

How to optimize Configurational Bias Monte Carlo?

Daan Frenkel and Germonda Mooij
FOM Institute for Atomic and Molecular Physics
Kruislaan 407
1098 SJ Amsterdam
The Netherlands

1 Introduction

Configurational-Bias Monte Carlo (CBMC) is a dynamic MC scheme that makes it possible to achieve large conformational changes in a single trial move that affects a large number of monomeric units of a chain molecule [1, 2, 3, 4]. The CBMC method is based on the Rosenbluth sampling scheme [5, 1, 2] for lattice systems. In this scheme, the molecular conformation is built up step-by-step, in such a way that, at every stage, the next monomeric unit is preferentially added in a direction that has a large Boltzmann weight. This increases the probability of generating a trial conformation that has no hard-core overlaps. As explained below, the probability of acceptance of the trial conformation is given by the ratio of the ‘Rosenbluth weights’ of the new and the old conformations. Whereas the original Rosenbluth scheme was devised for polymers on a lattice, the CBMC scheme will also work for chain molecules in continuous space. Unlike the reptation algorithm [6], CBMC can be used in the simulation of grafted chains and ring polymers. Recently, the CBMC method has been integrated with the Gibbs-ensemble technique to simulate liquid-vapour and fluid-fluid phase equilibria of chain molecules [7]. In Gibbs-ensemble simulations of phase coexistence, simulations of the two coexisting phases (e.g. liquid and vapor) are carried out in parallel. In addition to MC trial moves of the molecules within either system, we also allow the two systems to exchange volume and mass. CBMC trial moves are used to swap chain molecules between the two systems. Clearly this requires complete regrowth of the entire chain. For long chains this becomes expensive and, at present, Gibbs-ensemble simulations are limited to chain molecules with less than 50 carbon atoms [8]. For simple CBMC sampling the situation is less serious, because one can choose not to regrow the entire chain but only part thereof. In the limit that only one monomeric unit is regrown, CBMC reduces to the reptation algorithm, but in general it will be advantageous to regrow a larger number of monomeric units. Of course, the computational cost per trial move is higher for CBMC than for reptation and hence it becomes important to be able to construct the most efficient MC move for a given system.

The efficiency of the Rosenbluth sampling technique depends on the choice of a set of parameters, namely the number of trial insertions for a given segment i , k_i . As described in the next section, k_i can, in principle, be chosen freely. However, the choice of k_i affects the efficiency of the sampling scheme. The aim of the present contribution is to show that the choice of the parameters in a CBMC simulation is not a question of black magic. Rather, as we shall show, there are systematic techniques to optimize the algorithm. In particular, we show how the efficiency of a CBMC program can be optimized with respect to k_i . Although we apply our analysis to the CBMC scheme, it is in fact much more general, and can be used to optimize the efficiency of any MC trial move that can be decomposed into a sequence of elementary steps. Before discussing the optimization of CBMC sampling, we briefly review the basic idea behind the method.

2 Configurational-Bias Monte Carlo

2.1 Rosenbluth sampling

The Configurational-Bias Monte Carlo scheme for continuously deformable chain molecules [3], is based on Rosenbluth sampling [5, 1, 2] for lattice systems. Chain configurations are generated by successive insertion of the bonded segments of the chain. When the positions of the segments are chosen at random, it is very likely, that one of the segments will overlap with another particle in the fluid, which results in 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 can be tested. In continuous space, there are in principle an infinite number of positions that should be tested (e.g. in the case of a 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 of bond-lengths and bond-angles of the ‘ideal’ chain molecule. From here on, the procedure 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 (more precisely, the contributions of all those interactions that have not yet been accounted for in the generation of the trial positions). One of these trial positions is then selected with a probability proportional to its Boltzmann factor. 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 [5]. 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:

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

where $\beta = 1/k_B T$ and $U_{\Gamma_{ij}}^{\text{nb}}$ 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} :

$$W_\Gamma = e^{-\beta U_{\Gamma_0}} \prod_{i=1}^{\ell} W_{\Gamma_i}, \quad (2)$$

where ℓ is the chain length. In the original Rosenbluth scheme, every chain conformation Γ was given a statistical weight proportional to W_Γ . However, as explained in ref. [9], this approach fails when the largest contribution to the equilibrium properties of a chain molecule come from

conformations that have a large Rosenbluth weight W , but a very small probability $P(W)$ of being generated in the Rosenbluth sampling scheme. The Configurational-Bias MC scheme that we discuss below, was designed to avoid this problem.

2.2 CBMC: ‘Dynamic’ Rosenbluth sampling

The Configurational-Bias Monte Carlo method is a sampling scheme that employs the Rosenbluth method (extended to continuously deformable molecules [3]) to generate trial conformations. However, it does not suffer from the sampling problem of the original Rosenbluth scheme, because all chain conformations are generated with the correct statistical weight: hence, all averages obtained with the CBMC method are *unweighted* averages over MC configurations, and the problems that are associated with the re-weighting in the original Rosenbluth scheme, disappear.

The CBMC procedure for generating a new conformation of a chain is as follows. First, a chain is chosen at random. Next, a trial conformation for this chain is generated by means of the Rosenbluth sampling scheme and a Rosenbluth weight for this new conformation is calculated. Next, we should decide if we accept the proposed ‘move’. To this end, we must compare the Rosenbluth weight W_{new} of the trial conformation with W_{old} , the weight of the old conformation. In fact, the computation of the latter quantity is a bit subtle. In case of a lattice system it is obvious what the trial directions for the old conformation are, and hence its Rosenbluth weight can be evaluated unambiguously. In contrast, for continuously deformable chains the trial directions are chosen at random for every new conformation, and it is not immediately obvious what choice should be made for the calculation of the Rosenbluth weight of the old conformation. As shown in ref. [3], it can be proven that the following simple procedure satisfies detailed balance, and thereby fulfills a sufficient condition to ensure that all chain conformations are generated with a probability proportional to their Boltzmann weight: around every segment i of the old chain, $k_i - 1$ trial directions are drawn from the same probability distribution as the one from which the directions for the trial conformation are chosen. The old Rosenbluth weight is calculated, by treating the $k_i - 1$ trial directions *plus the direction in which the segment of the old chain is situated*, as the set of ‘trial’ directions for the existing conformation. Finally, we compute the ratio of the Rosenbluth weights of the new and the old conformations. We use a Metropolis-like criterion to decide on the acceptance of the trial move, i.e. the trial move is accepted with a probability P_{acc} ,

$$P_{\text{acc}} = \text{Min}\left(1, \frac{W_{\text{new}}}{W_{\text{old}}}\right). \tag{3}$$

The procedure sketched above is valid for a complete regrowth of the chain, but it is also possible to regrow only part of a chain, i.e. to cut a chain at a (randomly chosen) point and regrow the cut part of the chain either at the same site or at the other end of the molecule. Clearly, if only one segment is regrown and only one trial direction is used, CBMC reduces to the reptation algorithm (at least, for linear homo-polymers). It should be stressed that there are many possible ways to

generate a trial conformation. For instance, one can generalize the ‘pivot’ algorithm [10]. In the pivot algorithm a new conformation is generated by rotating a molecule over a random angle around a randomly selected ‘pivot’ segment. The pivot algorithm is very efficient for isolated chains, but becomes inefficient for molecules in dense media. However, with CBMC, one can introduce a larger number of pivots in a chain molecule, in such a way that the acceptance of the trial moves is enhanced (at the expense of additional computation). Of course, when CBMC is combined with Grand Canonical and Gibbs-ensemble MC simulations, where entire molecules are exchanged, it is necessary to include moves that attempt to (re)grow chains completely.

One choice remains to be made before applying the Rosenbluth sampling scheme for continuously deformable chain molecules to CBMC and chemical potential calculations, namely the choice for the number of trial directions k_i . Too many trial directions increase the cost of a simulation cycle, but too few trial directions lower the acceptance rate, and increase the simulation length. Clearly, we wish to have simple guidelines that allow us to select k_i for every segment such that it optimizes the efficiency of the simulation. In the following section we show how the optimal values for the set $\{k_i\}$ and the maximum efficiency achievable can be estimated.

3 Efficiency of Configurational-Bias Monte Carlo

In order for the Rosenbluth sampling scheme to work, it is essential to generate, on average, at least one trial position that has a non-negligible Boltzmann weight for every segment. If all trial positions have a small Boltzmann weight, the Rosenbluth weight of the new conformation is virtually zero, while the Rosenbluth weight of the existing conformation is necessarily finite, and the trial move will be rejected. The probability of finding at least one trial position with a non-negligible Boltzmann weight, depends on the choice for the value of k_i , i.e. the number of trial directions that are scanned when looking for an acceptable position of the next, i^{th} , segment. In discussing the efficiency of the CBMC scheme, it is convenient to consider monomeric units with a hard repulsive core because in that case the Boltzmann weight associated with conformations that have hard-core overlaps is strictly zero. However, the general conclusions carry over to systems with continuous intermolecular interaction potentials.

Two trends determine this choice for optimal k_i -values, k_i^{opt} . On the one hand, the probability of a successful chain insertion grows with increasing k_i . There is an upper limit to that, because when the space to insert another segment is simply not available, there is no point in generating more and more trial directions. Moreover, the computational cost also rises with increasing k_i . The optimal choice for k_i depends on density, temperature and the nature of the intermolecular interactions. For instance, at high densities a larger number of trial directions is needed to regrow a given number of segments than at low densities. It can also be expected that k_i^{opt} varies along a chain. After successful insertion of part of the chain, a larger number of trial directions should be chosen for the next segment, in order to minimize the probability that we waste the computational

effort that has already been invested in this trial move.

Below, we show how we can arrive at an estimate of the optimal values k_i^{opt} . To do so, we should first define what we mean by the ‘efficiency’ of a given CBMC trial move. Loosely speaking, we expect the efficiency to be proportional to the probability that a given trial conformation is successfully generated and inversely proportional to the computational cost of that trial move. For a chain of ℓ segments

$$\text{Eff}(\ell) = \frac{\langle P(\ell) \rangle}{\langle \text{Cost}(\ell) \rangle}, \quad (4)$$

where $\langle P(\ell) \rangle$ is the probability to find for every segment at least one trial direction with a non-negligible Boltzmann weight, in which case the chain can be inserted successfully. $\langle \text{Cost}(\ell) \rangle$ is the average cost for trying to insert the chain, measured in the number of times the energy of a trial direction is calculated. The extra cost for trying to insert a chain which is one segment longer, depends linearly on the number of trial directions and on the probability to insert ℓ segments successfully. So, the average cost for one trial insertion of a chain of length $\ell + 1$ is given by

$$\langle \text{Cost}(\ell + 1) \rangle = \langle \text{Cost}(\ell) \rangle + 2k_{\ell+1} \times \langle P(\ell) \rangle. \quad (5)$$

where we have introduced, as our unit of computational cost, the amount of computation needed to compute the energy for one trial segment. In the computational cost of a trial move in the CBMC scheme, we have included the cost of the energy calculations for the $k_{\ell+1}$ ‘trial’ directions of the old conformation, needed to compute the ‘old’ Rosenbluth weight W_{old} . The probability to find at least one acceptable position for the $\ell + 1^{\text{th}}$ segment, $\langle P_{\text{add}}(k_{\ell+1}) \rangle$, also increases with $k_{\ell+1}$. If we assume that subsequent insertions of segments are independent, $\langle P(\ell + 1) \rangle$ is given by

$$\langle P(\ell + 1) \rangle = \langle P(\ell) \rangle \times \langle P_{\text{add}}(k_{\ell+1}) \rangle. \quad (6)$$

Equations 5 and 6 can be combined with equation 4 to yield the following very simple recursive relation

$$\frac{\text{Eff}(\ell + 1)}{\text{Eff}(\ell)} = \frac{\langle P_{\text{add}}(k_{\ell+1}) \rangle}{1 + 2k_{\ell+1} \times \text{Eff}(\ell)}. \quad (7)$$

Together with the ‘boundary’ condition $\text{Eff}(\ell = 1)$, equation 7 allows us to compute the efficiency of a trial move for a given set of k_i -values. The values of the set $\{k_i\}$ affect both the numerator and the denominator of equation 7. Our aim is to vary all k_i -values until the optimum efficiency is reached.

The computational cost of the insertion of the first monomer of the chain is zero if we simply start regrowing part of an existing chain. However, if we must successfully insert one monomer before we can continue growing the rest of the chain, then the computational cost of the first insertion is non-negligible and this, in turn, will affect (increase) the optimal values for all subsequent

k_i 's. In addition to $\text{Eff}(1)$, we must know $\langle P_{\text{add}}(k_{\ell+1}) \rangle$ for all ℓ . $\langle P_{\text{add}}(k_{\ell+1}) \rangle$ can be determined numerically by calculating

$$\langle P_{\text{add}}(k_{\ell+1}) \rangle = 1 - \langle (1 - P_{\text{add}}(1))^{k_{\ell+1}} \rangle. \quad (8)$$

In words: the probability to generate at least one acceptable trial segment is equal to one minus the probability that not a single acceptable trial segment is generated in $k_{\ell+1}$ attempts. In equation 8, $P_{\text{add}}(1)$ is the probability that the insertion of a single trial segment will be successful. It should be noted that this probability is a fluctuating quantity: the angular brackets in equation 8 denote averaging over the equilibrium configurations of the fluid. Of course, we can make a crude estimate of $\langle P_{\text{add}}(k_{\ell+1}) \rangle$ by ignoring all fluctuations, in which case we get the ‘mean-field’ estimate

$$\langle P_{\text{add}}(k_{\ell+1}) \rangle = 1 - (1 - \langle P_{\text{add}}(1) \rangle)^{k_{\ell+1}}. \quad (9)$$

Although equation 9 is useful for order-of-magnitude estimates, we shall not use it in what follows. Rather, we shall compute $\langle P_{\text{add}}(k_{\ell+1}) \rangle$ by simulation. In stead of computing $\langle P_{\text{add}}(k_{\ell+1}) \rangle$ for all ℓ , we measured it for $\ell \leq 2$, and assume that for $\ell > 2$, the values for $\ell = 2$ can be used as an estimate. We verified this assumption under various conditions by calculating $\langle P_{\text{add}}(k_{\ell+1}) \rangle$ for all ℓ and we found no significant difference in the answers.

The procedure described above allows us to determine numerically the values for the set $\{k_i\}$ that maximize equation 7, and thereby the efficiency to generate an acceptable trial conformation for a chain in a CBMC move. Thus far we have ignored the fact that this trial conformation,

ℓ	0.3	0.4	0.5
2	5	10	29
3	9	18	55
4	12	27	86
5	15	35	> 100
6	18	43	> 100
7	20	51	> 100
8	22	59	> 100
9	25	66	> 100
10	27	73	> 100
11	29	80	> 100
12	30	87	> 100

Table 1: Optimal k_ℓ -values, for insertion of a chain by Rosenbluth sampling in the Configurational-Bias Monte Carlo scheme.

although acceptable in principle, may be rejected in practice. As stated before (equation 3), the

overall acceptance probability is determined by the ratio of the new and the old Rosenbluth weights: $W_{\text{new}}/W_{\text{old}}$. Below we shall show that this further attrition of trial conformation also affects the efficiency of the CBMC scheme.

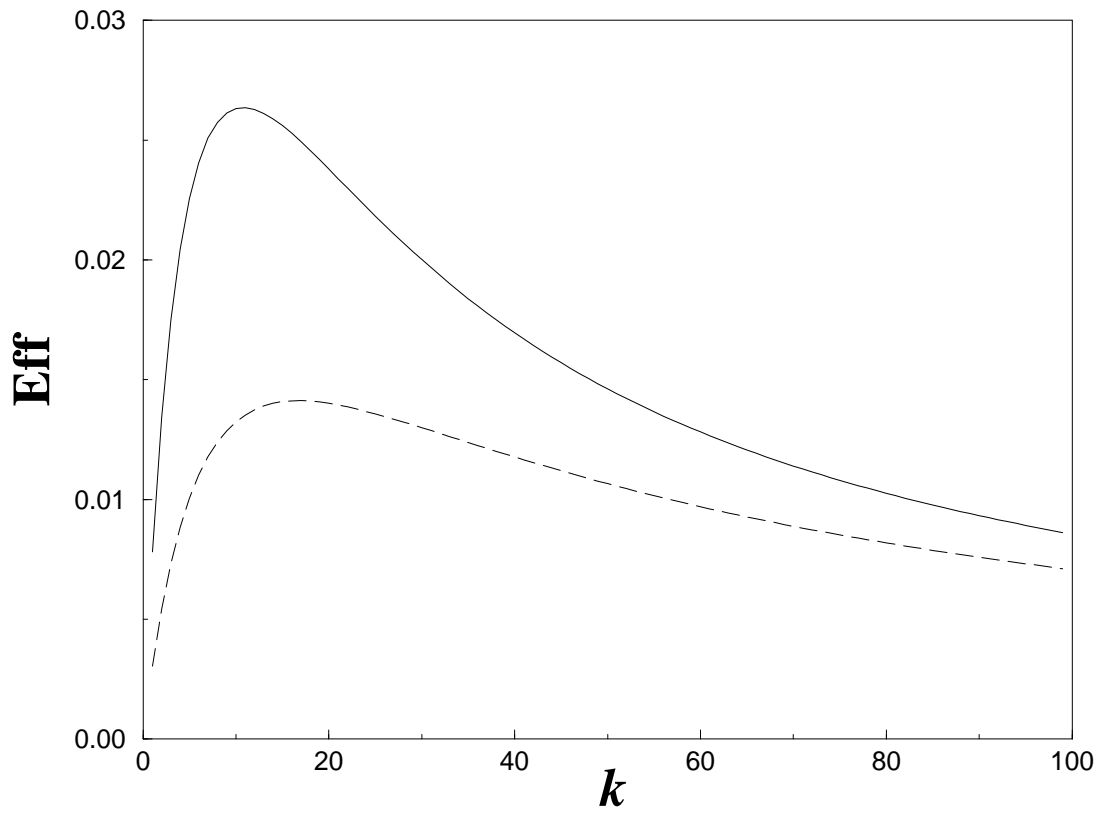


Figure 1: The efficiency, as defined by equation 7, for inserting a hard dimer (—) and a fully flexible trimer of hard spheres (---) into a fluid of hard spheres at several densities $\rho\sigma^3$, over a range of k -values.

4 Results

As an example we studied a system with only hard core interactions, but it should be noted that the efficiency analysis that we have presented above can be applied with minor modifications to systems with soft potentials [11].

Let us consider the case where, in a fluid of hard spheres with diameter σ at number density $\rho\sigma^3$, we insert a fully flexible chain of ℓ hard spheres with the same diameter, attached at a fixed bond length σ . The insertion probability of one segment (which for this particular system is given by the Carnahan-Starling equation [12]) gives $\text{Eff}(1)$ and by inserting a second segment $\langle P_{\text{add}}(k_2) \rangle$ is calculated from equation 8 for a range of k_2 -values. The efficiency for successfully adding another segment, $\text{Eff}(2)$, is calculated from equation 7, and the result is shown in Figure 1 for a fluid at density $\rho\sigma^3 = 0.4$. The maximum determines the value of k_2^{opt} . $\text{Eff}(3)$ for a fluid at the same density is plotted in the same Figure, which shows a shift of the maximum to a value for k_3^{opt} which is higher than k_2^{opt} . As already mentioned, k_i^{opt} is expected to increase with ℓ , because more and more effort is invested previously in the insertion of $\ell - 1$ segments, which will be wasted if all the trial directions result in a hard core overlap with spheres in the fluid.

In Table 1 the optimal k_i -values are listed for insertion of chains up to 12 segments long into a fluid at various densities. For adding a fifth segment or more in a fluid at the highest density, $\rho\sigma^3 = 0.5$, the optimal k_i -values fell out of the range of values that we considered. However, here the efficiency is already close to its optimal value for the highest k_i -values in our range. In Figure 2 we show the corresponding maximal values of $\text{Eff}(\ell)$, and in the same Figure we compare these efficiencies with the efficiencies of a random insertion, i.e. the limit $k_i = 1$ for all ℓ . It shows a considerable increase of efficiency using CBMC, and much longer chain lengths are feasible. In the same figure, we also indicate the effect of the attrition of acceptable trial conformations due to the acceptance criterion (equation 3). The decrease is estimated as $\langle W_{\text{new}}/W_{\text{old}} \rangle$, where W_{new} is now only averaged over chains already inserted successfully (i.e. only the suitable trial conformations are considered). It is possible to give a rough estimate of the maximum chain length that can be reached: if we impose that in a typical MC run we wish to limit the number of evaluations of the potential energy of a segment to a value of 10^8 (i.e. something that can be achieved in a reasonable amount of time on most workstations).

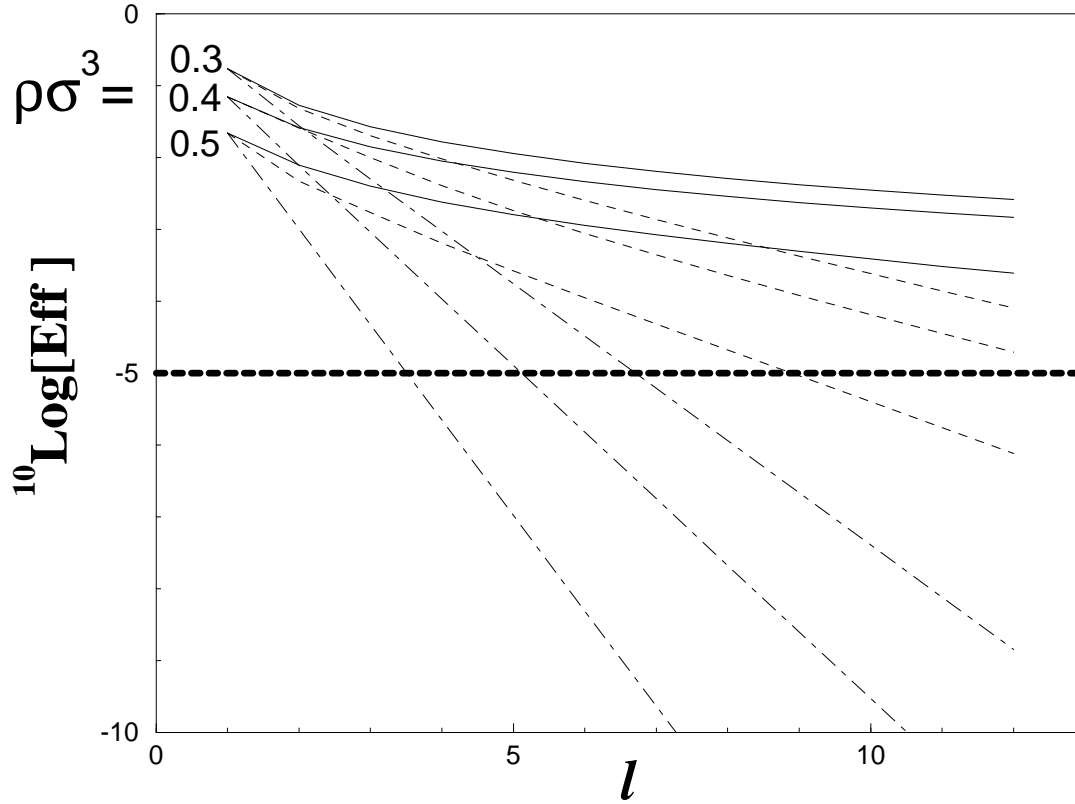


Figure 2: The efficiency (equation 7) for inserting a fully flexible chain of l hard spheres into a fluid of hard spheres at several densities $\rho\sigma^3$. Both the efficiency of a random insertion ($-\cdot-\cdot-$), i.e. $k_l = 1$ for all l , and the maximal efficiency (—), obtained by choosing the optimal k -values, are shown. In the same Figure we show the efficiency for acceptance of a CBMC move ($---$) by the acceptance criterion 3. The dashed horizontal line shows the minimal efficiency needed for a simulation of typical length.

Moreover, we assume that, in order to sample configuration space effectively, we need at least 10^3 successful chain insertions. This implies that the minimal efficiency needed is of the order of 10^{-5} . Figure 2 shows, that random insertion does not fulfil this requirement for chains longer than three segments at $\rho\sigma^3 = 0.5$, five segments at $\rho\sigma^3 = 0.4$ or seven segments at $\rho\sigma^3 = 0.3$. The CBMC scheme can be used at least up to $\ell = 12$ for $\rho\sigma^3=0.3$ and 0.4, and at the higher density $\rho\sigma^3 = 0.5$ it can be used up to $\ell = 9$. Note, however, that at still higher densities, where the probability of successful insertion of a monomer becomes small, the optimal value of ℓ shifts to lower values until, eventually, $\ell=1$. When this happens, CBMC reduces to reptation.

References

- [1] J. Harris and S.A. Rice, *J. Phys. Chem.* **88**, 1298 (1988)
- [2] J.I. Siepmann and D. Frenkel, *Mol. Phys.* **75**, 59 (1992)
- [3] D. Frenkel, G.C.A.M. Mooij, and B. Smit, *J. Phys.: Condens. Matter* **4**, 3053 (1992)
- [4] J. J. de Pablo and M. Laso and U. W. Suter, *J. Chem. Phys.* **96**, 2394 (1992).
- [5] M.N. Rosenbluth and A.W. Rosenbluth, *J. Chem. Phys.* **23**, 356 (1955)
- [6] A.K. Kron, *Polymer Sci. USSR* **7**, 1361 (1965), F.T. Wall and F. Mandel, *J. Chem. Phys.* **63**, 4529 (1975)
- [7] G.C.A.M. Mooij, D. Frenkel, and B. Smit, *J. Phys.: Condens. Matter* **4**, L255 (1992)
- [8] J.I Siepmann, S. Karaborni, and B. Smit, *Nature* **365** 330 (1993)
- [9] J. Batoulis and K. Kremer, *J. Phys. Chem.* **90**, 2370 (1989)
- [10] M. Lal, *Mol. Phys.* **17**, 57 (1969)
- [11] D. Frenkel “Advanced Monte Carlo Techniques” in the Proceedings of the NATO Advanced Study Institute on Computer Simulation in Chemical Systems (M.P. Allen and D.J. Tildesley, eds), Alghero (1992)
- [12] N.F. Carnahan and K.E. Starling, *J. Chem. Phys* **51**, 635 (1969)