Published in final edited form as:

J Comput Phys. 2013 October 1; 250: 1–12. doi:10.1016/j.jcp.2013.05.001.

Computing the partition function, ensemble averages, and density of states for lattice spin systems by sampling the mean

Dirk Gillespie^{1,*}

¹Department of Molecular Biophysics and Physiology, Rush University Medical Center, Chicago, IL

Abstract

An algorithm to approximately calculate the partition function (and subsequently ensemble averages) and density of states of lattice spin systems through non-Monte-Carlo random sampling is developed. This algorithm (called the sampling-the-mean algorithm) can be applied to models where the up or down spins at lattice nodes interact to change the spin states of other lattice nodes, especially non-Ising-like models with long-range interactions such as the biological model considered here. Because it is based on the Central Limit Theorem of probability, the sampling-the-mean algorithm also gives estimates of the error in the partition function, ensemble averages, and density of states. Easily implemented parallelization strategies and error minimizing sampling strategies are discussed. The sampling-the-mean method works especially well for relatively small systems, systems with a density of energy states that contains sharp spikes or oscillations, or systems with little *a priori* knowledge of the density of states.

INTRODUCTION

Lattice spin models are classic models of many physical phenomena, dating back at least to Ising in 1925 [1] and Bethe in 1931 [2]. Since then, these models have been applied to magnetism [1, 2], phase transitions [3, 4], quantum mechanics [5], and solid state physics [6, 7]. In this paper, a model from biology is introduced. In these models, N lattice nodes have spins (either up or down) that interact to change the spins of their neighbors.

Despite their age, lattice spin models continue to be useful tools in many fields, with thousands of recent papers using Ising models alone. Therefore, there continues to be interest in efficient numerical techniques to simulate traditional lattice spin models (e.g., [8, 9]). In addition, there is a growing interest in lattice spin models from biology. In particular, there are a number of vital phenomena that involve flipping the state of proteins based on their neighbors' states which can be modeled as lattice spin systems. These include calcium-induced calcium release that creates Ca^{2+} sparks and Ca^{2+} puffs [10], the propagation of calcium waves in arrhythmic heart muscle [11, 12], and action potential initiation and propagation both within a neuron [13] and across connexin-connected heart cells [14]. Many of these systems differ from traditional lattice spin system in that they involve a relative small number of lattice nodes (sometimes <100) and do not have a physically-defined Hamiltonian (although it is possible to define a pseudo-Hamiltonian). Moreover, the

© 2013 Elsevier Inc. All rights reserved.

*corresponding author: dirk_gillespie@rush.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

interactions are not only between nearest neighbors. In fact, because many biological systems involve diffusing ions as the interaction mechanism, the interactions can be very long-ranged. One of the things shown in this paper is that lattice spin systems with these properties can behave differently compared to traditional ones and that a new computational tool can be effective in studying them (and possibly some traditional lattice spin systems as well).

The most common numerical technique is Monte Carlo (MC) simulation where the spins of randomly-chosen nodes are flipped to determine thermodynamic properties [15]. In this paper, random sampling is used in a different way to approximately calculate the partition function of the system, a quantity Metropolis-Hastings MC algorithms [16–18] explicitly avoid. This is done by rearranging the partition sum into $N+1$ terms by considering only states with exactly n ($n=0,1, \dots, N$) up spins. Each term includes a mean (average) quantity that can be approximated using random sampling. It is therefore called the sampling-the-mean algorithm. After all the $N+1$ terms are computed separately, they are added together. With the same algorithm, it is also possible to approximately calculate ensemble averages and the density of energy states. Moreover, this highly parallelizable algorithm gives explicit estimates for the error of the approximation.

The sampling-the-mean algorithm is in many ways different from most MC methods. First, sampling-the-mean is based on the Central Limit Theorem of probability theory [19–21]; MC simulations are based on microscopic reversibility [18]. Secondly, there is no wasted computational effort; all the calculations are used, whereas in MC simulations many samples are rejected (with the Wang-Landau algorithm [22] being a notable exception). Lastly, all the random samples in the sampling-the-mean method are independent of each other; in MC simulations, the last accepted state is perturbed for the next attempted state, which sometimes can lead to the simulations being stuck in local energy minima.

However, the sampling-the-mean algorithm is also in many ways similar to some MC simulation techniques. For example, by holding the number of up spins constant for each set of random samples, the sampling-the-mean technique is similar to fixed-magnetization sampling [23] like that of Kawasaki [24] that interchanges the neighboring spins to keep the magnetization of the entire lattice constant [15].

The sampling-the-mean algorithm also shares some similarities with the Wang-Landau MC technique [22] in that it uses the density of states to compute thermodynamic properties. Specifically, Wang-Landau computes the number of states with energy in the interval $E + \Delta E_{WL}$, the density of energy states denoted $g(E)$. The partition function can then be computed by

$$Z = \sum_E g(E) e^{-E/kT} \quad (1)$$

where k is the Boltzmann constant and T the absolute temperature. Sampling-the-mean inverts this by picking sample states with a fixed number of up spins for which the density of up-spin states is known beforehand. Results are then combined in a sum similar to Eq. (1). Moreover, as will be shown, that the sampling-the-mean algorithm can also be used to compute the density of energy states.

In this paper, the sampling-the-mean algorithm is tested mainly on a new lattice spin model from biology. This model has several of the properties described above that make it different from many standard lattice spin models. First, it does not have a physically defined Hamiltonian so the transition from one state to another is not given by a Boltzmann factor.

Second, the interaction potential between nodes is not additive. Lastly, the interaction of the nodes is not only between nearest neighbors, but is long-ranged across the entire lattice. The model was chosen partly because of these properties to show that the sampling-the-mean algorithm does not rely on any special properties like nearest neighbor interactions. This model also highlights differences in the density of energy states between lattice spin models with long-range interactions and traditional nearest-neighbor models. Sampling-the-means appears to be more effective for the former, as compared to Wang-Landau.

In the next section, it is shown how the Central Limit Theorem can be used to (approximately) compute the partition function and ensemble averages and how the error in the calculation can be reduced most efficiently for a given number of random samples. Convergence and error trends are then analyzed for the biological test system to show that they are consistent with the predictions of the Central Limit Theorem. Later, sampling-the-mean is used to compute the density of energy states. When compared head-to-head with the Wang-Landau method, it is shown that the sampling-the-mean method is more efficient for relatively small systems (~ 400 nodes), especially for the test system because it produces a density of states $g(E)$ that is both discrete with large gaps between narrow spikes and highly oscillatory.

THEORY AND METHODS

General theory: Partition function

For a general lattice spin system, we first consider the state of the system α , defined by the vector $\alpha = (\alpha_1, \dots, \alpha_N)$ representing the spins of the N lattice nodes. The probability of the system being in state α is denoted P_α . Since the partition function Z is not known *a priori*, one can only calculate a quantity Q_α that is proportional to P_α . Then,

$$P_\alpha = Q_\alpha / Z \quad (2)$$

where the partition function is

$$Z = \sum_{\alpha} Q_{\alpha} \quad (3)$$

with the sum going over all possible states α of the system. For a system with a Hamiltonian $H(\alpha)$,

$$Q_{\alpha} = e^{-H(\alpha)/kT} \quad (4)$$

where k is the Boltzmann constant and T the absolute temperature.

To efficiently approximate Z through random sampling, Eq. (3) is rewritten by grouping all the 2^N possible states into those with 0, 1, 2, ..., N up spins. If, for a state α , there are N_α number of nodes with up spin, then

$$Z = \sum_{n=0}^N \sum_{\{\alpha | N_\alpha = n\}} Q_{\alpha} \quad (5)$$

Let $M(n)$ denote the number of states α so that $N_\alpha = n$. Then one can rewrite the inner sum in terms of the average Q_α for which $N_\alpha = n$, defined as

$$\bar{Q}(n) = \frac{1}{M(n)} \sum_{\{\alpha | N_\alpha = n\}} Q_\alpha. \quad (6)$$

Then

$$Z = \sum_{n=0}^N M(n) \bar{Q}(n). \quad (7)$$

In a lattice spin system, the number of nodes is constant and each node can only be in one of two states. Therefore, the number of possible states with exactly n of the N nodes having up spins is the number of ways to choose exactly n white balls from a bag that contains N number of black and white balls. This is the binomial coefficient N choose n so that

$$M(n) = \binom{N}{n} = \frac{N!}{n!(N-n)!} \quad (8)$$

and

$$Z = \sum_{n=0}^N \binom{N}{n} \bar{Q}(n). \quad (9)$$

The idea behind the proposed approximation of Z is that the average $\bar{Q}(n)$ can be sampled with randomly chosen states α with exactly n up spins. As discussed in further detail below, this formulation has a number of advantages from a computational point of view:

1. The calculations are relatively fast because $\bar{Q}(n)$ is a mean quantity.
2. It is possible to estimate how close the random sampling is to the correct result.
3. The sampling is easily highly parallelizable.

Estimating the $\bar{Q}(n)$ with random sampling is the sampling of a mean and through the Central Limit Theorem there is a theory for how such sampling behaves, at least asymptotically [19–21]. For now assume that n is fixed and that $\bar{Q}(n)$ is being sampled. Let $\bar{q}(\eta, n)$ denote *one* trial mean obtained by using η sample states; that is,

$$\bar{q}(\eta, n) = \frac{1}{\eta} \sum_{\alpha \in S(\eta, n)} Q_\alpha. \quad (10)$$

where $S(\eta, n)$ is a (uniformly) random subset of $\{\alpha | N_\alpha = n\}$ with η elements. (The standard nomenclature for $\bar{q}(\eta, n)$ is “sample mean” and not “trial mean”, but this would be cumbersome later on when we would discuss the relative merits of increasing the number of sample means versus increasing the number of samples used to compute the sample mean.) Now consider a set of ν such trial means $\{\bar{q}_1(\eta, n), \dots, \bar{q}_\nu(\eta, n)\}$ where each element $\bar{q}_k(\eta, n)$ uses a different random subset $S_k(\eta, n)$. Then, the Central Limit Theorem [19–21] states that, as $\eta \rightarrow \infty$, the set of trial means $\{\bar{q}_1(\eta, n), \dots, \bar{q}_\nu(\eta, n)\}$ is normally distributed (i.e., has a Gaussian distribution) about the mean of the full set ($\bar{Q}(n)$, in this case) with standard deviation

$$\tilde{\sigma}(n) = \sigma_n / \sqrt{\eta} \quad (11)$$

where σ_n is the standard deviation of the full set $\{Q_a\}_{\{a|N_a=n\}}$. It is important to note that the Gaussian distribution of the $\{\tilde{q}_1(\eta, n), \dots, \tilde{q}_\nu(\eta, n)\}$ is independent of the distribution of the numbers $\{Q_a\}_{\{a|N_a=n\}}$ whose mean we aim to approximate [19–21].

While this is asymptotically true when η is large and when the full set $\{Q_a\}$ has an infinite number of members, this result is useful in a number of ways. This is because $\tilde{\sigma}(n)$, the standard deviation of the trial means, depends only on η and not on ν . That is, it decreases as more random samples (η) from $\{Q_a\}_{\{a|N_a=n\}}$ are taken to make a single trial mean $\tilde{q}_k(\eta, n)$, but not as the number of such trial means (ν) increases. Since the number of sample Q_a to be computed for every n is $\eta\nu$, this means that for the same amount of computational work (i.e., for $\eta\nu$ constant), one can substantially decrease the standard deviation for the estimate of $\tilde{Q}(n)$ by using a large number η of random samples of $\{Q_a\}_{\{a|N_a=n\}}$ to compute each trial mean and a relatively few number ν of these trial means.

The Central Limit Theorem result then allows one to estimate how good the approximation of the partition function is. First, the mean of the trial means $\{\tilde{q}_1(\eta, n), \dots, \tilde{q}_\nu(\eta, n)\}$ is the approximation of $\tilde{Q}(n)$; that is,

$$\overline{Q}(n) \approx \tilde{Q}(n) \equiv \frac{1}{\nu} \sum_{k=1}^{\nu} \tilde{q}_k(\eta, n). \quad (12)$$

Second, the variance of $\{\tilde{q}_1(\eta, n), \dots, \tilde{q}_\nu(\eta, n)\}$ gives an estimate of the uncertainty in the estimate

$$\tilde{Z} = \sum_{n=0}^N \binom{N}{n} \tilde{Q}(n) \quad (13)$$

for Z . Since each term in the sum is independent of the other (each term is for a different n), this is given by the sum of the variances [19–21] so that

$$\varepsilon^2 = \sum_{n=0}^N \left[\binom{N}{n} \tilde{\sigma}(n) \right]^2. \quad (14)$$

Then,

$$Z \approx \tilde{Z} \pm \varepsilon. \quad (15)$$

Algorithm

This general theory can be implemented as an algorithm:

1. Pick a number W that is the maximum number of Q_a to be computed. This determines the maximum number of samples taken for each n .
2. For each n (perhaps computing each on an individual CPU):

- a. If $M(n)$ is small (e.g., when $M(n)$ is small for n close to 0 or N), then generate all possible states with n up spins and calculate $\bar{Q}(n)$ *exactly* through Eq. (6). Since the result is exact, $\tilde{\sigma}(n) = 0$.
- b. Otherwise, define ν in the range of 5 to 10 (see the example below for this specific choice) and let $\eta = W/\nu$. Generate η (uniformly) random sample states with exactly n up spins and compute the Q_α for each, making a set $\{Q_\alpha^{(\lambda)}\}_{\lambda=1,\dots,\eta}$. Compute $\bar{q}_1(\eta, n)$ sequentially as λ increases by $\bar{q}_1(\eta, n) \mapsto \bar{q}_1(\eta, n)[(\lambda-1)/\lambda] + Q_\alpha^{(\lambda)}/\lambda$. This prevents the addition of small numbers to large numbers to reduce round-off error. Repeat this sampling to generate the ν trial means $\{\bar{q}_1(\eta, n), \dots, \bar{q}_\nu(\eta, n)\}$. (This may be done on ν individual CPUs at the same time.) Compute the mean of the trial means ($\tilde{Q}(n)$) and their standard deviation ($\tilde{\sigma}(n)$).

3. Compute $\tilde{Z}$ from Eq. (13) and ε from Eq. (14).

General theory: Ensemble averages

The ensemble average of a property A that takes the value A_α depending on the state α is

$$\begin{aligned} \langle A \rangle &= \sum_{\alpha} A_{\alpha} P_{\alpha} \\ &= \frac{1}{Z} \sum_{n=0}^N \sum_{\{\alpha|N_{\alpha}=n\}} A_{\alpha} Q_{\alpha} \quad (16) \\ &= \frac{1}{Z} \sum_{n=0}^N \binom{N}{n} \overline{AQ}(n) \end{aligned}$$

where

$$\overline{AQ}(n) = \binom{N}{n}^{-1} \sum_{\{\alpha|N_{\alpha}=n\}} A_{\alpha} Q_{\alpha}. \quad (17)$$

The random sampling of the mean $\overline{AQ}(n)$ is then the same as for the partition function. In fact, both $\overline{AQ}(n)$ and $\bar{Q}(n)$ can be computed at the same time, from the same random samples, instead of two independent sampling-the-mean calculations. Moreover, the same kind of Central Limit Theorem analysis allows one to estimate error bars via Eq. (14) for the approximation of $\langle A \rangle$.

One simple example of this ensemble averaging is determining the average number of up-spin nodes. This is

$$\langle N_{\text{up}} \rangle = \sum_{\alpha} N_{\alpha} P_{\alpha} \quad (18)$$

so $A_{\alpha} = N_{\alpha}$. Then, by Eqs. (17) and (6),

$$\begin{aligned} \overline{AQ}(n) &= \binom{N}{n}^{-1} \sum_{\{\alpha|N_{\alpha}=n\}} N_{\alpha} Q_{\alpha} \\ &= n \bar{Q}(n) \end{aligned} \quad (19)$$

so that

$$\langle N_{\text{up}} \rangle = \frac{1}{Z} \sum_{n=0}^N n \binom{N}{n} \bar{Q}(n). \quad (20)$$

Therefore,

$$\langle N_{\text{up}} \rangle \approx \left(\sum_{n=0}^N n \binom{N}{n} \tilde{Q}(n) \pm \varepsilon_{\text{up}} \right) / (\tilde{Z} \pm \varepsilon) \quad (21)$$

where

$$\varepsilon_{\text{up}}^2 = \sum_{n=0}^N \left[n \binom{N}{n} \tilde{\sigma}(n) \right]^2. \quad (22)$$

Note that these are the $\tilde{Q}(n)$ that were already calculated for $\tilde{Z}$.

Test system

The example system is inspired by the biological phenomenon of calcium-induced calcium release in cardiac muscle [10]. In muscle, Ca^{2+} is stored in the sarcoplasmic reticulum organelle and is released through Ca^{2+} release channels into the cytoplasm where it binds to motor proteins to initiate muscle contraction. These channels are pores that conduct the Ca^{2+} across the sarcoplasmic reticulum membrane. There are ~10–100 channels in tight packed clusters (arrays) and they are usually closed, but are activated in the presence of micromolar Ca^{2+} . Therefore, when Ca^{2+} is released through one channel, that Ca^{2+} diffuses to neighboring channels in the array and activates them. The release channels then form a lattice spin system where each channel is at a node and the up spin corresponds to the channel being open and down spin to closed. An example of a 2×2 array of release channels is shown in Fig. 1.

Each channel has four activation sites that bind Ca^{2+} and four inactivation sites (Fig. 1) that also bind Ca^{2+} when the Ca^{2+} concentration is in the millimolar range (as opposed to the micromolar range for the activation sites). The probability of a channel being open (P_o) is given by [25]

$$P_o = P_o^{\text{max}} \left(1 + (K_d^{\text{act}} / [\text{Ca}]_{\text{act}})^H \right)^{-1} \left[1 - \left(1 + (K_d^{\text{inact}} / [\text{Ca}]_{\text{inact}})^H \right) \right]. \quad (23)$$

$P_o^{\text{max}} = 0.81$ sets the maximum open probability. The first term is for the activation sites with dissociation constant $K_d^{\text{act}} = 10^{-5}$ and the second for the inactivation sites with $K_d^{\text{inact}} = 6.3 \times 10^{-3} \text{ M}$. Each has a Hill coefficient of $H = 2.2$.

The Ca^{2+} concentrations $[\text{Ca}]_{\text{act}}$ and $[\text{Ca}]_{\text{inact}}$ are taken to be the average over the activation and inactivation sites, respectively. These concentrations contain both the constant background cytosolic Ca^{2+} and the concentration due to the Ca^{2+} going through the open channels and diffusing to neighboring channels. As function of distance r from the open channel center, this concentration is

$$\frac{i_{\text{Ca}}}{2\pi Dr} \quad (24)$$

where D is the diffusion coefficient of Ca^{2+} (taken to be $10^{-9} \text{ m}^2/\text{s}$) and i_{Ca} is the Ca^{2+} current through a single channel. For as-of-yet undetermined reasons, it is known that a channel cannot activate itself with the Ca^{2+} it releases [25] and so this empirical experimental finding is also included in the model.

The only so-far unspecified parameters are the cytosolic Ca^{2+} concentration and the Ca^{2+} current. For illustrative purposes in this paper, these are given nonphysiological values of $1 \mu\text{M}$ and 0.003 pA , respectively, so that the number of up-spin states contributing most to the partition function would also have large binomial coefficients. That is, these two parameters were chosen specifically so that those n with large binomial coefficients are the largest terms in the sum in Eq. (13). Truly physiological parameters are used in the Discussion.

To compute Q_a for an array of N channels in the spin state $a = (a_1, \dots, a_N)$, we have

$$Q_a = \prod_{k=1}^N p_k \quad (25)$$

where p_k is the probability that channel k is in its current state ($a_k = \uparrow$ or $\downarrow$). That is,

$$p_k = \begin{cases} P_o & \text{if channel } k \text{ is open} \\ 1 - P_o & \text{if channel } k \text{ is closed.} \end{cases} \quad (26)$$

This lattice spin system is then characterized by long-range interactions because the Ca^{2+} concentration due to diffusion is only inversely proportional to distance (Eq. (24)). This and the nonlinear definition for P_o means that the lattice nodes do not interact with an additive interaction potential. In fact, the interactions are highly nonlinear since p_k depends nonlinearly on the Ca^{2+} from all the open channels (Eq. (23)) through their reciprocal distance from each other (Eq. (24)). That is, for an array state a , p_k is a function of all the up spins in a that is significantly more complicated than for Ising models. There is also no physical energy or temperature associated with each spin state, only a Q_a .

RESULTS

To illustrate the effectiveness of the algorithm, consider the example system described in Theory and Methods. First, a 5×5 array is considered so that the sampling-the-mean algorithm can be compared to the exact answer which was determined by generating all possible $2^{25} \approx 3 \times 10^7$ spin states. This was the largest square system possible and took 2–4 hours depending on the computer clock speed. Later, a 10×10 array is considered to analyze the convergence of the sampling-the-mean algorithm for a system with so many states ($2^{100} \approx 10^{30}$) that there is no exact solution. All calculations were done with *Mathematica* 8.0.4 (Wolfram Research, Champaign, IL) using four parallel processors to divide the work of generating the ν trial means.

For the 5×5 array, the parameters of the system were intentionally chosen so that the states with $N_a = 12$ contribute most to the partition function; that is, the n with the largest binomial coefficient (5.3×10^6) is the largest term in the sum in Eq. (9). Therefore, the accuracy of the $n = 12$ term plays a large role in the accuracy of the final approximation of the partition function Z . The terms for all n are shown in Fig. 2.

Before assessing the accuracy of the algorithm, however, we first consider the validity of the Central Limit Theorem analysis, specifically whether Eq. (11) roughly holds. This is done by holding the total amount computational effort ($\eta\nu$) constant at 10^4 , 10^5 , and 10^6 and varying the ratio of η and ν . By Eq. (11), increasing η should decrease the standard deviation of the trial means while varying ν should do nothing. Fig. 3 shows that this is indeed the case. In each column, the larger η is, the smaller the circle is. In Fig. 3, the area of each circle is proportional to the standard deviation of the ν trial means. (The area rather than the radius was chosen because, with the radius, the smaller circles were not visible.) On the hand, if ν is increased while η is held constant (across columns at the same height), the size of the circles is approximately the same. This shows that it is the number of samples that make up each trial mean (η) rather than the number of trial means (ν) that determines the error bars for the estimate of Z . Therefore, it is best to use as large a number of random samples even if only 5–10 trial samples are actually averaged to compute $\tilde{Q}(n)$ in Eq. (12).

Next, we analyze the accuracy of the algorithm for the 5×5 test case. In Fig. 4, all the same calculations for $N_a = 12$ shown in Fig. 3 are compared to the exact answer. The results are very accurate for this test case. Most importantly, however, it is seen that ε gives an accurate estimate of the error in the approximation for this component of the partition function. This is further seen in Fig. 2 where all the terms in the sum in Eq. (13) are shown. There, the error bars are red rectangles. In Fig. 2, $\nu = 10$ and $\eta = 10^4$ and the standard deviations are small. Therefore, it seems that ε from Eq. (14) gives a reasonable bound for how close the approximation $\tilde{Z}$ is to the actual Z .

In Fig. 2, $\eta\nu = 10^5$ so that at least 2% of all the spin states were randomly sampled. However, for larger systems that one is actually interested in computing, usually 1% of all possible states is unattainably large. For example, for the 10×10 system that is considered next, there are $>10^{24}$ possible states with 72 up spins. With the same model parameters as for the 5×5 lattice, the largest contribution to the partition coefficient—as found by the sampling-the-mean algorithm—was with 72 up spins (shown in Fig. 5A). In Fig. 5A, a very small number of sample points was used— $\eta = 100$ —so that each point took only ~ 1.5 seconds on a 3.33 GHz Windows 7 machine using four cores to parallelize the $\nu = 10$ trial means.

Even with such a small number of samples, the error bars are reasonably small for many applications. However, one would like to know how many sample states are necessary for a good estimate of Z . Since there is no exact answer for such a large system, we use the convergence of the error bars as η increases from 10^3 to 10^7 (with $\nu = 10$) to estimate how close $\tilde{Q}(n)$ is to the actual $\tilde{Q}(n)$ (Fig. 5B). For this example system anyway, there is convergence and the error bars decrease in size by $\sim \sqrt{\eta}$ (specific numbers not shown), as predicted by Eq. (11). Moreover, even only $\eta = 1000$ samples gives an estimate within $\sim 2\%$ of the $\eta = 10^7$ result (Fig. 5B, right axis) in only ~ 14 seconds of calculation time. This is $\sim 1/3$ of the error shown in Fig. 5A where $\eta = 100$ was used.

DISCUSSION

The results on the test system show that for that system the sampling-the-mean algorithm is very efficient for biologically significant array sizes. To gauge how sampling-the-mean fares against other simulation methods, it was tested against the Wang-Landau algorithm [22] for both the test system and a 50×50 Ising model with periodic boundary conditions. The latter was chosen because an exact solution for the density of energy states exists [26] and because Wang and Landau considered this system [22].

To compare the methods, the sampling-the-mean technique was adapted to compute the density of energy states, as Wang-Landau does. This was done for several reasons. First, it allows a direct comparison to the Ising model's exact result. Second, direct computing of the partition function of the Ising system is numerically difficult because of the Boltzmann factor of the energy (Eq. (4)) whose smallest value is $-2N$. Moreover, the partition function is temperature dependent while the density of states is not.

To compute the density of states $g(E)$, first pick the energy interval length ΔE_{stm} and define

$$\delta_E(E_\alpha) = \begin{cases} 1 & \text{if } E \leq E_\alpha < E + \Delta E_{\text{stm}} \\ 0 & \text{otherwise} \end{cases} \quad (27)$$

Then,

$$\begin{aligned} g(E) &= \sum_{\alpha} \delta_E(E_\alpha) \\ &= \sum_n \sum_{\{\alpha | N_\alpha = n\}} \delta_E(E_\alpha) \\ &= \sum_n M(n) \bar{\delta}_E(n) \end{aligned} \quad (28)$$

where E_α is the energy of state α , $M(n)$ is the binomial factor of Eq. (8), and $\bar{\delta}_E(n)$ is computed by sampling-the-mean similar to Eq. (12). There is more to this sampling-the-mean than before, however, because, in the random sampling, the energy E is also important. Therefore, for each random state α with n up spins, E_α is computed and the corresponding energy level count for n up spins is incremented by 1. Algorithmically, for exactly n up spins and with ν and η given,

1. Pick η random states $\{\alpha_1, \dots, \alpha_\eta\}$ with n up spins and compute the energies $\{E_{\alpha_1}, \dots, E_{\alpha_\eta}\}$.
2. Calculate which energy levels the samples fall into. Suppose, for example, that these energies fall into 3 distinct energy level E_1, E_2 , and E_3 with η_1, η_2 , and η_3 of the η samples in the respective energy levels.
3. Make a vector d of the energies and the average occupancy. In this example,

$$d = \{(E_1, \eta_1/\eta), (E_2, \eta_2/\eta), (E_3, \eta_3/\eta)\}. \quad (29)$$

4. Repeat Steps 1–3 to create ν such vectors, $d_1, \dots, d_\nu$.
5. Combine these vectors according to energy levels and average over ν . In the example, if $\nu = 2$ and

$$\begin{aligned} d_1 &= \{(E_1, \eta_1/\eta), (E_2, \eta_2/\eta), (E_3, \eta_3/\eta)\} \\ d_2 &= \{(E_1, \eta'_1/\eta), (E_4, \eta'_4/\eta)\} \end{aligned} \quad (30)$$

then the combined and averaged vector is

$$\tilde{\delta}(n) = \{(E_1, (\eta_1 + \eta'_1)/\nu\eta), (E_2, \eta_2/\nu\eta), (E_3, \eta_3/\nu\eta), (E_4, \eta'_4/\nu\eta)\}. \quad (31)$$

At the same time, the standard deviation of the averaging can be computed to compute error bars later.

6. Each element of $\tilde{\delta}(n)$ is an approximation of $\tilde{\delta}_E(n)$. If E is not in $\tilde{\delta}(n)$, then $\tilde{\delta}_E(n) = 0$.

There are several things to note about this formulation. First, all the large numbers are in the binomial factor and $\tilde{\delta}_E(n)$ is only a running average of 0's and 1's. Second, no *a priori* knowledge of the energy spectrum is necessary; the range of energies is constructed as the sampling is done. Third, the energy interval length ΔE_{stm} can be chosen to be quite small (even 0) without affecting the calculation time; after the calculation is done, the intervals can be expanded into large ones whose new length (a multiple of ΔE_{stm}) is more appropriate by combining the energies of each $\tilde{\delta}(n)$.

These last two features are different from the Wang-Landau algorithm. Wang-Landau is a MC random walk where acceptance and rejection is based on the number of times an energy level has been previously visited [22]. Therefore, *a priori* knowledge of both the smallest and largest possible energies and an appropriate ΔE_{WL} is necessary. For traditional lattice spin models like Ising and Potts models, these are all well-known, but that is not necessarily true for other spin models like the biological system considered here.

To show how these differences affect performance, first consider the density of states for a 50×50 Ising model with periodic boundary conditions. This is shown in Fig. 6A. With 1.5×10^8 samples, there is divergence from the exact result. Moreover, with significantly fewer samples, Wang and Landau reported very small errors for this same system [22]. It appears that for this system Wang-Landau is more efficient than sampling-the-mean (although more work is needed to establish this definitively).

The story is reversed for the biological test system. For this system, a pseudo-energy for a state α can be defined as

$$E_\alpha = \ln Q_\alpha = - \sum_{k=1}^N \ln p_k. \quad (32)$$

Fig. 6B shows the *exact* density of pseudo-energy states $g(E)$ and all the terms $g(E)e^{-E}$ in the partition sum (Eq. (1)) for a 5×5 test system. This density of states is qualitatively different from the Ising system. Not only do the energy levels in the test system become more dense with E instead of being uniformly spaced, but $g(E)$ also has very narrow spikes and oscillations. Some of these are highlighted in Fig. 6C.

When sampling-the-mean was compared with Wang-Landau for various test system sizes, Wang-Landau generally required significantly more samples than sampling-the-mean for the same level of accuracy (data not shown). This is because of the very small ΔE_{WL} needed to resolve the spikes, and the smaller ΔE_{WL} , the more samples are needed in the Wang-Landau algorithm. Fig. 6C shows the approximate size of the energy intervals needed to have merely 2 digits of accuracy in the partition function (or $<10\%$ error since the partition function for the system in Fig. 6B is ~ 1.37). Therefore, without prior knowledge of required interval size, many trial simulations must be run to find the correct answer using Wang-Landau.

From these initial studies, it seems that the sampling-the-mean algorithm is relatively more efficient when the density of energy states has fine detail or where there is no prior knowledge of required energy interval size.

CONCLUSION

An non-Monte-Carlo random-sampling algorithm was presented to compute both the partition function and ensemble averages of N -node lattice spin models with discrete up and down spins. It was then extended to compute the density of states. This sampling-the-mean algorithm is based on an exact decomposition of the necessary sums (Eqs. (9) and (16)) into $N+1$ terms of the mean quantity $\bar{Q}(n)$ that may be estimated through random sampling. Because these $\bar{Q}(n)$ are averages (Eqs. (6) and (17)), a number (ν) of trial means made of many η random samples (Eq. (12)) approximate each term. Moreover, the Central Limit Theorem gives an estimate of the error ε in this approximation and how it decreases with the number of sample points η (Eq. (11)). This ε can not only be used to estimate the final error, but can also be computed “on the fly” to stop a calculation of $\bar{Q}(n)$. For example, for each term, instead of using a fixed η (as done in this paper) one can sample enough points until the mean of ν subsets of the samples gives a standard deviation below a specified threshold, at which point the calculation is stopped. By either method, this algorithm is highly—and trivially—parallelizable because, at the very least, each of the $N+1$ terms can be computed on separate CPUs, as can the ν trial means. However, other parallelizations over more CPUs and with better load balancing are certainly possible.

The sampling-the-mean algorithm was applied to a non-Hamiltonian model system from biology with long-range interactions for which the method is particularly efficient. The particular features of this model that seem to make sampling-the-mean relatively faster are that the density of energy states has wide gaps between sharp spikes that need to be sampled accurately because they make significant contributions to the partition function. For such systems—especially, sampling-the-mean can be a useful tool for lattice spin systems especially nontraditional ones—that complements existing MC simulation techniques.

This paper is a first introduction of the sampling-the-mean algorithm. Future work includes further benchmarking against existing MC methods, for both classic lattice spin systems of much larger size and for specific applications like calcium-induced calcium release models.

Acknowledgments

This work was supported by NIH grants R01-HL057832 and R01-AR054098 (Michael Fill, PI).

References

1. Ising E. Beitrag zur Theorie des Ferromagnetismus. Zeitschrift für Physik. 1925; 31:253–258.
2. Bethe H, Zur Theorie der Metalle I. Eigenwerte und Eigenfunktionen der linearen Atomkette. Z Phys A: Hadrons Nucl. 1931; 71:205–226.
3. Lee TD, Yang CN. Statistical theory of equations of state and phase transitions II. Lattice gas and Ising model. Phys Rev. 1952; 87:410–419.
4. Dyson FJ. Existence of a phase-transition in a one-dimensional Ising ferromagnet. Communications in Mathematical Physics. 1969; 12:91–107.
5. Dziarmaga J. Dynamics of a quantum phase transition: Exact solution of the quantum Ising model. Phys Rev Lett. 2005; 95:245701. [PubMed: 16384394]
6. Hubbard J. Electron correlations in narrow energy bands. Proceedings of the Royal Society of London. Series A Mathematical and Physical Sciences. 1963; 276:238–257.
7. Lieb, EH. The Hubbard model: Some rigorous results and open problems. In: Baeriswyl, D.; Campbell, DK.; Carmelo, JMP.; Guinea, F.; Louis, E., editors. The Hubbard Model: Its Physics and Mathematical Physics. Plenum Press; New York: 1995. p. 1-20.
8. Levy T, Cohen G, Rabani E. Simulating lattice spin models on graphics processing units. J Chem Theory Comput. 2010; 6:3293–3301.

9. Melchert O, Hartmann AK. A dedicated algorithm for calculating ground states for the triangular random bond Ising model. *Comput Phys Commun.* 2011; 182:1828–1832.
10. Bers, DM. *Excitation-Contraction Coupling and Cardiac Contractile Force.* 2. Kluwer Academic Publishers; Dordrecht, The Netherlands: 2001.
11. Hilliard FA, Steele DS, Laver D, Yang Z, Le Marchand SJ, Chopra N, Piston DW, Huke S, Knollmann BC. Flecainide inhibits arrhythmogenic Ca^{2+} waves by open state block of ryanodine receptor Ca^{2+} release channels and reduction of Ca^{2+} spark mass. *J Mol Cell Cardiol.* 2010; 48:293–301. [PubMed: 19835880]
12. Zhou Q, Xiao J, Jiang D, Wang R, Vembaiyan K, Wang A, Smith CD, Xie C, Chen W, Zhang J, Tian X, Jones PP, Zhong X, Guo A, Chen H, Zhang L, Zhu W, Yang D, Li X, Chen J, Gillis AM, Duff HJ, Cheng H, Feldman AM, Song LS, Fill M, Back TG, Chen SRW. Carvedilol and its new analogs suppress arrhythmogenic store overload-induced Ca^{2+} release. *Nat Med.* 2011; 17:1003–1009. [PubMed: 21743453]
13. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell.* 4. Garland Publishing; New York: 2002.
14. Rohr S. Role of gap junctions in the propagation of the cardiac action potential. *Cardiovasc Res.* 2004; 62:309–322. [PubMed: 15094351]
15. Newman, MEJ.; Barkema, GT. *Monte Carlo Methods in Statistical Physics.* Oxford University Press; London: 1999.
16. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state Calculations by fast computing machines. *J Chem Phys.* 1953; 21:1087–1092.
17. Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika.* 1970; 57:97–109.
18. Chib S, Greenberg E. Understanding the Metropolis-Hastings algorithm. *The American Statistician.* 1995; 49:327–335.
19. Cramér, H. *Mathematical Methods of Statistics.* Princeton University Press; Princeton: 1946.
20. Feller, W. *An Introduction to Probability Theory and Its Applications.* 3. Vol. 1. John Wiley and Sons; 1968.
21. Feller, W. *An Introduction to Probability Theory and Its Applications.* 2. Vol. 2. John Wiley and Sons; 1971.
22. Wang F, Landau DP. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys Rev Lett.* 2001; 86:2050–2053. [PubMed: 11289852]
23. Lebowitz JL, Percus JK, Verlet L. Ensemble dependence of fluctuations with application to machine computations. *Phys Rev.* 1967; 153:250–254.
24. Kawasaki K. Diffusion constants near the critical point for time-dependent Ising models. I. *Phys Rev.* 1966; 145:224–230.
25. Liu Y, Porta M, Qin J, Ramos J, Nani A, Shannon TR, Fill M. Flux regulation of cardiac ryanodine receptor channels. *J Gen Physiol.* 2010; 135:15–27. [PubMed: 20008518]
26. Beale PD. Exact distribution of energies in the two-dimensional Ising model. *Phys Rev Lett.* 1996; 76:78–81. [PubMed: 10060438]
27. Gillespie D, Fill M. Intracellular calcium release channels mediate their own countercurrent: The ryanodine receptor case study. *Biophys J.* 2008; 95:3706–3714. [PubMed: 18621826]

Figure 1. (color online) Geometry of the biological test system. A 2×2 array of Ca^{2+} release ion channels is shown. Each channel is a 40 nm square with four inactivation sites near the center (depicted as cyan circles 1.3 nm from center) and four activation sites in the corners (depicted as red circles 18 nm from center). The Ca^{2+} current comes from the center and spreads by diffusion to these sites on neighboring channels, changing the open (Ca^{2+} current conducting) state of the channel through Eq. (23).

Figure 2.

(color online) The terms in Eq. (9) for the 5×5 test system as a function of n , the number of

up states. The solid black curve connects the exactly calculated $\binom{N}{n} \bar{Q}(n)$ for each n . The red boxes are the standard deviations of the trial means for $\nu = 10$ and $\eta = 10^4$. Those without error bars were computed exactly by generating all possible states with n up spins because $\binom{N}{n} < \eta \nu = 10^5$.

Figure 3.

(color online) The standard deviation of the trial samples for the 5×5 test system with 12 up states as η and ν are changed. The area of the circles corresponds to the standard deviation. In each vertical column (represented by one color), the total amount computational effort ($\eta\nu$) is constant at 10⁴ (red), 10⁵ (green), or 10⁶ (blue) and η is varied on the vertical axis. The corresponding ν is shown next to each circle.

Figure 4.

(color online) The accuracy of the sampling-the-mean algorithm for the 5×5 test system with 12 up states as η and ν are changed. The horizontal black line shows the exact result. Within each of the three groups of colored points with error bars, the total amount computational effort ($\eta\nu$) is constant at 10^4 (red), 10^5 (green), or 10^6 (blue) and ν is varied. If error bars are not visible, then they are smaller than the symbol.

Figure 5.

(A) The terms in Eq. (13) for the 10×10 test system as a function of n , the number of up states, as calculated with the sampling-the-mean algorithm using $\nu = 10$ and $\eta = 100$. Note the very small number of samples per trial mean (η) that give a “quick and dirty” calculation of these terms. Each took on the order of 1.5 seconds using the cores and parallelization described in the main text. (B) Focusing on the $n = 72$ point in panel (A) and changing the number of samples η with $\nu = 10$ fixed. Note that the $\eta = 100$ result shown in panel (A) is

not included in this plot. The value of $\binom{N}{n} \tilde{Q}(n)$ and its standard deviation error bars are shown on the left axis. The percent deviation from the $\eta = 10^7$ result is shown on the right

axis. Below each error bar is the computation time using the cores and parallelization described in the main text. Each is 10 times longer than the previous, as expected.

Figure 6.

(color online) (A) The density of energy states for a 50×50 Ising model with periodic boundary conditions where the energy is given by the sum over the states of nearest neighbors (denoted $\langle i, j \rangle$); that is, $E(a) = -\sum_{\langle i, j \rangle} a_i a_j$. The exact result (black) was calculated using the freely-available program by Beale [26]. The sampling-the-mean result (red) used $\nu = 6$ and $\eta = 10^5$ and took ~ 9.8 hours. Energies above ~ 500 were not visited in the sampling. (B) Density of pseudo-energy states for a 5×5 test system with physiological $i_{Ca} = 0.35$ pA [27] (top, red) and elements of parameters from Ref. [25] and a physiological the corresponding partition sum (bottom, black). Note that the point for $g(0.00784) = 1$ that is far removed from the rest of the energy spectrum. Both results are exact by directly calculating every possible state. (C) Expanded view of selected elements of the partition sum from panel B. Note that the linear scale here versus the logarithmic scale in panel B.