On the force bias Monte Carlo simulation of simple liquids[a)]

M. Rao and B. J. Berne

Department of Chemistry, Columbia University, New York, New York 10027
(Received 12 February 1979; accepted 2 April 1979)

Two recently devised Monte Carlo schemes, the force bias method and the smart Monte Carlo method are compared using a simple Lennard-Jones fluid as an example. Using diffusion in the configuration space as a criterion, the efficiencies of the two methods are evaluated. When the diffusion is optimized the FB method is found to be more efficient.

I. INTRODUCTION

The Monte Carlo procedure of Metropolis *et al.*[1,2] is widely used to determine the equilibrium structural and thermodynamic properties of gases, liquids, solids, and mesophases. In a previous paper[3] we introduced a modification of the usual Metropolis procedure that gives more rapid convergence and thereby much more efficient Monte Carlo runs. In this new procedure each particle move is chosen with greater probability in the direction of the instantaneous force on the particle than in other directions. The particle moves therefore usually lead to a lowering of the overall potential energy and thereby to a higher acceptance probability than in the usual Metropolis procedure. The new procedure, appropriately called the force bias method, was then applied to a study of ST-2 water to very good effect.

It was clear at the outset that there were no objective criteria for comparing the efficacy of two different Monte Carlo procedures or for optimizing any given procedure. Thus, in a follow-up publication[4] we argued that a good measure is the diffusion in configuration space. It is clear that if the Monte Carlo procedure does not involve the exchange of particles, then the step sizes, etc., for the procedure should be chosen in such a way as to optimize the diffusion coefficient.

Moreover, that procedure is best which for optimized step sizes gives the largest diffusion coefficients. This will guarantee that more independent configurations are sampled in a given number of moves. An exhaustive study was made[4] of diffusion in the Metroplis scheme and in the force bias scheme and it was shown that for ST-2 water the force bias scheme gave a substantial improvement in convergence over the Monte Carlo scheme.

Recently a seemingly different Monte Carlo procedure[5] has been suggested. This method is based on Brownian dynamics, and is referred to by its inventors as the smart Monte Carlo method (SMC). This method also generates moves biased in the direction of the force, and is in many ways similar to the force bias (FB) procedure.

In this paper we compare these two methods. It is found that the force bias scheme gives larger acceptance probabilities and diffusion than the Brownian dynamic smart Monte Carlo scheme. In our view these two present algorithms represent only the beginning in the search for more rapidly convergent schemes, and *smarter* algorithms should soon be in the offing.

II. METHODOLOGY

In both the force bias (FB) and the Brownian dynamic smart Monte Carlo method (SMC), a new configuration of the N particle system $\mathbf{R}'=(\mathbf{R}_1', \ldots, \mathbf{R}_N')$ is generated from an old configuration $\mathbf{R}=(\mathbf{R}_1, \ldots, \mathbf{R}_N)$ by sampling $\mathbf{R}'$ from a transition probability $T(\mathbf{R}'|\mathbf{R})$. The transition probability is chosen such that it is normalized and such that if a state $\mathbf{R}'$ is accessible from $\mathbf{R}$, then the state $\mathbf{R}$ is accessible from $\mathbf{R}'$. The new state $\mathbf{R}'$ is then accepted with probability,

$$p=\min\{1, q(\mathbf{R}'|\mathbf{R})\} \tag{2.1}$$

or the old state $\mathbf{R}$ is kept with probability $(1-p)$ where

$$q(\mathbf{R}'|\mathbf{R}) \equiv T(\mathbf{R}|\mathbf{R}')P(\mathbf{R}')/T(\mathbf{R}'|\mathbf{R})P(\mathbf{R}) \tag{2.2}$$

and

$$P(\mathbf{R})=Z^{-1}e^{-\beta V(\mathbf{R})} \tag{2.3}$$

is the Boltzmann distribution. The sequence of configurations so generated will then be distributed according to the Boltzmann distribution, Eq. (2.3).

The choice of the transition probability is what distinguishes the various methods. In all of the methods discussed here, only one particle is moved at a time; so that $\mathbf{R}'$ differs from $\mathbf{R}$ only with respect to the position of a particular particle.

In the force bias (FB) method the transition probability is taken as[3,4]

$$T_{\mathrm{FB}}(\mathbf{R}'|\mathbf{R})=\begin{cases} C(\mathbf{F}(\mathbf{R}), \Delta)\exp[\lambda\beta\mathbf{F}(\mathbf{R})\cdot(\mathbf{r}'-\mathbf{r})]; & \mathbf{r}'-\mathbf{r}\in D \\ 0; & \mathbf{r}'-\mathbf{r}\notin D \end{cases} \tag{2.4}$$

where $\mathbf{r}'-\mathbf{r}$ stands for the displacement of the particle being moved, $\mathbf{F}(\mathbf{R})$ is the force acting on the particle before it is moved (when the whole system has the old configuration, $\mathbf{R}$), λ is a parameter to be discussed below, $\beta=(k_BT)^{-1}$, and $C(\mathbf{F}(\mathbf{R}), \Delta)$ is a normalization constant. From Eq. (2.4) we note that the transition probability is zero if the particle displacement $\Delta\mathbf{r}=\mathbf{r}'-\mathbf{r}$ falls outside of a certain domain D defined such that; $-\Delta/2\leq\Delta x, \Delta y, \Delta z\leq\Delta/2$. This means that the maximum step size allowed is $(\sqrt{3}/2)\Delta$. Clearly the normalization

[a)]Research supported by grants from the National Science Foundation (NSF CHE 76-11002) and the National Institutes of Health (NIH RO1 NS 12714-03).

 0021-9606/79/010129-04$01.00

constant C depends both on the instantaneous force on the particle and on the value of Δ. As we have shown elsewhere[4] for each system there is an optimum choice for Δ.

It is clear that Eq. (2.4) gives rise to a biased sampling of $\Delta\mathbf{r}$ in the direction of the force $\mathbf{F}$. If $\lambda=0$, the transition probability gives uniform sampling in the domain D and reduces to the usual Metropolis scheme.

In the Brownian dynamic smart Monte Carlo scheme,[5]

$$T_{\mathrm{SMC}}(\mathbf{R}'|\mathbf{R})=\frac{1}{(4\pi A)^{3/2}}\exp\left(-\frac{[(\mathbf{r}'-\mathbf{r})-\beta A\mathbf{F}(\mathbf{R})]^2}{4A}\right), \quad (2.5)$$

where A is a parameter that must be chosen. Here too the sampled particle displacement is biased in the direction of the instantaneous force.

Explicit evaluation of the exponent allows us to write Eq. (2.5) in the form

$$T_{\mathrm{SMC}}(\mathbf{R}'|\mathbf{R})=\frac{B(\mathbf{F},A)}{(4\pi A)^{3/2}}\exp\left(-\frac{(\mathbf{r}'-\mathbf{r})^2}{4A}\right)\times\exp[+\tfrac{1}{2}\beta\mathbf{F}(\mathbf{R})\cdot(\mathbf{r}'-\mathbf{r})], \quad (2.6)$$

where $B=e^{-\beta^2A^2F^2(R)/4}$. Thus, in reality the T_{SMC} is very similar to T_{FB} with one major difference and one minor difference. The major difference is that in the FB method there is an upper bound on the displacement whereas in the SMC there is no upper bound. For example, if in a certain initial configuration, $\mathbf{F}(\mathbf{R})=0$, then T_{FB} would sample $\Delta\mathbf{r}$ uniformly whereas T_{SMC} would sample $\Delta\mathbf{r}$ from a normalized Gaussian whose width is specified by the parameter A. From the properties of the Gaussian we note that in this case the mean square displacement on a move would be $\langle\Delta r^2\rangle=6A$. Clearly then $\sqrt{A}$ gives a measure of the kind of displacement generated in this methods, and is thus analogous to the parameter Δ in the FB. In applying the SMC method one should choose a value of A that optimizes the method.

The minor difference is the parameter λ. If this parameter λ in Eq. (2.4) is adjusted to be $\lambda=\frac{1}{2}$, then the force bias part of the sampling is the same for the two methods. The FB however allows freedom in the choice of λ whereas in the SMC it is fixed at $\frac{1}{2}$. When we proposed the force bias method we set $\lambda=1$, but in our subsequent work we varied λ to study its effect.

Let us compare the acceptance probabilities in the FB and SMC methods. Substitution of Eq. (2.4) into Eq. (2.1) gives

$$P_{\mathrm{FB}}=\min[1,\ \exp(-\beta\{\Delta W_{\mathrm{FB}}+V(\mathbf{R}')-V(\mathbf{R})+\lambda[\mathbf{F}(\mathbf{R}')+F(\mathbf{R})]\cdot\Delta\mathbf{r}\})] \quad (2.7)$$

whereas substitution of Eq. (2.6) into Eq. (2.1) gives

$$P_{\mathrm{SMC}}=\min[1,\ \exp(-\beta\{\Delta W_{\mathrm{SMC}}+V(\mathbf{R}')-V(\mathbf{R})+\tfrac{1}{2}\{\mathbf{F}(\mathbf{R}')+\mathbf{F}(\mathbf{R})\}\cdot\Delta\mathbf{r})] \quad (2.8)$$

where

$$\Delta W_{\mathrm{FB}}\equiv -kT\ln\left\{\frac{C(F(\mathbf{R}'),\Delta)}{C(F(\mathbf{R}),\Delta)}\right\}\equiv W_{\mathrm{FB}}(\mathbf{R}')-W_{\mathrm{FB}}(\mathbf{R}) \quad (2.9)$$

and

$$\Delta W_{\mathrm{SMC}}\equiv -kT\ln\left\{\frac{B(\mathbf{F}(\mathbf{R}'),A)}{B(\mathbf{F}(\mathbf{R}),A)}\right\}\equiv W_{\mathrm{SMC}}(\mathbf{R}')-W_{\mathrm{SMC}}(\mathbf{R}), \quad (2.10)$$

where C and B are defined in Eqs. (2.4) and (2.6). It follows from an explicit evaluation that

$$\Delta W_{\mathrm{SMC}}=(\beta A/4)\{2\mathbf{F}\cdot\Delta\mathbf{F}+(\Delta\mathbf{F})^2\} \quad (2.11)$$

and

$$\Delta W_{\mathrm{FB}}=-\beta^{-1}\sum_{\mu=x,y,z}\ln\left[\left(\frac{F_\mu+\Delta F_\mu}{F_\mu}\right)\times\left(\frac{\exp(\alpha F_\mu)-\exp(-\alpha F_\mu)}{\exp[\alpha(F_\mu+\Delta F_\mu)]-\exp[-\alpha(F_\mu+\Delta F_\mu)]}\right)\right], \quad (2.12)$$

where $\alpha\equiv\lambda\beta\Delta/2$, $\mathbf{F}\equiv\mathbf{F}(\mathbf{R})$, $\Delta\mathbf{F}\equiv\mathbf{F}(\mathbf{R}')-F(\mathbf{R})$, and the subscript μ denote the x, y, z components of the force.

For small step sizes (Δ, small) Eq. (2.12) reduces to the form

$$\Delta W_{\mathrm{FB}}\simeq\frac{\lambda^2\beta\Delta^2}{24}\{2\mathbf{F}\cdot\Delta\mathbf{F}+(\Delta\mathbf{F})^2\}. \quad (2.13)$$

Comparing Eqs. (2.11) and (2.13) we see that for the particular choice of $\lambda=\frac{1}{2}$, the only difference between P_{FB} and P_{SMC} springs from the difference between ΔW_{FB} and ΔW_{SMC}. For small step sizes these two quantities will be identical if A is adjusted to

$$A=\Delta^2/24. \quad (2.14)$$

Thus we conclude that for sufficiently small step sizes (Δ small) and for $\lambda=\frac{1}{2}$, the FB and SMC methods give identical acceptance probabilities.

For large step sizes (Δ, large) on the other hand, ΔW_{FB}, as given by Eq. (2.12) differs considerably from ΔW_{SMC} with the consequence that even for $\lambda=\frac{1}{2}$, P_{FB} and P_{SMC} are considerably different. Analysis of Eq. (2.12) shows that the FB method should give larger acceptance probabilities than the SMC—an observation that is confirmed in the next section where comparison of the two methods when applied to Lennard-Jonesium shows that the FB procedure is superior to the SMC scheme.

Comparison of the two methods for the case when $\lambda=1$ is more subtle. First we note that expansion of Eq. (2.7) to second order in the step size gives

$$P_{\mathrm{FB}}=\min\{1,\ \exp\{-\beta[(2\lambda-1)\mathbf{F}-\mathbf{F}_{\mathrm{FB}}]\cdot\Delta\mathbf{r}+\tfrac{1}{2}\Delta\mathbf{r}\cdot[(2\lambda-1)\nabla\mathbf{F}-\nabla\mathbf{F}_{\mathrm{FB}}]\cdot\Delta\mathbf{r}\}, \quad (2.15)$$

where $\nabla\mathbf{F}$ is the gradient of the force and $\mathbf{F}_{\mathrm{FB}}=-\nabla\Delta W_{\mathrm{FB}}$. Now for $\lambda=1$ we note that terms involving $\mathbf{F}$ and ∇F survive unlike what happens at $\lambda=\frac{1}{2}$. Nevertheless, the terms $\mathbf{F}_{\mathrm{FB}}$ and $\Delta\mathbf{F}_{\mathrm{FB}}$ depend nonlinearly on λ, and might well neutralize the effect of these surviving terms. Thus it is not clear what the situation is with respect to these terms. For small step sizes it is clear that these terms will generate a smaller acceptance probability than does the choice $\lambda=\frac{1}{2}$. This is precisely what is claimed in Ref. 5. For large step sizes, however, it is not clear at this point whether the FB with $\lambda=1$ should be more or less efficient than either the SMC or the FB with $\lambda=\frac{1}{2}$. In this paper we show that

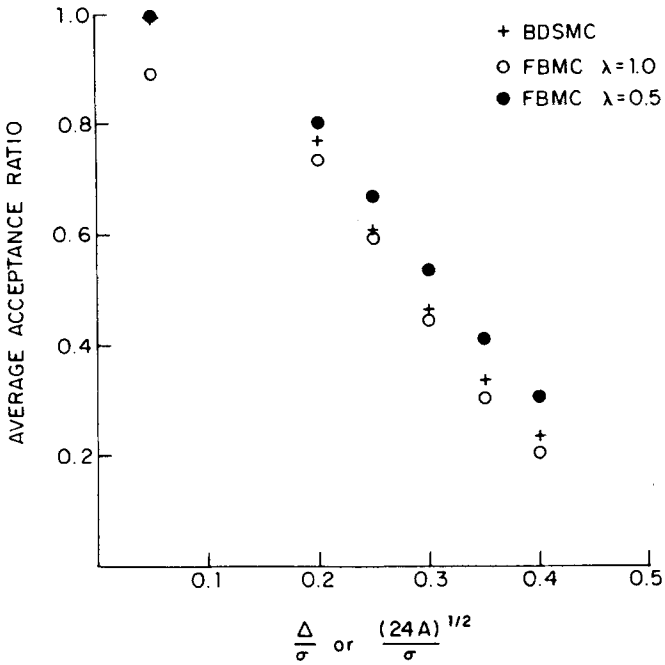

FIG. 1. The mean acceptance ratio during the Monte Carlo simulation of 200 passes as a function of the step size. The dots denote FB with $\lambda=\frac{1}{2}$. The circles denote FB with $\lambda=1$. The crosses denote SMC.

for large step sizes, where the diffusion is optimum, the FB with $\lambda=1$ is as good as the SMC and the FB with $\lambda=\frac{1}{2}$ is better yet.

In the FB method there are two freely adjustable parameters, namely, λ and the "step size" Δ, whereas in the SMC method there is only one parameter, namely A. Until recently such parameters were chosen in accordance with the rule of thumb: choose the parameters so that the mean acceptance probability is close to 0.5. Recently we have shown[4] that this criterion does not give the optimum choice for the parameters. Instead we have suggested[4] that a good measure of the method is the diffusion of the particles. Thus the parameters Δ (or A) can be optimized to give the largest possible diffusion in a given number of passes. When this is done it is found that the optimum parameters with respect to diffusion give mean acceptance probabilities considerably smaller than 0.5. This is discussed in great detail in Ref. 4. A measure of diffusion is the mean square displacement,

$$\langle \Delta r^2(\tau)\rangle \equiv \frac{1}{N}\sum_{j=1}^{N}\left|\sum_{\alpha=1}^{\tau}\delta r_{j\alpha}\right|^2 , \tag{2.16}$$

where $\delta r_{j\alpha}$ is the displacement of particle j during the αth pass, and the sum gives the total displacement of particle j in τ passes. A pass is by definition one cycle of moves in which a move is attempted on each of the N particles. It can be shown that for large τ,

$$\langle \Delta r^2(\tau)\rangle \to 6\,\Gamma\tau , \tag{2.17}$$

where Γ is a diffusion coefficient. Thus we must choose Δ (or A) to give either the largest possible $\langle \Delta r^2(\tau)\rangle$ for a preselected τ, or equivalently to give the largest possible diffusion coefficient.

If each particle move is uncorrelated then it is easy to show that

$$\Gamma \cong \tfrac{1}{6}\langle \delta r^2\rangle p , \tag{2.18}$$

where $\langle \delta r^2\rangle$ is the mean square displacement of a particle in one move and p is the mean acceptance probability. Then it is clear that the method which generates large moves with small acceptance ratio may well win out over a method that generates small moves with large acceptance ratio. Unfortunately, in condensed systems the displacements are highly correlated (see Ref. 4) so that the diffusion coefficient is much more complex than Eq. (2.18) and we cannot fully analyze it.

The point of view taken in the next section is to compare various features of the FB and SMC Monte Carlo walk using diffusion as a measure of their efficiencies.

III. RESULTS

Both FB and SMC methods are used to study a Lennard-Jones liquid at 84°K. It is assumed that all atoms interact via a pairwise potential,

$$\phi(r) = v(r) - v(r_0) ,$$

where

$$v(r) = 4\epsilon\left\{\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right\} \quad r \le r_0 = 2.5$$
$$= 0 \quad r \ge r_0 . \tag{3.1}$$

The particles are placed in a periodic box of size $(6.54\times 6.54\times 6.54)\sigma^3$ corresponding to a liquid density of $\rho\sigma^3 = 0.77$. The starting configuration was obtained from a previous equilibrated liquid simulation at this density

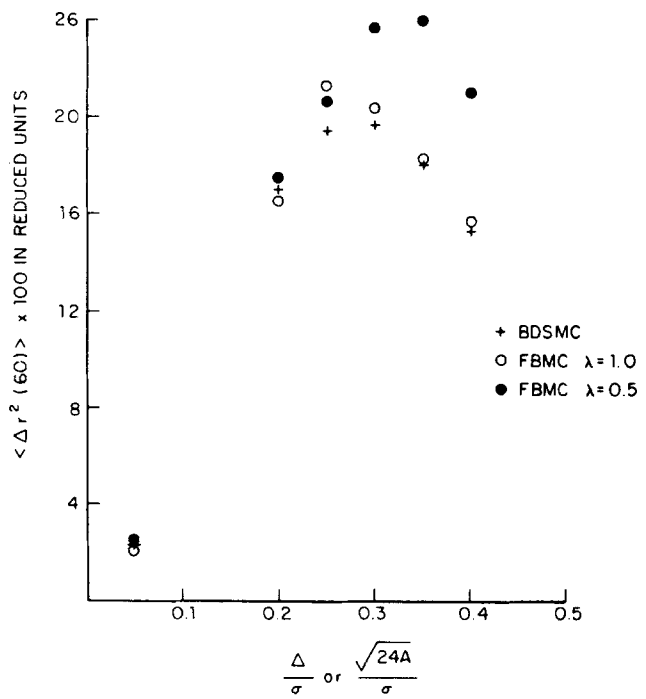

FIG. 2. The mean square displacement $\langle \Delta r^2(\tau)\rangle$ after 60 passes as a function of the step size Δ obtained from two MC simulations of length 200 passes each. The symbols used are the same as in Fig. 1.

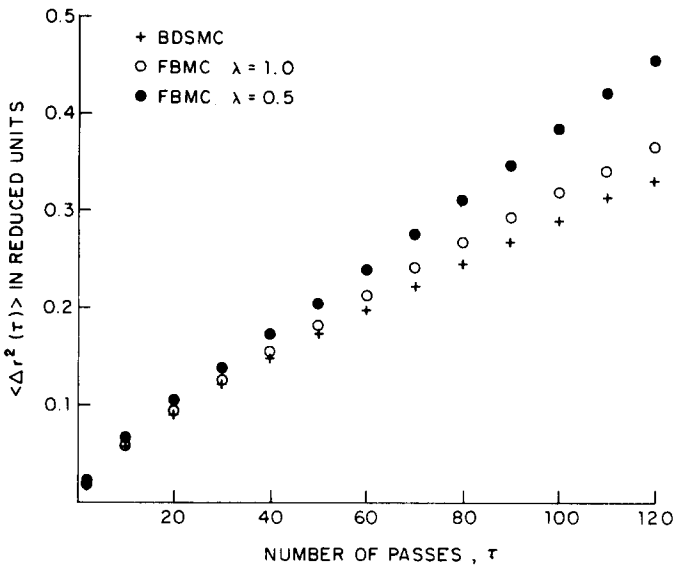

FIG. 3. The mean square displacement $\langle \Delta r^2(\tau) \rangle$ as a function of the number of passes τ. The symbols are the same as in Fig. 1.

and temperature. Starting from the same configuration each time, various Monte Carlo walks were carried out using the prescriptions given in Eqs. (2.4) and (2.5) for various values of the parameters λ, Δ and A. One particle was moved at a time. Each run consisted of 200 passes and in each run the average acceptance ratio and the mean square displacement $\langle \Delta r^2(\tau) \rangle$ defined in Eq. (2.16) were monitored.

Figure 1 shows the mean acceptance ratio P_{FB} and P_{SMC} as functions of Δ for the two methods where in the SMC we take $A = \Delta^2/24$ in accordance with Eq. (2.14). The circles denote FB with $\lambda = 1$, crosses denote SMC and the dots denote FB with $\lambda = \frac{1}{2}$. The acceptance ratio is larger for the FB method with $\lambda = \frac{1}{2}$ as discussed in the last section. However, the acceptance ratios for larger values of Δ (and thereby A) are not dramatically different for FB with $\lambda = 1$ and SMC. All distances are measured in units of the Lennard-Jones parameter σ.

In Fig. 2 we present $\langle \Delta r^2(\tau) \rangle$ as a function of Δ (or A) for $\tau = 60$, averaged over 60 origins. Indeed the trend is the same for all values of τ, though we chose to present only values for $\tau = 60$. The diffusion in all cases passes through a maximum, and optimal diffusion occurs for different values of Δ or A with different methods. A step size of $\Delta = 0.35$ and $\lambda = \frac{1}{2}$ with FBMC is found to be the most efficient method for this example. For the optimal values of Δ (and A) for the three methods, the mean square displacement is plotted in Fig. 3 as a function of τ. Again it is clear that asymptotically the FB walk with $\lambda = \frac{1}{2}$ wins over the other two methods. For the optimum walks, it is hard to distinguish the FB ($\lambda = 1$) from the SMC. A quick glance at the figures shows that when the diffusion is optimized the acceptance probability is less than 0.5.

These results show that even in a system as simple as the Lennard-Jones on the FB ($\lambda = \frac{1}{2}$) method generates a more efficient Monte Carlo walk than any method yet devised, including the SMC. What is meant by this is that for a given amount of computer time the FB walk covers a larger region of the accessible configuration space and thereby generates more statistically uncorrelated configurations than the other method. It is important to note that the diffusion is maximized for fairly large values of the step size and hence the gradient terms in Eqs. (2.11) and (2.12) cannot be neglected. Further improvements may be attainable by including higher derivatives of the potential in the sampling scheme.

In more complex systems like aqueous media, the efficiency achieved by the FB method, in comparison to the Metropolis method, has been demonstrated. For such systems where molecules must be translated and reoriented Brownian dynamics has not yet been cast in an applicable form. The comparison presented here for simpler systems leads us to believe that further efforts to apply SMC to more complex systems may not be worth the extra work.

In conclusion we note that any comparison of different Monte Carlo schemes is system dependent. We have shown in this and preceding papers[3,4] that the FB scheme has certain advantages over other schemes for aqueous solutions and for simple fluids. An important aspect of this work is the suggestion that diffusion can be used to optimize the choice of the step size. This enables us to use larger step sizes and to generate more efficient walks than in the past. It may well be that other criteria will prove to be even better. Unfortunately these do not yet exist.

ACKNOWLEDGMENT

The authors wish to acknowledge many useful discussions with C. Pangali.

[1] N. Metropolis, A. W. Metropolis, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[2] J. P. Valleau and S. G. Whittington, in: *Modern Theoretical Chemistry*, Vol. 5, "Statistical Mechanics, Part A: Equilibrium Techniques," ed. B. J. Berne (Plenum, New York, 1977.)
[3] C. Pangali, M. Rao, and B. J. Berne, Chem. Phys. Lett. **3**, 413 (1978).
[4] M. Rao, C. Pangali, and B. J. Berne, Mol. Phys. (in press).
[5] P. J. Rosky, J. D. Doll, and H. L. Friedman, J. Chem. Phys. **69**, 4628 (1978).