Estimation of the density of states by multicanonical molecular dynamics simulation

Hisashi Shimizu
Department of Physics, Faculty of Science, Shinshu University, 3-1-1 Asahi, Matsumoto, Nagano 390-8621, Japan
(Received 9 March 2004; revised manuscript received 23 June 2004; published 15 November 2004)

We formulate a procedure for calculating the density of states (DOS) from a multicanonical molecular dynamics (MMD) simulation. DOS cannot be obtained directly from the result of MMD simulation, because the Gaussian thermostat that is used in MMD simulation restricts the system to a spherical surface in momentum space. We perform MMD simulation for liquid Ar with Lennard-Jones potentials and evaluate DOS. Some physical quantities are estimated as a function of temperature from that DOS. The internal energy, entropy, and Helmholtz free energy are in good agreement with experiment. The quantity related to the fluctuation—the specific heat at constant volume—does not agree with experiment, which is ascribed to insufficient accuracy of DOS.

DOI: 10.1103/PhysRevE.70.056704 PACS number(s): 02.70.Ns, 65.20.+w

I. INTRODUCTION

The canonical molecular dynamics (MD) simulation has been a powerful tool for molecular simulations. The MD simulation has been applied to many systems to investigate the thermodynamic equilibrium state. But physically interesting systems generally have considerable freedom which causes the minimum rich structure in the potential energy surface. That potential energy structure makes it difficult for the system to access the entire phase space during the available simulation time. Thus it is difficult to investigate phase transitions, folding of protein, and chemical reactions by computer simulation. In order to overcome this drawback of the conventional canonical MD and Monte Carlo simulation, many new simulation techniques have been proposed in recent years [1–14]. These include the simulated annealing method [1], the histogram method [2,3], the umbrella sampling method [4], the simulated tempering method [5,6], the entropic sampling method [7], the $1/k$-sampling method [8], the replica exchange method [9,10], Tsallis statistics [11], and the multicanonical method [12–14]. Among these methods, the simulated tempering, entropic sampling, $1/k$-sampling, and multicanonical methods enable the system to have access over a wide area of phase space during a simulation time by considering a nonphysical ensemble. The aim of these methods is to obtain the density of states (DOS) or integrated DOS. It has been shown in Ref. [15] that those four methods are related to each other. Although umbrella sampling, the replica exchange method, and Tsallis statistics also consider a nonphysical ensemble to widen the accessible area of phase space, these methods are not intended to produce either DOS or integrated DOS. Recently, Wang and Landau proposed a new sampling method to enable a running estimate of DOS (DOS is concurrently estimated during Monte Carlo simulation) and to refine the accuracy of DOS. [16,17] This improvement has been applied to the Potts model, the spin glass model, Lennard-Jones fluid, and the Ising model [18–20]. DOS plays an important role in statistical physics in the estimation of entropy and free energy. Therefore, the simulation methods that enable the calculation of DOS are considered to be useful tools to investigate phase transitions and other interesting phenomena. The multicanonical method has been applied in many areas, particularly to the biomolecule systems (see Ref. [15] and references therein). Berg and Neuhaus proposed the multicanonical algorithm and they implemented that algorithm in Monte Carlo simulation [12]. Nakajima *et al.* implemented it in MD simulation [14]. In their work, MD simulation is performed by the canonical MD simulation with a modified Hamiltonian. Also note that the Gaussian thermostat [21–23] is more suitable than the Nosé-Hoover thermostat in performing the multicanonical MD (MMD) simulation. However, the Gaussian thermostat restricts the time development of momentum, which makes it impossible to obtain DOS directly from the simulation result. In this paper, we formulate how to estimate DOS from the MMD simulation result. DOS is calculated for liquid Ar with the Lennard-Jones potential. The accuracy of the estimated DOS is verified by calculating the internal energy, specific heat at constant volume, Helmholtz free energy, and entropy as a function of temperature. Section II is devoted to the formulation of how to estimate DOS from the MMD simulation result. In Sec. III, the model used in this paper is described. The simulation results and discussions are presented in Sec. IV.

II. THEORY

In Sec. II A, the MMD algorithm is explained briefly. In Sec. II B, by assuming the functional form of the Hamiltonian and the kinetic energy, the integration in the definition of DOS is partially carried out. And in Sec. II C we formulate how to estimate DOS from MMD simulation.

A. Multicanonical molecular dynamics simulation

The multicanonical algorithm was originally applied to the MD method by Nakajima *et al.* [14]. A brief explanation of the MMD method is given here. Suppose a system contains N identical particles. We may write the Hamiltonian of the system with a set of coordinates $\boldsymbol{r}=(\boldsymbol{r}_1,\boldsymbol{r}_2,\ldots,\boldsymbol{r}_N)$ and momenta $\boldsymbol{p}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\ldots,\boldsymbol{p}_N)$ as $H(\boldsymbol{r},\boldsymbol{p})$. In MMD simulation, the probability density is supposed to be

$$\rho(\boldsymbol{r},\boldsymbol{p}) \propto \exp[-\mathcal{W}(\boldsymbol{r},\boldsymbol{p})], \tag{1}$$

where $\mathcal{W}(\boldsymbol{r},\boldsymbol{p})$ depends on $\boldsymbol{r}$ and $\boldsymbol{p}$ only through Hamiltonian $H(\boldsymbol{r},\boldsymbol{p})$. In this ensemble, the probability distribution of energy is written as

$$\mathcal{P}(E) = \frac{1}{\mathcal{Z}} \int \delta(H(\boldsymbol{r},\boldsymbol{p}) - E)\rho(\boldsymbol{r},\boldsymbol{p})d\Gamma, \tag{2}$$

$$= \frac{1}{\mathcal{Z}} n(E) \exp[-\mathcal{W}(E)], \tag{3}$$

where $\mathcal{Z}$ is the partition function, $n(E)$ is the DOS, and the integral is carried out over the entire phase space. If $\mathcal{P}(E)$ is constant,

$$\mathcal{P}(E) = \text{const}, \tag{4}$$

then $\mathcal{W}(E) = \ln n(E)$ except for the constant term $-\ln \mathcal{P}(E)\mathcal{Z}$. In practical simulations, Eq. (4) is satisfied within a certain energy region. This energy region is determined according to what one is interested in. The purpose of MMD simulation is to obtain $\mathcal{W}(E)$ by MD simulations. If we write Eq. (3) as

$$\mathcal{P}(E) = \frac{1}{\mathcal{Z}} n(E) \exp\left(-\frac{\mathcal{H}(E)}{k_B T}\right), \tag{5}$$

it can be regarded as a canonical ensemble for a system with a modified Hamiltonian $\mathcal{H}$ at a given temperature T. k_B in Eq. (5) is the Boltzmann constant. T is just a parameter to mimic the canonical simulation. Although $\mathcal{H}$ is required before performing the MMD simulation, it is not known *a priori*. So we first approximate $\mathcal{H}$ from the result of conventional canonical MD simulation of the original system. DOS is approximately obtained as

$$n(E) = ZP(E) \exp\left(\frac{E}{k_B T}\right). \tag{6}$$

Taking the logarithm of Eq. (6), $\mathcal{H}(E)$ is approximately obtained as

$$\mathcal{H}(E) = E + k_B T \ln P(E), \tag{7}$$

where the term $\ln Z$ is neglected because it is independent of E. By performing the canonical MD simulation on the modified system, we obtain the energy probability distribution, $\mathcal{P}(E)$. If $\mathcal{P}(E)$ does not satisfy Eq. (4) during the specified energy region, $\mathcal{H}(E)$ is renewed as

$$\mathcal{H}_{\text{new}}(E) = \mathcal{H}_{\text{old}}(E) + k_B T \ln \mathcal{P}(E) \tag{8}$$

and another canonical MD simulation is performed. This procedure is iterated until Eq. (4) is satisfied over the specified energy region. DOS is evaluated by $\mathcal{H}(E)$ as

$$n(E) = \mathcal{Z}\mathcal{P}(E) \exp\left(\frac{\mathcal{H}(E)}{k_B T}\right). \tag{9}$$

Once DOS is obtained, the canonical distribution at any temperature $\mathcal{T}$ is generated by the reweighting technique as

$$P(E) \propto n(E) \exp\left(-\frac{E}{k_B \mathcal{T}}\right). \tag{10}$$

The prefix "multi-" signifies that the MMD simulation enables us to obtain probability distributions at any temperature.

B. Density of states

The definition of DOS is

$$n(E) = \int \delta(H(\boldsymbol{r},\boldsymbol{p}) - E)d\Gamma, \tag{11}$$

where the integral is carried out over the entire phase space. Suppose the Hamiltonian is separable into the kinetic energy and potential energy as

$$H(\boldsymbol{r},\boldsymbol{p}) = K(\boldsymbol{p}) + U(\boldsymbol{r}), \tag{12}$$

$$K(\boldsymbol{p}) = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m}, \tag{13}$$

where m is the particle mass and N is the number of particles. Then Eq. (11) can be rewritten as

$$n(E) = \int_0^{E-U_{\min}} \int \delta(H(\boldsymbol{r},\boldsymbol{p}) - E)\,\delta(K(\boldsymbol{p}) - K)d\Gamma dK, \tag{14}$$

$$= \int_0^{E-U_{\min}} \int \delta(U(\boldsymbol{r}) - [E - K(\boldsymbol{p})]) \times \delta(K(\boldsymbol{p}) - K)d\Gamma dK, \tag{15}$$

where $U_{\min}$ is the minimum value of $U(\boldsymbol{r})$. By carrying out the integration over momentum space, one obtains

$$n(E) = S_N \int_{U_{\min}}^{E} (E - U)^{(3N/2)-1} \tilde{n}(U)dU, \tag{16}$$

$$\tilde{n}(U) = \int \delta(U(\boldsymbol{r}) - U)d\boldsymbol{r}^{3N}, \tag{17}$$

where $S_N = (2m\pi)^{3N/2}/\Gamma(3N/2)$, $U = E - K$, and $\Gamma(x)$ is the Gamma function. In Eq. (17), $\tilde{n}(U)$ is the partial DOS, which is calculated by the integration over conformational space.

C. DOS estimation from MMD simulation

As described in Ref. [14], we adopt the Gaussian thermostat [21–23] to keep the temperature constant in performing the MMD simulation. The kinetic energy $K(\boldsymbol{p})$ is fixed at the specific value $K_0 = [(3N-1)/2]k_B T_0$ during the simulation. It has been shown in Ref. [23] that the probability density of a Gaussian thermostat becomes

$$\rho(\boldsymbol{r},\boldsymbol{p}) \propto \delta(K(\boldsymbol{p}) - K_0) \exp\left(-\frac{U(\boldsymbol{r})}{k_B T_0}\right), \tag{18}$$

where the Hamiltonian is assumed to be separable into kinetic energy $K(\boldsymbol{p})$ and potential energy $U(\boldsymbol{r})$. The modified

Hamiltonian, $\mathcal{H}(E)$, is also assumed to be a sum of kinetic energy and potential energy such as

$$\mathcal{H}(\boldsymbol{r},\boldsymbol{p}) = K(\boldsymbol{p}) + \Phi(\boldsymbol{r}) = K(\boldsymbol{p}) + U(\boldsymbol{r}) + k_B T \ln \mathcal{P}(U). \tag{19}$$

Because of the restriction on the kinetic energy, the additional term in the Hamiltonian, the third term of Eq. (19), can be regarded as a modification of the potential energy. Thus the probability distribution in MMD simulation becomes

$$\rho(\boldsymbol{r},\boldsymbol{p}) \propto \delta(K(\boldsymbol{p}) - K_0) \exp\left(-\frac{\Phi(\boldsymbol{r})}{k_B T_0}\right). \tag{20}$$

The energy probability distribution, $\mathcal{P}(E)$, becomes equivalent to the probability distribution of potential energy, $\mathcal{P}(U)$. $\mathcal{P}(U)$ is obtained by substituting Eq. (20) into Eq. (2) and carrying out the integration over the momentum space as follows:

$$\mathcal{P}(U) = \frac{1}{\mathcal{Z}} \exp\left(-\frac{\Phi(U)}{k_B T_0}\right)\tilde{n}(U), \tag{21}$$

where $\Phi(U)$ denotes $\Phi(\boldsymbol{r})$ at $\boldsymbol{r}=\boldsymbol{r}_0$ in which $U(\boldsymbol{r}_0)=U$, and $\mathcal{Z}$ is determined by the normalization of $\mathcal{P}(U)$. The condition of Eq. (4) is changed as

$$\mathcal{P}(U) = \text{const}, \tag{22}$$

and the procedure of MMD simulation is also changed as follows.

(i) Performing canonical MD simulation for the system of potential energy $\Phi(U)$ [at first time $\Phi(U)=U$].

(ii) Renewing $\Phi(U)$ as

$$\Phi_{\text{new}}(U) = \Phi_{\text{old}}(U) + k_B T \ln \mathcal{P}(U), \tag{23}$$

and go back to (i) until Eq. (22) is satisfied over the specified energy region.

(iii) Evaluating $\tilde{n}(U)$ by

$$\tilde{n}(U) = \mathcal{Z}\mathcal{P}(U) \exp\left(\frac{\Phi_{\text{new}}(U)}{k_B T}\right). \tag{24}$$

Even if Eq. (22) is satisfied, $n(E)$ is not obtained yet. The Gaussian thermostat restricts the accessible momentum space to a spherical surface of $\Sigma_i(p_i^2/2m)=[(3N-1)/2]k_B T_0$. Hence in order to get $n(E)$ from the MMD simulation, $\tilde{n}(U)$ should be substituted into Eq. (16) and integration carried out with U. Then we get the DOS as

$$n(E) = C_N \int_{U_{\min}}^{E} (E-U)^{(3N/2)-1} \exp\left(\frac{\Phi(U)}{k_B T_0}\right) dU, \tag{25}$$

where $C_N = S_N \mathcal{P}(U)\mathcal{Z}$. When Eq. (22) is satisfied over the specified energy region, the value of $\mathcal{P}(U)$ is approximately $1/\Delta U$, where ΔU is the width of the energy region. ΔU depends on how many MMD iterations are performed, but it is independent of both the kinetic energy and the potential energy. $\mathcal{Z}$ is also independent of U because the dependence of $\tilde{n}(U)$ on U is completely described by $\Phi(U)$. Therefore, both $\mathcal{P}(U)$ and $\mathcal{Z}$ can be taken out of the integral.

III. MODEL AND METHODS

We perform MMD simulation for liquid Ar with the Lennard-Jones potential,

$$U(\boldsymbol{r}) = \sum_{i=1}^{N} \sum_{j>i}^{N} 4\varepsilon \left\{ \left(\frac{\sigma}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}\right)^{12} - \left(\frac{\sigma}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}\right)^{6} \right\}, \tag{26}$$

where $N=500$ and the potential parameters are $\varepsilon/k_B = 120$ K and $\sigma = 3.4$ Å [24]. All particles are arranged in a cubic box to which the periodic boundary condition is applied. The density is set to 1.33 g/cm^3. As pointed out in Sec. II C, the Gaussian thermostat [21–23] is used for temperature control. The integration of the equation of motion is performed by the velocity Verlet [25] method with a time step $\delta t = 0.1$ fs. Although this δt is very short, by choosing this value, temperature is kept constant within the accuracy $\sim 10^{-5}$ K. In performing the MMD simulation, the force on each particle is modified according to the modification of the potential energy, such that

$$\boldsymbol{\mathcal{F}}_i = \left(1 + k_B T \frac{d}{dU} \ln \mathcal{P}(U)\right)[-\nabla_i U(\boldsymbol{r})]. \tag{27}$$

According to the logarithmic derivative of $\mathcal{P}(U)$, the multiplier enhances or suppresses the force. As pointed out in Ref. [14], this multiplier is uncongenial to the temperature fluctuation in the Nosé-Hoover thermostat. Too large a multiplier tends to corrupt the simulation, because it is equivalent to increasing the time step. In order to avoid this kind of corruption, we recommend adoption of a short δt. In Eq. (27), the logarithmic derivative of $\mathcal{P}(U)$ is required in calculating the force. Many other researchers have estimated this derivative by fitting the energy histogram to a polynomial in U and differentiating it. The fitting is performed over a selected energy region. However, when we perform the MMD simulation, we cannot neglect the possibility of an energy state appearing which is outside that selected energy region. Such a situation frequently occurs because the MMD simulation enhances the system to access an energy state that is outside the accessible area during the previous simulation. In this paper, we use the same method described in Ref. [26]. $\mathcal{P}(U)$ is approximated by a Lorenzian sum as follows:

$$\mathcal{P}(U) = \frac{1}{M} \sum_{n=0}^{n_U} h(U_n) \frac{1}{\pi} \frac{1}{1 + \left(\frac{U_n - U}{\sigma}\right)^2}, \tag{28}$$

where $h(U_n)$ is the energy histogram sampled at $U_n = U_0 + n\delta U$, ranging $n = 0, 1, \ldots, n_U$ and $M = \Sigma_n h(U_n)\delta U$. Then the analytical functional form of $d\mathcal{P}/dU$ can be obtained by differentiating Eq. (28). In Eq. (28), σ is an adjustable parameter, which is set to be $\sigma = \delta U$ in this present work. To be strict, the normalization factor $1/\pi$ in Eq. (28) must be $1/[(\pi/2) - \theta_0]$, with $\tan\theta_0 = (U_{\min} - U_n)/\sigma$ as $U_{\min}$ is the minimum of potential energy. In this work, θ_0 is approximated to $-\pi/2$, because we assume that the accessible range of potential energy is sufficiently larger than $U_{\min}$. In Refs. [18–20], the energy states that are outside the specified energy region are rejected in order to refine the efficiency of

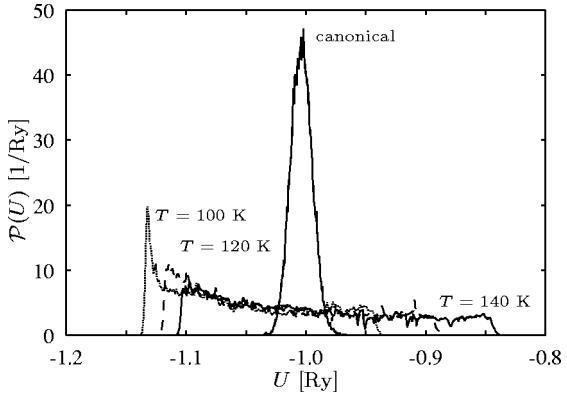

FIG. 1. The probability distribution of potential energy obtained by the conventional canonical MD simulation at $T=140$ K and the MMD simulation at $T=140$ K (solid line), 120 K (dashed line), and 100 K (broken line) after 16 MMD iterations.

simulation and the accuracy of DOS. It is impossible to reject any state that appears during the MD simulation because all states are the solution of the equation of motion. If we have enough information about the phase space or conformational space for the system under consideration, it is possible to suppress the appearance of the energy state outside the specified energy region by controlling the multiplier on the force. However, in general, we do not know details of the system. We cannot eliminate the possibility that the states corresponding to the specified energy region are scattered in the whole phase space and energy barriers higher than the maximum energy of the specified energy region lie between them. Thus care must be taken to suppress the energy region. In the present work, we neither control the multiplier on force nor restrict the energy region.

IV. RESULTS AND DISCUSSIONS

We perform the MMD simulation at three different temperatures ($T=100$ K, 120 K, and 140 K). All three temperatures are below the critical temperature of Ar, $T_c \sim 150$ K. The density ($1.33\ \mathrm{g/cm^3}$) is chosen to obtain the liquid state. For each temperature, we first perform the conventional canonical MD simulation with 4 million steps to prepare initial $\tilde{n}(U)$. After that, we iterate the MMD simulation 14 times with 4 million steps each and in the last two iterations 20 million steps each. Consequently, we perform 16 MMD iterations for each temperature. After 16 MMD iterations at $T=140$ K, the energy region covered by $\mathcal{P}(U)$ includes all the energy states that appear in the conventional canonical MD simulation at $T=100$ K. The energy region covered by $\mathcal{P}(U)$ at $T=100$ and 120 K also includes the energy states that appear in the conventional canonical MD simulation at $T=100$, 120, and 140 K. We expect these MMD results to reproduce the canonical distribution within the temperature range 100–140 K. The equilibration of each simulation is checked by the velocity autocorrelation function. Figure 1 shows the probability distribution of U from the conventional canonical MD simulation at $T=140$ K and MMD simulations after 16 iterations. This figure indicates that compared to the conventional canonical MD simulation, the MMD simulation enables the system to access a wide energy region during a single simulation. For each temperature, $\mathcal{P}(U)$ acquires a flat structure over the energy region $-1.1<U<-0.85$ for $T=140$ K, $-1.1<U<-0.9$ for 120 K, and $-1.1<U<-0.95$ for 100 K. By carrying the MMD iteration further, we expected that the energy range covered by

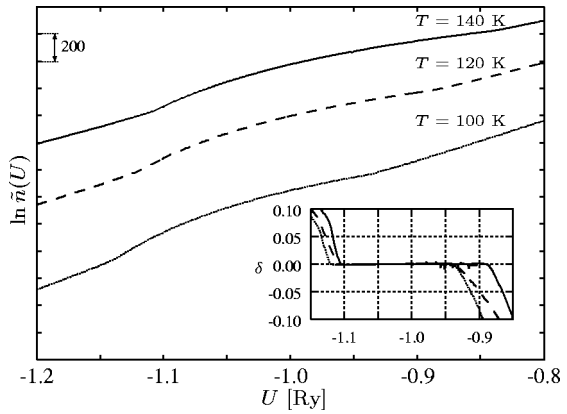

FIG. 2. $\ln \tilde{n}(U)$ from the results of 16 MMD iterations for $T=140$ K (solid line), 120 K (dashed line), and 100 K (broken line), respectively. The inset shows the difference in $\ln \tilde{n}(U)$ between 140 and 120 K (solid line), 140 and 100 K (dashed line), and 120 and 100 K (broken line). The definition of δ is Eq. (29).

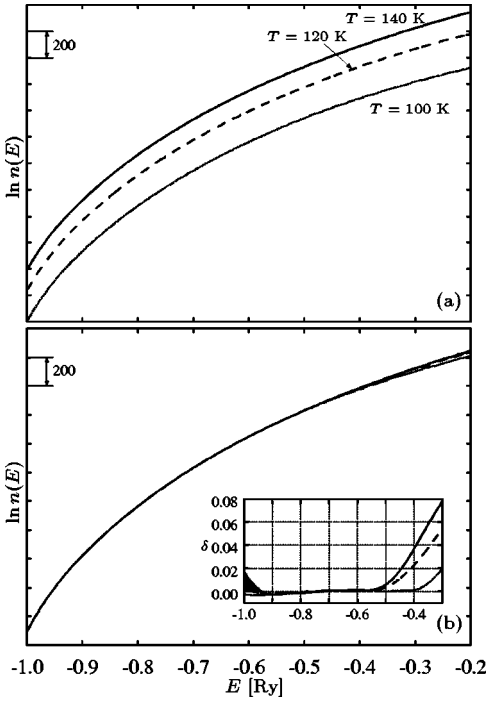

FIG. 3. (a) DOS estimated from the MMD simulation results at $T=140$ K (solid line), 120 K (dashed line), and 100 K (broken line), respectively. The lower panel (b) is obtained from the upper panel (a) by sliding $T=120$ K and 100 K lines in the vertical direction. The inset shows the difference in DOS between 140 and 120 K (solid line), 140 and 100 K (dashed line), and 120 and 100 K (broken line). δ is estimated by Eq. (29) by replacing $\tilde{n}(U)$ with n(E).

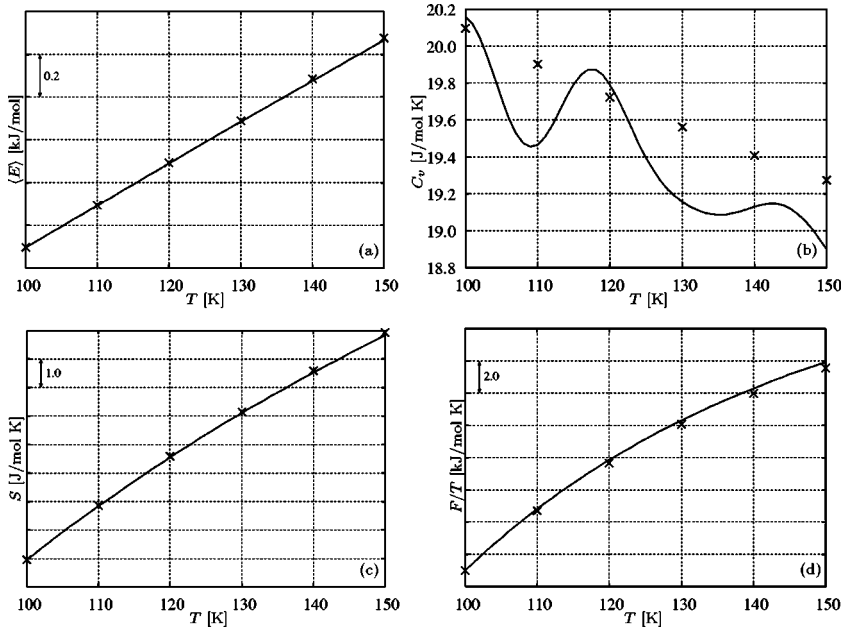

FIG. 4. Temperature dependence of (a) internal energy $\langle E\rangle$, (b) specific heat at constant volume C_v, (c) entropy S, and (d) Helmholtz free energy F/T. In each figure, the solid line is the MMD simulation result with $T=140$ K and $\times$ is the experimental data from NIST [27]. For $\langle E\rangle$, C_v, and F/T, the result of MMD is equated to the experiment at $T=100$ K.

the simulation will be increased. While some peak structures are recognized in $\mathcal{P}(U)$ for each temperature, they disappear with continuing MMD iteration. If the MMD iteration is carried further, an extra force originating from those peaks acts on each particle through Eq. (19). Figure 2 shows $\Phi(U)/k_BT \propto \ln \tilde{n}(U)$. Each of the three curves corresponds to $\ln \tilde{n}(U)$ at $T=140$ K, 120 K, and 100 K, respectively. The curves are not coincident. But this discrepancy can be ignored because it is ascribed to the omission of the term $\ln \mathcal{Z}$ in every iteration process on account of its being independent of U. If we vertically slide a curve, any curve coincides with others during the energy range in which the flat area of $\mathcal{P}(U)$ overlaps. During that energy range, the difference in $\ln \tilde{n}(U)$ between two different temperatures T and T' is $\delta<0.005$ for $(T,T')=(140,120)$, <0.003 for (140, 100), and <0.004 for (120, 100). δ is estimated by

$$\delta = \frac{\ln \tilde{n}_T(U) - \ln \tilde{n}_{T'}(U)}{\langle \ln \tilde{n}_T(U) - \ln \tilde{n}_{T'}(U)\rangle} - 1, \tag{29}$$

where the denominator of the first term is the average over the energy range in which $\mathcal{P}(U)$ at T and T' overlap. Note that both edges of $\ln \tilde{n}(U)$ are straight lines. This is because of the term U/k_BT in Eq. (19). It means that if $\ln \tilde{n}(U)$ is a straight line with a gradient $1/k_BT$ during some energy area, the energy state within that area never appeared during the MMD simulation. Thus in calculating DOS, both edges of $\tilde{n}(U)$ must be neglected. The important result of MMD simulation is the difference of $\ln \tilde{n}(U)$ from that straight line. The shape of each curve represents the density of equipotential states in the conformational space. Although $\mathcal{P}(U)$ has many peaks, all three curves of $\ln \tilde{n}(U)$ have a smooth shape. DOS is evaluated from these $\tilde{n}(U)$ as shown in the upper panel of Fig. 3. As was the case in Fig. 2, no curves are coincident. This discrepancy is also ascribed to the omission of the term $\ln C_N$. By sliding in the vertical direction, any curve can coincide with others as shown in the lower panel of Fig 3. In this figure, all three curves coincide very well between the energy region $-0.9<E<-0.5$ Ry. The error in DOS between two different temperatures is also estimated by Eq. (29) by replacing $\tilde{n}(U)$ with $n(E)$ as $\delta<0.002$ for $(T,T')=(140,120)$, <0.003 for (140, 100), and <0.006 for (120, 100). The denominator is evaluated during $[-0.9,-0.4]$ for $(T,T')=(140,120)$ and $[-0.9,-0.5]$ for (140, 100) and (120, 100). Due to omission of the term $\ln \mathcal{Z}$ and $\ln C_N$, it is impossible to estimate the absolute value of DOS from the MMD simulation results. However, they are sufficient to estimate the relative value of DOS for any different energy states over that energy region. By expressing the true DOS as $N(E)$, the present DOS, $n(E)$, can be expressed as

$$N(E) = \zeta n(E), \tag{30}$$

where ζ is a constant. The internal energy and the specific heat at constant volume can be expressed exactly by $n(E)$. On the other hand, Helmholtz free energy and entropy cannot be expressed in terms of uncertainty of ζ as follows.

(i) Internal energy,

$$\langle E\rangle = \frac{\int EN(E)e^{-\beta E}dE}{\int N(E)e^{-\beta E}dE} = \frac{\int En(E)e^{-\beta E}dE}{\int n(E)e^{-\beta E}dE}.$$

(ii) Specific heat at constant volume,

$$C_v = \frac{1}{k_B T^2}(\langle E^2\rangle - \langle E\rangle^2).$$

(iii) Helmholtz free energy,

$$\frac{F}{T} = -k_B \ln \int N(E) e^{-\beta E} dE$$

$$= -k_B \ln \int n(E) e^{-\beta E} dE - k_B \ln \zeta.$$

(iv) Entropy,

$$S = \frac{\langle E\rangle - F}{T}.$$

In these expressions, $\beta = 1/k_B T$. These four quantities are evaluated as a function of temperature and shown in Fig. 4 by using $n(E)$ obtained from the result of MMD simulation at $T = 140$ K. In those figures, the results are compared with experimental data from NIST [27]. The temperature range of Fig. 4 is chosen according to the energy range where Eq. (22) is satisfied in the MMD simulation at $T = 140$ K. The discrepancy between the MMD result and experimental data is large below 100 K and over 150 K. The internal energy agrees well with experiment except for the difference in energy origin. The entropy and Helmholtz free energy also agree with experiment except for the difference in the position of origin and the uncertainty of ζ. In spite of the good agreement of the internal energy, the specific heat at constant volume does not agree with experiment. If the probability distribution of energy has a unimodal structure, the internal energy is approximated by the peak position of the distribution. On the other hand, the specific heat at constant volume depends on its variance. It is affected strongly by the shape of the probability distribution, much more than the internal energy. The canonical distribution of energy is shown in Fig. 5 at some temperatures between 100 K and 150 K. They are estimated by multiplying $n(E)$ and the Boltzmann factor [see Eq. (10)] and normalizing it to 1. The shape of the canonical distribution is significantly different from the result of MMD simulation and conventional canonical MD simulation except for the peak position. The reason for the disagreement of C_v cannot be recognized in the canonical distribution. However, it must be ascribed to the inaccuracy of DOS.

In summary, the original MMD simulation proposed by Nakajima *et al.* [14] is not enough to calculate DOS by means of the Gaussian thermostat. That thermostat restricts the momenta onto a spherical surface of $3N$ dimensions with radius $\sqrt{[(3N-1)/2]k_B T}$. The formalism of how to calculate DOS by the MMD simulation is established if the Hamiltonian can be expressed as a sum of the kinetic energy and potential energy. We estimate DOS and some physical quantities as a function of temperature for liquid Ar. It is shown that the calculated DOS is independent of the parameter T, and the accuracy of the DOS is sufficient to estimate the statistically averaged value. The internal energy, entropy, and Helmholtz free energy, which are calculated by the statistical average, are in good agreement with experiment. On the other hand, the specific heat at constant volume, which is related to the fluctuation of energy, is in disagreement with experiment. In the MMD simulation, the accuracy of DOS can be improved by flattening the distribution. The flat distribution obtained in our simulation is not very accurate for estimating the fluctuation. The accuracy of the DOS can be checked by δ. The independence of $\tilde{n}(U)$ from the parameter T indicates the system accesses a sufficiently wide area in phase space during the simulation. But we did not systematically check the flatness of the distribution during the simulation. In MD simulation, we need to calculate the force. Hence the criterion proposed in Refs. [18–20] is not applicable. In order to refine the accuracy of the DOS in the MMD method, some criterion including how to estimate force is required.

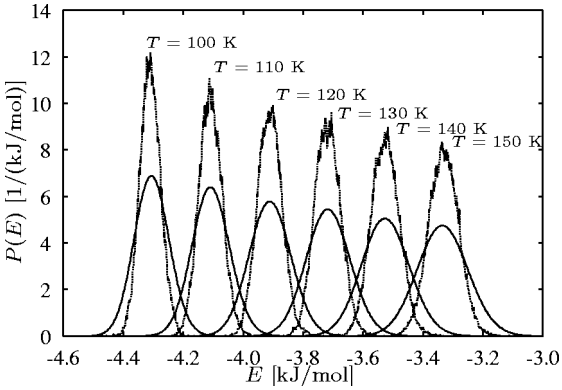

FIG. 5. The probability distribution of energy $P(E)$ for several temperatures which are estimated from the result of MMD simulation at $T = 140$ K (solid line) and the result of conventional canonical MD simulation (broken line).

[1] S. Kirkpatrick, J. C. D. Gelatt, and M. P. Vecchi, Science **220**, 671 (1983).
[2] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).
[3] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).
[4] G. M. Torrie and J. P. Valleau, J. Comput. Phys. **23**, 187 (1977).
[5] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
[6] A. P. Lyubartsev, A. A. Martsinnovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, J. Chem. Phys. **96**, 1776 (1992).
[7] J. Lee, Phys. Rev. Lett. **71**, 211 (1993).
[8] B. Hesselbo and R. B. Stinchcombe, Phys. Rev. Lett. **74**, 2151 (1995).
[9] K. Fukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).
[10] Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **314**, 141 (1999).
[11] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).

[12] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).
[13] U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **14**, 1333 (1993).
[14] N. Nakajima, H. Nakamura, and A. Kidera, J. Phys. Chem. B **101B**, 817 (1997).
[15] U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **18**, 920 (1997).
[16] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[17] F. Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001).
[18] Q. Yan, R. Faller, and J. J. de Pablo, J. Chem. Phys. **116**, 8745 (2002).
[19] Q. Yan and J. de Pablo, Phys. Rev. Lett. **90**, 035701 (2003).
[20] B. J. Schulz, K. Binder, M. Múller, and D. P. Landau, Phys. Rev. E **67**, 067102 (2003).
[21] W. G. Hoover, A. J. C. Ladd, and B. Moran, Phys. Rev. Lett. **48**, 1818 (1982).
[22] D. J. Evans, W. G. Hoover, B. H. Failor, B. Moran, and A. J. C. Ladd, Phys. Rev. A **28**, 1016 (1983).
[23] D. J. Evans and G. P. Morris, Phys. Lett. **98A**, 433 (1983).
[24] A. Rahman, Phys. Rev. **136**, A405 (1964).
[25] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, J. Chem. Phys. **76**, 637 (1982).
[26] H. Shimizu, K. Uehara, K. Yamamoto, and Y. Hiwatari, Mol. Simul. **22**, 285 (1999).
[27] E. Lemmon, M. McLinden, and D. Friend, NIST Chemistry WebBook, NIST Standard Reference Database Number 69 (URL http://webbook.nist.gov), edited by P.J. Linstrom and W.G. Mallard (2003).