Novel scheme to compute chemical potentials of chain molecules on a lattice

By G. C. A. M. MOOIJ and D. FRENKEL

FOM Institute for Atomic and Molecular Physics, PO Box 41883,
1009 DB Amsterdam, The Netherlands

We present a novel method that allows efficient computation of the total number of allowed conformations of a chain molecule in a dense phase. Using this method, it is possible to estimate the chemical potential of such a chain molecule. We have tested the present method in simulations of a two-dimensional monolayer of chain molecules on a lattice (Whittington–Chapman model) and compared it with existing schemes to compute the chemical potential. We find that the present approach is two to three orders of magnitude faster than the most efficient of the existing methods.

1. Introduction

Monte Carlo simulations of the polymer systems tend to be time consuming. The reason is that long chain molecules have a very large number of distinct conformations. Many algorithms to generate a representative sample of these conformations are based on a 'step-by-step' approach to transform one polymer conformation into another. In particular for dense, entangled polymers, the rate at which appreciable conformational changes can be achieved, however, decreases very rapidly with increasing chain length L (e.g., for the well known 'reptation' algorithm [1], the time for the renewal of a conformation is proportional to L^4). Not surprisingly, much attention has been devoted to the development of numerical schemes that allow a more efficient sampling of polymer conformations (for a recent review, see [1]). However, most of these improved schemes are either a smart version of a local algorithm that changes polymer conformations little-by-little or cut-and-paste algorithms that break and remake polymer bonds. The latter class of algorithms cannot be applied to mono-disperse polymer samples.

Even more problematic than the simulation of equilibrium properties of polymer systems, is the computation of the chemical potential of chain molecules. The chemical potential is a quantity that is of central interest for the study of phase equilibria. In computer simulations of simple atomic or molecular systems, the chemical potential is often evaluated with the particle-insertion method first described by Widom [2]. This method relates the chemical potential and the change in the potential energy of the system ΔV_{test}, due to the random insertion of a test particle, as

$$\mu^{\text{ex}} = -k_{\text{B}} T \ln \langle \exp(-\Delta V_{\text{test}}/k_{\text{B}} T) \rangle. \tag{1}$$

The angular brackets in (1) denote an ensemble average over all possible configurations of the molecules already present in the system. In Monte Carlo simulations this average has a very simple interpretation: it can be thought of as probability of acceptance of a (virtual) Monte Carlo move which consist of the random *insertion* of a test particle in the system of interest.

Whether or not (1) can be used in practice depends on the value of this insertion probability. If the latter is very low, the statistical accuracy is low, and long runs are needed to obtain a reliable estimate of μ^{ex}. This limitation of the method is particularly relevant for polymer simulations where a naive application of the particle-insertion method would result in extremely low insertion probabilities. Recently, Siepmann [3] introduced a method, based on the Rosenbluth scheme to generate (self-)avoiding polymer conformations [4], which makes it possible to use the particle-insertion method even for dense polymer systems, where the brute-force insertion method would fail completely. In the present paper we combine Siepmann's approach with a technique that has been developed to average over random walks in lattice-gases [5]. This results in a novel particle-insertion scheme that is very much faster than even the Siepmann method.

Below, we briefly describe our scheme to compute the excess chemical potential for a polymer system. We thereupon present results for the excess chemical potential of a two-dimensional polymer 'monolayer' (Whittington–Chapman model). The latter model was not chosen because of its intrinsic interest, but because this is the system for which Siepmann obtained accurate numerical estimates of the chemical potential.

2. Method

We consider a system of polymer chains on a lattice. For convenience, we assume that the interaction between polymer segments is a simple hard-core repulsion between segments on neighbouring lattice sites. The insertion probability of a trial chain is proportional to the total number of allowed conformations for the test particle in the system.

Before computing the insertion probability of a trial chain, let us first consider a simpler quantity, namely the insertion probability of a *non-self-avoiding* trial chain of the same length. In that case it is possible to evaluate the total number of allowed random walks starting from each lattice point in one calculation. To this end, we employ a procedure similar to the 'moment-propagation method' first used to compute time-correlation functions in lattice-gas cellular automata [5]. This method is best explained by induction. Suppose that we already know, $P_N(i)$, the total number of conformations of a polymer of length N that begin on site i. Then the number of conformations of a polmer of length $N + 1$ arriving at the same site i is calculated by adding the total number of conformations of a polymer of length N, arriving on any of the neighbouring sites j and multiplying it by the probability, $P_0(i)$, that site i is, in fact, accessible:

$$P_{N+1}(i) = \left(\sum_j^{nn} P_N(j) \right) P_0(i). \qquad (2)$$

Clearly, the number of polymer conformations of length 0 (points) arriving at any lattice point can be computed easily. Once these numbers are known, the number of conformations of a polymer of length N is obtained by iteration of (2). If we denote the coordination number of the lattice by c and the number of lattice sites by M, then the 'brute-force' sampling of all polymer conformations would require some Mc^N calculations. In contrast, the present method requires only of the order of $M \times c \times N$ calculations. It should be stressed that the above scheme is in no way limited to a hard repulsive potential. If, instead, we had considered a continuous inter-segment interaction then all walks would have been weighted by the product of the Boltzmann factors associated with the sites they pass through.

Thus far, we only considered ideal, non-self-avoiding random walks. it is, however, easy to extend the above scheme in such a way that 180° reversals are excluded. In the latter case the walks arriving at site i along bond $j \rightarrow i$ are not allowed to continue through bond $i \rightarrow j$. To take this into account, we keep track of the previous step in the random walk: i.e. if we wish to compute the *total* number of walks of length $N + 2$ that can be generated by extending a $N + 1$ step random walk from i to j, we use information about all N-step random walks that terminate on a neighbour site j' of i.

$$P_{N+2}(j) = \left(\sum_i^{nnj} \left(\sum_{j' \neq j}^{nnj} P_N(f') \right) P_0(i) \right) P_0(j). \tag{3}$$

With this bookkeeping device, we can count all non-reversing random walks. In this way we can account for all conformational properties that are determined locally by a small number of successive bonds.

We do not, however, only wish our polymer conformations to be non-reversing, but also self-avoiding. This latter condition cannot be implemented in a 'local' iterative algorithm based on (2). Rather, the test for self-overlap has to be made after a particular non-reversing polymer conformation has been selected. In practice, a large number of non-reversing random walks (NRRWs) are sampled and the fraction of self avoiding walks (SAWs) is computed. Multiplying this fraction with the total number of NRRWs, yields an estimate of the total number of SAWs. In order to arrive at a correct sampling of the fraction of SAWs, we must generate a number of non-self-avoiding polymer conformations with the appropriate Boltzmann weight. This is easily achieved because we already know the total number of random walks of length $N, N - 1, N - 2, \ldots, 1$ starting from any lattice site i, $(P_N(i), P_{N-1}(i), \ldots, P_1(i))$. We now use this information as follows. First we select the starting position i of a non-self-avoiding random walk of length N with a probability proportional to $P_N(i)$. Next we consider all neighbours j of site i. The total number of NSAWs of length $N - 1$ starting from site j is equal to $P_{N-1}(j)$. We now continue our random walk from i to j with a probability proportional to $P_{N-1}(j)$. Next we consider all NSAWs of length $N - 2$ starting from any of the sites neighbouring j, and so on.

It should be stressed that we can enumerate the total number of NSAWs exactly (and cheaply), but that sampling of SAWs is a 'biased' Monte Carlo process that can be more time-consuming.

3. Simulations

To facilitate comparison of the present technique to measure the chemical potential of chain molecules with existing methods, we carried out a series of simulations of the Whittington–Chapman model for a monolayer of chain-molecules, as this is the only model for which estimates of the chemical potential have been obtained both with the Widom method and with Siepmann's modification of the Rosenbluth scheme [3].

In this model, the polymer chains are confined to a two-dimensional hexagonal lattice and their headgroups are attached to a surface at a fixed spacing Δ (see figure 1). The first bond of each polymer is free to point in two directions. This is slightly different from the model described in [3]. Siepmann [6] has, however, also computed the chemical potential for the model described here, and it is these results of Siepmann rather than those of [3] that we compared with below. Periodic boundary conditions
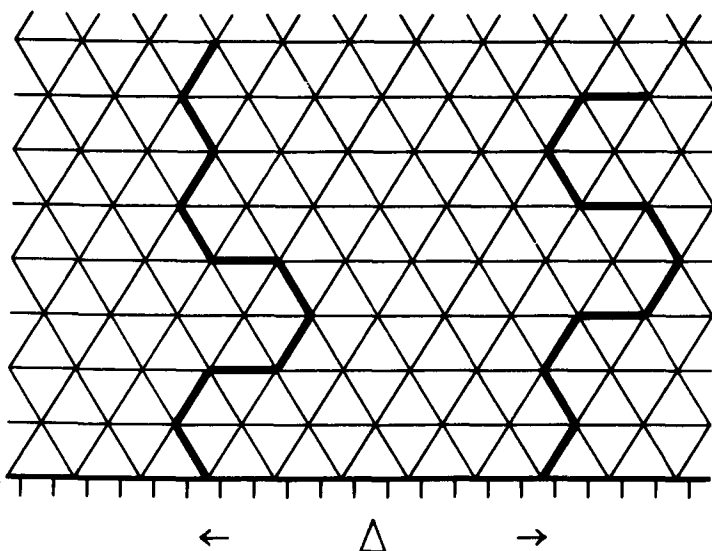
Figure 1.   An example of two polymers at spacing $\Delta = 5$.

are applied in the direction parallel to the surface. The segment–segment interaction is modelled by a hard-disc potential: this implies that a lattice site is available ony when all its nearest neighbours are empty. We do take into account that the chains have an internal energy associated with the trans-gauche isomerism. For example in figure 1 the left polymer starts with three steps in a gauche conformation while the other one starts with a trans configuration. In accordance with the simulations of [3], we chose the gauche–trans energy difference $\Delta_{TG} = 2900 \text{ kJ mol}^{-1}$. The temperature was fixed at $T = 273 \text{ K}$.

A trans or gauche configuration is locally defined by three subsequent bonds in the chains. Hence the Boltzmann factors associated with gauche-trans conformational changes can be built in to our sampling scheme for the NRRWs in much the same was as we have excluded the 180° reversals.

The simulation consists of two separate steps. The first step is the generation of a monolayer configuration. This layer is constructed using the Rosenbluth scheme [4]. The statistical weight of each generated layer is different, because the Rosenbluth weight-factor will, in general, be different. The frequency with which a specific configuration will be generated is proportional to the Boltzmann factor associated with the total number of gauche bonds. The second step of the simulation is the computation of the insertion probability of a test chain in the monolayer system. In order to facilitate comparison of the present simulations with those of [3] we use an ideal N segment chain on a hexagonal lattice as the reference state to which the excess chemical potential is referred. The total number of conformations of an ideal chain is clearly $p_N^{\text{ideal}}(i) = c^N$, where c is the coordination number of the lattice.

The average insertion probability $\langle I \rangle$ is obtained by averaging over a large number of trial insertions and a large number of different realizations of the monolayer.

4.   Results

Simulations were performed for a system of six chain molecules consisting of ten segments and with head-group spacings $\Delta = 5$, 10 and 15. In table 1 the insertion

Table 1. The results for the insertion probability, $\langle I \rangle$, and the CPU time needed, CPU, at three headgroupspacings, $\Delta = 5$, 10 and 15 for both Siepmann's method (S) and the 'moment-propagation' method (M). The percentages for the time used to compute the NRRW, part with the trans-gauche weighting factors, %NRRW, and those for the SAW part, %SAW, are given for the 'moment-propagation' method. The remaining time is used to build up the layer.

Δ	Method	$\langle I \rangle$	CPU/s	%NRRW	%SAW
5	S	$0 \cdot 3688 \pm 0 \cdot 0065 \times 10^{-3}$	15 000		
	M	$0 \cdot 3610 \pm 0 \cdot 0036 \times 10^{-3}$	11	25%	50%
10	S	$0 \cdot 2901 \pm 0 \cdot 0014 \times 10^{-2}$	2574		
	M	$0 \cdot 2926 \pm 0 \cdot 0022 \times 10^{-2}$	13	11%	75%
15	S	$0 \cdot 5492 \pm 0 \cdot 0013 \times 10^{-2}$	2574		
	M	$0 \cdot 5518 \pm 0 \cdot 0015 \times 10^{-2}$	69	9%	90%

probabilities computed using the present scheme are compared to the results obtained by Siepmann. The calculations were performed on an IBM 3090. It should be added that the method of [3] has been estimated to be a some 70 times more efficient than the 'brute-force' particle-insertion scheme. As can be seen from table 1, the scheme used in the present paper is another two to three orders of magnitude faster. In the case of $\Delta = 15$ we gain a factor 40. For $\Delta = 10$ the speed-up has increased to a factor of 200. For the highest densities ($\Delta = 5$) the present method is factor 1400 faster than Siepmann's method, i.e., 100 000 times faster than the 'brute-force' scheme. For this high density, figure 2 shows a plot of the accumulated averages of the insertion probabilities, $\langle I \rangle$. It can be seen that both methods have reached about the same convergence.

It should be noted that the part of the program that calculates the NRRW (including the trans-gauche energy difference) is completely vectorizable.
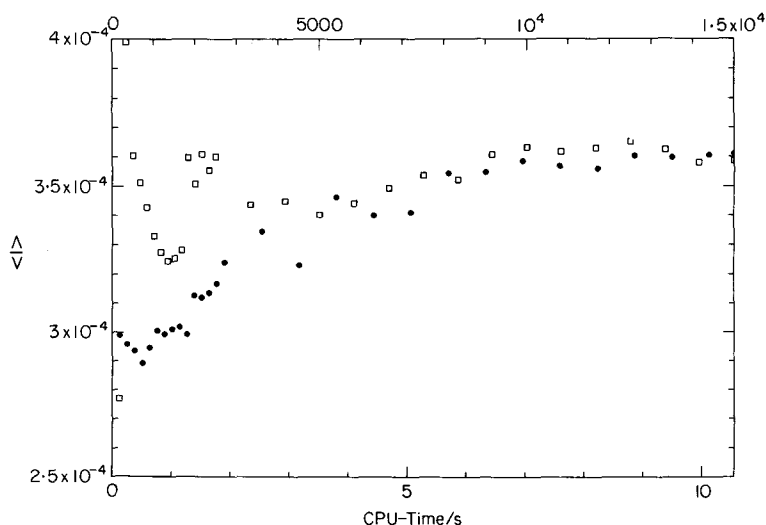

Figure 2. A plot of the accumulated averages of the insertion probabilities $\langle I \rangle$ against the CPU time needed at spacing $\Delta = 5$. Our results are indicated with ●s and the CPU time is plotted on the lower x axis. The upper axis and the □s correspond to the simulations using Siepmann's method.

For the present system, most NRRWs are also SAWs, because a chain must contain at least three gauche bonds to intersect itself. This reduces the Boltzmann weight of such a chain by a factor of $2 \cdot 2 \times 10^{-2}$. For the relatively short chains considered here, this small Boltzmann factor is not compensated by the large number of possible self-intersecting conformations.

5. Conclusions

We have presented an efficient method to enumerate the partition function of NRRWs of arbitrary length N, starting from any point on a lattice. Using this scheme we can compute the insertion probability, and thereby the excess chemical potential of interacting chain molecules. The actual performance of the method depends on the choice of the model system. We have tested the method for a two-dimensional monolayer of chain molecules. In that case we find that our method is some two or three orders of magnitude faster than the method of [3] (which, in turn is about two orders of magnitude faster than any brute-force particle insertion scheme). We should stress that the present model system was chosen because it allows comparison with [3], *not* because it is a particularly favourable case for our sampling method. On the contrary: the fact that all chains are constrained to end at the line of headgroups means that we can only use information about the number of walks terminating on this line. This means that we throw away a lot of information, because our algorithm computes the number of random walks terminating at *any* lattice site. Clearly, the comparison would have been even more favourable if we had considered a *bulk* polymer system. Even more so, if the simulations had been carried out in three dimensions.

Finally, we should point out that the present scheme is not limited to polymers on a lattice. The same scheme will work for chain molecules with continuous positions and orientations. The only thing that *is* essential is that each polymer backbone can be generated by a random walk on a suitably rotated and translated lattice. In other words, the present scheme can be applied to 'real' alkanes if:

1. we fix all dihedral angles at 0 or $\pm 120°$
2. we exclude any bond bending.

In that case the backbone of any alkane chain will fit on its *own* diamond lattice and the above techniques can be applied. It is also possible to use continuous rather than hard-core potentials. This does not affect the structure of the algorithm. However, it seems likely the computational effort will go up considerably as the interactions become more long-ranged. Of course, the method is particularly suited to study non-self-avoiding chain molecules. In that case the present algorithm guarantees that the probability of inserting a chain molecule in a given system will be computed exactly, no matter how small this probability may be.

References

[1] KREMER, K., and BINDER, K., 1988, *Comput. Phys. Rep.*, **7**, 259.

[2] WIDOM, B., 1963, *J. Chem. Phys.*, **39**, 2808.

[3] SIEPMANN, J. I., 1990, *Molec. Phys.*, **70**, 1145; SIEPMANN, J. I., and FRENKEL, D., *Molec. Phys.* (in press).

[4] ROSENBLUTH, M. N., and ROSENBLUTH, A. W., 1955, *J. chem. Phys.*, **23**, 356.

[5] FRENKEL, D., 1991, *J. Phys. Condens. Matter*, **2** SA 265; FRENKEL, D., 1989, *Cellular Automata and Modeling of Complex Physical Systems*, edited by P. Manneville, N. Boccara, G. Y. Vichniac and R. Bidaux, (Springer), p. 144; FRENKEL, D., and ERNST, M. H., 1989, *Phys. Rev. Lett.*, **63**, 2165.

[6] SIEPMANN, J. I., personal communication.