PHYSICAL REVIEW E **68**, 026122 (2003)

Variational formula for the free energy based on incomplete sampling in a molecular simulation

Nandou Lu,* Jhumpa Adhikari, and David A. Kofke†
Department of Chemical Engineering, University at Buffalo, The State University of New York, Buffalo, New York 14260-4200, USA

(Received 12 March 2003; published 21 August 2003)

Finite sampling in free-energy perturbation (FEP) calculations by molecular simulation leads to reproducible systematic errors, with sign shown to depend (in a known way) only on which system governs sampling in the simulation. Thus the result of a FEP calculation can be used as a bound on the true free energy. This inequality is of a wholly different nature from established forms such as the Gibbs-Bogoliubov inequality or the second law, in that its origins relate to the performance of a molecular simulation. If one can identify a suitable reference system having a free energy known as a function of some defining parameter, variational schemes based on the finite-sampling inequalities can be implemented. This idea is demonstrated through calculation of the free energy of a hard-sphere solid by perturbing from harmonic references and of a hard-sphere fluid by perturbing from infinitely polydisperse references. The tightness of the bounds increases with the amount of sampling in the simulation and correlates with the entropy difference between the target and reference systems. The bounds are tightest near the point where the entropy difference is least.

DOI: 10.1103/PhysRevE.68.026122 PACS number(s): 05.10.−a, 65.40.Gr, 02.70.Tt

I. INTRODUCTION

Variational methods are well established in statistical thermodynamics. They rely on rigorous expressions that establish bounds on the free energy in terms of quantities that are more tractable theoretically or computationally than the free energy itself. In a variational scheme, one adjusts a parameter (or several) to minimize an upper bound and/or maximize a lower one, and thereby obtain (one hopes) rather tight limits on the possible value of the free energy. Free-energy bounds can be useful in other ways. The relative thermodynamic stability of two systems is analyzed in terms of their free energy, and if a lower bound for system *A* can be established above an upper bound for system *B*, one can conclude that *B* is the more stable of the two, without measuring either free energy precisely.

Established approaches to the formulation of free-energy bounds proceed in either of two ways. In one, exact statistical-mechanical relations are analyzed to yield inequalities based on purely mathematical considerations. The Gibbs-Bogoliubov inequality, perhaps the best known variational formula, describes limits on the free energy *A* in terms of simple averages of the energy [1]

$$A_1 \leqslant A_0 + \langle U_1 - U_0 \rangle_0, \tag{1}$$

where the subscripts "0" and "1" represent the reference (unperturbed) and target (perturbed) states, the angular brackets indicate an ensemble average in the reference system, and $U_1 - U_0$ is the perturbation in energy from system 0 to system 1. This result is a consequence of the concavity of the exponential function [2]. As another example, Frenkel [3] recently presented this inequality:

$$A_1 \leqslant A_0 - 2kT \ln\left\langle \exp\left[-\frac{\beta}{2}(U_1 - U_0)\right]\right\rangle_0, \tag{2}$$

which was derived through an application of the Schwartz inequality.

The other common way to develop free-energy inequalities begins with the second law of thermodynamics [4,5]. A Hamiltonian system coupled to an external work source and thermal bath evolves from system 0 to system 1 via a finite-time (nonequilibrium) switching process [4]. From the second law, the free-energy change is bounded by the work *W* associated with the process:

$$A_1 \leqslant A_0 + \langle W \rangle. \tag{3}$$

The Gibbs-Bogoliubov inequality arises as a special case, when the change from 0 to 1 is instantaneous (i.e., very irreversible, with no equilibration at all), while the equality holds only for an infinitely slow (reversible) process. Free-energy bounds computed via molecular simulation using an inequality based on the second law will approach equality as additional sampling is performed, but only if this sampling is directed toward slowing down the rate of change. Then the added sampling better enables the system to remain in equilibrium with its surroundings and thus makes the process more reversible.

In an interesting development, Jarzynski [6] reestablished an equality between work and free energy for the general irreversible process:

$$\exp[-\beta(A_1 - A_0)] = \overline{\exp(-\beta W)}, \tag{4}$$

where the overbar indicates an ensemble of independent measurements—each based on an initial condition sampled from the 0 ensemble—for the irreversible process. In the same way that Eq. (3) reduces to Eq. (1) for an instantaneous change, in this situation the Jarzynski equality reduces to the well known free-energy perturbation (FEP) formula of Zwanzig [7]:

---

*Present address: Department of Physiology, School of Medicine, Johns Hopkins University, Baltimore, MD 21205, USA.

†Corresponding author. Electronic address: kofke@buffalo.edu

$$\exp[-\beta(A_1-A_0)]=\langle\exp[-\beta(U_1-U_0)]\rangle_0. \quad (5)$$

This idea has been proposed as the basis for experimental measurements of free-energy changes [8].

In the present work, an entirely different principle is employed to formulate free-energy bounds suitable for variational calculations. Although statistical mechanics provides exact formulas—such as Eqs. (4) and (5)—that prescribe how to calculate the free energy in a molecular simulation, in practice the simulation tends to sample inadequately all relevant molecular configurations, and consequently the free energy so calculated will be systematically in error. This error is highly reproducible in repeated simulations of a given length. Moreover, the nature of the error is such that one can conclude with near certainty that the actual free energy is greater (or less, depending on how the calculation is conducted) than the measured value. In what follows, it is shown in particular that these inaccurate data can provide free-energy bounds according to the following inequality:

$$(A_1-A_0)^{\mathrm{sim},1}\leqslant(A_1-A_0)^{\mathrm{exact}}\leqslant(A_1-A_0)^{\mathrm{sim},0}, \quad (6)$$

where the superscript ‘‘exact’’ indicates the exact value of free energy, and ‘‘sim, *k*’’ indicates the result of a simulation that samples (or begins with) the *k* system and gives the free-energy difference via Eq. (5) [or Eq. (4)]. Unlike for the second-law inequalities, equality is reached by performing additional sampling. However, this sampling is not directed at making the process infinitely slow and reversible; rather, it is applied to repeating the (irreversible) process an infinite number of times, sampling different initial configurations each time.

It is very well known [9,10] that FEP calculations exhibit a bias of the type described here, such that when performed in ‘‘one direction’’ they overestimate the free-energy difference, and in ‘‘the other direction’’ they underestimate it. There is also realization that this situation applies in connection with the nonequilibrium formula [6,8,11]. In recent work Lu and Kofke [10] indicated that the FEP calculation adheres to an inequality with the form given in Eq. (6). Zuckerman and Woolf [12] also observed an inequality related to the sampling size in their study of nonequilibrium free-energy calculations. In the present work we wish to emphasize the general nature of these statements. The direction of the inequality does not depend on the nature of the systems being compared, but instead depends only on which system governs sampling in the FEP calculation. A consequence of this observation is that the systematic error can be used as the basis for a variational method. Without the benefit of this analysis, one might expect that a sequence of inaccurate data taken in a FEP calculation as a function of a variational parameter could somewhere cross the correct value of the free energy. But such an outcome would be nearly useless, because there would be no indication of where the correct value lies. Instead, Eq. (6) shows that simulations of (or beginning with) the 0 and 1 systems for different values of the parameter can in fact be used to place bounds on the free energy.

The inequalities proposed here can be tightened and brought toward equality in two ways. One is by increasing the amount of sampling of the FEP average. Indications are that the entropy difference is relevant to the question of how much sampling is needed to approach equality (see below), so the necessary amount depends on the nature of the systems being studied. Thus for a fixed amount of sampling it might be possible to approach equality instead by varying features of the 0 system for a given definition of the 1 system, and this presents the second method to tighten the bounds. One can devise an effective variational scheme if a suitable reference system can be identified that has known (or readily evaluated) free energy, and which permits examination of a range of some variational parameter without requiring a separate simulation for each 0 and 1 pair.

Other variational methods and free-energy equalities when employed in molecular simulation must first overcome sampling limitations before they can yield a useful result. The practical validity of the Jarzynski equality and the Frenkel inequality, for example, is limited by the ability of the simulation to sample configurations relevant to both the reference and perturbed systems. Without supporting simulations and analysis, it can be very difficult to know if these limitations are overcome. The inequalities proposed here are immediately useful, because they are a statement of the consequences of these self-same sampling limitations. The single-step FEP inequality is particularly appealing in the variational context because, unlike most second-law methods or the general Jarzynski equality, the FEP calculation never requires sampling of any systems other than the target or reference.

In the next section we present a derivation of the inequality given by Eq. (6). Then in Sec. III we present examples of applications of the variational scheme, and in Sec. IV we discuss general features and a connection to the entropy. We conclude in Sec. V.

II. SAMPLING-BASED INEQUALITIES

The ensemble average in Eq. (4), or more specifically in Eq. (5), can be expressed as a one-dimensional integral over the work values *W* involved in transitioning from state 0 to state 1, weighted by the (normalized) distribution $p_0(W)$ of work encountered in simulations beginning from the 0 system [13–15],

$$\exp[-\beta(A_1-A_0)]=\int_{-\infty}^{\infty}dW\exp(-\beta W)p_0(W). \quad (7)$$

An analogous distribution $p_1(W)$ is defined for the perturbation conducted in the reverse direction, beginning from state 1 and evolving back to state 0:

$$\exp[+\beta(A_1-A_0)]=\int_{-\infty}^{\infty}dW\exp(+\beta W)p_1(W). \quad (8)$$

We retain the sense that the work is from 0 to 1, so must change the sign in this formula; so in the case of a FEP [Eq. (5)], W is always defined as $U_1 - U_0$. The distributions are related by [13–15]

$$p_0(W)\exp(-\beta W) = p_1(W)\exp[-\beta(A_1 - A_0)]. \tag{9}$$

When beginning from the 0 system, the exponential in Eq. (7) greatly weights the contributions to small values of W, perhaps even into the region where the sampling density p_0 is almost negligible. Accordingly, systematic error in the free energy arises largely from failure to sample W in this range. An effective way to model this inaccuracy is to assume that W is sampled perfectly above a minimum value W_0, and that it is not sampled at all below this value [10]. In this view, the (inaccurate) free-energy difference that will be measured in a simulation is

$$\exp[-\beta(A_1 - A_0)]^{\mathrm{sim},0} = \frac{\int_{W_0}^{\infty} dW \exp(-\beta W) p_0(W)}{\int_{W_0}^{\infty} dW p_0(W)}, \tag{10}$$

where the superscript "sim,0" indicates the result of a simulation that samples the 0 system. The denominator on the right-hand side will be nearly unity and can safely be ignored (barring sampling problems of a different nature from those considered here, in which a class of configurations important to the sampled distribution is not represented in the simulation average). If Eq. (10) is subtracted from Eq. (7) and Eq. (9) is applied, a simple expression for the systematic error results [10]:

$$\frac{\exp[-\beta(A_1 - A_0)]^{\mathrm{exact}} - \exp[-\beta(A_1 - A_0)]^{\mathrm{sim},0}}{\exp[-\beta(A_1 - A_0)]^{\mathrm{exact}}} = \int_{-\infty}^{W_0} dW\, p_1(W). \tag{11}$$

That is, the fractional error in $\exp[-\beta(A_1 - A_0)]$ measured in a simulation of the 0 system is given by the area under the complementary (p_1) distribution that lies below the minimum sampled energy difference.

The right-hand side of Eq. (11) is non-negative, and it follows that

$$A_1^{\mathrm{exact}} \leqslant A_0^{\mathrm{exact}} + (A_1 - A_0)^{\mathrm{sim},0}. \tag{12}$$

Thus, measurement of the free-energy difference by a simulation beginning from the 0 system using Eq. (4) or (5), when added to the exact free energy of the 0 system, provides a value that will be greater than or equal to the exact free energy of the 1 system. The simulation data provide an upper bound on the free energy of the 1 system. A complementary result can be derived following a similar path from Eq. (8) [or by simply switching the indices in Eq. (12)]:

$$A_1^{\mathrm{exact}} \geqslant A_0^{\mathrm{exact}} + (A_1 - A_0)^{\mathrm{sim},1}. \tag{13}$$

Simulation beginning from the 1 system provides a lower bound on this free energy.

For a given sample of $(U_1 - U_0)$, the mathematical inequality [16]

$$\exp[-\beta(A_1 - A_0)]^{\mathrm{sim}} = \langle \exp[-\beta(U_1 - U_0)] \rangle^{\mathrm{sim}} \geqslant \exp\langle -\beta(U_1 - U_0)^{\mathrm{sim}} \rangle \tag{14}$$

leads to the conclusion

$$(A_1 - A_0)^{\mathrm{sim}} \leqslant (U_1 - U_0)^{\mathrm{sim}}. \tag{15}$$

In other words, a given set of data taken for a FEP calculation via Eq. (5) will always yield a tighter bound than if the same data were used in the Gibbs-Bogoliubov inequality. A similar analysis will show that the Frenkel inequality will, for the same set of perturbation data, always yield a looser bound on the free energy than that given by the sampling-based bound.

It must be emphasized that the inequalities considered here are not inviolable, in the way that, e.g., the Gibbs-Bogoliubov inequality is when fully ensemble averaged. Instead, they are a statement of the likely outcome of a measurement. It is possible that in a given molecular simulation events will conspire to yield a result that violates the inequality. This outcome becomes increasingly probable as the FEP calculation approaches the correct value, i.e., as the equality limit is reached and the noise in the calculation becomes comparable to the free-energy difference. In this sense these inequalities have features in common with the second law, which can be irreproducibly violated, particularly in small systems; they are both more statistical than purely mathematical in nature [17,18]. For example, consider a calculation of the free energy of a system of 100 hard spheres (HS's) via Eq. (5) using an ideal-gas (IG) reference. In simulating the ideal gas, zero is contributed to the FEP average if the perturbation to the hard spheres finds no overlapping spheres; otherwise the contribution is unity. At high density the free-energy difference is large, $A_{\mathrm{HS}}^{\mathrm{exact}} \gg A_{\mathrm{IG}}^{\mathrm{exact}}$, and it will take very much sampling before a no-overlap perturbation is encountered. Until then, the FEP calculation gives the HS free energy as infinity [i.e., $-\ln(0)$], in accord with Eq. (12). However, it is possible (but improbable) that a no-overlap configuration is encountered early in the FEP calculation, say in the very first configuration sampled. At this point, before any further sampling is performed, the calculation indicates $A_{\mathrm{HS}} = A_{\mathrm{IG}}$, which means $A_{\mathrm{HS}}^{\mathrm{sim}} < A_{\mathrm{HS}}^{\mathrm{exact}}$ in violation of Eq. (12). The point of the inequality presented here is that, although such an outcome may be possible, it is extremely unlikely to be observed.

III. EXAMPLES

Two examples involving systems of hard spheres will demonstrate. We consider only applications involving (instantaneous) FEP measurements [Eq. (5)], rather than the more general nonequilibrium method offered by Jarzynski. We think this variational approach is more useful in the context of simple FEP, as it then involves only two systems and parameter variation is more easily implemented.

A. HS solid and harmonic references

Hard spheres of diameter σ are known to form a stable crystalline phase for densities $\rho\sigma^3$ greater than 1.04 [19,20]. The free energy of this phase has been measured very precisely, so it presents a suitable system to test and demonstrate the variational formulas. An appropriate reference is found in a system of coupled harmonic oscillators arranged on a fcc lattice. For the HS crystal, one can perform simulations at a given density to measure the covariance matrix $C_{ij}=\langle(\mathbf{r}_i-\mathbf{r}_{i,0})(\mathbf{r}_j-\mathbf{r}_{j,0})\rangle_{\mathrm{HS}}$ describing the correlation between the excursions of the spheres from their lattice sites $\mathbf{r}_{i,0}$. Then a suitable "base" harmonic system is defined such that it has the same covariance matrix [21]. A variational parameter can be identified as a scaling factor s that uniformly tightens or loosens all harmonic interactions with respect to the base system: $U_{\mathrm{harm}}=sU_{\mathrm{harm}}^{\mathrm{base}}$.

Variational calculations of this type were performed for systems of 32 fcc hard spheres and a corresponding harmonic reference. Three densities $\rho\sigma^3=1.04$ (melting), 1.2, and 1.3, respectively, were examined. When sampling the HS system, each configuration can be perturbed into a range of harmonic reference systems, and because the HS energy is always zero, the appropriate ensemble average is simply $\langle\exp(-\beta sU_{\mathrm{harm}}^{\mathrm{base}})\rangle_{\mathrm{HS}}$. Study of the reverse perturbation, sampling the harmonic system, requires a separate simulation for each value of s, but the simulations do not require a Markov chain for Monte Carlo (MC) sampling. Instead, each configuration of the harmonic system can be generated *de novo* by sampling the harmonic normal-mode coordinates from a Gaussian of width given by the corresponding normal-mode eigenvalue scaled by s. Each configuration so generated is then examined for HS overlaps, and if no overlap is found the quantity $\exp(+\beta sU_{\mathrm{harm}}^{\mathrm{base}})$ is added to the ensemble average (otherwise zero is added). The HS-based simulations sampled 5×10^6 MC cycles (1 cycle$=$32 MC trial sphere displacements, with one contribution to the FEP average for each s performed at the end of the cycle), while for the harmonic-based simulations 6×10^6 independent configurations were used (each contributing to the FEP average). Scale factors s were considered on a logarithmic scale from 10^{-3} to 10^6.

The results are presented in Fig. 1. The data are presented as the simulation-measured free energy minus the correct value for each density [20,22]. The curves above the zero line are data from FEP calculations sampling the harmonic reference (0) and thus are a test of Eq. (12), while those found below the zero line sample the target system (1) of hard spheres and test Eq. (13). Adherence to the variational formulas is clear: sampling the 0 system gives free-energy estimates that bound the true free energy from above, while sampling the 1 system gives a lower free-energy bound. The extrema in the measurements are prominent, showing how the tightness of the bounds is affected by the choice of the reference system. The accuracy of the measurements does not depend much on density, and all series give their most accurate results at about the same value of the scale factor, approximately $s=100$. In considering this fact it should be remembered that the base harmonic reference system is different for each density. Nevertheless, it is interesting to find that the optimal accuracy is obtained not when the covariances of the reference and target match (the base system, $s=1$), but when a significantly tighter harmonic system is used.

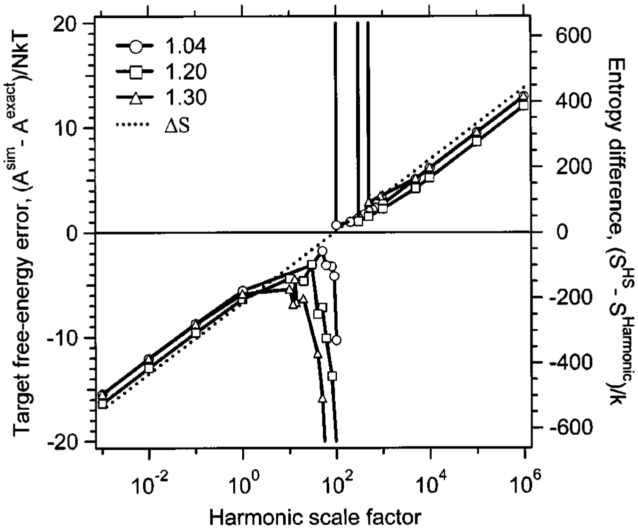

FIG. 1. Error in free energy (per particle) of pure fcc hard-sphere crystal at several densities, given as a function of the scale factor of the harmonic reference system. Values plotted are free energy as measured by molecular simulation, minus the correct value as given by the Hall equation of state [22] and the data of Frenkel and Ladd [20]. Results above the zero line are data in which the FEP simulations sampled configurations of the harmonic reference and perturbed into the hard-sphere system, while results falling below the zero line were taken from simulations sampling the hard-sphere crystal and perturbing into the harmonic reference. The dotted line is the difference in the (constant) entropy of the hard-sphere target system and the (s-dependent) entropy of the reference (extensive entropies, not per particle), for the system of density 1.3.

B. HS fluid and infinitely polydisperse references

The second example considers the hard-sphere *fluid*. A suitable variational reference for this system can be found in a system of infinitely polydisperse (IP) hard spheres [23]. An IP mixture is formulated in a semigrand ensemble, where the distribution of species chemical potential differences $\mu(\sigma)-\mu(\sigma_o)$ is an independent quantity (here, species are designated by their hard-sphere diameter, and σ_o is an arbitrarily selected reference species). In particular, in an IP mixture

$$\mu(\sigma)-\mu(\sigma_o)=c_0\ln(\sigma/\sigma_o) \tag{16}$$

where $c_0>-1$ is a dimensionless parameter that can be treated as a state variable. This mixture is interesting because it has no intrinsic molecular length scale. Consequently, all properties have a trivial scaling dependence on the density (akin to the way hard spheres depend on temperature). The pressure P obeys [23]

$$Z\equiv\beta P/\rho=(4+c_0)/3, \tag{17}$$

and the semigrand free energy Y is of the form

$$\beta Y = \ln(\rho\Lambda^3) + (c_0+1)\ln(\rho^{1/3}\sigma_o) + \ln K(c_0), \tag{18}$$

where $K(c_0)$ is independent of density. It is worth emphasizing that these results are not limited to hard-sphere systems, but are true for any spherically symmetric potential polydisperse in the sphere size.

The density-scaling properties of this system make it appealing as a reference for the fluid phase, including mixtures. A simulation performed at one density can be used to perturb into a target system (a pure fluid or a conventional mixture) at the same or any other density. The match between the IP reference and the target can be adjusted with the parameter c_0. Small values of this parameter cause the IP mixture to be populated by small particles, thus giving a low-density system in the sense of the total volume occupied (low packing fraction). As c_0 is increased (at fixed number density), the distribution of diameters moves to larger spheres, and the mixture becomes suitable for the description of high-volume-fraction phases.

Simulations were performed of systems of 32 IP hard spheres with values of c_0 ranging from 0 to 40. The system of interest, pure hard spheres, was represented by a system of hard spheres polydisperse in ‘‘fictitious’’ diameters that are completely decoupled from the interparticle potential. This approach permits perturbing between two semigrand-ensemble systems without dealing with the singularity that arises in relating monodisperse and polydisperse fluids [24]. Decoupled from the potential, the fictitious diameters have nothing constraining their size, so it is necessary to adopt a different chemical potential distribution for them. It is helpful to select a distribution that leads to a diameter distribution similar to the one adopted by the IP hard spheres. The following choice is suitable [25–27]:

$$[\beta\mu(\sigma)-\beta\mu(\sigma_o)]_1 = c_0\ln(\sigma/\sigma_o) - \frac{5}{4}\left[\frac{4+c_0}{3}\rho\right]^{1/3}\sigma - \frac{5}{4}\left[\frac{4+c_0}{3}\rho\right]^{2/3}\sigma^2 - \frac{\pi}{6}\,\frac{4+c_0}{3}\rho\sigma^3. \tag{19}$$

The real quantity of interest, the Helmholtz free energy A_{HS} of the pure hard spheres, is simply related to the simulation-measured semigrand free energy Y_1 of the hard-sphere/fictitious-diameter system:

$$\beta A_{\mathrm{HS}} = \beta Y_1 + N\ln\left[\int_0^\infty \exp[\beta\mu(\sigma)-\beta\mu(\sigma_o)]_1 d\sigma\right]. \tag{20}$$

The IP mixture is simulated at any convenient density. Contributions to the FEP ensemble average, Eq. (5) (written with free energy Y) are made as follows. The IP system is scaled to the density of interest, which entails rescaling all particle positions and diameters and thus causes no change in the energy (no overlaps result). Then for the FEP trial, all diameters are converted to the size of the pure hard-sphere target, and if no overlap then results, the quantity $\exp\{[\beta\mu(\sigma)-\beta\mu(\sigma_o)]_1 - c_0\ln(\sigma/\sigma_o)\}$ is added to the FEP ensemble average (if overlap results, zero is added). In this manner averages are easily taken for multiple densities of the target system. Parallel tempering [28,29] was used to enhance sampling of IP systems at different values of c_0—40 systems were simulated at once, with values of c_0 selected such that ‘‘adjacent’’ systems were in constant ratios of (c_0+1) [30]. Results were tabulated for perturbation into five pure-fluid densities $\rho\sigma^3$: 0.1, 0.3, 0.5, 0.7, 0.9. Approximately 10^7 elementary Monte Carlo trials were performed for each IP system, and about 300 000 FEP samples were taken in each.

Complementary simulations of 32 pure hard spheres were performed to demonstrate the inequality Eq. (13). Each pure-fluid density was simulated independently. Perturbation into the full range of IP references (c_0 varying from 0 to 40) was performed simultaneously by selecting the ‘‘fictitious’’ diameters from the distribution $\exp\{[\beta\mu(\sigma)-\beta\mu(\sigma_o)]_1\}$ using a rejection method [31] (choosing a different realization of the distribution for each c_0). With the diameters thus assigned, the system was examined for overlaps, reflecting the realization of the diameters in the IP system. If no overlap was detected, the quantity $\exp\{-[\beta\mu(\sigma)-\beta\mu(\sigma_o)]_1 + c_0\ln(\sigma/\sigma_o)\}$ was added to the ensemble average (zero added otherwise). About 10^8 elementary trials were performed at each density and about 3×10^6 FEP measurements. The IP-reference-system free energy as a function of c_0 can be evaluated from a previously determined value by integrating the log-moment of the distribution of diameters [23].

The results are presented in Fig. 2. All free-energy measurements taken by FEP from the IP reference (the 0 system) lie above the correct value (here given by the Carnahan-Starling equation [32]) at each density [in accord with Eq. (12)], while those measured in simulations of the pure hard-sphere fluid (the 1 system) lie below [in agreement with Eq. (13)]. The IP reference at low density (0.1) is best suited to describe its pure-fluid counterpart: it gives the tightest bound and provides a (noninfinite) result over the entire range of c_0, with the best result (the minimum free energy) reached at small c_0. At high density nonoverlap perturbations are rarely encountered, and noninfinite results arise only at the largest values of c_0.

Opposing this, the FEP calculations from sampling the pure fluid produce the broadest range of non-(negative)-infinite results for the highest-density system. But the low-density calculations again provide the tightest bounds. In fact, none of these FEP averages will converge to the exact result, regardless of the amount of sampling performed. The IP system always has some particles with diameters smaller than the pure fluid and thus separated by distances smaller than can be possible in the pure fluid of hard spheres. The inability to sample these configurations contributes to the systematic error in the FEP calculation. In accord with the theme of this work, the inadequate sampling manifests itself as a free-energy estimate lying below the true value.

IV. DISCUSSION

In both examples (fluid and solid), the upper and lower bounds given by the inaccurate averages are asymmetric—

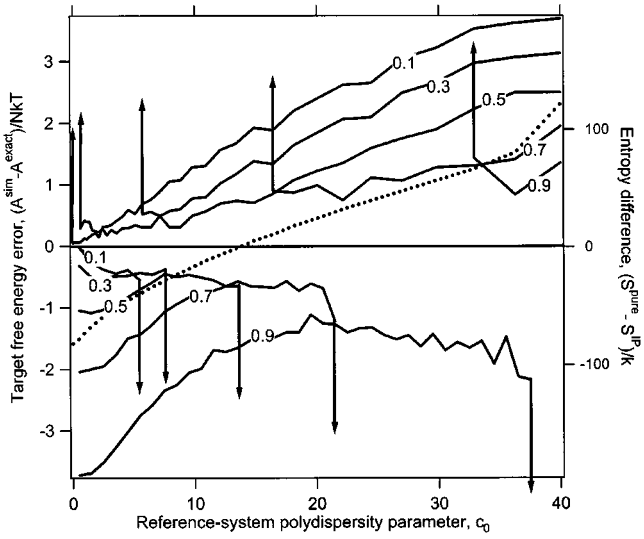

FIG. 2. Error in free energy (per particle) of pure hard spheres at several densities, given as a function of the reference-system polydispersity parameter c_0. Values plotted are free energy as measured by molecular simulation, minus the correct value as given by the Carnahan-Starling equation of state. Results above the zero line are data in which the FEP simulations sampled configurations of an IP mixture and perturbed into pure hard spheres, while results falling below the zero line were taken from simulations sampling pure hard spheres and perturbing into IP references. Numbers on each line indicate the density $\rho\sigma^3$ of the target system of pure hard spheres. Arrows indicate values going to ± infinity. The dotted line is the (extensive) entropy difference (as in Fig. 1) for systems of density 0.7.

they do not mirror each other. In other words, for a given value of the variational parameter the errors from the two calculations are not equal in magnitude, and thus it is not a good practice to average the two results to estimate the free-energy difference [33]. Recently, it has been suggested that the magnitude of the error is related to the sign of the entropy difference between the systems [10]. To examine this point, the difference in entropy between the target and reference systems for one density is included in Figs. 1 and 2. It is seen that the averages in both directions are most accurate in the range where ΔS crosses zero. The FEP averages can converge for finite sampling only when perturbation is performed as a "generalized insertion"—from the high-entropy system to the low-entropy one [10,34]. Thus when the variational parameter is to the left of the zero of ΔS, perturbation must go from the reference to the hard-sphere target; when to the right of zero, one should sample the hard-sphere system and perturb into the reference. However, the necessary amount of sampling for an accurate result scales as $\exp(-|\Delta S|/k)$ [34]. The perturbation in both examples involves the whole system, so the entropy difference is extensive and quite large even for the small systems used here. So for these FEP calculations it is difficult to generate sufficient sampling to approach the equality in Eq. (6).

V. CONCLUSION

This work derives and demonstrates a free-energy bound that originates in the way inadequate sampling in FEP calculations produces systematic error in the calculated free energy. The direction of the bound (i.e., upper or lower) is determined only by the direction of perturbation (i.e., the choice of which system governs sampling in the simulation), and not by the nature of the systems. In addition, the bound can be tightened by (1) increasing the sample size and/or (2) adjusting the variational parameter. Relatively small systems were used in the demonstration of variational schemes based on the principle, but one can expect that the proposed inequalities will apply *a fortiori* for larger systems, where sampling inadequacies are exacerbated. The examples here focus on the hard-sphere potential, but there is nothing about the inequality that limits it to this system. Solid-phase free energies for other models can be bounded using direct application of the ideas presented in the example here, using the same type of harmonic reference as employed for hard spheres. Extension of the fluid-phase example to other models is less straightforward, because the IP reference would be different and thus requires some effort to characterize it before applying it in a variational scheme. One might also consider other types of reference-target pairs. It would probably be possible to obtain very tight bounds if this methodology were applied to describe an intensive (size-independent) free-energy change, such as that associated with the insertion (or deletion) of a molecule in a system. For the particle-insertion free-energy change, the target and reference are fixed (defined as systems differing only in the presence of one molecule), and thus they provide no variational degree of freedom. Instead, a staging scheme can be applied [35,36], with an intermediate (designated *W*) that has the needed variational flexibility. Then bounds on the free-energy difference between the 0 and 1 systems can be derived from free-energy bounds applying to each stage, leading to

$$\begin{aligned}(A_1-A_W)^{\mathrm{sim},1}-(A_0-A_W)^{\mathrm{sim},W}\\ \leqslant(A_1-A_0)^{\mathrm{exact}}\leqslant(A_1-A_W)^{\mathrm{sim},W}-(A_0-A_W)^{\mathrm{sim},0}.\end{aligned} \tag{21}$$

The *W* system might, for example, interact as a real particle does, but with a diameter that is scaled as a variational parameter. The inequalities in Eq. (21) apply for any particular value of the parameter, which can be adjusted to tighten each inequality separately. Moreover, it is valid for any staged free-energy calculation, not just for this particle-insertion example.

Another place where a result of this type could find application is in the evaluation of free energies for systems modeled with computationally expensive potentials, such as encountered in the Car-Parinello method [37] or other *ab initio* techniques that are likely to come into wider use in the future. Simulations of these systems usually work with small numbers of molecules, and the expense involved in multistage FEP calculations, thermodynamic integration, or nonequilibrium paths make it especially difficult to measure free energies in such systems. This application might be particularly important for solid-phase systems, where particle-insertion methods cannot be applied.

ACKNOWLEDGMENTS

This work is supported by 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). Computational resources were provided by the University at Buffalo Center for Computational Research. We are grateful to Daan Frenkel for providing helpful comments.

[1] J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic, London, 1986).

[2] R. B. Griffiths, J. Math. Phys. **5**, 1215 (1964).

[3] D. Frenkel, in *Proceedings of the Como School on Monte Carlo and Molecular Dynamics Simulation*, edited by K. Binder and G. Ciccotti (SIF, Bologna, 1996), p. 4.

[4] K. Denbigh, *Principles of Chemical Equilibrium* (Cambridge University Press, Cambridge, England, 1971).

[5] Of course, the second law can be derived mathematically, so one might argue that there should be no distinction between these approaches and the ‘‘purely mathematical’’ ones.

[6] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).

[7] R. W. Zwanzig, J. Chem. Phys. **22**, 1420 (1954).

[8] G. Hummer and A. Szabo, Proc. Natl. Acad. Sci. U.S.A. **98**, 3658 (2001).

[9] A. E. Mark, in *Encyclopedia of Computational Chemistry*, edited by P. v. R. Schleyer (Wiley, New York, 1998), Vol. 2, p. 1070.

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

[11] D. A. Hendrix and C. Jarzynski, J. Chem. Phys. **114**, 5974 (2001).

[12] D. M. Zuckerman and T. B. Woolf, Phys. Rev. Lett. **89**, 180602 (2002).

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

[14] G. Crooks, J. Stat. Phys. **90**, 1481 (1998).

[15] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic, San Diego, 2002).

[16] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, Oxford, 1987).

[17] W. G. Hoover, *Time Reversibility, Computer Simulation, and Chaos* (World Scientific, Singapore, 1999).

[18] J. E. Hunter III, W. P. Reinhardt, and T. F. Davis, J. Chem. Phys. **99**, 6856 (1993).

[19] W. G. Hoover and F. H. Ree, J. Chem. Phys. **49**, 3609 (1968).

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

[21] J. R. Morris and K. M. Ho, Phys. Rev. Lett. **74**, 940 (1995).

[22] K. R. Hall, J. Chem. Phys. **57**, 2252 (1972).

[23] D. A. Kofke and E. D. Glandt, J. Chem. Phys. **90**, 439 (1989).

[24] J. J. Salacuse and G. Stell, J. Chem. Phys. **77**, 3714 (1982).

[25] D. A. Kofke, Ph.D. thesis, University of Pennsylvania, 1989.

[26] D. A. Kofke and E. D. Glandt, J. Chem. Phys. **92**, 658 (1990).

[27] D. A. Kofke and E. D. Glandt, J. Chem. Phys. **92**, 4417 (1990).

[28] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **57**, 2607 (1986).

[29] U. H. E. Hansmann, Chem. Phys. Lett. **281**, 140 (1997).

[30] D. A. Kofke, J. Chem. Phys. **117**, 6911 (2002).

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

[32] N. F. Carnahan and K. E. Starling, J. Chem. Phys. **51**, 635 (1969).

[33] N. Lu, J. K. Singh, and D. A. Kofke, J. Chem. Phys. **118**, 2977 (2003).

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

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

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

[37] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).