Energy landscape of a spin-glass model: Exploration and characterization

Qing Zhou[1,*] and Wing Hung Wong[2,†]
[1]*Department of Statistics, University of California, Los Angeles, California 90095, USA*
[2]*Department of Statistics and Department of Health Research and Policy, Stanford University, Stanford, California 94305, USA*
(Received 30 March 2009; published 18 May 2009)

The disconnectivity graph (DG) is widely used to represent energy landscapes. Although powerful numerical methods have been developed to construct DGs for continuous potential-energy surfaces, they have difficulties in applications to discrete Hamiltonians as the case of spin-glass models. When the configuration space is large, brute force enumeration of all configurations to build a DG is not practical. We propose an alternative approach to construct DGs based on recursive partition of Monte Carlo samples from microcanonical ensembles. To characterize energy landscapes, we define the local density of states (LDOS) on a DG, with which one can compute many thermodynamic properties over local energy basins for any temperature. Estimation of LDOS is developed with DG construction. We further propose the concepts of tree entropy and local escape probability, both of which are functions of local density of states, to capture the symmetry and the roughness of a Boltzmann distribution, respectively. Our approach is applied to a study of the Sherrington-Kirkpatrick spin-glass model with N varying between 20 and 100 spins. We observe that the energy landscape is extremely asymmetric and there exists a sharp increase in local escape probability preceding the transition from spin glass to paramagnetic phase.

DOI: 10.1103/PhysRevE.79.051117 PACS number(s): 05.30.−d, 05.10.Ln, 02.50.Ng, 75.10.−b

I. INTRODUCTION

To understand complex systems, theoretical and numerical studies have been conducted to characterize the energy landscapes of spin-glass models [1–3]. Theoretical results on some characteristics of the energy landscapes were obtained, such as properties of ground states and metastable states of Ising spin glasses [4–6]. Numerical analyses have also been performed to capture various aspects of spin glasses, including local energy minima [7–9], valley structures [10,11], low-energy dynamics [12], and folding properties [13] among others. To explore a global picture, disconnectivity graphs (DGs) or barrier trees [14–16] were constructed for spin glasses [17–20] with further extensions to disconnectivity networks [21] and free-energy landscapes [22–24]. However, due to the discrete nature of the Hamiltonians [1,2] and the exponential complexity of spin glasses, there are difficulties in constructing and characterizing the DGs of their energy landscapes for large systems. Although numerical approaches to DG construction have been developed via the identification of stationary points on a potential-energy surface in chemical physics (e.g., [25,26]), they are not directly applicable to spin glasses since gradients and eigenvectors are not well defined for a discrete Hamiltonian. Therefore, previous works on DGs of spin-glass models have been focused on moderate system sizes of 20–30 spins, e.g., [17,18,20,21]. For such systems, it is possible to detect all critical states (local minima and energy barriers) via enumeration of all configurations or all low-energy configurations. Valuable insights on the energy landscapes of spin glasses have been provided by these studies.

Clearly, extensive enumeration of all local minima and barriers to construct a complete DG of a spin-glass model is not feasible for systems with more than 40 spins. In this work, we make an attempt to study main features of spin-glass energy landscapes rather than building a full DG. We simulate Monte Carlo samples of configurations at a wide energy range via the equienergy (EE) sampler [27], which is able to construct microcanonical ensembles of a Hamiltonian. Then we estimate a DG with the samples from all microcanonical ensembles through a bottom-up partition (BUP) algorithm. In this way, the estimated DG represents a meaningful approximation to the underlying energy landscape because it is constructed with Monte Carlo samples and thus naturally emphasizes on more important regions (with high Boltzmann probabilities) in the configuration space. Furthermore, we link the concept of the density of states to a DG via the definition of local density of states. The local density of states (LDOS) gives the number of configurations located in a particular energy basin or superbasin for a given energy level (to be defined more precisely later). Given LDOS, one can conveniently compute Boltzmann averages over various local basins for any temperature, which provides a useful means to characterize statistical and thermodynamic properties of the underlying Hamiltonian. We define tree entropy and local escape probability based on LDOS for a global characterization of the symmetry and the roughness of energy landscapes, respectively. These statistics are expected to convey more quantitative information for energy landscape characterization than classifying DGs by their shapes and topologies (e.g., [28]).

In this study we focus on the energy landscape of the Sherrington and Kirkpatrick (SK) spin glass [2,29]. In this model, a piece of spin glass is represented by N spins, $\boldsymbol{S}=\{S_1,S_2,\ldots,S_N\}$, each of which has two orientations $\{\pm 1\}$. A random interaction J_{ij} from a Gaussian distribution $\mathcal{N}(J_0,J^2)$ is assigned independently to each pair of spins S_i and S_j. Given the interactions, the Hamiltonian of the system is

*zhou@stat.ucla.edu
†whwong@stanford.edu

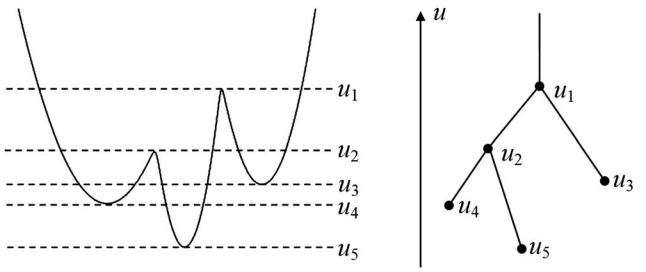

FIG. 1. A hypothetical energy function (left) and its DG (right).

$$h(\boldsymbol{S}) = -\sum_{i<j} J_{ij} S_i S_j, \tag{1}$$

in which the summation is running over all pairs of spins (not only nearest neighbors). Accordingly, the Boltzmann distribution of the SK spin-glass model at temperature T is

$$p_B(\boldsymbol{S};T) = \frac{1}{Z(T)} \exp[-h(\boldsymbol{S})/T], \tag{2}$$

where $Z(T)=\Sigma \exp[-h(\boldsymbol{S})/T]$ is the partition function. Throughout the paper, the Boltzmann constant k_B is absorbed into the temperature T as in Eq. (2).

This paper is organized into six sections. We introduce LDOS, tree entropy, and local escape probability in Sec. II after a brief review of the disconnectivity graph. The bottom-up partition algorithm for estimating DGs and LDOS from microcanonical ensembles is described in Sec. III. We illustrate our method in Sec. IV on a series of energy functions for which theoretical calculations are available. In Sec. V we study the energy landscape and thermodynamics of SK spin glasses for N ranging between 20 and 100. The paper concludes with discussions in Sec. VI.

II. LOCAL DENSITY OF STATES

In this section we first review the disconnectivity graph of an energy function and introduce necessary notations. The local density of states is then defined on a DG based on which we will propose two quantities, the tree entropy and the local escape probability, to characterize energy landscapes.

As introduced in [14], a DG summarizes the hierarchy among various basins and superbasins in the configuration space. The topology of a DG is a tree as illustrated in Fig. 1. To connect a DG with the density of states, we reformulate an equivalent definition in terms of the disconnectedness among sublevel sets of the energy function,

$$A(u) \triangleq \{\boldsymbol{x} | h(\boldsymbol{x}) < u\}, \tag{3}$$

for all energy levels u. Suppose that the sublevel set $A(u)$ contains $K(u)$ connected components $\{A_k(u) | k=1,\ldots,K(u)\}$. The root of the tree is defined at the energy level u_1 where the sublevel set $A(u_1)$ becomes disconnected, i.e., $u_1 = \inf\{u | A(u) \text{ is connected}\}$. We then define its kth child node, for $k=1,\ldots,K(u_1)$, at the energy level u_{1k} such that $A(u_{1k}) \cap A_k(u_1)$ becomes disconnected or reaches a local minimum. That is, $u_{1k} = \inf\{u | A(u) \cap A_k(u_1)$ is connected and nonempty$\}$. In the figure, the child nodes of u_1 are denoted by u_2 and u_3. A recursive application of the above definition to every internal node gives the tree topology of a DG. The leaves (terminal nodes) of the tree correspond to the local minima of $h(\boldsymbol{x})$ and the internal nodes correspond to the energy barriers that separate the minima. In this paper, a DG and its corresponding tree are used as interchangeable terms.

A. Definition of LDOS

One of the most important concepts in statistical physics is the density of states $\Omega(u)$, i.e., the number of all configurations for an energy level u. Many thermodynamic quantities can be calculated directly with $\Omega(u)$ for any temperature T. For example, the partition function $Z(T)$, the internal energy $U(T)$, and the specific heat $C(T)$ can be calculated by

$$Z(T) = \int \Omega(u) e^{-u/T} du,$$

$$U(T) = \frac{1}{Z(T)} \int u \Omega(u) e^{-u/T} du,$$

$$C(T) = \frac{\partial U(T)}{\partial T} = \frac{1}{T^2} \left[\frac{\int u^2 \Omega(u) e^{-u/T} du}{Z(T)} - \{U(T)\}^2 \right], \tag{4}$$

in which only univariate integrals, or summations for discrete Hamiltonians, are involved once $\Omega(u)$ is provided.

Although extremely useful, the density of states can only facilitate the computation of overall Boltzmann averages such as Eq. (4). To obtain a full picture of the energy landscape of a Boltzmann distribution, it is desired to compute thermodynamic and statistical quantities over various local domains in the configuration space. Such local domains can be defined by edges on a disconnectivity graph. For example, an edge with a terminal node as one of its end points corresponds to a unique local domain of a local minimum. On the tree in Fig. 1, the edge between nodes u_1 and u_3, denoted by $E(u_1,u_3)$, represents a local basin around the local minimum u_3. We call it the (connected) basin of the local minimum. Statistical properties of connected basins are useful for characterizing the behaviors of local minima (metastable states). For this purpose, we introduce the definition of the local density of states. Index all the edges on a tree by a set of integers $\mathcal{I}=\{1,\ldots,K_e\}$. For any configuration $\boldsymbol{x}$, we denote by $I(\boldsymbol{x})$ the index of the edge to which $\boldsymbol{x}$ belongs. For example, a configuration in the connected basin of the local minimum u_3 belongs to the edge $E(u_1,u_3)$. Note that the vertical axis of a DG represents energy level u (Fig. 1). All the configurations in a microcanonical ensemble,

$$\mathcal{X}(u) \triangleq \{\boldsymbol{x} | h(\boldsymbol{x}) = u\}, \tag{5}$$

can be partitioned according to their connectivity on the tree. That is, configurations in $\mathcal{X}(u)$ that can be connected via pathways with energy $\leq u$ are identified into one group,

which maps to a point on an edge with index i at energy u. Such a group of configurations is called a connected microcanonical ensemble (CME)

$$\mathcal{X}^{(i)}(u) \triangleq \{\boldsymbol{x}|h(\boldsymbol{x})=u, I(\boldsymbol{x})=i\}, \quad i \in \mathcal{I}. \tag{6}$$

The local density of states $\Omega^{(i)}(u)$ is defined as the number of configurations in $\mathcal{X}^{(i)}(u)$. On the tree of Fig. 1, if the index of the edge $E(u_1,u_3)$ is i and $u_3<u<u_1$, then $\Omega^{(i)}(u)$ is the number of configurations in the connected basin of the local minimum u_3 for the energy level u.

If local density of states is provided for all energies and edges, one can compute directly many useful statistical and thermodynamic quantities of local basins and domains defined by an energy landscape at any temperature. Denote the two end-point energies of edge i by $L_i<B_i$. The local partition function $Z^{(i)}(T)$, the probability mass $P^{(i)}(T)$, and the local energy $U^{(i)}(T)$ of the edge at temperature T can be defined and computed by

$$Z^{(i)}(T) = \int_{L_i}^{B_i} \Omega^{(i)}(u) e^{-u/T} du, \tag{7}$$

$$P^{(i)}(T) = \frac{Z^{(i)}(T)}{Z(T)}, \tag{8}$$

$$U^{(i)}(T) = \frac{1}{Z^{(i)}(T)} \int_{L_i}^{B_i} u\Omega^{(i)}(u) e^{-u/T} du. \tag{9}$$

Such quantities, called local Boltzmann averages, provide basic information for characterizing the underlying energy landscape. Furthermore, we define

$$V^{(i)} = \int_{L_i}^{B_i} \Omega^{(i)}(u) du \tag{10}$$

as the volume of edge i. For example, the volume of the edge $E(u_1,u_3)$ in Fig. 1 gives the number of configurations in the connected basin of u_3 and its local probability mass (8) reports the probability of observing configurations in this basin under the Boltzmann distribution at temperature T.

B. Tree entropy

We define the tree entropy to measure the degree of symmetry of energy landscapes via the LDOS. From Eq. (6) one sees that the microcanonical ensemble $\mathcal{X}(u)$ is partitioned into a collection of CMEs $\mathcal{X}^{(i)}(u)$ for each energy level u. The number of CMEs for the energy level u equals the number of connected components of the sublevel set $A(u)$ (3), denoted by $K(u)$. On a DG, $K(u)$ is the number of intersections between the tree and the horizontal line at the energy level u. For example, $K(u)=2$ for $u_2<u<u_1$ on the tree of Fig. 1. If we use $\mathcal{I}(u)$ to denote the collection of indices of the edges that intersect with the horizontal line at energy u, then its size $|\mathcal{I}(u)|=K(u)$ and we have the following relationship:

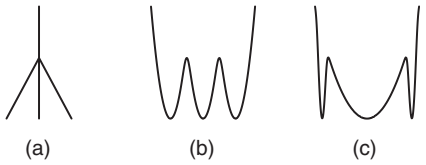

FIG. 2. An illustration of the tree entropy.

$$\Omega(u) = \sum_{i \in \mathcal{I}(u)} \Omega^{(i)}(u). \tag{11}$$

Since a Boltzmann distribution $p_B(\boldsymbol{x};T) \propto \exp[-h(\boldsymbol{x})/T]$ restricted to any microcanonical ensemble $\mathcal{X}(u)$ is uniform for any temperature, the probability that a configuration in $\mathcal{X}(u)$ belongs to the CME $\mathcal{X}^{(i)}(u)$ is simply $\Omega^{(i)}(u)/\Omega(u)$ for $i \in \mathcal{I}(u)$. The entropy of this discrete distribution measures the degree of symmetry of the tree at energy u in terms of the relative sizes of the CMEs. Using bits as the unit, we define the entropy of the tree at energy u as

$$S_{tr}(u) = -\sum_{i \in \mathcal{I}(u)} \frac{\Omega^{(i)}(u)}{\Omega(u)} \log_2 \left\{ \frac{\Omega^{(i)}(u)}{\Omega(u)} \right\}, \tag{12}$$

which varies between 0 and $\log_2 K(u)$ bits. If the LDOS $\Omega^{(i)}(u)$ is exactly identical for all CMEs, $S_{tr}(u)$ reaches its maximal value $\log_2 K(u)$. If the density of states $\Omega(u)$ is dominated by one of the CMEs, then $S_{tr}(u)$ is close to 0.

We further define the (average) entropy of the whole tree as

$$S_{tr} = \frac{1}{B-u_0} \int_{u_0}^{B} S_{tr}(u) du, \tag{13}$$

where u_0 is the energy of the global minimum and B is the energy of the highest barrier. The tree entropy S_{tr} captures symmetry beyond the shape of a tree. The tree in Fig. 2(a) has three terminal edges of identical end-point energies and, thus, the topology is symmetric. The energy functions in Figs. 2(b) and 2(c) have exactly the same tree topology as shown in Fig. 2(a), but the levels of their tree entropy are very different. The tree entropy of Fig. 2(b) is high because the volumes of the three basins [Eq. (10)] are comparable, but the tree entropy of Fig. 2(c) is close to zero because the volume of the middle basin is much larger than the volumes of the other two basins.

C. Local escape probability

Next we propose an index to quantify the roughness of a Boltzmann distribution at a given temperature T. Suppose that we are sampling from $p_B(\boldsymbol{x};T) \propto \exp[-h(\boldsymbol{x})/T]$ via a Metropolis algorithm [30]. If a proposed Metropolis move results in an energy increase of Δu, the probability of accepting this proposal is $\exp(-\Delta u/T)$. Assuming that our proposal gives equal chance to reach all possible levels of energy increment, the probability distribution for an (accepted) energy increment is

$$p(\Delta u;T) = \frac{1}{T}\exp(-\Delta u/T), \quad \Delta u > 0. \tag{14}$$

Consider edge i on the tree of a DG with end-point energies $L_i < B_i$. Suppose that the current configuration $\boldsymbol{x}$ is on this edge with energy u, i.e., $\boldsymbol{x} \in \mathcal{X}^{(i)}(u)$. Under the assumption of Eq. (14), the probability for the Metropolis sampler to escape from the edge and reach an energy level higher than the barrier B_i by an energy increment is $\exp[-(B_i-u)/T]$. We call such an event a local escape. Averaged over all possible energy levels on edge i, the probability of a local escape on this edge at temperature T is

$$P_{LE}^{(i)}(T) = \int_{L_i}^{B_i} e^{-(B_i-u)/T}\frac{1}{Z^{(i)}(T)}\Omega^{(i)}(u)e^{-u/T}du = \frac{V^{(i)}e^{-B_i/T}}{Z^{(i)}(T)}, \tag{15}$$

where $Z^{(i)}(T)$ is the local partition function (7) and $V^{(i)}$ is the volume (10) of the edge.

Because the edge with the smallest local escape probability will be the bottleneck for the communication of the Metropolis sampler between different local basins, the minimum local escape probability

$$P_{LE}(T) = \min_{i\in\mathcal{I}} P_{LE}^{(i)}(T) \tag{16}$$

gives a meaningful index for the roughness of the distribution at temperature T. The smaller the $P_{LE}(T)$, the higher the roughness of the Boltzmann distribution. In addition, we define

$$T_E(\alpha) = \max\{T|P_{LE}(T) \le \alpha\} \tag{17}$$

as the escape temperature at level $\alpha \in (0,1]$. Intuitively, the chance of a local escape will be greater than α if the temperature is as high as $T_E(\alpha)$.

III. METHODS

A. Equienergy sampler

To construct a reasonable estimate of the DG of an underlying energy landscape, we need to have Monte Carlo samples of configurations at a wide range of energy levels, which ideally cover all important portions of the energy space (say, from the global minimum to the highest energy barrier). Examples of Monte Carlo methods which may achieve this goal include the multicanonical ensemble methods [31–35], the flat histogram methods [36,37], stochastic approximation for Monte Carlo [38,39], parallel tempering [40], exchange Monte Carlo [41], and the EE sampler [27] among others. In this work, we employ the EE sampler to simulate Monte Carlo samples for the construction of a DG. Given an energy ladder $H_1 < \cdots < H_L$ and a temperature ladder $T_1 < \cdots < T_L$, the EE sampler generates samples from a sequence of Boltzmann distributions

$$p_B(\boldsymbol{x};T_i,H_i) \propto \exp\left\{-\frac{h(\boldsymbol{x}) \vee H_i}{T_i}\right\}, \tag{18}$$

for $i=1,\ldots,L$. In the above equation, $h(\boldsymbol{x}) \vee H_i = \max[h(\boldsymbol{x}),H_i]$, which flattens the part of the Hamiltonian $h(\boldsymbol{x})$ below H_i. Sampling from this truncated Hamiltonian enables the sampler to explore high-energy portion ($\ge H_i$) of the configuration space and to avoid being trapped to a deep local energy basin. The energy and temperature ladders are chosen such that one can obtain samples from most of the energy levels. In practice, we set H_1 and H_L close to the lower and the upper bounds of the energy range of interest, respectively. The intermediate truncation energies are then determined via a geometric progression. A global move, the equienergy jump, is designed to enhance the sampling of low-order distributions based on the empirical distributions generated by high-energy high-temperature chains. Samples from the EE sampler can be used to estimate the density of states $\Omega(u)$ via an iterative approach. For more details on the EE sampler, please see Ref. [27].

B. Constructing DGs

In what follows, we assume that we have simulated samples of configurations from a sequence of tempered-truncated distributions (18) and for each configuration we have recorded its energy. Denote the configurations simulated from the ith distribution in Eq. (18) by $\{X_t^{(i)}\}_{t=1}^n$ for $1 \le i \le L$. For every energy level u, if we partition the empirical sublevel set

$$\hat{A}(u) \triangleq \{X_t^{(i)}|h(X_t^{(i)}) < u, 1 \le i \le L, 1 \le t \le n\} \tag{19}$$

into connected components, we will then obtain an estimated DG of $h(\boldsymbol{x})$. Thus, after discretizing the energy space into small bins with a sequence of energy levels $u_1 < \cdots < u_M$ (here, M is often large), our goal is to identify the connected components of the empirical sublevel sets $\hat{A}^{(m)} \triangleq \hat{A}(u_m)$ for $m=1,\ldots,M$. However, there are two technical difficulties to consider. First, at a high-energy level, samples of configurations are usually sparse since the configuration space increases very fast with the increase of energy. Consequently, one cannot determine connectivity simply based on directed neighbors, but needs to rely on clustering algorithms and statistical criteria to identify clusters of configurations. Second, it will be more efficient to cluster empirical level sets

$$\hat{C}^{(m)} \triangleq \{X_t^{(i)}|h(X_t^{(i)}) \in [u_{m-1},u_m), 1 \le i \le L, 1 \le t \le n\} \tag{20}$$

for $m=1,\ldots,M$, where $u_0 \equiv -\infty$. The reasons are that configurations in a level set are approximately uniformly distributed and that the size of a level set is much smaller than that of a sublevel set, which reduces the computational burden in clustering.

These considerations motivate our development of the bottom-up partition algorithm for constructing a DG from Monte Carlo samples. Given a metric of the configuration space we employ single-linkage clustering (SLC) [42] to partition an empirical level set (20) into clusters. Starting from singleton configurations, SLC recursively merges two closest subsets of configurations according to nearest-neighbor distances (NND). Clusters may then be identified with a chosen distance threshold. Define the maximum NND of a cluster by the nearest-neighbor distance between the two subsets of the

cluster that are merged at the last step during SLC, i.e., the maximal merging distance within the cluster. The BUP algorithm is outlined as follows.

(1) Initialization: Perform SLC on $\hat{C}^{(1)}$ to obtain clusters $\{\hat{C}_k^{(1)}\}_1^{K_1}$ and their respective maximum NNDs $\{d_k^{(1)}\}_1^{K_1}$. Let $\{\hat{A}_k^{(1)}\}_1^{K_1}=\{\hat{C}_k^{(1)}\}_1^{K_1}$ be the clusters of $\hat{A}^{(1)}$.

(2) Induction: For $m=2,\ldots,M$,

(a) perform SLC on $\hat{C}^{(m)}$ to obtain clusters $\{\hat{C}_i^{(m)}\}_1^{K_m^*}$ and their maximum NNDs $\{r_i^{(m)}\}$;

(b) connect $\hat{C}_i^{(m)}$ and $\hat{A}_j^{(m-1)}$ if the NND between them is $\leq \max(r_i^{(m)}, d_j^{(m-1)})$ for $i=1,\ldots,K_m^*$ and $j=1,\ldots,K_{m-1}$; and

(c) merge the resulting connected clusters to obtain $\{\hat{A}_k^{(m)}\}_1^{K_m}$, the clusters of $\hat{A}^{(m)}$, and update their maximum NNDs $d_k^{(m)}=\max\{r_i^{(m)}, d_j^{(m-1)} | \hat{C}_i^{(m)}, \hat{A}_j^{(m-1)} \subset \hat{A}_k^{(m)}\}$ for $k=1,\ldots,K_m$.

Clearly, the disconnectivity among $\{\hat{A}_k^{(m)} | k=1,\ldots,K_m, m=1,\ldots,M\}$ is equivalent to the estimated structure of the tree. Asymptotic properties and statistical applications of the BUP algorithm are discussed in [43]. The exact method for identifying clusters from SLC of spin-glass configurations will be discussed in Sec. V.

C. Estimating local density of states

As we have mentioned, the density of states $\Omega(u)$ can be estimated by the EE sampler. Denote by $\hat{\Omega}_m$ the estimate for the mth discretized energy bin $(1\leq m\leq M)$. Since the range of an energy bin $\Delta u_m=u_m-u_{m-1}$ is designed to be small, the empirical level set $\hat{C}^{(m)}$ [Eq. (20)] can be regarded as a sampled microcanonical ensemble $\mathcal{X}(u_m)$ (5). Given an estimated tree from the BUP algorithm, we can estimate the local density of states for the mth energy bin from the empirical level set $\hat{C}^{(m)}$ and its connectivity on the estimated tree. The BUP algorithm partitions $\hat{C}^{(m)}$ into K_m subsets, each of which belongs to a cluster of the empirical sublevel set $\hat{A}^{(m)}$. Intuitively, the K_m subsets of $\hat{C}^{(m)}$, i.e., $\hat{C}^{(m)}\cap\hat{A}_k^{(m)}$ $(k=1,\ldots,K_m)$, correspond to the intersections between the estimated tree and the horizontal line at energy u_m. They represent our sampled version of the CMEs $\mathcal{X}^{(i)}(u_m)$ for $i\in\mathcal{I}(u_m)$ (6). Define

$$\hat{\Omega}_{m,k}=\hat{\Omega}_m n_k^{(m)}/n^{(m)}, \quad \text{for } k=1,\ldots,K_m, \tag{21}$$

where $n^{(m)}$ and $n_k^{(m)}$ are the numbers of configurations in $\hat{C}^{(m)}$ and $\hat{C}^{(m)}\cap\hat{A}_k^{(m)}$, respectively. Because configurations in $\hat{C}^{(m)}$ are roughly uniform in distribution, $\hat{\Omega}_{m,k}(1\leq k\leq K_m)$ gives a reasonable estimate of the LDOS $\Omega^{(i)}(u_m)$, $i\in\mathcal{I}(u_m)$, after reindexing $\hat{\Omega}_{m,k}$ by the edges on the estimated tree. For notational consistency, we denote the estimated local density of states by $\{\hat{\Omega}_m^{(i)}, m=1,\ldots,M, i\in\hat{\mathcal{I}}\}$, where $\hat{\mathcal{I}}$ is the collection of the edge indices of the estimated tree.

Our definition and estimation of LDOS have connections to the superposition approximation approach, reviewed in [44,45], which dates back to early work on homogeneous nucleation (e.g., [46–48]). However, the purpose and the computation employed are different. In a superposition approach, the configuration space is partitioned into attraction domains of local minima. Local density of states is approximated for each local minimum, and thermodynamic quantities, such as the partition function, are then calculated by the summation of contributions from different local minima. In this work, we start with an estimation of density of states from the EE sampler and decompose the density of states for every energy level into the LDOS of connected microcanonical ensembles. The goal is to calculate with LDOS statistical and thermodynamic properties of various local minima and local basins for any temperature. Such calculations will be exact if the size of Monte Carlo samples approaches infinity.

IV. ILLUSTRATIVE EXAMPLE

In this section we illustrate the ideas and intuitions behind LDOS, tree entropy, and local escape probability with an example for which theoretical calculations are available. The Hamiltonian of this example is given by

$$h(\boldsymbol{x})=\min_{1\leq i\leq K}\left(w_i+\frac{\|\boldsymbol{x}-\boldsymbol{\mu}_i\|^2}{2\sigma_i^2}\right), \tag{22}$$

where $\boldsymbol{x},\boldsymbol{\mu}\in\mathbf{R}^3$ and $K=4$. This function contains K local minima centered at $\boldsymbol{\mu}_i$ with energy w_i. The Hamiltonian restricted to the connected basin of $\boldsymbol{\mu}_i$ is a quadratic form and the parameter σ_i^2 controls its local variance.

For all the following realizations of the Hamiltonian (22), we applied the same parameter setup in the EE sampler with ten chains, each of 500 K samples. The energy ladder was set between h_0 and h_0+50, where h_0 is the global minimum of $h(\boldsymbol{x})$, and the temperature ladder between 0.5 and 20, both with geometric progression. We partitioned the energy space $[h_0,h_0+50]$ into $M=200$ small bins to estimate the density of states. Then we randomly sampled 20% of the configurations (to save computation) from each chain to form the input of the BUP algorithm to construct DGs and estimate LDOS. The above computation was repeated ten times independently for each realization of Eq. (22).

A. Local Boltzmann averages

In this section, we set $w_i=i-1$ for $i=1,\ldots,4$, $\sigma_1^2=\sigma_2^2=1$, $\sigma_3^2=\sigma_4^2=2$, $\boldsymbol{\mu}_1=(0,0,0)$, $\boldsymbol{\mu}_2=(10,0,0)$, $\boldsymbol{\mu}_3=(10,10,0)$, and $\boldsymbol{\mu}_4=(10,10,10)$. The theoretical tree of this Hamiltonian is shown in Fig. 3 and our estimated tree from the BUP algorithm gives an identical topology.

Index the connected basin (terminal edge) of the minimum $\boldsymbol{\mu}_i$ by i (Fig. 3) and denote its barrier (upper end-point energy) by B_i for $i=1,\ldots,4$. The volume of basin i for energy lower than a given level $u\in(w_i,B_i)$ can be calculated up to a multiplicative constant from the quadratic form in Eq. (22) via

$$V^{(i)}(u)\propto[\sigma_i^2(u-w_i)]^{p/2},$$

in which $p=3$ is the dimension of the space. Then the local density of states

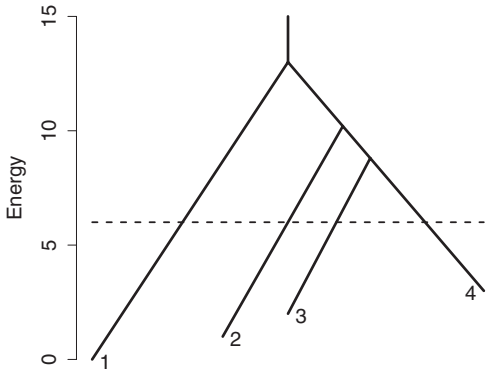

FIG. 3. The tree of the continuous Hamiltonian with the indices of terminal edges.

$$\Omega^{(i)}(u) = \frac{d}{du}V^{(i)}(u) \propto (\sigma_i^2)^{p/2}(u - w_i)^{p/2-1}, \tag{23}$$

based on which one can compute local Boltzmann averages over the basin. In particular, the sublevel set $A(H)=\{\boldsymbol{x}|h(\boldsymbol{x})<H\}$, with $H=6$, has four connected components, denoted by $A_i(H)$, each containing a $\boldsymbol{\mu}_i$, for $i=1,\ldots,4$. They can be understood as four separate energy wells (the part of the tree below the dashed line in Fig. 3). Suppose that we are interested in characterizing these energy wells. We may compute various local Boltzmann averages, such as the local average energies and the relative probabilities of these energy wells as functions of the temperature T as follows:

$$\begin{aligned} U_i(T) &\triangleq \frac{\int_{A_i(H)} up_B(\boldsymbol{x};T)d\boldsymbol{x}}{\int_{A_i(H)} p_B(\boldsymbol{x};T)d\boldsymbol{x}} \\ &= \frac{\int_{w_i}^{H} u\Omega^{(i)}(u)e^{-u/T}du}{\int_{w_i}^{H} \Omega^{(i)}(u)e^{-u/T}du} \\ &= w_i + T\frac{\Gamma([H-w_i]/T, p/2+1)}{\Gamma([H-w_i]/T, p/2)}, \end{aligned}$$

$$\begin{aligned} P_i(T) &\triangleq \frac{\int_{A_i(H)} p_B(\boldsymbol{x};T)d\boldsymbol{x}}{\int_{A(H)} p_B(\boldsymbol{x};T)d\boldsymbol{x}} \\ &= \frac{\int_{w_i}^{H} \Omega^{(i)}(u)e^{-u/T}du}{\int_{0}^{H} \Omega(u)e^{-u/T}du} \\ &= \frac{(\sigma_i^2)^{p/2}e^{-w_i/T}T^{p/2}\Gamma([H-w_i]/T, p/2)}{\sum_{k=1}^{4}(\sigma_k^2)^{p/2}e^{-w_k/T}T^{p/2}\Gamma([H-w_k]/T, p/2)}, \end{aligned}$$

where $\Gamma(y,a)=\int_0^y t^{a-1}e^{-t}dt$ is the incomplete gamma function and $p_B(\boldsymbol{x};T)\propto\exp[-h(\boldsymbol{x})/T]$ is the Boltzmann distribution. The closed-form solutions are obtained by substituting the expression of $\Omega^{(i)}(u)$ [Eq. (23)] into the above integrals. Such quantities are very useful for understanding the neighborhood of a local minimum, but need estimation in general.

We can estimate these local Boltzmann averages as functions of temperature T directly with estimated LDOS,

$$\hat{U}_i(T) = \frac{\sum_{u<H} u\hat{\Omega}_u^{(i)}e^{-u/T}\Delta u}{\sum_{u<H} \hat{\Omega}_u^{(i)}e^{-u/T}\Delta u},$$

$$\hat{P}_i(T) = \frac{\sum_{u<H} \hat{\Omega}_u^{(i)}e^{-u/T}\Delta u}{\sum_{k}\sum_{u<H} \hat{\Omega}_u^{(k)}e^{-u/T}\Delta u},$$

in which the LDOS $\hat{\Omega}_u^{(i)}$ is indexed by the discretized energy level u and Δu is the length of the corresponding energy bin. We took 25 values of T evenly spaced between 0 and 5 to compare our estimates against theoretical curves (Fig. 4). Our estimates are seen to be very accurate for all the temperatures. The maximal standard deviations across estimates from ten independent runs over all the temperatures were 0.02, 0.03, 0.006, and 0.01 for $U_i(T)$, and were 6×10^{-3}, 5×10^{-3}, 5×10^{-3}, and 3×10^{-3} for $P_i(T)$, for $i=1,\ldots,4$, respectively. Please note that we obtained these local Boltzmann averages for all temperature levels with just one run of EE sampling and BUP tree construction and there is no need to sample from the Boltzmann distribution at all 25 different temperatures.

B. Tree entropy

To help understand the intuition behind tree entropy, we design the following parameters in the Hamiltonian (22): $w_i=0$, $\sigma_i^2=1$ for $i=1,\ldots,4$, and $\boldsymbol{\mu}_1=(x,0,0)$, $\boldsymbol{\mu}_2=(a,0,0)$, $\boldsymbol{\mu}_3=(a,a,0)$, and $\boldsymbol{\mu}_4=(a,a,a-x)$, where $a=10$ and $x\in[0,a]$. Under this setup, the theoretical value of the tree entropy as a function of x is given by $S_{tr}(x)=1+[(a-x)/a]^2$, which is a monotone decrease function for $x\in[0,a]$. If $x=0$, the tree of the DG is composed of four symmetric edges with a common root (left panel of Fig. 5). The entropy of the tree $S_{tr}(0)=2$ bits, which means that one needs two binary variables to code the four edges. If x is close to a, then $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are close to each other and so are $\boldsymbol{\mu}_3$ and $\boldsymbol{\mu}_4$. The tree is composed of two long internal edges each of which further splits into two short terminal edges (right panel of Fig. 5). In this case the tree entropy is expected to be slightly greater than one bit. For example, $S_{tr}(8)=1.04$. The tree entropy for a few typical values of x is tabulated in Table I.

Given estimated $\hat{\Omega}_u^{(i)}$ and $\hat{\Omega}_u$, one can approximate the tree entropy at energy u by substituting them into Eq. (12) and then calculate the average entropy by discretizing the integral in Eq. (13). We estimated the tree entropy of this example for $x=0,2,4,6,8$ as reported in Table I, which includes the average and standard deviation (in parentheses)

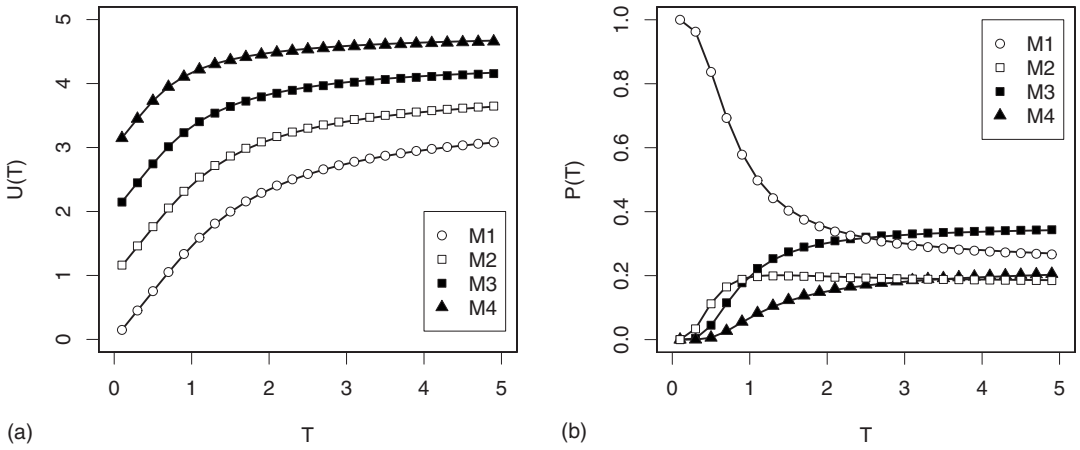

FIG. 4. Local Boltzmann averages of the energy wells of the four local minima, indexed by M1, M2, M3, and M4. (a) Estimated and theoretical curves of $U_i(T)$; (b) Estimated and theoretical curves of $P_i(T)$. Theoretical curves are plotted as solid lines. Symbols of different shapes report average estimates over ten runs.

across ten independent runs. Since the local density of states $\{\Omega^{(i)}(u)\}$ is identical across all the edges for every energy level u, the entropy of the tree at energy u [Eq. (12)] reaches its maximum. Consequently, it is not surprising that the tree entropy was slightly underestimated compared to the theoretical values.

C. Local escape probability

To elucidate the insight of local escape probability, we manipulate the distances between local minima and their local variances σ_i^2 in Eq. (22). If the minima are moved further away from each other or the local variances decrease, we expect the local escape probabilities to decrease due to the increase of energy barriers. We define a baseline parameter set as $w_i=0$ for all i, $\boldsymbol{\mu}_1=(0,0,0)$, $\boldsymbol{\mu}_2=(a,0,0)$, $\boldsymbol{\mu}_3=(a,a,0)$, and $\boldsymbol{\mu}_4=(a,a,a)$. In the first test, we set all $\sigma_i^2=1$ and choose $a=5,10,15$ to obtain three Hamiltonians with different distances between neighboring $\boldsymbol{\mu}_i$'s. In the second test, we fix $a=10$ and let the local variances $(\sigma_1^2,\ldots,\sigma_4^2)=(1,1,1,2)$, (1,1,2,3), or (1,2,3,4). This gives rise to three groups of parameters indexed as groups 1, 2, and 3 with increasing local variances.

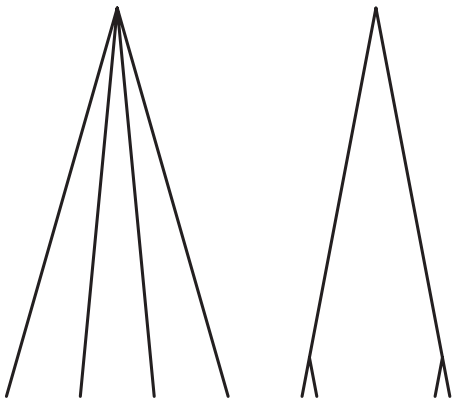

FIG. 5. Two trees with different entropy.

The estimation of the local escape probability of edge i at temperature T [Eq. (15)] is straightforward with $\hat{\Omega}_u^{(i)}$,

$$\hat{P}_{LE}^{(i)}(T)=\frac{e^{-\hat{B}_i/T}\sum_{u\in[\hat{L}_i,\hat{B}_i]}\hat{\Omega}_u^{(i)}\Delta u}{\sum_{u\in[\hat{L}_i,\hat{B}_i]}\hat{\Omega}_u^{(i)}e^{-u/T}\Delta u},$$

where $\hat{L}_i$ and $\hat{B}_i$ are the estimated lower and upper end-point energies of the edge on the tree. For each parameter setting in the two tests, we estimated the minimum local escape probability $P_{LE}(T)$ for $T\in(0,20)$ (Fig. 6). Two expected patterns are observed: (1) The minimum local escape probability increases with the temperature. (2) For a fixed temperature, $P_{LE}(T)$ is higher when the local minima are closer to each other [with smaller a in Fig. 6(a)] or when the local variances are larger [Fig. 6(b)].

A side-by-side comparison between the estimated escape temperature $\hat{T}_E(0.5)$ [Eq. (17)] and the length of the longest edge on the theoretical tree, $\max_{i\in\mathcal{I}}(B_i-L_i)$, provides some interesting quantitative understanding of the local escape probability (Table II). For all sets of parameters in this example, $\hat{T}_E(0.5)$ is roughly half of the length of the longest edge. Intuitively, the longest edge (in terms of B_i-L_i) on a tree is most likely the bottleneck for the mixing of Metropolis-based sampling and optimization (e.g., simulated annealing [49]). If the current energy is at the midpoint of the

TABLE I. Entropy of the tree.

x	Theory	Estimation
0	2	1.943 (0.010)
2	1.64	1.604 (0.017)
4	1.36	1.330 (0.008)
6	1.16	1.132 (0.004)
8	1.04	1.016 (0.001)

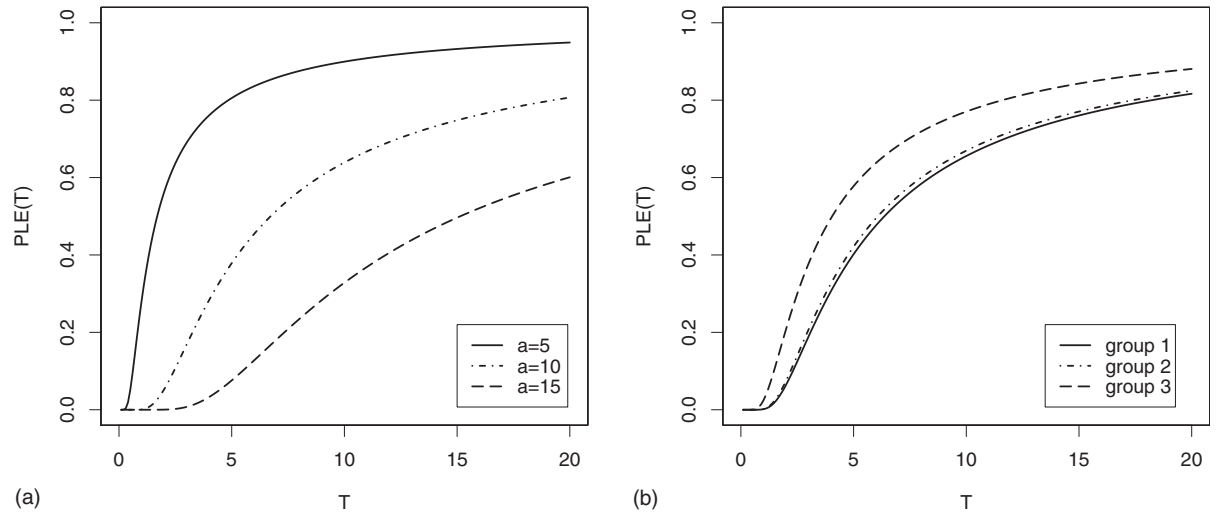

FIG. 6. The estimated $P_{LE}(T)$ as a function of temperature T for (a) test 1 and (b) test 2. The curves report average estimates from ten independent runs.

edge, the chance to accept a symmetric proposal that results in a local escape, i.e., raising energy higher than the barrier, is $e^{-1} \approx 0.4$ when the temperature T equals half of the edge length. This crude calculation matches the definition of $T_E(0.5)$, the escape temperature at level $\alpha=0.5$. This result also suggests that a Metropolis sampler for a Boltzmann distribution at $T \approx T_E(0.5)$ is expected to move freely between different local basins of the Hamiltonian $h(\boldsymbol{x})$. Thus, an estimated escape temperature can be used to set the highest temperature for sampling and optimization.

V. SK SPIN-GLASS MODEL

A. EE sampling

We study the SK spin-glass model for $N=20$, 40, and 100 spins and choose $J_0=0$, $J^2=1/N$ to generate the interactions $\{J_{ij}\}$ in Eq. (1). Under these parameters, there exists a phase transition from spin glasses to paramagnetic phase at $T_c=1$ as $N \to \infty$ [29]. We generated 20 independent realizations of the interaction parameters $\{J_{ij}\}$ $(1 \le i,j \le N)$ for each value of N. The spin-glass model with a particular realization of interactions is called a realization or a realized system. All following computations were performed on 20 independent realizations for each N.

TABLE II. The maximum edge length and the escape temperature. Note that the edge length is calculated from the theoretical tree and $\hat{T}_E(0.5)$ is the average estimate over ten runs with standard deviation in parentheses.

Test 1			Test 2		
a	$\max_i(B_i-L_i)$	$\hat{T}_E(0.5)$	Group	$\max_i(B_i-L_i)$	$\hat{T}_E(0.5)$
5	3.125	1.66(0.052)	1	12.5	6.33(0.157)
10	12.5	6.67(0.095)	2	12.5	6.04(0.084)
15	28.125	15.09(0.166)	3	8.58	4.01(0.129)

To study the energy landscape, we simulated spin-glass configurations from the Boltzmann distribution (2) via EE sampling. We compared a few key thermodynamic quantities estimated from our Monte Carlo sampling to their theoretical values to verify the unbiasedness of our samples. The range of the rescaled energy $h(\boldsymbol{S})/N$ that is of interest for thermodynamic simulation is between -0.8 and 0.2. Accordingly, we applied the EE sampler to this problem with energy truncation between $[-0.8N \text{ and } 0.2N]$ and temperature between $[0.4 \text{ and } 0.1N]$, both with geometric progression. The number of configurations in each chain and the total number of chains were set in proportion to the size of a system as $100K \times 10$, $200K \times 20$, and $1000K \times 50$ for $N=20$, 40, and 100, respectively. We then divided the energy space into $M=40$, 80, and 500 bins to estimate the density of states for $N=20$, 40, and 100, respectively. Given estimated $\hat{\Omega}_u$, we calculated the internal energy via

$$\hat{U}(T) = \frac{\sum_u u \hat{\Omega}_u e^{-u/T} \Delta u}{\sum_u \hat{\Omega}_u e^{-u/T} \Delta u} \tag{24}$$

and the specific heat via

$$\hat{C}(T) = \frac{\widehat{U^2}(T) - [\hat{U}(T)]^2}{T^2} \tag{25}$$

for $T \in (0,4)$, where $\widehat{U^2}(T)$ was calculated with $\hat{\Omega}_u$ similarly to Eq. (24). Please note that the energy u is rescaled in the unit of N for both calculations. The estimated internal energy and specific heat are compared to the theoretical curves as $N \to \infty$ [29] in Fig. 7. Although the number of spins in our simulation is not too large, the general tendency here is very clear: the estimated Boltzmann averages approach consistently the theoretical limits with the increase of N. From the specific heat, we approximated the critical temperature by $\hat{T}_c = \arg\max_T \hat{C}(T)$ and obtained $\hat{T}_c = 0.780 \pm 0.131$,

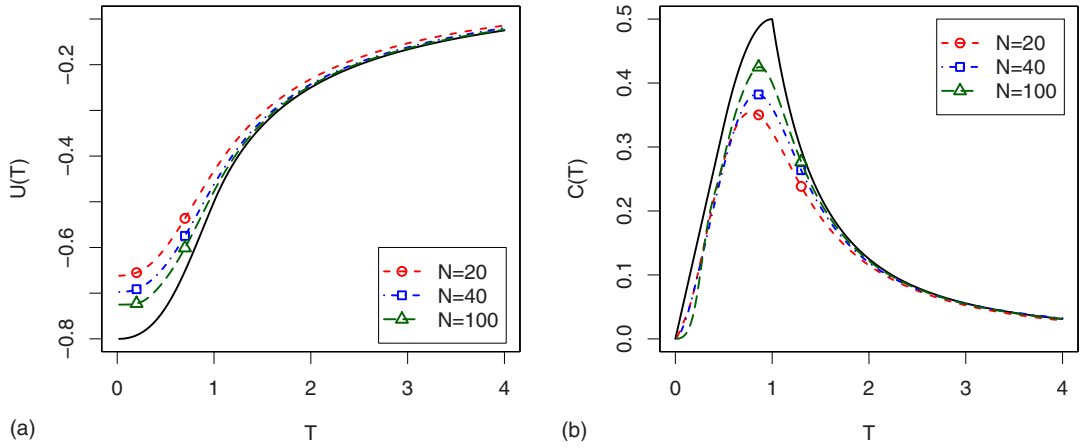

FIG. 7. (Color online) (a) Internal energy and (b) specific heat as a function of T. Black solid lines report the theoretical curves as $N\to\infty$. Other curves are estimated by the averages over 20 realizations. Symbols of different shapes are used only to distinguish curves.

0.840 ± 0.118, and 0.860 ± 0.140 for $N=20$, 40, and 100 respectively, where $\pm$ indicates one standard error. As expected, the difference between the estimated critical temperature and the theoretical $T_c=1$ decreases with the increase of N. These results demonstrate that the configurations simulated by the EE sampler formed a good representation for the microcanonical ensembles and the density of states was accurately estimated. They provided the raw materials for understanding the energy landscape of the spin-glass model.

B. Connectivity and clustering

We define connectedness for the configuration space of the spin-glass model via the Hamming distance, i.e., the number of different spins between two configurations. Two configurations are called connected neighbors if they only differ by one spin. Such definition of connectedness has been used in many previous studies (e.g., [11,12]). Furthermore, if one flips the orientations of all the N spins, the system shows exactly the same statistical properties determined by the Hamiltonian (1). Thus, in this work we identify two configurations with completely reverse spins. This can be achieved by modifying the distance between two configurations $\boldsymbol{S}_a$ and $\boldsymbol{S}_b$ to

$$d(\boldsymbol{S}_a,\boldsymbol{S}_b)=\min\{d_H(\boldsymbol{S}_a,\boldsymbol{S}_b),d_H(\boldsymbol{S}_a,-\boldsymbol{S}_b)\}, \tag{26}$$

in which d_H denotes the Hamming distance.

Suppose that we have simulated $n+1$ configurations, $\boldsymbol{S}_1,\ldots,\boldsymbol{S}_{n+1}$, from a microcanonical ensemble $\mathcal{X}(u)$. Applying single-linkage clustering on the $n+1$ configurations generates n NNDs, denoted by $d_1,d_2,\ldots,d_n$, which join closest subsets sequentially. Motivated by the observation that the histogram of d_i $(i=1,\ldots,n)$ decays exponentially if the microcanonical ensemble is connected, we model them by a geometric distribution,

$$P(d_i|\beta)=\beta^{d_i}(1-\beta),\quad d_i=0,1,\ldots, \tag{27}$$

where $\beta\in(0,1)$ is an unknown parameter. We rank d_i to obtain the order statistics $d_{(1)}\le d_{(2)}\le\cdots\le d_{(n)}$. If $\mathcal{X}(u)$ consists of $K+1$ connected components, one expects the largest K NNDs to be significantly greater than the remaining ones when n is large. Consequently, our goal is translated to detect overly large distances (between-cluster NNDs) that are very unlikely to observe from the geometric distribution (27) determined by the other distances (within-cluster NNDs). To avoid the bias introduced by always removing the largest observations, we can estimate β in Eq. (27) assuming that the largest k $(k\ge0)$ distances are missing (but still exist). This gives the following maximum-likelihood estimator:

$$\hat{\beta}_k=\frac{\sum_{i=1}^{n-k}d_{(i)}+kd_{(n-k)}}{\sum_{i=1}^{n-k}d_{(i)}+kd_{(n-k)}+(n-k)},$$

and the mean of the geometric distribution can be estimated by

$$\hat{\theta}_k=\frac{\hat{\beta}_k}{1-\hat{\beta}_k}=\frac{\sum_{i=1}^{n-k}d_{(i)}+kd_{(n-k)}}{n-k}.$$

Note that if one simply removes the largest k distances, the mean will be $\Sigma_{i=1}^{n-k}d_{(i)}/(n-k)$, which is always underestimated compared to $\hat{\theta}_k$. From the memoryless property of the geometric distribution, the expected value of any of the largest k distances given $\hat{\theta}_k$ is $d_{(n-k)}+\hat{\theta}_k$. Let $K_{\max}$ be a predetermined maximal number of components. For $k=1,\ldots,K_{\max}-1$, we compute $\hat{\theta}_k$ and define $\gamma_k=(d_{(n-k+1)}-d_{(n-k)})/\hat{\theta}_k$ as a statistic to test whether the observed largest k NNDs are significantly greater than that expected from the other $n-k$ distances. Find $K_H=1+\max\{k|\gamma_k>\alpha\}$, with $\alpha=10$, as an upper bound for the number of clusters. Starting from K_H potential clusters determined by SLC, we prune out those potential clusters which contain less than 50 configurations to generate the final clusters. These final clusters not only have significantly large distances between each other but also contain enough samples of configurations. This increases the robustness of the cluster detection procedure in the BUP algorithm. It is helpful to clarify that configurations in a cluster are deter-

TABLE III. Estimation of DGs of the spin-glass model. Note that estimates are followed by ± one standard deviation computed from 20 realizations. Energies of minima and barriers are reported in the unit of N.

Statistics	$N=20$	$N=40$	$N=100$
Number of local minimum	10.2 ± 4.6	20.1 ± 10.1	99.5 ± 59.5
Lowest minimum	-0.663 ± 0.045	-0.700 ± 0.040	-0.731 ± 0.021
Median minimum	-0.574 ± 0.046	-0.667 ± 0.045	-0.712 ± 0.022
Highest minimum	-0.514 ± 0.051	-0.640 ± 0.048	-0.704 ± 0.020
Number of barriers	6.8 ± 3.2	10.2 ± 3.6	52.8 ± 26.7
Lowest barrier	-0.600 ± 0.056	-0.676 ± 0.041	-0.724 ± 0.022
Median barrier	-0.542 ± 0.040	-0.654 ± 0.042	-0.711 ± 0.021
Highest barrier	-0.487 ± 0.046	-0.626 ± 0.045	-0.702 ± 0.020
S_{tr} (bits)	0.461 ± 0.194	0.762 ± 0.482	0.707 ± 0.437
max $S_{tr}(u)$ (bits)	1.41 ± 0.472	1.99 ± 0.782	1.96 ± 0.744
$T_E(0.5)$	0.905 ± 0.335	0.660 ± 0.157	0.550 ± 0.110

mined by the above statistical procedure and thus may not be connected neighbors. This is especially so at a high-energy level, for which one can at most obtain sparse samples.

C. Energy landscapes and characterization

We resampled 20% of the configurations from all Monte Carlo chains in the above EE sampling to estimate the tree for each realized system. We note that the highest barriers (the roots) of all the estimated trees for $N=100$ turned out to be less than six units from the global minima $\hat{u}_0$. Since the configuration space is huge for 100 spin systems, we focused on a more detailed estimation of the low-energy part between $\hat{u}_0$ and $\hat{u}_0+6$. For each realized system of $N=100$, we reapplied the EE sampler with energy truncation between $[\hat{u}_0$ and $\hat{u}_0+6]$ and temperature between $[0.4,1]$ so as to generate denser configurations in the lower portion of the energy space. Similarly, 20% of these configurations were utilized to construct an estimated DG. Summary statistics of the estimated trees are reported in Table III. For a spin system, a local minimum is defined as a configuration for which energy increases if any of the spins is flipped. We verified all the identified local minima by this definition and found 98% of them were true ones. For large N values, our approach only reconstructed a small portion of the full energy landscape, mostly in low-energy ranges as indicated by the summary statistics of critical energies. However, we want to emphasize that the low-energy range is the most interesting and important part for understanding and characterizing spin-glass behaviors of the model. In particular, we have estimated DGs for energy between $[-0.700N$ and $-0.626N]$ for $N=40$ and $[-0.731N$ and $-0.702N]$ for $N=100$. We note that enumerating configurations for these energy ranges is practically impossible for $N \geq 40$ and thus, simulating configurations from microcanonical ensembles seems a reasonable alternative.

As an illustration, we plot four examples of the estimated trees for various system sizes in Fig. 8. A few general characteristics are observed: (1) the structure of the tree becomes much more complicated for larger N; (2) the topologies of the trees are highly asymmetric for all system sizes; and (3) the depths of the energy barriers seem to decrease for larger systems. In order to quantify some of these observations, we use the defined statistics based on LDOS to characterize the energy landscape.

From the estimated local density of states, we calculated the average tree entropy over 20 realizations for each system size. It turns out that the average entropy is less than 1.0 bit for $N=20$, 40, and 100 (Table III). In other words, when averaged over different energy levels, the uncertainty in determining to which connected microcanonical ensemble a configuration belongs is lower than the uncertainty in the outcome of tossing a fair coin, even for a tree with about 100 leaves (such as those with $N=100$). Furthermore, the maximum tree entropy over all the energy levels we sampled, $\max_u S_{tr}(u)$ [Eq. (12)], is less than two bits for all the three sizes of spin glasses (Table III). These results imply that the microcanonical ensemble for any given energy level is mostly dominated by a small number of components with a much larger volume than the remaining ones. Although there exist a large number of local minima and local basins, many of them only occupy a very small portion of the configuration space.

We computed the minimum local escape probability $P_{LE}(T)$ [Eq. (16)] from the estimated trees and local density of states for temperature $T \in (0,4)$ and plotted the average curves over 20 realizations in Fig. 9(a). As expected, when $T \to 0$ all escape probabilities are close to zero since the system will be trapped in any local minimum with probability one. This is consistent with the spin-glass behaviors of the model at low temperatures. If the temperature increases, the Boltzmann distribution becomes less rough in the sense that the probability of a local escape increases. An interesting phenomenon is that at a given temperature the P_{LE} actually increases with the system size N. Consequently, the escape temperature at level 0.5, $T_E(0.5)$, shows a decreasing trend with the increase of N (Table III). This implies that in spite of the existence of many more edges on the tree of a large-size system, the energy barriers are actually lower compared to those of a small-size system. In particular, the minimum

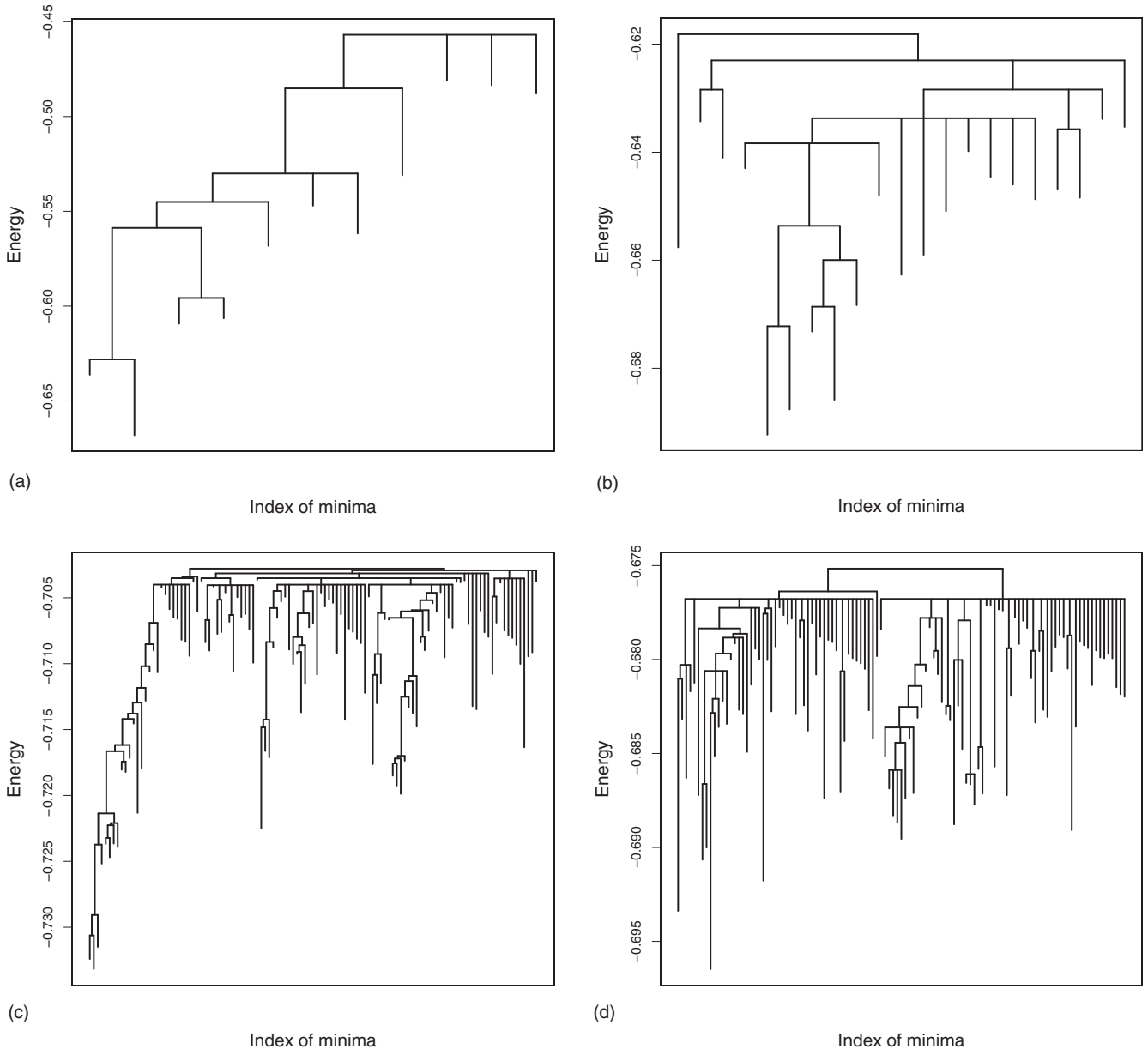

FIG. 8. Examples of estimated trees of the spin-glass model. The energy is reported in the unit of N. The system sizes are 20, 40, 100, and 100 for (a), (b), (c), and (d), respectively.

local escape probability is around 0.7 at the critical temperature $T_c=1$ for $N=100$. This suggests that the spin system is unlikely to be trapped in a local basin at the critical temperature, which is consistent with the phase transition from spin glass to paramagnetic state.

There seems to be a transition point $T^* \in (0,0.5)$ for local escape probabilities. For $T<T^*$, $P_{LE}(T)$ is close to zero and the roughness of the Boltzmann distribution is high. Once beyond the transition point, the system shows a dramatic increase in local escape probabilities, especially for large N [Fig. 9(a)]. To quantify this phenomenon, we investigate the derivative of $P_{LE}(T)$. From Eq. (15), the derivative of the local escape probability on edge i with respect to T is

$$\frac{d}{dT}P_{LE}^{(i)}(T) = V^{(i)}\frac{d}{dT}\left\{\frac{e^{-B_i/T}}{Z^{(i)}(T)}\right\} = \frac{P_{LE}^{(i)}\{B_i - U^{(i)}(T)\}^{\Delta}}{T^2} = DP_{LE}^{(i)}(T), \tag{28}$$

where $U^{(i)}(T)$ is the local energy of the edge (9). Let $m_T = \arg\min_i P_{LE}^{(i)}(T)$ be the index of the edge that gives the minimum local escape probability at temperature T. We define the derivative of $P_{LE}(T)$ as

$$DP_{LE}(T) = DP_{LE}^{(m_T)}(T). \tag{29}$$

Then the maximum of $DP_{LE}(T)$ indicates the transition point for the minimum local escape probability. As shown in Fig. 9(b), the peak of the estimated $DP_{LE}(T)$ appears when tem-

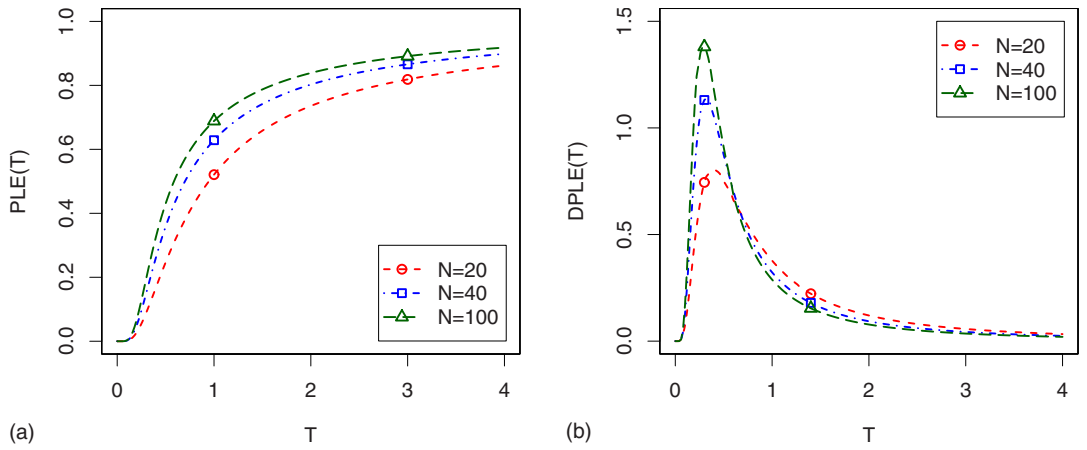

FIG. 9. (Color online) (a) Average curves of $P_{LE}(T)$ and (b) its derivative for the spin-glass model. Symbols of different shapes are used only to distinguish curves.

perature approaches $T^*=0.3$ and the transition is sharper for larger N with a higher maximum derivative. The relation between this transition temperature and the critical temperature T_c for phase transition is interesting for future studies. The observation that $T^*<T_c$ may be due to the assumption that a proposed energy increment is uniform in the derivation of Eq. (14), which may result in a higher probability for large energy increase compared to thermal perturbation.

VI. DISCUSSION

From this study, we gained some understanding of the energy landscape and its relationship to the thermodynamics of the SK spin-glass model. First, the tree of the energy landscape remains very asymmetric with small tree entropy (<1 bit) even for $N=100$. Second, although there exists a large number of local minima, the basins of many local minima only occupy a small fraction of the configuration space and their barriers are usually low. This is one of the reasons for the observed low escape temperature $T_E(0.5)$. Third, if we heat up the system from $T=0$ to a high temperature, there exists a steep increase in local escape probabilities before the phase transition at the critical temperature. This observation links the features of energy landscapes to the thermodynamics of the spin system. Some of the characteristics we discovered here, such as asymmetry and low barriers of the DGs, are consistent with previous studies on smaller systems of $N\leq 25$ (e.g., [17,20]). Finally, obvious differences are seen from a crude visual comparison between the DGs of the SK spin-glass model and those of a structural glass former [50,51]. But a concrete characterization of structural glass energy landscapes with tree entropy and local escape probability awaits further detailed analysis.

The concepts of local density of states, tree entropy, and local escape probability are very useful for characterizing the energy landscapes of complicated distributions in different scientific fields. For the past two decades, many powerful Monte Carlo methods have been developed and numerous Monte Carlo samples have been generated to solve various complicated problems. The concepts and methods proposed in this paper can be applied to a wide range of problems in physics, statistics, and computational biology, with the availability of large-size Monte Carlo samples and the increasing computing capacity. An interesting future work is to link the disconnectivity graph and local density of states to Monte Carlo sampling. For example, the local escape probability is potentially connected to the convergence and efficiency of a Metropolis sampler. Given a roughly estimated tree, one may design a more efficient sampling algorithm to re-explore the configuration space. Ideally, iteration among sampling, tree construction, and local density of states estimation has a great potential to improve the efficiency of Monte Carlo sampling and energy landscape exploration. However, the computationally intensive nature of these methods will present great challenges. Another direction for future work is to simplify energy landscapes by regrouping local minima based on thresholding barrier heights [52], which may be connected to construction of transition networks for modeling molecular dynamics [53].

ACKNOWLEDGMENT

This work was supported by NSF Grants No. DMS-0805491 and DMS-0505732.

[1] S. F. Edwards and P. W. Anderson, J. Phys. F: Met. Phys. **5**, 965 (1975).
[2] D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792 (1975).
[3] G. Parisi, Phys. Rev. Lett. **43**, 1754 (1979).
[4] F. Tanaka and S. F. Edwards, J. Phys. F: Met. Phys. **10**, 2769 (1980).
[5] A. J. Bray and M. A. Moore, J. Phys. C **13**, L469 (1980).

[6] B. Waclaw and Z. Burda, Phys. Rev. E **77**, 041114 (2008).
[7] K. Nemoto, J. Phys. A **21**, L287 (1988).
[8] G. Hed, A. K. Hartmann, D. Stauffer, and E. Domany, Phys. Rev. Lett. **86**, 3148 (2001).
[9] S. Y. Kim, S. J. Lee, and J. Lee, Phys. Rev. B **76**, 184412 (2007).
[10] T. Klotz and S. Kobe, J. Phys. A **27**, L95 (1994).
[11] Z. Burda, A. Krzywicki, O. C. Martin, and Z. Tabor, Phys. Rev. E **73**, 036110 (2006).
[12] J. Krawczyk and S. Kobe, Physica A **315**, 302 (2002).
[13] T. X. Hoang, N. Sushko, M. S. Li, and M. Cieplak, J. Phys. A **33**, 3977 (2000).
[14] O. M. Becker and M. Karplus, J. Chem. Phys. **106**, 1495 (1997).
[15] C. L. Brooks III, J. N. Onuchic, and D. J. Wales, Science **293**, 612 (2001).
[16] D. J. Wales, Philos. Trans. R. Soc. London, Ser. A **363**, 357 (2005).
[17] P. Garstecki, T. X. Hoang, and M. Cieplak, Phys. Rev. E **60**, 3219 (1999).
[18] F. Ferreira, J. F. Fontanari, and P. F. Stadler, J. Phys. A **33**, 8635 (2000).
[19] J. F. Fontanari and P. F. Stadler, J. Phys. A **35**, 1509 (2002).
[20] W. Hordijk, J. F. Fontanari, and P. F. Stadler, J. Phys. A **36**, 3671 (2003).
[21] H. Seyed-allaei, H. Seyed-allaei, and M. R. Ejtehadi, Phys. Rev. E **77**, 031105 (2008).
[22] T. Aspelmeier, R. A. Blythe, A. J. Bray, and M. A. Moore, Phys. Rev. B **74**, 184411 (2006).
[23] S. V. Krivov and M. Karplus, J. Chem. Phys. **117**, 10894 (2002).
[24] D. A. Evans and D. J. Wales, J. Chem. Phys. **118**, 3891 (2003).
[25] P. N. Mortenson and D. J. Wales, J. Chem. Phys. **114**, 6443 (2001).
[26] P. N. Mortenson, D. A. Evans, and D. J. Wales, J. Chem. Phys. **117**, 1363 (2002).
[27] S. C. Kou, Q. Zhou, and W. H. Wong, Ann. Stat. **34**, 1581 (2006).
[28] D. J. Wales, M. A. Miller, and T. R. Walsh, Nature (London) **394**, 758 (1998).
[29] S. Kirkpatrick and D. Sherrington, Phys. Rev. B **17**, 4384 (1978).
[30] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[31] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).
[32] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).
[33] B. A. Berg and T. Celik, Phys. Rev. Lett. **69**, 2292 (1992).
[34] W. Janke and S. Kappler, Phys. Rev. Lett. **74**, 212 (1995).
[35] B. Hesselbo and R. B. Stinchcombe, Phys. Rev. Lett. **74**, 2151 (1995).
[36] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[37] F. Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001).
[38] F. Liang, J. Stat. Phys. **122**, 511 (2006).
[39] F. Liang, C. Liu, and J. Carroll, J. Am. Stat. Assoc. **102**, 305 (2007).
[40] C. J. Geyer, in *Computing Science and Statistics*, Proceedings of the 23rd Symposium on Interface, edited by E. M. Keramidas (Interface Foundation, Fairfax Station, VA, 1991), p. 156.
[41] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).
[42] J. A. Hartigan, *Clustering Algorithms*, (Wiley New York, 1975).
[43] Q. Zhou and W. H. Wong, Ann. Appl. Stat. **2**, 1307 (2008).
[44] D. J. Wales and T. V. Bogdan, J. Phys. Chem. B **110**, 20765 (2006).
[45] D. J. Wales, *Energy Landscapes* (Cambridge University Press, Cambridge, England, 2003).
[46] J. J. Burton, J. Chem. Phys. **52**, 345 (1970).
[47] D. J. McGinty, J. Chem. Phys. **55**, 580 (1971).
[48] W. G. Hoover, A. C. Hindmarsh, and B. L. Holian, J. Chem. Phys. **57**, 1980 (1972).
[49] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Science **220**, 671 (1983).
[50] F. Calvo, T. V. Bogdan, V. K. de Souza, and D. J. Wales, J. Chem. Phys. **127**, 044508 (2007).
[51] V. K. de Souza and D. J. Wales, J. Chem. Phys. **129**, 164507 (2008).
[52] B. Strodel and D. J. Wales, Chem. Phys. Lett. **466**, 105 (2008).
[53] F. Noé and S. Fischer, Curr. Opin. Struct. Biol. **18**, 154 (2008).