Multicanonical Monte Carlo ensemble growth algorithm

Graziano Vernizzi,[1,*] Trung Dac Nguyen,[2] Henri Orland,[3,4] and Monica Olvera de la Cruz[5]
[1] *Department of Physics and Astronomy, Siena College, Loudonville, New York 12211, USA*
[2] *Department of Chemical and Biological Engineering, Northwestern University, Evanston, Illinois 60208, USA*
[3] *Institut de Physique Théorique, CEA Saclay, 91191 Gif-sur-Yvette Cedex, France*
[4] *Beijing Computational Science Research Center, No. 10 East Xibeiwang Road, Beijing 100193, China*
[5] *Department of Materials Science and Engineering, Northwestern University, Evanston, Illinois 60208, USA*

(Received 4 December 2019; accepted 3 February 2020; published 26 February 2020)

We present an ensemble Monte Carlo growth method to sample the equilibrium thermodynamic properties of random chains. The method is based on the multicanonical technique of computing the density of states in the energy space. Such a quantity is temperature independent, and therefore microcanonical and canonical thermodynamic quantities, including the free energy, entropy, and thermal averages, can be obtained by reweighting with a Boltzmann factor. The algorithm we present combines two approaches: The first is the Monte Carlo ensemble growth method, where a "population" of samples in the state space is considered, as opposed to traditional sampling by long random walks, or iterative single-chain growth. The second is the flat-histogram Monte Carlo, similar to the popular Wang-Landau sampling, or to multicanonical chain-growth sampling. We discuss the performance and relative simplicity of the proposed algorithm, and we apply it to known test cases.

DOI: 10.1103/PhysRevE.101.021301

Over the past two decades, chain-growth algorithms proved to be among the most powerful methods for sampling the equilibrium configuration of polymer systems, in different environments and with various interactions [1–4]. References [5–9] report several cases where such methods are more efficient than classic move-set-based random-walk Monte Carlo (MC) sampling, both in the configuration space and in the energy space. As the name indicates, a chain-growth algorithm grows the polymer chain one monomer at a time, while avoiding occupied locations (in the case of self-avoiding random walks), and by correcting the corresponding sampling bias with suitable weight factors. Nowadays, several efficient chain-growth algorithms are available, and many of them show impressive performances, such as pruned-enriched Rosenbluth method (PERM) [2], or flatPERM [3], including a large family of variations and refinements (for reviews, see, e.g., Refs. [10,11]).

An important advancement in the field was the realization that the difficulties encountered by MC growth in the canonical ensemble at fixed temperature T (poor sampling at low temperatures, critical slowdown, or energy barrier trapping, to name a few), can be avoided by microcanonical MC chain-growth methods in the *energy space* [1,3]. There, the density of states $g(E)$ is sampled by performing a series of single-polymer chain growths using a PERM algorithm (or some variations thereof) until all energy states have been visited an approximatively equal number of times. Such methods not only capture the benefits of the so-called flat-histogram techniques, but they are also very robust and have been shown to produce reliable results. From $g(E)$ one can obtain several thermodynamic quantities, such as the partition function, the free energy, the entropy, the specific heat, and other thermal averages [12].

A different implementation of the MC chain-growth method in the canonical ensemble was proposed in Refs. [4,13]. It was later called *breadth-first* implementation in Refs. [2,14,15] (the list is not exhaustive) to distinguish it from the more popular *depth-first* implementations, where a single polymer chain (or a family of chains, recursively) is grown all the way to the last monomer, and the procedure is repeated a number of times until sufficient statistics is accumulated. In contrast, in breadth-first implementations the chains are grown one monomer at a time in parallel, with a probability proportional to the Boltzmann distribution, and in such a way that the ensemble remains at thermodynamic equilibrium *at every step*. Depth-first methods require prior knowledge on the "attrition" weights, and historically have been shown to be the fastest, due also to the limited memory requirements. Breadth-first methods require less external parameters, and no prior knowledge on the weights, but suffer from large-memory requirement (to store all chain configurations) and from slower performance on single-core computational platforms. In this Rapid Communication, we extend the breadth-first algorithm from Refs. [4,13] to the microcanonical ensemble (multicanonical), and at the same time we merge it with the Wang-Landau sampling [9]. The latter computes the microcanonical density of states $g(E)$ via a random sampling that produces a flat histogram in the energy space. In the next section we describe our algorithm in detail and we test its performance on some classic systems.

The algorithm. For the sake of conciseness, we do not review here the original MC ensemble-growth method in the canonical ensemble [4], nor the multicanonical chain-growth method [1], for which we refer to the original papers. To fix some notation, we define a *configuration* of a polymer chain

[*] Corresponding author: gvernizzi@siena.edu

of length N to be the sequence of $N+1$ monomers (or beads): $X^{(N)} = \{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$. In case of a chain on a lattice, or of freely jointed rods, one can additionally request that each link has constant length $|\mathbf{x}_{i+1} - \mathbf{x}_i| = a$, for $i = 0, \ldots, N-1$. Let $\mathscr{C}_N = \{X_\alpha^{(N)}\}$ be the *ensemble* of all configurations $X_\alpha^{(N)}$ of a polymer chain of length N (i.e., the population). Moreover, let $E_\alpha^{(N)} = \mathscr{H}(X_\alpha^{(N)})$ be the energy of each configuration, with $\mathscr{H}$ being the Hamiltonian. The canonical partition function for such a system is

$$\mathscr{Z}_N \equiv \int_{\mathscr{C}_N} \mathscr{D}X \, e^{-\beta \mathscr{H}(X)}, \tag{1}$$

where $\beta = \frac{1}{k_B T}$, k_B is the Boltzmann constant, and T the absolute temperature. Although we use a continuous notation, the equations in this Rapid Communication are equally valid for discrete systems (such as a lattice polymer), by simply substituting the integration symbols with discrete sums, $\int \mathscr{D}X \to \sum_X$. By introducing the density of states $g_N(E)$ in the energy space,

$$g_N(E) \equiv \int_{\mathscr{C}_N} \mathscr{D}X \, \delta(E - \mathscr{H}(X)), \tag{2}$$

where $\delta(x)$ is the Dirac delta function, the partition function $\mathscr{Z}_N$ can be expressed as an integral over all energies,

$$\mathscr{Z}_N = \int dE \, g_N(E) e^{-\beta E}. \tag{3}$$

Since $g_N(E)$ is temperature independent, its single evaluation allows the computation of $\mathscr{Z}_N$ at *any* temperature by direct Boltzmann reweighting. From Eq. (2) it follows that $\int dE \, g_N(E) = \int_{\mathscr{C}_N} \mathscr{D}X = \mathrm{Vol}(\mathscr{C}_N)$, which is the population size, i.e., the total number of states. Moreover, $\int_{\mathscr{C}_N} \mathscr{D}X \frac{1}{g_N[\mathscr{H}(X)]} = \sigma_N$, where σ_N is the support of $g_N(E)$.

The algorithm we propose grows a sample of the ensemble $\mathscr{C}_n$ given a sample of the ensemble $\mathscr{C}_{n-1}$, both ensembles being at thermodynamic equilibrium. By iterating over $n = 1, \ldots, N$, one obtains a statistical ensemble of configurations for a polymer chain with $N+1$ monomers. Namely, given a sample $\mathscr{S}_{n-1} = \{X_\alpha^{(n-1)}\} \subset \mathscr{C}_{n-1}$, $\alpha = 1, \ldots, M_{n-1}$, where M_{n-1} is the sample size, one generates new "daughter" configurations $X_\beta^{(n)}$ by adding an nth monomer at the end of each "parent" configuration $X_\alpha^{(n-1)}$. Next, each daughter configuration is replicated with a frequency that is proportional to the reciprocal of the density of states, in the same spirit of Wang-Landau sampling. Namely, if $\mu_n(X)$ is the number of times a configuration X with energy E is present in the sample $\mathscr{S}_n$, then the goal is to choose $\mu_n(X) \propto 1/g_n(E)$. That can be verified by computing the histogram $H_n(E)$ of the number of samples with energy E [9],

$$H_n(E) \equiv \int_{\mathscr{S}_n} \mathscr{D}X \, \mu_n(X) \delta(E - \mathscr{H}(X)). \tag{4}$$

When $\mu_n(X) \propto 1/g_n(E)$, Eq. (4) gives $H_n(E) \propto 1$, i.e., a flat histogram [1,9]. In this Rapid Communication we prove that one can produce a flat histogram *at each step* of the growth, if each daughter configuration is replicated precisely w_n times,

$$w_n = \frac{g_{n-1}(E_p)}{g_n(E)} \gamma_n, \tag{5}$$

where E is the energy of a daughter chain, E_p is the energy of her parent, and γ_n is a suitable "population control" factor [4], that we discuss below. By starting from a sample with M_0 single monomers at the origin, corresponding to a density of states $g_0(E) = \delta(E)$, the number of times a given configuration $X^{(N)}$ with energy E appears in the sample $\mathscr{S}_N$ is

$$\mu_N(X^{(N)}) = M_0 \prod_{i=1}^{N} w_i = M_0 \prod_{i=1}^{N} \frac{g_{i-1}}{g_i} \gamma_i = M_0 g_0 \frac{\prod_{i=1}^{N} \gamma_i}{g_N(E)}, \tag{6}$$

which is proportional to $1/g_n(E)$ indeed. By inserting Eq. (6) in Eq. (4) one has $H_n(E) = M_0 \prod_{i=1}^{n} \gamma_i$, that is, a flat histogram. Moreover, for sufficiently large samples, the integral of $H_n(E)$ over E gives $M_n = M_0 \sigma_n \prod_{i=1}^{n} \gamma_i = M_{n-1} \sigma_n / \sigma_{n-1} \gamma_n$.

In general, w_n is a real number and therefore the replication, which can occur in discrete units only, must be achieved statistically: Each daughter is replicated an integer number of times $m = \mathrm{Int}(w_n) + [r < \mathrm{Frac}(w_n)]$, where r is a uniform random number in the interval [0,1), and $\mathrm{Int}(x)$ and $\mathrm{Frac}(x)$ are the integer part and the fractional part of x, respectively. The daughter configuration is going to be present in the sample an average number of times $\langle m \rangle = w_n$ (we indicate with $\langle x \rangle$ averages over different samples, while $\bar{x}$ indicates averages within a sample). The growth step can be repeated multiple times, depending on the model: For a polymer on a lattice with coordination number z, the addition of the new monomer is repeated z times, whether randomly or systematically (in our code, we attempt the growth in each one of the z directions). In the continuum case, such as for a freely jointed chain, the new monomer is placed at a random position on a sphere centered on the nth monomer, and again we indicate with z the number of times this operation is repeated. In case the added monomer violates steric constraints, one has $E_2 = \infty$, and can set directly $w_n = 0$ in Eq. (5) so that the chain is rejected effectively.

The expression for the weights in Eq. (5) contains two unknown quantities: the population control factor γ_n, and the "new" density of states $g_n(E)$. Similarly to Refs. [4,13], the factor γ_n can be chosen to maintain an "ideal" number of samples M_0, depending on the memory capacity and speed performance of the computational platform. In general, after the monomer growth one obtains a sample size $M_n \neq M_{n-1}$, with fluctuations around $\langle M_n \rangle = z M_{n-1} w_n$. In order to avoid an exponential explosion (or depletion) of the total number of sample configurations, a "population control procedure" is required [4,13]. In our case, the condition $M_n \simeq M_{n-1}$ (close to the set value M_0) is obtained by choosing $\gamma_n = \sigma_{n-1}/\sigma_n$.

To determine the unknown density of states $g_n(E)$, we suggest the following scheme. First, for $n > 1$ set $g_n^*(E) = 1$ (for all E) and $\gamma_n^* = 1$. Then, generate z daughter configurations for each parent ($z = 4$ for a square lattice, or $z = 6$ for a cubic lattice), and compute their energy $E_\beta^{(n)} = \mathscr{H}(X_\beta^{(n)})$. Next, we compute the weighted energy histogram $H_n^*(E)$ from the daughter configurations, where each new daughter configuration counts with a guestimated weight $w_n^* = g_{n-1}(E_p)$, where E_p is the energy of her parent [see Eq. (5)]. Since the goal is the determination of the *true* values for $g_n(E)$, and γ_n, we do not replicate the daughter configurations at this point and the weights do not need to be rounded to an integer number. $H_n^*(E)$ is generally not flat because $g_n^*(E) = 1$ and $\gamma_n^* = 1$ are not correct. However, we can use $H_n^*(E)$ to determine both $g_n(E)$ and γ_n by observing that the *true* replication weight should be

$$w_n = \frac{g_{n-1}(E_p)}{g_n(E)}\gamma_n = w_n^* \frac{\gamma_n}{g_n}. \qquad (7)$$

It means that if a daughter configuration X with energy E is present in the sample $\mu^*(X)$ times after being replicated with a factor w_n^*, then it should have been present $\mu(X) = \mu^*(X)\gamma_n/g_n(E)$ times instead. Hence Eq. (4) states that the *true* histogram is $H_n(E) = H_n^*(E)\gamma_n/g_n(E)$, which we know is flat: $H_n(E) = M_0 \prod_{i=1}^n \gamma_i$. For a sufficiently representative sample, the support σ_n of $H_n^*(E)$ is the support of $g_n(E)$, and $\gamma_n = \sigma_{n-1}/\sigma_n$. Consequently, the true density of states and the factor γ_n can be determined *in a single step*,

$$g_n(E) = \frac{H_n^*(E)}{M_0 \prod_{i=1}^{n-1} \gamma_i}, \qquad \gamma_n = \frac{\sigma_{n-1}}{\sigma_n}. \qquad (8)$$

We summarize the whole procedure:

(1) Start with a sample of M_0 copies of a single monomer with energy $E = 0$, and set the density of states to $g_0(E) = \delta(E)$. Then, for each $n = 1, \ldots, N$, repeat the following steps:

(2) Set $g_n(E) = 1$ for all E, and $\gamma_n = 1$. Grow all daughter configurations, and evaluate the energy E of each daughter.

(3) Compute the energy histogram $H_n(E)$, by counting each daughter with the weights in Eq. (5). Compute also the support σ_n of $H_n(E)$. (For discrete histogram distributions, σ_n is the number of nonzero energy bins times the size of a bin.)

(4) Set $g_n(E) = H_n(E)/(M_0 \prod_{i=1}^{n-1} \gamma_i)$ and $\gamma_n = \sigma_{n-1}/\sigma_n$, and apply the replication process with Eq. (5) explicitly. This time, the resulting energy histogram is guaranteed to be flat, and the sample size is close to M_0 (on average, over several samples). At this point all daughters become the new parents for the next growth step.

Since $g(E)$ becomes large for large values of N, it is convenient to work with $\ln[g(E)]$, i.e., the entropy. The rescaling at step (4) reads $\ln(g_n) \to \ln(H_n) - \ln(H_{n-1})$, and the weights in Eq. (5) are computed by $w_n = \gamma_n \exp[\ln(g_{n-1}) - \ln(g_n)]$.

The main advantage of this algorithm is that it produces the correct density of states $g_n(E)$ *at each* n, and when M_0 is sufficiently large it does not require additional iterations or corrections.

Examples. We validate our algorithm by simulating an interacting self-avoiding walk (ISAW) on a lattice, where the interaction energy is proportional to the number of nearest-neighbor contacts, and compare our results with known exact enumeration values, available for $N = 21$ on a square lattice [16] and $N = 28$ on a cubic lattice [17]. All curves have been obtained after averaging over 10 runs, with a sample size of $M_0 = 10^5$ each. As explained above, sample size fluctuations are expected: Ref. [13] shows that the best estimate of a quantity O is the weighted average $O = \sum_S M^{(S)}\bar{O}_S / \sum_S M^{(S)}$, where $\bar{O}_S$ is the mean value of O over all $M^{(S)}$ configurations in the sample S, and the sum is over all different samples. We also computed the specific heat C_V, directly from $g(E)$ by using a shifted reweighting technique [18,19] [see Supplemental Material (SM) [20] for details]. The statistical averages for $g(E)$ and C_V are very stable, with small error bars (see Fig. 1).

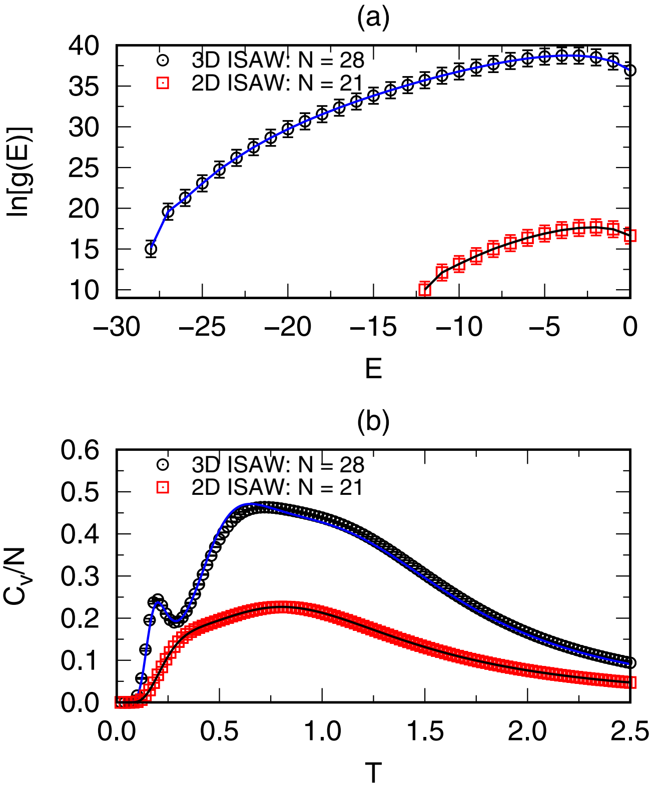

FIG. 1. Density of states (top) and specific heat (bottom) for ISAW chains with $N = 21$ on a square lattice, and $N = 28$ on a cubic lattice. Markers represent our simulations, while dots (connected by solid lines, for eye guidance) are from exact enumeration. The error bars (reported) are smaller than marker sizes. The agreement between the simulation and the exact values is remarkable.

Moreover, the statistical uncertainty due to correlated chains in an ensemble growth algorithm is inversely proportional to the sample size M_0 and the number of samples, or equivalently, the number of independent runs [13]. We monitored the flatness of the histogram $H_n(E)$ by using the Kullback-Leibler divergence test [21], and we found that the scaling in Eq. (8) is sufficiently robust to produce flat histograms in a single step when $M_0 \gg N$. The comparison shows that our algorithm is precise and accurate, with a relative error on C_v of the order 10^{-3}–10^{-2} for temperatures that are not too low. In Fig. 2 we show the results for ISAW on a square lattice with lengths $N = 40, 41, 42$, which, to our knowledge, are the longest ISAW that have been numerated exactly to date [16,17]. Also in this case the agreement with the exact enumeration results is remarkable, although a larger sample size was necessary ($M_0 = 3 \times 10^5$). A common critique that has been advanced in the past against ensemble growth Monte Carlo methods is that the memory requirements may be prohibitive in practical applications. Nevertheless, our algorithm has the advantage to be easily parallelizable for calculations across multiple nodes. In particular, the parallelization scheme we implemented is not limited by the per-node memory resources. For the $N = 28$ ISAW on a cubic lattice the whole calculation with $M_0 = 10^5$ replicas per Message Passage Interface (MPI) rank takes about 10 s to complete 10 independent runs with 8 MPI ranks on an Intel(R) Xeon(R) E5-2680 2.4-GHz CPU node. We have

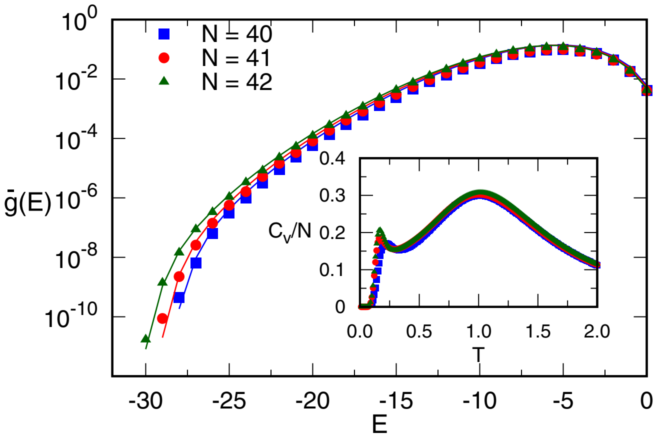

FIG. 2. Normalized density of states (and specific heat, inset) of an ISAW chain on a square lattice, for lengths $N = 40, 41, 42$. The number of replicas is $M_0 = 3 \times 10^5$. The solid lines connect the exact enumeration results from Refs. [16,17].

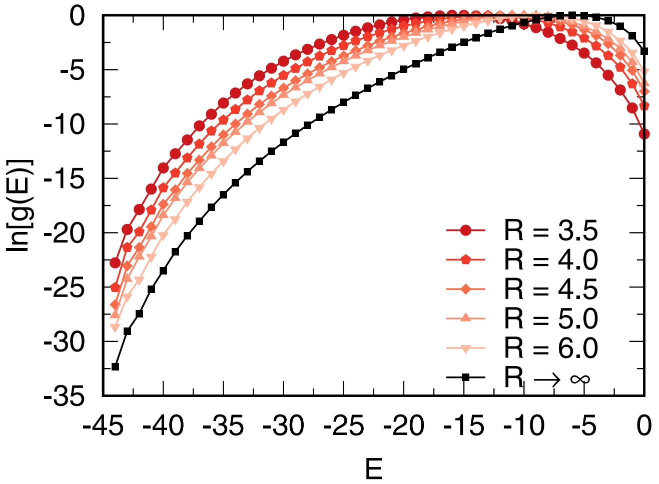

FIG. 3. Density of states of a $N = 40$ ISAW confined in a spherical three-dimensional cavity, as a function of the cavity radius R. The curves are shifted to zero at their maximum value.

also found agreement between the specific heat curves of the ISAW on a cubic lattice for $N = 35, 36, 37$ (see Fig. S1 in the SM [20]) with those reported in Ref. [22]. For lengths up to $N = 200$, the results are robust across different runs using $M_0 = 8 \times 10^8$ replicas (see Fig. S2 in the SM [20]). We find also remarkable agreement on the location of the peaks of $C_v(T)$ and the ground-state degeneracy for the hydrophobic-polar (HP) models in Refs. [1,12,23] (see Fig. S3 in SM [20]). Finally, we show in Fig. 3 that our algorithm can be readily extended to systems under confinement. A spherical cavity is modeled as a spherical hard wall which effectively allows only daughter chains that remain inside the cavity.

We conclude this Rapid Communication with a few comments. Similarly to the canonical counterpart [13], the ensemble growth method in the microcanonical ensemble is nondynamical, and therefore is not affected by slowing-down effects at phase transitions that are typical of dynamical Monte Carlo methods (such as for subefficient reptation algorithms for collapsed polymer chains). It profits also from sampling in the energy space, hence it is relatively insensitive to energy barriers, and it does sample both low-energy and high-energy configurations. Moreover, it shares similar advantages of analogous microcanonical sampling schemes [1,9]: From a *single* simulation, one can determine the density of states $g_n(E)$ for all chain lengths $n = 1, \dots, N$, which can be used to compute any thermal average at any temperature. Finally, we emphasize the generality of this method, which can be adapted to generate statistical ensembles with a density of states that is filtrated by parameters other than the energy.

Acknowledgments. G.V., T.D.N., and M.O.d.l.C. are thankful for the support by the National Science Foundation through Grant No. DMR-1611076, and by the Sherman Fairchild Foundation.

[1] M. Bachmann and W. Janke, Phys. Rev. Lett. **91**, 208105 (2003).
[2] P. Grassberger, Phys. Rev. E **56**, 3682 (1997).
[3] T. Prellberg and J. Krawczyk, Phys. Rev. Lett. **92**, 120602 (2004).
[4] T. Garel and H. Orland, J. Phys. A: Math. Gen. **23**, L621 (1990).
[5] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[6] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
[7] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).
[8] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).
[9] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[10] W. Janke and W. Paul, Soft Matter **12**, 642 (2016).
[11] H. P. Hsu and P. Grassberger, J. Stat. Phys. **144**, 597 (2011).
[12] M. Bachmann and W. Janke, J. Chem. Phys. **120**, 6779 (2004).
[13] P. G. Higgs and H. Orland, J. Chem. Phys. **95**, 4506 (1991).
[14] P. Grassberger, Comput. Phys. Commun. **147**, 64 (2002)
[15] D. Aldous and U. Vazirani, in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science* (IEEE, Piscataway, NJ, 1994), pp. 492–501.
[16] J. Lee, Comput. Phys. Commun. **228**, 11 (2018).
[17] Y. H. Hsieh, C. N. Chen, and C. K. Hu, Comput. Phys. Commun. **209**, 27 (2016).
[18] D. P. Landau, T. Shan-Ho, and M. Exler, Am. J. Phys. **72**, 1294 (2004).
[19] J. Kim, J. E. Straub, and T. Keyes, J. Chem. Phys. **126**, 135101 (2007).
[20] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.101.021301 for additional information on the ISAW models, the HP models, and the shifted reweigthing technique we used in the examples.
[21] S. Kullback and R. A. Leibler, Ann. Math. Stat. **22**, 79 (1951).
[22] T. Vogel, M. Bachmann, and W. Janke, Phys. Rev. E **76**, 061803 (2007).
[23] M. Bachmann and W. Janke, Acta Phys. Pol., B **34**, 4689 (2003).