Absolute entropy and free energy of fluids using the hypothetical scanning method. II. Transition probabilities from canonical Monte Carlo simulations of partial systems

Ronald P. White and Hagai Meirovitch[a)]
Center for Computational Biology and Bioinformatics and Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, W1058 BST, Pittsburgh, Pennsylvania 15261

(Received 1 August 2003; accepted 22 September 2003)

A variant of the hypothetical scanning (HS) method for calculating the absolute entropy and free energy of fluids is developed, as applied to systems of Lennard-Jones atoms (liquid argon). As in the preceding paper (Paper I), a probability P_i approximating the Boltzmann probability of system configuration i, is calculated with a reconstruction procedure based on adding the atoms gradually to an initially empty volume, where they are placed in their positions at i; in this process the volume is divided into cubic cells, which are visited layer-by-layer, line-by-line. At each step a transition probability (TP) is calculated and the product of all the TPs leads to P_i. At step k, $k-1$ cells have already been treated, where among them N_k are occupied by an atom. A canonical metropolis Monte Carlo (MC) simulation is carried out over a portion of the still unvisited (future) volume thus providing an approximate representation of the $N-N_k$ as yet untreated (future) atoms. The TP of target cell k is determined from the number of visits of future atoms to this cell during the simulation. This MC version of HS, called HSMC, is based on a relatively small number of *efficiency* parameters; their number does not grow and their values are not changed as the number of the treated future atoms is increased (i.e., as the approximation improves); therefore, implementing HSMC for a relatively large number of future atoms (up to 40 in this study) is straightforward. Indeed, excellent results have been obtained for the free energy and the entropy. [DOI: 10.1063/1.1625920]

I. INTRODUCTION

In the preceding paper,[1] called here Paper I, the hypothetical scanning (HS) (Refs. 2 and 3) method was applied to a system of Lennard-Jones particles (argon atoms) at different densities and temperatures, and very good results have been obtained for the absolute entropy, S, and the absolute free energy, F. With this method a Monte Carlo (MC) (Ref. 4) or molecular dynamics (MD) (Refs. 5 and 6) sample is analyzed by reconstructing each of the sample's configurations step-by-step and calculating the corresponding transition probabilities (TPs). More specifically, the volume of the NVT system is divided into L^3 cells, which are visited in a linear order, row-by-row, layer-by-layer, starting from an empty volume. The TPs of the empty and occupied cells are calculated and their product leads to an approximation of the Boltzmann probability density. At step k of the reconstruction process the system consists of an already visited volume of $k-1$ cells containing N_k atoms and an unvisited (or future) volume of L^3-k+1 cells; one has to calculate the TP of the "target" cell k. Obtaining an exact TP would require calculating future partition functions based on all of the configurations of the $N-N_k$ remaining future atoms within the still unvisited future volume in the presence of the N_k already fixed atoms—an unfeasible task for a large system. Therefore, in Paper I we have calculated *approximate* future grand partition functions in a gridlike fashion based on up to two future atoms, which can occupy the target cell and its three nearest neighbor future cells. While very satisfactory results have been obtained, we have argued there that further improvement is feasible through a consideration of several more future cells populated by up to three particles.

However, one would like to be able to apply even better approximations, based on a significantly larger number of the future atoms, which might be necessary for treating more complex systems such as water. Therefore, in this paper we propose an alternative technique for calculating the TPs, where at each step the future atoms considered are simulated by the Metropolis MC method[4] within a limited future volume around the target cell, and the TP for a vacant (occupied) cell is obtained from the fraction of MC steps for which the target cell is vacant (occupied). Thus, the deterministic grand-canonical treatment of Paper I is replaced by a stochastic procedure carried out within the framework of the canonical ensemble. While the present stochastic process produces some noise, the number of future argon atoms treated can be increased dramatically (up to 40 in this study), leading to a significant improvement in the accuracy for the lower and upper bounds of F^A and the results for S^A as compared to those obtained in Paper I.

To distinguish between the two methods we name the grand-canonical (GC) HS procedure of Paper I, HSGC, and the present Monte Carlo-based HS technique, HSMC. As described in Paper I, HSGC requires optimization of a rela-

a)Author to whom correspondence should be addressed.

tively large number of chemical potential and temperature parameters. In this respect HSMC is much simpler and straightforward because the number of parameters to be optimized is minimal, and the number of atoms is constant. However, to make HSMC efficient a well suited MC process should be applied and the main part of the next section is devoted to discussing this process.

The HSMC method is based on the HS theory developed in Paper I. We assume that the reader is familiar with the philosophy of the HS method and the details of its step-by-step reconstruction procedure (in particular, the BC4 boundary conditions) described in the Introduction and Methods of Paper I. For clarity the numbers of equations of Paper I referred to in this paper will be preceded by the letter I. For comparison we apply HSMC to a system of 216 argon atoms in the two highest densities studied in Paper I.

II. THEORY AND IMPLEMENTATION

As mentioned above, with HSMC (as with HSGC in Paper I) we seek to calculate the entropy from a sample of argon configurations simulated by MC or MD in the framework of the NVT ensemble. The crux of the HSMC method is the new way for calculating the TPs for a vacant or an occupied target cell at each step of the HS reconstruction process. Thus, at step k a fixed partial structure is defined by the atoms and vacant cells reconstructed in previous steps. To define the TP of target cell k, we consider a nearby region (a portion of the total future volume that contains the target cell) and place within it a number, N_f, of movable future atoms. These future atoms, which interact with each other and with the atoms of the fixed structure, are moved within this limited region by a Metropolis MC simulation. The TP is determined from the number of visits of the future atoms to the target cell counted during the MC simulation.

A. A basic HSMC procedure

To calculate a TP we denote by M_{tot} the total number of *attempted* moves (steps) in the MC run for the kth target cell and by M_{cell} the number of MC steps in which a future atom is located in the target cell; the probability $P^+(P^-)$ for the target cell being occupied (vacant) is estimated by

$$P^+=\frac{M_{\text{cell}}}{M_{\text{tot}}}, \quad P^-=1-P^+. \tag{1}$$

Moreover, for an occupied target cell we need to calculate the TP density of the atom to be at the specific atom location exhibited in the configuration. In Paper I, the cell was divided into smaller subcells termed cubes; here, we center such a cube at the atom position, and count the number of times an atom was observed in this cube during the MC simulation. From these counts we estimate the TP density ρ^m [compare with Eq. (I.26)],

$$\rho^m=P^+\left(\frac{M_{\text{cube}}}{M_{\text{cell}}}\right)\left(\frac{1}{V_{\text{cube}}}\right)=\left(\frac{M_{\text{cube}}}{M_{\text{tot}}}\right)\left(\frac{1}{V_{\text{cube}}}\right), \tag{2}$$

where V_{cube} is the cube volume and M_{cube} is the total number of counts observed in the cube. In this equation the probability density over the cube volume is assumed to be uniform, a condition, which will be changed later.

The total product of all of the $(\text{TP})_k$, the transition probability densities ρ^m and the transition probabilities for the empty cells, provides an estimate for the probability density of the configuration. The protocol described here [as in Paper I, Eq. (I.29)] yields,

$$\prod_k (\text{TP})_k = N!\,\rho(\mathbf{x}^N,N_f,\alpha)\approx N!\,\frac{\exp[-E(\mathbf{x}^N)/k_BT]}{Z_N}, \tag{3}$$

where $\rho(\mathbf{x}^N,N_f,\alpha)$ is an approximate probability density that depends on the number of future atoms, N_f, and a set of parameters, α; $\mathbf{x}^N$ is a $3N$ vector of the Cartesian coordinates of the N atoms. [Notice that N_f, appearing in Eq. (3) above replaces the number of future cells, f in Eq. (I.29) defined for the HSGC procedure in Paper I.] The (Boltzmann) average of the log of this estimate for each configuration leads to the functional, S^A [Eq. (I.7)] that is an upper bound for the correct entropy.

B. The simulation region

At step k, N_f future atoms are simulated in the vicinity of the target cell k. More specifically, the future atoms move in a portion, V_f, of the total future volume, which includes the target cell and a nearby region. The geometry of this region is partially determined by the $k-1$ previously visited cells that are excluded. In accordance with the homogeneous boundary conditions, BC4, described in Paper I, boundary cells, vacant or containing fixed image atoms are also treated as previously visited. Thus, excluded for the future atoms are: (1) the layers ''below'' the current layer where the target cell resides (i.e., those in the $-z$ direction), (2) the rows of the current layer already visited by the reconstruction process (in the $-y$ direction), and (3) the part of the already treated current row (in the $-x$ direction). Defining the center of the target cell as the origin of the coordinates, the $+x$, $+y$, and $+z$ directions all point toward the future volume. A two-dimensional representation of the simulation region is shown in Fig. 1.

To keep the future atoms close to the target cell, they are restricted by an infinitely repulsive hemispherical shell of radius R_c measured from the center of the target cell (see Fig. 1). The volume of this region is somewhat arbitrary, but it should be close to the average volume occupied by N_f particles in the liquid. We generally choose R_c such that the density of the future atoms in the hemispherical volume is 75% of the actual liquid density.

C. The basic Monte Carlo simulation

In its simplest form, the simulation of the future atoms is carried out by the standard Metropolis MC method[4] within the restricted future volume. Thus, a trial move is generated by selecting a future atom at random and displacing it at random within a small Cartesian cube. All of the interactions between the displaced atom and the other movable future

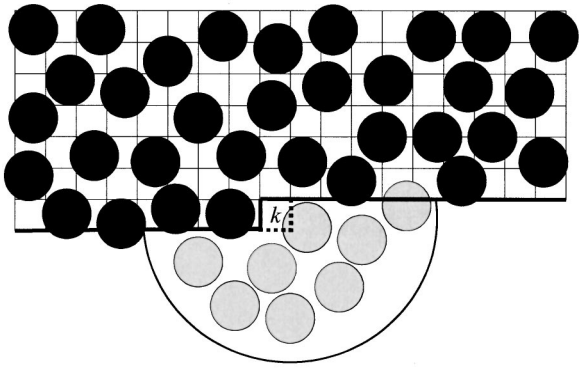

FIG. 1. A two-dimensional (2D) illustration of the kth step of the HSMC reconstruction procedure. The 2D "volume" is divided into cells, where $k-1$ of them have already been considered in previous steps; only these $k-1$ cells are shown in the figure. A cell can include the coordinates of only a single atom. Thus, these cells can be divided into two groups, the empty ones and the N_k cells occupied by previously reconstructed atoms (denoted by full black circles), which are kept fixed at their positions. (Some of these cells/atoms could be image cells/atoms defined by the BC4 boundary conditions.) The partial future 2D volume considered (V_f) is defined by a half-circle, which includes the target cell, k, depicted by dotted lines. This future volume contains the N_f future atoms (denoted by full gray circles) that are moved during the Monte Carlo simulation. These atoms can visit the target cell k and their counts in this cell lead to the transition probability of an empty cell or the transition probability density of an occupied one.

atoms and the atoms in the fixed partial structure (within the cutoff distance of 10.8 Å—half the box length; as in Paper I) are calculated before and after the trial move which is accepted or rejected according to the standard Metropolis criterion. A move that would carry a future atom outside of the future volume defined above is rejected.

The initial configuration of the future atoms is generated randomly within the hemispherical future volume, and the atoms are then subjected to an equilibration period during which the step size is periodically adjusted to achieve approximately 50% acceptance. Following the equilibration period, counts inside the target cell and inside the cube are accumulated after every attempted move. To increase efficiency several enhancements have been applied as discussed below.

D. The "near neighbor rule"

Because the number, N_f, of future atoms treated is relatively small and the future volume beyond the hemispherical region is empty, simulating these atoms can lead to an inhomogeneous environment (especially near the boundaries) typified by expanded on partially evaporated structures We would like, as much as possible, to emulate the liquidlike environment around the target cell and not have the transition probabilities strongly influenced by such surface effects. This can partially be achieved by varying the radius R_c or optimizing the temperature to minimize the entropy functional, S^A as was done in Paper I.

It should be noted however, that simply choosing a small constraining volume is undesirable. Given the same constraining value, R_c, there can still be great variability in the effective amount of (energetically accessible) volume available to the future atoms from one reconstruction step to another. For example, the neighboring previously defined atoms can at times be positioned toward the "outside edge" of their (previously defined) cells. In these cases their van der Waals radii extend well into the future volume making an appreciable fraction of this volume energetically inaccessible. At other times, the previously defined atoms can on average be backed away from the edges of their cells leaving a much larger proportion of the future volume accessible to the future atoms. A small constraining volume, that would more effectively suppress evaporation in the latter case, would be completely inappropriate for the former case due the high energy compressed environment it would create. For this reason, the constraining volume must always be chosen to give a future atom density (within the hemispherical future volume) that is somewhat less than the liquid density. Yet despite this, we would like the future atoms to behave such that the local density in the vicinity of the target cell is close to the true liquid density.

We have found that an effective remedy to this problem is to implement what we call a "near neighbor rule." The goal is to make the limited number of future atoms (with its inevitable surface) look like it was "carved out" of a real bulk liquid environment. The environment around any single atom in a bulk solid of many substances contains 12 near neighbors. The situation changes somewhat for the liquid, but it is still reasonable to imagine that an atom will have close to this number of near neighbors, especially if one increases the definition of the near neighbor distance to allow for stretched pair separations. Thus, if the future atoms were a sample "carved out" of the bulk liquid, one would expect the surface atoms of this sample to have at least a few near neighbors (in the sample). Therefore, in our MC simulation we impose that every future atom must have a minimum number, $nn_{\min}$, of near neighbors, defined by a separation of r_{nn} or less. Any MC trial move that would cause a violation of this condition is simply rejected. In practice the number of additional rejections due to this rule is small, therefore the mobility of the future atoms will not be drastically reduced.

Reasonable values for $nn_{\min}$ and r_{nn} can be easily chosen through a consideration of simple physical properties of the liquid system. A reasonable choice for r_{nn} is anywhere in the vicinity of the first minimum in the radial distribution function. Thus, for argon, we define $r_{nn}=5$ Å, and most commonly use $nn_{\min}=4$, where the choices, $nn_{\min}=3$ or 5 also work well. Note that these choices are consistent with the $\sim 5-8$ near neighbors that should be expected to be missing (for a surface atom) after the sample of future atoms is "carved out" of the bulk.

The near neighbor rule has the desirable effect of keeping the future atoms intact and lessens the importance of the hemispherical constraining shell, in that there are fewer encounters with it. This widens the range of reasonable choices for R_c. For example, very similar results are obtained when the density in the (hemispherical) future volume is varied in the range of 0.4–0.8 times the liquid density. Furthermore, because the atoms tend to stay "condensed" at all temperatures, the optimal simulation temperature is nearly identical to the real liquid temperature. In other words, one can simply

choose the liquid temperature as the HSMC simulation temperature.

E. Utilization of an ensemble averaged weighting factor

We have found that the performance of HSMC is strongly sensitive to the choice of the cube size. As has been pointed out earlier, Eq. (2) is based on a uniform probability density over the entire cube. Given the likelihood (in continuum fluids) for repulsive van der Waals overlaps upon a small change in a atom position, it should not be surprising that the assumption of a uniform probability density will eventually prevent further improvement in the results. We have found, for example, that halving the cube volume compared to the typical volume used in the HSGC calculations, results in significantly lower S^A values. Complications are encountered, however, if one tries to make this volume extremely small, because the counts accumulated in the cube become statistically unreliable. Here we describe a simple prescription which allows one to determine more accurately the probability density at a specific point, that in our case is the exact location of an atom at the center of the cube, without using extremely small cube sizes which would scarify the cube counts.

It is helpful to use a more refined notation for the transition probability density, ρ^m [Eq. (2)]; here it is written as $\rho_i(\mathbf{x}^{N_k},\mathbf{x}')$. This is the probability density for any future atom, i, to be located at the (exact) position, $\mathbf{x}'$, the exact location of the atom to be reconstructed in the target cell, where the configuration of the previously determined N_k atoms is $\mathbf{x}^{N_k}$; it should be pointed out that from now on N_k also includes the fixed image atoms at step k. As in the simplest form of the implementation described above, a cube (cube$'$) is centered around point, $\mathbf{x}'$. Counts are accumulated in cube$'$ during the MC simulation, and the probability is determined for cube$'$ to be occupied by any future atom, i. We write this probability explicitly here as $P_{i;\text{cube}'}(\mathbf{x}^{N_k},\mathbf{x}')$.

Over the course of the simulation, the following modification is introduced. Every time an atom i is recorded in cube$'$, we imagine a virtual displacement, which transports it to the exact location, $\mathbf{x}'$ (keeping all other atoms fixed). The energy, $E'(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1})$, is calculated for all of the pair interactions of the displaced atom, at $\mathbf{x}'$. That is, its interactions with the N_k defined atoms and the remaining N_f-1 future atoms. This energy is compared with, $E_i(\mathbf{x}^{N_k};\mathbf{x}_i,\mathbf{x}^{N_f-1})$, the corresponding ''undisplaced'' energy value of atom, i, as it appears in the HSMC simulation. The Boltzmann factor for this energy difference is ensemble averaged over all cases where an atom is located anywhere in cube$'$. With this ensemble average, the transition probability density, $\rho_i(\mathbf{x}^{N_k},\mathbf{x}')$, can be computed as

$$\rho_i(\mathbf{x}^{N_k},\mathbf{x}') = P_{i;\text{cube}'}(\mathbf{x}^{N_k},\mathbf{x}')\frac{1}{V_{\text{cube}}} \times\langle\exp[-(E'(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1}) - E_i(\mathbf{x}^{N_k};\mathbf{x}_i,\mathbf{x}^{N_f-1}))/k_BT]\rangle_{i;\text{cube}'}. \quad (4)$$

It is seen that the ensemble average acts as a weighting factor, which corrects the coarser estimate given for ρ^m in Eq. (2). Unsurprisingly, typical values for the ensemble average are on the order of 1. These corrections are important, however. They provide a far more accurate and efficient means to obtain the transition probability and improve the overall results significantly. A detailed derivation of Eq. (4) is given in the Appendix.

F. Preferential Monte Carlo sampling

Another way to significantly improve the HSMC results is to employ preferential sampling in the MC simulation,[7–10] a procedure that is commonly used in simulations of a solute in solution. Because solution properties are most strongly determined by the nearby environment of the solute molecule, it can be very beneficial to invest more computational resources on moving solvent molecules that are in the vicinity of the solute. A very similar situation occurs with HSMC, where the most important atoms to be moved are the closest ones to the target cell. These atoms most directly determine the cell and cube counts [M_{cell} and M_{cube}, see Eqs. (1) and (2), respectively], and sampling them more frequently indeed has provided a substantial enhancement in the reliability of the counting statistics.

A variety of forms of preferential sampling procedures are available, where their common feature is that the normal Metropolis acceptance criteria must be changed to properly account for the preferential (i.e., more frequent) trial probabilities given to the atoms close to certain locations of interest. The schemes given in Ref. 7 are relatively easy to implement and the appropriate references therein can be consulted for details. We have used the smooth preferential weighting scheme of Owicki,[8] in which the trial probability is weighted by a smooth function in the form

$$f(r) = 1/(r^{\upsilon}+c), \quad (5)$$

where υ is a constant often chosen to be an integer and c is a constant that sometimes is taken to be zero; in our case r is the distance of the atom from the center of the target cell. Using $c=0$, this distance can approach zero as an atom comes very close to the center of the target cell, thus causing a complete dominance of the weighting by this atom until it moves away. (This scenario does not arise in applications where r is a solvent–solute distance, because the van der Waals repulsion keeps r from becoming small.) For this reason, we use $c>0$. We have found that for argon $\upsilon=2$ works very well along with values for c within the range 2–4 Å^2 (where r is in Å). Alternatively, we have also used a $1/r^2$ weighting that becomes flat for values of r^2 that are less than $\sim$3 Å^2.

G. Effects of sampling convergence

The magnitude of the cells' TPs (or TP densities) can vary considerably. The TP of an empty target cell that is mostly covered by a van der Waals overlap (from a nearby atom in a previously defined cell) is relatively large (close to 1), thus contributing very little to $\ln\Pi_k(\text{TP})_k$,

$$\ln\left[\prod_k (\mathrm{TP})_k\right] = \sum_k \ln(\mathrm{TP})_k . \tag{6}$$

On the other hand, a reconstructed atom at a relatively ''uncommon'' position has a very small TP density, contributing thereby very strongly to $\ln \Pi_k(\mathrm{TP})_k$. Furthermore, $\ln(\mathrm{TP})_k$ is most strongly affected by statistical aberrations, δM_k in the counts of a small TP, where

$$\delta \ln(\mathrm{TP})_k \approx \frac{1}{(\mathrm{TP})_k}\delta(\mathrm{TP})_k = \frac{1}{(\mathrm{TP})_k}\frac{1}{M_{\mathrm{tot}}}\delta M_k . \tag{7}$$

It is therefore sensible not to allot the same number of MC steps to each target cell. Preliminary estimates of the TPs can be determined from counts accumulated during the MC equilibration period, which allows categorization of the extent of difficulty in estimating the different TPs, and development of a reasonable budgeting of their simulation lengths. Because empty cells are generally ''easier'' to handle than occupied ones, we have considered the empty and occupied cells separately and divided each into about seven difficulty categories, where the preliminary counting period is conducted only after the first 5000 equilibration steps.

Due to the stochastic nature of HSMC, increasing the amount of sampling will systematically improve the estimate of S^A [Eq. (I.7)] (whenever the sampling is not completely converged). This means that $\ln \Pi_k(\mathrm{TP})_k$ will systematically increase with increased sampling. Statistical uncertainties in the counts (if we imagine them to be roughly symmetrically distributed about the mean value) tend to lower the approximated values of $\ln(\mathrm{TP})_k$ because the second term in the expansion of $\delta \ln(\mathrm{TP})_k$ is always negative,

$$\delta \ln(\mathrm{TP})_k \approx \frac{1}{(\mathrm{TP})_k}\delta(\mathrm{TP})_k - \frac{1}{(\mathrm{TP})_k^2}\frac{1}{2}(\delta(\mathrm{TP})_k)^2 . \tag{8}$$

Thus, a negative fluctuation in the measured TP_k will lower $\delta \ln(\mathrm{TP})_k$ by a greater amount than an equivalent positive fluctuation would raise it, meaning that reduced sampling will lead to an overestimation of the upper bound, S^A.

III. RESULTS AND DISCUSSION

To compare the present results to those of Paper I, we applied HSMC to MC samples of Lennard-Jones atoms (liquid argon) in the NVT ensemble generated in Paper I. Thus, we study systems of $N=216$ atoms in the two highest (reduced) densities, $\rho^*=0.846$ and $\rho^*=0.75$, at $T=96.53$ and 137.77 K, respectively (see Table I). The results are (again) compared with thermodynamic integration (TI) values calculated in Paper I. Unless specified otherwise, the HSMC running conditions adhere to those given in Table I. The MC sample at each reconstructed step was generated in accordance with the requirement of a minimum of four near neighbor atoms discussed in the Theory section, and preferential MC sampling was employed in all cases. The TPs for occupied cells were calculated according to Eq. (4) using the ensemble averaged weighting factor. The length of an MC run depends on the estimated sampling difficulty of the corresponding target cell; therefore, we quote the *average* number of MC steps per cell, denoted, m_f.

TABLE I. The systems studied and the HSMC running conditions.[a,b]

$\rho^*=N\sigma^3/V$	T (K)[c]	Box length (Å)	Cell length (Å)	Cube length (Å)	ρ_f^*/ρ^*[d]
0.846	96.53	21.6	2.40	0.3429	0.75
0.75	137.77	22.49	2.50	0.3569	0.75

[a]In all cases, system size is 216 atoms.
[b]Unless specified otherwise, the average HSMC run length is ~350 000 MC steps per cell.
[c]The HSMC simulation temperature is the liquid temperature.
[d]The reduced density in the hemispherical future volume is $\rho_f^*=N_f\sigma^3/V_f = N_f\sigma^3/(2/3)\pi R_c^3$.

We have applied HSMC approximations based on $N_f = 5$, 10, 20, and 40 future atoms. The results for $\rho^*=0.846$ and $T=96.53$ K are presented in Tables II and III and those for $\rho^*=0.75$ in Tables IV and V. The averages of F^A and S^A given in these tables are averages over the generated samples and the reported uncertainties are taken as the standard deviations of these averages.

A. Results for $\rho^*=0.846$

Table II contains results for various approximations for the configurational Helmholz free energy, A_c [Eq. (I.21)]. As expected, the HSMC results for $A_c/\epsilon N(F^A)$ are lower than the TI value, $F(\mathrm{TI})$, which is considered to be exact they also increase systematically as the approximation is improved, i.e., in going from $N_f=5$ to 40. The free energy of the best approximation ($N_f=40$) underestimates the correct

TABLE II. HSMC results for $\rho^*=0.846$ at $T=96.53$ K.[a]

N_f	$n_{\mathrm{config.}}$	$-A_c/\epsilon N(F^A)$	σ^A	$-A_c/\epsilon N(F^B)$	$-A_c/\epsilon N(F^{M1})$
5	1871	4.2328(5)	0.0231(2)		
10	769	4.200(1)	0.0191(5)		
20	924	4.169(1)	0.0199(5)	4.125(3)	4.147(2)
40[b]	389	4.158(1)	0.0166(5)	4.105(4)	4.131(3)
''Exact'' (TI)[c]		4.120(1)		4.120(1)	4.120(1)
HSGC (4/6)[c]		4.3464(5)	0.0331(5)		

[a]N_f is the number of future atoms considered, n_{config} is the sample size, A_c is the configurational free energy [Eq. (I.21)], F^A is a lower bound of the free energy [Eq. (I.8)], and σ^A is its fluctuation [Eq. (I.14)]. F^B is an upper bound of the free energy [Eqs. (I.9) and (I.10)], and F^{M1} is the average of F^A and F^B [Eq. (I.13)]; the statistical error appears in parentheses, $0.0331(5)=0.0331\pm0.0005$.
[b]For $N_f=5$, 10, and 20 $m_f=350\,000$; for $N_f=40$ m_f is doubled to ~700 000 MC steps per cell.
[c]Results taken from Tables III and V of Paper I. TI stands for the thermodynamic integration results.

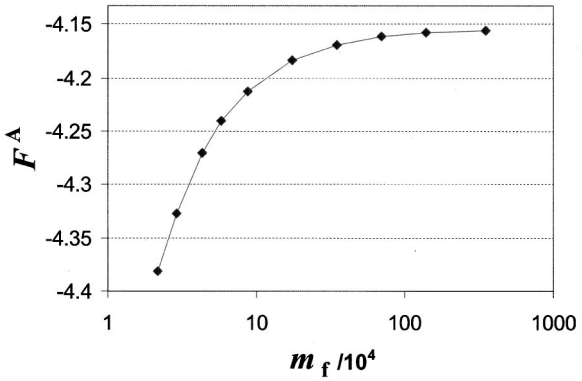

FIG. 2. The free energy functional, $A_c/\epsilon N(F^A)$ [Eq. (I.21)] vs m_f, the average number of Monte Carlo steps per cell for the approximation based on $N_f=20$ future atoms.

value by only ~0.9% as compared to an underestimation of 5.5% of the best result for $A_c/\epsilon N(F^A)$ (approximation 4/6) obtained in Paper I (see the bottom of Table II). Moreover, the latter approximation is inferior even to the worst approximation in Table II ($N_f=5$), which underestimates the correct value by ~2.7%. The fluctuation of F^A that is expected to vanish for the correct F, as expected, decreases with increasing the approximation (besides for $N_f=20$); these results are significantly smaller than the fluctuation of the 4/6 approximation of Paper I. [The "misbehavior" of the fluctuation for $N_f=20$ is a result of the insufficient sampling (350 000 MC steps on average) at each target cell, as discussed later.] All these results demonstrate that the probability density of the system configurations obtained by HSMC is significantly more accurate that that obtained by HSGC in Paper I.

For the two best approximations, $N_f=20$ and 40, we have also obtained results for $A_c/\epsilon N(F^B)$ [Eqs. (I.9) and (I.10)] that constitute upper bounds of the correct value. However, these results are based on relatively small samples and are characterized by small acceptance rate values of the reversed Schmidt procedure, $R=0.034$ and 0.024 [Eq. (I.12)]; therefore, the error bars provided for $A_c/\epsilon N(F^B)$ are probably too low. The corresponding values of $A_c/\epsilon N(F^{M1})$ [Eq. (I.13)] constitute very good approximations, where for $N_f=20$ and 40, they underestimate the correct value by 0.6% and 0.18%, respectively.

In Table III results are presented for the excess entropy, S_e [Eq. (I.19)] and the excess free energy, F_e [Eq. (I.20)], where the results for the average potential energy appear in Table VI of Paper I. Again, the results for the approximations S_e^A and F_e^A improve as N_f is increased, where for the best approximation ($N_f=40$) their deviations from the correct values are ~1%, as compared to 7.9% and 7% obtained with the best approximation (4/6) in Paper I. The corresponding deviations of the S_e^{M1} and F_e^{M1} values are, ~0.8% for $N_f=20$, and smaller than 0.3% for $N_f=40$. It should be pointed out that the entropy value for $N_f=5$ deviates from the TI value by no more than 4%, i.e., the lowest approximation captures 96% of the entropy relative to the ideal gas. This implies that the per-atom properties of the liquid are largely determined from the consideration of a region that is within one "solvation shell" about the target cell. It will be interesting to see how well a treatment of such a small number of particles can perform for other conditions or substances such as water.

TABLE III. HSMC results for the excess free energy, F_e, and the excess entropy, S_e, for $\rho^*=0.846$ at $T=96.53$ K.[a]

	$-S_e$ [cal/(K mol)]		$-F_e/T$ [cal/(K mol)]	
N_f	S^A	S^{M1}	F^A	F^{M1}
5	6.642 (3)		8.153 (1)	
10	6.724 (5)		8.073 (2)	
20	6.802 (6)	6.857 (4)	7.997 (2)	7.942 (4)
40[b]	6.839 (8)	6.896 (5)	7.968 (2)	7.902 (5)
"Exact" (TI)	6.923 (4)	6.923 (4)	7.875 (3)	7.875 (3)
HSGC (4/6)	6.364 (3)		8.430 (2)	

[a]S_e is defined in Eq. (I.19); F_e is defined in Eq. (I.20). S^{M1} is calculated using F^{M1} and the average potential energy. The other quantities are defined in the caption of Table II.
[b]m_f is doubled to ~700 000 MC steps per cell.

B. Results for $\rho^*=0.75$

The HSMC results for $\rho^*=0.75$ in Tables IV and V show the same behavior as those for $\rho^*=0.846$ discussed above but, as expected, the accuracy in general is better. Thus, the best approximation of $A_c/\epsilon N(F^A)$ ($N_f=40$) underestimates the correct (TI) value by ~0.6% as compared to the ~3.2% deviation obtained by the best result of Paper I, where the latter is still larger than the ~1.6% deviation of the worst HSMC result ($N_f=5$). The free energy fluctuations decrease systematically as the approximation improves and they are significantly smaller than the fluctuation obtained for the best approximation of Paper I, in accord with the significantly better HSMC results for $A_c/\epsilon N(F^A)$ discussed

TABLE IV. HSMC results for $\rho^*=0.75$ at $T=137.77$ K.[a]

N_f	$n_{\text{config.}}$	$-A_c/\epsilon N(F^A)$	σ^A	$-A_c/\epsilon N(F^B)$	$-A_c/\epsilon N(F^{M1})$	R
5	1296	3.7045(6)	0.0222(4)	3.638(3)	3.671(2)	0.033
10	1052	3.6845(6)	0.0193(5)	3.631(4)	3.658(2)	0.037
20	810	3.6759(6)	0.0184(4)	3.614(4)	3.645(2)	0.053
40[b]	292	3.670(1)	0.0151(5)	3.63(1)	3.650(5)	0.079
"Exact" (TI)		3.645(1)		3.645(1)	3.645(1)	
HSGC (4/6)		3.7662(3)	0.0317(5)	3.641(5)	3.704(3)	

[a]The various parameters are defined in the caption of Table II. R is the acceptance rate of the reversed Schmidt procedure [Eq. (I.12)].
[b]m_f is doubled to ~700 000 MC steps per cell.

TABLE V. HSMC results for the excess free energy, F_e, and the excess entropy, S_e, for $\rho^*=0.75$ at $T=137.77$ K.[a]

	$-S_e$ [cal/(K mol)]		$-F_e/T$ [cal/(K mol)]	
N_f	S^A	S^{M1}	F^A	F^{M1}
5	4.956(3)	5.010(3)	3.876(1)	3.818(3)
10	4.990(4)	5.033(4)	3.841(1)	3.795(4)
20	5.005(4)	5.056(4)	3.826(1)	3.772(4)
40[b]	5.024(7)	5.05 (1)	3.816(2)	3.78 (1)
"Exact" (TI)	5.056(3)		3.772(2)	
HSGC (4/6)	4.848(2)		3.981(1)	

[a]The various parameters are defined in the caption of Table III.
[b]m_f is doubled to ~700 000 MC steps per cell.

above. HSMC defines a more accurate probability density for $\rho^*=0.75$ than for $\rho^*=0.846$, which is also reflected in the larger values of the acceptance rate, R [Eq. (I.12)] obtained for $\rho^*=0.75$ ($0.033 \leqslant R \leqslant 0.079$) than for $\rho^*=0.846$ ($R \leqslant 0.024$), and in the fact that for $\rho^*=0.75$ the results for $A_c/\epsilon N(F^B)$ are upper bounds for all the approximations. The largest deviation of $A_c/\epsilon N(F^{M1})$ [Eq. (I.13)] from the correct result is by 0.7% ($N_f=5$), where the best result (for $N_f=40$), is equal to the TI value within the error bars. These results for $A_c/\epsilon N(F^A)$ and $A_c/\epsilon N(F^{M1})$ should be considered as excellent.

As in Table III, Table V provides results for the excess entropy, S_e [Eq. (I.19)] and the excess free energy, F_e [Eq. (I.20)], where the results for the average potential energy appear in Table VI of Paper I. For S^A and F^A the results improve as N_f is increased, where for the best approximation ($N_f=40$) S^A and F^A deviate from the correct values by 0.45% and 1.1%, respectively. The results for S^{M1} and F^{M1} for both $N_f=20$ and 40 are equal to the correct values within the error bars.

C. The effect of the MC sample size

It is of interest to check the effect on the free energy results by increasing, m_f, the average number of MC steps per target cell. Thus, we applied HSMC to the most dense system ($\rho^*=0.846$) using the approximation $N_f=20$ and ten m_f values increasing from 21 875 to 3.5×10^6. Table VI shows that as m_f is increased, $A_c/\epsilon N(F^A)$ increases as well in a systematic way from -4.382 ($m_f=21\,875$) to -4.156 (3.5×10^6) reaching a plateau for the larger values of m_f [i.e., converging, see Fig. 2 and the discussion related to Eq. (8). Also, note that any choice of m_f is always consistent with an upper bound, S^A, and lower bound F^A]. In conjunction with the behavior of $A_c/\epsilon N(F^A)$, the fluctuations are also steadily decreasing as m_f is increased. Likewise, for $m_f \geqslant 87\,500$, the results for $A_c/\epsilon N(F^B)$ become upper bounds, and the corresponding values for $A_c/\epsilon N(F^{M1})$ are better than their $A_c/\epsilon N(F^A)$ counterparts, becoming ~ -4.14 for $m_f=3.5\times10^6$ (these results are not shown in the table). This suggests that the approximations presented thus far (i.e., the results in the previous Tables II–V) might still be improved by using values of m_f for which convergence has been obtained.

TABLE VI. Sampling and convergence results for the $N_f=20$, 10, and 5 approximations of $\rho^*=0.846$ and $T=96.53$ K.[a]

m_f	$-A_c/\epsilon N(F^A)$	σ^A	$n_{\text{config.}}$
	$N_f=20$		
21 875	4.382(2)	0.055(1)	640
29 167	4.327(1)	0.0486(6)	2460
43 750	4.270(1)	0.0402(5)	1504
58 333	4.241(1)	0.0342(4)	2278
87 500	4.2129(6)	0.0290(4)	2692
175 000	4.1839(6)	0.0225(4)	1619
350 000	4.169(1)	0.0199(5)	924
700 000	4.162(1)	0.0160(9)	303
1 400 000	4.158(2)	0.0145(15)	131
3 500 000	−4.156(2)	0.013(2)	119
	$N_f=10$		
350 000	4.200(1)	0.0191(5)	769
700 000	4.197(1)	0.0181(10)	336
1 400 000	4.193(1)	0.0167(15)	174
3 500 000	4.192(2)	0.013(3)	69
	$N_f=5$		
350 000	4.2328(5)	0.0231(2)	1871
700 000	4.232(1)	0.0224(5)	507
1 400 000	4.231(2)	0.0236(7)	260
3 500 000	4.229(3)	0.024(1)	102

[a]The various parameters are defined in Table II.

To check this point further we have also reconstructed the sample of $\rho^*=0.846$ with $N_f=10$ and 5 future atoms, increasing m_f from 350 000 (Table II) to 700 000, 1.4×10^6, and 3.5×10^6. For $N_f=10$, $A_c/\epsilon N(F^A)$ increased by 0.008, from -4.200 to the converged value -4.192, as compared to the larger increase of 0.013, from -4.169 to -4.156 observed for $N_f=20$. On the other hand, for $N_f=5$ the slight increase in the values of $A_c/\epsilon N(F^A)$ is within the error bars., meaning that the result for $m_f=350\,000$ (of Table II) is already converged. As discussed later, for $N_f=40$ increasing m_f to 3.5×10^6 and treating a small sample has led to a further increase of $A_c/\epsilon N(F^A)$ from -4.158 to ~ -4.141; however, because a systematic convergence study for $N_f=40$ is very time consuming we have not pursued this line of research. These results show that as N_f is increased, the required number of MC steps per (future) atom, m_f/N_f, to reach convergence increases as well, e.g., from ~60 000 for $N_f=5$ to 175 000 for $N_f=20$. Therefore, the larger fluctuation for $N_f=20$ than for $N_f=10$ (0.0199 versus 0.0191, respectively) obtained in Table II stems from the relatively smaller sample, m_f/N_f, generated for $N_f=20$; indeed, when m_f is increased the fluctuations of both $N_f=20$ and 10 decrease, but they cannot be compared faithfully due to the large error bars.

The effect of the sample size on the convergence discussed above for $\rho^*=0.846$ is expected to apply also to the results for $\rho^*=0.75$, though to a lesser extent, because this more dilute system is easier to handle by HSMC. Indeed, the fluctuations in Table III decrease monotonically as N_f is increased and their correlation with the corresponding values for $A_c/\epsilon N(F^A)$ leads to very good approximations for the free energy, by best-fitting the parameters of Eq. (I.15). This analysis summarized in Table VII, shows that two of the

TABLE VII. Upper bound and extrapolated values for $A_c/\epsilon N$ based on the correlation between A_c and its fluctuation [Eqs. (I.14)–(I.16)].[a]

Approx.	$-A_c/\epsilon N$ (F^{up})	$-A_c/\epsilon N$ (F^{M2})	$-A_c/\epsilon N$ (F^{extp})	$-A_c/\epsilon N$ (TI)
		$\rho^*=0.75$ $N=216$		
10,20,40	3.661	3.666	3.669	
5,10,20,40	3.646*	3.658*	3.665	
5,10,40	3.638*	3.654*	3.663	
5,20,40	3.658	3.664	3.668	
Average		3.656(2)*	3.661(3)	3.645(1)

[a]F^{up} stands for the upper bound of F [Eq. (I.16)]; only two results are equal or larger than the correct value, and they are denoted by an asterisk. F^{M2} denotes the average free energy [Eq. (I.16)] between F^{up} and F^A ($N_f=40$). F^{extp} is the extrapolated free energy [Eq. (I.15)], and TI stands for the free energy obtained by thermodynamic integration. The values in the last row are averages of the results in the corresponding columns, where the average for F^{M2} is taken only over the values that appear with an asterisk. The presentation of the errors is defined in the caption of Table II.

results for $A_c/\epsilon N(F^{\mathrm{up}})$ are equal or larger than the correct value and the averages of the results for $A_c/\epsilon N(F^{M2})$ and $A_c/\epsilon N(F^{\mathrm{extp}})$, $-3.656(2)$ and $-3.661(3)$ underestimate the correct value by 0.44% and 0.3%, respectively.

IV. DISCUSSION

The results obtained with HSMC, in particular those for $A_c/\epsilon N(F^A)$, are significantly better than those obtained with HSGC in Paper I; however, this is achieved with some price—significantly larger computer time. For example, for the m_f values used in Table II, the number of structures reconstructed per day (24 h) are ~128, 107, 80, and 31 for $N_f=5$, 10, 20, and 40, respectively, using a single Athlon processor of 2.6 GHz. Thus, analyzing hundreds or thousands of configurations using HSMC is admittedly more computationally demanding than some of the established methods (such as TI). It should be noted, however, that the HS approach is quite unique in that it seeks to evaluate the correct Boltzmann probability density (denoted here as P_i^B) related to *any individual* configuration i, estimating the free energy through $F=E_i+k_BT\ln P_i^B$, which manifests the fact that the fluctuation of F is zero. While, in practice P_i^B is not known exactly, for a good enough approximation the fluctuation of F^A is small and a relatively small sample is required for evaluating F^A, the better the approximation the smaller the sample.

In this context it should be pointed out that for the best approximations ($N_f=20$ or 40) the accuracy obtained by HSMC for $\ln P_i^B$ of a single liquid configuration is striking. Here the fluctuations are small (Tables II and IV), and therefore the uncertainty in the free energy obtained from any single configuration is low as well. Thus, if just one configuration is analyzed with HSMC using $N_f=40$, the typical result for $A_c/\epsilon N(F^A)$ ($\rho^*=0.846$) will fall somewhere between -4.14 and -4.175. If one pursues the single configuration philosophy further by increasing m_f (for $N_f=40$) from 700 000 to 3.5×10^6, a calculation that only requires a few hours with an Athlon processor, will give for a single configuration a very good value between -4.13 and -4.15.

It should also be noted that for a system of 216 atoms the approximation of F^A cannot be improved indefinitely because of the difference between the periodic boundary conditions used to simulate the system and calculate its energy, E_i, and the boundary condition BC4 (see Paper I) used to reconstruct the system by HSMC and calculate the probability density. In particular, for a relatively small system reconstructed with a large m_f, the BC4 boundary conditions might be too restrictive; therefore, S^A might become a lower bound and F^A an upper bound close to the correct F. For the same reason, increasing N_f beyond a certain value (i.e., improving the approximation) might not lead to the expected decrease of the fluctuation of F^A, because the fluctuation of S^A might not approach that of the energy. Obviously, these effects will practically vanish for a large system.

V. SUMMARY AND CONCLUSIONS

In this paper we have developed HSMC—a new variant of the HS method, which constitutes an alternative to the HSGC procedure developed in Paper I. The transition probabilities of HSGC are based on future partition functions obtained by systematic numerical integrations over future local configurations. With HSMC, on the other hand, the TPs are calculated stochastically from canonical MC simulations of future atoms by counting their visits to the target cell. HSMC is based on a relatively small number of *efficiency* parameters that are not changed as the number of the future atoms treated is increased; therefore, implementing HSMC for a large number N_f of future atoms (maximum of 40 atoms in this study) is straightforward. Indeed, excellent results, even for the lower and upper bounds, F^A and S^A, respectively were obtained. While with HSGC the close environment of the target cell is treated more efficiently than with HSMC, the number of chemical potential parameters of HSGC grows strongly as the number of future cells or future atoms (or both) increase, and their optimization for each approximation is mandatory for obtaining improved results for S^A. Therefore, in Paper I we have not attempted to examine approximations beyond 4/6, which already has yielded very satisfactory results. Finally, it should be pointed out that one can envisage HS procedures that are hybrids of HSGC and HSMC, which might be more efficient than HSGC and HSMC individually; such hybrid methods will be studied in the future.

The development of the HSGC and HSMC methods as applied to argon is only the first step in their application to more complex fluids, in particular, water. Our goal is to be able to apply this approach to MC or MD trajectories of a peptide or a small protein soaked in a box of water molecules, where the contribution of the polymer chain to the total free energy is calculated by the local states method,[11,12] for example, and that of the water by HSMC. For that HSMC is being applied now to the TIP3P (Ref. 13) potential of water, where the first objective is to improve the efficiency of the MC simulations.

ACKNOWLEDGMENTS

This work was supported by NIH Grant No. R01 GM61916 and partially by NIH Grant No. R01 GM66090.

APPENDIX: DERIVATION OF THE WEIGHTING FACTOR EXPRESSION

We begin by considering (for clarity) the probability density for a single labeled atom, atom No. 1, to be located at a specific coordinate position, $\mathbf{x}'$, during a HSMC simulation for the reconstruction of an occupied (target) cell, in the presence of a fixed partial structure of N_k defined atoms. This probability density, $\rho_1(\mathbf{x}^{N_k},\mathbf{x}')$, is given by

$$\rho_1(\mathbf{x}^{N_k},\mathbf{x}')=\frac{\int_{N_f-1;V_f}\exp[-\beta E(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1})]d\mathbf{x}^{N_f-1}}{\int_{N_f;V_f}\exp[-\beta E(\mathbf{x}^{N_k};\mathbf{x}^{N_f})]d\mathbf{x}^{N_f}}, \tag{A1}$$

where $\beta=1/k_BT$, N_f is the number of future atoms, and V_f is the future volume in which they move. $\mathbf{x}^{N_k}$, $\mathbf{x}^{N_f}$, and $\mathbf{x}^{N_f-1}$ represent the coordinate sets for N_k, N_f, and N_f-1 atoms, respectively. The integral in the denominator is the overall configuration integral for the building step, where integrations are carried out over the future volume for all N_f atoms. In the numerator, integrations are carried out (over the future volume) for N_f-1 atoms with atom No. 1 fixed at the position, $\mathbf{x}'$.

We also consider the probability, $P_{1;\mathrm{cube}'}(\mathbf{x}^{N_k},\mathbf{x}')$, for atom No. 1 to be located anywhere inside a cube (cube$'$) centered around the position, $\mathbf{x}'$, which is given by

$$P_{1;\mathrm{cube}'}(\mathbf{x}^{N_k};\mathbf{x}')=\frac{\int_{1;\mathrm{cube}'}\int_{N_f-1;V_f}\exp[-\beta E(\mathbf{x}^{N_k};\mathbf{x}_1,\mathbf{x}^{N_f-1})]d\mathbf{x}^{N_f-1}d\mathbf{x}_1}{\int_{N_f;V_f}\exp[-\beta E(\mathbf{x}^{N_k};\mathbf{x}^{N_f})]d\mathbf{x}^{N_f}}. \tag{A2}$$

Here, integrations in the numerator involve all N_f atoms, however, atom No. 1 ranges only over the smaller volume inside cube$'$. We will ultimately construct an ensemble average expression by taking the ratio of Eqs. (A1) and (A2), given by

$$\frac{\rho_1(\mathbf{x}^{N_k},\mathbf{x}')}{P_{1;\mathrm{cube}'}(\mathbf{x}^{N_k},\mathbf{x}')}=\frac{\int_{N_f-1;V_f}\exp[-\beta E(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1})]d\mathbf{x}^{N_f-1}}{\int_{1;\mathrm{cube}'}\int_{N_f-1;V_f}\exp[-\beta E(\mathbf{x}^{N_k};\mathbf{x}_1,\mathbf{x}^{N_f-1})]d\mathbf{x}^{N_f-1}d\mathbf{x}_1}. \tag{A3}$$

The numerator in Eq. (A3) must be rewritten. We begin with a multiplication of this integral by one in the form

$$1=\left[\frac{1}{V_{\mathrm{cube}'}}\int_{1;\mathrm{cube}'}d\mathbf{x}_1\right]. \tag{A4}$$

Furthermore, the energy function in this integral is rewritten as

$$E(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1})=E'(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1})+E_{\mathrm{rest}}(\mathbf{x}^{N_k};\mathbf{x}^{N_f-1}), \tag{A5}$$

thus obtaining

$$\int_{N_f-1;V_f}\exp[-\beta E(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1})]d\mathbf{x}^{N_f-1}=\frac{1}{V_{\mathrm{cube}'}}\int_{1;\mathrm{cube}'}\int_{N_f-1;V_f}\exp[-\beta(E'(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1})+E_{\mathrm{rest}}(\mathbf{x}^{N_k};\mathbf{x}^{N_f-1}))]d\mathbf{x}^{N_f-1}d\mathbf{x}_1. \tag{A6}$$

$E'(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1})$ is the energy for all of the pair interactions of atom No. 1, fixed at $\mathbf{x}'$. In other words, its interactions with the N_k defined atoms and the remaining N_f-1 future atoms. $E_{\mathrm{rest}}(\mathbf{x}^{N_k};\mathbf{x}^{N_f-1})$ is the remaining energy of all pair interactions which do not involve atom No. 1.

We also use a similar energy function, $E_1(\mathbf{x}^{N_k};\mathbf{x}_1,\mathbf{x}^{N_f-1})$, which is the energy for all of the pair interactions of atom No. 1, but for cases where atom No. 1 could be located anywhere inside cube$'$. $E_1(\mathbf{x}^{N_k};\mathbf{x}_1,\mathbf{x}^{N_f-1})$ is defined by

$$E(\mathbf{x}^{N_k};\mathbf{x}_1,\mathbf{x}^{N_f-1})=E_1(\mathbf{x}^{N_k};\mathbf{x}_1,\mathbf{x}^{N_f-1})+E_{\mathrm{rest}}(\mathbf{x}^{N_k};\mathbf{x}^{N_f-1}). \tag{A7}$$

We now insert one as

$$1=\exp[+\beta E_1(\mathbf{x}^{N_k};\mathbf{x}_1,\mathbf{x}^{N_f-1})]\times\exp[-\beta E_1(\mathbf{x}^{N_k};\mathbf{x}_1,\mathbf{x}^{N_f-1})], \tag{A8}$$

into the integral in Eq. (A6), and substitute this form of the integral into the numerator of Eq. (A3). Upon regrouping of the energy terms, we gain an ensemble average expression for the Boltzmann factor of the energy difference resulting from the virtual displacement of atom No. 1 into the exact location, $\mathbf{x}'$,

$$\frac{\rho_1(\mathbf{x}^{N_k},\mathbf{x}')}{P_{1;\mathrm{cube}'}(\mathbf{x}^{N_k},\mathbf{x}')}$$

$$= \frac{\frac{1}{V_{\mathrm{cube}'}}\int_{1;\mathrm{cube}'}\int_{N_f-1;V_f}\exp[-\beta(E'(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1})-E_1(\mathbf{x}^{N_k};\mathbf{x}_1,\mathbf{x}^{N_f-1}))]\exp[-\beta E(\mathbf{x}^{N_k};\mathbf{x}_1,\mathbf{x}^{N_f-1})]d\mathbf{x}^{N_f-1}d\mathbf{x}_1}{\int_{1;\mathrm{cube}'}\int_{N_f-1;V_f}\exp[-\beta E(\mathbf{x}^{N_k};\mathbf{x}_1,\mathbf{x}^{N_f-1})]d\mathbf{x}^{N_f-1}d\mathbf{x}_1}$$

$$= \frac{1}{V_{\mathrm{cube}'}}\langle\exp[-\beta(E'(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1})-E_1(\mathbf{x}^{N_k};\mathbf{x}_1,\mathbf{x}^{N_f-1}))]\rangle_{1;\mathrm{cube}'}. \tag{A9}$$

This ensemble average is accumulated over all cases where atom No. 1 is located anywhere in cube′. It can be used together with $P_{1;\mathrm{cube}'}(\mathbf{x}^{N_k},\mathbf{x}')$, which would be estimated from the cube counts for atom No. 1, to solve for $\rho_1(\mathbf{x}^{N_k},\mathbf{x}')$.

We are actually interested in the (total) probability density, $\rho_i(\mathbf{x}^{N_k},\mathbf{x}')$, where i is any atom. We note, however, that

$$\rho_i(\mathbf{x}^{N_k},\mathbf{x}') = \rho_1(\mathbf{x}^{N_k},\mathbf{x}') + \rho_2(\mathbf{x}^{N_k},\mathbf{x}') + \rho_3(\mathbf{x}^{N_k},\mathbf{x}') + \cdots$$
$$= N_f\rho_1(\mathbf{x}^{N_k},\mathbf{x}') \tag{A10}$$

and that

$$P_{i;\mathrm{cube}'}(\mathbf{x}^{N_k},\mathbf{x}') = P_{1;\mathrm{cube}'}(\mathbf{x}^{N_k},\mathbf{x}') + P_{2;\mathrm{cube}'}(\mathbf{x}^{N_k},\mathbf{x}') + P_{3;\mathrm{cube}'}(\mathbf{x}^{N_k},\mathbf{x}') + \cdots$$
$$= N_f P_{1;\mathrm{cube}'}(\mathbf{x}^{N_k},\mathbf{x}'). \tag{A11}$$

Thus,

$$\frac{\rho_i(\mathbf{x}^{N_k},\mathbf{x}')}{P_{i;\mathrm{cube}'}(\mathbf{x}^{N_k},\mathbf{x}')} = \frac{\rho_1(\mathbf{x}^{N_k},\mathbf{x}')}{P_{1;\mathrm{cube}'}(\mathbf{x}^{N_k},\mathbf{x}')}$$
$$= \frac{1}{V_{\mathrm{cube}'}}\langle\exp[-\beta(E'(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1}) - E_1(\mathbf{x}^{N_k};\mathbf{x}_1,\mathbf{x}^{N_f-1}))]\rangle_{1;\mathrm{cube}'}. \tag{A12}$$

We also note that the value of the ensemble average in Eq. (A9) for atom No. 1 will be identical to the more general ensemble average which is taken over any case where any atom, i, is in cube′. In other words,

$$\langle\exp[-\beta(E'(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1})-E_1(\mathbf{x}^{N_k};\mathbf{x}_1,\mathbf{x}^{N_f-1}))]\rangle_{1;\mathrm{cube}'}$$
$$= \langle\exp[-\beta(E'(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1}) - E_i(\mathbf{x}^{N_k};\mathbf{x}_i,\mathbf{x}^{N_f-1}))]\rangle_{i;\mathrm{cube}'}. \tag{A13}$$

Thus the (total) probability density, $\rho_i(\mathbf{x}^{N_k},\mathbf{x}')$, is obtained from

$$\rho_i(\mathbf{x}^{N_k},\mathbf{x}') = P_{i;\mathrm{cube}'}(\mathbf{x}^{N_k},\mathbf{x}')\frac{1}{V_{\mathrm{cube}'}} \times\langle\exp[-\beta(E'(\mathbf{x}^{N_k},\mathbf{x}';\mathbf{x}^{N_f-1}) - E_i(\mathbf{x}^{N_k};\mathbf{x}_i,\mathbf{x}^{N_f-1}))]\rangle_{i;\mathrm{cube}'}, \tag{A14}$$

where $P_{i;\mathrm{cube}'}(\mathbf{x}^{N_k},\mathbf{x}')$ is obtained from the counts for any atom being located (anywhere) in cube′ and where the ensemble average is computed for all of these cases.

[1] A. Szarecka, R. P. White, and H. Meirovitch, J. Chem. Phys. **119**, 12084 (2003), preceding paper.
[2] H. Meirovitch, J. Phys. A **16**, 839 (1983).
[3] H. Meirovitch, Phys. Rev. A **32**, 3709 (1985).
[4] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[5] B. J. Alder and T. E. Wainwright, J. Chem. Phys. **31**, 459 (1959).
[6] J. A. McCammon, B. R. Gelin, and M. Karplus, Nature (London) **267**, 585 (1977).
[7] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarenden, Oxford, 1987).
[8] J. C. Owicki, in *Computer Modeling of Matter*, ACS Symposium Series, edited by P. Lykos (American Chemical Society, Washington, 1978), Vol. 86, p. 159.
[9] W. L. Jorgensen, J. Phys. Chem. **87**, 5304 (1983).
[10] J. C. Owicki and H. A. Scheraga, Chem. Phys. Lett. **47**, 600 (1977).
[11] H. Meirovitch, Chem. Phys. Lett. **45**, 389 (1977).
[12] H. Meirovitch, S. C. Koerber, J. Rivier, and A. T. Hagler, Biopolymers **34**, 815 (1994).
[13] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).