

1990

Reducing Quasi-Ergodic Behavior in Monte Carlo Simulations by J-Walking: Applications to Atomic Clusters

D. D. Frantz
University of Rhode Island

David L. Freeman
University of Rhode Island, dfreeman@uri.edu

Jimmie D. Doll

Follow this and additional works at: https://digitalcommons.uri.edu/chm_facpubs

Citation/Publisher Attribution

Frantz, D. D., Freeman, D. L. & Doll, J. D. (1990) Reducing Quasi-Ergodic Behavior in Monte Carlo Simulations by J-Walking: Applications to Atomic Clusters. *Journal of Chemical Physics*, 93(4), 2769-2784. doi: 10.1063/1.458863
Available at: <http://dx.doi.org/10.1063/1.458863>

This Article is brought to you by the University of Rhode Island. It has been accepted for inclusion in Chemistry Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu. For permission to reuse copyrighted content, contact the author directly.

Reducing Quasi-Ergodic Behavior in Monte Carlo Simulations by J-Walking: Applications to Atomic Clusters

Publisher Statement

© 1990 American Institute of Physics.

Terms of Use

All rights reserved under copyright.

Reducing quasi-ergodic behavior in Monte Carlo simulations by *J*-walking: Applications to atomic clusters

D. D. Frantz and D. L. Freeman

Department of Chemistry, University of Rhode Island, Kingston, Rhode Island 02881

J. D. Doll

Department of Chemistry, Brown University, Providence, Rhode Island 02912

(Received 11 April 1990; accepted 3 May 1990)

A method is introduced that is easy to implement and greatly reduces the systematic error resulting from quasi-ergodicity, or incomplete sampling of configuration space, in Monte Carlo simulations of systems containing large potential energy barriers. The method makes possible the jumping over these barriers by coupling the usual Metropolis sampling to the Boltzmann distribution generated by another random walker at a higher temperature. The basic techniques are illustrated on some simple classical systems, beginning for heuristic purposes with a simple one-dimensional double well potential based on a quartic polynomial. The method's suitability for typical multidimensional Monte Carlo systems is demonstrated by extending the double well potential to several dimensions, and then by applying the method to a multiparticle cluster system consisting of argon atoms bound by pairwise Lennard-Jones potentials. Remarkable improvements are demonstrated in the convergence rate for the cluster configuration energy, and especially for the heat capacity, at temperatures near the cluster melting transition region. Moreover, these improvements can be obtained even in the worst-case scenario where the clusters are initialized from random configurations.

I. INTRODUCTION

We present a method for reducing errors occurring in Metropolis¹ walks because of insufficient sampling of configuration space. This problem, which remains one of the fundamental concerns of Monte Carlo simulations, arises in systems where the sample space contains two or more regions marked by a very low transition probability between them, resulting in bottlenecks that effectively confine the sampling to only one of the regions.² The dichotomy of time scales characterizing the walks produces rapid motion within a confined region and very slow movement between the regions. This results in errors that are systematic in the sense that they have a definite sign and do not diminish with an increasing number of walks. The errors arise as a consequence of the finite length of the walks and hence diminish with increasing walk length, disappearing in the limit of infinitely long walks. A prototypical example is the double well problem where the wells are separated by a large barrier. For sufficiently low temperatures, the random walker is unable to cross over the barrier within the duration of the walk and hence never samples from the other well. Valleau and Whittington³ refer to this phenomenon as *quasi-ergodicity*. The problem is especially insidious because it is difficult to detect; the quasi-ergodic simulation can give all appearances of being a good one, having low asymptotic variance and rapid convergence, yet yield results that are completely wrong.

Despite the clear need for ensuring ergodicity in Monte Carlo simulations, there have been relatively few direct studies on identifying and reducing quasi-ergodicity. Within the context of free-energy determinations, attempts have been made to widen the distribution using umbrella sampling ap-

proaches.⁴ In a similar context, Voter⁵ has developed methods for determining the Helmholtz free-energy difference between two systems widely separated in configuration space. Although quite powerful, Voter's method requires the predetermination of the regions of configuration space that must be accessed. Recently, Cao and Berne⁶ have developed antiforce bias and variable step algorithms that substantially reduced quasi-ergodicity in the one-dimensional double well potential they tested by accelerating the barrier crossings. Their methods are distinguished by the use of techniques that increase the likelihood of walking over barriers. We have instead concentrated on the development of algorithms that are characterized by "jumping over" barriers with attempted moves made directly to the wells. Because Metropolis stepsizes ranges are implicitly scaled with temperature to maintain the usual move acceptance rate of approximately 50%, quasi-ergodicity in a given system becomes an issue only at low temperatures where the stepsize is too small to surmount the high barrier. We propose that a low temperature random walker making the usual Metropolis moves within a largely confined region of configuration space be occasionally transported to another region by periodically attempting jumps to another Metropolis walker whose temperature is sufficiently high to fully access the whole configuration space. This jump-walking (or *J*-walking) scheme couples the typically small stepsizes necessary for adequately sampling within a region of configuration space to the large scale moves necessary to escape the bottlenecks. The Boltzmann distribution generated by the high temperature walker (*J*-walker) becomes the sampling distribution for the low temperature walker's attempted jumps. Because its distribution's peaks correspond to the potential minima, the *J*-

walker's motion remains biased about the minima, greatly increasing the likelihood an attempted jump will be accepted. Furthermore, the *J*-walker finds these regions automatically; no *a priori* knowledge of the potential's wells and barriers is necessary.

This coupling of different distributions is similar in spirit to single and multiple histogram methods⁷ where the equilibrium probability distribution at one value of some parameter is used to determine the probability distribution at another value of the parameter. While these methods have been developed to extend the parameter space obtainable from a particular simulation and hence reduce the total calculation time for the applied range, they do not treat quasi-ergodicity *per se*, which is the intent of *J*-walking. The two methods are complementary and could in principle be combined.

The easiest way to implement the *J*-walker algorithm is to run the two walkers in tandem. However, this can lead to large correlations between the walkers that also results in large systematic errors. There are several methods useful for breaking these correlations. For example, the *J*-walker can be moved an extra number of steps whenever a jump is attempted, or several *J*-walkers can be run simultaneously with jumping attempted to each in a random fashion. We have found in general that the most efficient method is to generate a high temperature distribution of adequate size beforehand and store it in an array so that jumps can be attempted to a configuration within the distribution simply by generating a random index.

We have tested the *J*-walker method on a number of simple systems. In Sec. II we introduce for illustrative purposes a simple model of the prototypical one-dimensional double well potential based on a quartic polynomial. The potential consists of one well of fixed depth containing the global minimum and a second well of variable depth. The variable depth of the second well allows us to examine in a systematic manner quasi-ergodic behavior as a function of well depth as well as a function of temperature. A solution of high accuracy can be easily obtained using standard numerical integration so that errors arising from quasi-ergodicity can be quantified. We next develop our motivation for the *J*-walker method by examining some simple jumping schemes before presenting the method in detail. We systematically test the method in Sec. III, beginning with an analysis of the functional dependence of the classical average potential energy for the double well on the temperature and variable well depth. Although useful heuristically, one dimensional systems can be misleading since good results can be obtained from algorithms that are useless in typical Monte Carlo multidimensional applications. Hence we further test the utility of the method by extending the double well potential to several dimensions. In Sec. IV, we apply the *J*-walker method to a multiparticle cluster system consisting of classical argon atoms bound by pairwise Lennard-Jones potentials. We illustrate the great increase in the rate of convergence engendered by the method by obtaining good results even from clusters that have been initialized randomly. Finally, in Sec. V we summarize our experience and suggest future applications.

II. THEORY

A. Quasi-ergodicity in the double well potential

We begin to study quasi-ergodicity with a one dimensional double well potential defined by the quartic polynomial

$$V(x) = 3\delta x^4 + 4\delta(\alpha - 1)x^3 - 6\delta\alpha x^2 + 1, \quad (1)$$

where

$$\delta = \frac{1}{2\alpha + 1}. \quad (2)$$

This function has a fixed minimum located at $x = 1$, a fixed barrier of unit height at $x = 0$ and a variable minimum located at $x = -\alpha$. By restricting α to the range $0 < \alpha < 1$, the $x = 1$ well remains the global minimum and the potential varies from a single well and barrier ($\alpha = 0$) to a symmetric double well ($\alpha = 1$). We can recast α in terms of a more physically meaningful potential parameter by considering the ratio of the well depths,

$$\gamma = \frac{V(0) - V(-\alpha)}{V(0) - V(1)} = \alpha^3 \left(\frac{\alpha + 2}{2\alpha + 1} \right). \quad (3)$$

Since the barrier height is fixed at unity, γ is also a direct measure of the variable well depth. For a given γ , α can be easily obtained from Eq. (3) by simple iteration. Figure 1(a) shows the potentials for some representative values of γ .

The problems arising from quasi-ergodicity in these potentials can be seen by examining their classical average potential energy,

$$\langle V \rangle = \frac{\int V e^{-\beta V} dx}{\int e^{-\beta V} dx}, \quad (4)$$

as a function of γ and the temperature parameter $\beta = 1/kT$. In the usual Metropolis method,¹ a random walker samples the configuration space making moves from an initial coordinate x_i to a final coordinate x_f with a probability of acceptance,

$$p = \min\{1, q(x_f|x_i)\}, \quad (5)$$

where

$$q(x_f|x_i) = \frac{T(x_i|x_f)\rho(x_f)}{T(x_f|x_i)\rho(x_i)}, \quad (6)$$

$\rho(x) = Z^{-1} \exp\{-\beta V(x)\}$ is the Boltzmann distribution with Z the standard configuration integral, and $T(x'|x)$ the sampling distribution. Figure 1(b) displays the Boltzmann distributions for the $\gamma = 0.9$ potential for some representative values of β .

The sampling distribution is usually generated⁸ from uniform deviates ξ over a finite stepsize range Δ to give

$$T(x'|x) = \begin{cases} \frac{1}{\Delta} & \text{for } |x' - x| < \frac{\Delta}{2} \\ 0 & \text{otherwise,} \end{cases} \quad (7)$$

and

$$q(x'|x) = \exp\{-\beta [V(x') - V(x)]\}. \quad (8)$$

Attempted steps are generated according to $x' = x + (\xi - \frac{1}{2})\Delta$. The maximum stepsize $\Delta/2$ is usually adjusted to give acceptance probabilities of approximately

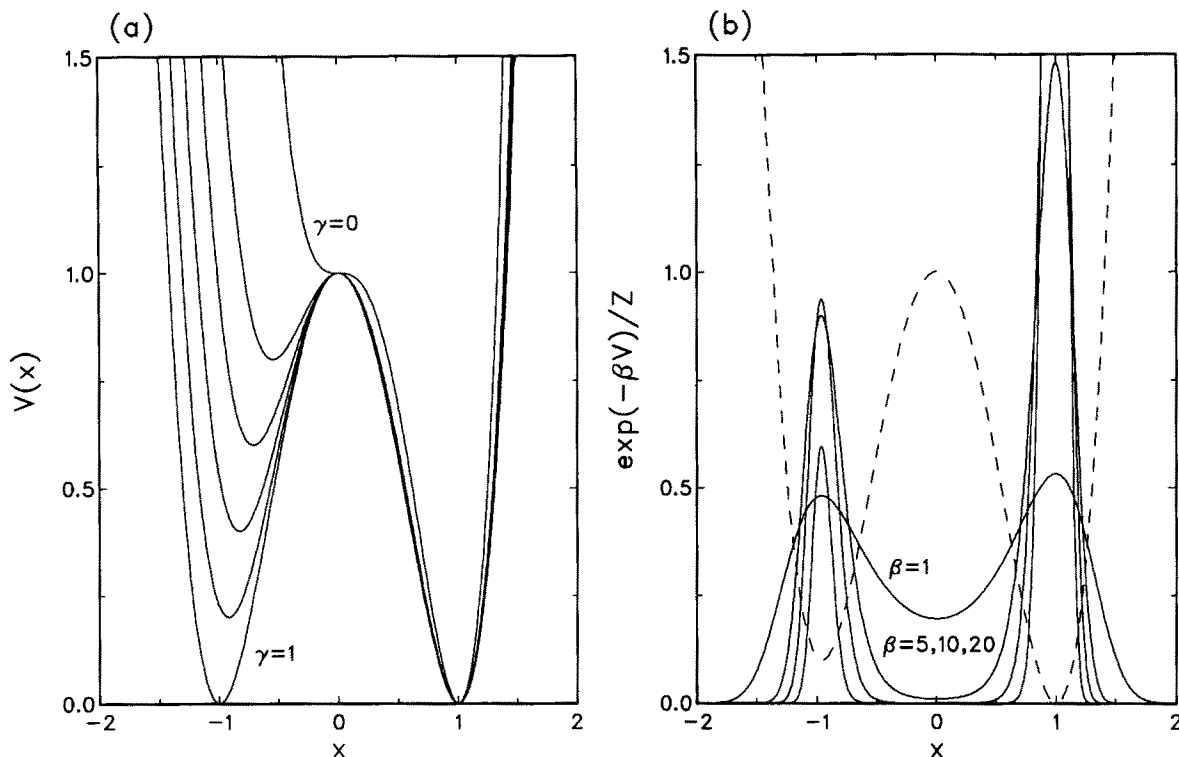


FIG. 1. (a) One dimensional double well potentials for $\gamma = 0, 0.2, \dots, 1$, where γ is the depth of the variable well as given by Eq. (3) in the text. The potential with $\gamma = 0$ corresponds to a single well with a barrier, while $\gamma = 1$ represents the symmetric double well. (b) Boltzmann distribution functions for the $\gamma = 0.9$ potential at $\beta = 1, 5, 10$ and 20 , where $\beta = 1/kT$ is the temperature parameter. The dashed curve is the potential.

50%. The required size decreases with increasing β . We have found that for potentials nearly quadratic about their minima, the stepsize range has a temperature dependence $\Delta \sim c/\sqrt{\beta}$, where c is a constant dependent upon the well curvature. This dependence on β can lead to quasi-ergodic behavior for systems with barriers too high to ascend with a sequence of several small steps and too broad to jump with a single step; the walker becomes trapped in one of the wells and never samples the remaining configuration space. This can be seen in Fig. 1(b). The energy for $\beta = 1$ is comparable to the barrier height and hence a significant fraction of the distribution function is found in the barrier region so that the walker has little trouble moving from well to well. An acceptance rate of approximately 50% can be obtained with a stepsize range of $\Delta = 2.5$, which is larger than the distance between the wells, $1 + \alpha$. At $\beta = 20$, the probability of penetrating the barrier with $\Delta = 0.5$ is so low that only the distribution in one of the wells can be effectively sampled.

The direction of the error resulting from this quasi-ergodic behavior depends upon the walk initialization. If the walks originate at the global minimum $x = 1$ then $\langle V \rangle$ will be too low since the distribution associated with the higher energy α well will be insufficiently sampled. The magnitude of the error depends upon the fraction of the distribution associated with the α well and hence varies with both γ and β . Likewise, initialization in the α well results in values of $\langle V \rangle$ that are too high. This behavior can be used to recognize quasi-ergodicity as can be seen in Fig. 2, which shows $\langle V \rangle$ obtained from standard Metropolis sampling as a function of β at constant γ [Fig. 2(a)] and as a function of γ at constant

β [Fig. 2(b)]. The two Monte Carlo curves represent initialization at the global minimum and random initialization. The third curve is the numerically exact solution obtained from conventional integration routines.⁹ Figure 2(a) reveals quasi-ergodic behavior in the $\gamma = 0.9$ potential beginning at $\beta \approx 6$. The error occurring with initialization fixed at the global minimum is mostly lower than that occurring with random initialization and decreases with increasing β at high β as the fraction of the distribution associated with the α well decreases. Figure 2(b) for $\beta = 10$ also shows a larger error for random initialization, with quasi-ergodic behavior beginning at $\gamma \approx 0.3$. For the fixed initialization case, the onset becomes noticeable at a greater value ($\gamma \approx 0.5$). Although the error is substantially less in this case, it is still significant (about 30%). The good agreement for low γ is due to the tiny fraction of the total distribution contributed by the small well. Note that quasi-ergodicity is not a problem for the classical average potential energy of the symmetric double well ($\gamma = 1$) since sampling from either half of the distribution gives the same result.

Table I shows the effects of quasi-ergodicity in the convergence for the $\gamma = 0.9$ potential for both initialization at the global minimum and for random initialization. The results obtained for fixed initialization for $\beta = 20$ illustrate the deceptive nature of quasi-ergodicity. The standard deviation has the expected $1/\sqrt{N}$ decrease with increasing walk length N , but $\langle V \rangle$ is converging to the wrong value. Quasi-ergodic behavior is more readily recognized when the walks are randomly initialized since the walkers can be confined in either well. This leads to a bimodal distribution in $\langle V \rangle$, and hence

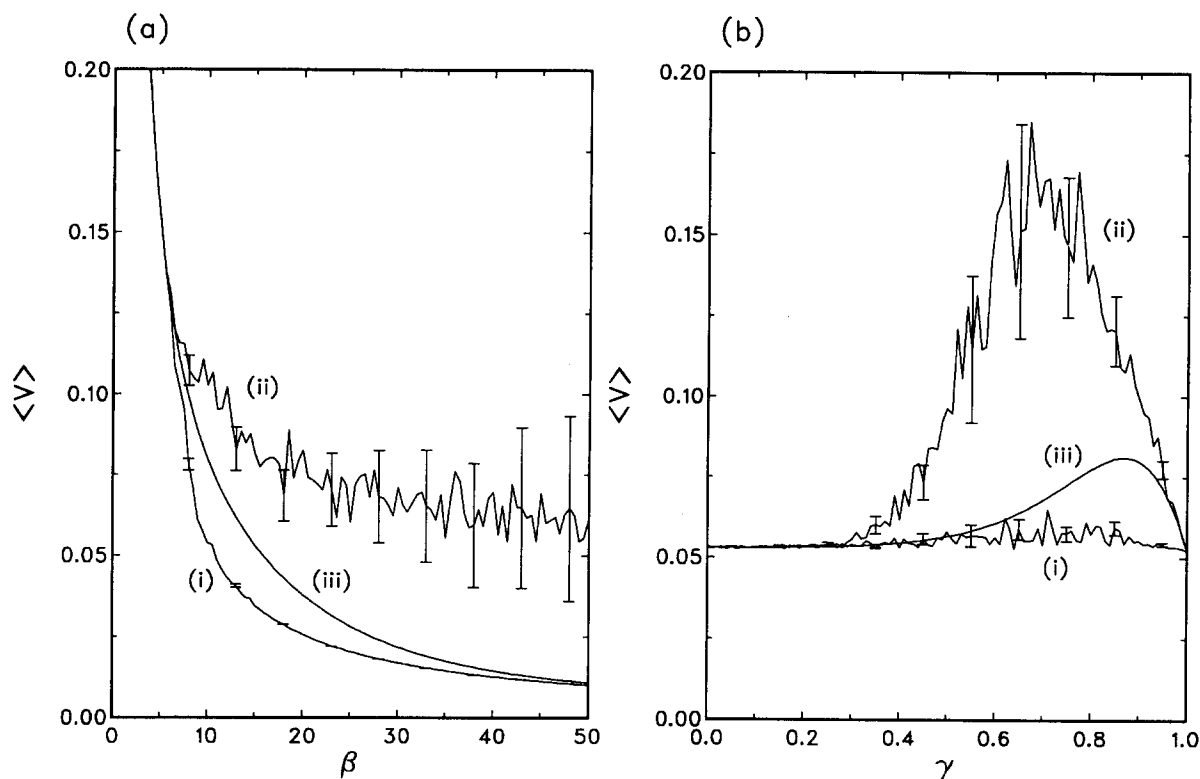


FIG. 2. Quasi-ergodicity in the double well potential. (a) The classical average potential energy $\langle V \rangle$ for the $\gamma = 0.9$ potential as a function of β . (b) $\langle V \rangle$ as a function of γ for $\beta = 10$. In both plots, curve (i) is the average $\langle V \rangle$ obtained from 100 independently initialized Metropolis walks, each originating at the global minimum $x = 1$ and consisting of 500 warmup steps followed by 10^4 steps with data accumulation. Curve (ii) was similarly obtained except that each walk was initialized at a random x and curve (iii) is the numerically exact solution. Some representative error bars have been included. As in subsequent figures, these correspond to single standard deviations.

the standard deviation reflects the difference in $\langle V \rangle$ for the two wells as well as the usual contributions from the variation about each well, which have the typical $1/\sqrt{N}$ behavior. For very low temperatures, the difference in $\langle V \rangle$ dominates and the standard deviation becomes nearly constant. This is evident in Table I where the relative standard deviation for $\beta = 20$ is constant to within 0.1% over a change in N of five orders of magnitude. At $\beta = 10$, the temperature is high enough to allow occasional barrier crossings in walks longer than about 10^4 steps, as can be seen by the decrease in the relative error for both cases. Again the lack of $1/\sqrt{N}$ behavior in the standard deviation indicates the persistence of some quasi-ergodicity in walks of $N = 10^6$.

B. Simple jumping schemes

The simplest way to deal with quasi-ergodicity in Metropolis sampling is to merely increase the stepsize range Δ sufficiently so that some of the steps become large enough to jump over the barriers. While this allows effective motion between the constricted regions of configuration space, the fraction of steps small enough to effectively sample the distribution within a region becomes too low. For example, we found using a stepsize range of $\Delta = 4$ for the one-dimensional double well at $\beta = 10$ gave good results despite step rejection rates of greater than 90%. This inherent inefficiency makes this scheme untenable for multidimensional systems where distributions are often typified by sparsely located

TABLE I. Quasi-ergodic behavior in convergence. The error and σ_R columns are the relative error ($Error = [calculated - exact] / calculated$) in the classical average potential energy $\langle V \rangle$ and the relative standard deviation ($\sigma_R = \sigma / exact$), respectively, for the $\gamma = 0.9$ potential, with each entry obtained from 100 independently initialized walks. The walks with fixed initialization all originated at the global minimum $x = 1$.

Steps	Fixed initialization				Random initialization			
	$\beta = 20$		$\beta = 10$		$\beta = 20$		$\beta = 10$	
	Error	σ_R	Error	σ_R	Error	σ_R	Error	σ_R
10^2	-0.31209	0.02229	-0.35889	0.02347	0.95714	0.13453	0.22280	0.06782
10^3	-0.33840	0.00531	-0.35054	0.01425	0.97497	0.13437	0.27064	0.06288
10^4	-0.32731	0.00185	-0.31786	0.01706	1.10802	0.13389	0.21808	0.06146
10^5	-0.32460	0.00060	-0.19033	0.02862	1.00336	0.13413	0.08846	0.04440
10^6	-0.32494	0.00020	-0.04045	0.01524	1.16313	0.13318	0.01445	0.02025

sharp peaks. A big improvement can be made by coupling the jumping scheme to standard Metropolis sampling. The stepsize range is selected with the usual criterion that the sampling have a acceptance rate of about 50%. Occasional jumps to the other well can then be attempted by magnifying the stepsize with

$$\Delta_J \rightarrow k_\alpha \Delta \text{ for } 0 \leq \xi_J < P_J, \quad (9)$$

where k_α is a constant dependent upon the distance between the wells and P_J is a constant that defines the jumping probability. Although this modified scheme provides good sampling within a well, it is still too inefficient since few of the jumps land near the other distribution peak.

A different method that greatly increases the jump acceptance rate is based upon the observation that at low temperatures the walker is confined to a small region about a minimum. Hence jumps to the other well have a greater likelihood of success if the step is translated symmetrically an amount equal to the distance between the wells,

$$x' = x + \left(\xi - \frac{1}{2} \right) \Delta + d(\xi_J), \quad (10)$$

where

$$d(\xi_J) = \begin{cases} -d_\alpha & \text{for } 0 \leq \xi_J < \frac{1}{2} P_J \\ d_\alpha & \text{for } \frac{1}{2} P_J \leq \xi_J < P_J \\ 0 & \text{otherwise,} \end{cases}$$

and $d_\alpha = 1 + \alpha$ is the distance between the wells. While this method works very well for the one-dimensional potential and continues to give good results when extended to the multidimensional case, it suffers from the severe limitation of having to know the minima locations beforehand. We will not pursue it here further except to note that its low computational overhead and ease of implementation can make it useful for certain problems.

C. J -walker method

The $\beta = 1$ Boltzmann distribution function shown in Fig. 1(b) encompasses the advantages of both simple jumping methods while overcoming their deficiencies. The range is large enough to cover both wells (equivalent to a stepsize range $\Delta_J \approx 3$) and much of the distribution is concentrated about the minima, implying a high likelihood of jump acceptance. Since this distribution can be easily obtained with a Metropolis walker (there are no quasi-ergodic problems at this temperature), prior knowledge of the locations of the potential minima is unnecessary. These features can be utilized then by having the low temperature walker attempt occasional jumps to a high temperature walker, or J -walker. The sampling distribution for the jumps is just the Boltzmann distribution at the higher temperature

$$T_J(x'|x) = Z^{-1} \exp\{-\beta_J V(x')\} \text{ for } 0 \leq \xi_J < P_J. \quad (11)$$

Using Eq. (6) gives

$$q_J(x'|x) = \exp\{(\beta_J - \beta)[V(x') - V(x)]\}. \quad (12)$$

In the limit $\beta_J \rightarrow 0$, $q_J(x'|x)$ reduces to the standard Metropolis form given in Eq. (8). Because the high temperature distribution broadens as β_J decreases, the J -walker method essentially reduces to simple jumping with a large stepsize range Δ_J in this limit. In the limit $\beta_J \rightarrow \beta$, $q_J(x'|x) \rightarrow 1$ since the low temperature walker is now effectively sampling from its own distribution.

There are two complementary methods of generating the J -walker distribution. The first method runs the J -walker in tandem with the low temperature walker with jumps attempted by

$$x' = x_J \text{ for } 0 \leq \xi_J < P_J. \quad (13)$$

As will be seen in the next section, this simple prescription leads to correlation errors that need to be overcome by taking extra J -walker steps whenever a jump is attempted or by employing several J -walkers in tandem and accessing them in random. A better alternative is to run the J -walker beforehand and save a given number of points from the distribution in an external array for subsequent sampling by the low temperature walker,

$$x' = x_J(\xi N) \text{ for } 0 \leq \xi_J < P_J, \quad (14)$$

where N is the size of the array. The former method has the advantage of lower memory requirements at a cost of increased computational time while the latter has the advantage of lower computational overhead at a cost of increased storage requirements.

III. RESULTS

A. One dimensional results

The first step in implementing the J -walker method is to find a suitable temperature for the high temperature walker. The temperature must be high enough to avoid quasi-ergodic behavior in the J -walker but low enough to provide a sufficient jump acceptance rate. Figure 3(a) shows the relative error in the classical average $\langle V \rangle$ as a function of the J -walker temperature parameter β_J for the $\gamma = 0.9$ double well potential with the low temperature walker at $\beta = 10$, for the case of the J -walker running in tandem with the low temperature walker. Results consistent with the standard deviation were obtained for both fixed and random initialization for β_J up to about 3, where correlation between the two walkers became a factor (both walkers were initialized at the same point). These curves were generated by allowing the J -walker to take an extra 100 steps to reduce the correlation whenever a jump was attempted; increasing the number of extra steps to 1000 increased the useful range for β_J up to about 5. This dependence on the number of extra J -walker steps can be seen in Fig. 3(b), which shows the relative error in $\langle V \rangle$ as a function of the number of extra steps for $\beta_J = 2$ and 3. The curve shows an exponential-like decrease in the error with an increasing number of extra steps. The number of extra steps required to remove the correlation increases with β_J ; it becomes prohibitively large as β_J approaches the onset of quasi-ergodicity at $\beta \approx 6$ [see Fig. 2(a)] because the time scales for intrawell and interwell motion in the J -walker become too different. Quasi-ergodicity can be seen in the J -

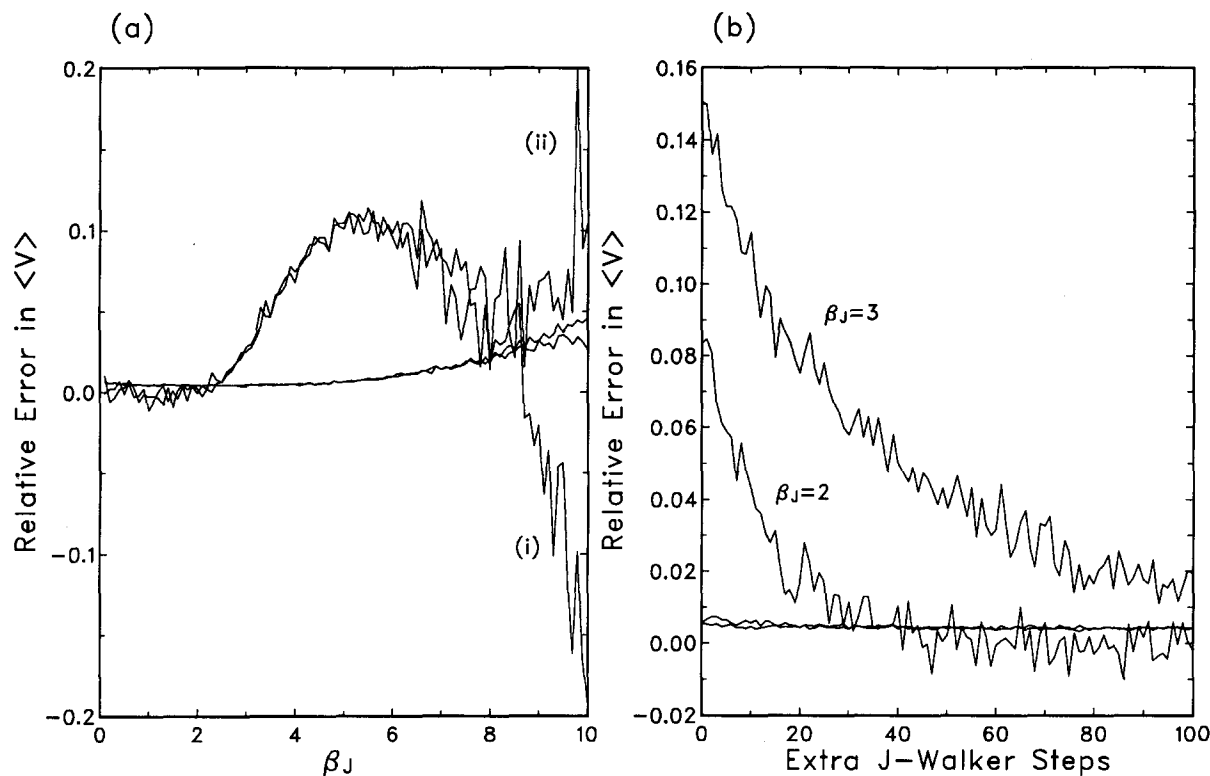


FIG. 3. J -walker temperature dependence for the case of the tandem walker. (a) The relative error (as defined in Table I) in $\langle V \rangle$ as a function of the J -walker temperature parameter β_J for the $\gamma = 0.9$ double well potential with the low temperature walker at $\beta = 10$. As in Fig. 2, the average $\langle V \rangle$ was obtained from 100 independently initialized walks consisting of 500 warmup steps followed by 10^4 steps with data accumulation. For curve (i), initialization was fixed at the global minimum, while for curve (ii), it was random. The two smoother curves are the relative standard deviations (also defined in Table I) for the 100 walks. The jumping probability was $P_j = 0.1$ and correlation between the walkers was reduced by allowing the J -walker 100 extra steps whenever a jump was attempted. (b) The reduction in the correlation error as a function of the number of extra J -walker steps for $\beta_J = 2$ and 3. Again, the two smoother curves are the relative standard deviation.

walker for β_J larger than about 7 by the divergence of the two error curves in Fig. 3(a).

The correlation problem can be overcome by generating the J -walker distribution beforehand and periodically storing points for subsequent sampling. Correlation between the walkers is no longer an issue since the steps are now accessed randomly and not sequentially. Figure 4(a) shows the relative error in $\langle V \rangle$ as a function of β_J for this case. Results consistent with the standard deviation were obtained for β_J up to the onset of quasi-ergodicity, at which point the systematic errors in the J -walker distribution lead to statistically incorrect sampling during jump attempts. Because the J -walker distribution was initialized randomly at each β_J , the quasi-ergodicity is made manifest here by the increasingly large fluctuations in the error for $\beta_J > 7$. Starting the low temperature walker at the global minimum gave similar results since the interwell motion is dominated by the J -walker distribution.

The elimination of the correlation error with the use of the external distribution is at the cost of the memory necessary to store the distribution. Hence some attention must be paid to minimizing the array size as much as possible. For example, successive moves should not be stored since the differences between them are minor. By storing every tenth or hundredth move, a more representative distribution can be generated for a given array size. The distribution will be

more coarse grained of course, but since it is used primarily for transporting the low-temperature walker to different regions of configuration space, this is incidental. Figure 4(b) displays the dependence on the J -walker distribution size for the $\gamma = 0.9$ potential. Each distribution consists of points sampled every 100 moves. It is clear for this case that a distribution of the order of 10^4 points is necessary to give proper convergence for 100 low temperature walks of 10^4 steps/walk each.

Table II further examines the dependence of the convergence on the J -walker distribution size for the $\gamma = 0.9$ potential. As in Fig. 4, each J -walker distribution of a given size was generated from one walk at $\beta_J = 2$ by sampling every 10 or 100 steps and was subsequently used for 100 walks at $\beta = 10$. The relative error in the average potential energy at β_J , $\langle V \rangle_J$, for each J -walker distribution is listed in the J -error column of Table II. The substantial reduction in quasi-ergodicity in the low temperature walks is evident in their standard deviations, which now show the expected $1/\sqrt{N}$ decline, even when distributions as small as 10^3 points were used (compare this with the random initialization case in Table I). Although the fluctuations are large, the relative error for each set of low temperature walks appears consistent with the respective standard deviation when the number of steps is of the same order as the J -walker distribution size or less. Decreasing the J -walker sampling rate to one point

TABLE II. J -walker convergence for one-dimensional potential. The Error and σ_R columns are the relative error in the classical average potential energy $\langle V \rangle$ and the relative standard deviation, respectively, for the $\gamma = 0.9$ potential, with each entry obtained from 100 independently initialized walks. The J -walker distributions were at $\beta_J = 2$, and the jumping probability was $P_J = 0.1$.

J -walker distribution		Random initialization			
Size	Steps/Pt.	J -error	Steps/Walk	Error	σ_R
10^3	10	0.01500	10^2	-0.01688	0.03807
		0.04951	10^3	0.00604	0.01449
		0.06461	10^4	0.01530	0.00464
		0.00931	10^5	-0.00698	0.00140
		-0.03467	10^6	-0.03593	0.00043
		-0.01664	10^2	0.00657	0.03754
10^4	10	0.01608	10^3	-0.02113	0.01377
		0.00133	10^4	0.00430	0.00414
		0.01399	10^5	0.01266	0.00154
		0.02075	10^6	0.00026	0.00041
		0.00603	10^2	-0.07335	0.03780
		-0.00487	10^3	0.00170	0.01399
10^5	10	0.00031	10^4	0.00080	0.00422
		0.00041	10^5	-0.00856	0.00116
		0.00093	10^6	-0.00341	0.00040
		-0.06586	10^2	-0.08968	0.03251
		-0.03217	10^3	0.02065	0.01345
		0.07866	10^4	-0.00868	0.00422
10^3	10^2	-0.05580	10^5	-0.00511	0.00123
		0.00230	10^6	-0.01528	0.00044
		0.00216	10^2	0.05088	0.04208
		-0.01872	10^3	0.00889	0.01362
		0.02146	10^4	-0.00793	0.00405
		0.01333	10^5	0.01654	0.00150
10^4	10^2	-0.00299	10^6	-0.00934	0.00037
		0.00181	10^2	0.03518	0.03851
		0.00385	10^3	0.01702	0.01290
		-0.00237	10^4	0.00695	0.00384
		0.00072	10^5	0.00035	0.00115
		0.00430	10^6	-0.00314	0.00042

per 100 steps mostly improved the accuracy of $\langle V \rangle$, but did not improve the accuracy in the low temperature walks. Even though a J -walker distribution of size 10^5 sampled every tenth step and a distribution of size 10^4 sampled every hundredth step were both generated from J -walks of 10^6 steps, the distribution with the larger size gave the better results. The large fluctuations in the error in $\langle V \rangle$, for J -walker distributions of a given size and the lack of correlation with the resultant error in $\langle V \rangle$ for the low temperature walks indicates that the error in $\langle V \rangle$ is not a good indicator of the quality of the J -walker distribution. It should also be noted that there is a strong correlation between systematic errors in the J -walker distribution and subsequent systematic errors in the low temperature walks. For example, using a distribution generated at $\beta_J = 10$ from a quasi-ergodic walk that is confined to one well will result in systematic errors for a $\beta = 20$ walker since all jumps are within the confined region.

Figure 5 shows the dependence on the jumping probability P_J for the $\gamma = 0.9$ potential where the J -walker distribution was generated at $\beta_J = 3$. A dramatic drop in the relative error in $\langle V \rangle$ occurs even for $P_J < 0.01$ (the error for

$P_J = 0$ was 0.3) because only a few successful jumps are needed for movement between the two wells. Results consistent with the standard deviation occur for a jumping probability greater than about 0.03. Since the computational cost of attempting a jump is about the same as attempting a regular Metropolis move (the computational overhead is mostly in the generation of the J -walker distribution), we usually used a jumping probability of $P_J = 0.1$.

Figure 6 shows the results obtained over the entire range of γ for a low-temperature walker at $\beta = 10$ using J -walking from an external distribution at $\beta_J = 2$. Figure 6(a) shows $\langle V \rangle$ as a function of γ and Fig. 6(b) the relative error in $\langle V \rangle$. Again, the fluctuations in the error are consistent with the standard deviation throughout. Comparison with Fig. 2(b) reveals the substantial improvement rendered by the J -walking. Similarly good results were obtained at the same β_J using a tandem J -walker taking 100 extra steps for each jump attempt.

B. Multidimensional results

While the one-dimensional case is useful for heuristic purposes, it is of limited use as a test since a method that gives good results in one dimension can fail badly when extended to several dimensions. The major advantage of the one-dimensional case is that the exact solution can be determined to high accuracy, allowing for a quantitative determination of the errors. We can maintain this advantage by forming a D -dimensional version of the double well potential,

$$V(x) = \sum_{k=1}^{D_w} V_1(x_k) + \sum_{k=D_w+1}^D x_k^2, \quad (15)$$

where $V_1(x_k)$ is the one-dimensional double well potential given by Eq. (1) and D_w is the number of coordinates having this potential. Since the number of wells in the potential hypersurface is 2^{D_w} , the addition of $D - D_w$ harmonic oscillator modes allows the number of dimensions to be increased independently of the number of wells. The global minimum well depth remains unity, while the other wells have depths of $\gamma, \gamma^2, \gamma^3, \dots, \gamma^{D_w}$. The average potential energy becomes

$$\langle V(x) \rangle = D_w \langle V_1(x) \rangle + (D - D_w) \frac{\beta}{2}. \quad (16)$$

Although this form still allows $\langle V(x) \rangle$ to be calculated to high accuracy by numerical integration, the separability of the modes implies the Monte Carlo results will be essentially similar to the one-dimensional Monte Carlo results if the Metropolis moves are attempted on individual coordinates. So better to mimic typical Monte Carlo simulations, attempted moves were made on all D coordinates in concert.

Figure 7(a) depicts the multidimensional version of Fig. 4(a), which plots the relative error in $\langle V \rangle$ as a function of β_J . While the two curves are similar for higher β_J , the multidimensional curve shows a large systematic error for very low β_J that is not present in the one-dimensional case. As $\beta_J \rightarrow 0$, the J -walker distribution broadens considerably so that it becomes essentially similar to the simple jumping scheme described by Eq. (9), which relies only on an expanded stepsize range. As the effective stepsize range in-

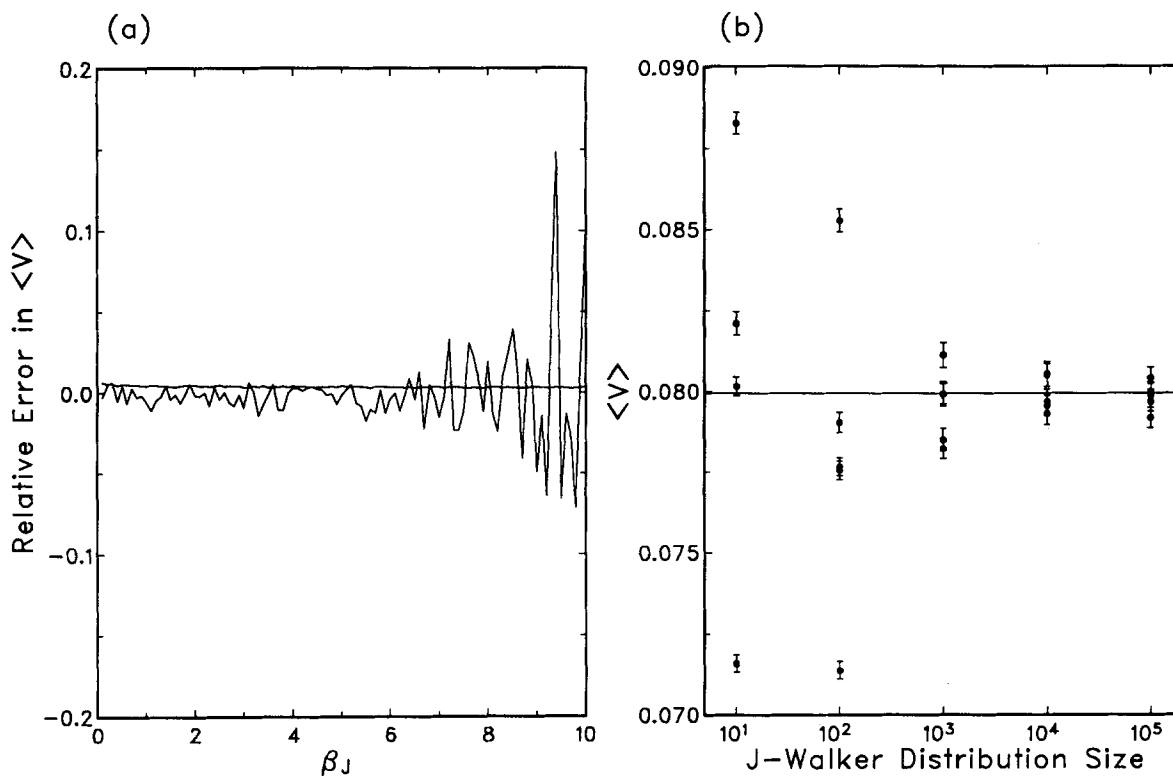


FIG. 4. *J*-walker temperature dependence for the external distribution case. (a) The plot for the $\gamma = 0.9$ potential was obtained as in Fig. 3 except that configurations for attempted jumps were sampled from a *J*-walker distribution generated once at each β_J point from one long *J*-walk and subsequently used for all 100 walks at that point. The distribution consisted of 10^5 points with 100 steps taken per point. The two curves are the relative error in $\langle V \rangle$ and the relative standard deviation for the case of random initialization of each walk; initialization at the global minimum gave similar results since the *J*-walker was initialized randomly. (b) Dependence on the *J*-walker distribution size. Each point is the mean of 100 walks. All the distributions consist of points sampled every 100 steps after 500 warmup steps. The solid line is the numerically exact result.

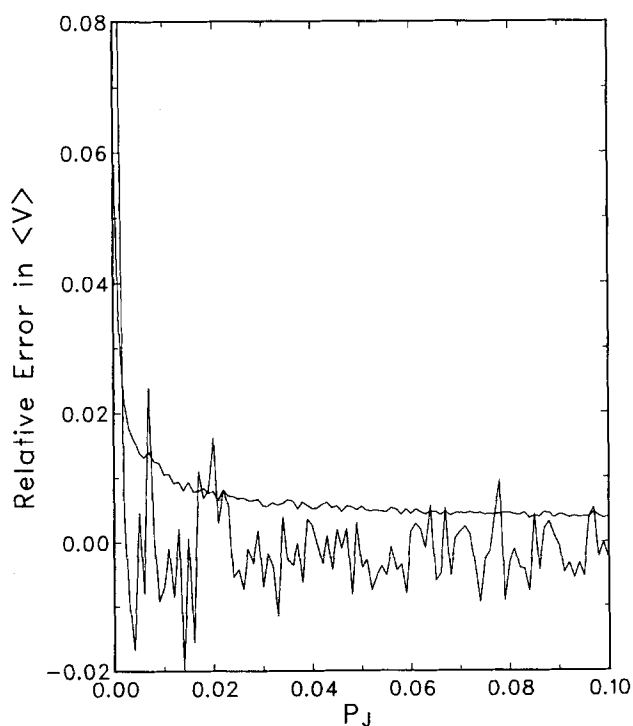


FIG. 5. Variation in the relative error in $\langle V \rangle$ with the jumping probability P_J for the external distribution case. The *J*-walker distribution was generated for the $\gamma = 0.9$ potential as in Fig. 4 for $\beta_J = 3$. Again, the two curves are the relative error and standard deviation for the case of random initialization. The error for no jumping is off scale at 0.3 for $P_J = 0$.

creases in the multidimensional case, the likelihood of accepting an attempted jump diminishes rapidly since most of the jumps will not be near a minimum. This can be seen Fig. 7(b), which plots the jump acceptance rate as a function of β_J for $D = 1, 5$ and 20. The good results obtained for the one-dimensional case for β_J as low as 0.1 can be seen to arise from the substantial jump acceptance still present at that point. This is an intrinsic feature of the one-dimensional double-well potential, since the constraint imposed by the single degree of freedom entails that almost any motion over the barrier has a good likelihood of being accepted. Note that this narrowing of the usable range in β_J , away from low β_J , with increasing dimension also diminishes the efficacy of the tandem walker since the number of extra moves required to break the correlation increases with increasing β_J .

Figure 8 shows the multidimensional analog of Fig. 6(b), which plots the dependence of the relative error in $\langle V \rangle$ on the variable well depth γ . Here results consistent with the standard deviation were obtained for all γ except for some points near $\gamma \approx 0.8$ where the error rose to about three standard deviations. The *J*-walking substantially reduced the quasi-ergodic error in this region, but did not entirely eliminate it. This illustrates the importance of still having to heed the usual caveats associated with standard Metropolis Monte Carlo when using *J*-walking—low asymptotic variance and rapid convergence are no guarantee of accuracy. While these fluctuations in the error are larger than expected relative to the standard deviations, they are unlike the sys-

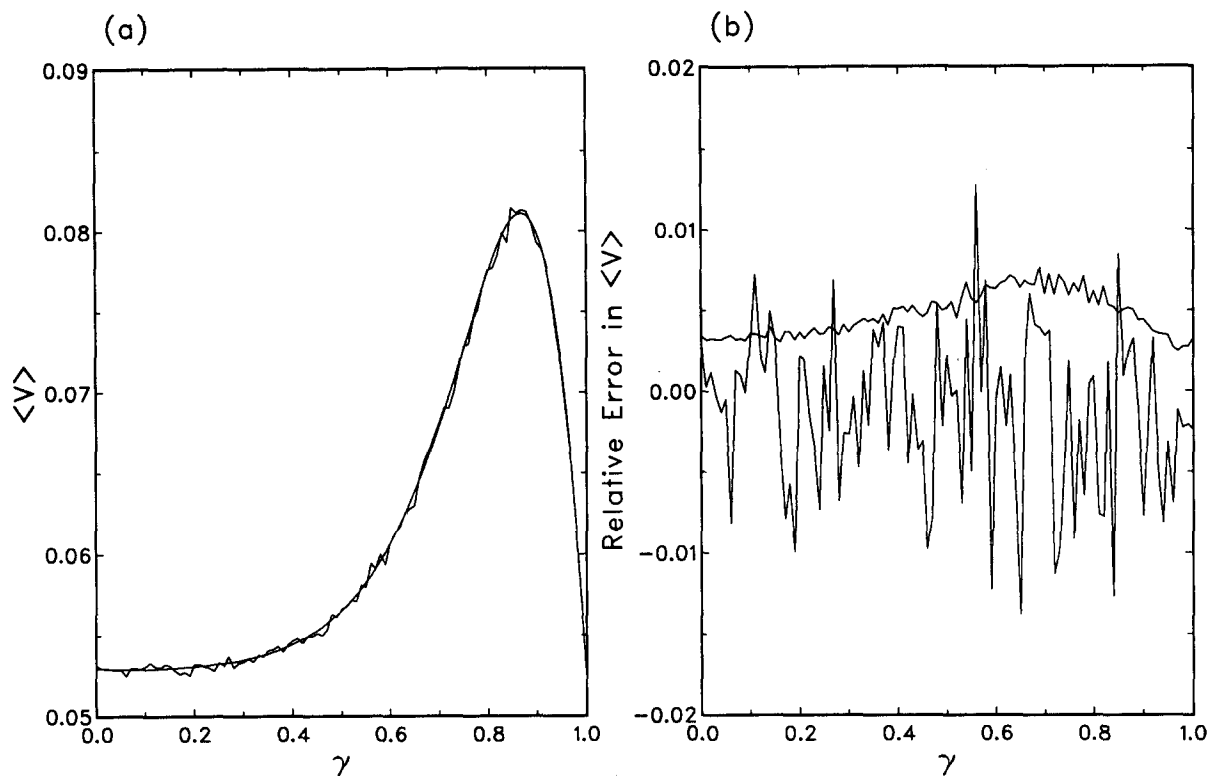


FIG. 6. Variation in $\langle V \rangle$ with the variable well depth γ . (a) As in Fig. 2(b), the plot shows the average $\langle V \rangle$ at $\beta = 10$ for 100 walks independently initialized from random configurations, each consisting of 500 warmup steps followed by 10^4 steps of data accumulation. The jumping probability was $P_j = 0.1$ and an external J -walker distribution generated as in Fig. 4 for $\beta_j = 2$ was used for each batch of walks. The smooth curve is the numerically exact result. (b) The relative error and standard deviation in $\langle V \rangle$ as a function of γ . Similar results were obtained for the same β_j using a tandem J -walker with 100 extra steps per jump.

tematic errors arising from quasi-ergodicity in standard Metropolis, which have a characteristic sign. As in the one-dimensional case shown in Table II, their magnitude is indicative of a J -walker distribution size that is insufficient for the given walk length.

IV. J -WALKING IN CLUSTERS

A. Background

Cluster melting is a prototypical example of a physical phenomenon where quasi-ergodicity can be troublesome for Monte Carlo simulations. Because of their small size, clusters do not have the sharp first-order transition characteristic of bulk melting, but instead show transitions occurring over a range of temperatures. Recent studies¹⁰ on the connection between structure and the dynamics of melting in small clusters have found the large separation in the time scales between intrawell and interwell motion (which is often indicative of quasi-ergodicity) occurring in the transition region. For temperatures well below the transition range, classical clusters are confined to the potential well associated with the isomer most stable for their size where they undergo small oscillations about the minimum energy configuration. They take on a very rigid, solid-like form, marked by small fluctuations in the bond length and no diffusion; there is essentially no motion between different wells. For temperatures well above the transition range, the clusters have easy access to numerous higher energy wells and

hence take on a very floppy, liquid-like form with large fluctuations and much diffusion. There is nearly unhindered motion over the barriers and between the wells. However, in the transition region, the solid-like forms exhibit highly anharmonic motions and are no longer rigid. The bond length fluctuations increase sharply and some diffusion appears as the clusters undergo hindered isomerization.

An especially interesting discovery has been that certain clusters can exhibit different melting and freezing temperatures, implying a temperature range where both solid-like and liquid-like forms coexist.^{11,12} This coexistence range then corresponds to the transition range. It decreases as the cluster size n increases, with the freezing and melting temperatures coalescing to the bulk transition temperature as n approaches infinity. This behavior is a very irregular function of n ; for example, argon clusters with "magic number" sizes $n = 7, 13,$ and 19 have especially high melting temperatures and pronounced coexistence ranges, while clusters of size $n = 8, 14$ and 20 melt at much lower temperatures and exhibit no coexistence. Beck and Berry¹⁰ have shown that in addition to the separation of time scales for the interwell and intrawell motions, another condition necessary for the coexistence of solid-like and liquid-like forms is the existence of a very stable potential minimum relative to the nearest accessible minima. For example, the minimum energy configuration for the Ar_{13} cluster is the icosahedron,¹³ which is very stable relative to any of its other isomers, while the most stable structure for Ar_{14} is the icosahedron with one atom

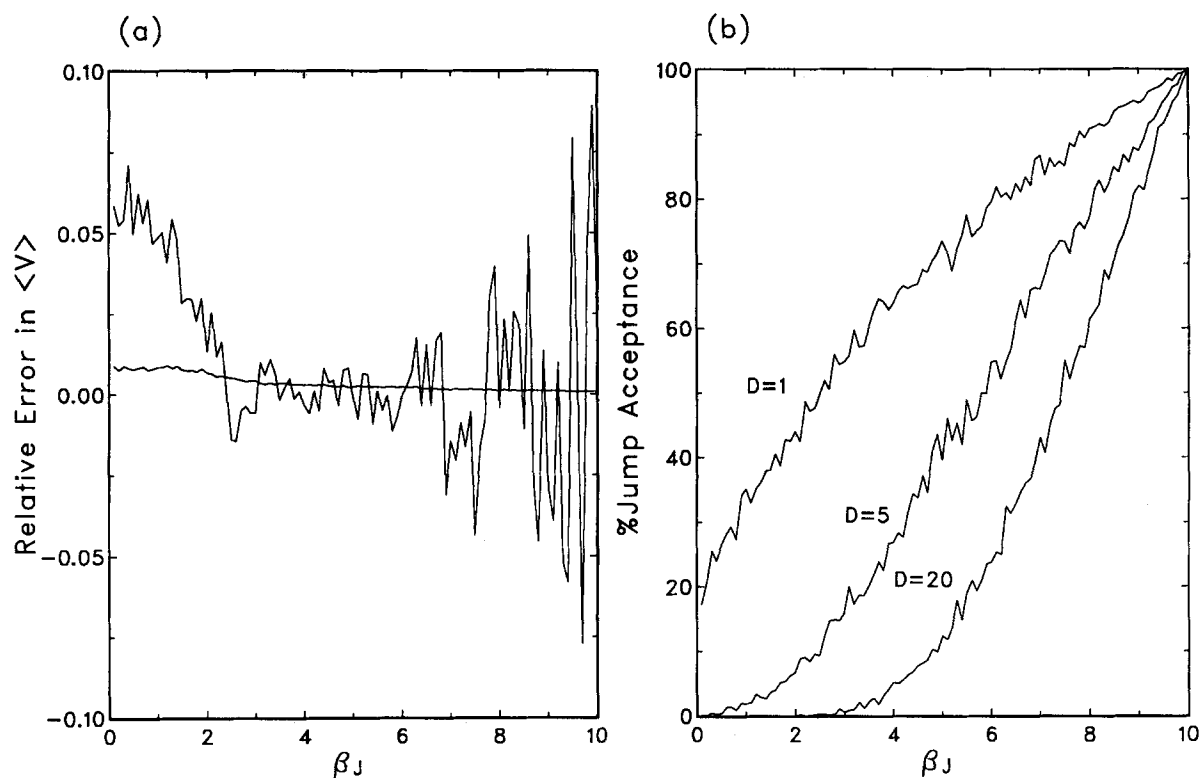


FIG. 7. (a) *J*-walker temperature dependence for the multidimensional extension of the double well potential given by Eq. (15), where $D = 20$ and $D_w = 3$ to give eight wells of depth $1, \gamma, \gamma^2$ and γ^3 , with $\gamma = 0.9$. The plot was obtained as in the one-dimensional case shown in Fig. 4(a) except that the *J*-walker distribution generated at each β_J consisted of 20 000 configurations sampled every 100 passes, the warmup period was 5000 passes, and the jumping probability was $P_J = 0.2$. Again, the two curves are the relative error in $\langle V \rangle$ and the relative standard deviation for the case of random initialization of each walk. (b) Jump acceptance for the two dimensions. Also included is the jump acceptance for the $D = 5$ case with $D_w = 3$; the *J*-walker distribution for this case consisted of 10^5 configurations sampled every 10 passes.

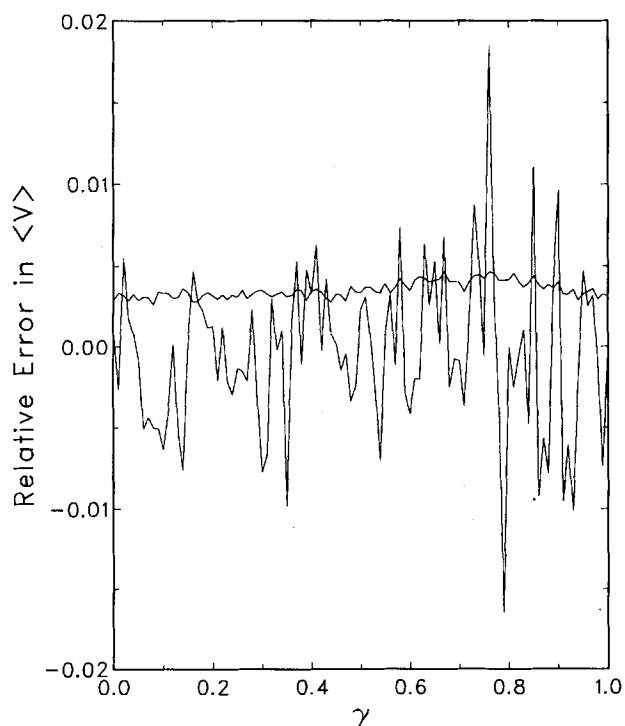


FIG. 8. Variation in the relative error in $\langle V \rangle$ with the variable well depth γ for the multidimensional potential. The plot was obtained as in the one-dimensional case shown in Fig. 6 except that the *J*-walker distribution consisted of 20 000 configurations sampled every 100 passes at $\beta_J = 3.5$. The warmup period was 5000 passes and the jumping probability was $P_J = 0.2$.

placed on one of its 20 equivalent faces. No coexistence range is found for Ar_{14} since nearly all of the accessible minima have the same energy with relatively small barriers between them. It is the large barriers associated with those clusters showing coexistence behavior that can be expected to produce appreciable quasi-ergodic behavior.

There has been some controversy in the literature concerning the solid-liquid transition in small rare-gas clusters. Various studies have produced conflicting conclusions on the sharpness of the transition and even whether two distinct forms are observable (Ref. 11 gives a detailed account). Some studies on Ar_{13} that used standard Metropolis Monte Carlo have made mention of difficulties in obtaining proper convergence and need for longer runs in this region.^{14,15} In addition, quasi-ergodicity has also been a problem with molecular dynamics simulations of argon clusters in the transition region¹⁶ and has been postulated to account for differences obtained between Monte Carlo and molecular dynamics calculations on Ar_{13} .¹⁷ The fundamental nature of these discrepancies, occurring in such a simple system as a small rare-gas cluster, testifies to the difficulties inherent in computer simulations in the transition region.

B. Method

The clusters were modeled by the usual pairwise additive Lennard-Jones potential

$$V(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (17)$$

with $\epsilon = 119.4$ K and $\sigma = 3.405$ Å for argon.¹⁸ Small clusters are known to become unstable beyond a threshold temperature T_b that varies with n .^{19,20} For the Lennard-Jones potential under free volume conditions, the average energy vanishes in the limit of an infinite number of configurations. Hence the choice of boundary conditions can have a pronounced effect on the properties of small clusters.¹¹ We have followed Lee, Barker, and Abraham²¹ and confined the clusters by a perfectly reflecting constraining potential of radius R_c centered on the cluster's center of mass. For all our cluster runs, the constraining radius was set to 4σ .

Some studies^{11,12} have shown a coexistence region for Ar_{13} clusters for temperatures from approximately 24 to 41 K. A *J*-walker temperature of 50 K would appear at first glance to be suitable for generating a sampling distribution since it is high enough to be fully liquid-like (and presumably ergodic) but low enough for acceptable jump acceptance ratios for walkers in the coexistence range. The situation is, however, slightly more complicated. Fig. 9 shows the distribution of potential energies for temperatures of 20, 30, 40 and 50 K. A 50 K *J*-walker accesses configurations having energies that range from about -35 to -10 (in reduced units V/ϵ). This range has appreciable overlap with that of a 40 K walker, only slight overlap with a 30 K walker and no overlap with a 20 K walker. This behavior suggests the *J*-walker distribution needs to be generated in stages. Thus the 50 K distribution was generated from one long Metropolis walk consisting of 10^6 warmup moves and 5×10^7 moves where a configuration was stored every 1000 moves. The long length of the walk and the infrequent sampling were used to ensure full sampling of the configuration space and to reduce the correlations in the distribution. The

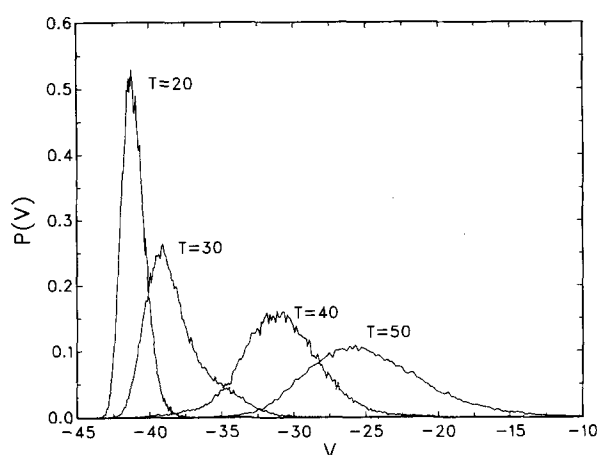


FIG. 9. Potential energy distributions for argon clusters of size $n = 13$, with V in reduced units. Each distribution consists of 50 000 configurations and was obtained from a single walk where each stored configuration was sampled every 1000 moves. The warmup period was 10^6 moves. The $T = 50$ K distribution was obtained using standard Metropolis Monte Carlo. The $T = 40$ K distribution was obtained by occasionally sampling from the $T = 50$ K distribution with $P_j = 0.1$. Similarly, the $T = 30$ K distribution was obtained by sampling from the $T = 40$ K distribution and the $T = 20$ K distribution by sampling from the $T = 30$ K distribution.

40 K distribution was then similarly generated except that occasional jumps were attempted to the 50 K distribution with a probability $P_j = 0.1$. The resultant 40 K distribution was used to generate the 30 K distribution, and the 30 K distribution used to generate the 20 K distribution. The distributions were stored piecewise in external files, each a size of about half the computer memory capacity, and used for subsequent calculations.²²

C. Results

Figure 10(a) shows the caloric curves, $\langle V \rangle$ as a function of temperature, for standard Metropolis and for *J*-walking using the distributions shown in Fig. 9. This plot illustrates the importance of initialization in standard Metropolis Monte Carlo. Curve (i) is the mean $\langle V \rangle$ obtained from 100 independently initialized walks, each consisting of 10^5 warmup moves followed by 10^5 moves where data was accumulated, where each walk was initialized from a random configuration. Curve (ii) was similarly obtained, except each walk was initialized at the minimum energy configuration, the icosahedron.¹³ The two curves are coincident only for temperatures greater than about 34 K; for low temperatures, the randomly initialized cluster quenches to one of the numerous higher potential isomers and becomes trapped. Mostly coincident with the icosahedral curve (on the scale of the plot) is curve (iii) for random initialization with *J*-walking. This curve was generated by sampling from the 20 K *J*-walker distribution for points in the range 0–20 K, sampling from the 30 K distribution for the 20 to 30 K range, and so on. Each distribution was stored piecemeal in several files because of memory constraints, with each walk sampling from a different random piece of the distribution. The *J*-walking has completely overcome the quasi-ergodicity in the randomly initialized walks at low temperature. This is a significant advantage in itself since often much time and energy can be extended in locating the global minimum to ensure proper initialization for reasonable warmup times and good convergence.

The good results for even very low temperatures near zero seems surprising at first glance because these temperatures are so low that no jumps to a 20 K distribution are accepted. The method works because the randomly initialized clusters almost always have energies higher than those in the *J*-walker distribution so that the first attempted jump during the warmup period is almost always accepted and the random initialization is for all practical purposes an initialization from the *J*-walker configuration. Since the vast majority of 20 K *J*-walker distributions are solid-like distorted icosahedra, the cluster quickly quenches to the global minimum, after which no jumps are accepted. In fact, the icosahedral wells are so dominant that even a *J*-walker distribution at 40 K, which contains a large proportion of liquid-like configurations, resulted in quenches to the global minimum more than 99% of the time.

The near agreement between the *J*-walker results and the icosahedral initialized results indicate that quasi-ergodicity is not a problem for the average potential energy $\langle V \rangle$. There are however some subtle advantages in *J*-walking that can be seen in Fig. 10(b), which shows the standard devi-

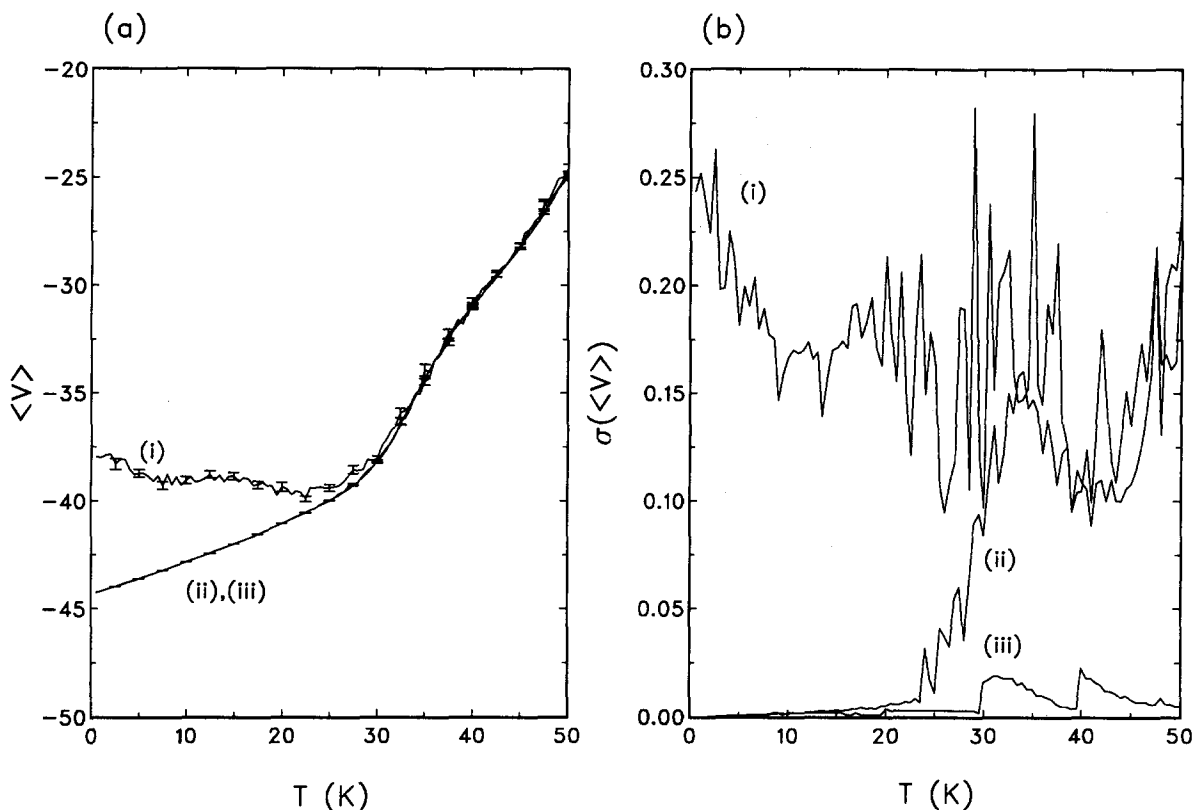


FIG. 10. (a) Variation of the classical average potential energy (in reduced units) with temperature for $n = 13$ argon clusters. Each point at a given temperature was obtained from 100 independently initialized walks, each consisting of 10^5 warmup moves followed by 10^5 moves where data was accumulated. Curve (i) was obtained with randomly initialized clusters, while curve (ii) represents initialization from the lowest energy (icosohedral) configuration. Nearly coincident with the icosohedral curve is curve (iii) obtained from random initialization using *J*-walking. The distributions in Fig. 9 were used in sequence, with walkers in the $T = 0$ to 20 K range jumping to the $T_j = 20$ K distribution, walkers in the 20 to 30 K range jumping to the 30 K distribution, and so on. The jumping probability P_j was 0.1 throughout. Some representative error bars corresponding to single standard deviations have been included. A more detailed account of the standard deviation is shown in (b). The icosohedral curve (ii) has low standard deviation until the onset of the coexistence range at approximately 24 K whereupon it rises sharply. The random initialization curve (i) has high standard deviation throughout. The lowest curve (iii) is for random initialization with *J*-walking.

ation as a function of temperature. The icosohedral initialized curve has very low standard deviation until about 24 K, where it rises rapidly with the onset of coexistence. It goes through a maximum at about 34 K and another minimum at about 41 K (corresponding to the end of the coexistence range) before rising again. Since each randomly initialized walk has a high probability of becoming trapped in a different well, the standard deviation for the low temperature points is considerably higher than that obtained when all walks are initialized at the global minimum. The curve has much larger fluctuations in the coexistence region, and also has a maximum there. It follows the icosohedral curve in the liquid-like region. The curve for random initialization with *J*-walking shows a substantial reduction in the standard deviation, becoming even lower than the icosohedral curve. Moreover, the standard deviation remains low throughout the coexistence region and the liquid-like region. The steps in the curve arise from the different stages of the *J*-walker distribution that were used.

A much more sensitive measure of quasi-ergodic behavior can be seen in the heat capacity. Fig. 11(a) shows the heat capacity as a function of temperature, where C_v is the sum of the configurational and kinetic heat capacities,

$$C_v = \frac{\langle V^2 \rangle - \langle V \rangle^2}{kT^2} + \frac{3nk}{2}. \quad (18)$$

The curves were calculated from the same data used for the caloric curves shown in Fig. 10. The lower curve represents initialization from the global minimum and the upper curve random initialization with *J*-walking; the curve representing random initialization without *J*-walking has been excluded for clarity. Again there is very good agreement for low temperatures, indicating the success obtained with *J*-walking in eliminating quasi-ergodicity in randomly initialized clusters at low temperatures. More remarkable is the improvement rendered by *J*-walking in the coexistence range. The *J*-walking curve is not only much smoother, it also shows a significantly higher maximum at 34 K and continues to be higher in the liquid-like region. The width of the heat capacity curve about the finite maximum is a measure of the transition range, with the maximum corresponding to the transition temperature.¹¹ Our curve shows a sharp rise at about 24 K and a minimum corresponding to the end of the transition region at about 41 K, in good agreement with the literature values¹² for the coexistence range of about 24–41 K. It should be noted that the shape and location of the finite

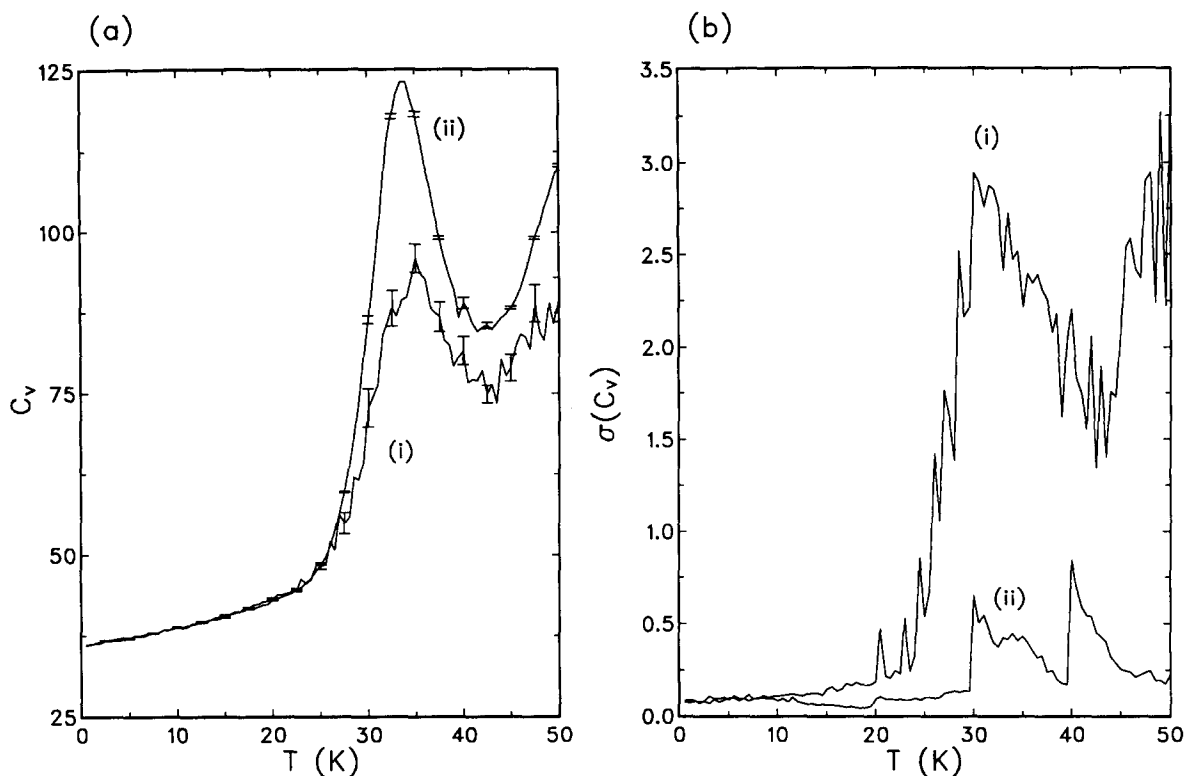


FIG. 11. (a) Heat capacity (in reduced units) for $n = 13$ argon clusters as a function of temperature. The curves were obtained as in Fig. 10. Curve (i) is the heat capacity for icosahedral initialization and curve (ii) for random initialization with *J*-walking. Substantial systematic error due to quasi-ergodicity is evident throughout the coexistence range (approximately 24–41 K) and continues in the liquid region ($T > 41$ K). Again, some representative error bars have been included and a more detailed account of the standard deviation is shown in (b). Curve (ii) is for random initialization with *J*-walking. The curves for random initialization without *J*-walking have not been included since they were mostly off scale.

maximum depend on the size of the constraining radius. Increasing R_c to 5σ , for example, moved the maximum to 35 K and raised the high temperature part of the curve considerably.

Figure 11(b) shows the standard deviation in the heat capacity as a function of temperature. Like the curve for the standard deviation of the average potential energy shown in Fig. 10(b), the curve for the icosahedral initialized clusters shows a very low standard deviation for low temperatures up to the onset of the coexistence region where it rises sharply to a maximum, falls off somewhat until the end of the coexistence region and then rises again in the liquid-like region. The standard deviation in the *J*-walker curve is also very low in the solid-like region, but remains much lower throughout the rest of the curve. Again the steps in the curve are due to the different stages used in the *J*-walker distribution.

Figure 12 demonstrates that the discrepancies between standard Metropolis Monte Carlo and *J*-walking are due to the limitations in standard Metropolis. It plots the convergence in the average potential energy and in the heat capacity as a function of the number of moves per walk N . Each point in the mean of 10 independently initialized walks, each consisting of 10^5 warmup moves followed by N moves of data accumulation. The standard Metropolis walks were initialized from the icosahedral configuration, while the *J*-walker enhanced walks were again initialized from a random configuration. The temperature selected was the maximum

of the heat capacity curve (34 K) since this point had the largest standard deviations. The *J*-walker distribution was at 40 K. Those points with the largest standard deviations have error bars drawn on them. The average of the five points at each N is also given in Table III for both cases. Figure 12(a) shows that both the standard Metropolis and *J*-walking methods are in agreement with the converged value when walks of about 10^4 or more moves are used, but that the spread of points and the standard deviation are significantly lower with *J*-walking. Figure 12(b) illustrates the great increase in convergence for the heat capacity attainable with *J*-walking. Standard Metropolis did not fully converge until 10^7 moves, while *J*-walking gave good accuracy by 10^4 moves. In order to ensure the warmup periods were sufficiently long, the runs were repeated using 10^6 warmup moves. The results obtained were similar.

The standard deviations for the two methods given in Table III were obtained by combining the five sets of data at each point into one set of 50 walks in order to reduce their fluctuations. For standard Metropolis Monte Carlo, the standard deviations have the expected $1/\sqrt{N}$ behavior only for $N > 10^5$, indicating that the shorter walks are not ergodic. However, proper reduction in the variance is no guarantee of ergodicity, as is evident in the heat capacity. The $N = 10^3$ point again illustrates how Monte Carlo results can be very misleading. The standard deviation for this point is quite

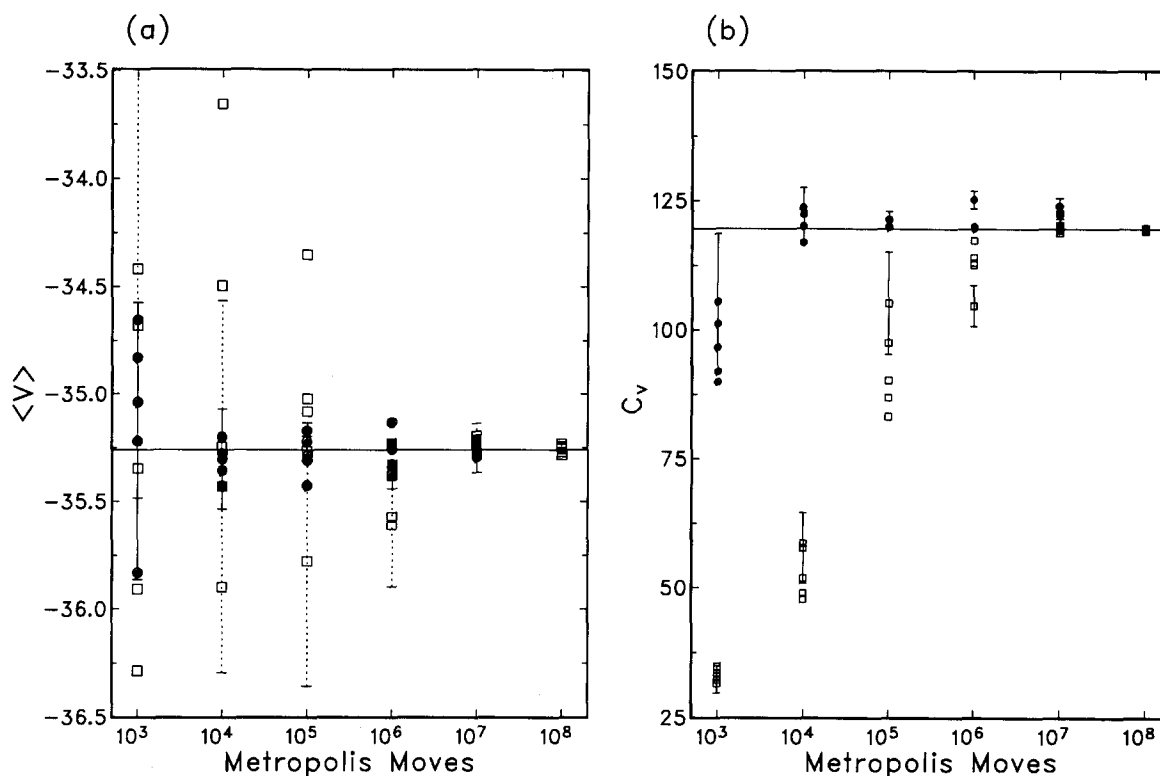


FIG. 12 Comparison of convergence rates between standard Metropolis (open squares) and *J*-walking (filled circles) for $n = 13$ argon clusters. (a) Convergence for the average potential energy. (b) Convergence for the heat capacity. Five sets of data were accumulated for each case, with each point the mean of 10 independently initialized walks having 10^5 warmup moves. The temperature used ($T = 34$ K) was taken from the maximum of the heat capacity curve in Fig. 11(a) where quasi-ergodicity was most evident. The *J*-walker distribution was generated at $T_j = 40$ K from a randomly initialized walk of 10^7 warmup moves followed by 5×10^7 moves where 1 in 10^3 configurations was saved. Error bars have been included for the points with the largest standard deviations. Averages of each set of five points are given in Table III.

low, but the heat capacity is less than a third of the converged value. The standard deviations for *J*-walking show proper $1/\sqrt{N}$ behavior already at $N = 10^3$, but, surprisingly, they level off at $N \approx 10^6$. At first glance, this appears to be simply a limitation of the finite size of the *J*-walker distribution. Each walk used a different random piece of the distribution, each consisting of only 10^4 points. With a jump probability of $P_j = 0.1$, a walk of $N = 10^7$ moves would have an average of 100 jump attempts for each point in the *J*-walker array. Thus the problem is not unlike that arising from the use of random numbers obtained from a random number generator with

too small a cycle length.²³ However, the one-dimensional results given in Table II clearly show proper behavior in the standard deviation, independent of the distribution size—it is the error that depends on the distribution size. Further investigation of these points is in progress.

V. DISCUSSION

These initial tests and preliminary applications of the *J*-walking method show it to have much potential for eliminating, or at least substantially reducing, errors arising from

TABLE III. *J*-walker convergence for Ar_{13} clusters. Averages of the data presented in Fig. 12. The standard deviations for each method were obtained by combining the five sets of data at each point into one large set of 50 walks. Both $\langle V \rangle$ and C_v are in reduced units. The values for the *J*-walker distribution were $T_j = 40$ K, $\langle V_j \rangle = -30.899 \pm 0.074$ and $C_v = 91.10 \pm 2.34$. These were obtained from a single randomly initialized walk consisting of 10^7 warmup moves followed by 5×10^7 moves from which 5×10^4 configurations were sampled.

Moves	$\langle V \rangle$		C_v	
	Metropolis	<i>J</i> -walker	Metropolis	<i>J</i> -walker
10^3	-35.327 ± 0.368	-35.113 ± 0.198	33.07 ± 0.76	97.03 ± 4.73
10^4	-34.944 ± 0.318	-35.312 ± 0.066	52.96 ± 1.96	121.28 ± 1.56
10^5	-35.098 ± 0.222	-35.282 ± 0.032	92.61 ± 3.91	120.57 ± 0.61
10^6	-35.423 ± 0.089	-35.263 ± 0.027	112.37 ± 1.45	120.96 ± 0.57
10^7	-35.230 ± 0.022	-35.260 ± 0.026	119.64 ± 0.36	121.76 ± 0.56
10^8	-35.258 ± 0.009		119.52 ± 0.14	

quasi-ergodicity in Monte Carlo simulations. The method has the added advantage of providing good results even with randomly initialized walks. It is easy to implement since it uses the same Metropolis algorithm to generate the J -walker distribution; when jumps are attempted, only the temperature parameter β needs to be changed to $\beta - \beta_J$. While much work remains to be done to investigate the limits of its applicability and to determine guidelines for optimal use, J -walking can already provide significant improvements when applied to small systems such as clusters. Our work up to this point has produced some generalizations that should be kept in mind when applying the method.

(1) Correlation between the low temperature walker and the J -walker can be a severe impediment when using tandem walkers. Tandem walkers also add substantial computational overhead to the simulation if several similar walks are run and then averaged. The prior generation of separately stored J -walker distributions eliminates the correlation problem and has no computational overhead after the distributions are generated, thus allowing many additional walks to be run from the distributions for good statistics at the same cost as standard Metropolis Monte Carlo. For example, the computer time needed to generate the distributions used for Figures 10 and 11 was approximately 10% of the time used to generate each curve. If memory constraints require tandem walkers, the efficiency can be improved by storing J -walker configurations periodically during the J -walk in a small buffer to provide at least some random access.

(2) Some care must be taken in selecting the J -walker temperature. It must be high enough to be fully ergodic but low enough to provide sufficient jump acceptance rates. The difficulty is that the size of this suitable temperature range can be quite small and difficult to detect. For some systems, both criteria cannot be met simultaneously since J -walker temperatures that provide good jump acceptances are still quasi-ergodic. In these cases, proper J -walker distributions can be generated by occasionally sampling from a yet higher temperature distribution that is itself ergodic.

(3) There is a strong correlation between systematic error in the J -walker distribution and subsequent systematic error in the low temperature walks. Hence the walk used to generate the distribution must be sufficiently lengthy to be representative. This is especially important if the J -walker distribution is generated at a temperature where quasi-ergodic behavior is substantial. Results obtained from such a distribution should be checked for consistency by generating additional J -walker distributions and repeating the simulations. Work still needs to be done to determine distribution parameters best suited for distinguishing good distributions from bad.

(4) Distribution sizes can be kept reasonably small by periodically storing configurations. Since most computers have much larger disk storage than memory storage, larger distributions can be stored as a collection of several files, each of which can then be randomly loaded into computer memory in a periodic manner. Additional study is required to determine the criteria that govern the appropriate distribution sizes.

(5) The optimum value for the jumping probability is system dependent. We have found $P_J \approx 0.1$ to be quite adequate for the systems we have studied. When attempting jumps from an external distribution, the computational overhead is insignificant and larger values can be used, especially if the jump acceptance rate is low. For the tandem walker case, attempted jumps are much more expensive since extra moves are needed to break the correlation, and so P_J should be smaller. This also gives the walkers more time to move apart after a jump is accepted until the next jump attempt, further reducing the correlation.

The use of externally stored distributions might appear at first glance to be a severe limitation for extending J -walking to larger systems. However, optical storage devices with capacities in the hundreds of megabytes are already commercially available at prices comparable to fixed magnetic disk drives and with access times only slightly slower.²⁴ Devices with terabyte (or 1 million megabytes) capacities are already in use in large government and business operations.²⁵ These kinds of capacities should prove sufficient even for the huge distributions that would be required for larger systems. In a similar vein, recent advances in parallel computer architectures offer much promise in overcoming the computational overhead of tandem J -walkers.

Work is currently underway in further optimizing the method and determining its suitability for other Monte Carlo problems. We are also extending the method to quantum systems by coupling it to Fourier path integral Monte Carlo methods.¹⁸ The ability of J -walking to cover large regions of configuration space within a single walk suggests a role as a general multidimensional minimization algorithm, akin to simulated annealing.

ACKNOWLEDGMENTS

Acknowledgment is made to the donors of the Petroleum Research Fund of the American Chemical Society for partial support of this work. One of us (J. D. D.) wishes to thank Brown University and the Department of Chemistry for its generous support of the present research. We would also like to thank Drs. David Leitner and Steve Rick for stimulating discussions.

¹ N. Metropolis, A. Rosenbluth, M. N. Rosenbluth, A. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).

² M. P. Allen and D. J. Tildesley, in *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1987), Ch. 4, pp. 114–119.

³ J. P. Valleau and S. G. Whittington, in *Statistical Mechanics*, edited by B. J. Berne (Plenum, New York, 1977), Ch. 4, p. 145.

⁴ J. P. Valleau and G. M. Torrie, in *Statistical Mechanics*, edited by B. J. Berne (Plenum, New York, 1977), Ch. 5, p. 169.

⁵ A. F. Voter, *J. Chem. Phys.* **82**, 1890 (1984).

⁶ J. Cao and B. J. Berne, *J. Chem. Phys.* **92**, 1980 (1990).

⁷ See A. M. Ferrenberg and R. H. Swendsen, *Comput. Phys.* **3**, 101 (1989) and references therein.

⁸ M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods* (Wiley, New York, 1986), Ch. 3, pp. 73–83.

⁹ W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge University Press, Cambridge, 1986), Ch. 4, pp. 114–120.

¹⁰ T. L. Beck and R. S. Berry, *J. Chem. Phys.* **88**, 3910 (1988); T. L. Beck and R. S. Berry, in *The Physics and Chemistry of Small Clusters*, edited by P. Jena, B. K. Rao, and S. N. Khanna (Plenum, New York, 1987).

- ¹¹ R. S. Berry, T. L. Beck, H. L. Davis, and J. Jellinek, in *Advances in Chemical Physics*, edited by I. Prigogine and S. A. Rice (Wiley, New York, 1988), Vol. 70B, p. 75.
- ¹² J. Jellinek, T. L. Beck, and R. S. Berry, *J. Chem. Phys.* **84**, 2783 (1986); T. L. Beck, J. Jellinek, and R. S. Berry, *J. Chem. Phys.* **87**, 545 (1987).
- ¹³ M. R. Hoare and P. Pal, *Adv. Phys.* **20**, 161 (1971).
- ¹⁴ J. B. Kaelberer and R. D. Etters, *J. Chem. Phys.* **66**, 3233 (1977); R. D. Etters and J. B. Kaelberer, *J. Chem. Phys.* **66**, 5112 (1977).
- ¹⁵ H. L. Davis, J. Jellinek, and R. S. Berry, *J. Chem. Phys.* **86**, 6456 (1987).
- ¹⁶ F. Amar and R. S. Berry, *J. Chem. Phys.* **85**, 5943 (1986).
- ¹⁷ C. L. Briant and J. J. Burton, *J. Chem. Phys.* **63**, 2045 (1975).
- ¹⁸ D. L. Freeman and J. D. Doll, in *Advances in Chemical Physics*, edited by I. Prigogine and S. A. Rice (Wiley, New York, 1988), Vol. 70B, p. 139.
- ¹⁹ R. D. Etters and J. B. Kaelberer, *Phys. Rev A* **11**, 1068 (1975).
- ²⁰ N. Quirke and P. Sheng, *Chem. Phys. Lett.* **110**, 63 (1984).
- ²¹ J. K. Lee, J. A. Barker, and F. F. Abraham, *J. Chem. Phys.* **58**, 3166 (1973).
- ²² All calculations were performed on a DEC3100 UNIX workstation with approximately 5 Mb of free memory.
- ²³ K. Binder and D. W. Heermann, in *Monte Carlo Simulation in Statistical Physics*, edited by P. Fulde (Springer, Berlin, 1988), Vol. 80, Ch. 3, pp. 77-79.
- ²⁴ S. Apiki and H. Eglowstein, *Byte*, October, 160 (1989).
- ²⁵ J. J. Burke and B. Ryan, *Byte*, October, 259 (1989).

The Journal of Chemical Physics is copyrighted by the American Institute of Physics (AIP). Redistribution of journal material is subject to the AIP online journal license and/or AIP copyright. For more information, see <http://ojps.aip.org/jcpo/jcpcr/jsp>
Copyright of Journal of Chemical Physics is the property of American Institute of Physics and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

The Journal of Chemical Physics is copyrighted by the American Institute of Physics (AIP). Redistribution of journal material is subject to the AIP online journal license and/or AIP copyright. For more information, see <http://ojps.aip.org/jcpo/jcpcr/jsp>