Multicanonical basin hopping: A new global optimization method for complex systems

Lixin Zhan, Bart Piwowar, and Wing-Ki Liu
Department of Physics, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1

P. J. Hsu and S. K. Lai
Department of Physics, National Central University, Chungli, Taiwan 320, Republic of China

Jeff Z. Y. Chen
Department of Physics, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1

(Received 19 November 2003; accepted 29 December 2003)

We introduce a new optimization algorithm that combines the basin-hopping method, which can be used to efficiently map out an energy landscape associated with minima, with the multicanonical Monte Carlo method, which encourages the system to move out of energy traps during the computation. As an example of implementing the algorithm for the global minimization of a multivariable system, we consider the Lennard-Jones systems containing 150–185 particles, and find that the new algorithm is more efficient than the original basin-hopping method. © *2004 American Institute of Physics.* [DOI: 10.1063/1.1649728]

I. INTRODUCTION

Global optimization of a multivariable problem has long been an intensive subject of research in many fields and is an important issue in the characterization of complex systems.[1,2] The algorithms of optimization can be classified into four overlapping categories:[3–5] (i) deterministic methods, (ii) stochastic methods, (iii) heuristic methods, and (iv) smoothing methods. Most of the techniques exhibit varying degrees of success in applying to corresponding physical systems, for which the numerical techniques are specifically designed. The main difficulty in global minimization is associated with the fact that multiple local minima may exist, with locations separated from each other in the variable space. Most numerical procedures are efficient in finding a minimum; however, not all numerical procedures are efficient in finding the global minimum.

Methods that employ the Monte Carlo procedure, which belong to the second category, have played an essential role in global optimization. For example, the simulated annealing method[6] and its variants[7,8] have been widely used for finding the minimum of a multivariable function. By treating the targeted function as an effective ''energy potential,'' finding the function minimum is translated into finding the ground-state energy of a physical system. In comparison with the thermodynamic concept described by the Boltzmann weight, such a ground state would be reached as the temperature is lowered to absolute zero. The numerical procedure is controlled by an effective temperature, simulating the slow cooling (annealing) of a physical system, in analogy with the transitions from the gaseous to liquid and liquid to solid states in the crystal growing process. Normally, the Metropolis algorithm is implemented in the numerical procedure, to deal with the weighting process, according to the Boltzmann weight.[9]

The basin-hopping (BH) method,[1] which is in principle the same as the ''Monte Carlo minimization'' method of Li and Scheraga,[10] incorporates the advantages from both the deterministic and the stochastic methods. During the search procedure for the global minimum, BH uses an efficient numerical scheme to precisely locate a local minimum in a relatively short time. Then, by comparing the Boltzmann weight (at a fixed temperature) for this minimum with that determined in the previous step, the BH method decides whether the system would undergo a transition to the current minimum from the previous one. The most significant feature of this method is the allowance for the system to navigate between local minima, in particular, from a lower energy minimum to a higher one, enabling the system to hop among them.

BH has been shown to be very successful in dealing with complex systems such as the crystalline clusters[1,11,12] and the native structure of polyalanine.[13] There have been a number of other successful applications of BH (see references in the overview[2]). However, for systems containing a large number of variables, BH still suffers from the problem of getting trapped in deep local minima, and requires multiple computational runs in order to select independent simulation trajectories. The deficiency is inherited from the use of the Boltzmann weight with a fixed temperature as the hopping probability: $\exp(-\Delta E/k_B T)$. Within a finite number of computational steps, it becomes exponentially difficult to hop from the current minimum to a higher minimum if the energy difference ΔE is much larger than the thermal energy $k_B T$ at the simulation temperature T. The undesirable trapping at the local basins prevents the simulation from accessing other parts of the phase space.

To overcome this deficiency, in this paper we introduce a new optimization algorithm that incorporates a multicanonical weight into the basin-hopping method. Instead of using the Boltzmann weight that relies on a fixed temperature, a multicanonical weight has the promise of allowing the sys-

tem to visit different energy levels randomly. As will be demonstrated below, the new algorithm is found to be more efficient in comparison with the original basin-hopping method.

The usage of a non-Boltzmann weighting scheme in Monte Carlo simulations was first suggested by Torrie and Valleau,[14] who have proposed the so-called umbrella method. Various implementations of the multicanonical (MUCA) sampling method are generalizations based on the principal concept behind the umbrella method, and have been successfully used to determine the thermodynamic properties (particularly at low temperatures) and to search for the approximate location of the global energy minimum of many complex systems.[15–23] The basic idea is to use a statistical weight that is independent of T and more biased towards the low-energy region, in lieu of the Boltzmann weight. Ideally, the inverse of the density of states can be directly used as the weight, so that the entire simulation trajectory corresponds to a random walk in the energy space. However, the density of states of the system is not known *a priori*, and in practice, the multicanonical weight, and hence the density of states, is approximated by statistics found iteratively through the procedure in a pre-production run.[17–22,24] An advantage of incorporating the multicanonical algorithm in BH can be seen in our algorithm presented below. If the system stays with the current configurations for a prolonged number of computational steps, the accumulated statistics from the presence of this particular minimum would automatically reduce the probability used in the next iteration loop. The system is thus encouraged to move out of the current state.

In Sec. II, we introduce the multicanonical basin-hopping (MUBH) method in detail by first giving a brief description of the BH and the MUCA methods themselves. In Sec. III, we use the clustering problem of the Lennard-Jones (LJ) particles as an example of the application of the MUBH method. The energy minimization problem of this system has been well studied previously.[1,11,25–27] Hence the LJ systems provide an ideal vehicle to benchmark the speed of our new algorithm. The advantages of MUBH over the original BH method can be noticed for large clusters—notably those systems with more than 150 LJ particles.

II. METHOD

A. Basin-hopping (BH) method

Consider the potential energy of a system, $E(\mathbf{a})$, which is a function of all molecular coordinates $\mathbf{a}$. The dotted curve in Fig. 1 schematically shows the presence of multiple minima of the system. Initially, a random configuration $\mathbf{a}$ is chosen as the starting point, from which the configuration of the local minimum $\mathbf{a}_{\min}$ is determined numerically from a minimization procedure, such as the conjugate gradient method. Then $\mathbf{a}_{\min}$ is given a small trial ''move'' to a new configuration $\mathbf{a}'$ not far from $\mathbf{a}_{\min}$, and a new local minimum $\mathbf{a}'_{\min}$ is again obtained from the minimization procedure. One determines the acceptance of this new minimum configuration according to the Metropolis scheme[9] by considering a Boltzmann weight $\exp(-\Delta E_{\min}/k_B T)$, where $\Delta E_{\min}$ is the energy difference between the new minimum and the minimum

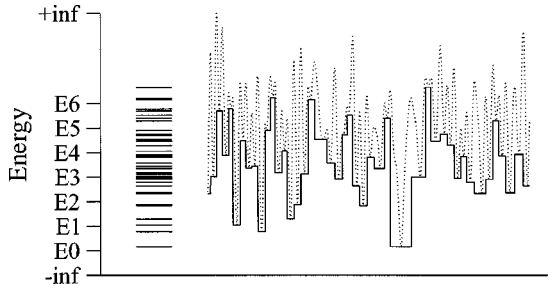

FIG. 1. A schematic illustration of the concept of the basin-hopping energy transformation in one dimension. The dotted curve shows an original function under consideration and the steps drawn by the solid line represent the local minima of the original curve, under the transformation in Eq. (1). The transformed energy landscape can be described by the energy spectrum to the left. Also shown is the division of the energy axis into various bins considered in the multicanonical procedure.

already found at the previous step. This procedure is repeated to search for the next local minimum. Computationally the local minimization procedure is the bottleneck of the BH method. We found that the limited-memory quasi-Newton optimization method[28] shows better performance when compared to other techniques discussed in Ref. 29. Any small improvement at this step would see repeated saving of the multi-iteration computational time.

Conceptually, the local-minimum searching step in the above algorithm is equivalent to transforming the energy landscape, represented by the dotted curve in Fig. 1, to a new reduced energy landscape, which consists of plateaus of energy minima only:

$$\tilde{E}(\mathbf{a}) = \min\{E(\mathbf{a})\} = E(\mathbf{a}_{\min}), \tag{1}$$

where $\min\{\cdots\}$ represents an energy minimization process by using $\mathbf{a}$ as the initial condition and $\mathbf{a}_{\min}$ is the configuration of the local minimum obtained from $\mathbf{a}$.[1] The energy maxima of the original function are discarded in the reduced energy landscape and are no longer of concern, as only the structure of the ''basins,'' represented by the solid curve in Fig. 1, is examined.

The reduced energy landscape is now a multistep function that has values of the local energy minima only. As can be seen in Fig. 1, there still exist energy barriers in the new landscape, where the new maxima, which are actually the minima of the original function, separate deeper minimum wells. The Monte Carlo part of the BH algorithm is used to handle the hopping of the system from one plateau to another under a thermal energy $k_B T$. The hopping probability depends highly on the reduced-energy difference between the plateaus of the two consecutive steps and the choice of $k_B T$. Hence the systems can still be trapped in the reduced energy landscape.

The basin-hopping method has been demonstrated to be superior to other techniques when applied to small size Lennard-Jones systems.[1] However, as remarked in Ref. 1, the efficiency of the approach ''could doubtless be improved by combining it with various other techniques.'' One of such improvements would be to use an annealing schedule to induce a temperature reduction process—a hibernate simulated

annealing and BH method; the other possibility would be to employ a non-Boltzmann weight in the Monte Carlo part of the simulation, as will be shown below.

B. Multicanonical Monte Carlo (MUCA) method

In a canonical ensemble, configurations at temperature T are weighted by the Boltzmann factor

$$w_B(E)=e^{-\beta E}, \tag{2}$$

where $\beta=(k_BT)^{-1}$. The resulting probability distribution of the energy, experienced by the system in the simulation, takes the form

$$P_B(E,T)\propto\rho(E)w_B(E), \tag{3}$$

where $\rho(E)$ is the density of states, which increases with energy rapidly. In the low-energy harmonic region and for a single minimum, $\rho(E)\propto(E-E_0)^{n_F/2}$, where n_F is the system degrees of freedom and E_0 its ground-state energy.[30] The Boltzmann factor decreases with energy exponentially for a given temperature. Thus $P_B(E,T)$ is strongly peaked at the average energy of the system corresponding to that temperature. Ideally, one uses a very low T to search for configurations near the ground state. However, a low T prevents the actual simulation from moving efficiently between the energy wells and barriers. Berg[24] has shown that $P_B(E,T)$ cannot be computed accurately due to the poor sampling in a canonical system.

The multicanonical ensemble, on the other hand, is designed differently, where the weight function used in the Monte Carlo simulation is directly related to the density of states by

$$w_{\mathrm{mu}}(E)\propto 1/\rho(E). \tag{4}$$

The resulting probability distribution of the energy is

$$P_{\mathrm{mu}}(E)\propto\rho(E)w_{\mathrm{mu}}(E)=\mathrm{const}, \tag{5}$$

i.e., the system is expected to move throughout the entire energy space by a random walk.

Without loss of generality, we simply set the constant of proportionality to unity in Eq. (4) because only the relative probabilities are required in the Metropolis scheme. However, the density of states of a physical system, and thus the multicanonical weight, is unknown *a priori*, and needs to be estimated via iterated numerical simulations itself. The starting iteration, the zeroth iteration, is usually performed by adopting the Boltzmann weight as the initial guess for $w_{\mathrm{mu}}(E)$, $w_{\mathrm{mu}}^{(0)}(E)\equiv w_B=\exp(-\beta^{(0)}E)$, with which an energy histogram $H^{(1)}(E)$ is constructed from the Monte Carlo sampling. Because we expect that $H^{(1)}(E)\propto w_{\mathrm{mu}}^{(0)}(E)\rho(E)$ or $\rho(E)\propto H^{(1)}(E)/w_{\mathrm{mu}}^{(0)}(E)$, Eq. (4) shows that an improved estimate of the multicanonical weight can be obtained from $w_{\mathrm{mu}}^{(1)}(E)=1/\rho(E)\approx w_{\mathrm{mu}}^{(0)}(E)/H^{(1)}(E)$. This procedure is repeated so that at the nth iteration, the simulation is carried out with the estimated weight $w_{\mathrm{mu}}^{(n)}(E)$, which yields a distribution $H^{(n+1)}(E)$ for E, collected within the nth iteration. A new estimate for the statistical weight

$$w_{\mathrm{mu}}^{(n+1)}(E)\approx w_{\mathrm{mu}}^{(n)}(E)/H^{(n+1)}(E) \tag{6}$$

is then used in the $(n+1)$th iteration.

In practice, one estimates a lower energy bound E_0 and an upper energy bound E_L from the first MC run, and divides the energy region of the interest into L bins having a bin width Δ. Each bin carries a label i and is characterized by its upper energy E_i. We further define the zeroth bin for energy $E\leqslant E_0$ and the $(L+1)$th bin for energy $E>E_L$. The statistics can now be collected for each bin, and $H^{(n+1)}(E_i)$ can be numerically defined as the number of states appearing in the ith bin during the nth iteration. To use the histogram effectively in the next iteration where a smooth $w_{\mathrm{mu}}^{(n+1)}(E)$ is required, the system entropy $S(E)\equiv-\ln w_{\mathrm{mu}}(E)$ in the ith bin can be parametrized as

$$S(E)=\beta_iE-\alpha_i, \quad \text{for} \quad E_{i-1}<E\leqslant E_i, \tag{7}$$

with $\beta_i=\partial S(E)/\partial E$ in bin i. The iteration equations of β and α based on the histogram distribution can be obtained from Eqs. (6) and (7) as

$$\beta_i^{(n+1)}=\beta_i^{(n)}+\left[\frac{\ln H^{(n+1)}(E_{i+1})-\ln H^{(n+1)}(E_i)}{\Delta}\right], \tag{8}$$

$$\alpha_i^{(n+1)}=\alpha_{i+1}^{(n+1)}+[\beta_i^{(n+1)}-\beta_{i+1}^{(n+1)}]E_i \tag{9}$$

for $i=0,1,\ldots,L-1$; and $\beta_i^{(n+1)}=\beta^{(0)}$, $\alpha_i^{(n+1)}=0$ for $i\geqslant L$.[31] Initially, $\beta_i^{(0)}=\beta^{(0)}$ and $\alpha_i^{(0)}=0$ for all i. The main updating procedure is reflected in Eq. (8) that relates the new choice of $\beta^{(n+1)}$ for the next iteration to the current choice by a correction term in the square brackets, which is directly related to the accumulated histogram. Berg has recently suggested that the correction should be re-weighted according to the variance of the distribution of each bin, in order to improve the significance of the correction term in the square brackets,[17,19,24]

$$\beta_i^{(n+1)}=\beta_i^{(n)}+\hat{g}_i^{(n+1)}\times\left[\frac{\ln H^{(n+1)}(E_{i+1})-\ln H^{(n+1)}(E_i)}{\Delta}\right], \tag{10}$$

where

$$\hat{g}_i^{(n+1)}=\frac{g^{(n+1)}(E_i)}{\Sigma_{k=1}^{n+1}g^{(k)}(E_i)} \tag{11}$$

and

$$\frac{1}{g^{(k)}(E_i)}=\frac{1}{H^{(k)}(E_{i+1})}+\frac{1}{H^{(k)}(E_i)}. \tag{12}$$

The major advantage of this improved parametrization, compared to Eq. (8), is that it leads to more stable simulating results. We have used Berg's improved expression, Eq. (10), together with Eq. (9) in this work. The updating procedure proceeds from the high-energy bins towards the low-energy bins, and if either $H^{(n+1)}(E_{i+1})=0$ or $H^{(n+1)}(E_i)=0$, we set $\beta_i^{(n+1)}=\beta_i^{(n)}$. The weight for the $(n+1)$th iteration is then calculated according to $w_{\mathrm{mu}}^{(n+1)}(E)=e^{-\beta_i^{(n+1)}E+\alpha_i^{(n+1)}}$ for E belonging to the ith bin.

C. Multicanonical basin-hopping (MUBH) method

MUCA is a very promising method that has particularly benefited the exploration of the low-energy landscape of

various complex systems in the calculation of thermodynamic properties at low temperatures.[15–18,20,22,31] Because MUCA contains an effective temperature $T_i=1/\beta_i$ which is *not* identically 0 for the lowest energy bin, thermal fluctuations deter the system from descending into the global energy minimum as many low-energy configurations still contribute prominently. To this extent, it suffers the same problem as the simulated annealing method, and can only be used to give an estimate, but not the exact value, of the ground-state energy. Considering the BH's merit in fast locating a minimum precisely and the MUCA's ability to surmount high-energy barriers, these two methods are complementary to each other.

The main idea of MUBH is to handle the hopping between the plateaus in the reduced-energy landscape with a probability function determined by the energy spectrum that contains all reduced-energy plateaus. The targeted multicanonical weight $w_{\mathrm{mu}}(E)$ in the original MUCA is replaced by $w_{\mathrm{mu}}(\tilde{E})=1/\rho(\tilde{E})$, with $\rho(\tilde{E})$ the density of minimal states.

MUBH contains the following steps. The first iteration is identical to the original BH procedure by carrying out a limited run that contains M canonical Monte Carlo steps for the transformed energy landscape, as described above. To do so, an initial temperature $T^{(0)}$ needs to be selected and the sensitivity of this selection on the efficiency of the algorithm will be addressed below. Upon finishing the first iteration, we collected the histogram of the reduced energy defined in Eq. (1), and started to consider a multicanonical Monte Carlo procedure for the acceptance of a new hopping. Most equations in the last section remain the same, provided that the energies E in these equations are replaced by the reduced counterparts $\tilde{E}$.

For a given system, we first selected a lower bound E_0, chosen close to the best estimate of the lowest energy, and established an upper bound E_L by identifying it with a value close to which the histogram attains its maximum in a limited MC run with the initial temperature $T^{(0)}$. The range $[E_0,E_L]$ is then divided into L equal segments with an increment Δ. Because we are only concerned about the energy plateaus in Fig. 1, which schematically gives us the energy spectrum on the left-hand side, the statistics collected during the simulation represent the frequency of visited plateaus. The statistical weight is characterized by the constants α_i and β_i for the ith bin. For $\tilde{E}>E_L$, the Boltzmann weight $w_B(\tilde{E})=\exp(-\tilde{E}/k_BT^{(0)})$ is used. In this weighting strategy, the final histogram distribution is expected to be smooth, if not completely flat, within $[E_0,E_L]$.

III. APPLICATION AND DISCUSSION

The Lennard-Jones cluster consisting of N particles (LJ_N) forms a crystal structure corresponding to an energy minimum. The LJ_N cluster was originally selected to demonstrate the effectiveness of the BH method, and the global energy minima for each given N have been relatively well determined.[1,26,27] Even for systems with cluster size as small as $N=98$,[11] there are of the order of 10^{40} minima, so that they are sophisticated enough for testing algorithms. Actually, the Lennard-Jones system has become a benchmark for checking the efficiency and accuracy of global optimization method recently.

The energy of the LJ_N system is given by

$$E=4\epsilon\sum_{i>j}^{N}\left[\left(\frac{\sigma}{r_{ij}}\right)^{12}-\left(\frac{\sigma}{r_{ij}}\right)^{6}\right],$$

where r_{ij} is the distance between particles i and j, ϵ and $2^{1/6}\sigma$ are the pair equilibrium well depth and separation, respectively, for a pair of particles. Hereafter we employ the reduced units, i.e., $\epsilon=\sigma=k_B=1$. We performed both BH and MUBH simulations on various LJ_N systems for comparison.

For the BH method, all of the calculations were performed with the temperature $T=0.8$. The MC displacement step size was initially set to 0.4, and was adjusted automatically during the simulations to maintain an acceptance ratio of 50%, averaged over every 100 MC steps. As for the MUBH calculations, two groups of runs based on different initial temperatures were performed. For $T^{(0)}=2.0$, the multicanonical update was performed after every $M=2000$ basin-hopping MC steps; while for $T^{(0)}=5.0$, the update was after every $M=1000$ basin-hopping MC steps. In a typical MUBH run, the entire energy range was divided into 12 bins—ten bins between E_0 and E_L, together with two additional bins $[-\infty,E_0]$ and $[E_L,+\infty]$. E_0 was chosen to be close to the global energies provided in Ref. 26. For the lower initial-temperature runs, $T^{(0)}=2.0$, the value of E_L will be smaller than that of the higher initial-temperature runs, $T^{(0)}=5.0$. Runs with a lower initial temperature had a narrower MUBH energy range.

Table I shows the results of BH and MUBH simulations on LJ_N. In the table, the number of runs failed ($\mathcal{F}$) to find the global minimum after a given number of steps, and the total number of runs ($\mathcal{T}$) are both shown. For any given system size N, $L_{\max}$ represents the maximal MC steps permitted in the simulation in case of any failed runs, or the actual maximal MC steps used to reach the global minimum if all runs are successful. For all the successful runs, the global energy minima obtained in our BH and MUBH calculations are identical to those reported in Ref. 26. The average MC steps to reach the global minimum listed in the table are plotted in Fig. 2 as a function of N. Compared to BH, MUBH requires less MC steps by factors of 2–5 in the global minimum exploring process. The figure also demonstrates that as N increases the reduction in computational time for finding the global minimum becomes more significant. The exact improvement, however, depends on the physical systems and hence cannot be quantified easily. The improvement of MUBH over BH for systems of size $N<150$ is not noteworthy, as for smaller systems, the energy landscape is relatively simple. In some cases, MUBH can even give worse results than BH because BH is already a very efficient method especially for low barrier cases, and it takes only several thousand MC steps to find the global minima for these systems, long before MUBH can reach a stable multicanonical weight. For example, we have tested both methods for systems of size $N=100-150$, and could not find saving in computational time in MUBH in comparison with BH.

TABLE I. The average number of MC steps to reach the global minimum for each N (Ave.) and their standard deviation (SD) in both BH and MUBH methods. $\mathcal{F}$ gives the number of runs failed to reach the global minimum after $L_{\max}$ steps and $\mathcal{T}$ is the total number of runs conducted. If $\mathcal{F}=0$, $L_{\max}$ represents the maximal number of steps to reach the global minimum among the $\mathcal{T}$ runs.

	BH ($T^{(0)}=0.8$)				MUBH ($T^{(0)}=2.0$)				MUBH ($T^{(0)}=5.0$)			
N	Ave.	SD	$L_{\max}$	$\mathcal{F}/\mathcal{T}$	Ave.	SD	$L_{\max}$	$\mathcal{F}/\mathcal{T}$	Ave.	SD	$L_{\max}$	$\mathcal{F}/\mathcal{T}$
150	20 114	4 522	100 000	1/15	13 967	4 394	69 166	0/15	9 862	1 573	22 641	0/15
155	62 674	17 964	500 000	1/16	35 027	8 004	500 000	1/15	22 799	4 789	73 832	0/20
160	157 940	80 003	1 500 000	5/16	58 327	15 871	253 582	0/16	39 294	8 149	200 000	1/16
165	635 236	124 619	2 000 000	4/16	92 362	18 704	300 000	1/16	141 610	25 654	500 000	1/18
170	333 800	85 854	1 400 000	1/16	119 980	32 125	1 300 000	2/16	51 200	8 180	200 000	2/20
175	290 970	106 290	1 268 539	0/16	81 277	29 282	500 000	1/16	52 306	15 685	1 400 000	1/14
180	234 310	77 099	2 000 000	1/15	127 800	27 984	400 000	1/15	47 353	9 478	200 000	1/15
181	259 620	54 799	816 121	0/16	132 740	27 636	2 000 000	1/16	54 762	6 342	92 671	0/14
182	265 280	59 801	1 500 000	2/16	137 510	28 219	2 000 000	4/16	64 606	15 709	600 000	1/18
183	1 000 972	206 894	2 603 728	0/16	217 200	54 835	1 000 000	1/16	207 690	49 353	650 000	2/18
185	1 194 140	170 688	4 000 000	1/16	357 920	92 595	1 594 614	0/18	419 380	91 730	2 000 000	2/19

The embedded capability of overcoming energy barriers in MUBH is the reason for the improvement in computational time. Figures 3(a) and 3(b) show the time trajectories of $\tilde{E}$ for both BH and MUBH, respectively. It is visible from Fig. 3(a) that the BH run can be trapped in a local minimum for a long computational time. In comparison, Fig. 3(b) shows that the computation in MUBH proceeds very differently. The trajectory covers a much wider energy range and the MUBH run rarely gets trapped in a local energy minimum. Hence, in comparison with BH, MUBH encourages better navigation in the reduced energy landscape in search of the global energy minimum.

One important parameter that influences the MUBH search efficiency is the initial temperature $T^{(0)}$. This is already apparent in Fig. 2 for the two choices of $T^{(0)}$. Using $N=150$ and $N=170$ as examples, we have further performed independent MUBH runs for a number of different choices of $T^{(0)}$. The average MC steps for finding the global minimum as a function of $T^{(0)}$ are listed in Table II and plotted in Fig. 4. Also shown in Table II are the corresponding values of E_L. From both the table and the figures, it is clear that when the initial temperature is low, it requires more MC loops on the average for systems to reach their global minima. When $T^{(0)}$ is high, less MC loops will be needed. A low $T^{(0)}$ produces a low E_L which discourages the anticipated navigation to the high-energy region. This potential pitfall becomes less serious as we use a higher $T^{(0)}$, resulting in a higher E_L. However, computationally, it is not always true that a higher $T^{(0)}$ necessarily leads to better efficiency. Indeed, the energy range is widened for the search with a high initial $T^{(0)}$, but consequently it also requires a larger MC displacement step. We find that such increase would demand longer computational time for the convergence of the conjugate gradient algorithm, in every MC step in MUBH. The ideal case would be to select a $T^{(0)}$ that corresponds to an E_L just above the energy barriers in the system—a largely unknown factor for any given complex system. Empirically, we found that $T^{(0)}$ in the range of 3.0–4.0 is probably optimal for conducting MUBH simulations for the $N=150$ and 170 LJ_N systems, as can be seen from

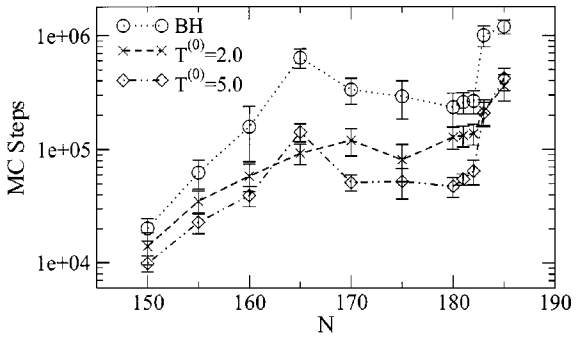

FIG. 2. BH and MUBH average Monte Carlo steps in finding global minima. The circles represent BH results, while the crosses and diamonds are MUBH results for $T^{(0)}=2.0$ and 5.0, respectively. The error bar represents the standard deviation of the $\mathcal{T}$ runs for each N.

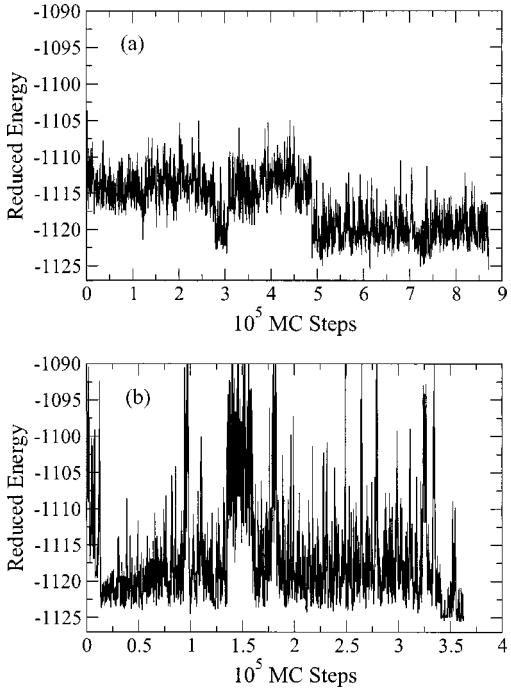

FIG. 3. Typical search trajectories of (a) BH with $T=0.8$ and (b) MUBH with $T^{(0)}=5.0$ for $N=185$.

TABLE II. The average number of MC steps correspond to different initial temperature for $N=150$ and $N=170$ systems. Ave., SD, $L_{\max}$, and $\mathcal{F}/\mathcal{T}$ are the same as those of Table I. E_L stands for the multicanonical upper bound energy with the unit of ϵ.

	$N=150$					$N=170$				
$T^{(0)}$	E_L	Ave.	SD	$L_{\max}$	$\mathcal{F}/\mathcal{T}$	E_L	Ave.	SD	$L_{\max}$	$\mathcal{F}/\mathcal{T}$
1.0	−870.4	13 765	4053	64 361	0/15	−1006.8	216 214	96 763	1 300 000	2/16
1.5	−870.4	13 880	3544	53 988	0/15	−1004.8	89 855	17 056	300 000	2/16
2.0	−870.4	13 967	4394	69 166	0/15	−1002.8	119 980	32 125	1 300 000	2/16
2.5	−868.4	11 714	2158	29 893	0/15	−1001.8	92 336	18 558	500 000	1/16
3.0	−868.4	13 881	3022	41 495	0/15	−1000.8	84 459	22 656	380 530	0/16
3.5	−865.4	10 410	1459	26 616	0/15	−999.8	66 016	9 763	140 733	0/16
4.0	−865.4	10 324	1258	21 957	0/15	−998.8	62 312	12 321	202 025	0/16
4.5	−864.4	12 529	1950	27 077	0/15	−997.8	80 160	24 820	416 724	0/16
5.0	−863.4	9 862	1573	22 641	0/15	−995.8	51 200	8 180	200 000	2/20
0.8(BH)		20 114	4522	100 000	1/15		333 800	85 854	1 400 000	1/16

Fig. 4. We have also developed the following strategy in selecting $T^{(0)}$. Before performing a production MUBH run, we try several short BH runs of one or two thousand MC steps for each of the several values of $T^{(0)}$. If all these runs for a $T^{(0)}$ give similar histogram distributions, this $T^{(0)}$ become one of the candidate. After we obtain several candidates of $T^{(0)}$, we choose the one with lowest value.

Yet another important parameter that influences the MUBH efficiency is the total MC steps M between each multicanonical update for α and β. A reduction of M would effectively reduce multiple steps of the conjugate gradient implementation in our algorithm. On the other hand, enough statistics for the energy bins should be accumulated before each update. For our study of the LJ_N systems, 1000–2000 MC steps between multicanonical update samplings on the $10+2$ bins gave satisfactory results. In general, the number of bins and the number of MC sweeps in each MUCA iteration should be carefully chosen.

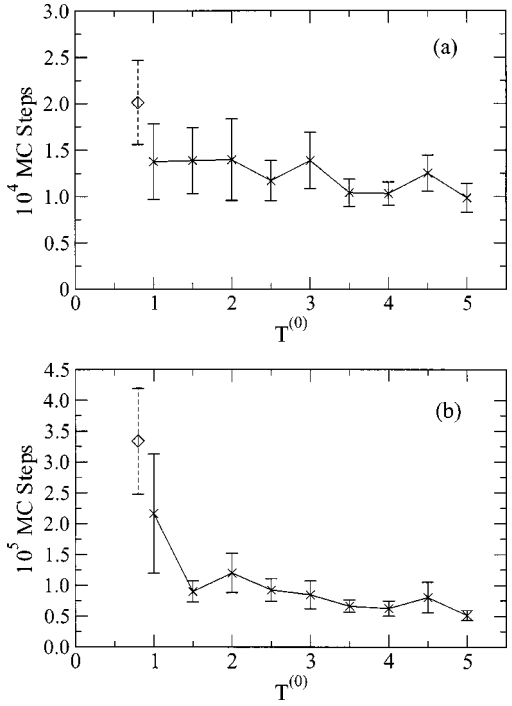

FIG. 4. Initial temperature dependence of the average number of steps to locate the global minimum using MUBH for (a) $N=150$ and (b) $N=170$. The points represented by diamonds at $T^{(0)}=0.8$ are the BH results. The error bar corresponds to the standard deviation for each run.

IV. SUMMARY

We have presented in this paper a new Monte Carlo method, the multicanonical basin-hopping (MUBH) method, as a practical approach to global optimization. We have combined the multicanonical Monte Carlo and the basin-hopping methods in order to make use of the advantages of both methods. To ascertain its efficiency, we have implemented it on the benchmark systems of the Lennard-Jones clusters. For small systems, $N<150$, the MUBH method gave no clear improvement over the BH method because the reduced potential-energy surface is relatively simple, so that BH could locate the global energy minimum before MUBH becomes effective. When the system size is increased to $N>150$, the improvement of MUBH over BH is dramatic. These observations suggest that MUBH is suitable for large systems. The efficiency of MUBH comes from the fact that not only can it hop between the local minima directly, an advantage from the basin-hopping method, but it can also easily overcome the plateaus energy barriers using the reweighting technique of the multicanonical method. It thus solves the problems of energy barriers in the reduced-energy landscape of basin hopping, and the insufficient sampling of low-energy landscape of the multicanonical method. Our paper also shows that the initial temperature setting is very important for the method. A suitable initial temperature will result in much better performance.

ACKNOWLEDGMENTS

We would like to thank NSERC for financial support and SHARCNET for computational time. S.K.L. wishes to acknowledge partial support from the National Science Council, Taiwan, Republic of China under Contract No. NSC92-2112-M-008-033.

[1] D. J. Wales and J. P. K. Doye, J. Phys. Chem. A **101**, 5111 (1997).
[2] D. J. Wales and H. A. Scheraga, Science **285**, 1368 (1999).
[3] A. R. Leach, Rev. Comput. Chem. **2**, 1 (1991).
[4] H. A. Scheraga, Rev. Comput. Chem. **3**, 73 (1992).

[5] R. V. Pappu, R. K. Hart, and J. W. Ponder, J. Phys. Chem. B **102**, 9725 (1998).
[6] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Science **220**, 671 (1983).
[7] Z. Wang and R. Pachter, J. Comput. Chem. **18**, 323 (1997).
[8] L. Ingber, http://www.ingber.com/ASA-README.html
[9] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[10] Z. Li and H. A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. **84**, 6611 (1987).
[11] R. H. Leary and J. P. K. Doye, Phys. Rev. E **60**, R6320 (1999).
[12] S. K. Lai, P. J. Hsu, K. L. Wu, W. K. Liu, and M. Iwamatsu, J. Chem. Phys. **117**, 10715 (2002).
[13] P. N. Mortenson and D. J. Wales, J. Chem. Phys. **114**, 6443 (2001).
[14] G. M. Torrie and J. P. Valleau, J. Comput. Phys. **23**, 187 (1977).
[15] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).
[16] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).
[17] B. A. Berg, in *Monte Carlo Methods*, edited by N. Madras, Fields Institute Communications Vol. 26 (American Mathematical Society, 2000).
[18] F. Yasar, T. Celik, B. A. Berg, and H. Meirovitch, J. Comput. Chem. **21**, 1251 (2000).
[19] B. A. Berg, Nucl. Phys. B (Proc. Suppl.) **63A-C**, 982 (1998).
[20] U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **14**, 1333 (1993).
[21] J. Lee, Phys. Rev. Lett. **71**, 211 (1993); **71**, 2353 (1993).
[22] M.-H. Hao and H. A. Scheraga, J. Phys. Chem. **98**, 4940 (1994).
[23] B. A. Berg, U. H. E. Hansmann, and Y. Okamoto, J. Phys. Chem. **99**, 2236 (1995).
[24] B. A. Berg, Comput. Phys. Commun. **153**, 397 (2003).
[25] R. H. Leary, J. Global Optim. **11**, 35 (1997).
[26] D. Romero, C. Barron, and S. Gomez, Comput. Phys. Commun. **123**, 87 (1999); we used the latest updated minimum value for $N=185$ reported in http://www.vcl.uh.edu/~cbarron/LJ_cluster/LJpottable.html
[27] D. J. Wales, J. P. K. Doye, A. Dullweber, M. P. Hodges, F. Y. Naumkin, F. Calvo, J. Hernandez-Rojas, and T. T. Middleton, the Cambridge Cluster Database, http://www-wales.ch.cam.ac.uk/CCD.html
[28] D. Liu and J. Nocedal, Math. Program. **45**, 503 (1989).
[29] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannary, *Numerical Recipes in C*, 2nd edition (Cambridge University Press, Cambridge, England, 1992).
[30] U. H. E. Hansmann and Y. Okamoto, Phys. Rev. E **56**, 2228 (1997).
[31] Y. Okamoto and U. H. E. Hansmann, J. Phys. Chem. **99**, 11276 (1995).