Fluctuation-based computer calculation of partial molar properties. I. Molecular dynamics simulation of constant volume fluctuations

Pablo G. Debenedetti
Department of Chemical Engineering, Princeton University, Princeton, New Jersey 08544

(Received 12 February 1987; accepted 12 March 1987)

Partial molar volumes, energies, and enthalpies are calculated via the (N,U,V) molecular dynamics implementation of their corresponding operational definitions in terms of constant volume fluctuations. The method is exact for the former property but only approximate for the latter two. The partial molar volume operational definition satisfies Euler's homogeneous function theorem identically. The corresponding energy definition is based upon a differential equation and does not, therefore, satisfy the homogeneity constraint identically. A numerically unimportant but conceptually profound correction to the recently derived partial molar energy operational definition is incorporated and its theoretical implications discussed.

INTRODUCTION

Partial molar quantities are among the most important properties in the experimental and theoretical study of nonideal mixtures. "Mechanical" partial molar quantities [i.e., those that can be defined without invoking the concept of entropy, namely, partial molar volume, energy (and hence enthalpy)] are central to our understanding of phenomena as diverse as thermal and volumetric effects upon mixing, cluster formation in supercritical mixtures,[1] and the effect of chemical equilibrium on the thermodynamic properties of reacting mixtures.[2]

Deterministic (molecular dynamics) and stochastic (Monte Carlo) computer simulations constitute extremely powerful tools in the study of the relationship between bulk properties and the details of such molecular characteristics as size, shape, electrical charge, and interaction potentials. Little work has been done in the area of computer-based study of partial molar volumes and enthalpies.

Previous researchers[3,4] have adopted a "literal" interpretation of the generic mechanical partial molar property. They have, in other words, considered the rate of change of an extensive quantity (λ) with respect to a particular component's mass (measured in moles or molecule numbers, N_i), at constant temperature, pressure, and masses of all other components (a constraint denoted here by $N[i]$),

$$\bar{\lambda}_i = \left[\frac{\partial \lambda}{\partial N_i}\right]_{T,P,N[i]} . \tag{1}$$

A literal computation of $\bar{\lambda}_i$ via a computer simulation, then, calls for the evaluation of the following quantity:

$$\Delta \equiv \lambda(N_1,...,N_i,...,N_n,T,P) - \lambda(N_1,...,N_i \pm 1,...,N_n,T,P) \simeq \mp \bar{\lambda}_i . \tag{2}$$

The partial molar property, therefore, is obtained from the difference between the values of the corresponding extensive quantity over two isobaric, isothermal simulations in which the composition of the system under study is changed by addition (or removal) of a single molecule. The main features of such an approach are:

(i) Two simulations per data point are required (this is probably the most important reason behind the above described dearth of work in the area).

(ii) The partial molar property is calculated as the (small) difference between two large quantities. This source of numerical inaccuracy is, of course, balanced by the fact that fluctuations in λ, an extensive property, decrease as the sample size increases.

(iii) The partial derivative $\partial/\partial N_i$ is replaced by a difference, computed over a finite sample. In this sense, the approach is exact when applied to the study of infinite dilution properties, if infinite dilution is simulated by imposing on a periodically bound system containing a single i-type molecule bulk conditions corresponding to an otherwise identical simulation in which such a molecule is absent (for example, pressure and temperature in an isobaric, isothermal simulation). This accuracy at infinite dilution is especially interesting in the light of the numerical limitations of the fluctuation method to be discussed below at high dilution of any particular species.

The literal approach has been successfully applied recently[5] to the study of a variety of infinitely dilute solute partial molar properties in binary mixtures. This interesting study, part of an ongoing comprehensive research effort in the area was, moreover, not limited to mechanical partial molar properties.

Although Eq. (1) is a rigorous definition of the generic partial molar quantity, it does not exhaust the possibilities of computer-based studies of such properties. Fluctuation-based methods such as the one to be discussed in this paper are among the most promising (and still unexplored) alternatives.

Any thermodynamic partial derivative can be expressed in terms of fluctuations. There results a fluctuation-based operational definition of the partial derivative: the idea is then to calculate the property in question not as a difference [i.e., through Eq. (2), for example] but via a computer simulation of the appropriate fluctuations. Among the distinctive features of this alternative approach, we cite

(i) the operational definition is not unique, since a given partial derivative can be expressed in many equivalent fluctuation-explicit ways.

(ii) Consequently, fluctuation-based methods offer a wide range of possibilities for the computer study of partial molar quantities, since the nature of the many possible oper-

0021-9606/87/127126-12$02.10

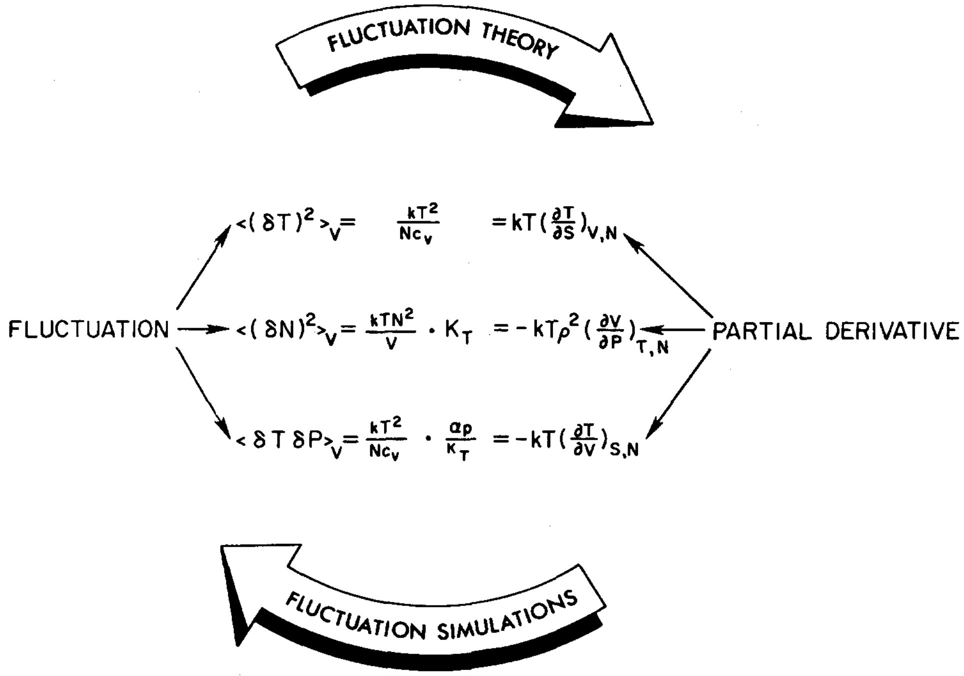

FIG. 1. Fluctuation theory allows the quantification of a fluctuation in terms of a partial derivative. Fluctuation simulations allow the calculation of partial derivatives from the simulation of the corresponding fluctuation(s). The compressibility relationship, as written, applies to a pure substance.

ational definition dictates the corresponding type of simulation.

(iii) Fluctuation-based methods require a single simulation per data point.

In this paper we consider constant volume fluctuations. Operational definitions for partial molar volumes, energies (and hence enthalpies) are presented and implemented within a molecular dynamics algorithm. The advantages and limitations of the approach are derived from an analysis of the calculations in the light of the homogeneity condition which must be obeyed by any partial molar property.

Interpreting thermodynamic quantities in terms of fluctuations, implementing the operational definitions as computer algorithms, and seeing statistical order emerge out of (simulated) molecular chaos is a permanent reminder of the great leaps of the imagination (Boltzmann, Gibbs, Einstein) that made our understanding of bulk matter in terms of its molecular constituents possible.

FLUCTUATION METHOD

Equilibrium states of thermodynamic systems can be characterized by the values of a finite number of properties. That such properties can, indeed, be associated with a time-invariant numerical value is an idealization. Quantities such as temperature and pressure fluctuate about their equilibrium values as a consequence of molecular motion. Fluctuation theory is the formalism that allows the quantification of the statistical regularities underlying this apparently chaotic behavior in terms of average (equilibrium) values of thermodynamic properties. An equilibrium state, in other words, can be characterized not only by the average value of properties, but also by the distribution of fluctuations of those same properties about their respective average values.[6]

In order to define a fluctuation, it is necessary to specify not only the fluctuating observable(s), but the extension of the region of space (subsystem) under consideration. We can, for example, focus our attention on a control volume (fixed but arbitrary volume, variable mass), or on a region of space so defined as to contain, at all times, a fixed mass of a particular chemical species. Subsystems defined by a single extensive constraint are of particular interest in the present context. The average value of any quantity ψ computed (or measured) either over an ensemble of subsystems defined by some extensive constraint X or over time, within a single, X-defined subsystem, will be denoted by $\langle \psi \rangle_X$.

The lowest-order, nontrivial moments which characterize the distribution of fluctuations in an equilibrium system are second moments. In this case, with $\delta\psi \equiv \psi - \langle \psi \rangle$, the theory[6-9] yields the completely general result ($\beta = 1/kT$; see Ref. 9 for a detailed analysis)

$$\beta \langle \delta\lambda_i \delta\lambda_j \rangle_{X_{n+2}} = \left[\frac{\partial \lambda_i}{\partial \bar{\lambda}_j} \right]_{\bar{\lambda}[j],X_{n+2}}, \tag{3}$$

where n denotes the number of chemically distinct components in the system under study, X_i ($i = 1,...,n+2$) are the independent extensive variables in terms of which the system's energy can be expressed [i.e., S, V, N_k ($k = 1,...,n$), or any of the $(n+2)!$ permutations of these variables], and $(\lambda_j,\bar{\lambda}_j)$ denotes a generic conjugate pair $[(S,T),(V,-P),(N_i,\mu_i)]$ in which λ is extensive (intensive) if $\bar{\lambda}$ is intensive (extensive). The notation $\bar{\lambda}[j]$ implies constancy of all $\bar{\lambda}$'s except for $\bar{\lambda}_j$. Except for the case $\lambda_j = N_j$, $\bar{\lambda}_j$ is not a partial molar quantity in Eq. (3).

Equation (3), then, relates fluctuation covariances (e.g., $\langle \delta T \delta P \rangle$, $\langle \delta S \delta V \rangle$) or mean-squared fluctuations [e.g., $\langle (\delta T)^2 \rangle$, $\langle (\delta N_i)^2 \rangle$] to partial derivatives. One can interpret this relationship between pairwise fluctuations and partial derivatives in two different ways,[9] one of which (i.e., the quantification of such a fluctuation in terms of a first-order partial derivative) is one of the important results of the theory of fluctuations. Conversely, one can invoke Eq. (3) in order to interpret a particular derivative property in terms of its corresponding fluctuation(s). The computer simulation of the appropriate fluctuation(s) is the central idea underlying the present work, and is illustrated in Fig. 1 for the case of three properties (specific heat, thermal expansion coefficient, isothermal compressibility) with extremely simple (and well known) fluctuation-explicit definitions (the K_T relationship, as written, is only valid for a pure substance).

The significance of this approach lies in the fact that many properties of interest in the study of nonideal mixtures are first order partial derivatives, and are thus amenable to a fluctuation interpretation. Such a fluctuation interpretation, however, is not unique, since a given partial derivative can be expressed in terms of as many types of fluctuations as there are subsystem definitions. Accordingly, a derivative property can be studied via a variety of computer simulations, some of which[10] have not been proposed before.

Out of the many possible definitions, we focus our attention here on fluctuations within a control volume, that is, an open, fixed volume region of space. The derivation of operational definitions for $\overline{V}_i$, $\overline{U}_i$ (and hence $\overline{H}_i$) in terms of such fluctuations has been discussed recently in detail.[10] For $\overline{V}_i$, the grand canonical treatment of Kirkwood and Buff[11] is invoked to obtain, for an n-component mixture,

$$\overline{V}_i \rho = \left[\sum_{j=1}^{n} x_j |B|_{ij}\right]\left[\sum_{j=1}^{n}\sum_{k=1}^{n} x_j x_k |B|_{jk}\right]^{-1}, \qquad (4)$$

where

$$\rho = \langle N \rangle / V = V^{-1} \sum_{j=1}^{n} \langle N_j \rangle , \qquad (5)$$

x_j is a mole fraction, and $|B|_{\alpha\beta}$ denotes the cofactor of the α,β element in the fluctuation determinant $|B|$, with elements

$$B_{ij} = \frac{\langle \delta N_i \delta N_j \rangle}{V}, \qquad (6)$$

where $\langle\ \rangle$ denotes thermodynamic averaging over an ensemble of open subsystems of volume V. As an example, for a binary mixture, we have

$$\overline{V}_2 \rho = \frac{\dfrac{\langle (\delta N_1)^2 \rangle}{\langle N_1 \rangle} - \dfrac{\langle \delta N_1 \delta N_2 \rangle}{\langle N_2 \rangle}}{1 + x_2\left[\dfrac{1 - x_2}{x_2}\left(\dfrac{\langle (\delta N_2)^2 \rangle}{\langle N_2 \rangle} - 1\right) + \left(\dfrac{\langle (\delta N_1)^2 \rangle}{\langle N_1 \rangle} - 1\right) - 2\,\dfrac{\langle (\delta N_1 \delta N_2 \rangle}{\langle N_2 \rangle}\right]}.$$

In the case of $\overline{U}_i$, a fluctuation-explicit operational definition follows from solving the linear system[10] ($j = 1,\ldots,n$; see Appendix for a derivation)

$$\overline{U}_j + \sum_{i \neq j} \frac{\langle \delta N_i \delta N_j \rangle}{\langle (\delta N_j)^2 \rangle}\,\overline{U}_i = \frac{\langle \delta U \delta N_j \rangle}{\langle (\delta N_j)^2 \rangle} + \frac{\beta P K'_T \langle \delta P \delta N_j \rangle}{\rho^2 \langle (\delta N_j)^2 \rangle}\left[\frac{T \langle \delta T \delta P \rangle}{P \langle (\delta T)^2 \rangle} - 1\right], \qquad (7)$$

where

$$K'_T \equiv \langle N \rangle \rho k T K_T = |B| V \Big/ \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j |B|_{ij}$$

which, for a binary mixture, becomes

$$\overline{U}_2 = \frac{[\langle \delta U \delta N_2 \rangle \langle (\delta N_1)^2 \rangle - \langle \delta U \delta N_1 \rangle \langle \delta N_1 \delta N_2 \rangle] + \phi[\langle \delta P \delta N_2 \rangle \langle (\delta N_1)^2 \rangle - \langle \delta P \delta N_1 \rangle \langle \delta N_1 \delta N_2 \rangle]}{\langle (\delta N_1)^2 \rangle \langle (\delta N_2)^2 \rangle - \langle \delta N_1 \delta N_2 \rangle^2},$$

where

$$\phi = \frac{\beta P}{\rho^2}\left[\frac{T}{P}\cdot\frac{\langle \delta T \delta P \rangle}{\langle (\delta T)^2 \rangle} - 1\right] \cdot \frac{(\langle N_1 \rangle + \langle N_2 \rangle)^2[\langle (\delta N_1)^2 \rangle \langle (\delta N_2)^2 \rangle - \langle \delta N_1 \delta N_2 \rangle^2]}{\langle N_2 \rangle^2 \langle (\delta N_1)^2 \rangle + \langle N_1 \rangle^2 \langle (\delta N_2)^2 \rangle - 2\langle N_1 \rangle \langle N_2 \rangle \langle \delta N_1 \delta N_2 \rangle}.$$

In the original derivation,[10] the relationship $K'_T = \langle (\delta N)^2 \rangle$ was incorrectly used. All of the results reported here were obtained using the correct form given above. Although numerically inconsequential (partial molar energies and enthalpies change by less than 1%) this difference has very interesting theoretical implications which we address in the Appendix.

Equation (4) is amenable to an algorithmic implementation via (N,U,V) molecular dynamics or (N,T,V) Monte Carlo simulations (U denotes internal energy); Eq. (7), being explicit in temperature fluctuations cannot, on the other hand, be computed via Monte Carlo techniques. In this paper we discuss molecular dynamics simulations exclusively.

The numerical implementation of the operational definitions within a molecular dynamics algorithm calls for the partition of the periodically bound computational cell into an appropriate number (L) of constant volume, open subcells. These subcells have no physical significance whatsoever, and are only used for "accounting" purposes (they do not, in other words, impose any constraint upon the molecules' motion). A possible subcell arrangement is illustrated in Fig. 2, where two of the eight space-filling units into which the periodically bound computational cell has been divided are shown.

At each time step during the simulation, subcell energies, temperatures, pressures, and compositions are calculated, and the appropriate averages, as per Eqs. (4)–(7), are then computed, yielding the desired partial molar property. The instantaneous configurational contributions to subcell energy and pressure (i.e., virial) were evenly divided among the subcells in which members of the particular pair under consideration were located at any given instant, an approximation consistent with the pairwise additivity assumption used throughout this work.

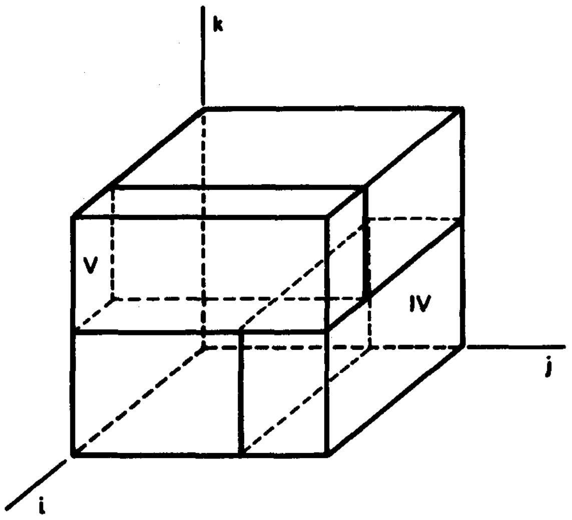

FIG. 2. Division of the periodic computational cell into eight space-filling, open subcells. Only subcells IV and V are shown.

TECHNICAL DETAILS

A binary Lennard-Jonesium was used as a model system. The following definitions apply throughout (1 = solute, 2 = solvent, m = mass, c_v = isochoric specific heat, K_T = isothermal compressibility, α_p = thermal expansion coefficient, H = enthalpy, v = velocity)

$$T^* = kT\epsilon_{22}^{-1}, \tag{8}$$

$$P^* = P\sigma_{22}^3\epsilon_{22}^{-1}, \tag{9}$$

$$V^* = V\sigma_{22}^{-3}, \tag{10}$$

$$\rho^* = N\sigma_{22}^3 V^{-1}, \tag{11}$$

$$m^* = mm_1^{-1}, \tag{12}$$

$$C_v^* = C_v k^{-1}, \tag{13}$$

$$K_T^* = \epsilon_{22}K_T\sigma_{22}^{-3}, \tag{14}$$

$$\alpha_p^* = \epsilon_{22}\alpha_p k^{-1}, \tag{15}$$

$$U^* = U\epsilon_{22}^{-1}; \quad H^* = H\epsilon_{22}^{-1}, \tag{16}$$

TABLE I. Fixed parameters.[a]

$\epsilon_{11}/\epsilon_{22}$	σ_{11}/σ_{22}	m_2^*	ρ^*
0.597	0.768	1.318	0.8594

[a] $\epsilon_{11} = \epsilon_{22}$, $\sigma_{11} = \sigma_{22}$, $m_2^* = 1$ in runs 7, 8, 10, 11, and 12 (see Table II).

$$v^* = v(m_1\epsilon_{22}^{-1})^{1/2}. \tag{17}$$

Parameters which were fixed throughout the simulations are listed in Table I. Cross interaction parameters were obtained from the Lorentz–Bertholet combining rules. A study of this highly asymmetric mixture via (N,T,V) Monte Carlo has recently been published[12]: it includes phase coexistence boundaries at or near the conditions explored in the present work.

The shifted force method[13] was used to truncate the potential "exactly" (i.e., without loss of energy conservation). The cutoff radius was set at $2.8\sigma_{22}$. All simulations followed relaxation runs in which the system was started from an fcc configuration in which solute molecules were randomly placed within the regular structure. A unimodal velocity distribution $(3T^*/m_i^*)^{1/2}$, with randomly oriented components was imposed upon each species, T^* being the "objective" temperature from which the system was started, and towards which it was brought, in the course of relaxation runs, by successive time rescaling. Relaxation runs varied in length, but were never shorter than 3.5×10^4 steps.

The equations of motion were integrated via the fourth order (Verlet) scheme[14]

$$x(t+\delta t) = 2x(t) - x(t-\delta t) + \ddot{x}(t)(\delta t)^2 + O(\delta t)^4. \tag{18}$$

In such an algorithm, the velocity is obtained from the equation

$$\dot{x}(t) = \left(\frac{1}{2\delta t}\right)[x(t+\delta t) - x(t-\delta t)] + O(\delta t)^2 \tag{19}$$

and, therefore, temperature corrections (during relaxation runs) were implemented via a rescaling of δt.

Equilibrium values of T^* and P^* and their associated fluctuations are shown in Table II for all of the simulations

TABLE II. Energy conservation, temperature, pressure, and time step for the simulations performed in this work.

Run[a]	N	N_1	$\overline{T}^*$	$\overline{P}^*$	rms(T^*)[b]	rms(P^*)[b]	10^4 rms(U^*)[b]	$10^4\,\delta t$[c]
1	256	77	0.923	0.842	0.0287	0.2473	3.391	7.82
2	256	77	0.861	0.561	0.0286	0.3416	1.711	7.54
3	256	77	0.801	0.309	0.0289	0.6096	2.384	7.22
4	256	64	0.930	1.051	0.0288	0.1929	1.806	6.77
5	256	51	0.934	1.275	0.0303	0.1610	3.054	6.35
6	256	38	0.913	1.469	0.0306	0.1402	2.951	13.25
7	256	77	0.974	2.936	0.0329	0.0735	1.394	4.52
8	256	77	0.975	2.945	0.0324	0.0724	4.039	4.52
9	500	150	0.930	0.871	0.0197	0.1602	1.907	5.96
10	256	51	0.974	2.936	0.0329	0.0735	1.394	4.52
11	256	38	0.974	2.936	0.0329	0.0735	1.394	4.52
12	256	26	0.974	2.936	0.0329	0.0735	1.394	4.52

[a] $\epsilon_{11} = \epsilon_{22}$, $\sigma_{11} = \sigma_{22}$, $m_2^* = 1$ in runs 7, 8, 10, 11, and 12.

[b] $\mathrm{rms}(\alpha) = [\,\overline{(\delta\alpha)^2}/\,\overline{(\alpha)}^2]^{1/2}$; overbars denote averaging over the whole computational cell.

[c] Time measured in units of $l(m_1/\epsilon_{22})^{1/2}$; l = size of computational cell.

TABLE III. Partial molar property calculations.

Run	N	L	N_1	$\overline{T}^*$	$\overline{V}_1^*$	$\overline{V}_2^*$	$\Sigma x_i \overline{V}_i^*$	V^*/N	$-\overline{U}_1^*$	$-\overline{U}_2^*$	$-\Sigma x_i \overline{U}_i^*$	$-U^*/N$	$-\overline{H}_1^*$	$-\overline{H}_2^*$	$-\Sigma x_i \overline{H}_i^*$	$-H^*/N$
1	256	16	77	0.923	0.8756	1.2874	1.1635	1.1636	1.4111	3.4913	2.8656	2.2966	0.6737	2.4071	1.8857	1.3167
2	256	16	77	0.861	0.9131	1.2713	1.1636	1.1636	1.4215	3.4647	2.8502	2.4338	0.9096	2.7519	2.1978	1.7815
3	256	16	77	0.801	0.8948	1.2792	1.1636	1.1636	1.5249	3.4632	2.8802	2.5678	1.2486	3.0683	2.5210	2.2085
4	256	16	64	0.930	0.9028	1.2505	1.1636	1.1636	1.7110	3.8723	3.3195	2.4770	0.7620	2.5579	2.1089	1.2540
5	256	16	51	0.934	0.8491	1.2418	1.1636	1.1636	1.7978	3.9862	3.5502	2.6654	0.7155	2.4033	2.0671	1.1821
6	256	16	38	0.913	0.8308	1.2216	1.1636	1.1636	2.1665	4.4379	4.1008	2.9063	0.9461	2.6435	2.4466	1.1971
7	256	8	77	0.974	1.1908	1.1518	1.1635	1.1636	6.6050	6.5305	6.5529	3.3219	3.1088	3.1487	3.1367	−0.0943
8	256	16	77	0.975	1.1789	1.1569	1.1635	1.1636	6.7972	6.7539	6.7669	3.3192	3.3256	3.3467	3.3103	−0.1074
9	500	16	150	0.930	0.8497	1.2981	1.1636	1.1636	1.5114	3.7815	3.1005	2.2857	0.7711	2.6505	2.0867	1.2719
10	256	8	51	0.974	1.1879	1.1575	1.1636	1.1636	6.7202	6.5172	6.5576	3.3219	3.2326	3.1188	3.1415	−0.0943
11	256	8	38	0.974	1.1790	1.1609	1.1636	1.1636	6.6468	6.5389	6.5549	3.3219	3.1853	3.1306	3.1387	−0.0943
12	256	8	26	0.974	1.1584	1.1641	1.1636	1.1636	6.6176	6.5468	6.5540	3.3219	3.2165	3.1289	3.1378	−0.0943

performed in this study (the notation $\overline{\overline{\alpha}}$ denotes averaging over a simulation, that is to say, over the entire periodically bound, constant energy computational cell, and should not be confused with $\langle\alpha\rangle$). Also shown in Table II is the algorithm's accuracy, as measured by the normalized root mean square energy fluctuation, and the magnitude of the corresponding time step δt [in units of $l(m_1/\epsilon_{22})^{1/2}$, where l is the computational cell's linear dimension]. All runs were 30 000 steps long, except for run 8, which lasted 40 000 steps. An explicitly vectorized Fortran program, and a Cyber 205 computer were used in all of the calculations. Each time step took approximately 0.1 s ($N = 256$).

RESULTS AND DISCUSSION

The simulations described in Table II were conducted so as to test the feasibility of the constant volume fluctuation method for the calculation of mechanical partial molar properties. The following discussion, accordingly, is intended to highlight the method's advantages and limitations: it does not constitute an exhaustive exploration of the thermodynamics of the particular mixture under study.

Partial molar property calculations (per molecule) are detailed in Table III. Calculated partial molar volumes satisfy the following thermodynamic identity exactly (N = total number of particles, V = total volume),

$$\sum_i x_i \overline{V}_i^* = 1/\rho^* = V/(N\sigma_{22}^3) = V^*/N. \tag{20}$$

Equation (20) follows from applying Euler's theorem for homogeneous functions[2] to the temperature-, pressure-, and molecule number-explicit expansion of any extensive property (volume in this particular case). The Kirkwood–Buff expression for $\overline{V}_i$ [see Eqs. (4)–(6)] satisfies this relationship identically. Euler's theorem is thus trivially verified when a constant volume fluctuation method is used to compute partial molar volumes. This is a very important advantage of the method but implies that Eq. (20), being an imposed consistency condition, cannot be used to test the calculation's numerical accuracy.

TABLE IV. Percent changes in partial molar properties due to sample and subcell size changes.

	$\Delta(\lvert\overline{U}_1^*\rvert)$	$\Delta(\lvert\overline{U}_2^*\rvert)$	$\Delta(\lvert\overline{H}_1^*\rvert)$	$\Delta(\lvert\overline{H}_2^*\rvert)$	$\Delta(\overline{V}_1^*)$	$\Delta(\overline{V}_2^*)$
N test[a]	+ 7.11	+ 8.31	+ 14.46	+ 10.11	− 2.96	+ 0.83
L test[b]	+ 2.91	+ 3.42	+ 6.97	+ 6.29	− 1.00	+ 0.44

[a] Normalized changes with respect to run 1.

[b] Normalized changes with respect to run 7.

Instead, a "trivial" mixture having identical solute and solvent molecules $(\sigma_{11}/\sigma_{22} = \epsilon_{11}/\epsilon_{22} = m_2/m_1 = 1)$ was simulated (run 8). In this particular case V^*/N is both the molar and the partial molar volume of both "species", and is known *a priori*. Numerical errors in the calculation of partial molar volumes can therefore be independently assessed. As can be seen from Table III, not only is Euler's theorem trivially satisfied, but the individual errors in solute and solvent partial molar volumes with respect to the theoretical value are 1.32% and − 1.01%, respectively.

In order to further test the numerical accuracy of the calculations, both N (sample size), and L (subcell number or reciprocal size) were varied at constant thermodynamic conditions $(x_1, \rho^*, U^*/N)$. The sample size dependence was tested by increasing N from 256 to 500 (runs 1 and 9); the subcell size, by decreasing L from 16 to 8 (or doubling subcell size; runs 8 and 7). Results are summarized in Table IV (the corresponding energy and enthalpy results will be discussed below). It can be seen that very good numerical accuracy results from an $O(10^2)$ sample size, and an $O(10)^{-1}$ subcell size. In addition, because of the Euler condition, errors in solute and solvent partial molar volumes are always of opposite signs.

If, then, the Kirkwood–Buff expression for $\overline{V}_i$ is written in terms of fluctuations instead of pair correlation integrals, it provides the basis for extremely accurate partial molar volume calculations via (N,U,V) molecular dynamics. This method cannot, however, be extended to infinite dilution: not only is Eq. (4) explicit in actual mole fraction, but quantities such as $\langle(\delta N_1)^2\rangle$ and $\langle\delta N_1 \delta N_2\rangle$ lose statistical significance as N_1 is decreased, the minimum solute mole fraction required being a function of sample and subcell size (N, L^{-1}).

The statistical significance of computed fluctuations as the solute concentration is decreased is addressed in Fig. 3. This illustration shows the calculated mixture and partial

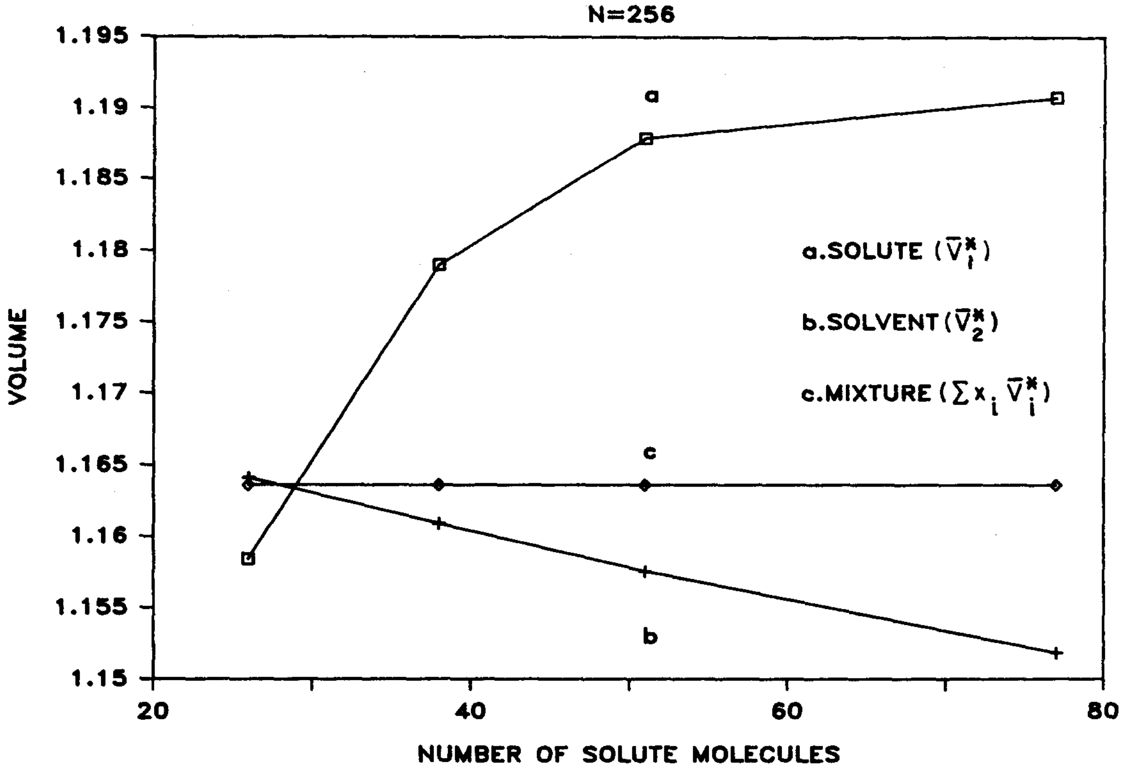

FIG. 3. Numerical solute-number dependence of $\overline{V}_1^*$, $\overline{V}_2^*$, and $\Sigma x_i \overline{V}_i^*$. Trivial mixture, $\langle T^* \rangle = 0.974$; $N_1 = 26,38,51,77$; eight subcells. $\Sigma x_i \overline{V}_i^*$ is indistinguishable from $V^* N^{-1}$.

molar volumes as a function of N_1 over four otherwise identical simulations of the trivial mixture (runs 7, 10, 11, and 12). Note that solute and solvent are now indistinguishable: the simulations yield identical values for U^*, $\overline{\overline{P}}^*$, $\overline{\overline{T}}^*$, etc. Differences between $\overline{V}_1^*$ and $\overline{V}_2^*$ are therefore due to N_1 dependence exclusively. A threefold variation in N_1 from 26 to 77 molecules gives rise to just 2.8% and 1.1% changes in $\overline{V}_1^*$ and $\overline{V}_2^*$, respectively. Also, within the range of compositions explored, these two quantities differ, at most, by 3.4%. Although the solute and solvent curves necessarily coincide (and are hence error free) when $x_1 = 0.5$, the accuracy at 10% mole fraction is significant, suggesting the possibility of extending the technique well into the $x_1 < 1\%$ region. Such a study is currently being done, and will be reported in future publications. The trivial mixture is therefore extremely useful for the simultaneous quantification of the method's accuracy (since the theoretical value of both partial molar properties is known), consistency (Euler condition), and usefulness for dilute mixture studies (through the N_1 dependence tests).

Figures 4 and 5 show selected results corresponding to Table III for nontrivial mixtures. Figure 4 illustrates the isothermal composition dependence of $\overline{V}_1^*$ and $\overline{V}_2^*$ over four simulations (runs 6, 5, 4, and 1) in which N was changed from 38 to 77. Temperature dependence is shown in Fig. 5. Also shown in Figs. 4 and 5 is the calculated value of $\Sigma x_i \overline{V}_i^*$, which is indistinguishable from the theoretical value (V^*/N) in all cases.

Table III also lists results for partial molar energy and enthalpy calculations, the latter being obtained from the equation

$$\overline{H}_i = \overline{U}_i + P\overline{V}_i \,. \tag{21}$$

Figures 6 and 7 illustrate some of the results. Contrary to partial molar volume calculations, the Euler condition is not satisfied identically. This follows from the nature of the energy operational definition[10] which is obtained from a temperature-, pressure-, and mass-explicit expansion for dU which contains no information whatsoever on U itself (see also the Appendix). This is not the case with the Kirkwood–Buff expression for $\overline{V}_i$, where the corresponding extensive quantity (volume) is imposed at the outset, both explicitly [Eq. (5)] and implicitly ($\langle\ \rangle$ denotes a fluctuation within a region of space having a specified volume).

That errors as small as 12.2% (energy) and 14.1% (enthalpy) can be obtained at all (run 3) is indeed remarkable: the Euler condition is now a relationship between completely unrelated quantities (computationally speaking). It is nevertheless clear from Table III that this particular operational definition for $\overline{U}_i$ does not in general lead to accurate predictions with sample sizes of $O(10^2)$. The attractive feature of a fluctuation approach, however, lies in the fact that a given derivative property is amenable to a wide variety of operational definitions, having identical theoretical validity but different computational implications. In the specific case of $\overline{U}_i$, alternative operational definitions have been recently derived,[10] and their numerical implementation will be reported in future publications. The ultimate goal is of course the derivation of an operational definition for $\overline{U}_i$ which incorporates the Euler relationship while, at the same time, being amenable to a computer implementation. Theoretical work in this direction has been recently completed.[15]

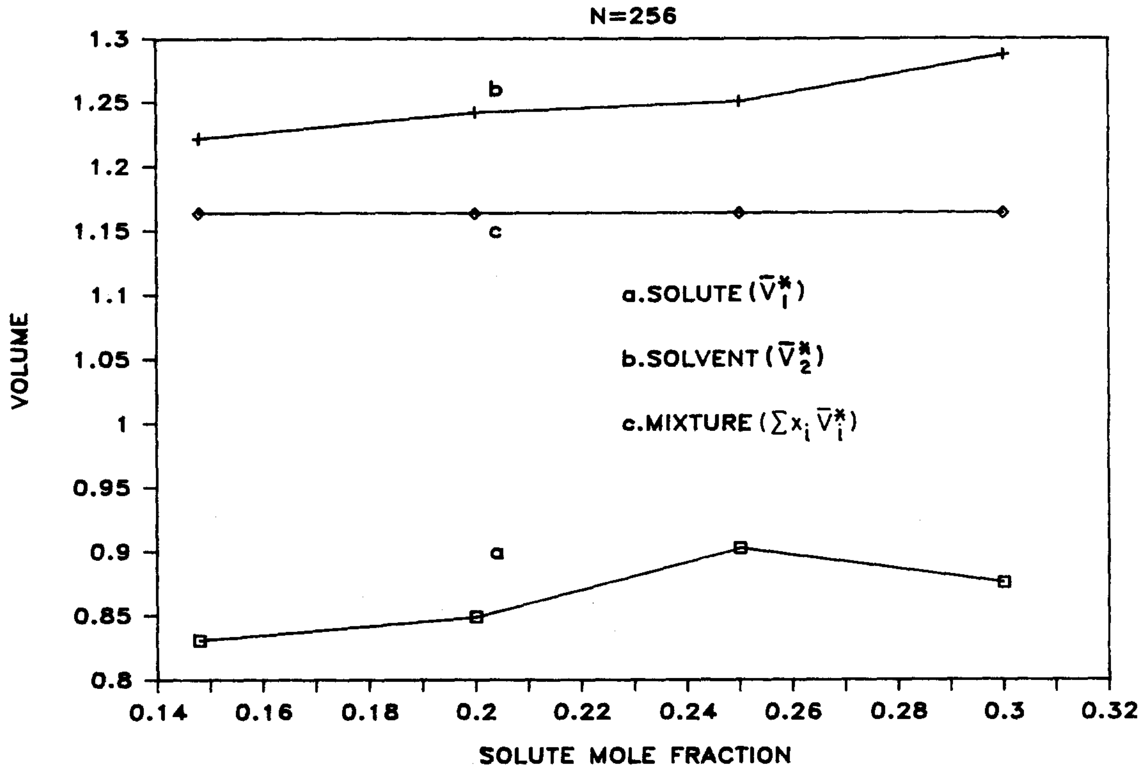

FIG. 4. Isothermal composition dependence of $\bar{V}_1^*$, $\bar{V}_2^*$, $\Sigma x_i \bar{V}_i^*$; 16 subcells. $(N_1, \langle T^* \rangle) = (38,0.913); (51,0.934); (64,0.930); (77,0.923)$. $\Sigma x_i \bar{V}_i^*$ is indistinguishable from $V^* N^{-1}$.

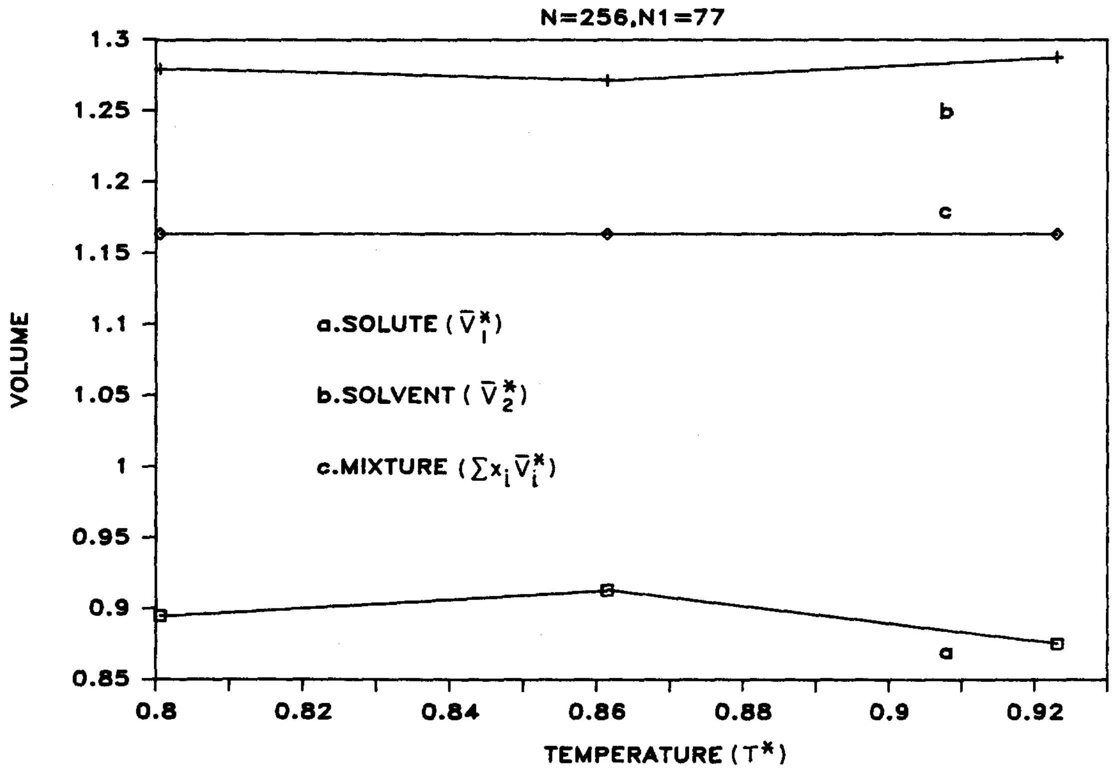

FIG. 5. Temperature dependence of $\bar{V}_1^*$, $\bar{V}_2^*$, $\Sigma x_i \bar{V}_i^*$; 16 subcells. $\langle T^* \rangle = 0.801, 0.861, 0.923$. $\Sigma x_i \bar{V}_i^*$ is indistinguishable for $V^* N^{-1}$.

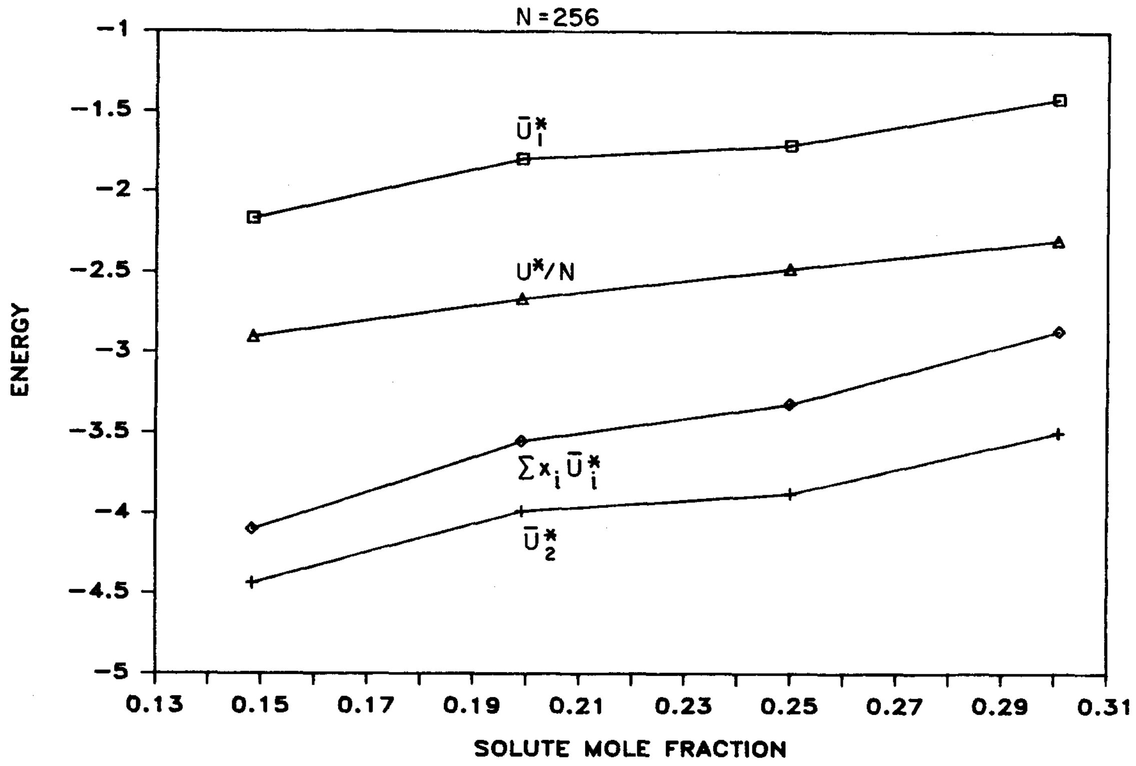

FIG. 6. Isothermal composition dependence of $\bar{U}_1^*$, $\bar{U}_2^*$, $\Sigma x_i \bar{U}_i^*$; 16 subcells. $(N_1, \langle T^* \rangle) = (38, 0.913); (51, 0.934); (64, 0.930); (77, 0.923)$. Note the difference between $\Sigma x_i \bar{U}_i^*$ and $U^* N^{-1}$.

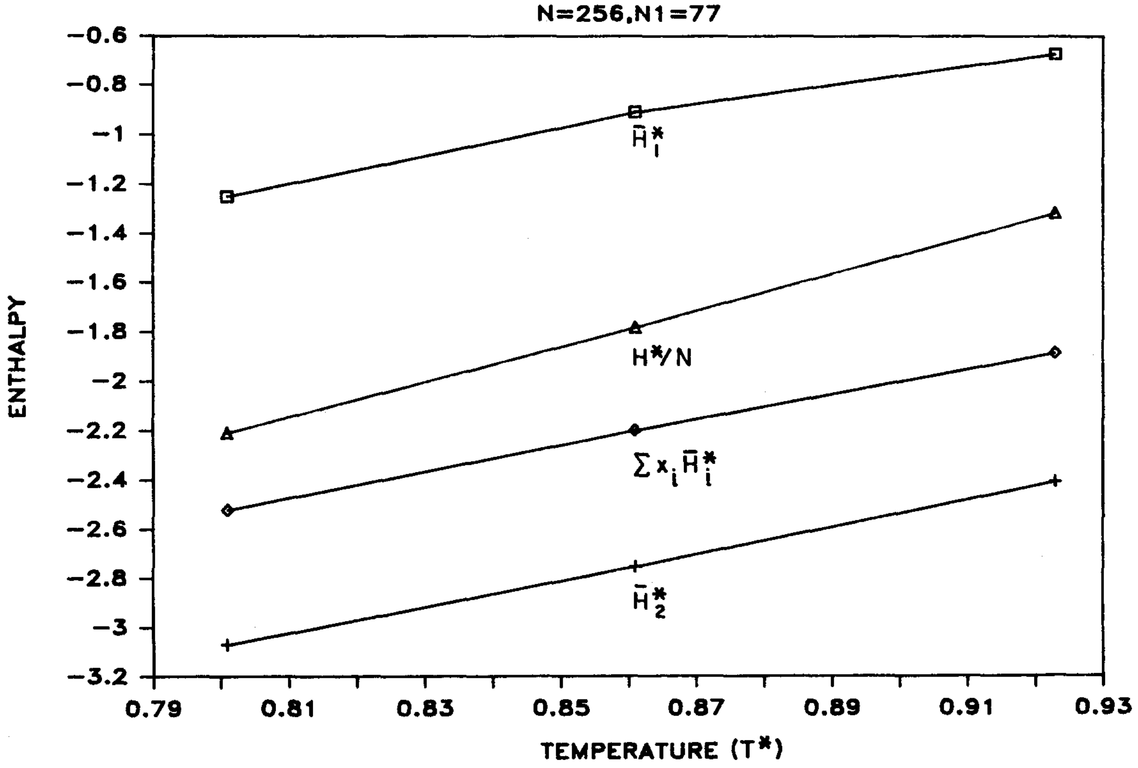

FIG. 7. Temperature dependence of $\bar{H}_1^*$, $\bar{H}_2^*$, $\Sigma x_i \bar{H}_i^*$; 16 subcells. $\langle T^* \rangle = 0.801, 0.861, 0.923$. Note the difference between $\Sigma x_i \bar{H}_i^*$ and $H^* N^{-1}$.

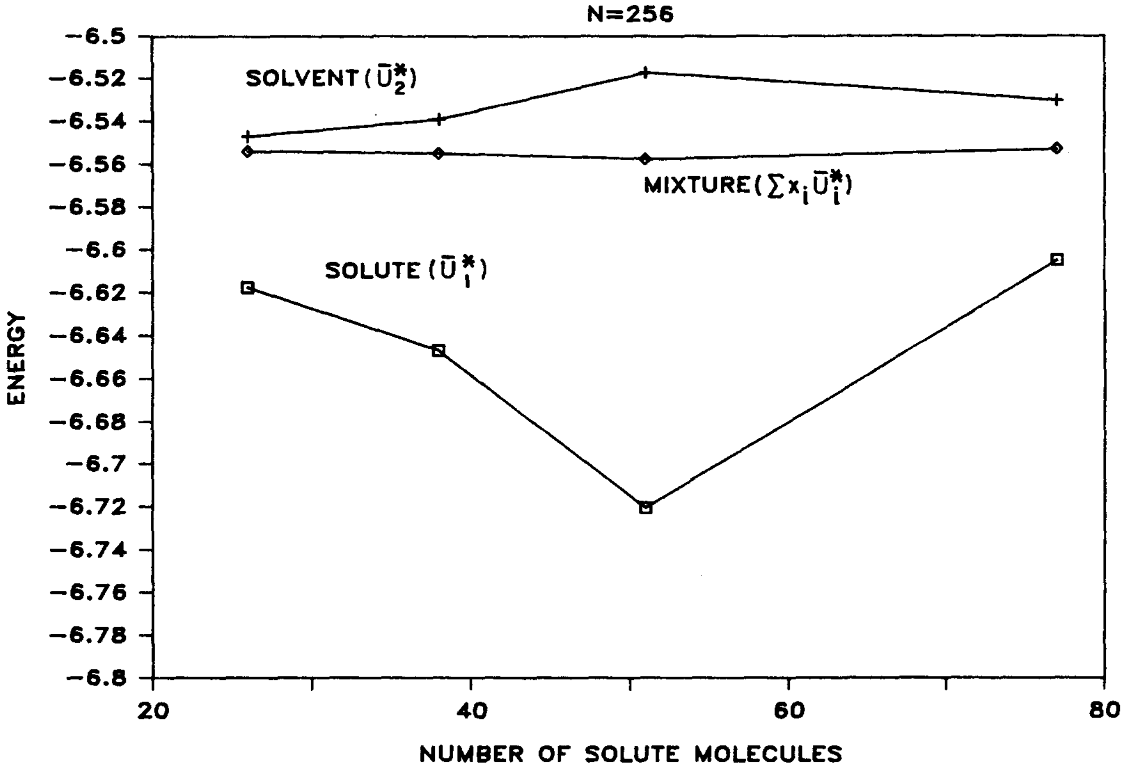

FIG. 8. Numerical solute-number dependence of $\bar{U}_1^*$, $\bar{U}_2^*$, and $\Sigma x_i \bar{U}_i^*$. Trivial mixture, $\langle T^* \rangle = 0.974$; $N_1 = 26,38,51,77$, eight subcells.

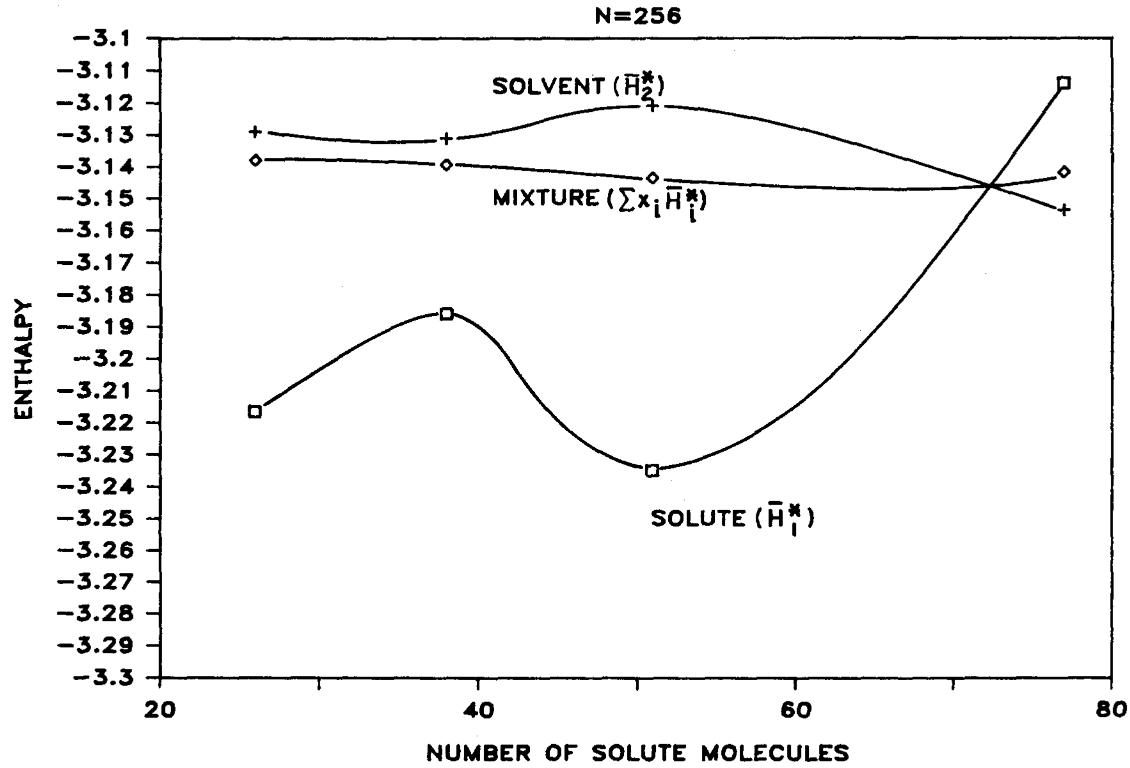

FIG. 9. Numerical solute-number dependence of $\bar{H}_1^*$, $\bar{H}_2^*$, and $\Sigma x_i \bar{H}_i^*$. Trivial mixture, $\langle T^* \rangle = 0.974$; $N_1 = 26,38,51,77$, eight subcells.

The energy and enthalpy calculations were also subjected to the same numerical tests used in the volume case, namely,

(i) N-dependence test (constant $\rho^*, x_1, U^* N^{-1}$; runs 1 and 9).

(ii) L-dependence test (constant $\rho^*, U^* N^{-1}$, trivial mixture; runs 7 and 8).

(iii) N_1-dependence test (constant $\rho^*, U^* N^{-1}$, trivial mixture; runs 7, 10, 11, and 12).

Results are shown in Table IV (N,L tests) and Figs. 8 and 9 (N_1 test). We note from Table IV that sample and subcell size dependence are significantly greater for energy and enthalpy than for volume calculations. In addition, there is no compensation between the errors in $\overline{U}_1^*$ (or $\overline{H}_1^*$) and $\overline{U}_2^*$ (or $\overline{H}_2^*$), since the Euler condition is not imposed.

From the N_1 dependence shown in Figs. 8 and 9, it can be seen that the computed properties are not significantly affected by a threefold change in N_1: maximum solute, solvent, and mixture property variations of 1.74%, 0.4%, and 0.07% (energy), and 3.98%, 0.95%, and 0.15% (enthalpy) occurred over simulations 7, 10, 11, and 12, in which N_1 was decreased from 77 to 26. Contrary to the volume case, though, this desirable numerical stability *vis-a-vis* low solute concentrations is not complemented by a corresponding accuracy (note the large errors shown in Table III for both energy and enthalpy with respect to the theoretical value); accordingly, the Euler condition is not satisfied and longer simulations are required.

The sample size question can be looked at from a slightly different perspective, which is of particular interest in the light of the considerable energy and enthalpy errors mentioned above. Equation (7) is a thermodynamic identity. It follows that its numerical implementation will yield results of arbitrary accuracy for a large enough sample size. This high-N limit, however, is not approached monotonically. To see this, we write, for a generic fluctuation,

$$\langle \delta\lambda_i \delta\lambda_j \rangle = \alpha N^a + \eta N^{-b}, \tag{22}$$

where the first term describes the true (thermodynamic) N dependence, and the second quantity is the numerical error which, for any fluctuation, is a decreasing function of N for a given run length. If we consider energy-component mass covariances, for example, we have ($u = UN^{-1}$)

$$\langle \delta U \delta N_i \rangle = NkT \left[\left(\frac{\partial u}{\partial \mu_i} \right)_{T,V,\mu[i]} + u \left(\frac{\partial \ln N}{\partial \mu_i} \right)_{T,V,\mu[i]} \right] + \text{error} \tag{23}$$

or, in other words, $a = 1$ for this particular fluctuation. Thus, although any given fluctuation approaches its thermodynamic limit monotonically as N is increased, the same is not true in general for a highly nonlinear combination of fluctuations, such as $\overline{U}_i$. That numerical accuracy is not approached monotonically can also be seen from Table III: the difference between $\overline{U}^*/N$ and $\Sigma x_i \overline{U}_i^*$ actually increases from 24.8% to 35.6% when N is increased from 256 to 500 at constant thermodynamic conditions (runs 1 and 9).

TABLE V. Specific heat, isothermal compressibility, and thermal expansion coefficient calculations.

Run	N	$\overline{\overline{T}}^*$	V^*/N	N_1	C_v^*	K_T^*	α_P^*
1	256	0.923	1.1636	77	1.5456	0.3119	0.2197
2	256	0.861	1.1636	77	1.5292	0.3297	0.2498
3	256	0.801	1.1636	77	1.5446	0.3699	0.2518
4	256	0.930	1.1636	64	1.5481	0.3097	0.1992
5	256	0.934	1.1636	51	1.5508	0.2825	0.2129
6	256	0.913	1.1636	38	1.5512	0.2811	0.1815
7	256	0.974	1.1636	77	1.6326	0.1815	0.0863
8	256	0.975	1.1636	77	1.5655	0.2399	0.1526
9	500	0.930	1.1636	150	1.5562	0.2740	0.2063
10	256	0.974	1.1636	51	1.6326	0.1817	0.0863
11	256	0.974	1.1636	38	1.6326	0.1818	0.0864
12	256	0.974	1.1636	26	1.6326	0.1818	0.0864

Finally, in Table V we show the behavior of the specific heat, isothermal compressibility, and thermal expansion coefficient, which were computed via the following equations ($\langle N \rangle = NL^{-1}$):

$$C_v^* \equiv \frac{C_v}{k} = \frac{\langle T^* \rangle^2}{\langle N \rangle \langle (\delta T^*)^2 \rangle}, \tag{24}$$

$$K_T^* \equiv \frac{\epsilon_{22} K_T}{\sigma_{22}^3} = \frac{K_T'}{\rho^* \langle N \rangle \langle T^* \rangle}, \tag{25}$$

$$\alpha_P^* \equiv \frac{\epsilon_{22} \alpha_P}{k} = \frac{K_T' \langle \delta T^* \delta P^* \rangle}{\rho^* \langle N \rangle \langle T^* \rangle \langle (\delta T^*)^2 \rangle}, \tag{26}$$

where, in all cases, $\langle T^* \rangle = \overline{\overline{T}}^*$, $\langle P^* \rangle = \overline{\overline{P}}^*$ and K_T' has already been defined. We thus conclude by emphasizing that the same principles and methods that underlie the well known fluctuation interpretation of C_v, K_T, and α_p can be successfully applied to the computer study of partial molar properties in both binary and multicomponent mixtures, a hitherto unexplored approach.

CONCLUSION

A given partial derivative is amenable to a variety of fluctuation-explicit operational definitions. The possible number of such operational definitions is at least as large as the corresponding ways of defining the extent of the region of space (subsystem) over which fluctuations are measured or calculated.

Operational definitions of thermodynamic derivative properties in terms of constant volume fluctuations can serve as a basis for binary and multicomponent mechanical partial molar quantity computations within a (N,U,V) molecular dynamics simulation. In the case of partial molar volumes, this method gives excellent results at finite solute mole fractions and requires one simulation per data point. Partial molar energy and enthalpy calculations are not exact. Contrary to the volumetric expression, the corresponding energy operational definition does not incorporate the homogeneity constraint (Euler's theorem).

Work has recently been completed[15] on the derivation of a fluctuation-explicit operational definition for $\overline{U}_i$ which incorporates the homogeneity constraint while at the same time being amenable to algorithmic implementation.

The distinguishing feature and main advantage of fluctuation-based methods is the variety of possible operational definitions corresponding to a given derivative property. In this paper we have explored the use of control volumes within a deterministic algorithm. Many other possibilities exist, the recently proposed[10] constant component mass method being but one example. We believe that fluctuation-based simulations combine theoretical challenge, computational interest and practical implications (i.e., the molecular-based study of nonideal mixture thermodynamics) to an uncommon degree.

APPENDIX: THE PARTIAL MOLAR ENERGY OPERATIONAL DEFINITION

We outline and correct a derivation originally presented elsewhere.[10] We start by writing

$$dU=\sum_{i=1}^{n}\overline{U}_i\,dN_i+\left(\frac{\partial U}{\partial T}\right)_{P,N_1\cdots N_n}dT + V(PK_T-T\alpha_P)dP \qquad \text{(A1)}$$

and, therefore, with $j=1,\ldots,n$, we have

$$\left(\frac{\partial U}{\partial N_j}\right)_{T,V,\mu[j]}=\overline{U}_j+\sum_{i\neq j}\overline{U}_i\left(\frac{\partial N_i}{\partial N_j}\right)_{T,V,\mu[j]} + V(PK_T-T\alpha_p)\left(\frac{\partial P}{\partial N_j}\right)_{T,V,\mu[j]} \qquad \text{(A2)}$$

which can be regarded as a system of n equations in n unknowns ($\overline{U}_j$). The coefficients in this system of equations, namely, K_T, α_p, and the three partial derivatives can all be expressed in terms of constant volume fluctuations[7–11]

$$\left(\frac{\partial U}{\partial N_j}\right)_{T,V,\mu[j]}=\frac{\langle\delta U\delta N_j\rangle}{\langle(\delta N_j)^2\rangle}, \qquad \text{(A3)}$$

$$\langle N\rangle\rho K_T=\beta|B|V/\sum_{i=1}^{n}\sum_{j=1}^{n}x_ix_j|B|_{ij}, \qquad \text{(A4)}$$

$$\left(\frac{\partial N_i}{\partial N_j}\right)_{T,V,\mu[j]}=\frac{\langle\delta N_i\delta N_j\rangle}{\langle(\delta N_j)^2\rangle}, \qquad \text{(A5)}$$

$$C_v=\frac{kT^2}{\langle N\rangle\langle(\delta T)^2\rangle}, \qquad \text{(A6)}$$

$$\left(\frac{\partial P}{\partial N_j}\right)_{T,V,\mu[j]}=\frac{\langle\delta P\delta N_j\rangle}{\langle(\delta N_j)^2\rangle}, \qquad \text{(A7)}$$

$$\alpha_P=\frac{\langle N\rangle K_T\langle\delta T\delta P\rangle C_v}{kT^2}, \qquad \text{(A8)}$$

where $|B|$ and $|B|_{ij}$ have already been defined.

Substitution of Eqs. (A3)–(A8) into Eq. (A2) leads straightforwardly to Eq. (7) which allows the solution of the linear system for the unknown $\overline{U}_j$ ($j=1,\ldots,n$) in terms of constant volume fluctuations.

Equation (A1) imposes no constraint upon U and N: the homogeneity condition $U=\Sigma N_i\overline{U}_i$ is obtained only upon Euler integrating[2] this differential relationship. The resulting fluctuation-explicit operational definition for $\overline{U}_i$, accordingly, contains no information on U and N_i and does not satisfy the homogeneity condition identically.

We now address the fluctuation-explicit operational definition for K_T. Equation (A4) is simply the Kirkwood-Buff relationship[11] for K_T: it was used here with B_{ij} expressed in terms of concentration fluctuations instead of pair correlation integrals. We can, on the other hand, write[8]

$$\beta\langle(\delta V)^2\rangle_{N_1,\ldots,N_n}=VK_T. \qquad \text{(A9)}$$

In the original derivation,[10] the above equation, after combination with the (correct) relationship

$$\langle(\delta V)^2\rangle_N=\frac{1}{\rho^2}\langle(\delta N)^2\rangle_V \qquad \text{(A10)}$$

yielded the (incorrect) equation $\beta\langle(\delta N)^2\rangle=\langle N\rangle\rho K_T$. The subtle point here is the realization of the difference between the $\langle\ \rangle_N$ and $\langle\ \rangle_{N_1,\ldots,N_n}$ constraints. The former merely specifies the total number of molecules; the latter, in addition, imposes the composition. Both constraints are of course identical for a pure substance. In the case of mixtures, however, Eqs. (A9) and (A10), though both correct, cannot be combined, since

$$\langle(\delta V)^2\rangle_N\neq\langle(\delta V)^2\rangle_{N_1,\ldots,N_n}. \qquad \text{(A11)}$$

This difference is conceptually profound (as can be seen from trying to implement the corresponding measurement processes in thought experiments) but numerically unimportant since, in an equilibrium system, the constraints are statistically (but never operationally) indistinguishable. To illustrate this point, we compare below the values of $\langle N\rangle\rho kTK_T$ calculated according to Eq. (A4) with the (incorrect) value, namely, $\langle(\delta N)^2\rangle$:

Run	$\dfrac{(\langle N_1\rangle+\langle N_2\rangle)^2[\langle(\delta N_1)^2\rangle\langle(\delta N_2)^2\rangle-\langle\delta N_1\delta N_2\rangle^2]}{\langle N_2\rangle^2\langle(\delta N_1)^2\rangle+\langle N_1\rangle^2\langle(\delta N_2)^2\rangle-2\langle N_1\rangle\langle N_2\rangle\langle\delta N_1\delta N_2\rangle}$	$\langle(\delta N)^2\rangle$
1	3.957 957	4.381 512
2	4.015 305	4.268 445
3	4.074 624	4.384 812
4	3.959 651	4.193 510
5	3.628 034	3.914 996
6	3.528 105	3.731 655
7	4.863 219	4.871 229
8	3.216 182	3.217 268
9	6.843 152	7.815 971
10	4.867 757	4.871 229
11	4.870 396	4.871 229
12	4.871 167	4.871 229

Note the equality between both quantities for the trivial mixture simulations (7,8,10,11,12), as demanded by the above discussion. Partial molar energies and enthalpies computed via the (incorrect) relationship $\beta \langle (\delta N)^2 \rangle = \langle N \rangle \rho K_T$ were always within 1% of the "correct" values computed via Eq. (A4).

ACKNOWLEDGMENTS

Supercomputer time grants from the National Science Foundation (CBT-8517739) and the John von Neumann Center National Allocations Committee are gratefully acknowledged. All calculations were performed on a Cyber 205 machine at the John von Neumann Center for Scientific Computing at Princeton. I wish to express my thanks to Carolyn Bolton and Amy Ciric for their help with communications and data transfer, and to Professor Athanassios Panagiotopoulos for helpful discussions.

[1] P. G. Debenedetti, Chem. Eng. Sci. (in press).
[2] M. Modell and R. C. Reid, *Thermodynamics and its Applications*, 2nd ed. (Prentice–Hall, Englewood Cliffs, 1983).
[3] J. Owicki and H. Scheraga, J. Am. Chem. Soc. **99**, 7413 (1977).
[4] G. Bolis and E. Clementi, Chem. Phys. Lett. **82**, 147 (1981).
[5] K. S. Shing and S. T. Chung, AIChE Annual Meeting, Miami, November 1986, paper 64C.
[6] L. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon, Oxford, 1980), Part 1.
[7] P. G. Debenedetti, J. Chem. Phys. **84**, 1778 (1986).
[8] R. F. Greene and H. B. Callen, Phys. Rev. **83**, 1231 (1951).
[9] A. Z. Panagiotopoulos and R. C. Reid, J. Chem. Phys. **85**, 4650 (1986).
[10] P. G. Debenedetti, Chem. Phys. Lett. **132**, 325 (1986).
[11] J. G. Kirkwood and F. P. Buff, J. Chem. Phys. **19**, 774 (1951).
[12] A. Z. Panagiotopoulos, U. W. Suter, and R. C. Reid, Ind. Eng. Chem. Fundam. **25**, 525 (1986).
[13] W. B. Streett, D. J. Tildesley, and G. Saville, in *Computer Modelling of Matter*, ACS Symp. Ser. 86, edited by P. Lykos (American Chemical Society, Washington, D.C., 1978), Chap. 13.
[14] L. Verlet, Phys. Rev. **159**, 98 (1967).
[15] P. G. Debenedetti, J. Chem. Phys. (in press).