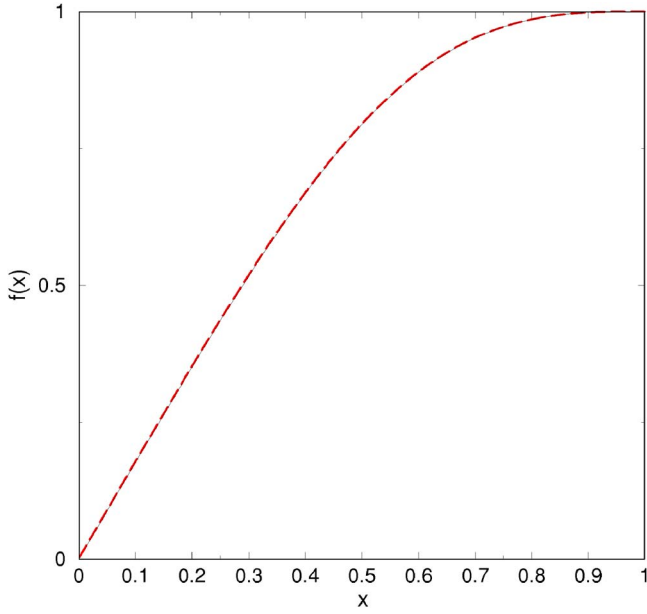Using the other boundary condition $f(1)=1$ determines λ explicitly,

FIG. 3. (Color online) The numerically obtained fixed point function $f(x)$ after ten iterations, for $\mu=2$, is compared to the analytical scaling function in Eq. (59) obtained using MATHEMATICA. The two curves are difficult to distinguish, indicating perfect agreement.

$$\lambda = \frac{1}{\mu+1}B\left(\frac{1}{\mu+1},\frac{1}{\mu+1}\right), \tag{58}$$

where $B(m,n)=\int_0^1 x^{m-1}(1-x)^{n-1}dx$ is the standard Beta function. The eigenfunction $f(x)$ in Eq. (57) can be expressed as the solution of the equation

$$f(x)F\left(\frac{1}{\mu+1},\frac{\mu}{\mu+1},\frac{\mu+2}{\mu+1},f^{\mu+1}(x)\right) = \frac{x}{\mu+1}B\left(\frac{1}{\mu+1},\frac{1}{\mu+1}\right) \tag{59}$$

where $F[a,b,c,z]$ is the hypergeometric function. The other eigenfunction $g(x)$ follows from Eq. (55), or simply from the symmetry $g(x)=f(1-x)$. To check that the solutions to the original recursion relations (46) and (47) indeed converge to these fixed solutions, we have numerically solved Eqs. (46) and (47) for $\mu=2$. We find that the numerical solutions [the ratio functions $f_{2m}(x)$ and $g_{2m+1}(x)$] converge to fixed point functions rather quickly after about three or four iterations. In Fig. 3, we compare the numerical fixed point solution (the solution after ten iterations) $f(x)$ with the analytical solution in Eq. (59) with $\mu=2$. The agreement is perfect.

To determine γ, we use Eq. (45) and the recursion relation (47), which show that

$$P(M_{\max},N) = \int_0^1 P_n(x)dx = q_{n+1}(0). \tag{60}$$

Then we take ratios to extract the bulk contribution of the Cayley tree, thereby removing artifacts coming from its surface [24]: as $m\to\infty$, we have $q_{2m-1}^\mu(0)/p_{2m}(1)\to\lambda$ and

$p_{2m}^\mu(1)/q_{2m+1}(0)\to\lambda$. Eliminating $p_{2m}(1)$ gives a recursion for large m,

$$q_{2m+1}(0) = \frac{1}{\lambda^{\mu+1}}q_{2m-1}^{\mu^2}(0). \tag{61}$$

Iterating it, we obtain, for large m, $q_{2m+1}(0) \sim \lambda^{-(\mu^{2m}-1)/(\mu-1)}$. Substituting this result in Eq. (60) and using $N=(\mu^n-1)/(\mu-1)$ gives our final result for large N:

$$P(M_{\max},N) \sim \gamma^{-N} \quad \text{with } \gamma=\lambda = \frac{1}{\mu+1}B\left(\frac{1}{\mu+1},\frac{1}{\mu+1}\right). \tag{62}$$

Note that the recursion relation (61) is valid only for large m. But in deriving the results above we have assumed that it holds even for small m. Even though the asymptotic result is not expected to change due to this "initial condition" effect, one can avoid this "surface effect" by appropriately defining γ as the ratio

$$\frac{1}{\gamma} = \lim_{n\to\infty} \frac{\int_0^1 \left[\int_x^1 P_n(y)dy\right]^\mu dx}{\left[\int_0^1 P_n(x)dx\right]^\mu}. \tag{63}$$

This definition follows from the following observation. The numerator on the RHS of (63) denotes the probability of a maximally packed configuration on a tree with $(n+1)$ generations. The denominator is the joint probability that the disconnected μ subtrees of n generations are all maximally packed. Thus, the ratio on the RHS is just the factor by which the probability of a maximally packed configuration per site decreases when one fuses the μ number of n-generation trees and one additional root to construct a newly maximally packed configuration on an $(n+1)$-generation tree. But, asymptotically for large n, this is precisely $1/\gamma$ by the orginal definition in Eq. (15). Physically, $\ln\gamma$ is the additional entropy change when packing an extra minimum in the tree. Using the recursion relations (43) and (44), in Eq. (63) one gets

$$\frac{1}{\gamma} = \lim_{n\to\infty} \int_0^1 \left[\frac{q_{n+1}(x)}{q_{n+1}(0)}\right]^\mu dx = \int_0^1 g_{n+1}^\mu(x)dx. \tag{64}$$

Substituting $g^\mu(x)=(1/\lambda)df/dx$, integrating, and using $f(1)=1$, we get

$$\gamma=\lambda = \frac{1}{\mu+1}B\left(\frac{1}{\mu+1},\frac{1}{\mu+1}\right). \tag{65}$$

Note that, for $\mu=1$, we recover the 1D result, $\gamma=\pi/2$. For $\mu=2$, we get $\gamma=B(1/3,1/3)/3=1.766\,64...$. As $\mu\to\infty$, γ converges slowly to $\gamma=2$. Since the $\mu\to\infty$ result should coincide with that on a hypercubic lattice in the infinite dimension limit, we expect that for hypercubic lattices in d dimensions γ is bounded, $\pi/2\leqslant\gamma\leqslant2$. As d increases from 1 to ∞, γ should increase monotonically from its $d=1$ value $\pi/2$
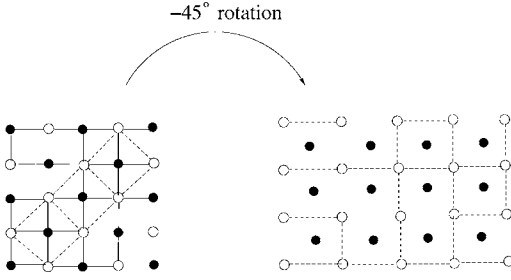
FIG. 4. A maximally packed configuration on a square lattice. On this checkerboard pattern, the black circles denote the locations of the local minimum. The corners of every plaquette around a minimum cannot contain a minimum and are shown by empty circles. On the right, the same pattern, rotated by an angle $-45°$.

$=1.560\,79\ldots$ to 2. Our numerical solutions on d-dimensional hypercubic lattices with $d=2,3,4,5$ are consistent with these bounds.

B. Maximally packed configuration on a bipartite lattice: An equivalent plaquette model

On a bipartite lattice, the maximally packed configuration is the one where one places a local minimum at every alternate site. For example, on a square lattice, a maximally packed configuration has a checkerboard pattern as shown in Fig. 4. Let us first focus on the neighborhood of a single local minimum. The sites at the corners of the plaquette containing this minimum are clearly not local minima. Let x_1, x_2, x_3, and x_4 denote the values of the random variables at the four corners of the plaquette. Then, given these four values, the probability that the site at the center of the plaquette is a local minimum is clearly

$$p_i(x_1, x_2, x_3, x_4) = \min(x_1, x_2, x_3, x_4) \tag{66}$$

where we have used the fact that the random variable at the center of the plaquette is drawn from a uniform distribution. So the probability of the full checkerboard configuration where every alternate site is a local minimum, given the values of the random variables $\{x_i\}$s at the sites of the other sublattice, is obtained by multiplying all the plaquettes,

$$\text{Prob}(M_{\max}, N | \{x_i\}) = \prod_{\text{plaquette } j} \min(x_1(j), x_2(j), x_3(j), x_4(j)) \tag{67}$$

where the plaquettes are labeled by j and $x_1(j)$, $x_2(j)$, $x_3(j)$, and $x_4(j)$ are the four random variables at the corners of the jth plaquette. The total number of plaquettes will also be denoted by M since the number of plaquettes is the same as the number of local minima. To simplify, it is convenient to rotate the lattice by $-45°$ (as shown in Fig. 4).

We then have a plaquette model where at each corner of a plaquette exists a random variable x_i drawn from a uniform distribution over $[0, 1]$, and we are interested in calculating the product in Eq. (67). Finally, the probability of the maximally packed configuration on the original lattice is obtained by averaging over the x_i's with uniform distribution in $[0, 1]$,

$$P(M_{\max}, N) = \left\langle \prod_{\text{plaquette } j} \min(x_1(j), x_2(j), x_3(j), x_4(j)) \right\rangle \tag{68}$$

where the angular brackets simply indicate integration over all the x variables from 0 to 1. Note that this plaquette model is very general and can be extended to any bipartite lattice. Also, the right-hand side of Eq. (68) is actually the probability of just one of the two checkerboard configurations, so the true $P(M_{\max}, N)$ is actually twice this amount, but to avoid unnecessary complications we shall keep to this notation.

1. Plaquette model in one dimension

As a simple example, for a 1D chain, one can reproduce the result $\gamma = \pi/2$ quite easily using the plaquette representation. The number of plaquettes is clearly $M = N/2$. In this case, Eq. (68) gives

$$P(M_{\max}, N) = \left\langle \prod_{i=1}^{M=N/2} \min(x_i, x_{i+1}) \right\rangle \tag{69}$$

where i runs over every alternate site of the original 1D chain. The quantity on the RHS of (69) can be evaluated using a simple transfer matrix approach. Defining a transfer matrix via $\langle x_i | \hat{T} | x_{i+1} \rangle = \min(x_i, x_{i+1})$, we have from Eq. (69), assuming a closed periodic chain,

$$P(M_{\max}, N) = \text{Tr}[\hat{T}^M] \tag{70}$$

where Tr is the trace. The eigenvalue equation of the transfer matrix is

$$\int_0^1 \min(x, y) \psi(y) dy = \lambda \psi(x) \tag{71}$$

where $\psi(x)$ is an eigenfunction with eigenvalue λ. Dividing the range of integration into $[0, x]$ and $[x, 1]$ and differentiating twice, one gets

$$\lambda \frac{d^2\psi}{dx^2} + \psi(x) = 0 \tag{72}$$

with the boundary conditions $\psi(x=0)=0$ and $\psi'(x=1)=0$ where $\psi'(x)=d\psi/dx$. The solution is simply $\psi(x) = A \sin(x/\sqrt{\lambda})$ where

$$\lambda = \frac{4}{\pi^2(2m+1)^2}, \quad m = 0, 1, 2 \ldots. \tag{73}$$

The largest eigenvalue is $\lambda = 4/\pi^2$ corresponding to $m=0$. Thus for large $M = N/2$, one gets from Eq. (70)

$$P(M_{\max}, N) \sim \left[\frac{4}{\pi^2} \right]^{N/2} \sim [\pi/2]^{-N}, \tag{74}$$

reproducing the results obtained in the previous section.

2. Plaquette model on a two-leg ladder

Another nontrivial solvable case of the plaquette model is on a two-leg ladder shown in Fig. 5. It can be thought of as
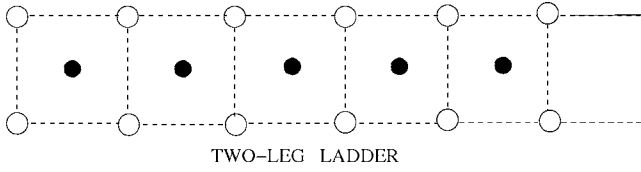
FIG. 5. The plaquette model on a two-leg ladder. At the center of each plaquette there is a local minimum shown by the black circles.

the first layer of the full 2D model shown on the right in Fig. 4. At the center of each plaquette there is a local minimum. Let M be the number of plaquettes. Then the total number of sites in the lattice (counting the minima at the centers) is $N = 3M$.

The probability of the maximally packed configuration in Eq. (67) can again be computed by defining the transfer matrix $\langle x_1, x_2 | \hat{T} | x_3, x_4 \rangle = \min(x_1, x_2, x_3, x_4)$ where (x_1, x_2) refers to the two corners on the left of a plaquette and (x_3, x_4) refers to the two corners on the right of a plaquette. Then,

$$P(M_{\max}, N) = \text{Tr}[\hat{T}^M]. \tag{75}$$

The corresponding eigenvalue equation is now two dimensional,

$$\int_0^1 dx_3 \int_0^1 dx_4 \, \min(x_1, x_2, x_3, x_4) \psi(x_3, x_4) = \lambda \psi(x_1, x_2) \tag{76}$$

which is considerably harder to solve compared to the 1D case. While one can solve the 2D eigenvalue equation directly, it is somewhat easier to first reduce it to an equivalent 1D eigenvalue problem by using the following trick.

Let us first define the random variable $z_i = \min(x_1(i), x_2(i), x_3(i), x_4(i))$ at the center of each plaquette i. Clearly

$$P(M_{\max}, N) = \langle z_1 z_2 \cdots z_M \rangle \tag{77}$$

where $M = N/3$ is the number of plaquettes. The random variables z_i are obviously correlated as two adjacent plaquettes will share two common random variables as shown in Fig. 6.

To evaluate the average on the RHS of Eq. (77) we need to know the joint distribution of the z_i's, which are correlated (see Fig. 6). This joint distribution can be explicitly computed. To see this, let us first compute the cumulative distri-
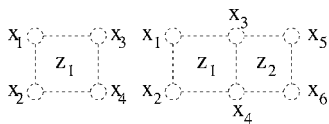

FIG. 6. On the left we have a single plaquette with the minimum $z_1 = \min(x_1, x_2, x_3, x_4)$ at the center. On the right we have two adjacent plaquettes with minima $z_1 = \min(x_1, x_2, x_3, x_4)$ and $z_2 = \min(x_3, x_4, x_5, x_6)$ at their respective centers. The variables z_1 and z_2 are correlated as they share a common bond with the two elements x_3 and x_4.

bution of one single z variable, say $z_1 = \min(x_1(1), x_2(1), x_3(1), x_4(1))$. For simplicity, we denote $x_1(1) = x_1$, $x_2(1) = x_2$, etc. (see Fig. 6). Evidently

$$\text{Prob}(z_1 > y) = \text{Prob}(\min(x_1, x_2, x_3, x_4) > y) = (1 - y)^4 \tag{78}$$

where we have used the fact that x_1, x_2, x_3, and x_4 are all independent random variables drawn from the uniform distribution over $[0, 1]$. Next, let us consider the joint distribution of two consecutive z's, say z_1 and z_2. Let us denote the random variables at the corners by x_1, x_2, x_3, x_4, x_5, and x_6 as shown in Fig. 6. Then it is easy to see that

$$\text{Prob}(z_1 > y_1, z_2 > y_2) = \text{Prob}(\min(x_1, x_2, x_3, x_4)$$
$$> y_1, \min(x_3, x_4, x_5, x_6) > y_2)$$
$$= (1 - y_1)^2 [1 - \max(y_1, y_2)]^2 (1 - y_2)^2. \tag{79}$$

In a similar way, one can construct the joint distribution of three adjacent plaquette minima,

$$\text{Prob}(z_1 > y_1, z_2 > y_2, z_3 > y_3) = (1 - y_1)^2 [1 - \max(y_1, y_2)]^2 [1 - \max(y_2, y_3)]^2 (1 - y_3)^2. \tag{80}$$

One can repeat the process above for a higher number of adjacent plaquettes and one immediately sees the pattern for the full ladder. Let us denote

$$F(y_1, y_2, y_3, \ldots, y_M) = \text{Prob}(z_1 > y_1, z_2 > y_2, z_3 > y_3, \ldots, z_M > y_M). \tag{81}$$

Then, for a ladder with open ends, we get

$$F(y_1, y_2, y_3, \ldots, y_M)$$
$$= (1 - y_1)^2 \left(\prod_{i=2}^M [1 - \max(y_{i-1}, y_i)]^2 \right) (1 - y_M)^2. \tag{82}$$

For a ladder with closed (periodic) ends, this is even simpler,

$$F(y_1, y_2, y_3, \ldots, y_M) = \prod_{i=1}^M [1 - \max(y_{i-1}, y_i)]^2, \tag{83}$$

where one identifies $y_0 = y_M$.

Once we have the joint distribution, it is easy to rewrite the average on the RHS of Eq. (77) in terms of the joint distribution

$$P(M_{\max}, N) = \langle z_1 z_2 \cdots z_M \rangle$$
$$= \int_0^1 \cdots \int_0^1 dy_1 dy_2 \cdots dy_M F(y_1, y_2, y_3, \ldots, y_M). \tag{84}$$

The latter identity can be easily derived using integration by parts. Let us, for simplicity, consider a plaquette with closed ends. Then, substituting the joint distribution from Eq. (83) into Eq. (84), we now have a one-dimensional multiple integral to perform:

$$P(M_{\text{max}},N) = \int_0^1 \cdots \int_0^1 dy_1 dy_2 \cdots dy_M$$
$$\times \prod_{i=1}^M [1 - \max(y_{i-1},y_i)]^2. \quad (85)$$

Using the identity $1 - \max(y_1,y_2) = \min(1-y_1,1-y_2)$ and making a change of variable $x_i = 1 - y_i$, the integral in Eq. (85) simplifies further:

$$P(M_{\text{max}},N) = \int_0^1 \cdots \int_0^1 dx_1 dx_2 \cdots dx_M \prod_{i=1}^M [\min^2(x_{i-1},x_i)]. \quad (86)$$

This one-dimensional integral can now be performed by a transfer matrix technique which operates in one dimension. Defining $\langle x_{i-1}|\hat{T}|x_i \rangle = \min^2(x_{i-1},x_i)$ we get $P(M_{\text{max}},N) = \text{Tr}[\hat{T}^M]$ and the eigenvalue equation is given by

$$\int_0^1 \min^2(x,y) \psi(y) dy = \lambda \psi(x). \quad (87)$$

Thus we have managed to reduce a two-dimensional eigenvalue problem in Eq. (76) to an equivalent but much simpler 1D eigenvalue problem in Eq. (87), which can then be solved exactly.

To proceed, we first divide the range of integration in Eq. (87) into $[0,x]$ and $[x,1]$. Next we differentiate once to get

$$\lambda \frac{d\psi}{dx} = 2x \int_x^1 \psi(y) dy. \quad (88)$$

Note the boundary conditions that emerge from Eqs. (87) and (88): $\psi(x=0)=0$ and $\psi'(x=0)=0$. But in addition one also has to satisfy $\psi'(x=1)=0$. Dividing Eq. (88) by x and differentiating once more we get an ordinary second-order differential equation

$$\psi''(x) - \frac{1}{x}\psi'(x) + \frac{2x}{\lambda}\psi(x) = 0 \quad (89)$$

with the boundary conditions $\psi(x=0)=0$, $\psi'(x=0)=0$, and $\psi'(x=1)=0$. The general solution to this equation, after a few changes of variables, can fortunately be obtained explicitly as a linear combination of two independent Bessel functions. The boundary condition $\psi(0)=0$ rules one of them out. Then, the most general solution satisfying $\psi(0)=0$ can be written as

$$\psi(x) = A x J_{2/3}\left(\sqrt{\frac{8x^3}{9\lambda}}\right) \quad (90)$$

where $J_\nu(x)$ is the ordinary Bessel function with index ν [25] and A is an arbitrary amplitude. Note that the other boundary condition at $x=0$, namely, $\psi'(0)=0$, is automatically satisfied by the solution in Eq. (90). This can be seen by using the small-x expansion of $J_\nu(x) \sim x^\nu$ which indicates $\psi(x) \sim x^2$ as $x \to 0$. Hence $\psi'(0)=0$.

To determine the eigenvalue λ, we have to use the other nontrivial boundary condition at $x=1$, namely, $\psi'(x=1)=0$. This condition, substituted in Eq. (90), gives us an implicit equation for λ that looks a bit complicated:

$$\frac{2}{3}J_{2/3}\left(\sqrt{\frac{8}{9\lambda}}\right) + \sqrt{\frac{8}{9\lambda}}J'_{2/3}\left(\sqrt{\frac{8}{9\lambda}}\right) = 0. \quad (91)$$

However, a nice simplification occurs when one uses the identity [25] $xJ'_\nu(x) + \nu J_\nu(x) = xJ_{\nu-1}(x)$. Then Eq. (91) simply gives

$$J_{-1/3}\left(\sqrt{\frac{8}{9\lambda}}\right) = 0. \quad (92)$$

The Bessel function oscillates on the positive axis of its argument, so each of its zeros (roots) would give an eigenvalue λ. However, for large M, we are interested only in the largest eigenvalue, which is then given by

$$\lambda = \frac{8}{9\alpha^2} \quad (93)$$

where α is smallest positive root of $J_{-1/3}(x)=0$. The root α is known [26], $\alpha = 1.866\,35\ldots$.

Using $P(M_{\text{max}},N) = \text{Tr}[\hat{T}^M] \sim \lambda^M$ for large M and using $M = N/3$, we finally get the exact result for the two-leg ladder,

$$P(M_{\text{max}},N) \sim \gamma^{-N} \quad \text{where } \gamma = \left[\frac{9\alpha^2}{8}\right]^{1/3} = 1.576\,57\ldots. \quad (94)$$

When compared with the 1D result, $\gamma = \pi/2 = 1.570\,79\ldots$, we see that γ changes by a very small amount as one goes from a chain to a ladder. It would be interesting to see if one can extend these calculations to ladders with more than two legs [27] and eventually to the full two-dimensional lattice.

V. NUMERICAL RESULTS FOR MAXIMALLY PACKED *d*-DIMENSIONAL LATTICES

Now we consider the dependence of γ on dimension. For simplicity, we have focused on d-dimensional hypercubic lattices as these lattices are simple to parametrize and are bipartite, allowing us to use the d-dimensional analog of the framework described in Sec. IV B. Our approach will be computational: we numerically estimate $P(M_{\text{max}},N)$ from an integral representation for lattices of increasing size and then try to extract the limit of large sizes. To avoid boundary effects which slow down the convergence to this limit, we have used L^d lattices having periodic boundary conditions in all directions.

The integral in Eq. (68) is based on decomposing the lattice into even and odd sites and imposing the local minima to be say on the even sites. Its generalization to d dimensions involves all the neighbors of a given even site j, the corresponding "star" set of $2d$ odd sites playing the role the plaquette had for the square lattice:

$$P(M_{max},N) = \left\langle \prod_{star j} \min(x_1(j), x_2(j), \ldots, x_{2d}(j)) \right\rangle. \quad (95)$$

Here the average is over all values of the random variables belonging to stars; these are uniform identical independently distributed (i.i.d.) in $[0, 1]$. Furthermore, as before, we neglect the factor 2 in $P(M_{max},N)$ coming from the fact that the minima could have been taken to be on the odd sites. The difficulty in computing these integrals is their high dimensionality. In Monte Carlo calculations as used in most statistical physics applications, it is straightforward to use importance sampling methods (e.g., the METROPOLIS algorithm) to get expectation values of observables; unfortunately here, the quantity to compute is the analog of the free energy and it is not directly accessible via such methods. We have thus used a different approach that is of the "variance reduction" type. It can be motivated as follows. In Eq. (95) we are to get the mean value of the integrand, sampling the random variables x_i uniformly. It is quite easy to see that the signal-to-noise ratio goes to zero exponentially with the number of lattice sites; to counteract this, we sample the x_i with a different density and then correct for this biased sampling. For this to be practical, we keep the x_i as i.i.d. variables, but optimize their individual distribution so as to maximize the signal-to-noise ratio. Let $\rho(x)$ be the probability density used for sampling the x_i. For any such distribution,

$$P(M_{max},N) = \left\langle \prod_{star\ j} \frac{\min(x_1(j), x_2(j), \ldots, x_{2d}(j))}{[\rho(x_1(j))\rho(x_2(j)) \cdots \rho(x_{2d}(j))]^{1/2d}} \right\rangle_\rho. \quad (96)$$

The denominator corrects for the modified measure of the random variables, taking into account the fact that each odd site appears in $2d$ stars. If ρ is well chosen, the numerator and denominator of the integrand will fluctuate together so that their ratio has a reduced variance. For our purposes, we parametrized ρ as follows:

$$\rho(x) = \begin{cases} Ax^\phi, & x \leqslant x^*, \\ B, & x \geqslant x^*, \end{cases} \quad (97)$$

where ϕ and x^* are arbitrary parameters while A and B are set so that ρ is continuous and is a normalized probability density. For each dimension, we adjust ϕ and x^* to minimize the variance of the integrand; the signal-to-noise ratio still decreases exponentially with the number of lattice sites, but with a smaller exponent. The motivation for this functional form is simple: the odd sites, all of which are maxima, have an *a posteriori* distribution that is strongly suppressed at low values of the random variable.

One last obstacle comes from the fact that the integrand typically takes on small values when one takes large lattices; this is expected, of course, since the integral itself is becoming exponentially small. To keep track of values that are smaller than what can be represented by the machine coding, (we used 96-bit representations of real numbers), we shifted multiplicatively each star term in the integrand and corrected for this shift when computing $\ln[P(M_{max},N)]$. The set of
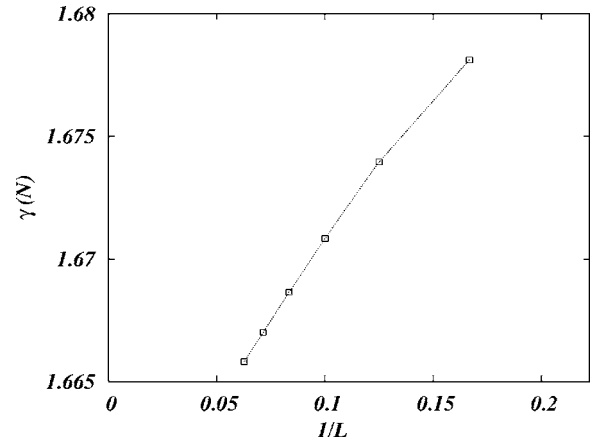

FIG. 7. The convergence to the large L limit of the estimates of $\gamma(N)$ for the square lattice, $N = L^2$.

these procedures then gave us values of $\ln[P(M_{max},N)]/N$ with measurable statistical errors for a range of $N = L^d$ from which we extrapolated to the large N limit.

In practice, we find that this strategy works very well in low dimensions. For instance, in dimension $d = 1$, we still have a very good precision for $L = 30$ sites, and the large L limit can be very reliably extracted, giving γ to better than five significant figures. (Of course, since the exact value is known at $d = 1$, this really only serves as a check of our procedures.) In dimension $d = 2$, the method gives better than four significant figures for $L \approx 16$. Our estimates of $\gamma(N) = P(M_{max},N)^{-1/N}$ as a function of $1/L = 1/\sqrt{N}$ are shown in Fig. 7. The convergence to the $L = \infty$ limit seems to follow a $1/L$ law, a property we also find for the higher dimensions investigated. Here we see that $\gamma(N)$ converges rapidly to a value close to 1.658. Unfortunately the signal-to-noise ratio decreases with L and with dimension; furthermore the accessible range of L decreases fast as d increases; because of this, we are able to extract γ reliably only for dimensions up to $d = 5$. Our results are summarized in Table I where the error estimates come from both statistical noise and uncertainties in the large-L extrapolation.

VI. DISCUSSION AND CONCLUSIONS

We have been concerned with the statistical properties of M, the number of local minima in a random energy landscape. The low-order moments of M can easily be obtained; higher-order moments could be computed by automated

TABLE I. The numerical estimates of γ for d-dimensional hypercubic lattices ($\gamma_{hypercube}$); error estimates come from statistical fluctuations and uncertainties in the large lattice size extrapolation. Bottom line: exact values for the Cayley tree with the same coordination number $\mu + 1 = 2d$.

	$d=1$	$d=2$	$d=3$	$d=4$	$d=5$
$\gamma_{hypercube}$	1.57082(6)	1.6577(6)	1.7152(10)	1.761(5)	1.806(10)
γ_{Cayley}	1.570796	1.854075	1.927621	1.956922	1.971464

counting of graphs. The atypical values of M follow a large-deviation principle as given in Eq. (13); we have been able to compute this function $\Phi(y)$ in one dimension analytically; it diverges as $y \rightarrow 0$ and goes to a finite limit when M reaches its maximum value where one-half of the lattice sites are local minima. This then led us to consider the limit of maximum packing (M takes on its largest possible value) for more general lattices. We focused on bipartite lattices where up to half of the sites can be local minima. We derived analytically the probability of this maximum packing $P(M_{max}, N)$ for the Cayley tree and for a two-leg ladder. We then tackled d-dimensional hypercubic lattices by computational techniques. For all these lattices, it is easy to see that $P(M_{max}, N) \geqslant 2^{-N}$ simply by forcing the x_i on the even (odd) sites to be less (greater) than $1/2$; this immediately leads to $\gamma(d) \leqslant 2$ for all dimensions. Furthermore, when d becomes large, it will be very rare to have maximum packing if the x_i do not very nearly satisfy this even-odd pattern so one expects $\gamma(d)$ to tend toward 2 in the large-d limit.

Given this large-d limit, it is natural to ask how $\gamma = 2$ is approached. In Fig. 8 we show that the Cayley tree case follows a very clear power correction law which can be derived analytically as being $1/d^2$. The case of the d-dimensional lattice is less clear but is compatible with a $1/d$ law.

The Cayley tree thus provides an approximation to the hypercubic case but not a very accurate one; this is probably because the nature of the correlations from site to site of the x_i (given that one has a maximally packed configuration) are quite different when considering the tree rather than the hypercube.

Our work can be extended in several ways. (i) We computed the exact large-deviation function in one dimension, but this function can also be determined for the Cayley tree [28]. (ii) One can also introduce a chemical potential z for each minimum and consider the thermodynamics of minima as a function of z; recent numerical results by Derrida in 2D [29] indicate the presence of a phase transition with z. (iii) Our variables x_i were i.i.d. random variables; does a large-deviation principle still hold if these variables have short-
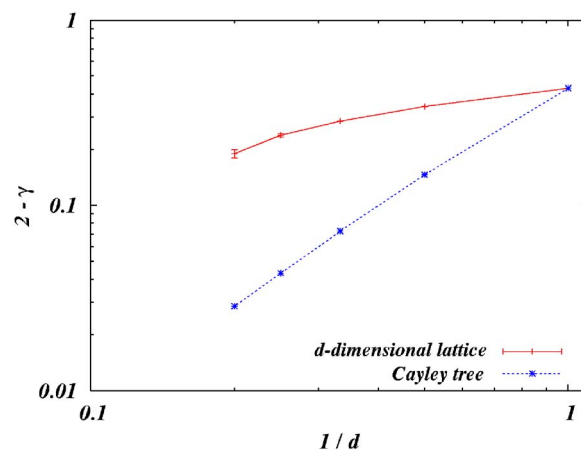

FIG. 8. (Color online) The convergence to the large-d limit of the estimates of $\gamma(d)$ vs $1/d$: plotted is the difference $2 - \gamma(d)$ which vanishes as a power at large d.

range correlations? One expects so. (iv) How do all these statistical properties generalize if one asks for minima within a given energy range? Finally, it would also be of interest to understand the similarities and differences between the statistics of local minima in continuous random energy landscapes [8,9,30] and in lattice models as presented here; the energies could be random or correlated as in the Sinai problem.

ACKNOWLEDGMENTS

[1] J.-L. Barrat, M. Feigelman, J. Kurchan, and J. Dalibard, *Slow Relaxations and Nonequilibrium Dynamics in Condensed Matter: Les Houches Session LXXVII* (Springer, Heidelberg, Berlin, Germany, 2002).

[2] L. Susskind, e-print hep-th/0302219.

[3] A. Aazami and R. Easther, J. Cosmol. Astropart. Phys. 03 (2006) 013.

[4] S. Gavrilets, *Fitness Landscapes and the Origin of Species* (Princeton University Press, Princeton, NJ, 2004).

[5] L. Mersini-Houghton, Class. Quantum Grav. **22**, 3481 (2005).

[6] T. Halpin-Healy and Y. C. Zhang, Phys. Rep. **254**, 215 (1995).

[7] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin-Glass Theory and Beyond*, Lecture Notes in Physics Vol. 9 (World Scientific, Singapore, 1987).

[8] Y. V. Fyodorov, Phys. Rev. Lett. **92**, 240601 (2004).

[9] A. Cavagna, J. P. Garrahan, and I. Giardina, Phys. Rev. E **59**, 2808 (1999).

[10] A. Cavagna, J. P. Garrahan, and I. Giardina, Phys. Rev. B **61**, 3960 (2000).

[11] D. S. Dean and S. N. Majumdar, Phys. Rev. Lett. **97**, 160201 (2006).

[12] D. Gross and M. Mézard, Nucl. Phys. B **240**, 431 (1984).

[13] B. Derrida, Phys. Rev. Lett. **45**, 79 (1980).

[14] D. Wakes, Science **285**, 1368 (1999).

[15] C. A. Angell, Science **267**, 1924 (1995).

[16] J. Fontanari and P. Stadler, J. Phys. A **35**, 1509 (2002).

[17] J. Stembridge, Trans. Am. Math. Soc. **349**, 763 (1997).

[18] G. Oshanin and R. Voituriez, J. Phys. A **37**, 6221 (2004).

[19] F. Hivert, S. Nechaev, G. Oshanin, and O. Vasilyev, e-print cond-mat/0509584.

[20] Z. Burda, A. Krzywicki, O. C. Martin, and Z. Tabor, Phys. Rev. E **73**, 036110 (2006).

[21] B. Derrida and E. Gardner, J. Phys. C **19**, 2253 (1986).

[22] S. N. Majumdar, Phys. Rev. E **65**, 035104(R) (2002).

[23] S. N. Majumdar and D. Dhar, Phys. Rev. E **64**, 046123 (2001).

[24] R. Baxter, *Exactly Solvable Models in Statistical Mechanics* (Academic Press, London, 1989).

[25] I. Gradshteyn and I. Ryzhik, *Tables of Integrals, Series, and Products* (Academic, San Diego, 1965).

[26] *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, edited by M. Abramowitz and I. Stegun, (Dover, New York, 1964).

[27] M. Evans, S. N. Majumdar, and K. Mallick (unpublished).

[28] P. Sollich, S. N. Majumdar, and A. Bray (unpublished).

[29] B. Derrida (unpublished).

[30] Y. Fyodorov, Acta Phys. Pol. B **36**, 2699 (2005).