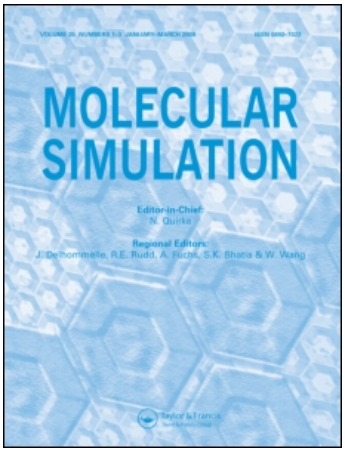

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

A New Version of the Insertion Particle Method for Determining the Chemical Potential by Monte Carlo Simulation

Ivo Nezbeda[a]; Jiří Kolafa[a]
[a] Institute of Chemical Process Fundamentals, Czechoslovak Academy of Sciences, Prague, Czechoslovakia

To cite this Article Nezbeda, Ivo and Kolafa, Jiří(1991) 'A New Version of the Insertion Particle Method for Determining the Chemical Potential by Monte Carlo Simulation', Molecular Simulation, 5: 6, 391 — 403
To link to this Article: DOI: 10.1080/08927029108022424
URL: http://dx.doi.org/10.1080/08927029108022424

Molecular Simulation, 1991, Vol. 5, pp. 391–403
Reprints available directly from the publisher
Photocopying permitted by license only

A NEW VERSION OF THE INSERTION PARTICLE METHOD FOR DETERMINING THE CHEMICAL POTENTIAL BY MONTE CARLO SIMULATION

IVO NEZBEDA and JIŘÍ KOLAFA

Institute of Chemical Process Fundamentals, Czechoslovak Academy of Sciences, 165 02 Prague, Czechoslovakia

(*Received March 1990, accepted March 1990*)

A new version of the test particle method for determining the chemical potential by Monte Carlo simulations is proposed. The method, applicable to any fluid at any density, combines the Widom's test particle insertion method with the ideas of the scaled particle theory, gradual insertion method and multistage sampling. Its applicability is exemplified by evaluating the chemical potential of the hard sphere fluid at a very high density in semi-grand-canonical and grand-canonical ensembles. A theory estimating the efficiency (i.e. statistical errors) of the method is proposed and the results are compared with the Widom's and gradual insertion methods, and the analytic results.

KEY WORDS: Monte Carlo simulation, chemical potential, scaled particle, gradual insertion

1. INTRODUCTION

While the evaluation of the mechanical properties (average energy, pressure, etc.) of classical N-particle systems is quite straightforward and has become the routine calculation in the last decade [1], the evaluation of the statistical (entropical) properties (e.g. the Helmholtz or Gibbs free energies) has posed severe problems. In addition to simulations in a grand-canonical ensemble [2, 3], special general techniques have been developed to overcome these problems: thermodynamic integration methods and the energy distribution method and its modifications leading to various non-Boltzmann samplings (for a review see e.g. [3–5]).

In recent years a good deal of attention has been paid especially to the chemical potential for its key role in determining fluid phase equilibria [6]. The most direct method for determining the chemical potential is the Widom's test particle insertion method [7]. It has been widely used because it calls only for minor changes to conventional Monte Carlo (MC) or molecular dynamics programs. The principal drawback of this technique is its apparent failure at liquid-like densities because of the predominance of configurations in which the test particle overlaps one or more of the other molecules. This drawback can be overcome by using special techniques (like restricted or full umbrella samplings or f–g sampling [8, 9]) or tricks (excluded volume map sampling [10]). Yet the problem remains if we deal, for instance, with hard body fluids, fluids at very high densities, or mixtures of significantly different components.

To overcome the above mentioned problem of determining the chemical potential we propose in this paper a new version of the test particle method which is conceptually very simple and is applicable to any fluid and their mixtures at any density. The

basic idea is borrowed from the scaled particle theory [11] in which the process of inserting a particle into the system of N real particles is divided into two steps: a point-wise particle is inserted first and then it is let gradually grow. Neither step poses any problem: the reversible work (and thus the chemical potential) associated with the first step is either known analytically (e.g. for hard body fluids) or is very simply determined from one MC run. The second step corresponds, in a certain sense, to the thermodynamic integration over a coupling parameter and may be realized on computer in three different ways: (A) by performing a series of common simulations (a sort of the multistage sampling [4]) each with the test particle of different size (a gradual insertion method [12]), (B) by performing a simulation in a NVT ensemble in which the size of the inserted particle fluctuates (we call it the semi-grand-canonical ensemble simulation), and (C) by carrying out a simulation on a modified (scaled) grand-canonical ensemble. While method A is a realization of the known methods, methods B and C are novel versions of the insertion particle method.

For fluids made up of molecules with hard cores usual methods fail already at intermediate densities or are even (for hard body fluids) inapplicable. For this reason the fluid of hard spheres seems a very suitable system to demonstrate applicability of the proposed method. Its principles and details are explained in section 2. In section 3 we derive approximate expressions for the statistical errors from which the optimum values of simulation parameters are obtained. Technical details are given in section 4. Results obtained for the fluid of hard spheres at one very high packing fraction ($\eta = 0.45$) via different routes are compared and discussed in section 5.

2. PRINCIPLE OF THE METHOD

Let us consider a system of N identical particles. For the sake of simplicity we assume that they interact via a pair-wise additive spherical potential $u(r; \sigma)$ where σ is a parameter which may (but may not) be identified with the diameter of a molecule (extension to angle-dependent interactions, mixtures, systems with non-additive forces etc. is straightforward). The total potential energy is thus

$$U_N = \sum_{i<j} u(r_{ij}; \sigma), \tag{1}$$

where r_{ij} is the separation between particles i and j. The chemical potential, μ, is given by the change of the Helmholtz free energy associated with adding an additional particle to the system and may be written as

$$\beta\mu_r = \beta(\mu - \mu_{id}) = -\ln \frac{Q_{N+1}}{VQ_N} = -\ln \langle \exp[-\beta(U_{N+1} - U_N)] \rangle_N, \tag{2}$$

where μ_{id} is the chemical potential of an ideal gas at the same temperature and density, Q_N is the configurational partition function of an N-particle system, $\langle \cdot \rangle_N$ means an ensemble average in a system of N particles, and $\beta = 1/kT$.

Direct evaluation of the ensemble average in (2) (Widom's method, [17]) is straightforward at low and medium densities: a system of N molecules is sampled and a (Boltzmann) probability of inserting a fictitious (ghost) particle into a random place in the fluid is calculated. With increasing density the probability rapidly decreases and the method becomes inefficient unless special samplings are used.

The problem of determining the thermodynamic properties by the insertion method

was addressed three decades ago by Reiss and his co-workers [11] in their scaled particle theory. The direct insertion of a real full-size particle is circumvented by inserting a point-wise particle first, which causes no problems, and then this point-wise guest particle is gradually let swell until it becomes identical with the remaining host particles. This gradual swelling corresponds to the thermodynamic integration over a coupling parameter [4]. In computer language, instead of inserting the whole particle at once, the insertion may be done piecemeal using the multistage sampling to bridge the states of the systems with N and $N + 1$ real indentical particles. We will denote this method, which is in fact an example of the gradual insertion method [12], as method A:

$$\frac{Q_{N+1}}{VQ_N} = \frac{Q_{N+\sigma_1}}{VQ_N}\frac{Q_{N+\sigma_2}}{Q_{N+\sigma_1}}\cdots\frac{Q_{N+\sigma_k}}{Q_{N+\sigma_{k-1}}}$$

$$= \langle\exp[-\beta(U_{N+\sigma_1} - U_N)]\rangle_N \prod_{i=1}^{k-1}\langle\exp[-\beta(U_{N+\sigma_{i+1}} - U_{N+\sigma_i})]\rangle_{N+\sigma_i}$$

$$\equiv q_0\prod_{i=1}^{k-1} q_i \tag{3}$$

Here $Q_{N+\sigma_i}$ denotes the configurational partition function of the system of $N + 1$ particles, one of which (we call it the scaled particle) has the diameter σ_i, and $\sigma_1 < \sigma_2 < \ldots < \sigma_k \equiv \sigma$. (Note that the pair potential of the scaled particle with the normal one is here $u(r; [\sigma + \sigma_i]/2)$; a different dependence on σ may be used for other meanings of σ.) In other words, it means to perform k simulations: one to obtain q_0 (i.e. the probability of inserting a point-wise or small particle; for $\sigma_1 = 0$ and some systems this is known analytically) and $k - 1$ simulations to obtain q_i, $i > 1$.

The above implementation of the insertion method seems rather cumbersome and sometimes, as it was already discussed in [12], also time consuming. It would be more convenient to determine the chemical potential in one MC run instead in k separate ones. This may be done by assembling all the canonical ensembles considered above into one ensemble by assigning them appropriate weights. A change of the size of the scaled particle thus becomes one of the usual MC trial steps. Let us consider an ensemble which may undergo the following transitions (method B)

$$[N] \underset{1}{\rightleftarrows} [N + \sigma_1] \underset{w_1}{\rightleftarrows} [N + \sigma_2] \underset{w_2}{\rightleftarrows} \ldots \rightleftarrows [N + \sigma_k \equiv N + 1], \tag{4}$$

where the brackets denote appropriate sub-ensembles, and w_i are weights assigned to the respective sub-ensembles. Transition $[N] \to [N + \sigma_1]$ means that a particle of diameter σ_1 is created at a random position in the fluid and $[N + \sigma_1] \to [N]$ that it is annihilated; the other transitions denote the change of the diameter of the particle. According to a general formula expressing the detailed balance condition (see e.g. [13]), the change (by one position in scheme (4) either to the right or to the left) is accepted with the probability

$$\min\left\{1, \frac{p_{\text{new}\to\text{old}}}{p_{\text{old}\to\text{new}}}\frac{w_{\text{new}}\exp(-\beta U_{\text{new}})}{w_{\text{old}}\exp(-\beta U_{\text{old}})}\right\}. \tag{5}$$

where $p_{\text{new}\to\text{old}}$ and $p_{\text{old}\to\text{new}}$ denote *a priori* (trial) transition probabilities of the respective changes. The desired residual chemical potential is then given by

$$\beta\mu_r = \ln (w_k \, \mathrm{Prob}[N]/\mathrm{Prob}[N+1]), \tag{6}$$

where Prob[N] is the probability of observing state [N] in the ensemble undergoing transitions (4). The way the appropriate weights w_i and transition probabilities are actually chosen and their explicit expressions are given in the following section.

The above method, which we call the semi-grand-canonical ensemble simulation, can be easily extended to a common grand-canonical ensemble (method C). In place of the finite transition scheme (4) we may adopt the following infinite one starting from state [0] corresponding to an empty system:

$$\begin{aligned} &[0] \underset{w_{[0]}}{\rightleftarrows} [0+\sigma_1] \underset{w_{[0+\sigma_1]}}{\rightleftarrows} [0+\sigma_2] \underset{w_{[0+\sigma_2]}}{\rightleftarrows} \ldots \rightleftarrows [1] \underset{w_{[1]}}{\rightleftarrows} \ldots \\ \ldots &\rightleftarrows [N] \underset{w_{[N]}}{\rightleftarrows} [N+\sigma_1] \underset{w_{[N+\sigma_1]}}{\rightleftarrows} [N+\sigma_2] \underset{w_{[N+\sigma_2]}}{\rightleftarrows} \ldots \rightleftarrows [N+1] \ldots \end{aligned} \tag{7}$$

where the weights assigned to the respective sub-ensembles should satisfy the same condition as for the common grand-canonical simulation ([2, 3]):

$$w_{[N+1]}/w_{[N]} = B/(N+1), \tag{8}$$

where B is a parameter of the simulation. The residual chemical potential μ_r is then given by

$$\beta\mu_r = \ln (B/\langle N \rangle), \tag{9}$$

where $\langle N \rangle$ is the grand-canonical average number of particles. It should be noted that the sequence of configurations generated according to (7) contains two classes of configurations: those in which a smaller scaled particle exists and 'normal' ones in which no scaled particle is present. The latter configurations are distributed according to the grand-canonical ensemble and all ensemble averages (including $\langle N \rangle$ entering (9)) should be taken only over these states.

Formula (2) and its consequences can be used for molecules both with and without hard cores. The complementary equation,

$$\beta\mu_r = \ln \frac{VQ_N}{Q_{N+1}} = \ln \langle \exp[\beta(U_{N+1} - U_N)] \rangle_{N+1}, \tag{10}$$

in which a real particle is virtually removed, is restricted only to ideally soft molecules (without a core) and although the original form is useless for realistic molecules, the idea is used in more sophisticated methods [8, 9]. A decomposition similar to (3) can be easily written also for (10) and it might serve as the basis for a method complementary to A. A combination with the original method A would be, however, more efficient because one simulation of system $[N + \sigma_i]$ can be used to obtain the chemical potential both for swelling $\sigma_i \to \sigma_{i+1}$ and shrinking $\sigma_i \to \sigma_{i-1}$. The latter transition is, however, feasible only in such cases in which any configuration important for ensemble $[N + \sigma_{i-1}]$ can be obtained (by shrinking) from a configuration which occurs in the simulated ensemble $[N + \sigma_i]$ with a reasonable probability. This is not satisfied for molecules with very steep repulsions and especially for molecules with hard cores. However, both these complementary approaches (3) and (10) are naturally included in methods B and C.

A continuous range of diameters $(0, \sigma)$ of the scaled particle might be used theoretically in method B instead of the discrete set. An extrapolation (or interpola-

tion because the range may exceed the interval $(0, \sigma)$) of the histogram of the density of states is necessary to obtain the chemical potential. We, however, cannot imagine any advantage of the continuous approach either in method B or methods A and C.

3. DETERMINATION OF PARAMETERS AND ESTIMATION OF ERRORS

All three above considered methods contain several free parameters: the number of scaled steps k, the sequence of diameters σ_i, and methods B and C also the weights and transition probabilities defining the Markov chain. In this section we first put forward criteria for a suitable determination of the diameters σ_i, weights w_i, and transition probabilities. Approximate expressions for expected statistical errors of the chemical potential in dependence on k for all three methods are then derived and their minimization gives finally the optimum value of k.

One may assume that the optimum process will be obtained if it is equally probable to swell a particle from σ_i to σ_{i+1} for all i, i.e. if the chemical potentials (reversible works) for these one-step swellings are the same. Being guided by the scaled particle theory we assume that $\sigma_i = (i/k)^{1/2}\sigma$ which means that the chemical potential is proportional to the square of σ_i which was also verified *a posteriori* (see Figure 1). To incorporate the reversible work associated with the insertion of a point-wise particle into the system, we set

$$\sigma_i = [(i/k - \alpha)/(1 - \alpha)]^{1/2}\sigma \tag{11}$$

for $1 \leqslant i \leqslant k - 1$ (method A) and $1 \leqslant i \leqslant k$ (methods B and C), where $0 < \alpha < 1$. A more accurate estimate of good σ_is can be obtained from the chemical potential (Henry constant) predicted by a suitable equation of state (if this is available) for the mixture at infinite dilution: the diameters σ_i should be specified so as to correspond to the same differences of the chemical potential.

Method A does not employ the weights w_i and transition probabilities so that an estimate of the error and hence the optimum value of k can be derived without their knowledge. For the sake of simplicity we assume that convergence properties of the simulations of all k systems are the same (they are given by a certain correlation length τ_A; see e.g. [14–15] for the definition) although this is not generally correct: a small scaled particle may travel faster through the surrounding liquid. Let us denote by n the total number of measurements of the quantities of interest available for all k sub-simulations.

Configurations in which the scaled particle "overlaps" a host particle (i.e. it lies within the range of strong repulsions) do not contribute to the average of q_i, Equation (3). The most important contributions, E, come from the configurations in which the scaled molecule is close to its equilibrium position in the liquid. (For fluids of hard particles $E = 1$ while for soft molecules we usually have $E > 1$ due to long range attractive forces.) A configuration giving a non-zero contribution to the average occurs with a probability $r_i \approx q_i/E$. In addition, we may assume that all r_is are the same if σ_is are appropriately chosen; $r_i = R^{1/k}$ where R is the probability that a randomly inserted molecule in the simple Widom's method does not overlap other molecules; $R = \exp(-\beta\mu_r)$ for hard bodies. The relative error of r_i or q_i is then

$$\left[\frac{1 + 2\tau_A}{n/k}\frac{(1 - r_i)}{r_i}\right]^{1/2}, \tag{12}$$

where n/k is the number of measurements for one sub-simulation. The total error of $\beta\mu_r$ for method A is thus

$$\delta_A(\beta\mu_r) = \left[k^2(R^{-1/k} - 1)\frac{1 + 2\tau_A}{n}\right]^{1/2}. \tag{13}$$

Since $R < 1$, $\delta_A(\beta\mu_r)$ reaches a minimum for a certain positive k.

As regards methods B and C, we have to specify first the weights w_i and transition probabilities. In addition to that, formula (6) for the chemical potential in method B is not in the form of a simple ensemble average so that some rearrangements are necessary which would allow us both to compute the statistical error and estimate it theoretically.

It may be reasonable to fix, along with (11), the weights so that $w_{i+1}/w_i \approx \exp(\beta\mu_r/k)$ in method B and analogously $w_{[N+\sigma_{i+1}]}/w_{[N+\sigma_i]} \approx [B/(N+1)]^{1/k}$ in method C. Probability to accept the trial move, Equation (5), contains a weighting factor $p_{\text{new}\to\text{old}}w_{\text{new}}/(p_{\text{old}\to\text{new}}w_{\text{old}})$. A practical advantage is gained if these weighting factors equal unity, especially for simulations on hard body systems: in that case the overlap test is recovered. One can easily achieve this by assigning appropriate transition probabilities.

In the following algorithm we assume that the weights increase in the respective chains (4) and (7) from the left to the right which is typical for dense or hard body fluids; extension to the opposite case is straightforward. Transition probability $p_1 = 1$ is assigned to the first step in (4) and (7), i.e. $[N] \to [N + \sigma_1]$ and $[0] \to [0 + \sigma_1]$, and

$$p_j = 1 - p_{j-1}/a^{1/k} \tag{14}$$

for the consecutive steps, $j = 2,3, \ldots$, in the respective chains. Here $a > 1$ denotes either $\exp(\beta\mu_r)$ (method B) or $B/(N+1)$ (method C). The transition probabilities of inverse changes (i.e. shrinkages of the scaled particle) are evidently $1 - p_j$. An exception occurs in method B because chain (4) is truncated at $[N+1]$ while (14) gives nonzero $p_{[N+1]\to[N+1+\sigma_1]}$; this transition should be evidently interpreted as an empty transition $p_{[N+1]\to[N+1]}$. The inverse transition remains unchanged: $p_{[N+1]\to[N+\sigma_{k-1}]} = 1 - p_{[N+1]\to[N+1+\sigma_1]}$.

Formula (6) for the chemical potential of method B contains the ratio of probabilities $\text{Prob}[N]/\text{Prob}[N+1]$ which is in the simulation estimated by

$$\sum_{i=1}^{n} \chi_i[N] \Big/ \sum_{i=1}^{n} \chi_i[N+1], \tag{15}$$

where $\chi_i[N]$ is unity if state $[N]$ is observed in the i-th MC step and zero otherwise. To estimate its error, the following trick may be used:

$$\text{Var}\left(\sum_{i=1}^{n} \chi_i[N] \Big/ \sum_{i=1}^{n} \chi_i[N+1]\right)$$

$$\approx \text{Var}\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\langle\chi[N]\rangle\,\chi_i[N+1] - \langle\chi[N+1]\rangle\,\chi_i[N]}{\langle\chi[N+1]\rangle^2}\right), \tag{16}$$

where the expectation values, $\langle\chi[N]\rangle \equiv \text{Prob}[N]$, may be replaced by the MC estimates. In other words, the error of (15) equals the error of the average of a single quantity. To be correct, Equation (6) with the ratio of probabilities estimated by (15)

is a biased estimator for $\beta\mu_r$ because its expectation value, in expansion to the second order, is

$$\beta\mu_r - \tfrac{1}{2}\varepsilon_r^2(\chi_i[N]) + \tfrac{1}{2}\varepsilon_r^2(\chi_i[N+1]), \tag{17}$$

where ε_r denotes the relative error of the respective quantity. Since in (17) only the second power of the errors appears, the correction is in practice negligible, as it was also confirmed by the simulations.

The acceptance ratio of swelling or inserting a particle in methods B and C is determined, similarly as the virtual acceptance ratio in method A, mainly by an overlap test and is approximately given by $R^{1/k}$. The corresponding transition probability is approximately $1/(1 + a^{-1/k})$ as follows from the iterative use of (14) for $j \to \infty$. The average probability of a move by one to the right in the chain is then

$$D = R^{1/k}/(1 + a^{-1/k}) \tag{18}$$

which is the same as that of one move to the left. (Note that $Ra = 1$ for hard bodies). The process of changing states in chains (4) and (7) is now described by a random walk on $k + 1$ (approximately) equivalent positions in method B and on positions which are occupied with an (approximately) Gaussian probability distribution in method C where D may be identified with the autodiffusion coefficient.

Neglecting the correlations between two consecutive steps, the development of the state in method B is given by

$$\begin{aligned}
\chi_0^{t+1} &= D\chi_1^t + (1 - D)\chi_0^t \\
\chi_i^{t+1} &= D\chi_{i-1}^t + D\chi_{i+1}^t + (1 - 2D)\chi_i^t, \qquad 0 < i < k \\
\chi_k^{t+1} &= D\chi_{k-1}^t + (1 - D)\chi_k^t,
\end{aligned} \tag{19}$$

where χ_i^t is the probability of observing the wandering particle at position i at time t. Eigenvectors of the stochastic operator defined by (19) (it is a $(k + 1) \times (k + 1)$ matrix) may be obtained by a method which is analogous to the solution of the continuous diffusion on a line segment. From formula (13) of [15] and Equation (16) above we then obtain an estimate of the error of $\beta\mu_r$ for method B:

$$\delta_B(\beta\mu_r) = \left\{\left[k\,\frac{1 + a^{-1/k}}{R^{1/k}}(1 + 2\tau_B) - 1\right]\frac{2(k+1)}{n}\right\}^{1/2}, \tag{20}$$

where the factor $1 + 2\tau_B$ has been introduced to account for the consecutive correlations in the series of configurations leading to a lower effective autodiffusion coefficient, $D/(1 + 2\tau_B)$, than (18). This correction is justified only in such simulations in which the structural changes in the neighbourhood of the scaled particle are fast enough in comparison with a random walk, i.e. in which the normal MC displacements, rotations and other changes are made sufficiently often.

Dynamic properties of simulation C may be approximated, similarly as for method B, by a random walk along chain (7). It may be supposed, for enough large N, that the probability distribution of N is Gaussian with a variance Var $N \gg 1$ and mean value $\langle N\rangle$. Since we assume that all sub-states $[N + \sigma_i]$ are equivalent (it is the same as considering $\tau_A = \text{const}$ for all sub-simulations in method A), the probabilities of all states in chain (7) have the Gaussian distribution, too. Further, the fluctuations at most of one particle (i.e. by k positions in chain (7)) are fast enough compared with

fluctuations over the typical accessible range (which is $(\mathrm{Var}\, N)^{1/2} \leqslant 1$) and thus the convergence properties (namely, the correlation length) of observable N are the same as those of the observable denoting the position in chain (7). In addition, we may adopt a continuous description, i.e. the diffusion in a one-dimensional quadratic potential hole. Similarly as χ in (19), the probability density $y \equiv y(N, t)$ of observing state $[N]$ at time t satisfies the equation in which the differences are replaced by derivatives:

$$D\frac{\partial^2 y}{\partial N^2} + d\frac{\partial(Ny)}{\partial N} = \frac{\partial y}{\partial t}, \tag{21}$$

where $D = k^{-2}R^{1/k}/(1+a^{-1/k})$ is the autodiffusion coefficient of a random walk (see method B), $d = D/\mathrm{Var}\, N$, and $\langle N \rangle = 0$ is assumed for the sake of simplicity. The spectrum of stochastic operator (21) is $\exp(-jd), j = 0,1, \ldots$. The time development of eigenvectors of (21) is, for $j < 2$, given by

$$y_j(N,t) \equiv y_j(N)\mathrm{e}^{-jdt} = \exp\left[-\frac{N^2}{2\,\mathrm{Var}\, N}\right] N^j \mathrm{e}^{-jdt}. \tag{22}$$

Function $y_0(N)$ is, after normalization, just the equilibrium probability distribution. The quantity of interest, N, is $y_1(N)/y_0(N) = N$ and thus only one term survives in sum (13) of [15] yielding the correlation length of N as $2\,\mathrm{Var}\, N/D$ (provided that $d \ll 1$) and the error of N is then

$$\delta N = \mathrm{Var}\, N\left[2k^2(1 + a^{-1/k})R^{-1/k}/n\right]^{1/2}. \tag{23}$$

It follows from (9) that the equivalent error of the chemical potential (if N is constant) is given by $\delta\mu_r = [\partial\mu_r/\partial N - 1/\beta N]\,\delta N$. Using the well known formula $\partial(\beta\mu_r)/\partial N = 1/\mathrm{Var}\, N$ and neglecting $1/\beta N$ with respect to $\partial\mu_r/\partial N$ (it is justified for large densities) we obtain the equivalent error of the chemical potential.

$$\delta_C(\beta\mu_r) = \left(2k^2\frac{1 + a^{-1/k}}{R^{1/k}}\frac{1 + 2\tau_C}{n}\right)^{1/2}, \tag{24}$$

where the factor $1 + 2\tau_C$ has the same origin as in method B; in addition $\tau_B \approx \tau_C$ may be assumed.

It follows from comparison of errors of methods B and C (Equations (20) and (24)) that the grand-canonical simulation C is more efficient. However, the difference is, for typical a and R, quite negligible; further, if either σ_is or corresponding weights are changed in method B so that the relative probabilities of the states of interest, i.e. $[N]$ and $[N + 1]$, are slightly greater in comparison with the probabilities of the states containing the scaled particle, slightly lower errors are expected. Methods B and C may be thus, for a pure fluid, considered as equivalent. Comparison with method A is more difficult: Although structural changes in the neighbourhood of the scaled particle may influence both the correlation length τ_A in method A and the factor τ_B (or τ_C) defining the effective autodiffusion coefficient in method B (or C) in a similar way, we are not able to make a detailed quantitative comparison and thus we refer to the results of particular MC simulations.

Another consequence of these considerations is that the measured sequence of N in method C forms the first order autoregressive process and that the data for method B have the exponential decay of autocorrelation coefficients not too different from the

first order autoregressive process, especially for lower k. These properties will be made use of for estimating the errors.

4. IMPLEMENTATION AND TECHNICAL DETAILS

For reasons mentioned already in section 1 the fluid of hard spheres has been chosen to test the porposed method. The thermodynamic properties of this fluid are known nowadays with high accuracy: The compressibility factor

$$z = \frac{1 + \eta + \eta^2 - \frac{2}{3}(\eta^3 + \eta^4)}{(1 - \eta)^3} \tag{25}$$

has the expected relative error less than 2×10^{-4} [16]. By integration we then obtain the residual chemical potential

$$\beta\mu_r = z - 1 + \frac{1}{6}\left(4\eta + 10\ln(1 - \eta) + \frac{20}{1 - \eta} + \frac{5}{(1 - \eta)^2} - 25\right). \tag{26}$$

This expression is used both as a benchmark against which the computed results are checked and for estimating some parameters of the simulation method.

To implement the proposed method we used the common Monte Carlo setup: a cubic box with periodic boundary conditions. The box contained 107 host hard spheres of unit diameter and one scaled sphere. The edge length of the box was set to 5.0011 which corresponds to a very high packing fraction, $\eta = 0.45$ for $N = 107.5$. In the basic cycle eight displacements and one swelling were attempted. One displacement was defined so that a scaled particle attempted a move with probability 0.25 while a randomly chosen normal size hard sphere with probability 0.75; no systematic study was made whether these numbers were optimum or not. Lengths of the displacements in all implementations were set so that the acceptance ratios lay around 0.2 (with the exception of the first simulation in method A where a point-wise particle moves in the liquid). The total number of 11.5×10^6 trial moves (i.e. displacements and size fluctuations) were performed for each method, which required about 100 hours on an IBM/PC-XT (4.77 MHz + 8087); the programs were written in Turbo-Pascal.

As regards method A, the expectation value of q_0 (see Equation (3)) is, for $\sigma_1 = 0$, given by the probability that a randomly placed point does not intersect any sphere, $q_0 = 1 - 107\eta/107.5$. The values $\alpha = 0$ and $0 \leqslant i \leqslant k - 1$ were used in (11) to obtain the sequence of the diameters where the value $k = 7$ was found to be optimum, see Equation (13). Since the convergence is better for a smaller scaled particle, the total number of generated configurations was divided, approximately, to the respective sub-simulations as follows: 4%, 8%, 11%, 14%, 18%, 21%, 24%; no check was made whether this was the optimum choice.

Similarly as in method A, adding a point-wise particle was omitted in the implementation of method B. The value $k = 6$ was found to be optimum, see (20), and σ_i, $0 \leqslant i \leqslant k$, were obtained from (11) with $\alpha = 0$. The weighting factor $w_k = 125000$ was estimated from (26). For the grand canonical simulation by method C we used the same box and the parameter $b = \exp(17.0328)$ corresponding to the expected average number of particles $\langle N \rangle = 107.5$, see Equations (8), (9) and (26). The values $k = 6$ (minimizing the error (24)) and $\alpha = 0.0484$ were used to define the sequence of diameters according to Equation (11).

The statistical errors have been obtained by the method published e.g. in [14–15] with the autocorrelation coefficients of the measured series of data extrapolated by the power decay with the exponent of $-3/2$ for method A (the reasons are discussed in [17]) and by an exponential decay for methods B and C, see the end of section 3. The mean errors (at the 68% confidence level) are used.

The proposed method of evaluating the chemical potential employs systems with one guest particle and is thus ideally suited for the application of a preferential sampling. Further, in the course of the simulations the scaled particle may disappear (method C) or become identical with other particles (both method B and C). Also this fact may be made use of in an effort to make the sampling of the configuration space more efficient.

If the particle disappears or becomes identical with the remaining particles its identification is no more relevant. The next shrinkage, $[N + 1] \rightarrow [N + \sigma_{k-1}]$, thus should involve a (random) choice of a new scaled particle which will attempt to shrink, otherwise the sampling of the configuration space around the scaled particle would be less efficient. Even if a single solute particle is simulated (i.e. Henry constant is computed), we recommend to interchange (swap) the (scaled) solute with a (randomly chosen) solvent molecule; the change is accepted according to the standard Metropolis test. This MC step may be, however, efficient only if the solute is, for certain σ_i, similar enough to the solvent molecule. For more details see [18].

The preferential sampling method (see e.g. [13]) is based on an idea that structural changes in a neighbourhood of a simple solute particle (the scaled particle in our case) are more important and, consequently, the simulation will be more efficient if the MC moves are performed more often with molecules in a vicinity of the scaled particle than with molecules in the bulk liquid. This preference is described by a function $f(r)$ denoting the relative probability that a host particle, a distance r apart from the scaled particle, attempts a trial move. We therefore performed also one simulation on the semi-grand-canonical ensemble (method B) to see to what extent the preferential sampling may affect the efficiency of the simulation.

In order to set the proper function $f(r)$, the following circumstances should be taken into account: (i) Fast convergence of the simulation in the case when the scaled particle is point-wise (or small) as well as when it is identical with the other molecules, and (ii) inefficiency caused by the fact that the fluctuations of the diameter of the scaled particle follow the structural changes. For these reasons we expect that the fast decaying (and easily computable) function $f(r) = (\sigma/2 + \sigma_i/2)^4/r^4$ will be close to the optimum one.

The smaller the scaled particle, the faster is convergence and less intensive shuffling of its neighbourhood is necessary. That is why we use the definition of a "preferred move" which consists of the following steps: (i) a host particle is chosen at random; (ii) with probability $1 - f(r)$ the preferred move is abandoned; (iii) a new position of the chosen particle is generated; (iv) with probability $1 - \min(1, f(r_{\text{new}})/f(r_{\text{old}}))$ the preferred move is abandoned; (v) the trial position is tested for an overlap. It means that host particles surrounding the big scaled particle make more moves than those around the small one. In addition, no move is made as soon as the scaled particle achieves its maximum diameter because it looses then its identity and its neighbourhood is no more relevant.

In the course of the MC run of method B with the preferential sampling we repeated the following cycle: (i) one attempted swelling/shrinkage; (ii) 50 steps which were either a displacement of the scaled particle (with probability 0.05) or a "preferred

move" (with probability 0.95). Totally 4.8×10^6 trial displacements of the scaled particle, 4.5×10^6 trial displacements of the host particle (i.e. those actually tested for an overlap), and 2.2×10^6 attempted swellings/shrinkages were made. No detailed study was made whether these arranagements were optimum.

Since the preferential function $f(r)$ was rapidly decreasing, about 10% of the total computer time was wasted by the procedure choosing the host particle to move. This number will be, however, quite negligible for simulations of more realistic intermolecular potentials because the time to compute the potential becomes the leading one; therefore this increase in time is not taken into account in assessing the results in the next section.

5. RESULTS AND DISCUSSION

The dependence of the chemical potential of a small solute particle in the hard sphere fluid on its diameter, as obtained from methods A and B, is drawn in Figure 1. The values of the chemical potential are related to a point-wise solute "molecule" as the standard state. It is seen that the chemical potential grows approximately linearly with the square of the diameter which *a posteriori* justifies the use of Equation (11) defining the series of diameters.

The results for the chemical potential from different methods are compared in Table 1. It is seen that all results are in agreement with the analytic estimate. The methods however differ in efficiency: Method A is only slightly better than the simplest Widom's insertion method (though it might be better if the preferential sampling were used) and cannot be recommended. This is also in agreement with observations of Mon and Griffiths [12]. The comparable runs of novel methods B and

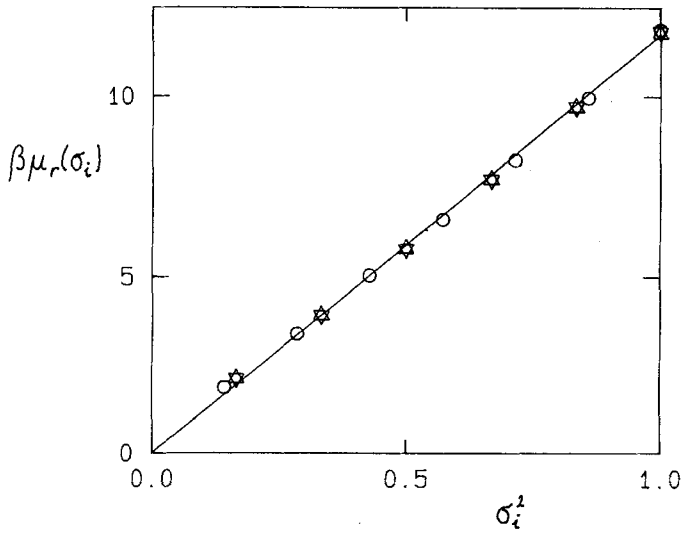

Figure 1 The residual chemical potential $\beta\mu_r(\sigma_i)$ of a scaled hard sphere of diameter σ_i in the fluid of hard spheres at the packing fraction $\eta = 0.45$, using $\sigma_i = 0$ as the standard state. o: method A; ∇: method B; Δ: method B with the preferential sampling; straight line is given by the chemical potential from Equation (26)

Table 1 Residual chemical potential of the fluid of hard spheres at the packing fraction $\eta = 0.45$ computed by various methods with the same number of MC steps. The last column contains the number of independent measurements.

method	$\beta\mu_r$	$\delta(\beta\mu_r)$	$n/(1 + 2\tau)$
Equation (26)	12.355		
Widom		$\gtrsim 0.152$	$\lesssim n$ (a)
A	12.47	0.135	200–300
B	12.43	0.083	1250
C	12.32	0.063	75 (b)
B	12.38	0.048	2700 (c)

(a) theoretical value
(b) equivalent values; the directly measured values were $\langle N \rangle = 107.63$, $\delta N = 0.20$
(c) with the preferential sampling

C give lower errors. In addition, method C gives by 24% smaller error than B, although the theory (Equations (20) and (24)) predicts the error of method C to be only by 2–3% smaller (the actual value depends on the value of $\tau_B = \tau_C$). A slightly better performance of method C may be expected also due to the fact that after reaching 'normal' states $[N]$ or $[N + 1]$ the transitions along chain (4) (method B) are allowed only in one direction while in method C they may go in both directions (chain (7)). Nonetheless, the discrepancy between the theory and experiment is too large. The main source of this discrepancy is likely inaccuracy of the value of error $\delta_C(\beta\mu_r)$ because the correlation length in the series of data is about 1% of the total length of the MC run and thus the expected uncertainty of the error is approximately 10%; in addition, the method used tends to underestimate τ_C in short runs.

From examining Table 1 it is seen that the use of the preferential sampling leads to further lowering of the error for method B; it may be expected that it would lead to a similar lowering of the errors also for methods A and C.

We may summarize that methods B (semi-grand-canonical) and C (grand-canonical) are almost equivalent for computing the chemical potential of a pure fluid (or a mixture of comparable components at comparable concentrations). The long correlation time of the grand-canonical simulation may, however, cause problems in simulations of large systems and it also reduces reliability of the estimate of the statistical errors (which are the indispensable part of the results) and we thus recommend method B rather than C. Further, method B is suitable for the evaluation of the Henry constant even if the solute molecule is larger than the solvent.

6. CONCLUSIONS

We have developed a new version of the insertion particle method which is especially suited for evaluation of the chemical potential in those cases where the up to date known methods fail, for instance, for dense hard body fluids and any very dense fluids in general, or for mixtures of significantly different components. The method is quite general and can be also combined with other methods, like e.g. with the Gibbs ensemble method for direct determining phase equilibria [6]. The method is currently being applied to determining the chemical potential of the Lennard-Jones fluid to clear up the existing discrepancies [10] and to examining phase equilibria of fluid mixtures in pores; these results will be reported in due course.

Acknowledgement

We would like to thank Dr. S. Labík for useful discussions.

References

[1] J.P. Valeau and S.G. Whittington, "A guide to Monte Carlo for statistical mechanics: 1. Highways", in *Statistical Mechanics*, Part A, B.J. Berne, ed, Plenum, New York, 1977, ch. 4; *Computer modeling of matter*, P. Lykos, ed, ACS Symposium Series, 1978.

[2] D.J. Adams, "Grand canonical ensemble Monte Carlo for a Lennard-Jones fluid", *Mol. Phys.*, **29,** 307 (1975).

[3] M.P. Allen and D.J. Tildesley, *Computer simulations of liquids*, Clarendon Press, Oxford, 1987.

[4] J.P. Valeau and G.M. Torrie, "A guide to Monte Carlo for statistical mechanics: 2. Byways", in *Statistical Mechanics, Part* A, B.J. Berne, ed, Plenum, New York, 1977, ch. 5.

[5] D. Frenkel, "Free-energy computation and first-oder phase transitions", in *Molecular-dynamics simulation of statistical-mechanical systems*, Soc. Italiana di Fisica, Bologna, 1986.

[6] K.E. Gubbins, "The role of computer simulation in studying fluid phase equilibria", *Molecular Simulation*, **2,** 223 (1989).

[7] B. Widom, "Some topics in the theory of fluids", *J. Chem. Phys.*, **39,** 2808 (1963).

[8] K.S. Shing and K.E. Gubbins, "The chemical potential from computer simualtion. Test particle method with umbrella sampling", *Mol. Phys.*, **43,** 717 (1981).

[9] K.S. Shing and K.E. Gubbins, "The chemical potential in dense fluids and fluid mixtures via computer simulations", *Mol. Phys.*, **46,** 1109 (1982); C.H. Bennett, "Efficient estimation of free energy differences from Monte Carlo data", *J. Comput. Phys.*, **22,** 245 (1976).

[10] G.L. Deitrick, L.E. Scriven and H.T. Davis, "Efficient molecular simulation of chemical potentials", *J. Chem. Phys.*, **90,** 2370 (1989).

[11] H. Reiss, H.L. Frisch and J.L. Lebowitz, "Statistical mechanics of rigid spheres", *J. Chem Phys.*, **31,** 369 (1959); E. Helfand, H. Reiss and H.L. Frisch, "Scaled particle theory of fluids", *J. Chem. Phys.*, **33,** 1379 (1960).

[12] K.K. Mon and R.B. Griffiths, "Chemical potential by gradual insertion of a particle in Monte Carlo simulations", *Phys. Rev. A*, **31,** 956 (1985).

[13] C. Pangali, M. Rao and B.J. Berne, "On a novel Monte Carlo scheme for simulating water and aqueous solutions", *Chem. Phys. Lett.*, **55,** 413 (1978).

[14] A. Beretti and A.D. Sokal, "New Monte Carlo method for the self-avoiding walk", *J. Stat. Phys.*, **40,** 483 (1985); T.P. Straatsma, H.J.C. Berendsen and A.J. Stam, "Estimation of statistical errors in molecular simulation calculations", *Mol. Phys.*, **57,** 89 (1986).

[15] J. Kolafa, "On optimization of Monte Carlo simulations", *Mol. Phys.*, **63,** 559 (1988).

[16] T. Boublik and I. Nezbeda, "P-V-T behaviour of hard body fluids. Theory and experiment", *Colln. Czech. Chem. Commun*, **51,** 2301 (1986).

[17] J. Kolafa, "Convergence properties of Monte Carlo simulations on fluids", *Molecular Simulation*, submitted.

[18] J. Kolafa, "A novel approach to improve convergence of Monte Carlo simulations on fluids", *Molecular Simulation*, to be submitted.