Population Annealing and Its Application to a Spin Glass

Cite as: AIP Conference Proceedings **690**, 200 (2003); https://doi.org/10.1063/1.1632130
Published Online: 06 November 2003

K. Hukushima, and Y. Iba

Population Annealing and Its Application to a Spin Glass

K. Hukushima* and Y. Iba†

*Department of Basic Science, University of Tokyo 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan
†The Institute of Statistical Mathematics, Minato-ku, Tokyo, Minami-Azabu, Japan

Abstract. A way to modify simulated annealing to a Monte Carlo algorithm for calculating canonical averages is presented. The proposed algorithm is based on the idea of population based Monte Carlo method where multiple replicas of the original system are used to represent a target distribution. Inspired by non-equilibrium work relation of Jarzynski, an appropriate weight of the replicas is introduced, which enable correct computation of the canonical averages in the limit of infinite number of replicas. The method is applied to a spin glass model and its efficiency is discussed.

INTRODUCTION

Since the Metropolis algorithm was introduced in 1953[1], Monte Carlo (MC) methods have been used intensively in a wide area of physics[2] and statistical sciences[3, 4]. Most of the MC methods in statistical physics are based on the Metropolis strategy, in which a Markov chain is constructed in order for its invariant distribution to coincide with the desired distribution. There exist various improvements on the Metropolis MC algorithms, which mainly categorized into two directions, i.e., non-local updating methods such as cluster algorithm[5] and extended ensemble methods[6]. Typical examples of the latter are the multicanonical method[7], the simulated tempering[8] and the exchange MC[9] or parallel tempering method. These methods have been applied to various complex systems, e.g. protein models and spin glasses, and turned out to be quite useful for simulating these systems.

An alternative class of MC methods which does not belong to the Metropolis stragety, on the other hand, has been studied from early times. This is called population Monte Carlo algorithm[10], in which multiple replicas are used to represent distributions to be studied. A most famous example in physics is diffusion Monte Carlo for quantum systems. Recently the population MC algorithms for polymer simulations have developed as the pruned-enriched Rosenbluth method[11]. For an interdisciplinary review, see Ref. [10].

In the repsent article, we propose an improved MC algorithm which belongs to the population MC method. This work was partly motivated by the recent discover of a non-equilibrium relationship of Jarzynski[12, 13]. In equilibrium thermodynamics, there is the minimum work principle $\Delta F \leq \langle W \rangle$, where ΔF denotes the equilibrium free-energy difference between the initial and final states of a system, W is the total work performed in switching an external parameter of the system, and $\langle \cdots \rangle$ is an average

CP690, *The Monte Carlo Method in the Physical Sciences*, edited by J. E. Gubernatis

over possible histories in such switching measurements. The equality holds when and only when the switching process is infinitely slowly perfomed. Jarzynski[12] presented the interesting equality between the equilibrium states for the finite-time switching, $\exp(-\beta\Delta F) = \langle \exp(-\beta W) \rangle$ where β is the inverse temperature. This relation is called Jarzynski equality. While the equality has been extended to other non-equilibrium systems, it is also regarded as a numerical tool for calculating the equilibrium free-energy difference[13, 14]. Neal[15] has independently proposed a MC scheme by applying the Jarzynski equality to a temperature annealing process such as simulated annealing, and explicitly show that not only the free-energy difference but also an equilibrium canonical average of any physical quantity is estimated by an appropriate weight factor. A straightforward application of this MC method, however, could not give a stable calculation in statistical-mechanical systems, because the weight factor is largely fluctuated in the simulation. In order to avoid such difficulty, we introduce a resampling technique in the switching process of the MC simulation, which plays a crucial role in an accurate estimate of the canonical average. The method, which we call population annealing (PA), is demonstrated in an application to a spin glass model which exhibits extremely slow dynamics, namely strong dependence of annealing process.

This article is organized as follows. In the next section, we present the proposed MC algorithm in detail. An application of the method to a spin glass model is shown in the subsequent section. The final section summarizes our conclusion.

POPULATION ANNEALING

We give an implementation of the population annealing for canonical distributions. In principle, the population annealing can be used with any family of distributions parameterized by a parameter, which plays the role of β in the canonical distribution.

Let us represent the state of the system by x and consider a family of canonical distributions parameterized by the inverse temperature β

$$P_\beta(x) = \frac{\exp(-\beta E(x))}{Z_\beta}, \tag{1}$$

where $E(x)$ is the energy of the state x and Z_β is the partition function at β. Consider K copies (replicas or particles) of the state x, which are randomly initialized. We will indicate them by $\{x^k\}(1 \le k \le K)$. These replicas $\{x^k\}$ are simulated in a parallel manner as described in the following description of the algorithm. We also associate weight W^k to each replica k, which is initialized by $W_0^k = 1$. Starting from sufficiently high temperature $1/\beta_0$ at which the randomized initial states are regarded as sampling from the canonical ensemble, we repeat the following procedures with decreasing sequences of temperatures $\{1/\beta_i\}$ with i being the index of the temperature.

Step 1: *Calculate weights of the replicas*:
For each replica k, the new weight is given in the recursive form by

$$W_i^k = W_{i-1}^k \exp\left(-(\beta_i - \beta_{i-1})E(x_i^k)\right).$$

The weight, which we call Neal-Jarzynski (NJ) factor, depends on the energy E of a given sequence of the states $X^k = \{x_0^k, x_1^k, \cdots, x_i^k\}$.

Step 2: *Resampling (Split/Remove) of the replicas*:
If $i \equiv 0 \,(\mathrm{mod}\, M)$, then the following procedure is performed. For all replicas k, the probability is set as

$$P^k = \frac{W_i^k}{\sum_k W_i^k}.$$

Then, the new replica x_i^k is resampled according to the probability P^k. In this procedure the total number K of the replicas is strictly preserved and a multiple selection of an old replica is allowed. Thus, a replica with a small weight W^k is removed with a high probability, while a replica with a large weight tends to have multiple "descendants" . After the resampling, all the weights are re-initialized to the unity,

$$W^k = 1.$$

The interval M of the resampling procedure should be properly chosen. In general, we can perform the procedure **2** with unequal intervals. But we should be careful to on-line adaptive implementation, because it can introduce systematic bias.

Step 3: *Monte Carlo update of the replicas*:
Each replica x_i^k is updated independently with $\beta = \beta_i$ using the energy $E(x)$ in *finite* number of Monte Carlo steps (MCS) (usually, a small number of steps is preferable, for example, 1 MCS). Any dynamical MC algorithm which can sample from the canonical distribution (1) with $\beta = \beta_i$ can be used for the update.

Step 4: *Calculate averages*:
The canonical average of any desired physical quantity A is then calculated by

$$\langle A \rangle_{\beta_i} = \frac{\sum_k A_i^k W_i^k}{\sum_k W_i^k}. \tag{2}$$

The free-energy difference $\Delta F = \beta_i F(\beta_i) - \beta_0 F(\beta_0)$ or the partition function is also calculated by

$$\exp(-\Delta F) = Z_{\beta_i}/Z_{\beta_0} = \frac{1}{K}\sum_k W_i^k. \tag{3}$$

This part of the algorithm can be done at the end of the simulation, when we store $\{A_i^k\}$ and $\{W_i^k\}$.

Step 5: *Set* $i := i+1$ *and return to step 1*.

In Fig. 1, an example of the "pedigree", i.e., the graph of ancestor - descendants relations is shown. For a replica in the population at the lowest temperature, its "ancestor" in the highest temperature $1/\beta_0$ is identified, and then all the descendants of it are drawn in the figure with a gray level. Repeating this procedure, we can visualize the branching process induced by the resampling process **2**. Note that the descendants of an ancestor are not necessarily in the same state, because Monte Carlo updates in the step **3** described below changes the states.

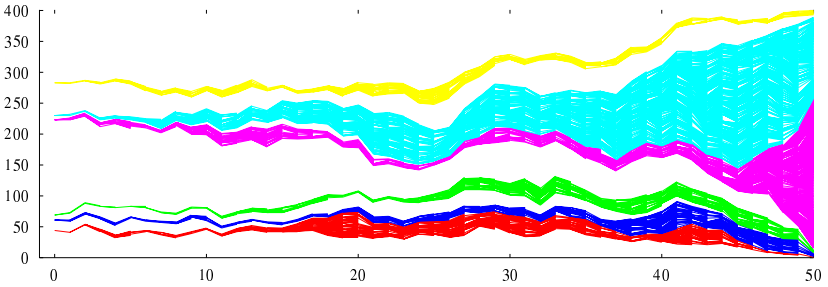

FIGURE 1. A pedigree of the replicas (particles). The leftmost line corresponds to the highest temperature, while the rightmost line corresponds to the lowest temperature. The present data is taken from a toy-simulation with a small number of replicas.

The algorithm without the steps **1** and **2** reduces *simulated annealing* (SA), which is a method for optimization, or equivalently a tool for finding the ground states of the system. As is well known, SA is not a sampling method for the canonical ensemble and does not correctly reproduce the canonical averages at finite temperatures. On the other hand, the one with the step **1** is formally correct even without the step **2**. Its efficiency is, however, severely affected by the increase of the variance of weights $\{W^k\}$. It would not work well in a complex and large-scale system.

The algorithm without the step **2** essentially the same as fast growth method proposed by Hendrix and Jarzynski[14], though they focused on computing the free-energy difference. Neal [15] discussed a similar idea named annealed importance sampling in a more transparent manner. From this viewpoint, the population annealing algorithm is regarded as an extension of the fast growth method or the annealed importance sampling.

The population annealing simulates multiple replicas in a parallel manner, which is similar to the exchange MC (parallel tempering) in this sense. However, these are essentially different methods. In the exchange MC, each replica has a different temperature which changes stochastically during the simulation, while in the population annealing the temperature is common for all the replicas and gradually decreases in the simulation.

APPLICATION TO SPIN GLASS

To illustrate the efficiency of our population annealing, we have tested it on a three-dimensional Ising spin glass model, which is a challenging problem since the system is expected to have many meta-stable states and consequently exhibits extremely slow dynamics[16]. We also perform SA and the exchange MC and compare these results to that of the present method.

The model Hamiltonian is given by

$$H = -\sum_{\langle ij \rangle} J_{ij} S_i S_j - H \sum_i S_i, \tag{4}$$

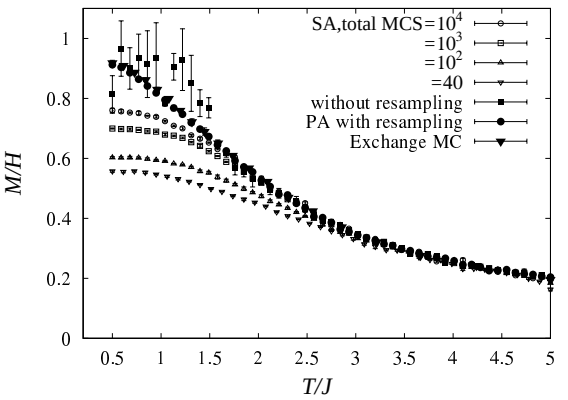

FIGURE 2. Comparison of the population annealing (PA), the simulated annealing (SA) with different cooling rates, an algorithm without resampling (with NJ factor, Neal-Jarzynski algorithm), and the exchange MC. The magnetization at each temperature is shown in the vertical axis.

where S_i denotes the Ising variable defined on a simple cubic lattice. The nearest neighbor coupling J_{ij} obeys the Gaussian distribution with a zero mean and a variance J^2 and H is an uniform external field.

During each annealing simulation in SA and the population annealing, the temperature T/J is decreased with a constant cooling rate from $T/J = 5.0$ to 0.5. The lowest temperature we have examined is well below the critical temperature $T_c(\simeq 0.95)$. In Fig. 2, we show temperature dependence of the magnetization $M = \frac{1}{N}\sum_i \langle S_i \rangle$ induced by the external field with $H/J = 0.1$. The total MC steps in SA are 40, 10^2, 10^3 and 10^4 MCS during the annealing process. We use the only one cooling rate with 10^3 MCS for PA. The total number of the replicas is chosen to be $K = 1600$ both in SA and PA. Furthermore we perform the exchange MC in order to check the equilibrium value of the induced magnetization.

As shown in Fig. 2, the results of SA considerably depend on the cooling rate at low temperatures and hardly saturate to the equilibrium values. This is a typical example of the slow dynamics in spin-glass simulations, demonstrating that SA gives incorrect canonical averages even with a very slow cooling rate. The weighted average of the magnetization using Eq. (2) provides us a correct estimate of the canonical averages *in principle*. The result of the algorithm without the resampling, i.e., step **2** explained above, however, shows large statistical error. The proposed method gives good results, which coincides with the results of the exchange MC method. That is, in this example, both of the steps **1** and **2** are crucial for the efficient calculation of canonical averages.

We also observe the spin-glass order parameter defined as $q^{(2)} = \left\langle \left(\frac{1}{N}\sum_i S_i^{(1)} S_i^{(2)} \right)^2 \right\rangle$ where the upper suffix denotes the replica index. In this simulation for calculating $q^{(2)}$, we prepare the two real replicas, not the same as the replicas in PA, with different initial

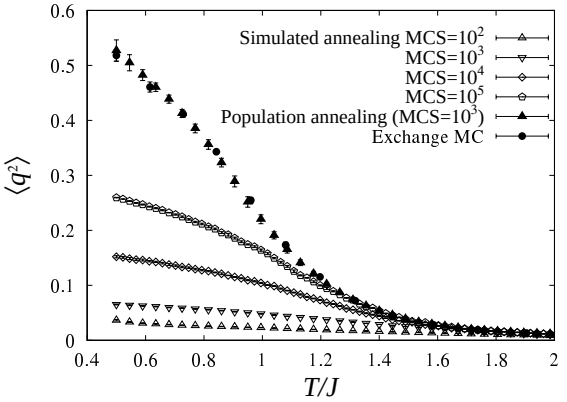

FIGURE 3. Temperature dependence of the spin-glass order parameter. These are obtained by the simulated annealing with four different cooling rates, the population annealing and the exchange MC method.

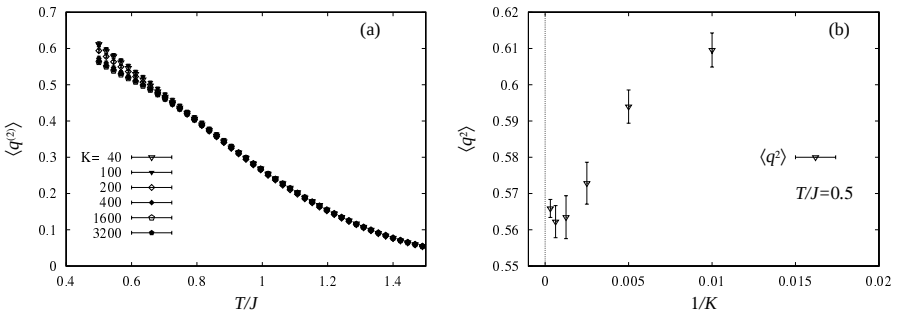

FIGURE 4. The spin glass order parameter q^2 calculated by using PA with 6 different numbers of replicas. (a) Temperature dependence of q^2. (b) The replica number dependence at temperature $T/J = 0.5$.

conditions and different sequences of random numbers. Fig. 3 shows the result of $q^{(2)}$ in the model (4) under the zero field. Again, the estimates by PA are consistent with those obtained by the exchange MC, showing the validity of the proposed method, while the results by SA even with the slowest annealing still largely deviate from the data by two other methods. The free energy (difference), not shown here, is also successfully calculated through Eq. (3) as a function of temperature.

Finally we discuss an effect of finite-replica-numbers(samplings) which leads to a systematic correction error to the infinite-number result[13, 17]. We perform a couple of simulations by varying the replica number K from 40 to 3200. We present in Fig. 4(a) temperature dependence of $q^{(2)}$. While the systematic error is not seen at higher temperatures, it becomes significant for small K at lower temperatures. In Fig.4(b), we plot the

data of $q^{(2)}$ as a function of the inverse replica number. It is found that the asymptotic behavior for large K is roughly linear in $1/K$. Although the systematic error due to the finite replica number has investigated for the free energy estimates in the fast growth method (or the NJ algorithm) theoretically[13, 17], a similar analysis is also required for the canonical averages in PA. In particular, it would be more complicated in PA because an effective number of independent replicas is reduced by the resampling procedure. The issue remains to be solved for future study.

SUMMARY

In this article, we have proposed the population annealing, which is regarded as a modified way of the simulated annealing to an algorithm for finite-temperature sampling. We introduce a resampling procedure which splits or removes the replicas according to the weights. This procedure is the key ingredient in the algorithm for calculating the canonical average accurately and stably.

We have applied the method to a three-dimensional spin glass model and obtained the canonical averages of the magnetization and the spin-glass order parameter as well as the free energy in relatively fast annealing simulation. Figures 2 and 3 demonstrate the validity of the method.

ACKNOWLEDGMENTS

This work was supported by the Grants-In-Aid (No. 14084204) for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

1. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
2. D. P. Landau and K. Binder, *A guide to Monte Carlo simulations in statistical physics*, Cambridge (2000).
3. J. S. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer, (2001).
4. A. Doucet et al. (eds.), *Sequential Monte Carlo in Practice*, Springer-Verlag (2001).
5. R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).
6. For a recent review, Y. Iba, Int. J. Mod. Phys. C**12**, 623 (2001).
7. B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).
8. E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
9. K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).
10. Y. Iba, Trans. Jpn. Soc. Artif. Intel. **16**, 279 (2001), e-print: cond-mat/0008226.
11. P. Grassberger, Phys. Rev. E **56**, 3682 (1997).
12. C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997)
13. C. Jarzynski, Phys. Rev. E **56**, 5018 (1997).
14. D. A. Hendrix and C. Jarzynski, J. Chem. Phys. **114**, 5974 (2001).
15. R. M. Neal, Statistics and Computing, **11**, 125 (2001).
16. A. P. Young ed. *Spin glasses and random fields*, World Scientific, Singapore, (1997).
17. D. M. Zuckerman and T. B. Woolf, Phys. Rev. Lett. **89**, 180602 (2002).