On path integral Monte Carlo simulations

M. F. Herman,[a] E. J. Bruskin, and B. J. Berne

Department of Chemistry, Columbia University, New York, New York 10027
(Received 21 October 1981; accepted 26 January 1982)

A Monte Carlo procedure based on a discrete point representation of the path integral for the density matrix is explored. It is found that the variance of the estimator used to evaluate the energy grows as the square root of the number of discrete points used, and is therefore to be avoided in highly quantum mechanical systems, where the number of discrete points must be large. A new energy estimator based on the virial theorem is proposed and shown to be well behaved. The main points of the paper are illustrated, using the harmonic oscillator as an example.

I. INTRODUCTION

There is considerable interest in devising Monte Carlo algorithms[1] to simulate quantum or mixed quantum-classical systems. In this note, we focus on path integral Monte Carlo. This method springs directly from Feynman's path integral formulation of the canonical density matrix.[2] Morita[3] and later Barker[4] used Monte Carlo methods to evaluate a discrete approximation to the full path integral for the density matrix. The discretized path integral of a quantum particle in which there are N discrete points is isomorphic to a classical polymer chain with a prescribed nearest neighbor harmonic interaction potential. Monte Carlo simulation of this "classical chain molecule" interacting with a classical solvent allows one to simulate the "approximate" quantum particle in a classical solvent. The larger the number of discrete points, that is, the larger the "classical polymer," the better will be the discrete point approximation to the full path integral. If the forces acting on the quantum particle do not vary much over the DeBroglie wavelength, and/or if the density of quantum states is high enough that many quantum states are excited at $kT=\beta^{-1}$, the number of discrete points required for an accurate determination of the density matrix is small. Unfortunately, there are many problems where this is not the case and a very large polymer is required. This will be the case for solvated electrons, or protonic motions, as well as for stiff oscillatory coordinates. Path integral Monte Carlo may then become impractical. Thus, efforts to devise algorithms to reduce the number of discrete points are well worth making, and recent progress in this area is welcome.[5,6] However, there still remains a very important problem. Once one has an algorithm for sampling the configuration of the classical polymer and solvent it is necessary to compute various averages. Thus, one must have estimators for the various properties of interest. For a given property, it is possible to devise more than one estimator such that the Monte Carlo average is the same for each choice, but the variance may be very different. Of course, one wants to choose that estimator for which the variance is as small as possible to minimize the error in the computation. A particularly dramatic illustration of this is the energy estimator used by Barker.[4] It is easy to show that the variance of this estimator, and therefore the relative error in the determination of the energy, grows as $\sqrt{N}$, where N is the number of discrete points. Thus, for highly degenerate systems, the mean energy based on this estimator will be very inaccurate. It behooves us to find a better estimator for the energy—a problem we address in this paper. It is very important to note that estimators for other properties like the configurational distribution functions may not have variances that grow strongly with N.

In this paper, we show how the foregoing problem arises in two exactly soluble systems. More importantly, we devise a different energy estimator for which there are no such problems.

II. THE ENERGY ESTIMATOR

The canonical density matrix in the position representation

$$\rho(r_1, r_1'; \beta) = \langle r_1 | e^{-\beta H} | r_1' \rangle \tag{2.1}$$

can be used to compute ensemble averages

$$\langle A \rangle = Q(\beta)^{-1} \int dr_1 \int dr_1' \, \rho(r_1, r_1'; \beta) \langle r_1' | A | r_1 \rangle \, , \tag{2.2}$$

where $Q(\beta)$ is the canonical partition function

$$Q(\beta) = \int dr_1 \, \rho(r_1, r_1; \beta) \tag{2.3}$$

and $\beta = (kT)^{-1}$. Expressing $e^{-\beta H}$ as the product $(e^{-\beta H/N})^N$ and using the closure relation allows one to express ρ and Q as

$$Q(\beta) = \int dr_1 \cdots \int dr_N \prod_{t=1}^{N} \langle r_t | e^{-\beta H/N} | r_{t+1} \rangle \, , \tag{2.4}$$

$$\rho(r_1, r_1'; \beta) = \int dr_2 \cdots \int dr_N \prod_{t=1}^{N} \langle r_t | e^{-\beta H/N} | r_{t+1} \rangle \, , \tag{2.5}$$

where in Eq. (2.4) $r_{N+1} = r_1$ and in Eq. (2.5) $r_{N+1} = r_1'$. The quantity

$$\rho(r_t, r_{t+1}, \beta/N) \equiv \langle r_t | e^{-\beta H/N} | r_{t+1} \rangle \tag{2.6}$$

is the density matrix at the temperature NT, or reciprocal temperature β/N. The above equations are exact for any integer value N.

Using the canonical partition function to calculate the mean energy

[a] Current address: Department of Chemistry, Tulane University, New Orleans, Louisiana, 70118.

 0021-9606/82/105150-06$02.10

$$\langle E\rangle = -\frac{\partial \ln Q(\beta)}{\partial \beta} \tag{2.7}$$

gives

$$\langle E\rangle = \int dr_1 \cdots \int dr_N P(r_1, \ldots, r_N; \beta)\epsilon_N(r_1, \ldots, r_N) , \tag{2.8}$$

where

$$P(r_1, \ldots, r_1) = \prod_{t=1}^{N} \rho(r_t, r_{t+1}; \beta/N) \tag{2.9}$$

and

$$\epsilon_N(r_1, \ldots, r_N) = -\sum_{t=1}^{N} \frac{\partial}{\partial \beta} \ln \rho(r_t, r_{t+1}; \beta/N) . \tag{2.10}$$

All of this is still exact. Clearly, $(r_1, \ldots, r_N)$ can be regarded as the configuration of a fictitious N particle classical system, $P(r_1, \ldots, r_N; \beta)$ can be regarded as a configurational distribution function, and ϵ_N can be regarded as the energy of the system in the particular configuration—albeit a β (or temperature) dependent quantity. ϵ_N consists of interactions between adjacently indexed particles.

Suppose that ρ is known exactly; then Eq. (2.8) can be used to determine the exact mean energy, since there are no approximations. Given the exact ρ, it is possible to evaluate $\langle E\rangle$ by a Monte Carlo procedure. This is easily done by sampling configurations from $P(r_1, \ldots, r_N; \beta)$ in precisely the way one does a classical Monte Carlo simulation; and for each configuration sampled, $\epsilon_N(r_1, \ldots, r_N)$ is evaluated. Summing these values and dividing by the number M of configurations sampled gives

$$\langle E\rangle_N = \frac{1}{M}\sum_{i=1}^{M} \epsilon_N(i) = \langle \epsilon_N\rangle , \tag{2.11}$$

where the subscript N indicates the number of discrete points taken in Eq. (2.8). For M large enough, we expect

$$\langle E\rangle_N = \langle E\rangle \tag{2.12}$$

for any N because no approximations have been made. In general, there will be a distribution in the values of $\epsilon_N(i)$ with a width given by the standard deviation σ_N, where

$$\sigma_N^2 = \langle \delta\epsilon_N^2\rangle = \frac{1}{M}\sum_{i=1}^{M} [\epsilon_N^2(i) - \langle \epsilon_N\rangle^2] \tag{2.13}$$

and the standard deviation in the mean [Eq. (2.11)] is

$$\sigma_\epsilon^2 = \sigma_N^2/M . \tag{2.14}$$

Despite the fact that this algorithm uses the exact density matrix, it is shown later that $\sigma_N^2 = O(N)$ so that

$$\sigma_\epsilon^2 = O(N/M) . \tag{2.15}$$

Since $\langle E\rangle_N$ is essentially independent of M, the relative error using the energy estimator given by Eq. (2.10) is $O(N/M)$. This will be illustrated using the exact density matrix for the harmonic oscillator. These results point out a shortcoming of using the above energy estimator [Eq. (2.10)], even in algorithms using the exact density matrix.

It is important to note that the mean square fluctuation in the energy $\langle \delta E^2\rangle$ can also be evaluated from the exact $Q(\beta)$ of Eq. (2.4), as

$$\langle \delta E^2\rangle = \frac{\partial^2 \ln Q(\beta)}{\partial \beta^2} = kT^2 C_V , \tag{2.16}$$

where C_V is the heat capacity. Substitution of Eq. (2.4) into Eq. (2.16) gives an estimator for δE^2 (which we denote δE_N^2). A MC calculation can be used to determine $\langle (\delta E)^2\rangle$ $(= \langle \delta E_N^2\rangle)$ and thereby C_V. It is important to note that the estimator δE_N^2 is different from $\delta\epsilon_N^2$; that is

$$\delta E_N^2 \neq \epsilon_N^2 - \langle \epsilon_N\rangle^2 \tag{2.17}$$

so that the divergence of σ_ϵ^2 with N does not lead to a divergence of the heat capacity.

III. THE PRIMITIVE ALGORITHM

In general, one does not know the exact form of the density matrix. Recognizing that for large enough N, β/N is very small, the short time approximation to the density matrix in Eq. (2.6) is often adopted;

$$\rho(r_t, r_{t+1}; \beta/N) = \exp\left[-\frac{\beta}{2N} V(r_t)\right] \times \rho_0(r_t, r_{t+1}; \beta/N) \exp\left[-\frac{\beta}{2N} V(r_{t+1})\right] , \tag{3.1}$$

where $V(r_t)$ is the potential energy of classical particle t and ρ_0 is the free particle density matrix. A substitution of this into Eq. (2.9) gives the distribution function in this short time primitive algorithm

$$P(r_1, \ldots, r_N; \beta) = Q^{-1} \prod_{t=1}^{N} \left(\frac{Nm}{2\pi\hbar^2\beta}\right)^{N/2} \times \exp\left\{-\beta \sum_{t=1}^{N} \left[\frac{Nm}{2\hbar^2\beta^2}(r_t - r_{t+1})^2 + \frac{1}{N} V(r)_t\right]\right\} , \tag{3.2}$$

where m is the mass of the quantum particle. This looks like the distribution function of a classical circular chain molecule, consisting of N particles joined by nearest neighbor harmonic bonds, with each particle experiencing a potential field $V(x)$. The partition function in this primitive algorithm is

$$Q(N, \beta) = \int dr_1 \cdots dr_N \left(\frac{Nm}{2\pi\hbar^2\beta}\right)^{N/2} \times \exp\left\{-\beta \sum_{t=1}^{N} \left[\frac{Nm}{2\hbar^2\beta^2}(r_t - r_{t+1})^2 + \frac{1}{N} V(r_t)\right]\right\} . \tag{3.3}$$

In this approximation the energy estimator [cf. Eq. (2.10)] is

$$\epsilon_N = \frac{N}{2\beta} - \alpha_N + \lambda_N = \frac{N}{2\beta} - \gamma_N , \tag{3.4}$$

where

$$\alpha_N \equiv \frac{Nm}{2\hbar^2\beta^2} \sum_{t=1}^{N} (r_t - r_{t+1})^2 , \tag{3.5}$$

$$\lambda_N \equiv \frac{1}{N} \sum_{t=1}^{N} V(r_t) , \tag{3.6}$$

$$\gamma_N = \alpha_N - \lambda_N . \tag{3.7}$$

This is the estimator used by Barker[4] and more recently by others. In Eq. (3.4), the quantity $N/2\beta - \alpha_N \equiv \tau_N$ is the estimator for the kinetic energy and λ_N is the estimator for the potential energy.

In order to gain some insight into the behavior of these different parts of the energy estimators, we now study the case of the linear harmonic oscillator

$$V(x_t) = \tfrac{1}{2} m\omega_0^2 x_t^2 \ . \tag{3.8}$$

The primitive algorithm partition function for given N can be evaluated analytically[6]

$$Q(N,\beta) = \frac{f^{N/2}}{f^N - 1} \equiv Q(N,\beta,m,\omega) \ , \tag{3.9}$$

where

$$f = 1 + \tfrac{1}{2}R^2 + \tfrac{1}{2}R(4+R^2)^{1/2} \tag{3.10}$$

and

$$R = \beta\hbar\omega_0/N \ . \tag{3.11}$$

The mean energy for N discrete points is

$$\langle E\rangle_N = -\frac{\partial \ln Q(N,\beta)}{\partial\beta} = -\frac{\partial \ln Q(N,b)}{\partial f}\frac{\partial f}{\partial R}\frac{\partial R}{\partial\beta} \tag{3.12}$$

and the mean square fluctuation in the energy is

$$\langle \delta E^2\rangle_N = kT^2 C_V(N) = \left(\frac{\partial^2 \ln Q(N,\beta)}{\partial\beta^2}\right) \tag{3.13}$$

or

$$\langle \delta E^2\rangle_N = \frac{\partial^2 \ln Q(N,\beta)}{\partial f^2}\left(\frac{\partial f}{\partial R}\right)^2\left(\frac{\partial R}{\partial\beta}\right)^2 + \frac{\partial \ln Q(N,\beta)}{\partial f}\left(\frac{\partial^2 f}{\partial R^2}\right)\left(\frac{\partial R}{\partial\beta}\right) . \tag{3.14}$$

Substitution of Eq. (3.9) then gives explicit dependence on N. These are exact results for the *primitive algorithm*; and using them we can investigate how large N must be before we get agreement with the exact quantum mechanical oscillator ($N \to \infty$).

It is of interest to investigate the distribution function of the various parts of the energy estimator [cf. Eqs. (3.5)–(3.7)]. To this end we evaluate

$$P_1(\alpha) \equiv \langle \delta(\alpha - \alpha_N)\rangle_N \ , \tag{3.15a}$$

$$P_2(\gamma) \equiv \langle \delta(\gamma - \gamma_N)\rangle_N \ , \tag{3.15b}$$

$$P_3(x) \equiv \langle \delta(\lambda - \lambda_N)\rangle_N \ . \tag{3.15c}$$

Laplace transforming these three functions with respect to α, β, and λ, respectively gives the corresponding characteristic function

$$\tilde{P}_1(s) = \langle e^{-s\alpha_N}\rangle_N \ , \tag{3.16a}$$

$$\tilde{P}_2(s) = \langle e^{-s\gamma_N}\rangle_N \ , \tag{3.16b}$$

and

$$\tilde{P}_3(s) = \langle e^{-s\lambda N}\rangle_N \ , \tag{3.16c}$$

where s is the Laplace transform, and $\langle\cdots\rangle_N$ denotes an average over the distribution function of the primitive algorithm, i.e., for example,

$$\tilde{P}_2(s) = \frac{1}{Q(N,m,\omega_0)}\left(\frac{N_m}{2\Pi\hbar^2\beta}\right)^{1/2}\int d\mathbf{r}_1\cdots\int d\mathbf{r}_N \times \exp\left\{-\beta\left[\left(\frac{s+\beta}{\beta}\right)\alpha_N + \left(\frac{\beta-s}{\beta}\right)\lambda_N\right]\right\} . \tag{3.17a}$$

If we define a renormalized mass m' and frequency ω_0' as

$$m' = \left(\frac{\beta+s}{\beta}\right)m \ ; \quad \omega_0' = \left(\frac{\beta-s}{\beta+s}\right)\omega_0 \ , \tag{3.17b}$$

then

$$\tilde{P}_2(s) = \left(\frac{\beta}{\beta+s}\right)^{N/2}\frac{Q(N,\beta,m',\omega_0')}{Q(N,\beta,m,\omega_0)} \ , \tag{3.18}$$

that is, the integral can be expressed in terms of the partition function of an oscillator of mass m' and frequency ω_0' in the primitive algorithm with N discrete points. Equation (3.19) can be used to evaluate this; however, now

$$R = \beta\hbar\omega_0'/N = \frac{\beta\hbar\omega_0}{N}\left(\frac{\beta-s}{\beta+s}\right) . \tag{3.19}$$

In like manner, we find

$$\tilde{P}_1(s) = \left(\frac{\beta}{\beta+s}\right)^{N/2}\frac{Q(N,\beta,m',\omega_0')}{Q(N,\beta,m,\omega_0)} \ , \tag{3.20a}$$

with

$$m' = \left(\frac{\beta+s}{\beta}\right)m \ ; \quad \omega_0' = \left(\frac{\beta}{\beta+s}\right)^{1/2}\omega_0 \ , \tag{3.20b}$$

and

$$\tilde{P}_3(s) = \frac{Q(N,\beta,m',\omega_0')}{Q(N,\beta,m,\omega_0)} \ , \tag{3.21a}$$

where

$$m' = m \ ; \quad \omega_0' = \left(\frac{\beta+s}{\beta}\right)^{1/2}\omega_0 \ . \tag{3.21b}$$

These characteristic functions enable us to determine various moments in the primitive algorithm. For example,

$$\langle\gamma_N\rangle = -\left(\frac{\partial \ln \tilde{P}_1(s)}{\partial s}\right)_{s\to 0} \ , \tag{3.22a}$$

$$\langle\delta\gamma_N^2\rangle = \left(\frac{\partial^2 \ln \tilde{P}_1(s)}{\delta s^2}\right)_{s\to 0} \ , \tag{3.22b}$$

with equivalent formulas for the other distributions. Substitution of Eq. (3.18) into this followed by some lengthy algebra gives

$$\langle\gamma_N\rangle = \frac{N}{2\beta} - \langle E\rangle_N \ , \tag{3.23a}$$

$$\langle\delta\gamma_N^2\rangle = \frac{N}{2\beta^2} + \langle\delta E^2\rangle_N - \frac{\langle E\rangle_N}{\beta} \ . \tag{3.23b}$$

Similarly, for the other distributions,

$$\langle \alpha_N \rangle = \frac{N}{2\beta} - \tfrac{1}{2}\langle E \rangle_N \,, \tag{3.23c}$$

$$\langle \delta\alpha_N^2 \rangle = \frac{N}{2\beta^2} + \tfrac{1}{4}\langle \delta E^2 \rangle_N - \frac{3}{4}\frac{\langle E \rangle_N}{\beta} \,, \tag{3.23d}$$

$$\langle \lambda_N \rangle = \tfrac{1}{2}\langle E \rangle_N \,, \tag{3.23e}$$

$$\langle \delta\lambda_N^2 \rangle = \tfrac{1}{4}\langle E^2 \rangle_N + \frac{1}{4}\frac{\langle E \rangle_N}{\beta} \,. \tag{3.23f}$$

These formulas illustrate several points.

(a) From Eq. (3.23a), it follows that

$$\langle E \rangle_N = \frac{N}{2\beta} - \langle \gamma_N \rangle \,,$$

as required; that is, the average of the energy estimator gives the exact mean energy expected in the N particle primitive algorithm.

(b) From Eqs. (3.23c) and (3.23e) we see that the average of the kinetic energy part of the estimator

$$\langle \tau_N \rangle = \langle N/2\beta - \alpha_N \rangle = \tfrac{1}{2}\langle E \rangle_N \tag{3.24}$$

and the average of the potential energy estimator

$$\langle \lambda_N \rangle = \tfrac{1}{2}\langle E \rangle_N \tag{3.25}$$

satisfy the virial theorem for the N discrete point oscillator (see Appendix A).

(c) It follows from Eq. (3.4) that $\langle \delta\epsilon_N^2 \rangle = \langle \delta\gamma_N^2 \rangle$ and, therefore, from Eq. (3.23b) that

$$\langle \delta\epsilon_N^2 \rangle = N/2\beta^2 + \langle \delta E^2 \rangle_N - \langle \mathbf{E} \rangle_N/\beta \,.$$

Since, as we shall see, both $\langle \delta E_N^2 \rangle$ and $\langle E \rangle_N$ become independent of N as $N \to \infty$, the mean square fluctuation in the energy estimator as well as in the kinetic energy $\langle \delta\tau_N^2 \rangle$ grow as $O(N)$.

(d) In marked contrast to (c), the mean square fluctuation in the potential energy $\langle \delta\lambda_N^2 \rangle$ is not very sensitive to N.

The precise dependence of these quantities on N for $\beta\hbar\omega_0 = 3$ is given in Figs. 1 and 2. These figures illustrate that large fluctuations are expected in the energy estimator defined by Eqs. (2.10) and (3.1) with the consequence that the mean energy cannot be determined accurately in a primitive algorithm Monte Carlo run using this energy estimator. The trouble arises from the kinetic energy part of the energy estimator.

Clearly a *different* energy estimator is required. We have observed that the potential energy fluctuations do not depend strongly on N. This suggests that we seek an energy estimator that is determined only by the potential energy and its derivatives. In Appendix A, we show that a good estimator is

$$\epsilon_N^V = \frac{1}{N}\sum_{t=1}^{N}\left[V(x_t) + \tfrac{1}{2}\, x_t \frac{\partial V(x_t)}{\partial x_t}\right] . \tag{3.27}$$

This springs from the virial theorem applied to path integrals.

For the linear harmonic oscillator this "virial" energy estimator is particularly simple,

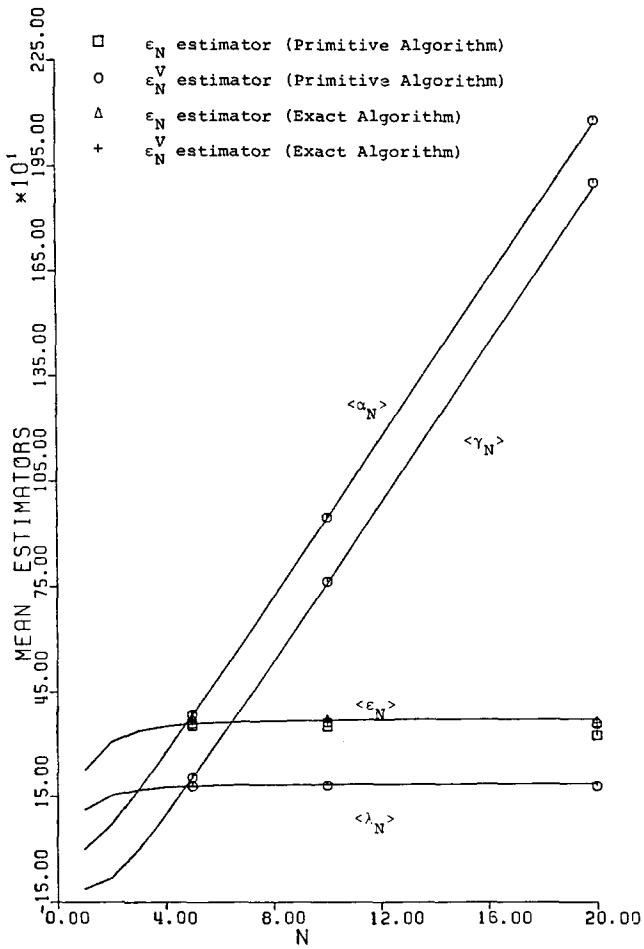

FIG. 1. This figure shows how the mean values of the various terms in the energy estimator in °K [cf. Eq. (3.4)] for the harmonic oscillator ($\beta\hbar\omega_0 = 3$) depend on the number of discrete points used in the algorithm. The solid lines correspond to the analytical results presented in Eqs. (3.23a)–(3.23e). In particular, the solid line labeled $\langle \epsilon_N \rangle$ is identical to the analytical calculation of $\langle E \rangle_N$ [cf. Eq. (3.12)]. The squares (□), circles (○), triangles (△), and pluses (+) give these same quantities using various Monte Carlo simulations all with the same number of configurations. The squares (□) and circles (○) correspond to Monte Carlo runs using the primitive algorithm [cf. Eq. (3.2)]. The squares (□) are based on the energy estimator of Eq. (3.4) and the circles (○) are based on the virial estimator, Eq. (3.27). The triangles and pluses correspond to Monte Carlo runs using the *exact* density matrix for the harmonic oscillator, [cf. Eqs. (2.9) and (4.1)]. The triangles (△) correspond to the use of the energy estimator in Eq. (2.10), while the pluses correspond to the use of the virial estimator.

$$\epsilon_N^V = \frac{2}{N}\sum_{t=1}^{N} V(x_t) = 2\lambda_N \,, \tag{3.28}$$

so that its mean square fluctuation is simply

$$\langle (\delta\epsilon_N^V)^2 \rangle = 4\langle \delta\lambda_N^2 \rangle = \langle \delta E^2 \rangle_N + \frac{\langle E \rangle_N}{\beta} \,, \tag{3.29}$$

where the last equality follows from Eq. (3.23f). Obviously, since the quantities on the right-hand side do not depend strongly on N for large N, this estimator is well behaved (cf. Figs. 1 and 2). As shall be seen in the next section, the virial estimator is generally to be preferred.

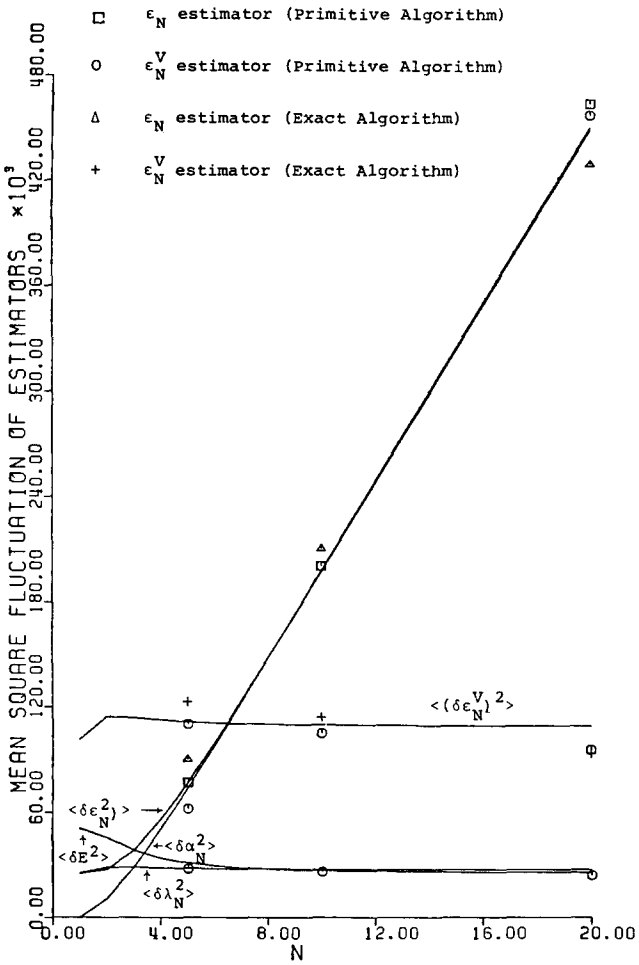

FIG. 2. This figure shows how the mean square flucutations in $(°K)^2$ of the properties for the harmonic oscillator ($\beta\hbar\omega_0 = 3$) indicated in Fig. 1 depend on the number of discrete points used in the algorithm. The solid lines correspond to the analytical results presented in Eqs. (3.23b), (3.23d), (3.23f), and (3.29). The squares (□), circles (○), triangles (Δ), and pluses (+), give these same quantities using the various Monte Carlo simulations all with the same numbers of configurations. These symbols denote the same runs defined in the caption to Fig. 1. This figure shows clearly that (a) $\langle\delta E_N^2\rangle$ is a weak function of N, (b) the energy estimator defined by Eq. (3.4) has a mean square fluctuation that grows $O(N)$ for large N, (c) the virial energy estimator defined by Eq. (3.27) has a mean square fluctuation that is a very weak function of N, and (d), the simulations agree very closely with the exact analytical theory for the N point primitive algorithm for all values of N.

IV. MONTE CARLO SIMULATIONS

To explore the accuracy of the various energy estimators discussed, and to study the convergence properties of path integral Monte Carlo, we first study the linear harmonic oscillator (LHO) of mass m and frequency ω_0. The exact density matrix for the LHO is[2]

$$\rho\left(x_t, x_{t+1}; \frac{\beta}{N}\right) = \left(\frac{\mu\omega_0}{2\pi\hbar\sinh(R)}\right)^{1/2} \times \exp\left(-\frac{\mu\omega_0}{2\hbar\sinh(R)}\right)[(x_t^2 + x_{t+1}^2)\cosh R - 2x_t x_{t+1}] \,, \quad (4.1)$$

where $R = \beta\hbar\omega_0/N$. Substitution of this into Eqs. (2.9) and (2.10) leads to an explicit expression for the energy estimator, and the distribution function $P(x_1, \ldots, x_N)$. The standard Metropolis MC algorithm was used to generate M configurations, and M corresponding values of ϵ_N for a LHO with $\beta\hbar\omega_0 = 3$. It was found that convergence was rather slow. Since $P(x_1, \ldots, x_N)$ is a multivariate Gaussian, one can transform to normal modes. When this is done, each mode has a different force constant and it is possible to adopt a step size appropriate to that force constant; that is, a small step size is chosen for modes of high force constant, whereas a large step size is chosen for modes corresponding to small force constants. This is done very systematically; but we shall not give the details here. Suffice it to say that this led to much more rapid convergence. In Fig. 1, the triangles (Δ) indicate the results [based on the estimator ϵ_N of Eq. (2.10)], corresponding to this exact algorithm for $N = 5$, 10, and 20. All of these runs were performed for $M = 2\times10^6$ so that the N dependence could be tested. In addition, the pluses (+) indicate the results [based on the virial estimator ϵ_N^V of Eq. (3.27)] for $N = 5$, 10, and 20 and for fixed M. The data show very clearly that the mean energy is the same for both estimators; however, the standard deviations $\langle(\delta\epsilon_N)^2\rangle$ and $\langle(\delta\epsilon_N^V)^2\rangle$, shown in Fig. 2, are very different, with $\langle(\delta\epsilon_N)^2\rangle$ increasing linearly with the N, while $\langle(\delta\epsilon_N^V)^2\rangle$ is basically insensitive to N. It is interesting that for $N < 6$, the error associated with the virial estimator ϵ_N^V is greater than that corresponding to the estimator ϵ_N. For large N, there is such a dramatic difference that it would be foolhardy not to use ϵ_N^V instead of ϵ_N. From Eq. (2.15), it is clear that if N is increased, so too must the length of the run M be increased to maintain the same accuracy. This has very important implications for algorithms based on the short time approximation [cf. Eq. (3.2)]. The foregoing conclusions are based on an algorithm using the exact density matrix, and consequently is an inherent defect of the energy estimator based on Eq. (2.10).

Equation (3.22) gives analytical results for the LHO using the primitive algorithm with N "classical particles." In Figs. 1 and 2, these quantities are plotted as a function of N for an oscillator with $\beta\hbar\omega_0 = 3$. Equation (3.2) is used with $V(x) = 1/2m\omega_0^2x^2$, as the sampling function in Monte Carlo simulation of the primitive algorithm with $N - 5$, 10, and 20. Each simulation generated $M = 2\times10^6$. The results of these simulations are indicated in Figs. 1 and 2 by squares (□) for the estimator ϵ_N [cf. Eq. (3.4)], and by circles (○) for the virial estimator. Several features are worth noting:

(i) The results of the simulation agree with analytical results.

(ii) The mean energies computed from ϵ_N and ϵ_N^V agree.

(iii) $\langle\delta\epsilon_N^2\rangle$ increases linearly with N, whereas $\langle(\delta\epsilon^V)^2\rangle$ is insensitive to N.

(iv) There is excellent agreement between the exact algorithm and the short time algorithm for all N considered.

For convenience, the analytical values of the mean en-

ergy $\langle E\rangle_N$ and the mean square fluctuation of the energy $\langle \delta E^2\rangle$ corresponding to the primitive algorithm are also given in Figs. 1 and 2. Those results go to the exact values expected for a LHO in the limit $N\to\infty$, but deviate from these for small N. For this oscillator ($\beta\hbar\omega_0=3$), $N\gtrsim 8$ is required to get good agreement with the exact quantum oscillator. Nevertheless, the results of the MC algorithm for given N should be compared with the exact results for that N and not with the $N\to\infty$ results. If $\beta\hbar\omega_0$ is very large, then the primitive algorithm will only agree with the exact quantum oscillator ($N\to\infty$), if N is very large. Thus, it is sometimes necessary to use very large N in a primitive algorithm.

In Sec. II, we showed that the standard deviation in the mean for the energy estimator ϵ_N of Eq. (2.10) is given by Eq. (2.15).

$$\sigma_\epsilon^2 = O(N/M)\ , \tag{2.15}$$

where N is the number of discrete points and M is the total number of independent configurations generated in the MC run. This should be compared with the standard deviation $\sigma_\epsilon v$ corresponding to the virial energy estimator. Since $\langle(\delta\epsilon_N^V)^2\rangle$ is only a very weak function of N for large N, it follows that

$$\sigma_{\epsilon V}^2 = \frac{\langle(\delta\epsilon_N^V)^2\rangle}{M} = O\left(\frac{1}{M}\right).$$

Clearly then for given N, the same accuracy is achieved for the virial estimator in M_V configurations (moves) as for the energy estimator of Eq. (3.4) in M moves where

$$M_V = \frac{M}{N}\ .$$

Thus, when N is large, as it must be in the simulation of highly degenerate systems, $M_V \ll M$ and the virial estimator is greatly to be preferred. Since the time required for an algorithm is of $O(NM)$, to achieve comparable accuracy the virial energy estimator requires a time of order $O(nN)$, compared to that using the usual energy estimator which requires a time $O(nN^2)$, where n is the number of moves per particle ($n\equiv M/N$).

Although the analysis presented here pertains to the harmonic oscillator, we believe that the conclusions are more far reaching. The derivation in the Appendix is general and the virial energy estimator is a perfectly legitimate estimator in all applications. In a recent simulation[7] of an electron in liquid helium, argon, etc., where $N=750$, we have compared the two estimators and find the virial estimator to be far superior.

ACKNOWLEDGMENTS

The authors have benefited from discussions with Dr. M. Rao and Mr. Randall W. Hall. This work was supported by grants from NSF, (NSF CHE 79-07820) and NIH (9R01 GM 26588-06).

APPENDIX A

Let us consider the average quantity in the primitive algorithm

$$\left\langle \sum_{t=1}^N x_t \frac{\partial U_N}{\partial x_t}\right\rangle = \frac{\int dx_1\cdots dx_N \left(\sum_{t=1}^N x_t \frac{\partial U_N}{\partial x_t}\right) e^{-\beta U_N}}{\int dx_1\cdots dx_N\, e^{-\beta U_N}}\ , \tag{A1}$$

where [cf. Eqs. (3.5) and (3.6)]

$$U_N = \alpha_N + \lambda_N\ . \tag{A2}$$

Since $x_t(\partial U_N/\partial x_t)\, e^{-\beta U_N} = -\beta^{-1} x_t(\partial/\partial x_t)\, e^{-\beta U_N}$; an integration by parts gives

$$\left\langle \sum_{t=1}^N x_t \frac{\partial U_N}{\partial x_t}\right\rangle = \beta^{-1}\sum_{t=1}^N \left\langle \frac{\partial x_t}{\partial x_t}\right\rangle = N\beta^{-1} \tag{A3}$$

or

$$\left\langle \sum_{t=1}^N x_t \frac{\partial \alpha_N}{\partial x_t}\right\rangle_N + \left\langle \sum_{t=1}^N x_t \frac{\partial \lambda_N}{\partial x_t}\right\rangle = N\beta^{-1}\ . \tag{A4}$$

Now α_N is a homogeneous function of $\{x_1,\ldots,x_N\}$ of degree 2, so that Euler's theorem gives $\sum_t x_t(\partial\alpha_N/\partial x_t) = 2\alpha_N$, and Eq. (A4) can be expressed as

$$\left\langle \frac{N}{2\beta} - \alpha_N\right\rangle = \left\langle \frac{1}{2N}\sum_{t=1}^M x_t \frac{\partial V(x_t)}{\partial x_t}\right\rangle, \tag{A5}$$

where we have substituted the definition of λ_N [cf. Eq. (3.6)].

The left-hand side of Eq. (A5) is the mean kinetic energy in the primitive algorithm. Since

$$\langle\epsilon_N\rangle = \left\langle \frac{N}{2\beta} - \alpha_N + \lambda_N\right\rangle\ . \tag{A6}$$

Substitution of Eq. (A5) into Eq. (A6) gives

$$\langle\epsilon_N\rangle = \left\langle \frac{1}{N}\sum_{t=1}^N \left[V(x_t) + \tfrac{1}{2}x_t\frac{\partial V(x_t)}{\partial x_t}\right]\right\rangle. \tag{A7}$$

Thus, another perfectly good energy estimator is

$$\epsilon_N^V = \frac{1}{N}\sum_{t=1}^N \left[V(x_t) + \tfrac{1}{2}x_t\frac{\partial V(x_t)}{\partial x_t}\right]. \tag{A8}$$

In some applications it is necessary to determine the mean kinetic energy $\langle\tau_N\rangle$ in the primitive algorithm; a good estimator for this is thus

$$\tau_N^V = \frac{1}{2}\frac{1}{N}\sum_{t=1}^N x_t\frac{\partial V(x_t)}{\partial x_t}\ . \tag{A9}$$

Both ϵ_N^V and τ_N^V have finite variances as opposed to ϵ_N and $\tau_N = (N/2\beta - \alpha_N)$.

[1] M. H. Kalos, Phys. Rev. A **2**, 250 (1970); M. H. Kalos, D. Levesque, and L. Verlet, Phys. Rev. A **9**, 2178 (1974); P. A. Whitlock, M. H. Kalos, J. Comp. Phys. **30**, 361 (1979); D. M. Ceperly and M. H. Kalos, *Monte Carlo Methods in Statistical Physics*, edited by K. Binder (Springer, Berlin, 1981). See also M. F. Herman and B. J. Berne, Chem. Phys. Lett. 77, 163 (1981).
[2] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw–Hill, New York, 1965).
[3] T. Morita, J. Phys. Soc. Jpn. **35**, 98 (1973).
[4] J. Barker, J. Chem. Phys. **70**, 2914 (1979).
[5] D. Chandler and P. G. Wolynes, J. Chem. Phys. **74**, 4078 (1981).
[6] K. S. Schweizer, R. M. Stratt, D. Chandler, and P. G. Wolynes, J. Chem. Phys. **75**, 1347 (1981).
[7] Randall W. Hall, M. F. Herman, and B. J. Berne, (in preparation).