PHYSICAL REVIEW E, VOLUME 65, 036709

# Self-referential method for calculation of the free energy of crystals by Monte Carlo simulation

C. Daniel Barnes and David A. Kofke
*Department of Chemical Engineering, University at Buffalo, State University of New York, Buffalo, New York 14260-4200*

(Received 15 October 2001; published 6 March 2002)

We propose a Monte Carlo simulation method for the evaluation of free energies in crystalline systems. In principle, the method involves evaluating the free-energy difference between systems of $N$ molecules and $2N$ molecules. This difference, coupled with the assumption that the free energy is extensive and thus proportional to $N$, provides sufficient information to obtain the absolute free energy of the crystal. The approach to doubling the system size does not involve insertion or removal of molecules in the system. Instead, the configurations of the molecules are expressed in terms of the normal-mode coordinates of a harmonic lattice. By decoupling certain of these coordinates from the molecule configurations, we obtain a transformation that in effect yields the system-size doubling. The method is examined via application to a system of hard rods in one dimension. This simple model is considered principally because of the availability of an analytic solution for its free energy, which permits accurate testing of the performance and correctness of the proposed method. In using the hard-rod model we also avoid other complications related to treatment of the temperature, and application of normal-mode coordinate decoupling in higher dimensions. The proposed method is shown to be able to provide good results for the free-energy calculation, but further development will be needed before it can be considered practical for general-purpose use.

DOI: 10.1103/PhysRevE.65.036709 PACS number(s): 02.70.Rr, 05.10.Ln, 65.40.Gr, 63.20.Dj

## I. INTRODUCTION

The development of methods for the calculation of the free energy of solids continues to be of interest for several reasons. Improved understanding of conditions promoting crystallization is important in the petrochemical industry where wax formation in oil wells, storage tanks, and pipelines causes regular shutdowns at substantial expense. Crystallization is becoming increasingly important in the specialty chemical and biochemical industries as a method of separation for components that have low-melting temperatures or where distillation would break down the products before the boiling point is reached. Product design can also benefit from an understanding of equilibria in solid phases. Semiconductor alloys exhibit miscibility gaps that strongly influence the quality of the product, depending on production conditions. All of these phenomena are governed by the free energy of solid phases relative to fluids, or other crystalline solid forms. Molecular modeling of such systems can be advanced through the improvement of methods for evaluating solid-phase free energies by molecular simulation.

We begin by pointing out some of the difficulties involved in measurement of solid-phase free energies by molecular simulation, with the idea that a discussion highlighting the problems will aid in the formulation of their solution. However, in the present paper, we do not succeed in overcoming these difficulties, but we see these efforts as proving a path that might achieve this aim.

Existing methods for measuring free energies of crystalline phases are not applied as easily as those developed for fluids. Fluid-phase systems can rely on the Euler relation [1] for free-energy calculation

$$A \equiv U - TS = -pV + \mu N, \tag{1.1}$$

where $A$ is the Helmholtz free energy, $U$ and $S$ are the internal energy and the entropy, respectively, $p$ is the pressure, $V$ the volume, and $\mu$ and $N$ are the chemical potential and number of molecules. In this equation, $p$, $V$, and $N$ are known or easily measured in simulation [2,3], and $\mu$ too can be evaluated reliably, albeit with more computational effort [4]. The amount of computation required for measurement of a free-energy difference scales exponentially with the entropy difference between the target and reference systems [5,6]. The chemical potential is a tractable quantity because in a fluid it can be determined by a process in which a molecule is inserted into the simulated system [7–9]. The accompanying entropy change does not scale with the system size, and consequently the free-energy measurement can be accomplished with a manageable amount of computation regardless of the number of molecules simulated. In contrast, direct measurement of the free energy itself, for example by computing its difference with respect to an ideal gas, yields a value that scales as $N$ because the free energy is extensive. Since the entropy difference too scales with $N$, we find that extensive free-energy measurements can be quite expensive. Thus for fluids, the ability to evaluate an extensive free energy via measurement of an intensive free-energy difference is of great utility.

Crystalline phases cannot rely on the Euler equation to evaluate the free energy, because the chemical potential is not available as it is for fluids. Insertion of a molecule into a crystalline phase leads to a defect, and for the small systems studied by simulation this single defect corresponds to a concentration that is not representative of that expected in a bulk crystal. Thus, solid phase free energies must be measured directly, by calculation of a difference with a reference system [10]. Usually this reference is selected to be some sort of perfect noninteracting crystal, and the free-energy difference is evaluated by thermodynamic integration along a reversible path joining the reference and target systems. The most widely used technique of this type is due to Frenkel and

1063-651X/2002/65(3)/036709(9)/$20.00

Ladd [11], who employ an Einstein crystal. More recently methods have been put forth to join phases of different crystalline symmetry [12,13], which can be useful to study polymorphism.

One concern with the existing methods for free-energy measurement in crystals is the possible occurrence of a phase transition along the path of integration. The presence of such a transition invalidates the use of thermodynamic integration, and can make other free-energy difference methods difficult or impossible to apply. Another issue is one of efficiency. The simulations devoted to traversing the path from the reference to the target are of no use for other property measurements; such simulations have no intrinsic interest aside from their use in measuring the crystalline free energy. This contrasts with the situation in the fluid, in which the molecule-insertion free-energy measurement can in many cases be accomplished while simulating the system of interest. Related to this is the ability to perform direct measurement of phase coexistence in fluid phases via the Gibbs ensemble [3,14,15]. Such a method is inapplicable to equilibria involving solids.

In this paper we describe a free-energy calculation method for solids. The approach is based on measuring the free-energy difference between a $2N$-particle system and an $N$-particle system plus an $N$-particle Debye harmonic lattice. The trick in doing this is to use harmonic normal-mode coordinates to parameterize the particle positions in the simulated system. In this way we can in effect increase the system size without inserting particles or introducing spurious boundary effects. The complete set of harmonic normal modes for an $N$-particle system forms a subset of the normal modes of the $2N$-particle system, making it possible to construct a free-energy path between them by switching on the modes not common to both system sizes. Further, since the free energy is an extensive quantity, and presuming we can ignore finite-size effects (as is reasonable for sufficiently large systems), this free-energy difference is equivalent to the free energy of the $N$-particle system.

This approach requires sampling of configurations in which certain normal modes of the $2N$-particle system are zero, and consequently it requires trial displacements in normal-mode space instead of the more conventional trial particle displacements in one-dimensional Cartesian space. This does not affect the sampling of the $2N$-hard-rod configuration space except to fix the system's center-of-mass (c.m.), which is addressed below. Although the sampling occurs in normal mode space, the test for accepting or rejecting a configuration is done conventionally by computing the energy expressed via Cartesian coordinates.

This self-referential method to evaluate the free energy is appealing because it circumvents some of the problems outlined above for the existing methods. The perturbation path never really departs from the target system, so we need not be concerned with encountering a phase transition along the way, and at all times other properties of the target system remain accessible to measurement. However, this method does not avoid the problem with measuring an extensive free energy, and moreover, the need to sample in normal-mode space is a complicating factor. We believe that the ideas underlying this method can lead to alternatives for solids that are analogous to insertion methods that work so well for fluids. If the method can be shown to be viable for general systems, the programming and computational expense involved with sampling normal-mode coordinates would be worthwhile.

Our approach is developed and demonstrated with Monte Carlo simulation applied to a canonical (NVT) ensemble of one-dimensional hard rods with periodic boundaries. The method can be extended to other model systems and to higher dimensions. The hard-rod system is convenient because there exists an analytical solution for its free energy [16], correct for any system size, which enables direct comparison of results of the proposed method with the exact free energy

$$e^{-\beta A_N(\beta,L)}=Q_N(\beta,L)=\frac{L(L-N\sigma)^{N-1}}{N!}. \tag{1.2}$$

Here, $A_N$ and $Q_N$ are the Helmholtz free energy and the canonical partition function, respectively, for a system of $N$ hard rods each of length $\sigma$ in a periodic container of length $L$, and $\beta=1/kT$ with $k$ Boltzmann's constant and $T$ the absolute temperature. The ensemble-averaged potential energy of this model is zero, so the free energy and entropy are equal; this feature eliminates some complications that we prefer to avoid at this point in the development of the methodology. The one-dimensional (1D) system does not undergo any phase transitions, so one might argue that there is no solid phase, but the system does exhibit the crucial feature that the particles occupy a "lattice," inasmuch as for all relevant configurations each particle may be unambiguously associated with a 1D lattice site.

In the following section we develop the formalism for the proposed method. Monte Carlo simulation details are described in Sec. III. We present and discuss the results in Sec. IV and give our conclusions in Sec. V.

## II. FORMALISM

As is well known [17], the energy of a system of $N$ particles interacting harmonically with force constant $\omega$

$$U(\gamma^N)=\sum_{i=1}^{N}\sum_{j=nbrs}\frac{\omega}{2}[\gamma_i-\gamma_j]^2, \tag{1.3}$$

where $\gamma_i$ is the Cartesian position vector of particle $i$ relative to a reference (lattice site) position, can be recast as a sum of $DN$ noninteracting harmonic oscillators, where $D$ is the spatial dimension of the system

$$U(\eta^{DN})=\frac{\omega}{2}\sum_{m=1}^{DN}\lambda_N(m)\,\eta_m^2. \tag{1.4}$$

The coordinates $\eta_m$ in this reformulation describe collective motions of the particles, such that the $k$th Cartesian deviation $(k=1\cdots DN)$ is given as

$$\gamma(k)=\sum_{m=1}^{DN}\eta_m\phi_N(m,k), \tag{1.5}$$

where the vectors $\phi_N(m,k)$ are the normal modes for the system. These modes can be expressed as the eigenvectors of the harmonic interaction matrix, and their corresponding eigenvalues $\lambda_N(m)$ are the harmonic force constants for the collective motion described by the normal-mode coordinate $\eta_M$ in a harmonic system of unit force constant. If the minimum-energy configuration of the harmonic oscillators forms a lattice, the normal-mode coordinates are given by the reciprocal lattice vectors [18]. Each plane-wave vector in a small system will have a counterpart in a larger system that points in the same direction and has the same wavelength. We can in effect "grow" the system by coupling to the small system those modes that are found only in the larger one. To fix ideas, and for the purpose of demonstrating and testing the methodology, we henceforth restrict our consideration to periodic one-dimensional systems, for which

$$\phi_N(m,k)=N^{-1/2}[\sin(2\pi mk/N)+\cos(2\pi mk/N)] \quad m,k=1...N \tag{1.6}$$

and

$$\lambda_N(m)=2\left[1-\cos\left(\frac{2\pi m}{N}\right)\right]. \tag{1.7}$$

We categorize the normal modes of a one-dimensional system of $2N$ particles as *even* or *odd*. The even modes are symmetric about the center of the system, while the odd ones are antisymmetric; we number them accordingly with even and odd indices

$$\begin{aligned}\phi_{2N}(m,k)&=\phi_{2N}(m,k+N) \quad m \text{ even},\\ \phi_{2N}(m,k)&=-\phi_{2N}(m,k+N) \quad m \text{ odd}.\end{aligned} \tag{1.8}$$

Each normal mode, or eigenvector, of the *N*-particle system is directly proportional to an *even* numbered normal mode of the $2N$ system. This is pictured in Fig. 1, showing that mode 1 of the *N*-particle system has precisely the same frequency as that of mode 2 of the $2N$ system. Moreover, the amplitude of mode 1 of the $N$ system is that of mode 2 of the $2N$ system multiplied by $2^{1/2}$. All modes of the $N$ system have exactly this relationship to a corresponding mode in the $2N$ system, specifically, mode *m* of the $N$ system corresponds to mode $2m$ of the $2N$ system, so that $\phi_N(m,k)=2^{1/2}\phi_{2N}(2m,k)$. The odd modes of the $2N$ system have no counterparts in the $N$ system, and contributions to the Cartesian $\gamma$ from these modes are what set the configurations of the $2N$ system apart from those of the $N$ system.

Let us consider the partition function of the $2N$ system, expressed in terms of normal-mode coordinates

$$Q_{2N}=\frac{1}{\Lambda^{2N}}\int d\eta^E d\eta^O e^{-\beta U(\eta^E,\eta^O)}, \tag{1.9}$$

where the superscripts "$E$" and "$O$" on $\eta$ indicate the set of even and odd normal-mode coordinates, respectively, and $\Lambda$ is the thermal de Broglie wavelength. The limits of integration are over all normal-mode coordinate values that have all particles in the system volume, and we note that the Jacobian for this transformation is unity. Let us now decouple the odd-mode coordinates $\eta^O$ from the Hamiltonian of the system of particles, and have these coordinates instead couple to a collection of independent harmonic oscillators. In the resulting *hybrid* system formed from the original particles and these oscillators, the particle configurations will be constrained to a subset of those originally available. These configurations will exhibit the symmetry of the even modes, and consequently they will have particles in one half of the system occupying the same relative positions as particles in the other half. In addition, because all modes of the $N$ system are represented by the even modes of the $2N$ system, the complete set of $2N$ even modes can describe all configurations of the *N*-particle system in each half. But because each half moves identically, the configuration-set accessible to the odd-mode-decoupled (hybrid) $2N$ system is exactly that of the $N$ system. Since the system size is doubled, the energy of each configuration will be exactly double that of either half, assuming periodic boundaries are employed. Consequently the hybrid $2N$-partition function (superscript*) is related to that of an unmodified *N*-particle partition function

$$\begin{aligned}Q_{2N}^{*}(\beta,2L)&\equiv\lambda^{-2N}\int_{2L}d\eta^E\int d\eta^O\exp[-\beta U_{2N}(\eta^E)]\\ &\quad\times\exp\left[-\frac{1}{2}\beta\omega\sum_{m,\mathrm{odd}}\lambda_N(m)\,\eta_m^2\right]\\ &=\Lambda^{-2N}\int_L d\eta^E\exp[-2\beta U_N(\eta^E)]\int d\eta^O\\ &\quad\times\exp\left[-\frac{1}{2}\beta\omega\sum_{m,\mathrm{odd}}\lambda_N(m)\,\eta_i^2\right]\\ &=2^{-N/2}Q_N(2\beta,L)Q_N^O(\beta),\end{aligned} \tag{1.10}$$

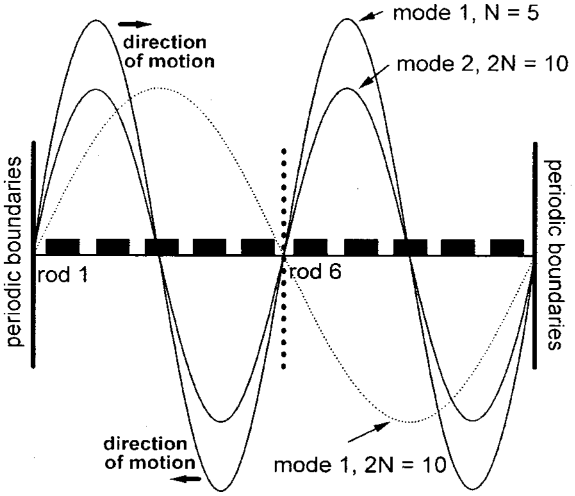

FIG. 1. Examples eigenvectors for a system of $N$ and $2N$ hard rods which describe the motion of the hard rods. Each curve describes the amount the rod at any position is displaced as the coordinate $\eta$ for the mode varies. The frequencies of mode 1 of the $N$ system and mode 2 of the $2N$ system are identical. Mode 1 and the other odd modes of the $2N$ system have no counterparts in the $N$ system.

where the factor of $2^{-N/2}$ arises from the Jacobian associated with expressing the configurations of the $N$ system in terms of the normal mode coordinates of the $2N$ system. $Q_N^O$ is the partition function for the odd-mode harmonic oscillators, including a factor of $\Lambda^{-N}$, it is given exactly as

$$Q_N^O(\beta)=\frac{1}{\Lambda^N}\prod_{m,\mathrm{odd}}\left(\frac{2\pi}{\beta\omega\lambda_N(m)}\right)^{1/2}. \tag{1.11}$$

Thus, the free energy of an $N$-rod system can be given in terms of the hybrid $2N$ system at twice the temperature. Presently we will circumvent the temperature issue by dealing only with the hard-rod potential, where the configurational energy may have values of zero or infinity only. In Sec. V we discuss extensions from this simplifying case.

If it may be taken that the free energy is an extensive quantity, we have

$$A_N(\beta,L)=A_{2N}(\beta,2L)-A_N(\beta,L). \tag{1.12}$$

Combined with Eq. (1.8) via the canonical-ensemble bridge equation [19]

$$\begin{aligned}\exp[-\beta A_N(\beta,L)]&=\frac{Q_{2N}(\beta,2L)}{Q_N(\beta,L)}\\&=2^{-N/2}\frac{Q_N^O(\beta/2)Q_{2N}(\beta,2L)}{Q_{2N}^*(\beta/2,2L)}.\end{aligned} \tag{1.13}$$

Equation (1.13) forms the basis of the proposed method for calculating the free energy of a crystalline phase. The ratio of partition functions can be evaluated by molecular simulation via any of a number of standard methods. We will consider the application of free-energy perturbation.

We note also that the desired free energy can be obtained from the simulation data via this relation

$$\begin{aligned}\exp[+\beta A_N(\beta,L)]&=\exp[+\beta A_{2N}(\beta,2L)]\frac{Q_{2N}(\beta,2L)}{Q_N(\beta,L)}\\&=\exp[+\beta A_{2N}(\beta,2L)]2^{-N/2}\\&\quad\times\frac{Q_N^O(\beta/2)Q_{2N}(\beta,2L)}{Q_{2N}^*(\beta/2,2L)}.\end{aligned} \tag{1.14}$$

This formulation has the advantage of not relying on the extensive nature of the free energy, but it is useless practically since it requires prior knowledge of $A_{2N}$. We present it here only because, given the exact expression for $A_{2N}$ for hard rods, it can be used to test whether the simulation provides correct values of the partition-function ratio without invoking also the extensive free-energy approximation.

## III. SIMULATION METHOD AND DETAILS

We examine a system of hard rods in one dimension, for which the free energy for any $N$ is given exactly by Eq. (1.2).

The ratio of partition functions in Eq. (1.13) describes a system of $2N$ particles with complete freedom, relative to the same system with the odd-mode coordinates decoupled from the particle configurations and instead coupled to a set of harmonic oscillators (there is also the temperature difference, but we can ignore this for the hard-rod potential that forms our present focus). We evaluate this ratio via a multistage free-energy perturbation (FEP) calculation. The procedure may be applied in either of two directions. In one approach, which we will designate as "upstaging," the series proceeds from the hybrid system to the full $2N$ system; in the other, which we designate "downstaging," FEP proceeds from the full $2N$ system to the hybrid. We use $N$ stages, one for each odd-numbered mode of the $2N$ system. In each stage, a single odd normal mode $m$ is selected for a FEP calculation. Once the mode is coupled (upstaging) or decoupled (downstaging) in a FEP calculation, it remains so for the rest of the stages of the series.

If we are upstaging, the perturbation consists of coupling this mode to the configuration of rods, and turning off its harmonic potential. The FEP ensemble average gives the free-energy difference for one stage as

$$\exp[-\beta\Delta A_m]=\left\langle\Theta(m)\exp\left(+\frac{1}{2}\omega\lambda_m\eta_m^2\right)\right\rangle_{m-1}. \tag{1.15}$$

In this formula, $\Theta(m)$ is a type of Heaviside function: it is zero if coupling mode $m$ to the hard-rod system causes any rod overlaps or misorderings, and it is unity otherwise; this implements the hard-rod potential. The "$m-1$" subscript on the ensemble average indicates that hard-rod configurations are sampled using all even normal modes and all odd modes that became coupled in previous FEP stages in the series, and that mode $m$ and all others are decoupled from the rod configurations. Of these decoupled modes, only mode $m$ is relevant to the FEP calculation. When the perturbation is performed, the value of the coordinate $\eta_M$ is selected from a Gaussian distribution, with probability density proportional to $\exp(-\omega\lambda_m\eta_m^2/2)$. The total free-energy difference for converting the hybrid system to a fully coupled $2N$ system is obtained by summing the Eq. (1.15) free energies over all stages.

If we are downstaging, the perturbation consists of decoupling the mode from the configuration of rods, and turning on its harmonic potential. The appropriate ensemble average is

$$\exp[-\beta\Delta A_m]=\left\langle\Theta(m)\exp\left(-\frac{1}{2}\omega\lambda_m\eta_m^2\right)\right\rangle_m. \tag{1.16}$$

Here, $\Theta(m)$ functions again to eliminate from the average any perturbations that lead to hard-rod overlaps or misorderings. The system samples configurations obtained from all even modes, as well as the odd modes (including $m$) that have not yet been decoupled from the rods in the perturbation series. The Gaussian term for normal-mode coordinate $\eta_m$ in this average uses whatever value of $\eta_m$ happens to be in place when the perturbation is applied to decouple $\eta_m$ from the hard-rod configuration. For a downstaging perturbation, we can easily examine a large range of force constants $\omega$ for the harmonic reference with each perturbation. Because $\eta_m$ for the perturbation is given by its present value, examination of many force constants simply entails evaluation of the Gaussian term in Eq. (1.16) for each. This procedure contrasts with an upstaging perturbation, where to examine many force constants at once, we must generate for each a random value of $\eta_m$ and couple it to the hard rods and check for overlap. It is helpful to be able to consider many harmonic references, because force constants that are too small or too large will lead to inaccurate FEP averages due to poor sampling, and the optimum force constant may be difficult to identify *a priori*.

There are two types of moves used in each simulation, the perturbation move between the reference and target systems mentioned above, and a trial normal-mode displacement move used to generate configurations of the rods. The normal-mode displacement is the random displacement of a randomly chosen mode, and is analogous to the typical trial particle displacement in Cartesian space. All rods move collectively in response to a normal-mode displacement trial. This is easily accomplished by displacing each rod in accordance with the normal-mode vector's contribution to its position (as opposed to regenerating the configuration by summing all the coupled modes). Candidate modes include all the even modes and those odd modes that are presently coupled to the rods configuration. The even mode $m=0$, which corresponds to uniform translation of the system of rods, is excluded from sampling. This has the consequence of fixing the center of mass of the rods. The effect of this constraint on the free energy can be evaluated exactly for the hard-rod system, and we include this consideration in our comparison of the simulation and exact free energies.

The choice to perform a FEP trial or a trial in the Markov chain is made at random, with probability such that one of every $2N$ trials is a FEP trial move. If the move selected is a Markov trial phonon displacement, a mode from the set coupled to the hard-rods configuration is selected, each with equal probability. The trial displacement for that mode is determined randomly according to a uniform distribution over some step size. In a conventional Monte Carlo displacement trial, the maximum step size is the same for identical particles. This is not appropriate for normal-mode displacements because each mode contributes differently to the particle motion. The long-wavelength modes describe a motion where adjacent particles move in a similar fashion, (i.e., for the translation mode, all particles move an equal distance in the same direction) while the short-wavelength modes represent a motion where the distances adjacent particles move may vary significantly, and possibly in opposite directions. Applying the same step size to all modes would not result in an equal acceptance rate of trial configurations between the modes. We showed [20] that the standard deviations of the singlet probability density distributions of the hard-rod and harmonic systems are identical and proportional to the inverse square root of their eigenvalues. Correspondingly, we have also observed that the set of step sizes giving the same trial configuration acceptance rate between the modes to be proportional to the inverse square of the eigenvalues. This allows us to use a single global step size which when divided by the square root of the eigenvalue of the selected mode yields an appropriate trial step size. We chose a global step size to give an acceptance rate of 35%. The current value of the selected mode is shifted by the trial displacement to generate a trial configuration, similar to adding a trial particle displacement to the current particle position in Cartesian coordinates. The trial configuration is transformed into Cartesian coordinates to test for overlap of the hard rods, rejecting the configuration if overlap or misordering occurs, and accepting it otherwise.

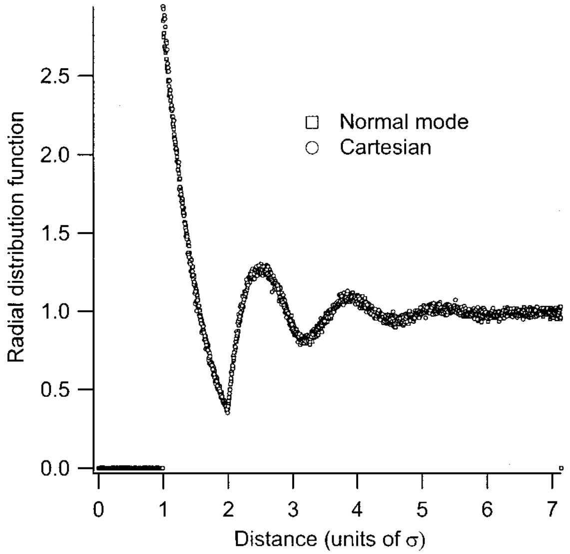

FIG. 2. Radial distribution function for a system of ten hard rods at density $\rho^*=0.7$, evaluated using conventional Monte Carlo sampling in Cartesian coordinates, and sampling in normal-mode coordinates.

We performed simulations to measure the free energy of a system of hard rods using the proposed method, considering both upstaging and downstaging approaches. We examined system sizes ranging from $N=4$ to $N=100$. For each system, we performed a series of $N$ FEP stages. In each stage we performed $10^5$ Monte Carlo cycles, where each cycle consists of $2N$ trials of a phonon displacement or FEP test trial in proportion as described above. We examined a very broad range of harmonic force constants with each FEP stage.

## IV. RESULTS AND DISCUSSION

Because we use a nonstandard method of sampling, we are interested in verifying that configurations important to the hard-rod system are properly sampled. To do this, we compare the radial distribution function measured in a normal-mode simulation with that from a conventional Monte Carlo simulation. The results for a system of ten rods at a density of $\rho^* \equiv N\sigma/L = 0.7$ are shown in Fig. 2. The curves of the two simulations are essentially indistinguishable, lending confidence to the normal-mode sampling method.

The approach to correcting the free energies for the c.m. motion is described in the Appendix. The correction requires knowledge of the c.m. distribution in an unconstrained sys-

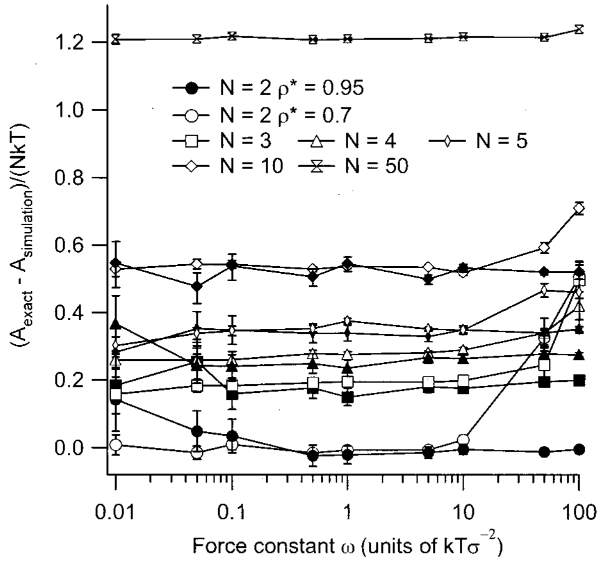

FIG. 3. Error in the free energy given by the upstaging self-referential method, as a function of harmonic force constant $\omega$ for hard-rod systems of various sizes at densities $\rho^* = 0.7$ and 0.95, respectively.

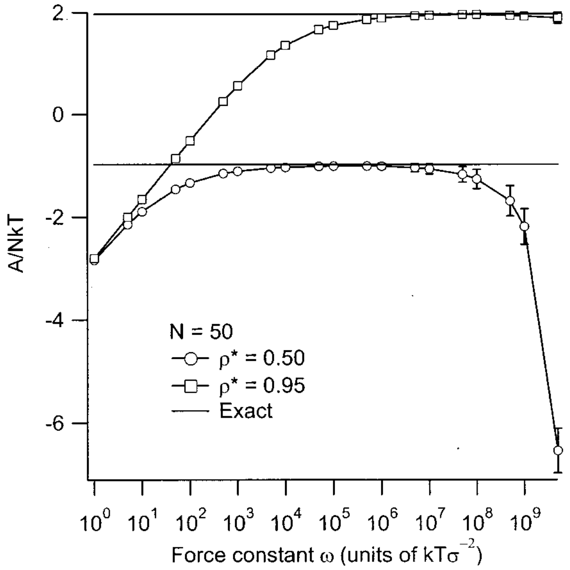

FIG. 4. Free energy given by the downstaging self-referential method, as a function of harmonic force constant $\omega$ for hard-rod systems of $2N = 100$ particles at densities $\rho^* = 0.7$ and 0.95, respectively.

tem. In the Appendix we compare the distribution obtained by standard, Cartesian-space Monte Carlo sampling with the analytic expressions used to apply the correction. The simulation data are noisy, but are consistent with the analytic forms. For $N$ larger than 50 or so, the c.m. correction is smaller than the confidence limits of the measured free energies.

Figure 3 plots the difference between the exact free energy given by Eq. (1.2), and the simulation result based on Eq. (1.15), as a function of the harmonic force constant $\omega$ for various system sizes at densities $\rho^*$ of 0.7 and 0.95 (we note that $\rho^* = 1.0$ is close packing). This is an upstaging FEP series. The simulation results exhibit a systematic error that increases with increasing system size, and which appears to be unaffected by density. A reasonable explanation for the error is that in some or all of the FEP stages there are configurations of the target that are not adequately sampled by the reference system. We examine this question further below. We have not pursued this method further as we consider upstaging to be a less effective approach than downstaging, owing to the inability of upstaging to sample as easily a broad range of force constants.

Figure 4 plots the free energy per particle as a function of the harmonic force constant for simulations using Eq. (1.16). This is a downstaging series. Here there exists a broad range of harmonic force constants for which the simulation results are in very good agreement with the exact free-energy value. We have compared the simulation data to the exact values expected from Eq. (1.2), using both Eq. (1.13) which is based on assuming an extensive free energy, and (not shown in Fig. 4) Eq. (1.14) which is less practical but invokes no such approximation. For $N = 100$ and the higher densities, the difference between these comparisons is not significant. The simulation data are consistently lower than the exact value taken from either formula, but over significant ranges of force constant they agree to within the confidence limits of the simulation data.

It is of interest to examine the free-energy change associated with each stage of the FEP series. One might consider whether the contribution from each stage depends on the wavelength of the normal-mode coordinate being decoupled from the rod configurations. Equation (1.6) indicates that each wavelength is doubly degenerate, with the middle modes having the shortest wavelength (mode 99 for $2N = 200$), and the end modes having the longest (modes 1 and 199). We ran three simulation series, changing the order of decoupling of the modes. First we turned them off in numerical order, from long to short wavelength for mode 99, back to long wavelength. We also ran simulations turning the modes off from long to short wavelength, and finally turning the modes off from short to long wavelength. In Fig. 5 we plot the entropy change per stage for these three simulations as a function of the number of modes still coupled, so that the data at the right correspond to the first perturbation of the series and the data on the left the last stage of the series. We present this in terms of the entropy because it is useful for discussion that follows below; the free-energy change is trivially related by a sign change and a vertical shift of 0.5 to account for the harmonic energy. We do not find a systematic dependence of the free-energy change on the frequency of the mode. Instead, we observe that the free-energy change associated with the decoupling of a mode depends on the order in which it lies in the sequence. The first decoupled mode, for example, contributes a constant amount to the overall free-energy difference, regardless of which mode it is.

Finally, we turn to the topic of the accuracy of these calculations. We have argued elsewhere [5,6,21,22] that FEP calculations must be performed in a direction such that con-

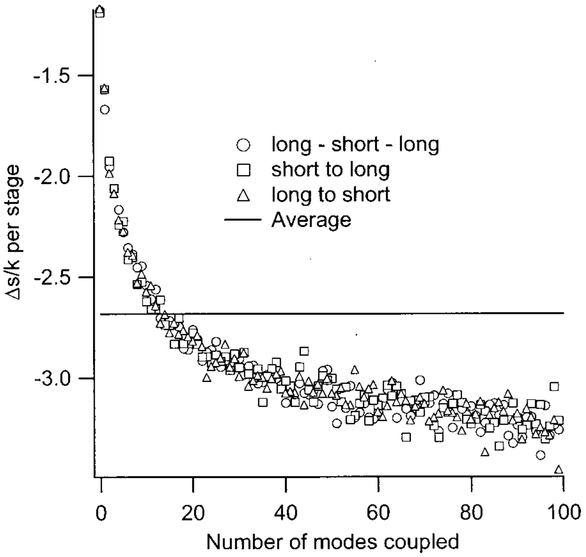

FIG. 5. Entropy change per stage for downstaging series, as a function of the number modes not yet uncoupled. System density is $\rho^*=0.95$ and harmonic force constant is $\omega=5\times10^5 kT\sigma^{-2}$; the number of rods in simulation is $N=100$.

figurations important to the reference system encompass all the configurations important to the target. A necessary condition for this to hold is that the entropy of the reference be greater than the entropy of the target, and thus the entropy change for the perturbation be negative. This condition is not sufficient however, inasmuch as it can happen that the reference and target have comparable-sized important configuration-space volumes which nevertheless do not completely overlap. However, the larger and more negative the entropy difference, the more likely it is that the target configuration set is wholly contained in the reference configuration set. In this FEP method, we have the ability to tune the harmonic system so that the target can always be made to have a smaller entropy than the reference. For the upstaging calculation, we require that the harmonic potential be sufficiently broad to encompass all the relevant contributions of the perturbation mode to the hard-rod system as it is coupled to it. Likewise, for the downstaging calculation, we need the harmonic potential to be narrow, so that it is well contained within the range explored in the hard-rod system as the mode is decoupled. Figure 6 shows an average measure of the entropy change per stage for several densities, computed from the exact formulas for the hard-rod and harmonic free energies, as a function of the harmonic force constant. The curve shows that the entropy difference for the highest-density system is substantially different from zero for force constants greater than about $10^5$ (appropriate for downstaging) or less than about $10^{-1}$ (appropriate for upstaging). Perhaps the poor quality of the upstaging results has something to do with the insufficiently soft harmonic systems examined in those simulations. The entropy change for each FEP stage, as already presented in Fig. 5, provides a more fine-grained view of the character of the FEP calculation. We must have all stages in an appropriate range of entropy difference to obtain completely accurate results. Clearly, some stages exhibit entropy changes that are about $1.5k$ greater than the average value, thus indicating that the appropriate range of force constant suggested in Fig. 6 provides only a rough guideline, and that more extreme values (softer or harder, depending on whether upstaging or downstaging is being performed) are appropriate.

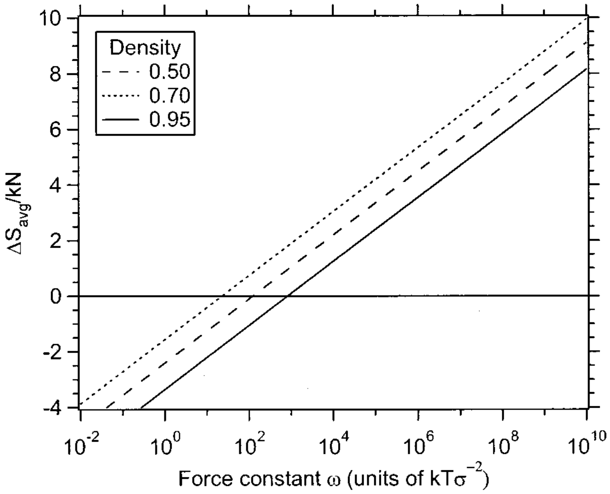

FIG. 6. Average entropy difference per particle for the downstaging process, given as a function of the harmonic force constant.

## V. CONCLUSIONS

We have shown this self-referential approach to be capable of calculating the free energy of a system of hard rods using the downstaging perturbation direction, from the $2N$-hard-rod system to the $N$-hard-rod plus harmonic systems. The sign of the entropy change for each of the perturbation stages must be negative to yield accurate results, and the harmonic system can be adjusted to ensure this criterion is met. We also demonstrated that the correction to the free-energy per stage is not a function of the wavelength of the perturbation mode, but only of the number of modes with which the hard-rod system still interacts. The proposed scheme can quickly become computationally expensive with increasing system size because it calculates $N$ independent averages, one for each stage.

One possible approach to increase the efficiency of the method is to conduct fewer FEP stages, and in each to decouple more than just one normal mode. However, in other work [6] we have shown that the optimal staging method uses intermediate stages with equal entropy differences of $\Delta S/k=-2$. Since the entropy difference for each stage is already in this range, this heuristic would indicate that we cannot improve efficiency by having fewer, coarser stages. Another approach would apply a “parameter hopping” method, in which Monte Carlo trials are performed that lead to fluctuations in the number of coupled modes. One might encounter difficulty with this approach in determining the appropriate weighting function to apply to encourage the system to explore the full range of coupling (from none to all of the odd modes), but this problem is tractable and its examination is worthwhile.

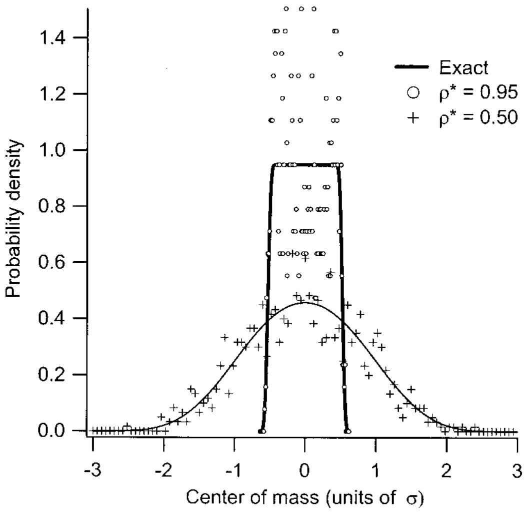

FIG. 7. Configurational center-of-mass distribution for a system of five hard rods. The dark line is the result of an analytic formula, and the light line describes data from a conventional Monte Carlo simulation. (a) Density $\rho^* = 0.5$; (b) $\rho^* = 0.95$.

Certain problems can be anticipated in extending this method to more practically interesting systems. First is the issue of the temperature. We were able to ignore this detail here because of our focus on a hard potential, but for the method to have any real value it must be applicable to soft potentials as well. Our view now is that the temperature must be stepped up or down as the modes are decoupled or coupled to the system of interest. The question that bears examining is whether this process would have any adverse effect on the measurement of other properties. One of the advantages of the proposed method is its avoidance of systems or states that are examined only as a bridge between the target and reference systems, and which are otherwise irrelevant. We do not anticipate problems in this direction, but further study is needed.

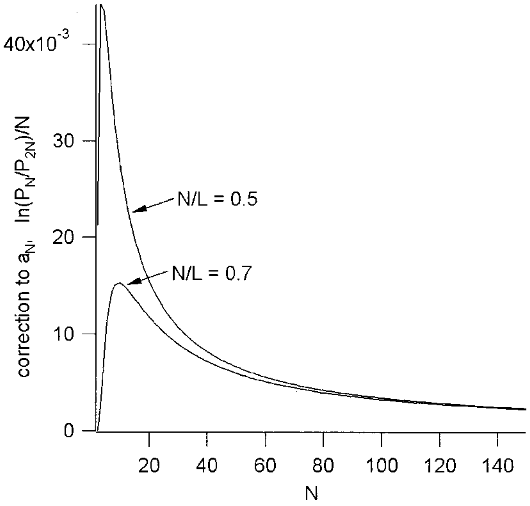

FIG. 8. Additive correction to the free-energy difference per particle for the neglect of center-of-mass motion, as a function of system size for two densities.

The second critical extension of the method is to higher dimensions. The appropriate way to proceed with this is to work directly with reciprocal-lattice vectors as the basis for the particle coordinates [17]. An appropriate set of wave vectors can be selected to double the system size in one direction (along a direct-lattice vector) via sequential coupling of basis functions corresponding to the appropriate ("odd") wave vectors. The overall approach would be much more valuable if it could be developed in a way that does not require complete system doubling, but instead could function by coupling or decoupling a few select modes. Presently it does not seem that this formulation can be accomplished, given the variation of free energy with perturbation mode that was observed in this paper.

The proposed method has demonstrated some promise for development of different avenues to solving the problem of calculating free energies of crystals by molecular simulation. However, further testing and perhaps reformulation of the methodology is warranted before considering the approach to be a viable solution to the problem. The method does at least open up new thinking toward this issue, and can point the way to variants that may be of some utility.

## ACKNOWLEDGMENT

This work has been performed with support from the Division of Chemical Sciences, Office of Basic Energy Sciences, Office of Energy Research of the U.S. Department of Energy (Contract No. DE-FG02-96ER14677).

## APPENDIX

In the normal-mode Monte Carlo simulations, the (even) mode corresponding to uniform translation of the system always remains uncoupled in both the $N$ and $2N$ systems. The contribution of this mode to the free energy differs slightly for the two systems, and for accurate calculations we should include it in the free-energy difference. One way to view this contribution is to consider the distribution $P(x)$ of the center of mass (c.m.) in an unconstrained system. This distribution is shown in Fig. 7 for a system of five hard rods at two densities. The figure compares simulation measurements of this distribution with an analytic form we developed previously [20]. The c.m. exhibits a peak at the center of the periodic system, which is simply a reflection of the fact that there are fewer ways to arrange the rods when the c.m. is constrained to be elsewhere. It is, for example, impossible to arrange the nonoverlapping rods in any way that causes the c.m. to be located at the edges of the boundary. The effect attenuates, and the distribution becomes flatter, with increasing density. The partition function $Q_N(x)$ having the c.m. constrained within $\Delta x$ of some point $x$ is related to the unconstrained partition function $Q_N$ by

$$Q_N(x)=Q_N P_N(x)\Delta x. \tag{1.17}$$

This formula indicates that we can correct for the unequal effect of the c.m. constraint on the $N\rightarrow 2N$ FEP calculation by multiplying by the ratio of the distributions, $P_{2N}/P_N$, where the distributions are evaluated at the center of the system. This correction is plotted in Figs. 7 and 8 as a function of system-size $N$, for two densities. On the whole, the correction is not significant.

[1] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics* (Wiley, New York, 1985).

[2] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).

[3] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic, New York, 1996).

[4] D. A. Kofke and P. T. Cummings, Mol. Phys. **92**, 973 (1997).

[5] N. Lu and D. A. Kofke, J. Chem. Phys. **114**, 7303 (2001).

[6] N. Lu and D. A. Kofke, J. Chem. Phys. **115**, 6866 (2001).

[7] B. Widom, J. Chem. Phys. **39**, 2808 (1963).

[8] J. L. Jackson and L. S. Klein, Phys. Fluids **7**, 228 (1964).

[9] K. S. Shing and K. E. Gubbins, Mol. Phys. **46**, 1109 (1982).

[10] P. A. Monson and D. A. Kofke, Adv. Chem. Phys. **115**, 113 (2000).

[11] D. Frenkel and A. J. C. Ladd, J. Chem. Phys. **81**, 3188 (1984).

[12] A. D. Bruce, N. B. Wilding, and G. J. Ackland, Phys. Rev. Lett. **79**, 3002 (1997).

[13] A. D. Bruce, A. N. Jackson, G. J. Ackland, and N. B. Wilding, Phys. Rev. E **61**, 906 (2000).

[14] A. Z. Panagiotopoulos, Mol. Phys. **61**, 813 (1987).

[15] A. Z. Panagiotopoulos, in *Observation, Prediction and Simulation of Phase Transitions in Complex Fluids*, Vol. 460 of *NATO Advanced Studies Institute Series C: Mathematical and Physical Sciences*, edited by M. Baus (Kluwer Academic, Dordrecht, 1995).

[16] L. Tonks, Phys. Rev. **50**, 955 (1936).

[17] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Clarendon, Oxford, 1954).

[18] L. A. Girifalco, *Statistical Physics of Materials* (Wiley, New York, 1973).

[19] D. A. McQuarrie, *Statistical Mechanics* (Harper & Row, New York, 1976).

[20] C. D. Barnes and D. A. Kofke, J. Chem. Phys. **110**, 11 390 (1999).

[21] N. Lu and D. A. Kofke, J. Chem. Phys. **111**, 4414 (1999).

[22] N. Lu and D. A. Kofke, AIChE Symp. Ser. **97**, 146 (2001).