Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Graphics and Modelling 22 (2004) 425–439
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Generalized-ensemble algorithms: enhanced sampling techniques for Monte Carlo and molecular dynamics simulations

Yuko Okamoto [a,b]

[a] *Department of Theoretical Studies, Institute for Molecular Science, Okazaki, Aichi 444-8585, Japan*
[b] *Department of Functional Molecular Science, The Graduate University for Advanced Studies, Okazaki, Aichi 444-8585, Japan*

Abstract

In complex systems with many degrees of freedom such as spin glass and biomolecular systems, conventional simulations in canonical ensemble suffer from the quasi-ergodicity problem. A simulation in generalized ensemble performs a random walk in potential energy space and overcomes this difficulty. From only one simulation run, one can obtain canonical ensemble averages of physical quantities as functions of temperature by the single-histogram and/or multiple-histogram reweighting techniques. In this article we review the generalized ensemble algorithms. Three well-known methods, namely, multicanonical algorithm (MUCA), simulated tempering (ST), and replica-exchange method (REM), are described first. Both Monte Carlo (MC) and molecular dynamics (MD) versions of the algorithms are given. We then present five new generalized-ensemble algorithms which are extensions of the above methods.

Keywords: Generalized-ensemble; Monte Carlo; Simulations

1. Introduction

Since the pioneering work of Metropolis et al. [1] half a century ago, computer simulations have been indispensable means of research in many fields of physical science. In the field of molecular science, for instance, a number of powerful simulation algorithms have been developed (for reviews see, e.g., Refs. [2–4]).

Canonical fixed temperature simulations of complex systems such as spin glasses and biopolymers are greatly hampered by the multiple-minima problem, or the quasi-ergodicity problem. Because simulations at low temperatures tend to get trapped in one of huge number of local-minimum-energy states, it is very difficult to obtain accurate canonical distributions at low temperatures by conventional Monte Carlo (MC) and molecular dynamics (MD) methods. One way to overcome this multiple-minima problem is to perform a simulation in a *generalized ensemble* where each state is weighted by an artificial, non-Boltzmann probability weight factor so that a random walk in potential energy space may be realized (for reviews see, e.g., Refs. [5–8]). The random walk allows the simulation to escape from any energy barrier and to sample much wider configurational space than by conventional methods. Monitoring the energy in a single simulation run, one can obtain not only the global-minimum-energy state but also canonical-ensemble averages as functions of temperature by the single-histogram [9] and/or multiple-histogram [10,11] reweighting techniques (an extension of the multiple-histogram method is also referred to as weighted histogram analysis method (WHAM) [11]). Besides generalized-ensemble algorithms, which are usually based on local updates, methods based on non-local updates such as cluster algorithms and their generalizations have also been widely used [12–14]. In this article, we focus our discussion on generalized-ensemble algorithms.

One of the most well-known generalized-ensemble methods is perhaps *multicanonical algorithm* (MUCA) [15,16] (for a review see, e.g., Ref. [17]). (The method is also referred to as *entropic sampling* [18], *adaptive umbrella sampling* [19] *of the potential energy* [20], *random walk algorithm* [21,22], and *density of states Monte Carlo* [23]. MUCA can also be considered as a sophisticated, ideal realization of a class of algorithms called *umbrella sampling* [24]. Also closely related methods are *transition matrix methods* reviewed in Refs. [8,25].) MUCA and its generalizations have been applied to spin systems (see, e.g., Refs. [26–30]). MUCA was also introduced to the molecular simulation field [31]. Since then MUCA and its generalizations have been extensively used in many applications in protein and related systems [32–60]. Molecular dynamics version of

E-mail address: okamotoy@ims.ac.jp (Y. Okamoto).
URL: http://konf2.ims.ac.jp.

1093-3263/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2003.12.009

MUCA has also been developed [38,41,20] (see also Refs. [38,61] for Langevin dynamics version). MUCA has been extended so that flat distributions in other parameters instead of potential energy may be obtained [27,28,37,42,44,59]. Moreover, multidimensional (or multicomponent) extensions of MUCA can be found in Refs. [37,42,43,60].

While a simulation in multicanonical ensemble performs a free 1D random walk in potential energy space, that in *simulated tempering* (ST) [62,63] (the method is also referred to as the *method of expanded ensemble* [62]) performs a free random walk in temperature space (for a review, see, e.g., Ref. [64]). This random walk, in turn, induces a random walk in potential energy space and allows the simulation to escape from states of energy local minima. ST has also been applied to protein folding problem [65,39,40,66].

The generalized-ensemble algorithm is powerful, but in the above two methods the probability weight factors are not a priori known and have to be determined by iterations of short trial simulations. This process can be non-trivial and very tedius for complex systems with many degreees of freedom. Therefore, there have been attempts to accelerate the convergence of the iterative process for MUCA weight factor determination [20,26,37,67–69] (see also Refs. [17,70]).

In the *replica-exchange method* (REM) [71–73], the difficulty of weight factor determination is greatly alleviated. (A closely related method was independently developed in Ref. [74]. Similar methods in which the same equations are used but emphasis is laid on optimizations have been developed [75,76]. REM is also referred to as *multiple Markov chain method* [77] and *parallel tempering* [64]. Details of literature about REM and related algorithms can be found in recent reviews [6,78].) In this method, a number of non-interacting copies (or replicas) of the original system at different temperatures are simulated independently and simultaneously by the conventional MC or MD method. Every few steps, pairs of replicas are exchanged with a specified transition probability. The weight factor is just the product of Boltzmann factors, and so it is essentially known.

REM has already been used in many applications in protein systems [79,80,66,81–91]. Other molecular simulation fields have also been studied by this method in various ensembles [92–97]. Moreover, REM was applied to cluster studies in quantum chemistry field [98]. The details of molecular dynamics algorithm have been worked out for REM in Ref. [80] (see also Refs. [79,95]). This led to a wide application of replica-exchange molecular dynamics method in the protein folding problem [99–106].

However, REM also has a computational difficulty: As the number of degrees of freedom of the system increases, the required number of replicas also greatly increases, whereas only a single replica is simulated in MUCA or ST. This demands a lot of computer power for complex systems. Our solution to this problem is: Use REM for the weight factor determinations of MUCA or ST, which is much simpler than previous iterative methods of weight determinations, and then perform a long MUCA or ST production run. The first example is the *replica-exchange multicanonical algorithm* (REMUCA) [83,88]. In REMUCA, a short replica-exchange simulation is performed, and the multicanonical weight factor is determined by the multiple-histogram reweighting techniques [10,11]. Another example of such a combination is the *replica-exchange simulated tempering* (REST) [84]. In REST, a short replica-exchange simulation is performed, and the simulated tempering weight factor is determined by the multiple-histogram reweighting techniques [10,11].

We have introduced two further extensions of REM, which we refer to as *multicanonical replica-exchange method* (MUCAREM) [83,88] (see also Refs. [107,108]) and *simulated tempering replica-exchange method* (STREM) [109]. In MUCAREM, a replica-exchange simulation is performed with a small number of replicas each in multicanonical-ensemble of different energy ranges. In STREM, on the other hand, a replica-exchange simulation is performed with a small number of replicas in "simulated tempering" ensemble of diffierent temperature ranges.

Finally, one is naturally led to a multidimensional (or, multivariable) extension of REM, which we refer to as *multidimensional replica-exhcange method* (MREM) [82] (see also Refs. [93,106,110,111]). Special realizations of MREM are *replica-exchange free energy perturbation* (REFEP) [82] and *replica-exchange umbrella sampling* (REUS) [82] and they are particularly useful in free energy calculations.

In this article, we describe the eight generalized-ensemble algorithms mentioned above. Namely, we first review three familiar methods: MUCA, ST, and REM. We then present the five new algorithms: REMUCA, REST, MUCAREM, STREM, and MREM (and REFEP and REUS).

2. Generalized-ensemble algorithms

2.1. Multicanonical algorithm and simulated tempering

Let us consider a system of N atoms of mass m_k ($k = 1, \ldots, N$) with their coordinate vectors and momentum vectors denoted by $q \equiv \{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N\}$ and $p \equiv \{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_N\}$, respectively. The Hamiltonian $H(q, p)$ of the system is the sum of the kinetic energy $K(p)$ and the potential energy $E(q)$:

$$H(q, p) = K(p) + E(q), \tag{1}$$

where

$$K(p) = \sum_{k=1}^{N} \frac{\boldsymbol{p}_k^2}{2m_k}. \tag{2}$$

In the canonical ensemble at temperature T each state $x \equiv (q, p)$ with the Hamiltonian $H(q, p)$ is weighted by the Boltzmann factor:

$$W_{\mathrm{B}}(x; T) = \exp(-\beta H(q, p)), \tag{3}$$

where the inverse temperature β is defined by $\beta = 1/k_{\mathrm{B}}T$ (k_{B} is the Boltzmann constant). The average kinetic energy

at temperature T is then given by

$$\langle K(p)\rangle_T = \left\langle \sum_{k=1}^{N} \frac{\boldsymbol{p}_k^2}{2m_k} \right\rangle_T = \frac{3}{2} N k_B T. \tag{4}$$

Because the coordinates q and momenta p are decoupled in Eq. (1), we can suppress the kinetic energy part and can write the Boltzmann factor as

$$W_B(x; T) = W_B(E; T) = \exp(-\beta E). \tag{5}$$

The canonical probability distribution of potential energy $P_B(E; T)$ is then given by the product of the density of states $n(E)$ and the Boltzmann weight factor $W_B(E; T)$:

$$P_B(E; T) \propto n(E) W_B(E; T). \tag{6}$$

Since $n(E)$ is a rapidly increasing function and the Boltzmann factor decreases exponentially, the canonical ensemble yields a bell-shaped distribution which has a maximum around the average energy at temperature T. The conventional MC or MD simulations at constant temperature are expected to yield $P_B(E; T)$. A MC simulation based on the Metropolis algorithm [1] is performed with the following transition probability from a state x of potential energy E to a state x' of potential energy E':

$$\begin{aligned} w(x \to x') &= \min\left(1, \frac{W_B(E'; T)}{W_B(E; T)}\right) \\ &= \min(1, \exp[-\beta(E' - E)]). \end{aligned} \tag{7}$$

A MD simulation, on the other hand, is based on the following Newton equation:

$$\dot{\boldsymbol{p}}_k = -\frac{\partial E}{\partial \boldsymbol{q}_k} = \boldsymbol{f}_k, \tag{8}$$

where $\boldsymbol{f}_k$ is the force acting on the k-th atom ($k = 1 \ldots, N$). This equation actually yields the microcanonical ensemble, and we have to add a thermostat such as Nosé–Hoover algorithm [112,113] and the constraint method [114,115] in order to obtain the canonical ensemble. However, in practice, it is very difficult to obtain accurate canonical distributions of complex systems at low temperatures by conventional MC or MD simulation methods. This is because simulations at low temperatures tend to get trapped in one or a few of local-minimum-energy states.

In the multicanonical ensemble [15,16], on the other hand, each state is weighted by a non-Boltzmann weight factor $W_{mu}(E)$ (which we refer to as the *multicanonical weight factor*) so that a uniform potential energy distribution $P_{mu}(E)$ is obtained:

$$P_{mu}(E) \propto n(E) W_{mu}(E) \equiv \text{constant}. \tag{9}$$

The flat distribution implies that a free random walk in the potential energy space is realized in this ensemble. This allows the simulation to escape from any local minimum-energy states and to sample the configurational space much more widely than the conventional canonical MC or MD methods.

The definition in Eq. (9) implies that the multicanonical weight factor is inversely proportional to the density of states, and we can write it as follows:

$$W_{mu}(E) \equiv \exp[-\beta_0 E_{mu}(E; T_0)] = \frac{1}{n(E)}, \tag{10}$$

where we have chosen an arbitrary reference temperature, $T_0 = 1/k_B\beta_0$, and the "*multicanonical potential energy*" is defined by

$$E_{mu}(E; T_0) \equiv k_B T_0 \ln n(E) = T_0 S(E). \tag{11}$$

Here, $S(E)$ is the entropy in the microcanonical ensemble. Since the density of states of the system is usually unknown, the multicanonical weight factor has to be determined numerically by iterations of short preliminary runs [15,16].

A multicanonical Monte Carlo simulation is performed, for instance, with the usual Metropolis criterion [1]: The transition probability of state x with potential energy E to state x' with potential energy E' is given by

$$\begin{aligned} w(x \to x') &= \min\left(1, \frac{W_{mu}(E')}{W_{mu}(E)}\right) = \min\left(1, \frac{n(E)}{n(E')}\right) \\ &= \min(1, \exp(-\beta_0 \Delta E_{mu})), \end{aligned} \tag{12}$$

where

$$\Delta E_{mu} = E_{mu}(E'; T_0) - E_{mu}(E; T_0). \tag{13}$$

The molecular dynamics algorithm in multicanonical ensemble also naturally follows from Eq. (10), in which the regular constant temperature molecular dynamics simulation (with $T = T_0$) is performed by solving the following modified Newton equation instead of Eq. (8): [38,41]

$$\dot{\boldsymbol{p}}_k = -\frac{\partial E_{mu}(E; T_0)}{\partial \boldsymbol{q}_k} = \frac{\partial E_{mu}(E; T_0)}{\partial E} \boldsymbol{f}_k. \tag{14}$$

From Eq. (11) this equation can be rewritten as

$$\dot{\boldsymbol{p}}_k = \frac{T_0}{T(E)} \boldsymbol{f}_k, \tag{15}$$

where the following thermodynamic relation gives the definition of the "effective temperature" $T(E)$:

$$\left.\frac{\partial S(E)}{\partial E}\right|_{E=E_a} = \frac{1}{T(E_a)}, \tag{16}$$

with

$$E_a = \langle E\rangle_{T(E_a)}. \tag{17}$$

If the exact multicanonical weight factor $W_{mu}(E)$ is known, one can calculate the ensemble averages of any physical quantity A at any temperature $T(= 1/k_B\beta)$ as follows:

$$\langle A\rangle_T = \frac{\sum_E A(E) P_B(E; T)}{\sum_E P_B(E; T)} = \frac{\sum_E A(E) n(E) \exp(-\beta E)}{\sum_E n(E) \exp(-\beta E)}, \tag{18}$$

where the density of states is given by (see Eq. (10))

$$n(E) = \frac{1}{W_{\mathrm{mu}}(E)}. \tag{19}$$

The summation instead of integration is used in Eq. (18), because we often discretize the potential energy E with step size ε ($E = E_i$; $i = 1, 2, \ldots$). Here, the explicit form of the physical quantity A should be known as a function of potential energy E. For instance, $A(E) = E$ gives the average potential energy $\langle E\rangle_T$ as a function of temperature, and $A(E) = \beta^2(E - \langle E\rangle_T)^2$ gives specific heat.

In general, the multicanonical weight factor $W_{\mathrm{mu}}(E)$, or the density of states $n(E)$, is not a priori known, and one needs its estimator for a numerical simulation. This estimator is usually obtained from iterations of short trial multicanonical simulations. The details of this process are described, for instance, in Refs. [26,34]. However, the iterative process can be non-trivial and very tedius for complex systems.

In practice, it is impossible to obtain the ideal multicanonical weight factor with completely uniform potential energy distribution. The question is when to stop the iteration for the weight factor determination. Our criterion for a satisfactory weight factor is that as long as we do get a random walk in potential energy space, the probability distribution $P_{\mathrm{mu}}(E)$ does not have to be completely flat with a tolerance of, say, an order of magnitude deviation. In such a case, we usually perform with this weight factor a multicanonical simulation with high statistics (production run) in order to get even better estimate of the density of states. Let $N_{\mathrm{mu}}(E)$ be the histogram of potential energy distribution $P_{\mathrm{mu}}(E)$ obtained by this production run. The best estimate of the density of states can then be given by the single-histogram reweighting techniques [9] as follows (see the proportionality relation in Eq. (9)):

$$n(E) = \frac{N_{\mathrm{mu}}(E)}{W_{\mathrm{mu}}(E)}. \tag{20}$$

By substituting this quantity into Eq. (18), one can calculate ensemble averages of physical quantity $A(E)$ as a function of temperature. Moreover, ensemble averages of any physical quantity A (including those that cannot be expressed as functions of potential energy) at any temperature $T(= 1/k_{\mathrm{B}}\beta)$ can now be obtained as long as one stores the "trajectory" of configurations (and A) from the production run. Namely, we have

$$\langle A\rangle_T = \frac{\sum_{k=1}^{n_0} A(x(k) W_{\mathrm{mu}}^{-1}(E(x(k))) \exp[-\beta E(x(k))]}{\sum_{k=1}^{n_0} W_{\mathrm{mu}}^{-1}(E(x(k))) \exp[-\beta E(x(E))]}, \tag{21}$$

where $x(k)$ is the configuration at the k-th MC (or MD) step and n_0 is the total number of configurations stored. Note that when A is a function of E, Eq. (21) reduces to Eq. (18) where the density of states is given by Eq. (20).

Eqs. (18) and (21) or any other equations which involve summations of exponential functions often encounter with numerical difficulties such as overflows. These can be overcome by using, for instance, the following equation [116,117]. For $C = A + B$ (with $A > 0$ and $B > 0$) we have

$$\begin{aligned} \ln C &= \ln\left[\max(A, B)\left(1 + \frac{\min(A, B)}{\max(A, B)}\right)\right], \\ &= \max(\ln A, \ln B) + \ln\{1 + \exp[\min(\ln A, \ln B) \\ &\quad - \max(\ln A, \ln B)]\}. \end{aligned} \tag{22}$$

We now briefly review the original *simulated tempering* (ST) method [62,63]. In this method temperature itself becomes a dynamical variable, and both the configuration and the temperature are updated during the simulation with a weight:

$$W_{\mathrm{ST}}(E; T) = \exp(-\beta E + a(T)), \tag{23}$$

where the function $a(T)$ is chosen so that the probability distribution of temperature is flat:

$$\begin{aligned} P_{\mathrm{ST}}(T) &= \int \mathrm{d}E\, n(E) W_{\mathrm{ST}}(E; T) \\ &= \int \mathrm{d}E\, n(E) \exp(-\beta E + a(T)) = \text{constant}. \end{aligned} \tag{24}$$

Hence, in simulated tempering the *temperature* is sampled uniformly. A free random walk in temperature space is realized, which in turn induces a random walk in potential energy space and allows the simulation to escape from states of energy local minima.

In the numerical work we discretize the temperature in M different values, T_m ($m = 1, \ldots, M$). Without loss of generality we can order the temperature so that $T_1 < T_2 < \ldots < T_M$. The lowest temperature T_1 should be sufficiently low so that the simulation can explore the global-minimum-energy region, and the highest temperature T_M should be sufficiently high so that no trapping in an energy-local-minimum state occurs. The probability weight factor in Eq. (23) is now written as

$$W_{\mathrm{ST}}(E; T_m) = \exp(-\beta_m E + a_m), \tag{25}$$

where $a_m = a(T_m)$ ($m = 1 \ldots, M$). Note that from Eqs. (24) and (25) we have

$$\exp(-a_m) \propto \int \mathrm{d}E\, n(E) \exp(-\beta_m E). \tag{26}$$

The parameters a_m are therefore "dimensionless" Helmholtz free energy at temperature T_m (i.e., the inverse temperature β_m multiplied by the Helmholtz free energy). We remark that the density of states $n(E)$ (and hence, the multicanonical weight factor) and the simulated tempering weight parameters a_m are related by a Laplace transform [39]. The knowledge of one implies that of the other, although in numerical work the inverse Laplace transform of Eq. (26) is nontrivial.

Once the parameters a_m are determined and the initial configuration and the initial temperature T_m are chosen, a simulated tempering simulation is then realized by alternately performing the following two steps [62,63]:

1. A canonical MC or MD simulation at the fixed temperature T_m (based on Eqs. (7) or (8)) is carried out for a certain steps.
2. The temperature T_m is updated to the neighboring values $T_{m\pm1}$ with the configuration fixed. The transition probability of this temperature-updating process is given by the Metropolis criterion (see Eq. (25)):

$$w(T_m \to T_{m\pm1}) = \min(1, \exp(-\Delta)), \tag{27}$$

where

$$\Delta = (\beta_{m\pm1} - \beta_m)E - (a_{m\pm1} - a_m). \tag{28}$$

Note that in Step 2 we exchange only pairs of neighboring temperatures in order to secure sufficiently large acceptance ratio of temperature updates.

As in multicanonical algorithm, the simulated tempering parameters $a_m = a(T_m)$ $(m = 1, \ldots, M)$ are also determined by iterations of short trial simulations (see, e.g., Refs. [40,64,65] for details). This process can be non-trivial and very tedius for complex systems.

After the optimal simulated tempering weight factor is determined, one performs a long simulated tempering run once. The canonical expectation value of a physical quantity A at temperature T_m $(m = 1, \ldots, M)$ can be calculated by the usual arithmetic mean as follows:

$$\langle A \rangle_{T_m} = \frac{1}{n_m} \sum_{k=1}^{n_m} A(x_m(k)), \tag{29}$$

where $x_m(k)$ $(k = 1, \ldots, n_m)$ are the configurations obtained at temperature T_m and n_m is the total number of measurements made at $T = T_m$. The expectation value at any intermediate temperature can also be obtained from Eq. (18), where the density of states is given by the multiple-histogram reweighting techniques [10,11] as follows. Let $N_m(E)$ and n_m be respectively the potential-energy histogram and the total number of samples obtained at temperature $T_m = 1/k_B\beta_m$ $(m = 1, \ldots, M)$. The best estimate of the density of states is then given by [10,11]

$$n(E) = \frac{\sum_{m=1}^{M} g_m^{-1} N_m(E)}{\sum_{m=1}^{M} g_m^{-1} n_m \exp(f_m - \beta_m E)}, \tag{30}$$

where we have for each $m(= 1, \ldots, M)$

$$\exp(-f_m) = \sum_E n(E) \exp(-\beta_m E). \tag{31}$$

Here, $g_m = 1 + 2\tau_m$, and τ_m is the integrated autocorrelation time at temperature T_m. For many systems the quantity g_m can safely be set to be a constant in the reweighting formulae [11], and so we usually set $g_m = 1$.

Note that Eqs. (30) and (31) are solved self-consistently by iteration [10,11] to obtain the density of states $n(E)$ and the dimensionless Helmholtz free energy f_m. Namely, we can set all the f_m $(m = 1 \ldots, M)$ to, e.g., zero initially. We then use Eq. (30) to obtain $n(E)$, which is substituted into Eq. (31) to obtain next values of f_m, and so on.

Moreover, ensemble averages of any physical quantity A (including those that cannot be expressed as functions of potential energy) at any temperature $T(= 1/k_B\beta)$ can now be obtained from the "trajectory" of configurations of the production run. Namely, we first obtain f_m $(m = 1, \ldots, M)$ by solving Eqs. (30) and (31) self-consistently, and then we have [88]

$$\langle A \rangle_T = \frac{\sum_{m=1}^{M}\sum_{k=1}^{n_m} A(x_m(k)) \left(g_m^{-1} / \left(\sum_{\ell=1}^{M} g_\ell^{-1} n_\ell \exp[f_\ell - \beta_\ell E(x_m(k))] \right) \right) \exp[-\beta E(x_m(k))]}{\sum_{m=1}^{M}\sum_{k=1}^{n_m} \left(g_m^{-1} / \left(\sum_{\ell=1}^{M} g_\ell^{-1} n_\ell \exp[f_\ell - \beta_\ell E(x_m(k))] \right) \right) \exp[-\beta E(x_m(k))]}, \tag{32}$$

where $x_m(k)$ $(k = 1, \ldots, n_m)$ are the configurations obtained at temperature T_m.

2.2. Replica-exchange method

The *replica-exchange method* (REM) [71–73] was developed as an extension of simulated tempering [71] (thus it is also referred to as *parallel tempering* [64]) (see, e.g., Ref. [80] for a detailed description of the algorithm). The system for REM consists of M *non-interacting* copies (or, replicas) of the original system in the canonical ensemble at M different temperatures T_m $(m = 1, \ldots, M)$. We arrange the replicas so that there is always exactly one replica at each temperature. Then there exists a one-to-one correspondence between replicas and temperatures; the label i $(i = 1, \ldots, M)$ for replicas is a permutation of the label m $(m = 1, \ldots, M)$ for temperatures, and vice versa:

$$\begin{cases} i = i(m) & \equiv f(m), \\ m = m(i) & \equiv f^{-1}(i), \end{cases} \tag{33}$$

where $f(m)$ is a permutation function of m and $f^{-1}(i)$ is its inverse.

Let $X = \{x_1^{[i(1)]}, \ldots, x_M^{[i(M)]}\} = \{x_{m(1)}^{[1]}, \ldots, x_{m(M)}^{[M]}\}$ stand for a "state" in this generalized ensemble. Each "substate" $x_m^{[i]}$ is specified by the coordinates $q^{[i]}$ and momenta $p^{[i]}$ of N atoms in replica i at temperature T_m:

$$x_m^{[i]} \equiv (q^{[i]}, p^{[i]})_m. \tag{34}$$

Because the replicas are non-interacting, the weight factor for the state X in this generalized ensemble is given by the product of Boltzmann factors for each replica (or at each temperature):

$$\begin{aligned} W_{\mathrm{REM}}(X) &= \prod_{i=1}^{M} \exp\{-\beta_{m(i)} H(q^{[i]}, p^{[i]})\} \\ &= \prod_{m=1}^{M} \exp\{-\beta_m H(q^{[i(m)]}, p^{[i(m)]})\} \\ &= \exp\left\{-\sum_{i=1}^{M} \beta_{m(i)} H(q^{[i]}, p^{[i]})\right\} \\ &= \exp\left\{-\sum_{m=1}^{M} \beta_m H(q^{[i(m)]}, p^{[i(m)]})\right\}, \end{aligned} \tag{35}$$

where $i(m)$ and $m(i)$ are the permutation functions in Eq. (33).

We now consider exchanging a pair of replicas in the generalized ensemble. Suppose we exchange replicas i and j which are at temperatures T_m and T_n, respectively:

$$\begin{aligned} &X = \{\ldots, x_m^{[j]}, \ldots, x_n^{[j]}, \ldots\} \rightarrow \\ &X' = \{\ldots, x_m^{[i]'}, \ldots, x_n^{[i]'}, \ldots\}. \end{aligned} \tag{36}$$

Here, i, j, m, and n are related by the permutation functions in Eq. (33), and the exchange of replicas introduces a new permutation function f':

$$\begin{cases} i &= f(m) \rightarrow j = f'(m), \\ j &= f(n) \rightarrow i = f'(n). \end{cases} \tag{37}$$

The exchange of replicas can be written in more detail as

$$\begin{cases} x_m^{[i]} \equiv (q^{[i]}, p^{[i]})_m &\rightarrow x_m^{[j]'} \equiv (q^{[j]}, p^{[j]'})_m, \\ x_n^{[j]} \equiv (q^{[j]}, p^{[j]})_n &\rightarrow x_n^{[i]'} \equiv (q^{[i]}, p^{[i]'})_n, \end{cases} \tag{38}$$

where the definitions for $p^{[i]'}$ and $p^{[j]'}$ will be given below. We remark that this process is equivalent to exchanging a pair of temperatures T_m and T_n for the corresponding replicas i and j as follows:

$$\begin{cases} x_m^{[i]} \equiv (q^{[i]}, p^{[i]})_m &\rightarrow x_n^{[i]'} \equiv (q^{[i]}, p^{[i]'})_n, \\ x_n^{[j]} \equiv (q^{[j]}, p^{[j]})_n &\rightarrow x_m^{[j]'} \equiv (q^{[j]}, p^{[j]'})_m. \end{cases} \tag{39}$$

In the original implementation of the replica-exchange method [71–73], Monte Carlo algorithm was used, and only the coordinates q (and the potential energy function $E(q)$) had to be taken into account. In molecular dynamics algorithm, on the other hand, we also have to deal with the momenta p. We proposed the following momentum assignment in Eq. (38) (and in Eq. (39)) [80]:

$$\begin{cases} p^{[i]'} \equiv \sqrt{\dfrac{T_n}{T_m}}\, p^{[i]}, \\ p^{[j]'} \equiv \sqrt{\dfrac{T_m}{T_n}}\, p^{[j]}, \end{cases} \tag{40}$$

which we believe is the simplest and the most natural. This assignment means that we just rescale uniformly the velocities of all the atoms in the replicas by the square root of the ratio of the two temperatures so that the temperature condition in Eq. (4) may be satisfied.

In order for this exchange process to converge towards an equilibrium distribution, it is sufficient to impose the detailed balance condition on the transition probability $w(X \rightarrow X')$:

$$\frac{W_{\mathrm{REM}}(X)}{Z} w(X \rightarrow X') = \frac{W_{\mathrm{REM}}(X')}{Z} w(X' \rightarrow X), \tag{41}$$

where Z is the partition function of the entire system. From Eqs. (1), (2), (35), (40) and (41), we have

$$\begin{aligned} &\frac{w(X \rightarrow X')}{w(X' \rightarrow X)} \\ &\quad = \exp\{-\beta_m[K(p^{[j]'}) + E(q^{[j]})] - \beta_n[K(p^{[i]'}) + E(q^{[i]})] \\ &\qquad +\beta_m[K(p^{[i]}) + E(q^{[i]})] + \beta_n[K(p^{[j]}) + E(q^{[j]})]\}, \\ &\quad = \exp\left\{-\beta_m \frac{T_m}{T_n} K(p^{[j]}) - \beta_n \frac{T_n}{T_m} K(p^{[i]}) + \beta_m K(p^{[i]}) \right. \\ &\qquad +\beta_n K(p^{[j]}) - \beta_m[E(q^{[j]}) - E(q^{[i]})] \\ &\qquad \left. -\beta_n[E(q^{[i]}) - E(q^{[j]})]\right\}, \\ &\quad = \exp(-\Delta), \end{aligned} \tag{42}$$

where

$$\Delta = \beta_m(E(q^{[j]}) - E(q^{[i]})) - \beta_n(E(q^{[j]}) - E(q^{[i]})) \tag{43}$$

$$= (\beta_m - \beta_n)(E(q^{[j]}) - E(q^{[i]})), \tag{44}$$

and i, j, m, and n are related by the permutation functions in Eq. (33) before the exchange:

$$\begin{cases} i = f(m), \\ j = f(n). \end{cases} \tag{45}$$

This can be satisfied, for instance, by the usual Metropolis criterion [1]:

$$w(X \rightarrow X') \equiv w(x_m^{[i]} | x_n^{[j]}) = \min(1, \exp(-\Delta)), \tag{46}$$

where in the second expression (i.e., $w(x_m^{[i]}|x_n^{[j]})$) we explicitly wrote the pair of replicas (and temperatures) to be exchanged. Note that this is exactly the same criterion that was originally derived for Monte Carlo algorithm [71–73].

Without loss of generality we can again assume $T_1 < T_2 < \ldots < T_M$. A simulation of the replica-exchange method [71–73] is then realized by alternately performing the following two steps:

1. Each replica in canonical ensemble of the fixed temperature is simulated *simultaneously* and *independently* for a certain MC or MD steps.
2. A pair of replicas at neighboring temperatures, say $x_m^{[i]}$ and $x_{m+1}^{[j]}$, are exchanged with the probability $w(x_m^{[i]}|x_{m+1}^{[j]})$ in Eq. (46).

Note that in Step 2 we exchange only pairs of replicas corresponding to neighboring temperatures, because the acceptance ratio of the exchange process decreases exponentially with the difference of the two β's (see Eqs. (44) and (46)). Note also that whenever a replica exchange is accepted in Step 2, the permutation functions in Eq. (33) are updated.

The REM simulation is particularly suitable for parallel computers. Because one can minimize the amount of information exchanged among nodes, it is best to assign each replica to each node (exchanging pairs of temperature values among nodes is much faster than exchanging coordinates and momenta). This means that we keep track of the permutation function $m(i;t) = f^{-1}(i;t)$ in Eq. (33) as a function of MC or MD step t during the simulation. After parallel canonical MC or MD simulations for a certain steps (Step 1), $M/2$ pairs of replicas corresponding to neighboring temperatures are simulateneously exchanged (Step 2), and the pairing is alternated between the two possible choices, i.e. (T_1, T_2), (T_3, T_4), ... ·and (T_2, T_3), (T_4, T_5), ...

The major advantage of REM over other generalized-ensemble methods such as multicanonical algorithm [15,16] and simulated tempering [62,63] lies in the fact that the weight factor is a priori known (see Eq. (35)), while in the latter algorithms the determination of the weight factors can be very tedius and time-consuming. A random walk in "temperature space" is realized for each replica, which in turn induces a random walk in potential energy space. This alleviates the problem of getting trapped in states of energy local minima. In REM, however, the number of required replicas increases as the system size N increases (according to $\sqrt{N}$) [71]. This demands a lot of computer power for complex systems.

2.3. Replica-exchange multicanonical algorithm and replica-exchange simulated tempering

The *replica-exchange multicanonical algorithm* (REMUCA) [83,88] overcomes both the difficulties of MUCA (the multicanonical weight factor determination is non-trivial) and REM (a lot of replicas, or computation time, is required). In REMUCA we first perform a short REM simulation (with M replicas) to determine the multicanonical weight factor and then perform with this weight factor a regular multicanonical simulation with high statistics. The first step is accomplished by the multiple-histogram reweighting techniques [10,11]. Let $N_m(E)$ and n_m be respectively the potential-energy histogram and the total number of samples obtained at temperature $T_m (= 1/k_B\beta_m)$ of the REM run. The density of states $n(E)$ is then given by solving Eqs. (30) and (31) self-consistently by iteration.

Once the estimate of the density of states is obtained, the multicanonical weight factor can be directly determined from Eq. (10) (see also Eq. (11)). Actually, the density of states $n(E)$ and the multicanonical potential energy, $E_{mu}(E; T_0)$, thus determined are only reliable in the following range:

$$E_1 \le E \le E_M, \tag{47}$$

where

$$\begin{cases} E_1 = & \langle E \rangle_{T_1}, \\ E_M = & \langle E \rangle_{T_M}, \end{cases} \tag{48}$$

and T_1 and T_M are respectively the lowest and the highest temperatures used in the REM run. Outside this range we extrapolate the multicanonical potential energy linearly: [83]

$$\varepsilon_{mu}^{\{0\}}(E) \equiv \begin{cases} \left.\dfrac{\partial E_{mu}(E;T_0)}{\partial E}\right|_{E=E_1} (E - E_1) + E_{mu}(E_1;T_0), & \text{for } E < E_1, \\ E_{mu}(E;T_0), & \text{for } E_1 \le E \le E_M, \\ \left.\dfrac{\partial E_{mu}(E;T_0)}{\partial E}\right|_{E=E_M} (E - E_M) + E_{mu}(E_M;T_0), & \text{for } E > E_M. \end{cases} \tag{49}$$

The multicanonical MC and MD runs are then performed respectively with the Metropolis criterion of Eq. (12) and with the modified Newton equation in Eq. (14), in which $\varepsilon_{mu}^{\{0\}}(E)$ in Eq. (49) is substituted into $E_{mu}(E; T_0)$. We expect to obtain a flat potential energy distribution in the range of Eq. (47). Finally, the results are analyzed by the single-histogram reweighting techniques as described in Eq. (20) (and Eq. (18)).

Some remarks are now in order. From Eqs. (11), (16), (17), and (48), Eq. (49) becomes

$$\varepsilon_{mu}^{\{0\}}(E) = \begin{cases} \dfrac{T_0}{T_1}(E - E_1) + T_0 S(E_1) = \dfrac{T_0}{T_1} E + \text{constant}, & \text{for } E < E_1 \equiv \langle E \rangle_{T_1}, \\ T_0 S(E), & \text{for } E_1 \le E \le E_M, \\ \dfrac{T_0}{T_M}(E - E_M) + T_0 S(T_M) = \dfrac{T_0}{T_M} E + \text{constant}, & \text{for } E > E_M \equiv \langle E \rangle_{T_M}. \end{cases} \tag{50}$$

The Newton equation in Eq. (14) is then written as (see Eqs. (15)–(17))

$$\dot{\boldsymbol{p}}_k = \begin{cases} \dfrac{T_0}{T_1}\boldsymbol{f}_k, & \text{for } E < E_1, \\ \dfrac{T_0}{T(E)}\boldsymbol{f}_k, & \text{for } E_1 \le E \le E_M, \\ \dfrac{T_0}{T_M}\boldsymbol{f}_k, & \text{for } E > E_M. \end{cases} \tag{51}$$

Because only the product of inverse temperature β and potential energy E enters in the Boltzmann factor (see Eq. (5)), a rescaling of the potential energy (or force) by a constant, say α, can be considered as the rescaling of the temperature by $1/\alpha$ [38,95]. Hence, our choice of $\varepsilon_{\mathrm{mu}}^{\{0\}}(E)$ in Eq. (49) results in a canonical simulation at $T = T_1$ for $E < E_1$, a multicanonical simulation for $E_1 \le E \le E_M$, and a canonical simulation at $T = T_M$ for $E > E_M$. Note also that the above arguments are independent of the value of T_0, and we will get the same results, regardless of its value.

For Monte Carlo method, the above statement follows directly from the following equation. Namely, our choice of the multicanonical potential energy in Eq. (49) gives (by substituting Eq. (50) into Eq. (10))

$$\begin{aligned} W_{\mathrm{mu}}(E) &= \exp[-\beta_0 \varepsilon_{\mathrm{mu}}^{\{0\}}(E)] \\ &= \begin{cases} \exp(-\beta_1 E + \text{constant}), & \text{for } E < E_1, \\ \dfrac{1}{n(E)}, & \text{for } E_1 \le E \le E_M, \\ \exp(-\beta_M E + \text{constant}), & \text{for } E > E_M. \end{cases} \end{aligned} \tag{52}$$

We now present another effective method of the multicanonical weight factor determination [7], which is closely related to REMUCA. We first perform a short REM simulation as in REMUCA and calculate $\langle E\rangle_T$ as a function of T by the multiple-histogram reweighting techniques (see Eqs. (30) and (31)). Let us recall the Newton equation of Eq. (15) and the thermodynamic relation of Eqs. (16) and (17). The effective temperature $T(E)$, or the derivative $(\partial E_{\mathrm{mu}}(E; T_0)/\partial E)$, can be numerically obtained as the inverse function of Eq. (17), where the average $\langle E\rangle_{T(E)}$ has been obtained from the results of the REM simulation by the multiple-histogram reweighting techniques. Given its derivative, the multicanonical potential energy can then be obtained by numerical integration (see Eqs. (11) and (16)): [7]

$$E_{\mathrm{mu}}(E; T_0) = T_0 \int_{E_1}^{E} \frac{\partial S(E)}{\partial E} \mathrm{d}E = T_0 \int_{E_1}^{E} \frac{\mathrm{d}E}{T(E)}. \tag{53}$$

We remark that the same equation was used to obtain the multicanonical weight factor in Ref. [68], where $\langle E\rangle_T$ was estimated by simulated annealing instead of REM. Essentially the same formulation was also recently used in Ref. [58] to obtain the multicanonical potential energy, where $\langle E\rangle_T$ was calculated by conventional canonical simulations.

We finally present the new method which we refer to as the *replica-exchange simulated tempering* (REST) [84]. In this method, just as in REMUCA, we first perform a short REM simulation (with M replicas) to determine the simulated tempering weight factor and then perform with this weight factor a regular ST simulation with high statistics.

The first step is accomplished by the multiple-histogram reweighting techniques [10,11], which give the dimensionless Helmholtz free energy f_m (see Eqs. (30) and (31)).

Once the estimate of the dimensionless Helmholtz free energy f_m are obtained, the simulated tempering weight factor can be directly determined by using Eq. (25) where we set $a_m = f_m$ (compare Eq. (26) with Eq. (31)). A long simulated tempering run is then performed with this weight factor. Let $N_m(E)$ and n_m be respectively the potential-energy histogram and the total number of samples obtained at temperature $T_m(= 1/k_{\mathrm{B}}\beta_m)$ from this simulated tempering run. The multiple-histogram reweighting techniques of Eqs. (30) and (31) can be used again to obtain the best estimate of the density of states $n(E)$. The expectation value of a physical quantity A at any temperature $T(= 1/k_{\mathrm{B}}\beta)$ is then calculated from Eq. (18).

The formulations of REMUCA and REST are simple and straightforward, but the numerical improvement is great, because the weight factor determination for MUCA and ST becomes very difficult by the usual iterative processes for complex systems.

2.4. Multicanonical replica-exchange method and simulated tempering replica-exchange method

In the previous subsection we presented REMUCA, which uses a short REM run for the determination of the multicanonical weight factor. Here, we present two modifications of REM and refer to the new methods as *multicanonical replica-exchange method* (MUCAREM) [83,88] and *simulated tempering replica-exchange method* (STREM) [109]. In MUCAREM the production run is a REM simulation with a few replicas not in the canonical ensemble but in the multicanonical ensemble, i.e., different replicas perform MUCA simulations with different energy ranges. Likewise in STREM the production run is a REM simulation with a few replicas that performs ST simulations with different temperature ranges. While MUCA and ST simulations are usually based on local updates, a replica-exchange process can be considered to be a global update, and global updates enhance the sampling further.

We first describe MUCAREM. Let $\mathcal{M}$ be the number of replicas. Here, each replica is in one-to-one correspondence not with temperature but with multicanonical weight factors of different energy range. Note that because multicanonical simulations cover much wider energy ranges than regular canonical simulations, the number of required replicas for the production run of MUCAREM is much less than that for the regular REM ($\mathcal{M} \ll M$).

The weight factor for this generalized ensemble is now given by (see Eq. (35))

$$W_{\mathrm{MUCAREM}}(X) = \prod_{i=1}^{\mu} W_{\mathrm{mu}}^{\{m(i)\}}(E(x_{m(i)}^{[i]})) = \prod_{m=1}^{\mu} W_{\mathrm{mu}}^{\{m\}}(E(x_m^{[i(m)]})), \tag{54}$$

where we prepare the multicanonical weight factor (and the density of states) separately for m regions (see Eq. (10)):

$$W_{\mathrm{mu}}^{\{m\}}(E(x_m^{[i]})) = \exp[-\beta_m \varepsilon_{\mathrm{mu}}^{\{m\}}(E(x_m^{[i]})) \equiv \frac{1}{n^{\{m\}}(E(x_m^{[i]}))}. \tag{55}$$

Here, we have introduced $\mathcal{M}$ arbitrary reference temperatures $T_m = 1/k_{\mathrm{B}}\beta_m$ ($m = 1, \ldots, \mathcal{M}$), but the final results will be independent of the values of T_m, as one can see from the second equality in Eq. (55) (these arbitrary temperatures are necessary only for MD simulations).

Each multicanonical weight factor $W_{\mathrm{mu}}^{\{m\}}(E)$, or the density of states $n^{\{m\}}(E)$, is defined as follows. For each m ($m = 1, \ldots, \mathcal{M}$), we assign a pair of temperatures ($T_{\mathrm{L}}^{\{m\}}, T_{\mathrm{H}}^{\{m\}}$). Here, we assume that $T_{\mathrm{L}}^{\{m\}} < T_{\mathrm{H}}^{\{m\}}$ and arrange the temperatures so that the neighboring regions covered by the pairs have sufficient overlaps. Without loss of generality we can assume $T_{\mathrm{L}}^{\{1\}} < \ldots < T_{\mathrm{L}}^{\{\mu\}}$ and $T_{\mathrm{H}}^{\{1\}} < \ldots < T_{\mathrm{H}}^{\{\mu\}}$. We define the following quantities:

$$\begin{cases} E_{\mathrm{L}}^{\{m\}} = \langle E \rangle_{T_{\mathrm{L}}^{\{m\}}}, \\ E_{\mathrm{H}}^{\{m\}} = \langle E \rangle_{T_{\mathrm{H}}^{\{m\}}}, \quad (m = 1, \ldots, \mu). \end{cases} \tag{56}$$

Suppose that the multicanonical weight factor $W_{\mathrm{mu}}(E)$ (or equivalently, the multicanonical potential energy $E_{\mathrm{mu}}(E; T_0)$ in Eq. (11)) has been obtained as in REMUCA or by any other methods in the entire energy range of interest ($E_{\mathrm{L}}^{\{1\}} < E < E_{\mathrm{H}}^{\{\mu\}}$). We then have for each m ($m = 1, \ldots, \mu$) the following multicanonical potential energies (see Eq. (49)): [83]

$$\varepsilon_{\mathrm{mu}}^{\{m\}}(E) = \begin{cases} \left.\dfrac{\partial E_{\mathrm{mu}}(E; T_m)}{\partial E}\right|_{E=E_{\mathrm{L}}^{\{m\}}} (E - E_{\mathrm{L}}^{\{m\}}) + E_{\mathrm{mu}}(E_{\mathrm{L}}^{\{m\}}; T_m), & \text{for } E < E_{\mathrm{L}}^{\{m\}}, \\ E_{\mathrm{mu}}(E; T_m), & \text{for } E_{\mathrm{L}}^{\{m\}} \le E \le E_{\mathrm{H}}^{\{m\}}, \\ \left.\dfrac{\partial E_{\mathrm{mu}}(E; T_m)}{\partial E}\right|_{E=E_{\mathrm{H}}^{\{m\}}} (E - E_{\mathrm{H}}^{\{m\}}) + E_{\mathrm{mu}}(E_{\mathrm{H}}^{\{m\}}; T_m), & \text{for } E > E_{\mathrm{H}}^{\{m\}}. \end{cases} \tag{57}$$

Finally, a MUCAREM simulation is realized by alternately performing the following two steps.

1. Each replica of the fixed multicanonical ensemble is simulated *simultaneously* and *independently* for a certain MC or MD steps.
2. A pair of replicas, say i and j, which are in neighboring multicanonical ensembles, say m-th and $(m+1)$-th, respectively, are exchanged: $X = \{\ldots, x_m^{[i]}, \ldots, x_{m+1}^{[j]}, \ldots\} \to X' = \{\ldots, x_m^{[j]}, \ldots, x_{m+1}^{[i]}, \ldots\}$. The transition probability of this replica exchange is given by the Metropolis criterion:

$$w(X \to X') = \min(1, \exp(-\Delta)), \tag{58}$$

where we now have (see Eq. (43)) [83]

$$\Delta = \beta_m\{\varepsilon_{\mathrm{mu}}^{\{m\}}(E(q^{[j]})) - \varepsilon_{\mathrm{mu}}^{\{m\}}(E(q^{[i]}))\} - \beta_{m+1}\{\varepsilon_{\mathrm{mu}}^{\{m+1\}}(E(q^{[j]})) - \varepsilon_{\mathrm{mu}}^{\{m+1\}}(E(q^{[i]}))\}. \tag{59}$$

Here, $E(q^{[i]})$ and $E(q^{[j]})$ are the potential energy of the i-th replica and the j-th replica, respectively. Note that in Eq. (59) we need to newly evaluate the multicanonical potential energy, $\varepsilon_{\mathrm{mu}}^{\{m\}}(E(q^{[j]}))$ and $\mathcal{E}_{\mathrm{mu}}^{\{m+1\}}(E(q^{[i]}))$, because $\varepsilon_{\mathrm{mu}}^{\{m\}}(E)$ and $\varepsilon_{\mathrm{mu}}^{\{n\}}(E)$ are, in general, different functions for $m \neq n$.

In this algorithm, the m-th multicanonical ensemble actually results in a canonical simulation at $T = T_{\mathrm{L}}^{\{m\}}$ for $E < E_{\mathrm{L}}^{\{m\}}$, a multicanonical simulation for $E_{\mathrm{L}}^{\{m\}} \le E \le E_{\mathrm{H}}^{\{m\}}$, and a canonical simulation at $T = T_{\mathrm{H}}^{\{m\}}$ for $E > E_{\mathrm{H}}^{\{m\}}$, while the replica-exchange process samples states of the whole energy range ($E_{\mathrm{L}}^{\{1\}} \le E \le E_{\mathrm{H}}^{\{\mathcal{M}\}}$).

For obtaining the canonical distributions at any intermediate temperature T, the multiple-histogram reweighting techniques [10,11] are again used. Let $N_m(E)$ and n_m be respectively the potential-energy histogram and the total number of samples obtained with the multicanonical weight factor $W_{\mathrm{mu}}^{\{m\}}(E)$ ($m = 1, \ldots, M$). The expectation value of a physical quantity A at any temperature T ($= 1/k_{\mathrm{B}}\beta$) is then obtained from Eq. (18), where the best estimate of the density of states is obtained by solving the WHAM equations, which now read [83]

$$n(E) = \frac{\sum_{m=1}^{\mu} g_m^{-1} N_m(E)}{\sum_{m=1}^{\mu} g_m^{-1} n_m \exp(f_m) W_{\mathrm{mu}}^{\{m\}}(E)} = \frac{\sum_{m=1}^{\mu} g_m^{-1} N_m(E)}{\sum_{m=1}^{\mu} g_m^{-1} n_m \exp(f_m - \beta_m \varepsilon_{\mathrm{mu}}^{\{m\}}(E))}, \tag{60}$$

and for each m ($= 1, \ldots, \mathcal{M}$)

$$\exp(-f_m) = \sum_E n(E) W_{\mathrm{mu}}^{\{m\}}(E) = \sum_E n(E) \exp(-\beta_m \varepsilon_{\mathrm{mu}}^{\{m\}}(E)). \tag{61}$$

Note that $W_{\mathrm{mu}}^{\{m\}}(E)$ is used instead of the Boltzmann factor $\exp(-\beta_m E)$ in Eqs. (30) and (31).

Moreover, ensemble averages of any physical quantity A (including those that cannot be expressed as functions of potential energy) at any temperature T ($= 1/k_B\beta$) can now be obtained from the "trajectory"of configurations of the production run. Namely, we first obtain f_m ($m = 1, \ldots, \mathcal{M}$) by solving Eqs. (60) and (61) self-consistently, and then we have [88]

$$\langle A\rangle_T = \frac{\sum_{m=1}^{\mu}\sum_{k=1}^{n_m} A(x_m(k)) \left(g_m^{-1} / \left(\sum_{\ell=1}^{\mu} g_\ell^{-1} n_\ell \exp(f_\ell) W_{\mathrm{mu}}^{\{m\}}(E(x_m(k))) \right) \right) \exp[-\beta E(x_m(k))]}{\sum_{m=1}^{\mu}\sum_{k=1}^{n_m} \left(g_m^{-1} / \left(\sum_{\ell=1}^{\mu} g_\ell^{-1} n_\ell \exp(f_\ell) W_{\mathrm{mu}}^{\{m\}}(E(x_m(k))) \right) \right) \exp[-\beta E(x_m(k))]}, \tag{62}$$

where the trajectories $x_m(k)$ ($k = 1, \ldots, n_m$) are taken from each multicanonical simulation with the multicanonical weight factor $W_{\mathrm{mu}}^{\{m\}}(E)$ ($m = 1, \ldots, \mathcal{M}$) separately.

As seen above, both REMUCA and MUCAREM can be used to obtain the multicanonical weight factor, or the density of states, for the entire potential energy range of interest. For complex systems, however, a single REMUCA or MUCAREM simulation is often insufficient. In such cases we can iterate MUCA (in REMUCA) and/or MUCAREM simulations in which the estimate of the multicanonical weight factor is updated by the single- and/or multiple-histogram reweighting techniques, respectively.

To be more specific, this iterative process can be summarized as follows [88]. The REMUCA production run corresponds to a MUCA simulation with the weight factor $W_{mu}(E)$. The new estimate of the density of states can be obtained by the single-histogram reweighting techniques of Eq. (20). On the other hand, from the MUCAREM production run, the improved density of states can be obtained by the multiple-histogram reweighting techniques of Eqs. (60) and (61).

The improved density of states thus obtained leads to a new multicanonical weight factor (see Eq. (10)). The next iteration can be either a MUCA production run (as in REMUCA) or MUCAREM production run. The results of this production run may yield an optimal multicanonical weight factor that yields a sufficiently flat energy distribution for the entire energy range of interest. If not, we can repeat the above process by obtaining the third estimate of the multicanonical weight factor either by a MUCA production run (as in REMUCA) or by a MUCAREM production run, and so on.

We remark that as the estimate of the multicanonical weight factor becomes more accurate, one is required to have a less number of replicas for a successful MUCAREM simulation, because each replica will have a flat energy distribution for a wider energy range. Hence, for a large, complex system, it is often more efficient to first try MUCAREM and iteratively reduce the number of replicas so that eventually one needs only one or a few replicas (instead of trying REMUCA directly from the beginning and iterating MUCA simulations).

We now describe the *simulated tempering replica-exchange method* (STREM) [109]. Suppose that the simulated tempering weight factor $W_{\mathrm{ST}}(E; T_n)$ (or equivalently, the dimensionless Helmholtz free energy a_n in Eq. (25)) has been obtained as in REST or by any other methods in the entire temperature range of interest ($T_1 \le T_n \le T_M$). We devide the overlapping temperature ranges into $\mathcal{M}$ regions ($\mathcal{M} \ll M$). Suppose each temperature range m has $\mathcal{N}_m$ temperatures: $T_k^{\{m\}}$ ($k = 1, \ldots, \mathcal{N}_m$) for $m = 1, \ldots, \mathcal{M}$. We assign each temperature range to a replica; each replica i is in one-to-one correspondence with a different temperature range m of ST run, where $T_1^{\{m\}} \le T_k^{\{m\}} \le T_{\mathcal{N}_m}^{\{m\}}$ ($k = 1, \cdots, \mathcal{N}_m$). We then introduce the replica-exchange process between neighboring temperature ranges. This works when we allow sufficient overlaps between the temperature regions.

A STREM simulation is then realized by alternately performing the following two steps [109].

1. Each replica performs a ST simulation within the fixed temperature range *simultaneously* and *independently* for a certain MC or MD steps.
2. A pair of replicas, say i and j, which are at, say $T = T_k^{\{m\}}$ and $T = T_\ell^{\{m+1\}}$, in neighboring temperature ranges, say m-th and $(m+1)$-th, respectively, are exchanged:
   $X = \{\ldots, x_k^{[i]}, \ldots, x_\ell^{[j]}, \ldots\} \to X' = \{\ldots, x_k^{[j]}, \ldots, x_\ell^{[i]}, \ldots\}$. The transition probability of this replica exchange is given by the Metropolis criterion:

$$w(X \to X') = \min(1, \exp(-\Delta)), \tag{63}$$

where

$$\Delta \equiv (\beta_k^{\{m\}} - \beta_\ell^{\{m+1\}})(E(q^{[j]}) - E(q^{[i]})). \tag{64}$$

While in MUCAREM each replica performs a random walk in multicanonical ensemble of finite energy range, in STREM each replica performs a random walk by simulated tempering of finite temperature range. These "local" random walks are made "global" to cover the entire energy range of interest by the replica-exchange process.

2.5. Multidimensional replica-exchange method

We now present our multidimensional extension of REM, which we refer to as *multidimensional replica-exchange method* (MREM) [82]. The crucial observation that led to the new algorithm is: As long as we have *M non-interacting*

replicas of the original system, the Hamiltonian $H(q, p)$ of the system does not have to be identical among the replicas and it can depend on a parameter with different parameter values for different replicas.

Namely, we can write the Hamiltonian for the i-th replica at temperature T_m as

$$H_m(q^{[i]}, p^{[i]}) = K(p^{[i]}) + E_{\lambda_m}(q^{[i]}), \tag{65}$$

where the potential energy E_{λ_m} depends on a parameter λ_m and can be written as

$$E_{\lambda_m}(q^{[i]}) = E_0(q^{[i]}) + \lambda_m V(q^{[i]}). \tag{66}$$

This expression for the potential energy is often used in simulations. For instance, in umbrella sampling [24], $E_0(q)$ and $V(q)$ can be respectively taken as the original potential energy and the "biasing" potential energy with the coupling parameter λ_m. In simulations of spin systems, on the other hand, $E_0(q)$ and $V(q)$ (here, q stands for spins) can be respectively considered as the zero-field term and the magnetization term coupled with the external field λ_m.

While replica i and temperature T_m are in one-to-one correspondence in the original REM, replica i and "parameter set" $\Lambda_m \equiv (T_m, \lambda_m)$ are in one-to-one correspondence in the new algorithm. Hence, the present algorithm can be considered as a multidimensional extension of the original replica-exchange method where the "parameter space" is one-dimensional (i.e., $\Lambda_m = T_m$). Because the replicas are non-interacting, the weight factor for the state X in this new generalized ensemble is again given by the product of Boltzmann factors for each replica (see Eq. (35)):

$$\begin{aligned} W_{\mathrm{MREM}}(X) &= \exp\left\{-\sum_{i=1}^{M}\beta_{m(i)}H_{m(i)}(q^{[i]}, p^{[i]})\right\} \\ &= \exp\left\{-\sum_{m=1}^{M}\beta_m H_m(q^{[i(m)]}, p^{[i(m)]})\right\} \end{aligned} \tag{67}$$

where $i(m)$ and $m(i)$ are the permutation functions in Eq. (33). Then the same derivation that led to the original replica-exchange criterion follows, and the transition probability of replica exchange is given by Eq. (46), where we now have (see Eq. (43)) [82]

$$\begin{aligned} \Delta = {} & \beta_m(E_{\lambda_m}(q^{[j]}) - E_{\lambda_m}(q^{[i]})) \\ & -\beta_n(E_{\lambda_n}(q^{[j]}) - E_{\lambda_n}(q^{[i]})). \end{aligned} \tag{68}$$

Here, E_{λ_m} and E_{λ_n} are the total potential energies (see Eq. (66)). Note that we need to newly evaluate the potential energy for exchanged coordinates, $E_{\lambda_m}(q^{[j]})$ and $E_{\lambda_n}(q^{[i]})$, because E_{λ_m} and E_{λ_n} are in general different functions.

For obtaining the canonical distributions, the multiple-histogram reweighting techniques [10,11] are particularly suitable. Suppose we have made a single run of the present replica-exchange simulation with M replicas that correspond to M different parameter sets $\Lambda_m \equiv (T_m, \lambda_m)$ $(m = 1, \ldots, M)$. Let $N_m(E_0, V)$ and n_m be respectively the potential-energy histogram and the total number of samples obtained for the m-th parameter set Λ_m. The WHAM equations that yield the canonical probability distribution $P_{T,\lambda}(E_0, V) = n(E_0, V)\exp(-\beta E_\lambda)$ with any potential-energy parameter value λ at any temperature $T = 1/k_{\mathrm{B}}\beta$ are then given by [82]

$$n(E_0, V) = \frac{\sum_{m=1}^{M} g_m^{-1} N_m(E_0, V)}{\sum_{m=1}^{M} g_m^{-1} n_m \exp(f_m - \beta_m E_{\lambda_m})}, \tag{69}$$

and for each m $(= 1, \ldots, M)$

$$\exp(-f_m) = \sum_{E_0, V} n(E_0, V)\exp(-\beta_m E_{\lambda_m}). \tag{70}$$

Here, $n(E_0, V)$ is the generalized density of states. Note that $n(E_0, V)$ is independent of the parameter sets $\Lambda_m \equiv (T_m, \lambda_m)$ $(m = 1, \ldots, M)$. The density of states $n(E_0, V)$ and the "dimensionless" Helmholtz free energy f_m in Eqs. (69) and (70) are solved self-consistently by iteration.

Incidentally, these formulations of MREM give multidimensional extensions of REMUCA [83,88] and REST [84]. In the former, we obtain uniform distributions both in E_0 and V, whereas in the latter, the parameter sets Λ_m become dynamical variables and a uniform distribution in those parameters will be obtained. Namely, after a short MREM simulation, we can use the multiple-histogram reweighting techniques of Eqs. (69) and (70) to obtain $n(E_0, V)$ and f_m. Hence, we can determine the multidimensional multicanonical weight factor $W_{\mathrm{mu}}(E_0, V)$ and the multidimensional simulated tempering weight factor $W_{\mathrm{ST}}(E_0, V; \Lambda_m)$. The former is given by

$$W_{\mathrm{mu}}(E_0, V) = \frac{1}{n(E_0, V)}, \tag{71}$$

and the latter is given by (see Eq. (25))

$$W_{\mathrm{ST}}(E_0, V; \Lambda_m) = \exp(-\beta_m E_{\lambda_m} + f_m). \tag{72}$$

We can use MREM for free energy calculations. We first describe the free energy perturbation case. The method is referred to as *replica-exchange free energy perturbation* (REFEP) [82]. The potential energy is given by

$$E_\lambda(q) = E_I(q) + \lambda(E_F(q) - E_I(q)), \tag{73}$$

where E_I and E_F are the potential energy for a "wild-type" molecule and a "mutated" molecule, respectively. Note that this equation has the same form as Eq. (66).

Our replica-exchange simulation is performed for M replicas with M different values of the parameters $\Lambda_m = (T_m, \lambda_m)$. Since $E_{\lambda=0}(q) = E_I(q)$ and $E_{\lambda=1}(q) = E_F(q)$, we should choose enough λ_m values distributed in the range between 0 and 1 so that we may have sufficient acceptance of replica exchange. From the simulation, M histograms

$N_m(E_I, E_F - E_I)$, or equivalently $N_m(E_I, E_F)$, are obtained. The Helmholtz free energy difference of "mutation" at temperature T ($= 1/k_B\beta$), $\Delta F \equiv F_{\lambda=1} - F_{\lambda=0}$, can then be calculated from

$$\exp(-\beta\,\Delta F) = \frac{Z_{T,\lambda=1}}{Z_{T,\lambda=0}} = \frac{\sum_{E_I,E_F} P_{T,\lambda=1}(E_I, E_F)}{\sum_{E_I,E_F} P_{T,\lambda=0}(E_I, E_F)}, \tag{74}$$

where $P_{T,\lambda}(E_I, E_F) = n(E_I, E_F)\exp(-\beta E_\lambda)$ are obtained from the WHAM equations of Eqs. (69) and (70).

We now describe another free energy calculations based on MREM applied to umbrella sampling, which we refer to as *replica-exchange umbrella sampling* (REUS) [82]. The potential energy is a generalization of Eq. (66) and is given by

$$E_{\boldsymbol{\lambda}}(q) = E_0(q) + \sum_{\ell=1}^{L} \lambda^{(\ell)} V_\ell(q), \tag{75}$$

where $E_0(q)$ is the original unbiased potential, $V_\ell(q)(\ell = 1, \ldots, L)$ are the biasing (umbrella) potentials, and $\lambda^{(\ell)}$ are the corresponding coupling constants ($\boldsymbol{\lambda} = (\lambda^{(1)}, \ldots, \lambda^{(L)})$).

Introducing a "reaction coordinate" ξ, the umbrella potentials are usually written as harmonic restraints:

$$V_\ell(q) = k_\ell(\xi(q) - d_\ell)^2, (\ell = 1, \ldots, L), \tag{76}$$

where d_ℓ are the midpoints and k_ℓ are the strengths of the restraining potentials. We prepare M replicas with M different values of the parameters $\boldsymbol{\Lambda}_m = (T_m, \boldsymbol{\lambda}_m)$, and the replica-exchange simulation is performed. Since the umbrella potentials $V_\ell(q)$in Eq. (76) are all functions of the reaction coordinate ξ only, we can take the histogram $N_m(E_0, \xi)$ instead of $N_m(E_0, V_1, \ldots, V_L)$. The WHAM equations of Eqs. (69) and (70) can then be written as [82]

$$n(E_0, \xi) = \frac{\sum_{m=1}^{M} g_m^{-1} N_m(E_0, \xi)}{\sum_{m=1}^{M} g_m^{-1} n_m \exp(f_m - \beta_m E_{\boldsymbol{\lambda}_m})} \tag{77}$$

and for each m ($= 1, \ldots, M$)

$$\exp(-f_m) = \sum_{E_0,\xi} n(E_0,\xi)\exp(-\beta_m E_{\boldsymbol{\lambda}_m}). \tag{78}$$

The expectation value of a physical quantity A with any potential-energy parameter value $\boldsymbol{\lambda}$ at any temperature T ($= 1/k_B\beta$) is now given by

$$\langle A\rangle_{T,\boldsymbol{\lambda}} = \frac{\sum_{E_0,\xi} A(E_0,\xi) P_{T,\boldsymbol{\lambda}}(E_0,\xi)}{\sum_{E_0,\xi} P_{T,\boldsymbol{\lambda}}(E_0,\xi)} \tag{79}$$

where $P_{T,\boldsymbol{\lambda}}(E_0,\xi) = n(E_0,\xi)\exp(-\beta E_{\boldsymbol{\lambda}})$ is obtained from the WHAM equations of Eqs. (77) and (78).

The potential of mean force (PMF), or free energy as a function of the reaction coordinate, of the original, unbiased system at temperature T is given by

$$\mathcal{W}_{T,\boldsymbol{\lambda}=\{0\}}(\xi) = -k_B T \ln\left[\sum_{E_0} P_{T,\boldsymbol{\lambda}=\{0\}}(E_0,\xi)\right], \tag{80}$$

where $\{0\} = (0, \ldots, 0)$.

We now present two examples of realization of REUS. In the first example, we use only one temperature, T, and L umbrella potentials. We prepare replicas so that the potential energy for each replica includes exactly one umbrella potential (here, we have $M = L$).

Namely, in Eq. (75) for $\boldsymbol{\lambda} = \boldsymbol{\lambda}_m$ we set

$$\lambda_m^{(\ell)} = \delta_{\ell,m}, \tag{81}$$

where $\delta_{k,l}$ is Kronecker's delta function, and we have

$$E_{\boldsymbol{\lambda}_m}(q^{[i]}) = E_0(q^{[i]}) + V_m(q^{[i]}). \tag{82}$$

We exchange replicas corresponding to "neighboring" umbrella potentials, V_m and V_{m+1}.

The acceptance criterion for replica exchange is given by Eq. (46), where Eq. (68) now reads (with the fixed inverse temperature $\beta = 1/k_B T$) [82]

$$\Delta = \beta(V_m(q^{[j]}) - V_m(q^{[i]}) - V_{m+1}(q^{[j]}) + V_{m+1}(q^{[i]})), \tag{83}$$

where replica i and j respectively have umbrella potentials V_m and V_{m+1} before the exchange.

In the second example, we prepare N_T temperatures and L umbrella potentials, which makes the total number of replicas $M = N_T \times L$. We can introduce the following relabeling for the parameters that characterize the replicas:

$$\underset{(m=1,\ldots,M)}{\boldsymbol{\Lambda}_m = (T_m, \boldsymbol{\lambda}_m)} \rightarrow \underset{(I=1,\ldots,N_T, J=1,\ldots,L)}{\boldsymbol{\Lambda}_{I,J} = (T_I, \boldsymbol{\lambda}_J)}. \tag{84}$$

The potential energy is given by Eq. (82) with the replacement: $m \rightarrow J$. We perform the following replica-exchange processes alternately:

1. Exchange pairs of replicas corresponding to neighboring temperatures, T_I and T_{I+1} (i.e., exchange replicas i and j that respectively correspond to parameters $\boldsymbol{\Lambda}_{I,J}$ and $\boldsymbol{\Lambda}_{I+1,J}$). (We refer to this process as T-exchange.)
2. Exchange pairs of replicas corresponding to "neighboring" umbrella potentials, V_J and V_{J+1} (i.e., exchange replicas i and j that respectively correspond to parameters $\boldsymbol{\Lambda}_{I,J}$ and $\boldsymbol{\Lambda}_{I,J+1}$). (We refer to this process as λ-exchange.)

The acceptance criterion for these replica exchanges is given by Eq. (46), where Eq. (68) now reads [82]

$$\Delta = (\beta_I - \beta_{I+1})(E_0(q^{[j]}) + V_J(q^{[j]}) - E_0(q^{[i]}) - V_J(q^{[i]})), \tag{85}$$

for T-exchange, and

$$\Delta = \beta_I(V_J(q^{[j]}) - V_J(q^{[i]}) - V_{J+1}(q^{[j]}) + V_{J+1}(q^{[i]})), \tag{86}$$

for λ-exchange. By this procedure, the random walk in the reaction coordinate space as well as in the temperature space can be realized.

3. Conclusions

In this article we have reviewed uses of generalized-ensemble algorithms for both Monte Carlo simulations and molecular dynamics simulations. A simulation in generalized ensemble realizes a random walk in potential energy space, alleviating the multiple-minima problem that is a common difficulty in simulations of complex systems with many degrees of freedom.

Detailed formulations of the three well-known generalized-ensemble algorithms, namely, multicanonical algorithm, simulated tempering, and replica-exchange method, were given.

We then introduced five new generalized-ensemble algorithms that combine the merits of the above three methods. We refer to these methods as replica-exchange multicanonical algorithm, replica-exchange simulated tempering, multicanonical replica-exchange method, simulated tempering replica-exchange method, and multidimensional replica-exchange method, the last of which also led to replica-exchange free energy perturbation and replica-exchange umbrella sampling.

The question is then which method is the most recommended. We have recently studied the effectiveness of MUCA, REM, REMUCA, and MUCAREM in the protein folding problem [88]. Our criterion for the effectiveness was how many times the random walk cycles between the high-energy region and low-energy region are realized within a fixed number of total MC (or MD) steps. We found that once the optimal MUCA weight factor is obtained, MUCA (and REMUCA) is the most effective (i.e., has the most number of random walk cycles), and REM is the least [88]. We also found that once the optimal ST weight factor is obtained, ST (and REST) has more random walk cycles than REM [84,109]. Moreover, we compared the efficiency of Berg's recursion [69], Wang–Landau method [21,22], and REMUCA/MUCAREM as methods for the multicanonical weight factor determination in two-dimensional 10-state Potts model and found that the three methods are about equal in efficiency [118–120].

Hence, the answer to the above question will depend on how much time one is willing to (or forced to) spend in order to determine the MUCA or ST weight factors. Given a problem, the first choice is REM because of its simplicity (no weight factor determination is required). If REM turns out to be insufficient or too much time-consuming (like the case with first-order phase transitions), then other more powerful algorithms such as those presented in the present article are recommended.

Acknowledgements

The author would like to thank his co-workers for useful discussions. In particular, he is grateful to Drs. B.A. Berg, U.H.E. Hansmann, A. Mitsutake, and Y. Sugita for collaborations that led to the new generalized-ensemble algorithms described in the present article. This work was supported, in part, by grants from the Research for the Future Program of the Japan Society for the Promotion of Science (JSPS-RFTF98P01101) and the NAREGI Nanoscience Project, Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

[1] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, J. Chem. Phys. 21 (1953) 1087–1092.
[2] M. Vásquez, G. Némethy, H.A. Scheraga, Chem. Rev. 94 (1994) 2183–2239.
[3] B.J. Berne, J.E. Straub, Curr. Opin. Struct. Biol. 7 (1997) 181–189.
[4] U.H.E. Hansmann, Y. Okamoto, Curr. Opin. Struct. Biol. 9 (1999) 177–183.
[5] U.H.E. Hansmann, Y. Okamoto, in: D. Stauffer (Ed.), Annual Reviews of Computational Physics VI, World Scientific, Singapore, 1999, pp. 129–157.
[6] A. Mitsutake, Y. Sugita, Y. Okamoto, Biopolymers (Peptide Science) 60 (2001) 96–123.
[7] Y. Sugita, Y. Okamoto, in: T. Schlick, H.H. Gan (Eds.), Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin, 2002, pp. 304–332; cond-mat/0102296.
[8] B.A. Berg, Comp. Phys. Commun. 147 (2002) 52–57.
[9] A.M. Ferrenberg, R.H. Swendsen, Phys. Rev. Lett. 61 (1988) 2635–2638;
A.M. Ferrenberg, R.H. Swendsen, Phys. Rev. Lett. 63 (1989) 1658.
[10] A.M. Ferrenberg, R.H. Swendsen, Phys. Rev. Lett. 63 (1989) 1195–1198.
[11] S. Kumar, D. Bouzida, R.H. Swendsen, P.A. Kollman, J.M. Rosenberg, J. Comput. Chem. 13 (1992) 1011–1021.
[12] R.H. Swendsen, J.S. Wang, Phys. Rev. Lett. 58 (1987) 86–88.
[13] U. Wolff, Phys. Rev. Lett. 62 (1989) 361–364.
[14] H.G. Evertz, G. Lana, M. Marcu, Phys. Rev. Lett. 70 (1993) 875–879.
[15] B.A. Berg, T. Neuhaus, Phys. Lett. B267 (1991) 249–253.
[16] B.A. Berg, T. Neuhaus, Phys. Rev. Lett. 68 (1992) 9–12.
[17] B.A. Berg, Fields Inst. Commun. 26 (2000) 1–24; also see cond-mat/9909236.
[18] J. Lee, Phys. Rev. Lett. 71 (1993) 211–214;
J. Lee, ibid. 71 (1993) 2353.
[19] M. Mezei, J. Comput. Phys. 68 (1987) 237–248.
[20] C. Bartels, M. Karplus, J. Phys. Chem. B 102 (1998) 865–880.
[21] F. Wang, D.P. Landau, Phys. Rev. Lett. 86 (2001) 2050–2053.
[22] F. Wang, D.P. Landau, Phys. Rev. E 64 (2001) 056101.
[23] Q. Yan, R. Faller, J.J. de Pablo, J. Chem. Phys. 116 (2002) 8745–8749.
[24] G.M. Torrie, J.P. Valleau, J. Comput. Phys. 23 (1977) 187–199.
[25] J.S. Wang, R.H. Swendsen, J. Stat. Phys. 106 (2002) 245–285.
[26] B.A. Berg, T. Celik, Phys. Rev. Lett. 69 (1992) 2292–2295.

[27] B.A. Berg, U.H.E. Hansmann, T. Neuhaus, Phys. Rev. B 47 (1993) 497–500.
[28] W. Janke, S. Kappler, Phys. Rev. Lett. 74 (1995) 212–215.
[29] B.A. Berg, W. Janke, Phys. Rev. Lett. 80 (1998) 4771–4774.
[30] N. Hatano, J.E. Gubernatis, Prog. Theor. Phys. (Suppl.) 138 (2000) 442–447.
[31] U.H.E. Hansmann, Y. Okamoto, J. Comput. Chem. 14 (1993) 1333–1338.
[32] U.H.E. Hansmann, Y. Okamoto, Physica A212 (1994) 415–437.
[33] M.H. Hao, H.A. Scheraga, J. Phys. Chem. 98 (1994) 4940–4948.
[34] Y. Okamoto, U.H.E. Hansmann, J. Phys. Chem. 99 (1995) 11276–11287.
[35] A. Kolinski, W. Galazka, J. Skolnick, Proteins 26 (1996) 271–287.
[36] N. Urakami, M. Takasu, J. Phys. Soc. Jpn. 65 (1996) 2694–2699.
[37] S. Kumar, P. Payne, M. Vásquez, J. Comput. Chem. 17 (1996) 1269–1275.
[38] U.H.E. Hansmann, Y. Okamoto, F. Eisenmenger, Chem. Phys. Lett. 259 (1996) 321–330.
[39] U.H.E. Hansmann, Y. Okamoto, Phys. Rev. E 54 (1996) 5863–5865.
[40] U.H.E. Hansmann, Y. Okamoto, J. Comput. Chem. 18 (1997) 920–933.
[41] N. Nakajima, H. Nakamura, A. Kidera, J. Phys. Chem. B 101 (1997) 817–824.
[42] C. Bartels, M. Karplus, J. Comput. Chem. 18 (1997) 1450–1462.
[43] J. Higo, N. Nakajima, H. Shirai, A. Kidera, H. Nakamura, J. Comput. Chem. 18 (1997) 2086–2092.
[44] Y. Iba, G. Chikenji, M. Kikuchi, J. Phys. Soc. Jpn. 67 (1998) 3327–3330.
[45] A. Mitsutake, U.H.E. Hansmann, Y. Okamoto, J. Mol. Graphics Mod. 16 (1998) 226–238; 262–263.
[46] U.H.E. Hansmann, Y. Okamoto, J. Phys. Chem. B 103 (1999) 1595–1604.
[47] H. Shimizu, K. Uehara, K. Yamamoto, Y. Hiwatari, Mol. Sim. 22 (1999) 285–301.
[48] S. Ono, N. Nakajima, J. Higo, H. Nakamura, Chem. Phys. Lett. 312 (1999) 247–254.
[49] A. Mitsutake, Y. Okamoto, J. Chem. Phys. 112 (2000) 10638–10647.
[50] K. Sayano, H. Kono, M.M. Gromiha, A. Sarai, J. Comput. Chem. 21 (2000) 954–962.
[51] F. Yasar, T. Celik, B.A. Berg, H. Meirovitch, J. Comput. Chem. 21 (2000) 1251–1261.
[52] A. Mitsutake, M. Kinoshita, Y. Okamoto, F. Hirata, Chem. Phys. Lett. 329 (2000) 295–303.
[53] M.S. Cheung, A.E. Garcia, J.N. Onuchic, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 685–690.
[54] N. Kamiya, J. Higo, H. Nakamura, Protein Sci. 11 (2002) 2297–2307.
[55] S.M. Jang, Y. Pak, S.M. Shin, J. Chem. Phys. 116 (2002) 4782–4786.
[56] J.G. Kim, Y. Fukunishi, H. Nakamura, Phys. Rev. E 67 (2003) 011105.
[57] N. Rathore, T.A. Knotts IV, J.J. de Pablo, J. Chem. Phys. 118 (2003) 4285–4290.
[58] T. Terada, Y. Matsuo, A. Kidera, J. Chem. Phys. 118 (2003) 4306–4311.
[59] B.A. Berg, H. Noguchi, Y. Okamoto, Phys. Rev. E 68 (2003) 036126.
[60] H. Okumura, Y. Okamoto, Chem. Phys. Lett. 383 (2004) 391–396.
[61] T. Munakata, S. Oyama, Phys. Rev. E 54 (1996) 4394–4398.
[62] A.P. Lyubartsev, A.A. Martinovski, S.V. Shevkunov, P.N. Vorontsov-Velyaminov, J. Chem. Phys. 96 (1992) 1776–1783.
[63] E. Marinari, G. Parisi, Europhys. Lett. 19 (1992) 451–458.
[64] E. Marinari, G. Parisi, J.J. Ruiz-Lorenzo, in: A.P. Young (Ed.), Spin Glasses and Random Fields, World Scientific, Singapore, 1998, pp. 59–98.
[65] A. Irbäck, F. Potthast, J. Chem. Phys. 103 (1995) 10298–10305.
[66] A. Irbäck, E. Sandelin, J. Chem. Phys. 110 (1999) 12256–12262.
[67] G.R. Smith, A.D. Bruce, Phys. Rev. E 53 (1996) 6530–6543.
[68] U.H.E. Hansmann, Phys. Rev. E 56 (1997) 6200–6203.
[69] B.A. Berg, Nucl. Phys. B (Proc. Suppl.) 63A-C (1998) 982–984.
[70] W. Janke, in: B. Dünweg, D.P. Landau, A. Milchev (Eds.), Computer Simulations of Surfaces and Interfaces, NATO Science Series, II. Mathematics, Physics and Chemistry, vol. 114, Proceedings of the Advanced Study Institute, Kluwer Academic Publishers, Dordrecht, 2003, pp. 137–157.
[71] K. Hukushima, K. Nemoto, J. Phys. Soc. Jpn. 65 (1996) 1604–1608.
[72] K. Hukushima, H. Takayama, K. Nemoto, Int. J. Mod. Phys. C 7 (1996) 337–344.
[73] C.J. Geyer, in: E.M. Keramidas (Ed.), Computing Science and Statistics, Proceedings of the 23rd Symposium on the Interface, Interface Foundation, Fairfax Station, 1991, pp. 156–163.
[74] R.H. Swendsen, J.-S. Wang, Phys. Rev. Lett. 57 (1986) 2607–2609.
[75] K. Kimura, K. Taki, in: R. Vichnevetsky, J.J.H. Miller (Eds.), Proceedings of the 13th IMACS World Congress on Computation and Appl. Math. (IMACS'91), vol. 2, 1991, pp. 827–828.
[76] D.D. Frantz, D.L. Freeman, J.D. Doll, J. Chem. Phys. 93 (1990) 2769–2784.
[77] M.C. Tesi, E.J.J. van Rensburg, E. Orlandini, S.G. Whittington, J. Stat. Phys. 82 (1996) 155–181.
[78] Y. Iba, Int. J. Mod. Phys. C 12 (2001) 623–656.
[79] U.H.E. Hansmann, Chem. Phys. Lett. 281 (1997) 140–150.
[80] Y. Sugita, Y. Okamoto, Chem. Phys. Lett. 314 (1999) 141–151.
[81] M.G. Wu, M.W. Deem, Mol. Phys. 97 (1999) 559–580.
[82] Y. Sugita, A. Kitao, Y. Okamoto, J. Chem. Phys. 113 (2000) 6042–6051.
[83] Y. Sugita, Y. Okamoto, Chem. Phys. Lett. 329 (2000) 261–270.
[84] A. Mitsutake, Y. Okamoto, Chem. Phys. Lett. 332 (2000) 131–138.
[85] D. Gront, A. Kolinski, J. Skolnick, J. Chem. Phys. 113 (2000) 5065–5071.
[86] G.M. Verkhivker, P.A. Rejto, D. Bouzida, S. Arthurs, A.B. Colson, S.T. Freer, D.K. Gehlhaar, V. Larson, B.A. Luty, T. Marrone, P.W. Rose, Chem. Phys. Lett. 337 (2001) 181–189.
[87] H. Fukunishi, O. Watanabe, S. Takada, J. Chem. Phys. 116 (2002) 9058–9067.
[88] A. Mitsutake, Y. Sugita, Y. Okamoto, J. Chem. Phys. 118 (2003) 6664–6675;
A. Mitsutake, Y. Sugita, Y. Okamoto, J. Chem. Phys. 118 (2003) 6676–6688.
[89] A. Sikorski, P. Romiszowski, Biopolymers 69 (2003) 391–398.
[90] C.Y. Lin, C.K. Hu, U.H.E. Hansmann, Proteins 52 (2003) 436–445.
[91] G. La Penna, A. Mitsutake, M. Masuya, Y. Okamoto, Chem. Phys. Lett. 380 (2003) 609–619.
[92] M. Falcioni, M.W. Deem, J. Chem. Phys. 110 (1999) 1754–1766.
[93] Q. Yan, J.J. de Pablo, J. Chem. Phys. 111 (1999) 9509–9516.
[94] T. Nishikawa, H. Ohtsuka, Y. Sugita, M. Mikami, Y. Okamoto, Prog. Theor. Phys. (Suppl.) 138 (2000) 270–271.
[95] R. Yamamoto, W. Kob, Phys. Rev. E 61 (2000) 5473–5476.
[96] D.A. Kofke, J. Chem. Phys. 117 (2002) 6911–6914.
[97] T. Okabe, M. Kawata, Y. Okamoto, M. Mikami, Chem. Phys. Lett. 335 (2001) 435–439.
[98] Y. Ishikawa, Y. Sugita, T. Nishikawa, Y. Okamoto, Chem. Phys. Lett. 333 (2001) 199–206.
[99] A.E. Garcia, K.Y. Sanbonmatsu, Proteins 42 (2001) 345–354.
[100] R.H. Zhou, B.J. Berne, R. Germain, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 14931–14936.
[101] A.E. Garcia, K.Y. Sanbonmatsu, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 2782–2787.
[102] R.H. Zhou, B.J. Berne, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 12777–12782.
[103] M. Feig, A.D. MacKerell, C.L. Brooks III, J. Phys. Chem. B 107 (2003) 2831–2836.
[104] Y.M. Rhee, V.S. Pande, Biophys. J. 84 (2003) 775–786.
[105] J.W. Pitera, W. Swope, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 7587–7592.

[106] M.K. Fenwick, F.A. Escobedo, Biopolymers 68 (2003) 160–177.
[107] H.F. Xu, B.J. Berne, J. Chem. Phys. 112 (2000) 2701–2708.
[108] R. Faller, Q. Yan, J.J. de Pablo, J. Chem. Phys. 116 (2002) 5419–5423.
[109] A. Mitsutake, Y. Okamoto, in preparation.
[110] K. Hukushima, Phys. Rev. E 60 (1999) 3606–3614.
[111] T.W. Whitfield, L. Bu, J.E. Straub, Physica A 305 (2002) 157–171.
[112] S. Nosé, Mol. Phys. 52 (1984) 255–268.
[113] W.G. Hoover, Phys. Rev. A 31 (1985) 1695–1697.
[114] W.G. Hoover, A.J.C. Ladd, B. Moran, Phys. Rev. Lett. 48 (1982) 1818–1820.
[115] D.J. Evans, G.P. Morris, Phys. Lett. A98 (1983) 433–436.
[116] B.A. Berg, Markov Chain Monte Carlo Simulations and Their Statistical Analysis I and II, in preparation.
[117] B.A. Berg, Comp. Phys. Commun. 153 (2003) 397–406.
[118] T. Nagasima, Y. Sugita, A. Mitsutake, Y. Okamoto, in preparation.
[119] T. Nagasima, Y. Sugita, A. Mitsutake, Y. Okamoto, Comp. Phys. Commun. 146 (2002) 69–76.
[120] Y. Okamoto, in: J.E. Gubernatis (Ed.), The Monte Carlo Method in the Physical Sciences: Celebrating the 50th Anniversary of the Metropolis Algorithm, American Institute of Physics, Melville, 2003, pp. 248–260; cond-mat/0308119.