An Efficient Simulation Method for Determining the Density of States of Complex Systems

Takuya Hayashi and Yuko Okamoto

1 Department of Physics, Graduate School of Science, Nagoya University, Nagoya, Aichi 464-8602, Japan
2 Structural Biology Research Center, Graduate School of Science, Nagoya University, Nagoya, Aichi 464-8602, Japan
3 Center for Computational Science, Graduate School of Engineering, Nagoya University, Nagoya, Aichi 464-8603, Japan
4 Information Technology Center, Nagoya University, Nagoya, Aichi 464-8601, Japan
5 JST-CREST, Nagoya, Aichi 464-8602, Japan

By combining two generalized-ensemble algorithms, Multicanonical Replica-Exchange Method and Replica-Exchange Wang-Landau method, we have developed an effective simulation method to determine the density of states of large and complex systems. In order to verify the effectiveness of our algorithm, we performed simulations of a square-lattice Ising model by the three methods. The results showed that the density of states obtained by the present method is more accurate than that is estimated by the two methods separately.

I. INTRODUCTION

The statistical mechanical expectation value of a physical quantity can be accurately calculated if the density of states (DOS) is given. However, in many cases, we do not know DOS a priori and it is often difficult to obtain it theoretically or experimentally. In recent decades, many methods were developed for the determination of DOS by using Monte Carlo (MC) and/or molecular dynamics (MD) simulations [1–17]. These methods have been successfully applied to a wide range of problems in condensed matter and statistical physics including spin glasses, liquid crystals, polymers, and proteins. Nevertheless, the problem still remains that the exact calculation of DOS cannot be achieved when the systems become large and complex. In this article, we propose an efficient simulation method to get the most precise DOS by combining the Multicanonical Replica-Exchange Method (MUCAREM) [11–13] and the Replica-Exchange Wang-Landau (REWL) method [15, 16].

This article is organized as follows. In Sec. II, we explain the methods. In Sec. III, the computational details are given. In Sec. IV, the results and discussion are presented. Sec. V is devoted to conclusions.

II. COMPUTATIONAL METHODS

We first introduce three basic generalized-ensemble algorithms. The Multicanonical Algorithm (MUCA) [1, 2] is one of the representative methods. A simulation in multicanonical ensemble is based on a non-Boltzmann weight factor, which we refer to as the multicanonical weight factor. This is inversely proportional to DOS of the system, and a free random walk in potential energy space is realized so that a wide configurational space may be sampled. However, the DOS is often not known a priori. The multicanonical weight factor is usually determined by iterations of short trial simulations [3, 18]. After a production run with the determined MUCA weight factor, the single-histogram reweighting techniques [19] are employed to obtain an accurate DOS. However, the weight factor determination process can be very tedious and difficult. The Wang-Landau (WL) method [4, 5] solved this problem drastically. In the WL sampling, the weight factor, which is also inversely proportional to DOS, is updated during the simulation by adding a constant to the weight factor. This procedure ultimately leads to a uniform histogram in potential energy space, and the modified weight factor converges to the inverse of the DOS. Another powerful algorithm is the Replica-Exchange Method (REM) [6, 7] (it is also referred to as parallel tempering [20]). Closely related method was independently developed in [21]. In this method, several copies (replicas) of the original system at different temperatures are simulated independently and simultaneously by conventional canonical MC or MD. Every few steps, pairs of
replicas are exchanged with a specified transition probability. This exchange process realizes a random walk in temperature space, which in turn induces a random walk in potential energy space. After a production simulation, the multiple histogram reweighting techniques [22, 23] (an extension of which is also referred to as the Weighted Histogram Analysis Method (WHAM) [23]) are used in order to determine the most accurate DOS from all the histograms of sampled potential energy at different temperatures.

These basic simulation methods can be combined for more effective sampling. One method is referred to as the Multicanonical Replica-Exchange Method (MUCAREM) [11–13]. In this method, the total energy range where we want to calculate the DOS is divided into smaller regions, each corresponding to a replica, and MUCA simulations are performed independently and simultaneously in each replica. Every few steps, a pair of neighboring replicas are exchanged like REM. The configurations can be sampled more effectively than ordinary MUCA because of replica exchange. The final, most accurate estimation of DOS is obtained by the multiple-histogram reweighting techniques again [12, 13]. A similar method is the Replica-Exchange Wang-Landau (REWL) method [15, 16]. The idea is almost the same as in MUCAREM except for using WL instead of MUCA for each replica. After a REWL simulation, DOS pieces are obtained for different energy regions. Connecting these pieces at the point where the slope of DOS is coincident, we can obtain the final estimation of DOS over the entire energy range.

We found that the DOS with the highest accuracy can be obtained by combining these two methods. The REWL is employed in the first half of the total number of MC (or MD) steps in order to get a rough estimate of MUCAREM weight factor and the MUCAREM is performed in the second half in order to refine the DOS. We refer to this new method as REWL-MUCAREM. The DOS thus obtained has higher accuracy than that is estimated by the two methods separately.

A brief explanation of MUCA is now given here. The multicanonical probability distribution of potential energy $P_{\text{MUCA}}(E)$ is defined by

$$P_{\text{MUCA}}(E) \propto w(E_{\text{MUCA}}(E)) \equiv \text{const} ,$$

where $W_{\text{MUCA}}(E)$ is the multicanonical weight factor and the function $g(E)$ is the DOS. $E$ is the total potential energy of a system. By omitting a constant factor, we have

$$W_{\text{MUCA}}(E) = \frac{1}{g(E)} .$$

In MUCA MC simulations, the trial moves are accepted with the following Metropolis transition probability $w(E \to E')$:

$$w(E \to E') = \min \left[ 1, \frac{W_{\text{MUCA}}(E')}{W_{\text{MUCA}}(E)} \right] = \min \left[ 1, \frac{g(E)}{g(E')} \right] .$$

(3)

Here, $E$ is the potential energy of the original configuration and $E'$ is that of a proposed one. After a long production run, the best estimate of DOS can be obtained by the single-histogram reweighting techniques:

$$g(E) = \frac{H(E)}{W_{\text{MUCA}}(E)} ,$$

(4)

where $H(E)$ is the histogram of sampled potential energy. Practically, the $W_{\text{MUCA}}(E)$ is set $\exp[-\beta E]$ at first and modified by repeating sampling and reweighting. Here, $\beta$ is the inverse of temperature $T$ ($\beta = 1/k_B T$ with $k_B$ being the Boltzmann constant).

The WL also uses $1/g(E)$ as the weight factor and the Metropolis criterion is the same as in Eq. (3). However, $g(E)$ is updated dynamically as $g(E) \to f \times g(E)$ during the simulation when the simulation visits a certain energy value $E$. $f$ is a modification factor. We continue the updating until the energy histogram becomes flat. If $H(E)$ is flat enough, a next simulation begins after resetting the histogram to zero and reducing the modification factor (usually, $f \to \sqrt{f}$). The flatness evaluation can be done in various ways. In this article we considered that the histogram is sufficiently flat when

$$\frac{H_{\text{min}}}{H_{\text{max}}} > 0.5 ,$$

(5)
where $H_{\text{min}}$ and $H_{\text{max}}$ are the least number and the largest number of nonzero entries in the histogram, respectively \cite{24}. This process is terminated when the modification factor attains a predetermined value $f_{\text{final}}$ and $\exp(10^{-8}) \approx 1.00000001$ is often used as $f_{\text{final}}$. Hence, the estimated $g(E)$ tends to converge to the true DOS of the system within this much accuracy set by $f_{\text{final}}$.

In MUCAREM, the entire energy range of interest $[E_{\text{min}}, E_{\text{max}}]$ is divided into $M$ sub-regions, $E_{\text{min}}^{(m)} \leq E \leq E_{\text{max}}^{(m)}$ ($m = 1, 2, \ldots, M$), where $E_{\text{min}}^{(1)} = E_{\text{min}}$ and $E_{\text{max}}^{(M)} = E_{\text{max}}$. There should be some overlaps between the adjacent regions. MUCAREM uses $M$ replicas of the original system. The weight factor for sub-region $m$ is defined by

$$W_{\text{MUCAREM}}^{(m)}(E) = \begin{cases} e^{-\beta^{(m)}_{L}E}, & \text{for } E < E_{\text{min}}^{(m)}, \\ g_{m}(E), & \text{for } E_{\text{min}}^{(m)} \leq E \leq E_{\text{max}}^{(m)}, \\ e^{-\beta^{(m)}_{H}E}, & \text{for } E > E_{\text{max}}^{(m)}, \end{cases}$$

(6)

where $g_{m}(E)$ is the DOS for $E_{\text{min}}^{(m)} \leq E \leq E_{\text{max}}^{(m)}$ in sub-region $m$, $\beta^{(m)}_{L} = d \log [g_{m}(E)] / dE$ ($E = E_{\text{min}}^{(m)}$) and, $\beta^{(m)}_{H} = d \log [g_{m}(E)] / dE$ ($E = E_{\text{max}}^{(m)}$). The MUCAREM weight factor $W_{\text{MUCAREM}}(E)$ for the entire energy range is expressed by the following formula:

$$W_{\text{MUCAREM}}(E) = \prod_{m=1}^{M} W_{\text{MUCAREM}}^{(m)}(E).$$

(7)

After a certain number of independent MC steps, replica exchange is proposed between two replicas, $i$ and $j$, in neighboring sub-regions, $m$ and $m+1$, respectively. The transition probability, $w_{\text{MUCAREM}}$, of this replica exchange is given by

$$w_{\text{MUCAREM}} = \min \left[ 1, \frac{W_{\text{MUCAREM}}^{(m)}(E_{i}) W_{\text{MUCAREM}}^{(m+1)}(E_{j})}{W_{\text{MUCAREM}}^{(m)}(E_{j}) W_{\text{MUCAREM}}^{(m+1)}(E_{i})} \right],$$

(8)

where $E_{i}$ and $E_{j}$ are the energy of replicas $i$ and $j$ before the replica exchange, respectively. If replica exchange is accepted, the two replicas exchange their weight factors $W_{\text{MUCAREM}}^{(m)}(E)$ and $W_{\text{MUCAREM}}^{(m+1)}(E)$ and energy histogram $H_{m}(E)$ and $H_{m+1}(E)$. The final estimation of DOS can be obtained from $H_{m}(E)$ after a simulation by the multiple-histogram reweighting techniques or WHAM. Let $n_{m}$ be the total number of samples for the $m$-th energy sub-region. The final estimation of DOS, $g(E)$, is obtained by solving the following WHAM equations self-consistently by iteration \cite{12}:

$$\begin{aligned}
g(E) & = \frac{\sum_{m=1}^{M} H_{m}(E)}{\sum_{m=1}^{M} n_{m} \exp (f_{m}) W_{\text{MUCAREM}}^{(m)}(E)} , \\
\exp (-f_{m}) & = \sum_{E} g(E) W_{\text{MUCAREM}}^{(m)}(E) .
\end{aligned}$$

(9)

These MUCAREM sampling and WHAM reweighting processes can, in principle, be repeated to obtain more accurate DOS \cite{13}. We remark that REM is often used to obtain the first estimate of DOS in the MUCAREM iterations. We also remark that when REM instead of MUCAREM is performed, the best estimate of DOS can be obtained by solving Eq. (9), where $W_{\text{MUCAREM}}^{(m)}(E)$ is replaced by $\exp(-\beta^{(m)}_{m}E)$ with temperature $T_{m}$ ($\beta_{m} = 1/k_{B}T_{m}$) for $(m = 1, 2, \ldots, M)$.

The REWL method is based on essentially the same weight factors as in MUCAREM, while the WL simulations replace the MUCA simulations for each replica. This simulation is terminated when the modification factors on all sub-regions attain a certain minimum value $f_{\text{final}}$. After a REWL simulation, $M$ pieces of DOS fragments with overlapping energy intervals are obtained. The fragments need to be connected in order to
TABLE I. Conditions of simulations to estimate the density of states.

| Methods       | Number of spins $N$ | Number of replicas $M$ | Frequency (in MC sweeps) of flatness evaluation in Eq. (5) | Total MC sweeps per replica |
|---------------|---------------------|------------------------|----------------------------------------------------------|-----------------------------|
| MUCAREM       | 64                  | 4                      | NA                                                       | 200 000                     |
|               | 256                 | 8                      |                                                          | 200 000                     |
|               | 1024                | 16                     |                                                          | 200 000                     |
|               | 4096                | 32                     |                                                          | 500 000                     |
|               | 16384               | 64                     |                                                          | 3 000 000                   |
| REWL          | 64                  | 4                      | 1000                                                     | 200 000                     |
|               | 256                 | 8                      |                                                          | 200 000                     |
|               | 1024                | 16                     |                                                          | 200 000                     |
|               | 4096                | 32                     |                                                          | 500 000                     |
|               | 16384               | 64                     |                                                          | 3 000 000                   |
| REWL–MUCAREM  | 64                  | 4                      | 1000–NA                                                  | 100 000–100 000             |
|               | 256                 | 8                      |                                                          | 100 000–100 000             |
|               | 1024                | 16                     |                                                          | 100 000–100 000             |
|               | 4096                | 32                     |                                                          | 250 000–250 000             |
|               | 16384               | 64                     |                                                          | 1 500 000–1 500 000        |

determine the final DOS in the entire energy range $[E_{\text{min}}, E_{\text{max}}]$. The joining point for any two overlapping DOS pieces is chosen where the inverse microcanonical temperature $\beta = d \log [g(E)]/dE$ coincides best.

III. COMPUTATIONAL DETAILS

In order to compare the effectiveness of the REWL-MUCAREM with other methods, we performed simulations of a 2-dimensional Ising model with periodic boundary conditions.

In a square-lattice Ising model, the total energy $E$ is defined by

$$E = -J \sum_{\langle i,j \rangle} S_i S_j,$$

where $i$ and $j$ are labels for lattice points. $J$ is the magnitude of interaction between neighboring spins. In this article, $J$ and $k_B$ are set to one for simplicity. $\langle i,j \rangle$ represents pairs of nearest-neighbor spins. $S_i$ is the state of spin on a lattice point $i$ and takes on values of $\pm 1$. Beale calculated the exact DOS of the model of finite sizes [25, 26].

Table I lists the conditions of our simulations. The total number of spins $N$ is $L^2$, where $L$ is the length of a side of the square lattice. The total number of spins considered was $N = 64, 256, 1024, 4096,$ and $16384$. One MC sweep is defined as an evaluation of Metropolis criteria $N$ times. The cost of computations (for example, the total number of MC sweeps) was set equal. However, we should point out that while the ordinary REWL algorithm is terminated when the recursion factor $f$ converged to $f_{\text{final}}$, our REWL simulations were finished after a certain fixed number of flatness evaluations had been made.

A Marsaglia random number generator was employed and we used the program code on open source [18, 27]. The number of replicas was set equal to $L/2$. Each replica performed a MUCA simulation in MUCAREM or a WL simulation in REWL within their energy sub-regions, which had an overlap of about 80 percent between neighboring sub-regions. In the cases of REWL and REWL-MUCAREM simulations, the WL flatness criterion in Eq. (5) was tested every 1000 MC sweeps. If the histogram of energy distribution is sufficiently flat at this time, the WL recursion factor was reduced. Replica exchange was tried every 100 MC sweeps. This means that replica exchange was tried 2000 times altogether for the systems up to $N = 1024$, 5000 times for $N = 4096$, and 30000 times for $N = 16384$ (see Table I). The cost of calculation in our simulations was measured by the total number of MC sweeps because we spend most of computational time...
to perform MC simulations. With the conditions in Table I, we made \( n = 25 \) independent runs with different initial random number seeds in order to estimate errors (and we obtained DOS estimates 25 times for each set of conditions). In this work, we did not iterate the DOS evaluation during the MUCAREM simulations for simplicity. In the present MUCAREM simulation, the first half of the total MC sweeps was run with REM and the remaining of the simulation was MUCAREM with the DOS obtained from the REM simulation. In the REM simulation, \( M \) temperature values were evenly distributed between \( \beta_1 = 1.0 \) and \( \beta_M = 0.01 \).

Finally, in REWL-MUCAREM, MUCAREM simulation was performed with the DOS determined by REWL.

**Fig. 1.** The specific heat. (a) is the exact solutions which were calculated by the exact DOS \([25, 26]\). (b), (c), and (d) were obtained by simulations with \( L = 8, 16, 32, 64, \) and 128 by MUCAREM, REWL, and REWL-MUCAREM, respectively.

### IV. RESULTS AND DISCUSSION

Fig. 1 shows the specific heat which was calculated from the estimated DOS by using the following equation:

\[
C(T) = \frac{\langle E^2 \rangle_T - \langle E \rangle_T^2}{T^2},
\]  

(11)
where

\[
(A)_T = \frac{\sum E A(E) g(E) e^{-\beta E}}{\sum E g(E) e^{-\beta E}},
\]

(12)

and \(A(E)\) is any physical quantity that depends on \(E\). The specific heat in Figs. 1(b)–1(d) obtained from the simulations is defined by

\[
C_{\text{sim}}(T) = \frac{1}{n} \sum_{i=1}^{n} C_{\text{sim}}^{(i)}(T),
\]

(13)

where \(C_{\text{sim}}^{(i)}(T)\) is the specific heat calculated from Eqs. (11) and (12) for the \(i\)-th simulation \((i = 1, 2, \cdots, n)\).

The errors in Fig. 1 (and Fig. 2 below) were estimated by the standard error:

\[
\varepsilon_C(T) = \sqrt{\frac{\sum_{i=1}^{n} \left( C_{\text{sim}}^{(i)}(T) - C_{\text{sim}}(T) \right)^2}{n(n-1)}}.
\]

(14)

FIG. 2. The differences of specific heat between the simulation results and exact one in Eq (15).
FIG. 3. Mean local flatness $G(E)$ in Eq. (16) (red curves). The total number of spins is $N = 32 \times 32$. $G(E)$ were obtained from the simulations by (a) MUCAREM, (b) REWL, and (c) REWL-MUCAREM. The error bars (vertical yellow bars) were obtained from Eq. (17). The best estimated DOS will give $G(E) = 1$.

The exact values of specific heat in finite sizes were obtained by Ferdinand and Fisher [28]. Here, we calculated the exact specific heat in Fig. 1(a) from the exact DOS, $g_{\text{EXACT}}(E)$, of Beale [25, 26] by using Eqs. (11) and (12). We used the Mathematica code, which is given in [26], for the calculations of $g_{\text{EXACT}}(E)$. All the algorithms could reproduce the exact solutions almost correctly.

The differences between exact values and simulation results are shown in Fig. 2. It was calculated by the following equation:

$$\Delta C(T) = C_{\text{EXACT}}(T) - C_{\text{sim}}(T).$$

Note that $|\Delta C(T)|$ takes maximum values around the phase transition temperature $T_c = 2/\log(1 + \sqrt{2}) \simeq 2.269$ in each method. The Fig. 2 implies that the results of the three methods agree with the exact results in the order of REWL-MUCAREM, REWL, and MUCAREM. It means that REWL-MUCAREM could get more accurate DOS than the other two methods.

In order to directly compare the accuracy of DOS among the three methods, we show the mean local
FIG. 4. Global flatness $F$ defined in Eq. (18) as a function of the total number of spins, $N$. If $g_{\text{sim}}^{(i)}(E)$ is equal to $g_{\text{exact}}(E)$ in the entire energy range, $F$ takes a value one.

Flatness $G(E)$ in Fig. 3 for the system of $N = 32 \times 32$, where

$$
\begin{align*}
    G(E) &= \frac{1}{n} \sum_{i=1}^{n} G^{(i)}(E), \\
    G^{(i)}(E) &= \frac{g_{\text{exact}}(E)}{g_{\text{sim}}^{(i)}(E)}
\end{align*}
$$

(16)

Here, $g_{\text{sim}}^{(i)}$ is the DOS estimated from the $i$-th simulation ($i = 1, 2, \cdots, n$). In Ising model, there are two states at the ground state energy $E_0$, which are all spins up or all spins down. Hence, $\log g_{\text{exact}}(E_0)$ takes the value $\log 2$. We matched $\log g_{\text{exact}}(E)$ and $\log g_{\text{sim}}^{(i)}(E)$ at $E_0$. If $g_{\text{sim}}^{(i)}(E)$ is equal to $g_{\text{exact}}(E)$, $G^{(i)}(E)$ becomes flat (= 1) ideally in the entire energy range. The red curves and the yellow vertical bars in Fig. 3 are the values of $G(E)$ and the error bars, respectively. The errors in Fig. 3 were also estimated by the standard error:

$$
\varepsilon(E) = \sqrt{\frac{\sum_{i=1}^{n} \left( G^{(i)}(E) - G(E) \right)^2}{n(n-1)}}.
$$

(17)

Because the error bars are the smallest, REWL-MUCAREM could obtain more precise DOS than REWL and MUCAREM.

In order to examine the accuracy of DOS further, we define the degree of global flatness $F$ by the following formula:

$$
F \equiv \frac{G_{\min}}{G_{\max}},
$$

(18)
TABLE II. Conditions of MUCAREM simulations.

| Methods   | Number of spins | Number of replicas | Number of MC sweeps per replica | Total MC sweeps |
|-----------|-----------------|--------------------|---------------------------------|-----------------|
|           | \(N\) iterations |                    |                                 |                 |
| REM MUCAREM | 100 000         | 100 000            | 200 000 \(\times 4\)          |                 |
| REM MUCAREM | 100 000         | 200 000 \(\times 8\) |                                 |                 |
| REM MUCAREM | 250 000         | 250 000 \(\times 16\) |                                 |                 |
| REM MUCAREM | 500 000         | 500 000 \(\times 32\) |                                 |                 |
| REM MUCAREM | 1 500 000       | 1 500 000 \(\times 64\) |                                 |                 |
| MUCAREM1   | 64              | 4                  | 1 000 000 \(\times 4\)        |                 |
| MUCAREM1   | 256             | 8                  | 1 000 000 \(\times 8\)        |                 |
| MUCAREM1   | 1024            | 16                 | 1 000 000 \(\times 16\)       |                 |
| MUCAREM1   | 4096            | 32                 | 25 000 \(\times 32\)          |                 |
| MUCAREM1   | 16384           | 64                 | 1 500 000 \(\times 64\)       |                 |
| MUCAREM2   | 64              | 8                  | 1 000 000 \(\times 4\)        |                 |
| MUCAREM2   | 256             | 16                 | 1 000 000 \(\times 8\)        |                 |
| MUCAREM2   | 1024            | 32                 | 1 000 000 \(\times 16\)       |                 |
| MUCAREM2   | 4096            | 64                 | 25 000 \(\times 32\)          |                 |
| MUCAREM2   | 16384           | 128                | 1 500 000 \(\times 64\)       |                 |
| MUCAREM3   | 64              | 8                  | 1 000 000 \(\times 4\)        |                 |
| MUCAREM3   | 256             | 16                 | 1 000 000 \(\times 8\)        |                 |
| MUCAREM3   | 1024            | 32                 | 1 000 000 \(\times 16\)       |                 |
| MUCAREM3   | 4096            | 64                 | 25 000 \(\times 32\)          |                 |
| MUCAREM3   | 16384           | 128                | 1 500 000 \(\times 64\)       |                 |

where \(G_{\text{min}}\) is the minimum value of \(G(E)\) over the entire energy range and \(G_{\text{max}}\) is the maximum one. \(F\) takes on values between 0 and 1. The closer the calculated \(g_{\text{sim}}(E)\) is to \(g_{\text{EXACT}}(E)\) globally, the closer \(F\) is to 1. Fig. 4 shows the measured flatness \(F\). Here, as a measure of errors, we define the minimum value of \(F\) and maximum one in Fig. 4 by

\[
\begin{align*}
F_{\text{min}} &= \frac{G_{\text{min}} - \varepsilon(E_{\text{glmin}})}{G_{\text{max}} + \varepsilon(E_{\text{glmax}})}, \\
F_{\text{max}} &= \frac{G_{\text{min}} + \varepsilon(E_{\text{glmin}})}{G_{\text{max}} - \varepsilon(E_{\text{glmax}})},
\end{align*}
\]

(19)

Here, \(\varepsilon(E)\) is the error calculated by Eq. (17), and \(E_{\text{glmin}}\) and \(E_{\text{glmax}}\) are the energy values where \(G_{\text{min}}\) and \(G_{\text{max}}\) are obtained, respectively. It is obvious that the value deteriorates as the size of system gets larger. This means that it was difficult to estimate the DOS of large systems because of the large degrees of freedom. We needed more samples in order to obtain DOS for larger systems. We could not find much differences among the three methods up to \(N = 32 \times 32\). If we consider the system larger than \(N = 32 \times 32\), these three methods gave good results in the order of REWL-MUCAREM, REWL, and MUCAREM.

In the present implementation of MUCAREM, we performed two multiple-histogram reweighting (WHAM) operations: one after the REM simulation in the first half of the run and second after the MUCAREM simulation in the second half of the run. Although a second WHAM operation converges quickly, the calculation cost of the first WHAM operation can become non-negligible in large systems. The REWL-MUCAREM uses only the second WHAM and this WHAM converges even more quickly than in MUCAREM, because a good estimate of DOS is already prepared by the preceding REWL simulation. Hence, REWL first and MUCAREM second is the order that we want to adopt in REWL-MUCAREM. Note also that in REWL-MUCAREM, we do not need the piece-connecting process of DOS required in REWL, because WHAM automatically gives DOS in the entire energy range of interest.

In Figs. 3 and 4, we saw that MUCAREM was the least effective among the three methods. Here, we discuss how we can improve the MUCAREM simulation with the same computational cost (i.e., the same total number of MC sweeps). We performed two more MUCAREM simulations with different conditions. Table II lists the conditions of the additional simulations (MUCAREM2 and MUCAREM3) together with the first MUCAREM in Table I. The MUCAREM1 simulation in Table II is the same as the MUCAREM simulation in Table I. The major differences of the additional MUCAREM simulations from the previous MUCAREM simulation (MUCAREM1) lies in the following: number of MC sweeps for REM and MUCAREM, number of
replicas used for REM, and number of iterations of MUCAREM. In the additional MUCAREM simulations, the 10% of the total MC sweeps was run with REM and the remaining 90% of the simulation was MUCAREM with the DOS obtained from the preceding REM simulation.

It is often said that the DOS obtained by MUCAREM simulations becomes better by iterating MUCAREM and WHAM reweighting [13]. We iterated MUCAREM simulations once for MUCAREM3. In REM simulations in MUCAREM2 and MUCAREM3, replica exchange was proposed every 10 MC sweeps. On the other hand, for MUCAREM in MUCAREM2 and MUCAREM3, replica exchange was proposed every 100 MC sweeps. Twenty-five independent runs with different initial random number seeds were performed (i.e., \( n = 25 \), and we obtained estimated DOS 25 times). It should be mentioned that because we obtained clearly wrong DOS, we performed one extra run for each system \( N = 32 \times 32 \) and \( N = 128 \times 128 \) in MUCAREM2, and simply discarded these apparently bad runs. Giving apparently wrong results suggests that MUCAREM simulations are unstable compared to REWL and REWL-MUCAREM simulations, which implies that the total number of MC sweeps and/or the number of runs \( n \) should be larger for MUCAREM simulations to conclude with confidence.

![Graph](image)

**Fig. 5.** Mean local flatness \( G(E) \) in Eq. (16) (red curves). The total number of spins is \( N = 32 \times 32 \). \( G(E) \) were obtained under the conditions by (a) MUCAREM1, (b) MUCAREM2, and (c) MUCAREM3. The error bars (vertical yellow bars) were obtained from Eq. (17). The best estimated DOS will give \( G(E) = 1 \).

Fig. 5 shows the results of mean local flatness \( G \) in Eq. (16) for the system \( N = 32 \times 32 \). It is obvious that MUCAREM2 and MUCAREM3 gave the DOS with smaller error bars than MUCAREM1. This means that we can estimate more accurate DOS if the conditions are selected carefully. It is worth saying that the width of the error bars of MUCAREM2 and MUCAREM3 were almost the same as in REWL-MUCAREM.
in Fig. 3.  

Fig. 6 shows the global flatness $F$ in Eq. (18). Although there are only small differences in $F$ among the three MUCAREM simulations up to the system $N = 64 \times 64$, we find a large difference in the system $N = 128 \times 128$. MUCAREM3, where MUCAREM simulations were iterated once, could obtain a very good estimate of DOS compared with MUCAREM1 and MUCAREM2. It implies that the DOS became better by iterating MUCAREM simulations rather than by using more sampling obtained from a single, long run of MUCAREM. Note that the estimation of DOS under the conditions of MUCAREM3 is even better than that of REWL-MUCAREM in Fig. 4. We also remark that REWL-MUCAREM would give better DOS if its MUCAREM simulation was iterated. Note also that the error bar for the system $N = 64 \times 64$ is abnormally large in MUCAREM3. We found a bad result from one run out of the 25 runs, which made the error bar very large. These suggest again that MUCAREM simulations are unstable compared to REWL and REWL-MUCAREM simulations.

We can obtain a good estimate of DOS under appropriate conditions and the DOS becomes more accurate by iterating the MUCAREM and WHAM reweighting operations. Although the most suitable conditions will depend on systems and methods, the combination of REWL and MUCAREM can give accurate DOS with less calculation cost (less number of replicas and less total number of MC sweeps).

V. CONCLUSIONS

In this article, we investigated the two existing methods, MUCAREM and REWL, to estimate the density of states in large systems. We proposed an effective method, REWL-MUCAREM, that combines the advantages of REWL and MUCAREM, where REWL is performed first and MUCAREM is performed next. We compared the three methods with a square-lattice Ising model, and we found that REWL-MUCAREM gave the most accurate density of states.

REWL-MUCAREM is effective with other systems. We have calculated the residual entropy of Ice Ih by
REWL-MUCAREM \[29\]. We have also applied it to helix-coil transitions of homo-polymers \[30\]. REWL-MUCAREM can easily be extended to the MD method, because MUCA MD \[31,32\] and WL MD \[33\] have already been developed. REWL-MUCAREM MD simulations of protein folding are now under way.

**Acknowledgements:**
Some of the computations were performed on the supercomputers at the Supercomputer Center, Institute for Solid State Physics, University of Tokyo.

[1] B. A. Berg and T. Neuhaus, Phys. Lett. B 267, 249 (1991).
[2] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. 68, 9 (1992).
[3] W. Janke, Physica A 254, 164 (1998).
[4] F. Wang and D. P. Landau, Phys. Rev. Lett. 86, 2050 (2001).
[5] F. Wang and D. P. Landau, Phys. Rev. E 64, 056101 (2001).
[6] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. 65, 1604 (1996).
[7] C. J. Geyer, in Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface, E. M. Keramidas (ed.), (Interface Foundation, Fairfax Station, 1991) pp. 156-163.
[8] U. H. E. Hansmann, Chem. Phys. Lett. 281, 140, (1997).
[9] Y. Sugita and Y. Okamoto, Chem. Phys. Lett. 314, 141 (1999).
[10] Y. Sugita, A. Kitao, and Y. Okamoto, J. Chem. Phys. 113, 6042 (2000).
[11] Y. Sugita and Y. Okamoto, Chem. Phys. Lett. 329, 261 (2000).
[12] A. Mitsutake, Y. Sugita, and Y. Okamoto, J. Chem. Phys. 118, 6664 (2003).
[13] A. Mitsutake, Y. Sugita, and Y. Okamoto, J. Chem. Phys. 118, 6676 (2003).
[14] K. Hukushima and Y. Iba, AIP Conf. Proc. 690, 200 (2003).
[15] T. Vogel, Y. W. Li, T. Wüst, and D. P. Landau, Phys. Rev. Lett. 110, 210603 (2013).
[16] T. Vogel, Y. W. Li, T. Wüst, and D. P. Landau, Phys. Rev. E 90, 023302 (2014).
[17] F. Yaşar, N. A. Bernhardt, and U. H. E. Hansmann, J. Chem Phys. 143, 224102 (2015).
[18] B. A. Berg, Markov Chain Monte Carlo Simulation and Their Statistical Analysis (World Scientific, Singapore, 2004).
[19] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. 61, 2635 (1988).
[20] E. Marinari, G. Parisi, and J. J. Ruiz-Lorenzo, in Spin Glasses and Random Fields, A. P. Young (ed.) (World Scientific, Singapore, 1997) pp. 59-98.
[21] R. H. Swendsen and J-S. Wang, Phys. Rev. Lett. 57, 2607 (1986).
[22] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. 63, 1195 (1989).
[23] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, J. Comput. Chem. 13, 1011 (1992).
[24] B. A. Berg, C. Muguruma, and Y. Okamoto, Phys. Rev. B 75, 092202 (2007).
[25] P. D. Beale, Phys. Rev. Lett. 76, 78 (1996).
[26] http://spot.colorado.edu/~beale/
[27] G. Marsaglia, A. Zaman, and W. W. Tsang, Stat. Prob. Lett. 8, 35 (1990).
[28] A. E. Ferdinand and M. E. Fisher, Phys. Rev. 185, 832 (1969).
[29] T. Hayashi, C. Muguruma, and Y. Okamoto, manuscript in preparation.
[30] T. Hayashi and Y. Okamoto, manuscript in preparation.
[31] U. H. E. Hansmann, Y. Okamoto, and F. Eisenmenger, Chem. Phys. Lett. 259, 321 (1996).
[32] N. Nakajima, H. Nakamura, and A. Kidera, J. Phys. Chem. B 101, 817 (1997).
[33] T. Nagasima, A. R. Kinjo, T. Mitsui, and K. Nishikawa, Phys. Rev. E 75, 066706 (2007).