

The overlapping distribution method to compute chemical potentials of chain molecules

G C A M Mooij and D Frenkel

FOM Institute for Atomic and Molecular Physics, Kruislaan 407, 1098 SJ Amsterdam, The Netherlands

Received 17 January 1994

Abstract. The chemical potential of continuously deformable chain molecules can be estimated by measuring the average Rosenbluth weight associated with the virtual insertion of a molecule. We show how to generalize the overlapping-distribution method of Bennett to histograms of Rosenbluth weights. In this way we arrive at a scheme to estimate chemical potentials of chain molecules that is a direct generalization of the Shing–Gubbins scheme for simple molecules. In particular, our overlapping-distribution method has the nice diagnostic feature that it can detect systematic sampling problems that may occur for long chains and high densities. We apply the method to the computation of the chemical potential of flexible chains of hard spheres and find that, for the systems studied, systematic sampling errors are less important than statistical errors.

1. Introduction

The chemical potential is a key quantity in the numerical study of phase coexistence. In equilibrium, the chemical potential of each species has to be the same in the coexisting phases. The most widely used method to compute chemical potentials in moderately dense fluids is the test-particle method of Widom [1]. This scheme is based on an expression that relates the excess chemical potential, μ^{ex} , and the average Boltzmann factor of the energy, U_{test} , of a ‘test’ particle that is inserted at a random position in the system:

$$\beta\mu^{\text{ex}} = -\ln\langle e^{-\beta U_{\text{test}}} \rangle \quad (1)$$

where the angular brackets denote ensemble averaging and $\beta = 1/k_{\text{B}}T$. For systems with repulsive intermolecular interactions, $\langle e^{-\beta U_{\text{test}}} \rangle$ can be interpreted as the probability of acceptance of a (virtual) Monte Carlo move in which a particle is added at random to the system. In principle, the Widom insertion technique can be used for any fluid, but in practice it is limited to small particles and to systems at low or moderate densities. To obtain a reliable estimate of the insertion probability, $\langle e^{-\beta U_{\text{test}}} \rangle$, and as such an accurate value for the chemical potential, it is necessary to reach regions of configurational space where the energy U_{test} is low. For systems at high densities, and in the case of large molecules even at moderate densities, the vast majority of random insertions results in an overlap with the repulsive core of a particle in the system, and contributes negligibly to the insertion probability. Under these circumstances, astronomically long simulations would be required in order to obtain μ^{ex} with good statistical accuracy (a more quantitative estimate of the computational effort required can be found in [2]). In a simulation that is too short, those fluid configurations that contribute most to $\langle \exp[-U_{\text{test}}/k_{\text{B}}T] \rangle$ would be inadequately sampled, and incorrect (too high) estimates for μ^{ex} result. Shing and Gubbins devised a

scheme to measure excess chemical potentials in such a way that sampling problems are automatically detected [3]. The Shing-Gubbins method is a special case of the overlapping-distribution method of Bennett [4]. The basic idea behind the Shing-Gubbins scheme is to combine particle insertions (that tend to probe high-energy states) with virtual trial moves in which a real particle is removed from the system (real particles will typically be situated in regions of space where the potential energy is not high). In the overlapping-distribution method, a histogram ($p_0(\Delta U)$) is made that measures the probability of finding a potential energy change ΔU on insertion of an $(N+1)$ th particle in a system of N particles. Similarly, a histogram $p_1(\Delta U)$ measures the probability of finding a potential energy change ΔU on particle removal from a system of $N+1$ particles. If there is a range of ΔU where these histograms overlap, we have [3, 4]

$$\begin{aligned} \ln p_1(\Delta U) &= \ln \left(\frac{\int d\mathbf{r}^N \exp[-\beta U(\mathbf{r}^N)] \int d\mathbf{r}_{N+1} \exp[-\beta U(\mathbf{r}_{N+1})] \delta(U(\mathbf{r}_{N+1}) - \Delta U)}{\int d\mathbf{r}^N \exp[-\beta U(\mathbf{r}^N)] \int d\mathbf{r}_{N+1} \exp[-\beta U(\mathbf{r}_{N+1})]} \right) \\ &= \ln \left(\frac{\int d\mathbf{r}^N \exp[-\beta U(\mathbf{r}^N)] \int d\mathbf{r}_{N+1} \exp[-\beta \Delta U] \delta(U(\mathbf{r}_{N+1}) - \Delta U)}{\Lambda^{3N} (N+1)! Z(N+1, V, T)} \right) \\ &= \ln \left(\frac{p_0(\Delta U) V Z(N, V, T)}{\Lambda^{3(N+1)} Z(N+1, V, T)} \right) - \beta \Delta U \\ &= \ln p_0(\Delta U) - \beta \Delta U + \beta \mu^{\text{ex}} \end{aligned} \quad (2)$$

where $U(\mathbf{r}^N)$ is the total energy of the system of N particles at positions \mathbf{r}^N , and $U(\mathbf{r}_{N+1})$ is the interaction energy of the $(N+1)$ th particle at position \mathbf{r}_{N+1} with the other N particles present. $Z(N, V, T) = (1/\Lambda^{3N} N!) \int d\mathbf{r}^N \exp[-\beta U(\mathbf{r}^N)]$ is the partition function of an N -particle system, with Λ^{-3N} a temperature-dependent factor that results upon integration over all momenta. For atomic systems, Λ is the thermal de Broglie wave length. This method has the same range of applicability as the test-particle method but, in addition, it is possible to judge the reliability of the results by checking whether $p_0(\Delta U)$ and $p_1(\Delta U)$ can indeed be measured accurately in a region of ΔU where they overlap.

The limitations of the particle insertion technique to calculate chemical potentials of atomic particles are even more serious for chain molecules. In fact, the test-particle method breaks down for chains longer than four or five segments, for all but the lowest densities [5]. Recently, we have shown that by using a biased sampling scheme for continuously deformable chain molecules [9], the test-particle method can be used to calculate chemical potentials for appreciably longer chains. This modified particle insertion is based on a generalization [7, 8] of the Rosenbluth sampling scheme for chain molecules [6]. To judge the reliability of the Rosenbluth sampling technique for continuously deformable chain molecules, it is desirable to use a diagnostic tool, such as the overlapping-distribution method, in combination with Rosenbluth sampling. In the present article we show how this can be achieved. In the following section we derive our overlapping-distribution method. In the third section we present the results of chemical-potential calculations for a fully flexible chain of hard spheres in a hard-sphere fluid.

2. Chemical-potential calculations by Rosenbluth sampling

2.1. Rosenbluth sampling

In Rosenbluth sampling, chain configurations are generated step by step by successive insertion of the bonded segments of the chain. When the positions are chosen at random,

it is very likely that one of the segments will overlap with another particle in the system, which results in a rejection of the trial move. The Rosenbluth sampling scheme increases the insertion probability by looking one step ahead. On lattices the availability (i.e. the Boltzmann factor) of all sites adjacent to the previous segment should be tested. In continuous space there is, in principle, an infinite number of positions that should be tested (e.g. in the case of a flexible chain molecule with rigid bonds, all points on the surface of a sphere with a radius equal to the bond length). Of course, it is not feasible to scan an infinite number of possibilities. Fortunately, however, it turns out that it is possible to construct a correct Monte Carlo scheme for off-lattice models in which only a finite number of trial segments (k) is selected either at random or, more generally, drawn from the distribution, P^{id} , of bond lengths and bond angles of the 'ideal' chain molecule. By an 'ideal' chain molecule we mean an isolated chain that has the same bond, bending, and torsion potentials as the real molecule under study, while all non-bonded interactions (intra- and inter-molecular) are set to zero. An ideal chain molecule will, in general, be non-self-avoiding. Later on, we shall take both the inter-molecular and the non-bonded intra-molecular interactions into account. For an ideal chain molecule, we can easily generate conformations with the correct Boltzmann weight. More generally, if we construct an ideal chain step by step, we can select several trial directions for the next segment with a probability that is determined by the bonded interactions, U^{id} , only:

$$P^{\text{id}}(\Gamma_{ij}) = e^{-\beta U_{ij}^{\text{id}}} / Z_i^{\text{id}} \quad (3)$$

where Γ_{ij} is the j th trial direction for the i th segment, and $Z_i^{\text{id}} = \int d\Gamma_{ij} e^{-\beta U_{ij}^{\text{id}}}$ which is the same for every j . From here on, the procedure that we follow is the same for lattices and continuous-space systems. For each of the trial positions, we compute the Boltzmann factor associated with the non-bonded interactions U^{nb} (both intra- and inter-molecular). One of these trial positions, Γ_i , is then selected with a probability, P , proportional to its Boltzmann factor

$$P(\Gamma_i) = e^{-\beta U_{\Gamma_i}^{\text{nb}}} / Z_i \quad (4)$$

where $Z_i = \sum_{j=1}^{k_i} e^{-\beta U_{\Gamma_j}^{\text{nb}}}$. In this way, regions of high potential energy, such as the hard core of another particle, are avoided and configurations with a non-vanishing Boltzmann weight are generated. To correct for the bias introduced by this very non-random sampling procedure, a weight has to be assigned to each conformation, Γ , called the Rosenbluth weight W_Γ [6]. The contribution of each i th segment to this Rosenbluth weight is equal to the average of the Boltzmann factors of the trial positions for this segment:

$$W_{\Gamma_i} = \frac{1}{k_i} \sum_{j=1}^{k_i} e^{-\beta U_{\Gamma_j}^{\text{nb}}} \quad (5)$$

where $U_{\Gamma_j}^{\text{nb}}$ is the non-bonded energy of the j th trial direction for the i th segment. The Rosenbluth weight of the total configuration, Γ , is the product of the weights of the individual segments, including the Boltzmann factor of the energy of the first segment, U_{Γ_0} :

$$W_\Gamma = e^{-\beta U_{\Gamma_0}} \prod_{i=1}^{\ell} W_{\Gamma_i} \quad (6)$$

where ℓ is the chain length.

2.2. Test-particle Rosenbluth method

As was shown in [9], the chemical potential of chain molecules in a fluid can be calculated from the average Rosenbluth weight of test chains inserted into the fluid. Chain configurations, Γ , are generated by first inserting a segment at a random position Γ_0 within the volume, V , of the fluid and by choosing for every following i th segment a set of trial directions $\{\Gamma_{ij}\}_{\text{trial}}$ ($j = 1, \dots, k_i$), from which the position Γ_i of the segment is chosen through the Rosenbluth procedure described above. The average Rosenbluth weight, $\langle W_\Gamma \rangle_N$, of these configurations is equal to

$$\begin{aligned} \langle W_\Gamma \rangle_N &= \left\langle \frac{1}{V} \int d\Gamma_0 e^{-\beta U_{r_0}} \int \prod_{i=1}^{\ell} d\{\Gamma_{ij}\}_{\text{trial}} P^{\text{id}}(\{\Gamma_{ij}\}_{\text{trial}}) W_{\Gamma_i} \sum_{j=1}^{k_i} P(\Gamma_{ij}) \right\rangle_N \\ &= \frac{1}{V} \left\langle \int d\Gamma_0 e^{-\beta U_{r_0}} \int \prod_{i=1}^{\ell} d\{\Gamma_{ij}\}_{\text{rest}} P^{\text{id}}(\{\Gamma_{ij}\}_{\text{rest}}) d\Gamma_i P^{\text{id}}(\Gamma_i) k_i P(\Gamma_i) W_{\Gamma_i} \right\rangle_N \\ &= \frac{1}{V Z^{\text{id}}} \left\langle \int d\Gamma e^{-\beta U_\Gamma} \int \prod_{i=1}^{\ell} d\{\Gamma_{ij}\}_{\text{rest}} P^{\text{id}}(\{\Gamma_{ij}\}_{\text{rest}}) \right\rangle_N \\ &= \frac{\Lambda^{3(N+1)} (N+1)! Z_{N+1}}{\Lambda^{3N} N! Z_N V Z^{\text{id}}} \end{aligned} \quad (7)$$

where the angular brackets $\langle \dots \rangle_N$ denote ensemble averaging over all configurations of the N particles in the fluid, and $\{\Gamma_{ij}\}_{\text{rest}}$ denotes the remaining set of $k_i - 1$ trial directions that have not been selected while building up the chain conformation. Z_N , Z_{N+1} , and $V Z^{\text{id}}$ are the partition functions of the fluid, the fluid plus the test chain, and the ideal chain molecule respectively. The partition function of the ideal molecule considered here is simply the product of the contributions from the individual segments. From equation (7) it follows that

$$\beta \mu^{\text{ex}} = \beta \mu - \beta \mu^{\text{id}} = -\ln \frac{Z_{N+1}}{Z_N} + \ln \frac{V Z^{\text{id}}}{\Lambda^3 (N+1)} = -\ln \langle W_\Gamma \rangle_N \quad (8)$$

which means that the test-particle method to calculate excess chemical potentials can be combined with Rosenbluth sampling for chain molecules, simply by replacing the average Boltzmann weight of the test chain by its average Rosenbluth weight.

2.3. Overlapping-distribution Rosenbluth method

Equation (8) can be applied to calculate chemical potentials of much longer chain molecules than the original test-particle method represented by equation (6) can handle. However, for high densities and for chains that are too long, it shares the weakness of the original method, namely that it may fail to sample regions of configurational space that have a high weight, in this case a high Rosenbluth weight. In fact, as is discussed in detail in the paper by Batoulis and Kremer [10], the original Rosenbluth scheme suffers from exactly the same drawback. To be able to detect this failure and to judge the reliability of the results, we wish to combine Rosenbluth sampling with the overlapping-distribution method. This involves inserting a chain in an N -particle system as well as removing a chain from an equilibrium system in which the chain is dissolved. Simply removing an existing chain only yields the Boltzmann factor and not the Rosenbluth factor of the configuration. But, in order to be able to judge the reliability, it is necessary to check whether there is enough overlap between the distribution of Rosenbluth weights of chains that are inserted, and those that are to be removed. In order to calculate the Rosenbluth weight of an existing chain, the same

simple procedure can be applied as used in the configurational-bias Monte Carlo method that was also introduced in [9]. In this scheme, a Monte Carlo trial move is generated by constructing a new configuration using Rosenbluth sampling. In order to decide whether to accept the trial move, the Rosenbluth weight of the old chain needs to be calculated. To do so, around every segment Γ_i of the existing chain, a set of $k - 1$ trial directions, $\{\Gamma_{ij}\}_{\text{rest}}$, is drawn from the same probability distribution P^{id} , from which the directions are chosen when a chain is inserted. The Rosenbluth weight of the existing chain is calculated, by treating the $k - 1$ trial directions *plus the direction in which the segment of the existing chain is situated*, as the set of k 'trial' directions, $\{\Gamma_{ij}\}_{\text{trial}}$, for the existing conformation. The distribution of Rosenbluth weights found for removing chains is directly related to the distribution found for inserting chains. More precisely, if we denote the logarithm of the Rosenbluth weight of a configuration by $\ln W \equiv x$, then we can measure the probability distributions $p_0(x)$ for chain insertions and $p_1(x)$ for chain removal. In the appendix we derive the relation between $p_0(x)$ and $p_1(x)$ explicitly. We show that from this relation the excess chemical potential can be obtained, because

$$\ln p_1(x) = x + \beta\mu^{\text{ex}} + \ln p_0(x). \quad (9)$$

Hence, by constructing a histogram of $\ln W$ both for inserting a chain into a system and for removing a chain dissolved in that system, we can derive the excess chemical potential by studying $\ln p_1(x) - \ln p_0(x)$. As in the original overlapping-distribution method of Bennett [4] and Shing and Gubbins [3], the method works best if there is a range of x values where we have good statistics on both $p_0(x)$ and $p_1(x)$. The advantage of this overlapping-distribution Rosenbluth scheme over the simple test-particle Rosenbluth method is that, with the present method, sampling problems for long chains and high densities will manifest themselves as a breakdown of the overlap of p_0 and p_1 .

Recently, de Pablo *et al* [12] proposed an alternative extension of the Shing-Gubbins method to chain molecules. De Pablo *et al* do not consider a histogram in $\ln W$ but a histogram in the energies, U_Γ , of the chain configurations Γ . In the appendix we derive the relation that de Pablo *et al* [12] used. In their approach, the average Rosenbluth weight found for inserted chains that have an energy $U_\Gamma = \Delta U$, $\langle W \rangle_{\Delta U, N}$, is measured. This quantity is related to the probability, $p_1(\Delta U)$, of finding a chain dissolved in the system with an energy ΔU :

$$\ln p_1(\Delta U) = \beta\mu^{\text{ex}} + \ln \langle W \rangle_{\Delta U, N}. \quad (10)$$

Equation (10) is, in principle, exact, as is equation (9). However, numerically the two expressions are rather different. To see this, consider a chain of hard-sphere segments. For long chains, the histograms used in equation (9) will no longer overlap, indicating that the results of the simulation are unreliable. In contrast, for hard-core particles equation (10) reduces to equation (8): the histograms of ΔU in equation (10) will *always* overlap at $\Delta U = 0$, as long as it is possible to insert a hard-sphere chain in the fluid. Hence the scheme of de Pablo *et al* lacks the additional diagnostic feature that characterizes the overlapping-distribution Rosenbluth method given by equation (9).

3. Results

As an example we applied both the test-particle Rosenbluth technique and the overlapping-distribution Rosenbluth method to calculate the chemical potential of fully flexible chains of ℓ hard spheres in a hard-sphere fluid at number density $\rho\sigma^3$, where σ is the hard-sphere diameter. Chains are inserted into the fluid by means of Rosenbluth sampling. The numbers

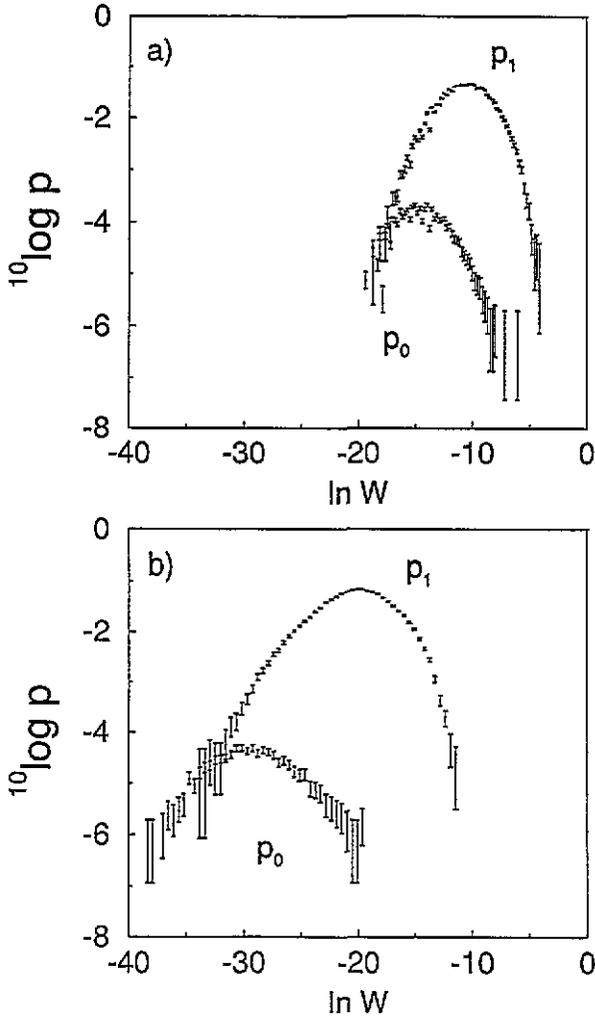


Figure 1. The distributions $p_0(\ln W)$ and $p_1(\ln W)$ for fully flexible hard-sphere chains of length (a) $\ell = 8$ and (b) $\ell = 14$ in a hard-sphere fluid at density $\rho\sigma^3 = 0.4$.

Table 1. The number of trial directions, k_i , from which the position Γ_i of the i th segment is selected. These numbers are chosen such that the efficiency of the chemical-potential calculations is optimized as explained in [2].

| i | $\rho^* = 0.4$ | $\rho^* = 0.5$ | i | $\rho^* = 0.4$ | $\rho^* = 0.5$ |
|-----|----------------|----------------|-----|----------------|----------------|
| 2 | 9 | 25 | 9 | 29 | 69 |
| 3 | 12 | 31 | 10 | 32 | 79 |
| 4 | 15 | 37 | 11 | 35 | 84 |
| 5 | 17 | 43 | 12 | 38 | 90 |
| 6 | 20 | 48 | 13 | 41 | 95 |
| 7 | 23 | 57 | 14 | 44 | 100 |
| 8 | 26 | 64 | | | |

of trial directions, k_i , from which the position Γ_i of the i th segment is taken, are listed in

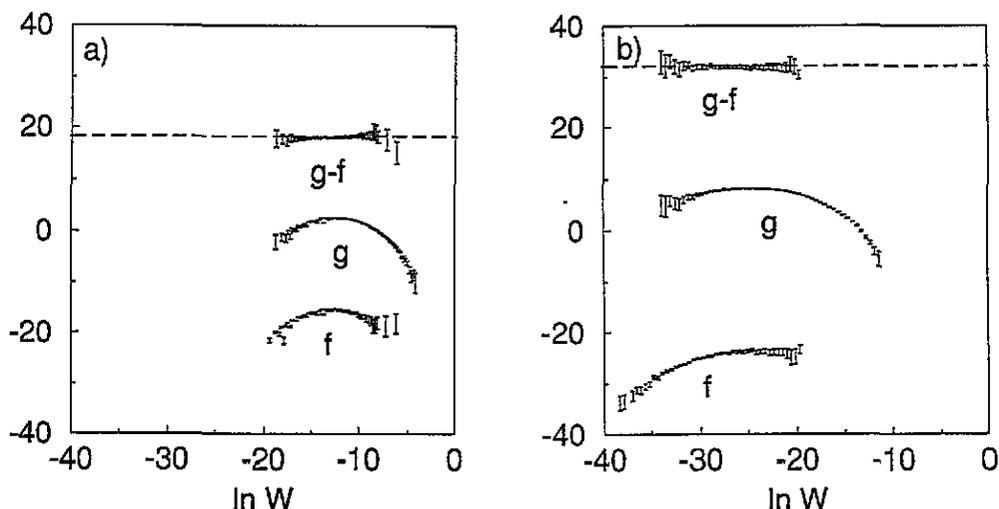


Figure 2. The functions $f(\ln W) = p_0(\ln W) + \frac{1}{2} \ln W$, $g(\ln W) = p_1(\ln W) - \frac{1}{2} \ln W$, and $g(\ln W) - f(\ln W)$ for fully flexible chains of hard spheres of length (a) $\ell = 8$ and (b) $\ell = 14$ in a hard-sphere fluid at density $\rho\sigma^3 = 0.4$. The dashed lines denote the value for $\beta\mu^{\text{ex}}$ that is calculated using the test-particle Rosenbluth method.

table 1 and are chosen such that the efficiency of the chemical-potential calculations is optimized. In a separate publication [2] we show how to determine the optimal k_i values. We used the test-particle Rosenbluth method (equation (8)) to calculate the excess chemical potential from the average of the Rosenbluth weights of the inserted chains. We have also calculated the excess chemical potential from equation (9), the overlapping-distribution Rosenbluth method. To this end, a histogram of the Rosenbluth weights of the inserted chains, $p_0(\ln W)$, was constructed for a range of values of $\ln W$. In addition, the fluid plus the chain was simulated and the Rosenbluth weight of the existing chain was calculated with the procedure described in the previous section. These Rosenbluth weights were also tabulated, in $p_1(\ln W)$. In figure 1 we show typical examples of the two distributions $p_0(\ln W)$ and $p_1(\ln W)$ for chains of lengths $\ell = 8$ and $\ell = 14$, in a hard-sphere fluid at density $\rho\sigma^3 = 0.4$. It can be seen that the amount of overlap between the two peaks decreases with chain length. For both chain lengths, there is sufficient overlap between the two peaks to be able to use equation (9) to calculate the chemical potential accurately. For longer chains the two peaks will shift further apart, and beyond a certain length the results will no longer be reliable. But the Rosenbluth sampling can break down for another reason, depicted by the fact that the distribution $p_0(\ln W)$ not only shifts to lower values for $\ln W$, but it also decreases. As $p(x)$ is normalized, this can only happen because a finite contribution to the integral of $p(x)$ comes from the value $x = -\infty$ (i.e. $W = 0$). Conformations with $W = 0$ are simply failed attempts to generate a trial conformation. The longer the chains, the larger the fraction of trial conformations that fails before the full length of the molecule is reached. Clearly, if insufficient trial conformations are generated, the statistical accuracy of the method breaks down. For the systems and length of simulation that we have studied here, the latter breakdown of both the test-particle Rosenbluth method and the overlapping-distribution Rosenbluth method occurs *before* the systematic sampling error plays a role. Finally, to calculate the excess chemical potential from equation (9), we follow the procedure that is used in the original overlapping-distribution method. Two functions, $f(\ln W) = p_0(\ln W) + \frac{1}{2} \ln W$ and $g(\ln W) = p_1(\ln W) - \frac{1}{2} \ln W$, are calculated and at

every $\ln W$ value their difference provides an estimate for the excess chemical potential, $\beta\mu^{\text{ex}}$. In figure 2 we show both the functions $f(\ln W)$ and $g(\ln W)$ and their difference $g(\ln W) - f(\ln W)$ for the same examples as in figure 1. In the same figure we show a line at the values corresponding to the values of $\beta\mu^{\text{ex}}$ that are calculated from the test-particle Rosenbluth method. The two estimates are in excellent agreement with each other, as they should be. We calculated chemical potentials for other chain lengths and densities, with both methods. In figure 3 we plot these results and we find that the chemical potential is linear with chain length, at least for the range of values that we looked at.

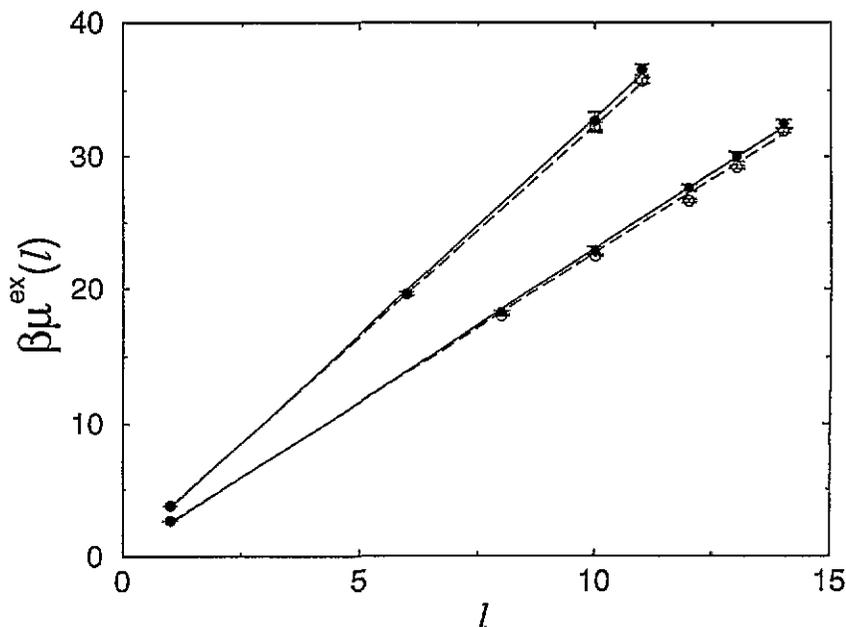


Figure 3. The chemical potential for a fully flexible chain of l hard spheres in a fluid at density $\rho\sigma^3 = 0.4$ and $\rho\sigma^3 = 0.5$, calculated both with the test-particle Rosenbluth method (\circ) and with the overlapping-distribution Rosenbluth method (\bullet). The lines are linear fits through the data.

4. Conclusions

We have shown how to combine the overlapping-distribution method for calculating chemical potentials with the Rosenbluth sampling technique for continuously deformable chain molecules. The advantage of this technique over the test-particle Rosenbluth method [9, 11] is that it is possible to detect systematic sampling errors that can occur when the sampling is restricted to regions of configurational space with a low Rosenbluth weight. In the example of fully flexible chains of hard spheres in a hard-sphere fluid, we did not find a systematic error. With increasing chain lengths a shift to lower Rosenbluth weights of the test-particle insertions does occur, but, even for the longest chains that could be inserted during the length of the simulations, this shift still did not cause a systematic error in the calculation of the chemical potential. On the contrary, the results of the test-particle and the overlapping-distribution Rosenbluth methods are in excellent agreement with each other. For longer chains a breakdown of both methods does occur, but this is primarily due to the fact that the insertion probability for long chains becomes too low and

the statistical error becomes too large to measure the chemical potential in a simulation.

In the present paper we only considered simple model polymers. However, the techniques that we discuss can, without loss of generality, be applied to realistic models of chain molecules. On the basis of other simulation results (see e.g. [13]), we know that Rosenbluth schemes for particle insertion can be applied to realistic models of alkanes as long as $C_{48}H_{98}$. The range of applicability of the overlapping-distribution method discussed above is not expected to exceed significantly that of the 'Widom-Rosenbluth' particle-insertion method. However, as emphasized above, overlapping distribution schemes have built-in diagnostics and are therefore safer to use.

Acknowledgments

The investigations reported in this paper were supported in part by Scheikundig Onderzoek Nederland (SON) with financial aid from NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek). The work of the FOM Institute is part of the research program of FOM and is supported by NWO.

Appendix

Here we derive the result presented in equation (9) on which the Rosenbluth-sampling overlapping-distribution method is based.

Chain conformations Γ of length ℓ are generated following the Rosenbluth sampling procedure described in section 2.1. First, a segment is inserted at a random position Γ_0 in the volume V . Then the orientations of the subsequent segments, Γ_i ($1 \leq i \leq \ell$), are selected from sets of k_i trial directions $\{\Gamma_{ij}\}_{\text{trial}}$ with a probability P proportional to the Boltzmann weight. These sets of trial directions are in turn chosen at random with a probability P^{id} (equation (3)). The Rosenbluth weight of a conformation is given by equation (6). The probability, $p_0(x)$, of generating a chain conformation in a fluid of N other chains, with a Rosenbluth weight W_Γ such that $\ln W_\Gamma = x$, is given by:

$$\begin{aligned}
 p_0(x) &= \frac{1}{V} \left\langle \int d\Gamma_0 \int \prod_{i=1}^{\ell} d\{\Gamma_{ij}\}_{\text{trial}} P^{\text{id}}(\{\Gamma_{ij}\}_{\text{trial}}) \sum_{j=1}^{k_i} P(\Gamma_{ij}) \delta(x - \ln W_\Gamma) \right\rangle_N \\
 &= \frac{1}{V} \left\langle \int d\Gamma_0 \int \prod_{i=1}^{\ell} d\{\Gamma_{ij}\}_{\text{rest}} P^{\text{id}}(\{\Gamma_{ij}\}_{\text{rest}}) d\Gamma_i P^{\text{id}}(\Gamma_i) k_i P(\Gamma_i) \delta(x - \ln W_\Gamma) \right\rangle_N \\
 &= \frac{1}{V Z^{\text{id}}} \left\langle \int d\Gamma e^{-\beta U_\Gamma} \int \prod_{i=1}^{\ell} d\{\Gamma_{ij}\}_{\text{rest}} P^{\text{id}}(\{\Gamma_{ij}\}_{\text{rest}}) \frac{1}{W_\Gamma} \delta(x - \ln W_\Gamma) \right\rangle_N \\
 &= \frac{1}{V Z^{\text{id}}} \left\langle \int d\Gamma e^{-\beta U_\Gamma} \int \prod_{i=1}^{\ell} d\{\Gamma_{ij}\}_{\text{rest}} P^{\text{id}}(\{\Gamma_{ij}\}_{\text{rest}}) \delta(x - \ln W_\Gamma) \right\rangle_N e^{-x}
 \end{aligned}
 \tag{A1}$$

where the symbols have the same meaning as in section 2.1. The term in the angular brackets on the last line of equation (A1) can be related to $p_1(x)$, the probability of choosing a chain Γ at random from $N + 1$ chains and of generating $k_i - 1$ trial directions $\{\Gamma_{ij}\}_{\text{rest}}$ around every segment from which a Rosenbluth weight W_Γ for the chain is obtained such that

In $W_\Gamma = x$. Namely, $p_1(x)$ is given by

$$p_1(x) = \left\langle \int \prod_{i=1}^{\ell} d\{\Gamma_{ij}\}_{\text{rest}} P^{\text{id}}(\{\Gamma_{ij}\}_{\text{rest}}) \delta(x - \ln W_\Gamma) \right\rangle_{N+1}. \quad (\text{A2})$$

If we write out the ensemble averages, denoted by the angular brackets, we obtain the desired result

$$\ln p_1(x) = x + \ln \frac{V Z^{\text{id}} Z_N}{\Lambda^3(N+1) Z_{N+1}} + \ln p_0(x) = x + \beta \mu^{\text{ex}} + \ln p_0(x) \quad (\text{A3})$$

where Z_N and Z_{N+1} are the partition functions for a system of N and $N+1$ particles respectively. Equation (A3) is the result we use in section 2.1.

In a similar manner, the method proposed by de Pablo *et al* [12], based on equation (10), can be derived. Instead of histograms in the Rosenbluth weights, histograms in energy are calculated. Upon inserting chains in a fluid of N chains, the average Rosenbluth factor of chains with a specific energy ΔU is calculated. This average, $\langle W_\Gamma \rangle_{N, \Delta U}$, is given by

$$\begin{aligned} \langle W_\Gamma \rangle_{N, \Delta U} &= \frac{1}{V} \left\langle \int d\Gamma_0 \int \prod_{i=1}^{\ell} d\{\Gamma_{ij}\}_{\text{trial}} \prod_{i=1}^{\ell} P^{\text{id}}(\{\Gamma_{ij}\}_{\text{trial}}) \sum_{j=1}^{k_i} P(\Gamma_{ij}) W_\Gamma \delta(\Delta U - U_\Gamma) \right\rangle_N \\ &= \frac{1}{V} \left\langle \int d\Gamma_0 \int \prod_{i=1}^{\ell} \left[d\{\Gamma_{ij}\}_{\text{rest}} \prod_{i=1}^{\ell} P^{\text{id}}(\{\Gamma_{ij}\}_{\text{rest}}) \right] \right. \\ &\quad \left. \times \int \prod_{i=1}^{\ell} d\Gamma_i P^{\text{id}}(\Gamma_i) k_i P(\Gamma_i) W_\Gamma \delta(\Delta U - U_\Gamma) \right\rangle_N \\ &= \frac{1}{V Z^{\text{id}}} \left\langle \int d\Gamma e^{-\beta U_\Gamma} \delta(\Delta U - U_\Gamma) \right\rangle_N. \end{aligned} \quad (\text{A4})$$

This average is related to the probability $p_1(\Delta U)$ of choosing a chain Γ at random from a fluid of $N+1$ chains, with an energy ΔU :

$$p_1(\Delta U) = \langle \delta(\Delta U - U_\Gamma) \rangle_{N+1}. \quad (\text{A5})$$

Again writing out the ensemble averages gives the relation

$$\ln p_1(\Delta U) = \ln \frac{V Z^{\text{id}} Z_N}{\Lambda^3(N+1) Z_{N+1}} + \ln \langle W \rangle_{\Delta U, N} = \beta \mu^{\text{ex}} + \ln \langle W \rangle_{\Delta U, N}. \quad (\text{A6})$$

This result is used in section 2.1.

References

- [1] Widom B 1963 *J. Chem. Phys.* **39** 2802
- [2] Mooij G C A M 1993 *Thesis* University of Utrecht
Mooij G C A M and Frenkel D 1994 to be published
- [3] Shing K S and Gubbins K E 1982 *Mol. Phys.* **46** 1109; 1983 *Mol. Phys.* **49** 1121
- [4] Bennett C H 1976 *J. Comput. Phys.* **22** 245
- [5] Allen M P and Tildesley D J 1987 *Computer Simulation of Liquids* (Oxford: Clarendon)
- [6] Rosenbluth M N and Rosenbluth A W 1955 *J. Chem. Phys.* **23** 356
- [7] Harris J and Rice S A 1988 *J. Phys. Chem.* **88** 1298
- [8] Siepmann J I and Frenkel D 1992 *Mol. Phys.* **75** 59
- [9] Frenkel D, Mooij G C A M and Smit B 1992 *J. Phys.: Condens. Matter* **4** 3053
- [10] Batoulis J and Kremer K 1988 *J. Phys. A: Math. Gen.* **21** 127
- [11] Frenkel D and Smit B 1992 *Mol. Phys.* **75** 983
- [12] de Pablo J J, Laso M and Suter U W 1992 *J. Chem. Phys.* **96** 6157
- [13] Siepmann J I, Karaborni S and Smit B 1993 *Nature* **365** 330