

# NUMERICAL TECHNIQUES TO STUDY COMPLEX LIQUIDS

D. FRENKEL

*FOM Institute for Atomic and Molecular Physics  
Kruislaan 407, 1098 SJ Amsterdam, The Netherlands  
email: Frenkel@amolf.nl*

## 1. Introduction

The subject of these lectures is computer simulation of complex liquids. Why should computer simulation of complex liquids be any different from computer simulation of simple liquids? After all, the great charm of computer simulation is that the complexity of the computer code does not increase with the number of degrees of freedom in a system, even though the computational effort does. Indeed, the existing Molecular Dynamics simulations of polymers, micelles or proteins are not basically different from simulations of simple atomic or molecular systems. They are simply longer. Usually much longer, because one of the features that make complex liquids complex is that interesting dynamical phenomena happen over a very wide range of time scales. For instance, in polymer dynamics the characteristic decay time of the velocity of a monomer is orders of magnitude shorter than the time it takes the polymer to relax in the 'tube' formed by its neighbors. And this time, in its turn, is much shorter than the time it takes the polymer to diffuse over a distance of the order of its own diameter. For atoms, the first and the third time are of the same order of magnitude.

But it is not only in their dynamics that complex liquids differ from simple liquids. Their structural properties too seem designed to make numerical simulation difficult. In particular, complex liquids may exhibit local order on at least two different length scales: namely on the scale of the atomic building blocks and on a length-scale comparable to the characteristic dimensions of the macromolecule. Finally, in order to study the phase behavior of complex liquids, we must either study phase coexistence directly using the Gibbs-ensemble technique (see the chapter by Panagiotopoulos) or by computing the chemical potential or, more generally, the free energy

of the different phases. Such calculations are subtle for simple liquids and even more so for complex liquids.

In these lectures, I will not discuss all aspects of computer simulation of complex liquids. Rather, I shall focus on three basic topics that are of particular relevance for the simulation of complex liquids. These topics are:

1. Free energy calculations.
2. 'Smart' Monte Carlo schemes.
3. Grand-canonical simulations.

I have selected these topics not only for their relevance for complex liquids, but also because they have evolved considerably over the past few years. I stress that many of the techniques that I discuss were not designed specifically for the study of complex liquids. Others were, but turn out to have interesting applications in other fields. In computer simulation, as in most areas of physics, the 'mobility' of tricks is quite high.

The reader may have noticed that I have not devoted a separate section to the discussion of dynamical properties. Rather, I have combined the discussion of statics and dynamics as much as possible. The reason is twofold. First of all, whereas we now have a wide range of very powerful tricks to study the static properties, including phase behavior, of complex liquids, most of the progress in the study of dynamics in complex liquids derives from the increased power of computers: the algorithms have not changed fundamentally since the fifties [1]. To give an example: the classic study by Kremer *et al.* of the reptation dynamics of dense polymer melts [2] (for a review, see [3]) employed conventional MD (and a simple model and lots of supercomputer time). Similarly, the study by Smit *et al.* [4, 5] of surfactant aggregation used conventional MD (and a simple model and very long runs on a massively parallel machine).

However, the focus of these lectures is computational techniques, not on advances in computer hardware. Nor shall I review the many very interesting results that have been obtained by 'standard' MD or MC simulations. To see where novel computational techniques can help in the study of dynamical phenomena in complex liquids, consider the factors that make dynamical simulations of complex fluids difficult.

1. The dynamics of macromolecules is often determined by topological constraints, such as entanglements.
2. Complex fluids can gelate to form effectively rigid structures.
3. Often the dynamics of complex fluids is determined by activated processes.
4. Hydrodynamic interactions can have a pronounced effect on the rheology of complex fluids.

In case 1, there exists (thus far) no easy way out: massive computation on as simple a model as possible appears to be the only way forward. In cases 2 and 3, appreciable speed-up can be achieved by the use of certain tricks. These tricks are not to be related to the dynamics itself, but to the way the equilibrium configurations of the system are sampled. In those cases, I shall refer to the dynamics in the context of the discussion of the relevant static technique. The problem of hydrodynamic effects (case 4) has been extensively studied by many authors, using a continuum description of the solvent. For the sake of coherence with the remainder of these lectures, I shall focus on techniques to simulate hydrodynamic effects that are ‘particle-based’, in the sense that the fluid medium is described as a collection of particles. In particular, I shall consider the simplest possible particle-based description of a fluid, namely a lattice gas. In that case, special tricks are possible to speed up the calculation of hydrodynamic effects. As the computational technique has its counterpart in the simulation of polymer-colloid mixtures, I shall discuss hydrodynamic effects in that context.

## 2. Free energy and Phase diagrams

### 2.1. INTRODUCTION

One of the aims of computer simulation is to predict the phase behavior of any given substance. The techniques that can be used to study phase diagrams numerically depend on the character of the phase transitions. In particular, there is quite a difference between the tools used to study first-order phase transitions and those applied to analyse critical phenomena near continuous phase transitions. In this section that is devoted to free energy calculations, I discuss only first order phase transitions. Continuous transitions are, from a computational point of view no less challenging than first-order phase transitions (see e.g. [6, 7, 8, 9]). However, the technical problems are not free-energy related.

### 2.2. FIRST ORDER PHASE TRANSITIONS

The most straightforward way to study phase behavior by computer simulation is to change the temperature or pressure of a given model system and then to simply *wait* for a phase transformation to occur. For instance, we can prepare a system in a stable crystalline phase, and then heat it until it melts. This approach can be used (and has been, in some cases [10, 11, 12, 13, 14, 15, 16]). But in general it suffers from a serious drawback: first-order phase transitions usually exhibit appreciable hysteresis. Hence the transformation to the new stable phase, if it happens at

all, will proceed irreversibly well beyond the coexistence point. The reason why hysteresis is common in first order phase transitions, is that there is a large free-energy barrier separating the two phases at, or near, coexistence. The height of this barrier is determined by the interfacial free energy of the interface separating the two coexisting phases. The larger the area of this interface, the higher the free-energy barrier. Direct simulations of first-order phase transitions therefore either start with a system that has been prepared such that the interface is already present, or by eliminating the interface altogether.

There exist several schemes to study phase coexistence without creating an interface. In fluids, the best-known method is the Gibbs-ensemble method of Panagiotopoulos [17, 18, 19, 20]. In the Gibbs-ensemble scheme, one performs two simulations in parallel. The two systems are held at the same temperature and can exchange volume and particles. At equilibrium, the pressures and chemical potentials in both systems will be equal. And hence the two systems satisfy the criteria for phase coexistence. The Gibbs-ensemble method is discussed extensively elsewhere in this volume. For what follows, it is important to point out the main limitation of the method: it breaks down when at least one of the two phase becomes so dense that it becomes effectively impossible to exchange particles. This happens for instance when one of the coexisting phases is a dense liquid or, *a fortiori*, a solid.

A direct method to study phase coexistence that does not suffer from this drawback is the so-called ‘Gibbs-Duhem’ integration method of Kofke [21, 22, 23]. This method, that will be discussed below, is designed to trace out a two-phase coexistence curve, once one point on that curve is known. As a consequence, the Kofke method can only be used in combination with another method that generates the initial point on the coexistence curve.

In the early eighties, Parrinello and Rahman [24, 25] designed a Molecular Dynamics scheme that was designed specifically to study solid-solid transitions. This method can be applied to those solid-solid transitions that cause the crystal unit cell to deform without much other rearrangement of the molecules in the unit cell. Even if these conditions are met, the Parrinello-Rahman method suffers from some hysteresis. Moreover, if the two solids have very different unit cells (such as, for instance, in the transition from the  $AB_2$  to  $AB_{13}$  structure), then the Parrinello-Rahman method cannot be used.

From the above discussion it is clear that there is a great need for ‘robust’ methods to locate phase-coexistence curves. In what follows, I shall briefly review several techniques, both ancient and modern, to study first-order phase transitions in computer simulation.

### 2.3. THERMODYNAMIC INTEGRATION

Let us briefly recall why free energies are important when we are interested in the relative stability of several phases. The second law of thermodynamics states that for a closed system with energy  $E$ , volume  $V$  and number of particles  $N$ , the entropy  $S$  is at a maximum when the system is in equilibrium. From this formulation of the second law it is simple to derive the corresponding equilibrium conditions for systems that can exchange heat, particles or volume with a reservoir. In particular, if a system is in contact with a heat bath, such that its temperature  $T$ , volume  $V$  and number of particles  $N$  are fixed, then the Helmholtz free energy  $F \equiv E - TS$  is at a *minimum* in equilibrium. Analogously, for a system of  $N$  particles at constant pressure  $P$  and temperature  $T$ , the Gibbs free energy  $G \equiv F + PV$  is at a minimum.

If we wish to know which of two phases (denoted by  $\alpha$  and  $\beta$ ) is stable at a given temperature and density, we should simply compare the Helmholtz free energies  $F_\alpha$  and  $F_\beta$  of these phases. It would seem that the obvious thing to do is simply to measure  $F_\alpha$  and  $F_\beta$  by computer simulation. Unfortunately, it is not possible to measure the free energy (or entropy) directly in a simulation. What we *can* measure in a simulation are averages of *mechanical* quantities, *i.e.* averages of functions of the coordinates and momenta of the molecules in the system, such as the pressure or the dielectric constant. If we denote such a mechanical quantity by  $A(\mathbf{p}^N, \mathbf{q}^N)$ , then the average of  $A$  that can be measured in a simulation at constant  $N$ ,  $V$  and  $T$  is

$$\langle A \rangle_{NVT} = \frac{\int d\mathbf{p}^N d\mathbf{q}^N A(\mathbf{p}^N, \mathbf{q}^N) \exp(-\beta H(\mathbf{p}^N, \mathbf{q}^N))}{\int d\mathbf{p}^N d\mathbf{q}^N \exp(-\beta H(\mathbf{p}^N, \mathbf{q}^N))} \quad (1)$$

where  $H$  is the Hamiltonian of the system expressed as a function of the momenta  $\mathbf{p}^N$  and coordinates  $\mathbf{q}^N$ .  $k_B$  is Boltzmann's constant and  $\beta = 1/(k_B T)$ .

However, entropy, free energy and related quantities are not simply averages of functions of the phase-space coordinates of the system. Rather they are directly related to the *volume* in phase space that is accessible to a system. For instance, in classical statistical mechanics, the Helmholtz free energy  $F$  is directly related to the canonical partition function  $Q_{NVT}$ :

$$F = -k_B T \ln Q_{NVT} \equiv -k_B T \ln \left( \frac{\int d\mathbf{p}^N d\mathbf{q}^N \exp(-\beta H(\mathbf{p}^N, \mathbf{q}^N))}{h^{dN} N!} \right) \quad (2)$$

where  $h$  is Planck's constant and  $d$  is the dimensionality of the system.

It is clear that  $Q_{NVT}$  is not of the form of a canonical average over phase space. And this is why  $F$  or, for that matter,  $S$  or  $G$ , cannot be measured

directly in a simulation. We call such quantities that depend directly on the available volume in phase *thermal* quantities.

In fact, there is nothing strange about the fact that thermal quantities cannot be measured directly in a simulation: the same problem occurs in the real world – thermal quantities cannot be measured directly.

When considering numerical schemes to determine the free energy, it is therefore instructive to look at the *experimental* techniques to determine entropies or free energies. In experiments, one does not measure the free energy directly, but rather a derivative of the free energy with respect to volume  $V$  or temperature  $T$ :

$$\left(\frac{\partial F}{\partial V}\right)_{NT} = -P \quad (3)$$

and

$$\left(\frac{\partial F/T}{\partial 1/T}\right)_{VT} = E. \quad (4)$$

As the pressure  $P$  and the energy  $E$  are mechanical quantities, they *can* be measured both in experiment and in simulation. In order to compute the free energy of a system at given temperature and density, we should find a reversible path in the  $V - T$  plane, that links the state under consideration to a state of known free energy. The change in  $F$  along that path can then simply be evaluated by integration of eqs. 3 and 4. There are only very few thermodynamic states where the free energy of a substance is known. One state is the ideal gas phase, the other is the perfectly ordered ground state at  $T = 0K$ .

In computer simulations, the situation is quite similar. In order to compute the free energy of a dense liquid, one may construct a reversible path to the very dilute gas phase. It is not really necessary to go all the way to the ideal gas. But at least one should reach a state that is sufficiently dilute that the free energy can be computed accurately, either from knowledge of the first few terms in the virial expansion of the compressibility factor  $PV/(NkT)$ , or that the chemical potential can be computed by other means (see section 2.5.1 and 2.5.2 below). For the solid, the ideal gas reference state is less useful (although techniques have been developed to construct a reversible path from a dense solid to a dilute (lattice) gas [26]). The obvious reference state for solids is the harmonic lattice. Computing the absolute free energy of a harmonic solid is relatively straightforward, at least for atomic and simple molecular solids. However, not all solid phases can be reached by a reversible route from a harmonic reference state. For instance, in molecular systems it is quite common to find a strongly anharmonic plastic phase just below the melting line. This plastic phase is not (meta-) stable at low temperatures.

### 2.3.1. Artificial paths

Fortunately, in computer simulations we do not have to rely on the presence of a ‘natural’ reversible path between the phase under study and a reference state of known free energy. If such a path does not exist, we can construct an artificial path. This is in fact a standard trick in statistical mechanics (see e.g. [27]). It works as follows: Consider a case where we need to know the free energy  $F(V, T)$  of a system with a potential energy function  $U_1$ , where  $U_1$  is such that no ‘natural’ reversible path exists to a state of known free energy. Suppose now that we can find another model system with a potential energy function  $U_0$  for which the free energy *can* be computed exactly. We define a generalized potential energy function  $U(\lambda)$ , such that  $U(\lambda = 0) = U_0$  and  $U(\lambda = 1) = U_1$ . The free energy of a system with this generalized potential is denoted by  $F(\lambda)$ . Although  $F(\lambda)$  itself cannot be measured directly in a simulation, we can measure its derivative with respect to  $\lambda$ :

$$\left(\frac{\partial F}{\partial \lambda}\right)_{NVT\lambda} = \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{NVT\lambda} \quad (5)$$

If the path from  $\lambda = 0$  to  $\lambda = 1$  is reversible, we can use eq. 5 to compute the desired  $F(V, T)$ . We simply measure  $\langle \partial U / \partial \lambda \rangle$  for a number of values of  $\lambda$  between 0 and 1.

In the numerical simulation of bio-molecular molecules, ‘artificial’ thermodynamic integration is often used to compute the difference in excess free-energy of similar but distinct molecules (see e.g. [28]). Such calculations are important in many different contexts. For instance, one can thus compute the relative binding strengths of closely related molecules to an enzyme. In such calculations, the thermodynamic integration involves a gradual replacement of part of the molecule by another building block (for instance, an  $H$  could be transformed into a  $CH_3$  group).

### 2.3.2. Adiabatic transformations

It should be noted that the thermodynamic integration method based on Eqn. 5 is intrinsically ‘static’, *i.e.* the derivative of the free energy is obtained in a series of equilibrium simulations (either MC or MD). An intrinsically *dynamic* scheme to perform free energy computations was suggested by Watanabe and Reinhardt [29]. This method also relies on the existence of a reversible path between the state-point of interest and a simple reference system. However, the method of ref. [29] is based on the concept of adiabatic changes. If the Hamiltonian of a closed system is changed sufficiently slowly (compared to the slowest natural time-scale in the system under consideration), then the entropy of the system does not change [30].

Clearly, the method of Watanabe and Reinhardt is designed for molecular dynamics, rather than Monte Carlo, simulations. However, the

method is not limited to the microcanonical ensemble. It can be extended to other ensembles (*e.g.*  $NVT$ ), in which case it becomes a technique to compute free-energy changes [29]. It should be noted that several authors have used Molecular Dynamics simulations with a time-dependent Hamiltonian to compute free-energy changes [31]. These simulations are close in spirit to the Monte Carlo scheme although, as pointed out in ref. [29], they can be ‘derived’ from the adiabatic transformation scheme.

### 2.3.3. *Reference systems*

Let us consider the computation of the free energy of a crystalline solid by thermodynamic integration. In such calculations it is important to select a reasonable reference system, *i.e.* a system that is, in some sense ‘close’ to the system of interest, and can be reached without crossing any first-order phase-transition boundaries. One of the safest choices for a reference system appears to be an Einstein crystal with the same structure as the phase under study [32]. This choice of reference system makes it extremely improbable that the path connecting  $\lambda = 0$  and  $\lambda = 1$  will cross an (irreversible) first order phase transition from the initial structure to another, only to go back to its original structure for still larger values of  $\lambda$ . Nevertheless, it is important that the parametrization of  $U(\lambda)$  be chosen carefully. Usually, a linear parametrization (*i.e.*  $U(\lambda) = \lambda U_1 + (1 - \lambda)U_0$ ) is quite satisfactory. But occasionally such a parametrization may lead to weak (and relatively harmless) singularities in eq. 5 for  $\lambda \rightarrow 0$ . There are, however, alternative thermodynamic integration schemes that do not suffer from this problem [33]. More details about such free energy computations can be found in ref. [33, 34, 35].

Similar techniques can be used to locate first-order phase transitions involving phases with partial order (*e.g.* liquid crystals). For details, the reader is referred to refs [36, 37, 38, 39, 40, 41]. Finally, thermodynamic integration techniques are particularly useful to compute the Gibbs free energy of mixtures. An example is the work of Kranendonk and Frenkel, who have computed the melting curve of a substitutionally disordered solid mixture of hard spheres with different size [42]. In this case, thermodynamic integration is used to slowly change the size-ratio of the particles at fixed volume fraction. The reference point is size ratio  $\alpha = 1$  (‘isotopic mixture’). The reversible work needed to change the size ratio from 1 to a given value of  $\alpha$  is a direct measure for the excess free energy of the solid solution at that composition. This technique was subsequently used by Eldridge et al. [43] to compute the relative stability of the  $AB_2$  and  $AB_{13}$  crystal phases of dissimilar hard spheres with a size ratio  $\alpha \approx 0.58$ . For details, the reader is referred to ref. [44, 45, 46].

## 2.4. TRACING COEXISTENCE CURVES

Once a single point on the coexistence curve between two phases is known, the rest of that curve can be computed without further free energy calculations. To map out such coexistence curves, Kofke [21, 22] proposed a scheme that is based on a numerical integration of the Clausius-Clapeyron equation (unfortunately, Kofke refers to the method as ‘Gibbs-Duhem’ integration). Let us briefly recall the Clausius-Clapeyron equation. If two phases coexist at a given temperature  $T$  and pressure  $P$ , then their chemical potentials must be equal. If we now change both the pressure and the temperature by infinitesimal amounts  $dP$  and  $dT$  respectively, then the difference in chemical potential of the two phases becomes

$$\mu_\alpha - \mu_\beta = -(S_\alpha - S_\beta)dT + (V_\alpha - V_\beta)dP \quad (6)$$

Along the coexistence curve  $\mu_\alpha = \mu_\beta$ , and hence

$$\frac{dP}{dT} = \frac{S_\alpha - S_\beta}{V_\alpha - V_\beta} = \frac{\Delta H}{T\Delta V} \quad (7)$$

where we have used the fact that, at coexistence,  $\Delta S = T\Delta H$ , where  $H_\alpha$  ( $H_\beta$ ) denotes the enthalpy of phase  $\alpha$  ( $\beta$ ). As  $\Delta H$ ,  $T$  and  $\Delta V$  can all be measured directly in a simulation,  $dP/dT$  can be computed from Eqn. 7. Kofke used a predictor-corrector algorithm to solve Eqn. 7 although other integration schemes may be more stable [47]. If one of the two coexisting phases is the (dilute) vapor phase, it is convenient to cast Eqn. 7 in a slightly different form:

$$\frac{d \ln P}{d1/T} = - \frac{\Delta H}{P\Delta V/T} \quad (8)$$

Examples of recent applications of the Kofke method can be found in refs. [48, 49, 50, 51, 47]. It should be stressed that the Kofke method is in no way limited to the computation of coexistence curves in the  $P, T$ -plane. A particularly important class of problems that can be treated in an analogous fashion are those where one studies the location of a phase transition as a function of the intermolecular interaction potential. For instance, Agrawal and Kofke have investigated the effect of a change of the steepness of the intermolecular potential in atomic systems on the melting point [23]. In the same spirit, Dijkstra and Frenkel studied the effect of a change in flexibility of rod-like polymers on the location of the isotropic nematic transition [52].

## 2.5. CHEMICAL POTENTIALS

The previous section discussed several direct, and very robust, techniques to compute ‘thermal’ properties of a many-body system. However, in many

cases there exist computationally cheaper ways to arrive at the same result. In particular, there exist several techniques to ‘measure’ the chemical potential of a given species in a single Monte Carlo or Molecular Dynamics simulation. This seems to contradict the statement made in section 2.3 that thermal properties cannot be measured directly in a simulation. However, on closer inspection it will appear that what we measure is not the chemical potential itself, but the *excess* chemical potential, *i.e.* the difference between the chemical potential of a given species and that of an ideal gas under the same conditions.

### 2.5.1. *The particle insertion method*

A particularly simple and elegant method to measure the chemical potential  $\mu$  of a species in a pure fluid or in a mixture is the ‘particle-insertion’ method (often referred to as the Widom-method [53]). The statistical mechanics that is the basis for this method is quite simple. Consider the definition of the chemical potential  $\mu_\alpha$  of a species  $\alpha$ . From thermodynamics we know that  $\mu$  is defined as:

$$\begin{aligned}\mu &= \left(\frac{\partial G}{\partial N}\right)_{PT} \\ &= \left(\frac{\partial F}{\partial N}\right)_{VT} \\ &= -T \left(\frac{\partial S}{\partial N}\right)_{VE}.\end{aligned}\quad (9)$$

Where  $G$ ,  $F$  and  $S$  are the Gibbs free energy, the Helmholtz free energy and the entropy, respectively. Here, and in the next few paragraphs we focus on a one-component system, and hence we drop the subscript  $\alpha$ . Moreover, to keep the notation compact, I shall initially assume that we deal with a system of  $N$  atoms in a cubic volume with diameter  $L$  and volume  $V = L^d$ , at constant temperature  $T$ . The classical partition function of such a system is given by:

$$Q(N, V, T) = \frac{1}{\Lambda^{dN} N!} \int_0^L \cdots \int_0^L d\mathbf{q}^N \exp(-\beta U(\mathbf{q}^N)), \quad (10)$$

where  $\Lambda = \sqrt{h^2/(2\pi mkT)}$  is the thermal de Broglie wavelength. It is convenient to rewrite eqn. 10 in a slightly different way by introducing scaled coordinates  $\mathbf{s}^N$ , as follows:

$$\mathbf{q}_i = L\mathbf{s}_i$$

for  $i = 1, 2, \dots, N$ . If we now insert these scaled coordinates in eqn 10 we obtain:

$$Q(N, V, T) = \frac{V^N}{\Lambda^{dN} N!} \int_0^1 \cdots \int_0^1 d\mathbf{s}^N \exp(-\beta U(\mathbf{s}^N; L)). \quad (11)$$

In equation 11, we have written  $U(\mathbf{s}^N; L)$  to indicate that  $U$  depends on the *real* rather than the *scaled* distances between the particles. The expression for the Helmholtz free energy of the system is:

$$\begin{aligned} F(N, V, T) &= -kT \ln Q \\ &= -kT \ln \left( \frac{V^N}{\Lambda^{dN} N!} \right) - kT \ln \left( \int d\mathbf{s}^N \exp(-\beta U(\mathbf{s}^N; L)) \right) \\ &= F_{id}(N, V, T) + F_{ex}(N, V, T). \end{aligned} \quad (12)$$

In the last line of the above equation we have identified the two contributions to the Helmholtz free energy on the previous line as the ideal-gas expression plus an excess part. It is obvious from eq. 9 that, for sufficiently large  $N$  the chemical potential is given by:  $\mu = -k_B T \ln(Q_{N+1}/Q_N)$ . If we use the explicit form (eq. 12) for  $Q_N$ , we find:

$$\begin{aligned} \mu &= -k_B T \ln(Q_{N+1}/Q_N) \\ &= -k_B T \ln \left( \frac{V/\Lambda^d}{(N+1)} \right) - k_B T \ln \left( \frac{\int d\mathbf{s}^{N+1} \exp(-\beta U(\mathbf{s}^{N+1}))}{\int d\mathbf{s}^N \exp(-\beta U(\mathbf{s}^N))} \right) \\ &= \mu_{id}(V) + \mu_{ex}. \end{aligned} \quad (13)$$

In the last line of eq. 13, we have separated the chemical potential in an ideal-gas contribution  $\mu_{id}$  and the excess part  $\mu_{ex}$ . As  $\mu_{id}(V)$  can be evaluated analytically, we focus on  $\mu_{ex}$ . We now separate the potential energy of the  $N + 1$ -particle system into the potential energy function of the  $N$ -particle system,  $U(\mathbf{s}^N)$ , and the interaction energy of the  $N + 1$ -th particle with the rest:  $\Delta U \equiv U(\mathbf{s}^{N+1}) - U(\mathbf{s}^N)$ . Using this separation, we can write  $\mu_{ex}$  as:

$$\mu_{ex} = -k_B T \ln \left\langle \int d\mathbf{s}_{N+1} \exp(-\beta \Delta U) \right\rangle_N, \quad (14)$$

where  $\langle \dots \rangle_N$  denotes canonical ensemble averaging over the configuration space of the  $N$ -particle system. The important point to note is that equation 14 expresses  $\mu_{ex}$  as an ensemble average that can be sampled by the conventional Metropolis scheme [54]. There is one aspect of the average in Eqn. 14 that makes it different from the quantities that are usually sampled in a computer simulation: we have to compute the average of an *integral* over the position of particle  $N + 1$ . This last integral can be sampled by brute-force (unweighted) Monte Carlo sampling. In practice the procedure is as follows: we carry out a perfectly normal constant- $NVT$  Monte Carlo simulation on the system of  $N$  particles. At frequent intervals during this simulation we randomly generate a coordinate  $\mathbf{s}_{N+1}$ , uniformly over the unit cube. With this value of  $\mathbf{s}_{N+1}$ , we then compute  $\exp(-\beta \Delta U)$ . By averaging the latter quantity over all generated trial positions, we obtain the

average that appears in eq. 14. So, in effect, we are computing the average of the Boltzmann factor associated with the random insertion of an additional particle in an  $N$ -particle system, *but we never accept any such trial insertions*, because then we would no longer be sampling the average needed in eq. 14. The Widom method provides us with a very powerful scheme to compute the chemical potential of (not too dense) atomic and simple molecular liquids. The extension of the Widom method to other ensembles, in particular  $NPT$  and  $NVE$  is relatively straightforward. However, it would be incorrect simply to apply equation 14 to these other ensembles. This point is discussed in some detail in the following references for the  $NPT$ -ensemble [55, 56, 57] and in [34, 58, 59] for the  $NVE$ -ensemble.

Extensive numerical studies of the excess chemical potential of finite periodic systems, have shown that this quantity is rather strongly dependent on the size of the system [60, 61]. Computer simulations are typically carried out for periodic systems in which the fundamental cell contains of order  $10^2$  to  $10^3$  particles, and the correction needed to give the infinite-system result can therefore be large. Of course, it is possible to estimate the finite-size correction empirically by carrying out simulations for different values of  $N$  (the number of particles), but this is very time consuming. It would clearly be much more convenient if the finite-size correction could be estimated directly, since it would then be possible to estimate the chemical potential in the thermodynamic limit on the basis of results obtained from simulations of relatively small systems. In fact, Siepmann, McDonald and Frenkel [62] have derived an expression for the leading ( $\mathcal{O}(1/N)$ ) system-size dependence of the excess chemical potential:

$$\Delta\mu_{ex}(N) = \frac{1}{2N} \left( \frac{\partial P}{\partial \rho} \right) \left[ 1 - k_B T \left( \frac{\partial \rho}{\partial P} \right) - \rho k_B T \frac{(\partial^2 P / \partial \rho^2)}{(\partial P / \partial \rho)^2} \right]. \quad (15)$$

As shown in ref. [62], this expression agrees with the exact result for hard rods in one dimension, and is in excellent agreement with numerical results for hard-disks in two dimensions.

The particle insertion method can be modified to measure the *difference* in chemical potential between two species  $\alpha$  and  $\beta$  in a mixture. In this case a trial move consists of an attempt to transform a particle of species  $\alpha$  into species  $\beta$  (without, of course, ever accepting such trial moves). For more details, the reader is referred to refs. [63, 57]. Finally it should be stressed that particle insertion and swapping techniques are not limited to the measurement of chemical potentials. In fact, a wide class of partial molar quantities (such as, for example the partial molar enthalpy  $h_\alpha$  or the partial molar volume  $v_\alpha$ ) can be measured in this way. For details, see refs. [57, 64]

### 2.5.2. Overlapping distribution method

The reader may wonder why, in the previous section, we have only been discussing a trial move that attempt to *add* a particle to the system, and not the reverse move. After all, the excess chemical potential can also be written as:

$$\begin{aligned}\mu &= +kT \ln(Q_N/Q_{N+1}) \\ &= \mu_{id} + kT \ln \langle \exp(+\beta\Delta U) \rangle_{N+1},\end{aligned}\quad (16)$$

where  $\Delta U$  denotes the interaction energy of particle  $N + 1$  with the remaining  $N$  particles. It would seem that eqn. 16 can be sampled by straightforward Metropolis Monte Carlo. In general, however, this is not true. The reason is that the function  $\exp(\beta\Delta U)$  is, in principle not bounded. It can become arbitrarily large, as  $\Delta U$  grows. (Incidentally, this is not true for  $\exp(-\beta\Delta U)$ , because one of the conditions that a system must satisfy in order to be describable by classical statistical mechanics is that its potential energy function has a lower bound). The problem with equation 16 is that very large values of the integrand coincide with very small values ( $\mathcal{O}(\exp(-\beta\Delta U))$ ) of the Boltzmann factor (that determines how often a configuration is sampled during a Monte Carlo run). As a consequence, an appreciable contribution to the average in eqn. 16 comes from a part of configuration space that is hardly ever, or indeed never, sampled during a run. Hard spheres offer a good illustration. As the potential energy function of non-overlapping hard spheres is always zero, a simple Monte Carlo sampling of eqn. 16 for a dense fluid of hard spheres would always yield the nonsensical estimate  $\mu_{ex} = 0$  (whereas, in fact, at freezing,  $\mu_{ex}/kT \sim 15$ ). The correct way to obtain chemical potentials from simulations involving both particle insertions and particle removals has been indicated by Shing and Gubbins [65, 66]. However, I find it convenient to discuss this problem in the context of a more general technique to measure free energy differences, first introduced by Bennett [67], called the overlapping distribution method.

Consider two  $N$ -particle systems, labeled 0 and 1 with partition functions  $Q_0$  and  $Q_1$ . For convenience we assume that both systems have the same volume  $V$ , but this is not essential. From eqn 12 it follows that the free energy difference  $\Delta F = F_1 - F_0$  can be written as:

$$\begin{aligned}\Delta F &= -kT \ln(Q_1/Q_0) \\ &= -kT \ln \left( \frac{\int ds^N \exp(-\beta U_1(\mathbf{s}^N))}{\int ds^N \exp(-\beta U_0(\mathbf{s}^N))} \right).\end{aligned}\quad (17)$$

Suppose that we are carrying out a (Metropolis) sampling of the configuration space of system 1. For every configuration visited during this

sampling of system 1 we can compute the potential energy of system 0 ( $U_0(\mathbf{s}^N)$ ) for the same configuration, and hence the potential energy difference  $\Delta U = U_1(\mathbf{s}^N) - U_0(\mathbf{s}^N)$ . We use this information to construct a histogram that measures the probability density for the potential energy difference  $\Delta U$ . Let us denote this probability density by  $p_1(\Delta U)$ . In the canonical ensemble,  $p_1(\Delta U)$  is given by:

$$p_1(\Delta U) = \frac{\int d\mathbf{s}^N \exp(-\beta U_1) \delta(U_1 - U_0 - \Delta U)}{q_1}, \quad (18)$$

where we have denoted the scaled, configurational part of the partition function by a lower-case  $q$  (e.g.  $q_1 = \int d\mathbf{s}^N \exp(-\beta U_1(\mathbf{s}^N))$ ). The  $\delta$ -function in eqn. 18 makes that we can substitute  $U_0 + \Delta U$  for  $U_1$  in the Boltzmann factor, hence:

$$\begin{aligned} p_1(\Delta U) &= \frac{\int d\mathbf{s}^N \exp(-\beta(U_0 + \Delta U)) \delta(U_1 - U_0 - \Delta U)}{q_1} \\ &= \frac{q_0}{q_1} \exp(-\beta \Delta U) \frac{\int d\mathbf{s}^N \exp(-\beta U_0) \delta(U_1 - U_0 - \Delta U)}{q_0} \\ &= \frac{q_0}{q_1} \exp(-\beta \Delta U) p_0(\Delta U), \end{aligned} \quad (19)$$

where  $p_0(\Delta U)$  is the probability density to find a potential energy difference  $\Delta U$  between systems 1 and 0, while Boltzmann-sampling the available configurations of system 0. As the free energy difference between systems 1 and 0 is simply  $\Delta F = -kT \ln(q_1/q_0)$ , we find from equation 19:

$$\ln p_1(\Delta U) = \beta(\Delta F - \Delta U) + \ln p_0(\Delta U). \quad (20)$$

In order to obtain  $\Delta F$  from eqn. 20 in practical cases, it is convenient to define two functions  $f_0$  and  $f_1$  by:

$$f_0(\Delta U) = \ln p_0(\Delta U) - \frac{\beta \Delta U}{2}$$

and

$$f_1(\Delta U) = \ln p_1(\Delta U) + \frac{\beta \Delta U}{2}$$

such that

$$f_1(\Delta U) = f_0(\Delta U) + \beta \Delta F.$$

Suppose that we have measured  $f_0$  and  $f_1$  in two separate simulations (one sampling system 0, the other system 1). We can then obtain  $\Delta F$  by fitting the functions  $f_0$  and  $f_1$  to two polynomials in  $\Delta U$  that are identical but for the constant term. The constant offset between the two polynomials

yields our estimate for  $\Delta F$ . Note that, in order to perform such a fit, it is not even necessary that there exists a range of  $\Delta U$  where both  $f_0$  and  $f_1$  can be measured. However, in the absence of such a range of overlap, the statistical accuracy of the method is usually poor.

Now consider the particle insertion/removal problem. Let us assume that system 1 is a system with  $N$  interacting particles, while system 0 contains  $N - 1$  interacting particles and 1 ideal-gas particle. The difference in free energy between these two systems is obviously equal to  $\mu_{ex}$ . Applying eqn. 20 to this particular case, we find:

$$\Delta\mu_{ex} = f_1(\Delta U) - f_0(\Delta U). \quad (21)$$

Eqn. 21 is equivalent to the result obtained by Shing and Gubbins. Using the overlapping distribution method it is possible to combine the results of simulations with trial insertions and trial removals to arrive at a more accurate estimate for the chemical potential. In section 3.1, we discuss the extension of the Bennett method to multiple histograms, and indicate the relation with recent developments in this field [68, 8].

## 2.6. BEYOND RANDOM INSERTION

The particle insertion scheme described above fails when the probability of ‘accepting’ a trial insertion becomes very small. One consequence is that the simple particle insertion method is less suited for molecular than for atomic systems. This is so because the probability of accepting the random trial insertion of a large molecule in a fluid is usually extremely small.

It is, however, possible to overcome this problem, at least partially, by performing non-random sampling. Below, I shall discuss several of the techniques that have been proposed to compute the chemical potential of chain molecules. Three different approaches have been proposed to improve the efficiency of the original Widom scheme. The most direct of these techniques are, in essence, thermodynamic integration schemes. Next, I will discuss the method that is based on (generalizations of) the Rosenbluth algorithm [69] to generate polymer conformations. And finally, I mention a recursive algorithm that has been proposed recently.

### 2.6.1. *Chemical potential as reversible work*

The excess chemical potential of a (chain) molecule is simply the reversible work needed to add such a molecule to a liquid in which  $N$  other (possibly identical) molecules are already present. If we choose to break up the insertion of the molecule into a number of steps, then clearly the reversible work needed to insert the whole molecule is equal to the sum of the contributions of the sub-steps. At this stage, we are still free to choose the elementary

steps, just as we are free to choose whatever reversible path we wish when performing thermodynamic integration. One obvious possibility is to start with an ideal (non-interacting) chain molecule and then slowly switch on the interaction of this molecule with the surrounding particles (and the non-bonded intra-molecular interactions). This could be done in the way described in sections 2.3. In fact, in a recent paper, Müller and Paul [70] have performed a simulation in which the polymer interaction is switched on gradually. Although this simulation could have been performed with straightforward thermodynamic integration, a multiple-histogram method (see section 3.1) was used instead, but this does not change the overall nature of the calculation. As stated before, the advantage of thermodynamic integration (and related techniques) is that it is robust. The disadvantage is that it is no longer possible to measure the excess chemical potential in a *single* simulation.

A closely related method to measure the chemical potential of a chain molecule has been proposed by Kumar *et al.* [71]. In this scheme, the chain molecule is built up step-by-step. The method of Kumar *et al.* resembles the ‘gradual insertion scheme’ to measure excess chemical potentials that was proposed several years earlier by Mon and Griffiths [72]. The reversible work involved in the intermediate steps is measured using the Widom method, *i.e.* the difference in excess free energy of a chain of length  $\ell$  and  $\ell + 1$  is measured by computing  $\Delta U(\ell \rightarrow \ell + 1)$ , the change in potential energy associated with the addition of the  $\ell + 1$ -st monomer. The change in free energy is then given by:

$$\begin{aligned} \Delta F_{ex}(\ell \rightarrow \ell + 1) &\equiv \mu_{ex}^{incr}(\ell \rightarrow \ell + 1) & (22) \\ &= -k_B T \ln \langle \exp(-\beta \Delta U(\ell \rightarrow \ell + 1)) \rangle . \end{aligned}$$

This equation defines the ‘incremental excess chemical potential’  $\mu_{ex}^{incr}(\ell \rightarrow \ell + 1)$ . The excess chemical potential of the complete chain molecule is simply the sum of the individual incremental excess chemical potentials. As the latter contributions are measured using the Widom method, the scheme of Kumar *et al.* is referred to as the ‘modified’ Widom method. This method is subject to the same limitations as the original Widom method (*i.e.* the insertion probability of the individual monomers should be appreciable). In this respect, it is slightly less general than thermodynamic integration. As with the method of ref. [70], the computation of the excess chemical potential may require many individual simulations.

## 2.7. ROSENBLUTH SAMPLING

During the past few years there have been several proposals to measure the chemical potential of a chain molecule in a *single* simulation. Harris and

Rice [73] and Siepmann [74] showed how to compute the chemical potential of chain molecules with discrete conformations using an algorithm to generate polymer conformations due to Rosenbluth and Rosenbluth [69]. A generalization to continuously deformable molecules was proposed by Frenkel, Smit and Mooij [75, 76] and by de Pablo *et al.* [79]. As the extension of the sampling scheme from molecules with discrete conformations to continuously deformable molecules is non-trivial, I shall discuss the two cases separately.

### 2.7.1. Chemical potential of macro-molecules with discrete conformations

It is instructive to first consider how we would compute  $\mu_{ex}$  of a chain molecule with the Widom technique. To this end, I introduce the following notation: the position of the first segment of the chain molecule is denoted by  $\mathbf{q}$  and the conformation of the molecule is described by  $\Gamma$ . The configurational part of the partition function of a system of chain molecules can be written as

$$Q_{chain}(N, V, T) = \frac{1}{N!} \int d\mathbf{q}^N \sum_{\Gamma_1, \dots, \Gamma_n} \exp(-\beta U(\mathbf{q}^N, \Gamma^N)). \quad (23)$$

The analogy with the previous sections suggests that the excess chemical potential of a chain molecule is obtained by considering the ratio

$$Q(N + 1, V, T) / [Q(N, V, T) Q_{non-interacting}(1, V, T)],$$

where the numerator is the (configurational part of) the partition function of a system of  $N + 1$  interacting chain molecules while the denominator the partition function for a system consisting of  $N$  interacting chains and one chain that does not interact with the others. The latter chain plays the role of the ideal gas molecule in the previous sections. Note, however, that although this molecule does not interact with any of the other molecules it *does* interact with itself, both through bonded and through non-bonded interactions. Unfortunately, this is not a particularly useful reference state, as we do not know *a priori* the partition function of an isolated self-avoiding chain. We therefore use another reference state, namely that of the isolated non-selfavoiding chain. It should be stressed that the choice of another reference system makes no difference whatsoever for the computation of any observable property.

Let us consider the case of a molecule that consists of  $\ell$  segments. Starting from segment 1, we can add segment 2 in  $b_2$  equivalent directions, and so on. Clearly, the total number of non-selfavoiding conformations is  $\Omega_{id} = \prod_{i=1}^{\ell} b_i$ . For convenience, I have assumed that for a given  $i$ , all  $b_i$  directions are equally likely (*i.e.* I ignore *gauche-trans* potential energy differences and I even allow the ideal chain to fold back on itself). These

limitations are not essential but they simplify the notation. Finally, I assume that all  $b_i$  are the same. Hence, for the simple model that we consider,  $\Omega_{id} = b^\ell$ . If we use such an ideal chain as our reference system, the expression for the excess chemical potential becomes

$$\begin{aligned} \beta\mu_{ex} &= -k_B T \ln \left( \frac{Q_{chain}(N+1, V, T)}{Q(N, V, T)Q_{ideal}(1, V, T)} \right) \\ &= -k_B T \ln \langle \exp[-\beta\Delta U(\mathbf{q}^N, \mathbf{\Gamma}^N; \mathbf{q}_{N+1}, \mathbf{\Gamma}_{N+1})] \rangle, \quad (24) \end{aligned}$$

where  $\Delta U$  denotes the interaction of the test chain with the  $N$  chains that are already present in the system *and with itself*, while  $\langle \dots \rangle$  indicates averaging over all starting positions and all ideal-chain conformations of a randomly inserted chain.

The problem with the Widom approach to eq. 24 is that almost all randomly inserted ideal chain conformations will overlap either with particles already present in the system, or internally. The most important contributions to  $\mu_{ex}$  will come from the extremely rare cases where the trial chain happens to be in just the right conformation to fit into the available space in the fluid. Clearly, it would be desirable if we could restrict our sampling to those conformations that satisfy this condition. If we do that, we introduce a bias in our computation of the insertion probability and we must somehow correct for that bias.

The ‘Rosenbluth’ approach used in refs. [73, 74] consists of two steps: in the first step a chain conformation is generated with a bias that ensures that ‘acceptable’ conformations are created with a high probability. The next step corrects for this bias by multiplying with a weight factor. A scheme that generates ‘acceptable’ chain conformations with a high probability was developed by Rosenbluth and Rosenbluth in the early fifties [69]. In the Rosenbluth scheme, a conformation of a chain molecule is constructed segment-by-segment. For every segment, we have a choice of  $b$  possible directions. In the Rosenbluth scheme, this choice is not random but favors the direction with the largest Boltzmann factor. To be specific, the probability ( $P$ ) to generate a polymer with a conformation  $\mathbf{\Gamma}$  using the Rosenbluth algorithm is given by

$$P_{\mathbf{\Gamma}} = \prod_{i=1}^{\ell} \frac{\exp[-\beta u^{(i)}(\mathbf{\Gamma}_i)]}{Z_i}, \quad (25)$$

where  $u^{(i)}(\mathbf{\Gamma}_i)$  denotes the energy of segment  $i$  of the chain with conformation  $\mathbf{\Gamma}$  (note that this energy excludes the contributions of segments  $i+1$  to  $\ell$ , so the total energy of the chain is given by:  $U_{\mathbf{\Gamma}} = \sum_{i=1}^{\ell} u^{(i)}(\mathbf{\Gamma}_i)$ ).  $Z_i$

in equation 25 is shorthand for

$$Z_i \equiv \sum_{j=1}^b \exp \left[ -\beta u^{(i)}(\Gamma_j) \right] .$$

where  $j$  enumerates all possible orientations from which the  $i$ -th segment of the chain can be chosen and  $u^{(i)}(\Gamma_j)$  denotes the potential energy of the  $i$ -th segment in orientation  $j$ . An important property of the probability given by eq. 25 is that it is normalized, i.e

$$\sum_{\Gamma} P_{\Gamma} = 1 .$$

The Rosenbluth weight factor that corrects for the bias in the selection of conformation  $\Gamma$  is given by

$$W_{\Gamma} \equiv \prod_{i=1}^{\ell} \frac{Z_i}{b} . \quad (26)$$

Now let us assume that we use the Rosenbluth scheme to generate a large number of chain conformations while keeping the coordinates of all other particles in the system fixed. For this set of conformations, we compute the average of the Rosenbluth weight factor  $W$ ,  $\overline{W}$ . If we also perform an ensemble average over all coordinates and conformations of the  $N$  particles in the system, we obtain

$$\langle W \rangle = \left\langle \sum_{\Gamma} P_{\Gamma}(\mathbf{q}^N, \Gamma^N) W_{\Gamma}(\mathbf{q}^N, \Gamma^N) \right\rangle \quad (27)$$

where the angular brackets denote the ensemble average over all configurations of the system  $\{\mathbf{q}^N, \Gamma^N\}$  of the 'solvent'. Note that the test polymer does not form part of the  $N$ -particle system. Therefore the probability to find the remaining particles in a configuration  $\{\mathbf{q}^N, \Gamma^N\}$  does not depend on the conformation  $\Gamma$  of the polymer.

In order to simplify the expression for the average in eq. 27, we first consider the average of the Rosenbluth factor for a given configuration  $\{\mathbf{q}^N, \Gamma^N\}$  of the solvent.

$$\overline{W}(\{\mathbf{q}^N, \Gamma^N\}) = \sum_{\Gamma} P_{\Gamma}(\mathbf{q}^N) W_{\Gamma}(\{\mathbf{q}^N, \Gamma^N\}), \quad (28)$$

Substitution of equations (25) and (26) yields

$$\overline{W} = \sum_{\Gamma} \left[ \prod_{i=1}^{\ell} \frac{\exp \left[ -\beta u^{(i)}(\Gamma_i) \right]}{Z_i} \right] \left[ \prod_{i=1}^{\ell} \frac{Z_i}{b} \right]$$

$$\begin{aligned}
&= \sum_{\Gamma} \prod_{i=1}^{\ell} \frac{1}{b} \exp \left[ -\beta u^{(i)}(\Gamma_i) \right] \\
&= \sum_{\Gamma} \frac{1}{b^{\ell}} \exp \left[ -\beta U_{\Gamma} \right], \tag{29}
\end{aligned}$$

where we have dropped all explicit reference to the solvent coordinates  $\{\mathbf{q}^N, \Gamma^N\}$ . Note that eq. 29 can be interpreted as an average over all *ideal* chain conformations of the Boltzmann factor  $\exp[-\beta U_{\Gamma}]$ . If we now substitute eq. 29 in eq. 28 we obtain

$$\langle W \rangle = \frac{\sum_{\Gamma} \langle \exp[-\beta \Delta U(\mathbf{q}^N, \Gamma^N; \mathbf{q}_{N+1}, \Gamma_{N+1})] \rangle}{\sum_{\Gamma}} \tag{30}$$

If we compare eq. 30 with eq. 24, we see that the ensemble average of the Rosenbluth factor is directly related to the excess chemical potential of the chain molecule,

$$\beta \mu_{ex} = -k_B T \ln \langle W \rangle. \tag{31}$$

This completes our demonstration that a measurement of the average Rosenbluth factor of a trial chain can indeed be used to estimate the excess chemical potential of a polymer in a dense fluid.

I should stress that the above method to measure the chemical potential is in no way limited to chain molecules on a lattice. What *is* essential is that the number of possible directions for each segment ( $b$ ) relative to the previous one is finite. The special case of chain molecules on a lattice was studied by Mooij and Frenkel [80, 81] who found that, using the techniques that are described in section 5.2, the calculation of the chemical potential can even be made 2-3 orders of magnitude more efficient than the original scheme of refs. [73, 74] which, in turn, is many orders of magnitude more efficient than the ‘brute force’ particle insertion method.

### 2.7.2. *Extension to continuously deformable molecules*

The numerical computation of the (excess) chemical potential of a flexible chain (with or without elastic forces that counteract bending), is rather different from the corresponding calculation for a chain molecule that has a large but fixed number of undeformable conformations. Yet this technique is of great practical use. For instance, Smit *et al.* have used it to study the adsorption of alkanes in zeolites with an accuracy that was hitherto impossible to achieve [77, 78].

Below, I consider the case of a flexible molecule *with* internal energy. Fully flexible chains are, of course, included as a special case. Consider a semi-flexible chain of  $\ell$  linear segments. The potential energy of a given conformation has two contributions:

1. The internal potential energy  $U_{int}$  is equal to the sum of the contributions of the individual joints. A joint between segments  $i$  and  $i + 1$  (say) has a potential energy  $u(\theta_i)$  that depends on the angle  $\theta_i$  between the successive segments. For instance,  $u(\theta_i)$  could be of the form  $u(\theta) = \alpha\theta^2$ . For realistic models for poly-atomic molecules,  $U_{int}$  would account for all local internal potential energy changes due to bending and torsion.
2. The 'external' potential energy  $U_{ext}$ . This energy accounts for all interactions with other molecules *and for the non-bonded intra-molecular interactions*. In addition, interactions with any external field that may be present are also included in  $U_{ext}$ .

In what follows I denote the chain in the absence of the 'external' interactions as the *ideal* chain. Clearly, the conformational partition function of the ideal chain is equal to

$$Z_{id} = c \int \cdots \int d\Gamma_1 \cdots d\Gamma_\ell \prod_{i=1}^{\ell} \exp(-\beta u_{id}(\theta_i)) \quad (32)$$

where  $c$  is a numerical constant. Our aim is to compute the effect of the external interactions on the conformational partition function. Hence, we wish to evaluate  $Z/Z_{id}$ , where  $Z$  denotes the partition function of the interacting chain. The excess chemical potential of the interacting chain is given by

$$\mu_{ex} = -k_B T \ln(Z/Z_{id}).$$

Before considering the 'smart' approach to compute  $\mu_{ex}$ , let us briefly review two not-so-smart methods.

The most naive way to compute the excess chemical potential of the interacting chain is to generate a very large number of completely random conformations of the freely jointed chain. For every conformation we compute both  $\exp(-\beta U_{int})$  and  $\exp(-\beta[U_{int} + U_{ext}])$ . The average of the former quantity is proportional to  $Z_{id}$ , while the average of the latter Boltzmann-factor is proportional to  $Z$ . The ratio of these two averages should therefore yield  $Z/Z_{id}$ .

The problem with this approach is that the overwhelming majority of randomly generated conformations correspond to semi-flexible chains with a very high internal energy (and therefore a very small Boltzmann weight). Hence the statistical accuracy of this sampling scheme will be very poor.

The second scheme is designed to alleviate this problem. Rather than generating conformations of a freely jointed chain, we now sample the internal angles in the chain in such a way that the probability of finding a

given angle  $\theta_i$ ; is given by the Boltzmann weight

$$P(\theta_i) = \frac{\exp(-\beta u(\theta_i))}{\int d\Gamma_i \exp(-\beta u(\theta_i))}.$$

Such sampling can be performed quite easily using a rejection method (see *e.g.* [82]). In what follows, I use the symbol  $\Gamma_i$  to denote the unit vector that specifies the orientation of the  $i$ -th segment of the chain molecule. For every conformation thus generated, we compute the Boltzmann factor  $\exp(-\beta U_{ext})$ . The average of this Boltzmann weight is then equal to

$$\begin{aligned} \langle \exp(-\beta U_{ext}) \rangle &= \frac{\int \prod d\Gamma \exp(-\beta[U_{id} + U_{ext}])}{\int \prod d\Gamma \exp(-\beta U_{id})} \\ &= Z/Z_{id} \end{aligned} \quad (33)$$

This approach is obviously superior to the first scheme. However, in many practical situations it will still yield poor statistics, because most ideal chain conformations will not correspond to energetically favourable situations for the interacting chain. Hence the Boltzmann weights will, again, be small for most conformations and the statistical accuracy will not be very good.

The problem with both schemes described above, is that neither allows us to focus on those conformations that should contribute most to  $Z$ , namely the ones for which the sum of the internal and external potential energies is not much larger than a few  $k_B T$  per degree of freedom. It would clearly be desirable to *bias* the sampling towards such favourable conformations.

One might guess that we can use a procedure similar to the one used in section 2.7.1 to compute the excess chemical potential of a chain molecule with many fixed conformations. Yet, there is an important difference precisely because the number of conformations is now, in principle, infinite. We can never hope to sample over *all* possible orientations of a new segment as we grow a chain. The next guess might be that we should then use a large but finite number of trial directions. Surprisingly, this is not necessary. It is possible to devise a *rigorous* algorithm to measure the excess chemical potential of a continuously deformable chain using an *arbitrary subset* of all possible trial directions. The answer that we get does *not* depend on the number of trial directions that we choose, but the statistical accuracy does. Below I describe this sampling scheme and show that it does indeed yield the desired average.

To compute  $\mu_{ex}$ , we apply the following 'recipe' to construct a conformation of a chain of  $\ell$  segments. The construction of chain conformations

proceeds segment by segment. Let us consider the addition of one such segment. To be specific, let us assume that we have already grown  $i$  segments, and that we are trying to add segment  $i + 1$ . This is done as follows:

1. Generate a fixed number (say  $b$ ) trial segments with orientations distributed according to the Boltzmann weight associated with the internal energy  $u(\theta)$ . We denote the different trial segments by indices  $1, 2, \dots, b$ .
2. For all  $b$  trial segments, we compute the 'external' Boltzmann factor  $\exp(-\beta u_{ext}(j))$ .
3. Select one of the trial segments, say  $j$ , with a probability

$$P_j = \frac{\exp(-\beta u_{ext}(j))}{Z_i},$$

where we have defined

$$Z_i \equiv \sum_{j'=1}^b \exp(-\beta u_{ext}(j')).$$

4. Add this segment as segment  $i + 1$  to the chain and store the corresponding partial Rosenbluth weight  $w_i = Z_i/b$ .

The desired ratio  $Z/Z_{id}$  is then equal to the average value (over many trial chains) of the product of the partial Rosenbluth weights:

$$Z/Z_{id} = \langle \prod_{i=1}^{\ell} w_i \rangle \quad (34)$$

The advantage of this scheme is that step 3 biases the sampling towards energetically favourable conformations.

To show that equation 34 is, in fact, correct, let us consider the probability with which we generate a given chain conformation. This probability is the product of a number of factors. Let us first consider these factors for one segment, and then later extend the result to the complete chain. The probability to generate a given set of  $b$  trial segments with orientations  $\Gamma_1$  through  $\Gamma_b$  is

$$P_{id}(\Gamma_1) \cdot P_{id}(\Gamma_2) \cdots P_{id}(\Gamma_b) d\Gamma_1 \cdots d\Gamma_b \quad (35)$$

The probability of selecting any one of these trial segments, say segment  $j$ , is

$$P_{sel}(j) \equiv \frac{\exp(-\beta u_{ext}(j))}{Z_i} \quad (36)$$

We wish to compute the average of a quantity, say  $w$ , over all possible sets of trial segments and all possible choices of the segment. To this end, we

must sum over all  $j$  and integrate over all orientations  $\prod_{j'=1}^b d\Gamma_{j'}$  (i.e, we average over the normalized probability distribution for the orientation of segment  $i + 1$ ):

$$\langle w \rangle = \int \left[ \prod_{j'=1}^b d\Gamma_{j'} P_{id}(\Gamma_{j'}) \right] \sum_{j''=1}^b \frac{\exp(-\beta u_{ext}(j''))}{Z_i} w(1, 2, \dots, b). \quad (37)$$

Now we make use of the fact that  $w_i(1, 2, \dots, b)$  is equal to  $Z_i/b$  (see step 4 of the 'recipe' above). Inserting this expression in Eq. 37, we obtain:

$$\langle w \rangle = \int \left[ \prod_{j'=1}^b d\Gamma_{j'} P_{id}(\Gamma_{j'}) \right] \sum_{j''=1}^b \frac{\exp(-\beta u_{ext}(j''))}{b}. \quad (38)$$

But the labeling of the trial segments is arbitrary. Hence, all  $b$  terms in the sum in Eq. 38 yield the same contribution, and Eq. 38 simplifies to

$$\langle w \rangle = \int d\Gamma P_{id}(\Gamma) \exp(-\beta u_{ext}(\Gamma)) \quad (39)$$

$$= \frac{\int d\Gamma \exp(-\beta[u_{id}(\Gamma) + u_{ext}(\Gamma)])}{\int d\Gamma \exp(-\beta u_{id}(\Gamma))} \quad (40)$$

$$= \frac{Z^{(1)}}{Z_{id}^{(1)}}, \quad (41)$$

which is indeed the desired result, but for the fact that the expression in Eq. 41 refers to 1 segment (as indicated by the superscript in  $Z^{(1)}$ ). The extension to a chain of  $\ell$  segments is straightforward, be it that the intermediate expressions become a little unwieldy.

### 2.7.3. Overlapping-distribution Rosenbluth method

Although the Rosenbluth particle-insertion scheme described in section 2.7 is correct in principle, it may run into practical problems when the excess chemical potential becomes large. The reason why the method breaks down is a direct consequence of the fact that the chain conformations generated by the Rosenbluth algorithm have a non-Boltzmann distribution. In fact, as was already indicated by the Rosenbluths [69] one can correct for this bias by attributing to every generated conformation a statistical weight proportional to  $w$ . Typically, compact chains have a low Rosenbluth weight (because the fraction of 'open' trial directions is low), while more expanded chains have a higher Rosenbluth weight. Hence, the fact that we must use the Rosenbluth weight  $w$  to compensate for the bias in our distribution of chain conformations indicates that the Rosenbluth scheme generates chain conformations that are too compact. Multiplying by  $w$  supposedly generates

the unbiased distribution. However, as was already discussed in great detail by Batoulis and Kremer [83], this correction will only work if we have good statistics on the resulting distribution. But for long chains, or chains that are inserted in dense media, the Rosenbluth scheme will not generate a significant number of chain conformations in the relevant range of  $w$ . Hence the Rosenbluth scheme will run into trouble under those conditions.

Fortunately, it is possible to combine the Rosenbluth scheme with the overlapping-distribution method to obtain a technique with built-in diagnostics. This scheme is explained in appendix A.

As with the original overlapping-distribution method, the scheme described in appendix A constructs two histograms (but now as a function of the logarithm of the Rosenbluth weight, rather than the potential energy difference). If the sampled distributions do not overlap, then one should expect the estimate of the excess chemical potential of chain molecules to become unreliable and the Rosenbluth method should not be used. In fact, recent simulations by Mooij and Frenkel [84] using this overlapping distribution method, show that indeed there is a tendency for the two distributions to move apart when long chains are inserted into a moderately dense fluid. Yet, these simulations also show that, at least in this case, the statistical errors in  $\mu_{ex}$  become important before the systematic errors due to inadequate sampling show up.

#### 2.7.4. Recursive sampling

In view of the preceding discussion, it would seem attractive to have unbiased sampling schemes to measure the chemical potential. Of course, ‘thermodynamic integration methods’ [70] are unbiased and the modified Widom scheme [71], although biased at the level of the insertion of a single monomer (like the original Widom scheme), is less biased than the Rosenbluth method. Yet, these method cannot be used to measure  $\mu_{ex}$  in a single simulation.

Surprisingly, recent work of Grassberger and Hegger [85, 86, 87] has shown that it *is* possible to perform unbiased sampling of  $\mu_{ex}$  in a single simulation. Below, I briefly sketch the basic idea behind this method.

Like the Rosenbluth and modified Widom schemes, the recursive sampling approach is based on a segment-by-segment growth of the polymer. But that is about where the similarity ends. In recursive sampling the aim is to generate a *population* of trial conformations. The excess chemical potential of a chain molecule is directly related to the average number of molecules that have survived the growth process.

The first step of the procedure is to attempt a trial insertion of a monomer in the system. Suppose that the Boltzmann factor associated with this trial insertion is  $b_0 \equiv \exp(-\beta u_0(\mathbf{r}^N))$ . We now allow the monomer

to make multiple copies of itself, such that the average number of copies,  $\langle n_0 \rangle$ , is equal to

$$\langle n_0 \rangle = \pi_0 b_0$$

where  $\pi_0$  is a constant multiplicative factor that remains to be specified. A convenient rule to determine how many copies should be made is the following. Denote the fractional part of  $\pi_0 b_0$  by  $f_0$  and the integer part by  $i_0$ . Our rule is then to generate  $i_0$  ( $i_0 + 1$ ) copies of the inserted particle with a probability  $1 - f_0$  ( $f_0$ ). Clearly if  $i_0 = 0$ , there is a probability  $1 - f_0$  that the monomer will 'die'. Assume that we have generated at least 1 copy of the monomer. Every copy from now on proceeds independently to generate off-spring. For instance, to generate a dimer population, we add a segment to every surviving monomer. We denote the Boltzmann weight associated with these trial additions by  $b_1(i)$ , where the index  $i$  indicates that every surviving monomer will give rise to a different dimer. As before, we have to decide how many copies of the dimers should survive. This is done in exactly the same way as for the monomer, *i.e.* the average number of dimers that descends from monomer  $i$  is given by

$$\langle n_1(i) \rangle = \pi_1 b_1(i)$$

where  $\pi_1$ , just like  $\pi_0$  before, is a constant to be specified later. The number of dimers generated may either be larger or smaller than the original number of monomers. We now proceed with the same recipe for the next generation (trimers) and so on. In fact, as with the semi-flexible molecules discussed in section 2.7.1, it is convenient to include the intra-molecular bond-bending, bond-stretching and torsional energies in the probability distribution that determines with what orientation new segments should be added.

The average number of surviving molecules at the end of the  $\ell$ -th step is equal to

$$\langle N_\ell \rangle = \left( \prod_{i=0}^{\ell} \pi_i \right) \langle \exp(-\beta U_\ell(\mathbf{r}^N)) \rangle$$

where  $U_\ell(\mathbf{r}^N)$  is the total interaction of the chain molecule with the  $N$  'solvent' molecules (and the non-bonded intra-molecular interactions). The angular brackets denote a canonical average over the coordinates and over the intra-molecular Boltzmann factors of the ideal (non-self-avoiding) chain. In other words,

$$\langle N_\ell \rangle = \left( \prod_{i=0}^{\ell} \pi_i \right) \exp(-\beta \mu_{ex}(\ell))$$

Hence, the excess chemical potential is given by

$$\mu_{ex}(\ell) = -k_B T \ln \left( \frac{\langle N_\ell \rangle}{\prod_{i=0}^{\ell} \pi_i} \right). \quad (42)$$

The constants  $\pi_i$  should be chosen such that there is no population explosion, nor mass extinction. If we have a good guess for  $\mu_{ex}(\ell)$  then we can use this to estimate  $\pi_i$ . In general, however,  $\pi_i$  must be determined by trial and error.

This recursive algorithm has several nice features. First of all, it is computationally quite efficient (in some cases, more than an order of magnitude faster than the Rosenbluth scheme, for the same statistical accuracy). In fact, in actual calculations, the algorithm searches in *depth* first, rather than in *breadth*. That is to say: we try to grow a polymer until it has been completed (or has died). We then continue from the last ‘branch’ of the tree from where we are allowed to grow another trial conformation. In this way we work our way back to the ‘root’ of the tree. The advantage of this scheme is that the memory requirements are minimal. Moreover, the structure of the program is very simple indeed (in languages that allow recursive calls). Last but not least, the recursive scheme generates an *unbiased* (*i.e.* Boltzmann) population of chain conformations. This is shown in Appendix B.

### 3. Other free energy methods

The present section deals with free-energy methods that are of particular importance for ‘barrier-crossing’ problems. By this I mean, either the computation of the free energy barrier between two states (phases, conformations) or the computation of the free energy of a system that cannot be equilibrated by conventional means because its free energy landscape consists of many valleys separated by high barriers. The numerical techniques described in this section are therefore not exclusively designed for ‘free-energy’ calculations but also for equilibration of ‘glassy’ systems. This is a very frequent problem in complex liquids and deserves therefore special attention. I start, however, with a simple extension of the overlapping-distribution method of section 2.5.2. In that section, I focused on the application of the technique to compute the free-energy *difference* between two systems. Now I will discuss the method, in its generalized form, in a different context: namely as a technique to study high free-energy barriers. Such barriers cannot be studied using conventional MC or MD simulations because the probability that a spontaneous fluctuation will bring the system to the top of the barrier is vanishingly small.

#### 3.1. MULTIPLE HISTOGRAMS

As explained in section 2.5.2, for the overlapping distribution-function method to work, it is not strictly necessary that the two distributions overlap. However, they should not be too far apart. In case there is a large

gap between  $f_0$  and  $f_1$ , it is often useful to perform additional simulations for intermediate values of the Hamiltonian. We thus obtain a sequence of distribution functions  $f_0, f_1, f_2, \dots, f_n$ , such that  $f_0$  overlaps with  $f_1$ ,  $f_1$  with  $f_2$ , etc. The free energy difference between system 0 and system  $n$  can then be obtained by adding  $\Delta F_{0,1} + \Delta F_{1,2} + \dots + \Delta F_{n-1,n}$ . The problem with a naive implementation of this approach is that the statistical errors in the individual free-energy differences  $\Delta F_{i,i+1}$  add up quadratically in the final result. Fortunately, there exist techniques that prevent such error propagation. One approach, proposed by Ferrenberg and Swendsen [68] is based on the idea that it is possible to construct a self-consistent estimate for the histograms without assuming any specific functional form.

For what follows, it is convenient to define the distributions that we are considering in a slightly more general form. We assume that the original system that we are studying is characterized by a potential energy function  $U_0$ . We now define a sequence of  $n$  related models that are characterized by a potential energy function  $U_i \equiv U_0 + W_i$  ( $i = 1, \dots, n$ ). For example,  $W_i$  could be  $\lambda_i(U_1 - U_0)$ , with  $0 < \lambda_i \leq 1$ . Alternatively,  $W_i$  could be a function of some order parameter or reaction coordinate  $Q(\mathbf{r}^N)$ . For the sake of generality, we take the latter point of view, because  $W_i = \lambda_i(U_1 - U_0)$  can be interpreted as a function of the 'order parameter'  $Q(\mathbf{r}^N) \equiv U_1 - U_0$ . We now consider histograms  $p_i(Q)$  defined as follows

$$p_i(Q) = \frac{\int \exp(-\beta(U_0 + W_i))\delta(Q - Q(\mathbf{r}^N))d\mathbf{r}^N}{Z_i},$$

where  $Z_i$  is given by

$$Z_i \equiv \int \exp(-\beta(U_0 + W_i))d\mathbf{r}^N. \quad (43)$$

The original system is characterized by an order-parameter distribution

$$p_0(Q) = \frac{\int \exp(-\beta U_0)\delta(Q - Q(\mathbf{r}^N))d\mathbf{r}^N}{Z_0},$$

In some case, we are interested in the ratio  $Z_0/Z_n$ , because that determines the free-energy difference between systems 0 and  $n$ . But often it is the distribution function  $p_0(Q)$  itself that is of central interest, because it allows us to compute the Landau free energy of the original system, as a function of the order parameter  $Q$ :

$$F_{\text{Landau}}(Q) = -k_B T \ln p_0(Q).$$

### 3.1.1. Self-consistent histogram method

In a simulation, the histograms  $p_i(Q)$  are estimated by measuring  $H_i(Q)$ , the number of times that a system with potential energy function  $U_0 + W_i$  has a value of the order parameter between  $Q$  and  $Q + \Delta Q$ . If the total number of points collected in histogram  $i$  is denoted by  $M_i$ , then

$$p_i(Q)\Delta Q = \langle H_i(Q) \rangle / M_i ,$$

where the  $\langle \dots \rangle$  denote ensemble averaging. Of course, in a simulation of finite length, the ratio  $H_i(Q)/M_i$  will fluctuate around  $p_i(Q)\Delta Q$ . If we assume that the number of points in bin  $i$  is determined by a Poisson distribution, then it follows that the variance in  $p_i^{est}(Q)\Delta Q$ , *i.e.* our estimate of  $p_i(Q)\Delta Q$ , is given by

$$\begin{aligned} \frac{\langle H_i^2(Q) \rangle - \langle H_i(Q) \rangle^2}{M_i^2} &= \frac{\langle H_i(Q) \rangle}{M_i^2} \\ &= \frac{p_i(Q)\Delta Q}{M_i} . \end{aligned} \quad (44)$$

In what follows, we shall assume, without loss of generality, that our units are chosen such that  $\Delta Q = 1$ . This choice is, of course, not essential, but it simplifies the notation and does not affect the final results. Once we have measured a set of histograms, we can try to combine this information to arrive at an estimate of  $p_0(Q)$  or, equivalently, of  $F_{\text{Landau}}(Q)$ . First of all, we should note that in principle, although not in practice,  $p_0$  can be reconstructed from every individual histogram  $p_i(Q)$

$$p_0(Q) = \exp(+\beta W_i) \frac{Z_i}{Z_0} p_i(Q) . \quad (45)$$

In practice, this approach will hardly ever work because the range of  $Q$  values where  $p_0(Q)$  and  $p_i(Q)$  differ significantly from zero need not overlap. Hence the most important contribution to  $p_0(Q)$  would come from a range of  $Q$  values where  $\exp(+\beta W_i)$  is very large, but  $p_i(Q)$  is vanishingly small. We shall therefore attempt to construct our estimate  $p_0^{est}(Q)$  by a linear combination of the estimates based on the different histograms:

$$p_0^{est}(Q) = \sum_{i=1}^n w_i(Q) \exp(+\beta W_i) \frac{Z_i}{Z_0} p_i^{est}(Q) , \quad (46)$$

where  $w_i(Q)$  is an, as yet undetermined weight function, subject to the condition

$$\sum_{i=1}^n w_i(Q) = 1 . \quad (47)$$

Note that, at this stage, the  $Z_i/Z_0$  are also still unknown. Let us now choose the weights  $w_i(Q)$  such that the estimated variance in  $p_0^{est}(Q)$  is minimized. Using the fact that fluctuations in different simulations are uncorrelated, this variance is given by

$$\begin{aligned}
 & \langle (p_0^{est}(Q))^2 \rangle - \langle p_0^{est}(Q) \rangle^2 & (48) \\
 &= \sum_{i=1}^n w_i^2(Q) \exp(+2\beta W_i) \left(\frac{Z_i}{Z_0}\right)^2 \langle (p_i^{est}(Q))^2 \rangle - \langle p_i^{est}(Q) \rangle^2 \\
 &= \sum_{i=1}^n w_i^2(Q) \exp(+2\beta W_i) \left(\frac{Z_i}{Z_0}\right)^2 p_i(Q)/M_i \\
 &= p_0(Q) \sum_{i=1}^n w_i^2(Q) \exp(+\beta W_i) \frac{Z_i}{Z_0} / M_i,
 \end{aligned}$$

where, in the last line, we have used Eqns.44 and 45. Next we must determine the values for  $w_i(Q)$  that minimize the variance of  $p_0^{est}(Q)$ . Differentiating Eqn.44, under the constraint given by Eqn.47, yields

$$w_i(Q) = \alpha \exp(-\beta W_i) M_i \frac{Z_0}{Z_i}, \quad (49)$$

where  $\alpha$  is an undetermined multiplier, fixed by the condition that  $p_0^{est}(Q)$  is normalized.

$$\alpha = \frac{1}{\sum_{i=1}^n \exp(-\beta W_i) M_i Z_0 / Z_i}, \quad (50)$$

Inserting Eqn.49 in Eqn.46 yields

$$p_0^{est}(Q) = \frac{\sum_{i=1}^n H_i(Q)}{\sum_{i=1}^n \exp(-\beta W_i) M_i Z_0 / Z_i}. \quad (51)$$

Finally, we must determine the  $Z_i$ . We do this by inserting Eqns. 51 and 45 in the definition of  $Z_i$  (Eqn. 43).

$$\begin{aligned}
 Z_i &= \int \exp(-\beta(U_0 + W_i)) d\mathbf{r}^N \\
 &= \int Z_0 p_0(Q) \exp(-\beta W_i) dQ \\
 &= \int dQ \exp(-\beta W_i) \frac{\sum_{j=1}^n H_j(Q)}{\sum_{k=1}^n \exp(-\beta W_k) M_k / Z_k}.
 \end{aligned} \quad (52)$$

This is an implicit equation for the  $Z_i$  that must be solved self-consistently. In fact, we cannot determine the absolute value of all  $Z_i$  but only their ratio. Therefore, we can arbitrarily fix one of the  $Z_i$  (e.g.  $Z_1$ ) at a constant value.

The (non-linear) set of equations 52 is then solved for the remaining  $Z_i$  until self-consistency is reached. The free energy difference between system  $n$  and system 1, can then simply be computed as

$$\Delta F = -k_B T \ln(Z_n/Z_1) .$$

To give a specific example, assume that we have simulated only two systems 1 and 2, with potential energy functions  $U_1$  and  $U_2 = U_1 + \Delta U$ . Moreover, let us assume for the sake of simplicity, that both simulations resulted in the same number of histogram entries ( $M_1 = M_2$ ). Then Eqn. 52 becomes

$$\begin{aligned} Z_2 &= \int d\Delta U \exp(-\beta\Delta U) \frac{H_1(\Delta U) + H_2(\Delta U)}{M(1/Z_1 + \exp(-\beta\Delta U)/Z_2)} \\ \frac{Z_2}{Z_1} &= \int d\Delta U \exp(-\beta\Delta U) \frac{p_1(\Delta U) + p_2(\Delta U)}{(1 + \exp(-\beta\Delta U))(Z_1/Z_2)} . \end{aligned} \quad (53)$$

Using  $\beta\Delta F = -\ln(Z_2/Z_1)$  we can rewrite Eqn. 53 as

$$\int d\Delta U \frac{p_1(\Delta U)}{1 + \exp(-\beta(\Delta U - \Delta F))} = \int d\Delta U \frac{p_2(\Delta U)}{1 + \exp(-\beta(\Delta F - \Delta U))} ,$$

which is equivalent to the equation that must be solved self-consistently for  $\Delta F$  in Bennett's acceptance-ratio method (ref. [67]). This is somewhat surprising because the acceptance ratio method was devised to minimize the estimated error in the free-energy difference between two systems. In contrast, Eqn.52 minimizes the error in our estimate of  $p_0(Q)$ . If the number of histograms is larger than two, this expression is not completely equivalent to the one obtained by minimizing the error in the free-energy difference between system 1 and  $n$  (say). But it is straightforward to derive the equations for the set  $\{Z_i\}$  that minimize the error in the estimate of a particular free-energy difference.

### 3.2. UMBRELLA SAMPLING

In section 2.5.2, we introduced the distribution functions  $p_0(\Delta U)$  and  $p_1(\Delta U)$  that measure the probability of finding system 0 (1) in an equilibrium configuration  $\mathbf{s}^N$  where the potential energy of system 1 and 0 differ by an amount  $\Delta U$ . At first sight it would seem that knowledge of either  $p_0$  or  $p_1$  is sufficient to estimate the free energy difference between systems 0 and 1. After all, eqn. 20 states that

$$p_1(\Delta U) = p_0(\Delta U) \exp(\beta(\Delta F - \Delta U)) .$$

If we integrate over  $\Delta U$  on both sides of this equation, we obtain:

$$\begin{aligned} \int_{-\infty}^{\infty} p_1(\Delta U) d\Delta U &= \exp(\beta\Delta F) \int_{-\infty}^{\infty} p_0(\Delta U) \exp(-\beta\Delta U) d\Delta U \\ 1 &= \exp(\beta\Delta F) < \exp(-\beta\Delta U) >_0 , \end{aligned} \quad (54)$$

or

$$\exp(-\beta\Delta F) = \langle \exp(-\beta\Delta U) \rangle_0 .$$

Although eqn. 54 is very useful to estimate free energy differences between two systems that are not too dissimilar, its applicability is limited. The problem is that, in many cases of practical interest, the largest contributions to the average  $\langle \exp(-\beta\Delta U) \rangle_0$  come from region of configuration space where  $p_0(\Delta U)$  is very small, while  $\exp(-\beta\Delta U)$  is very large. As a result, the statistical error in  $\Delta F$  is large.

One method to achieve a more accurate estimate of  $\Delta F$  is the ‘umbrella sampling’ scheme suggested by Torrie and Valleau [88]. The basic idea behind umbrella sampling is that, in order to obtain an accurate estimate of the free energy difference between two system (0 and 1), one should sample both the part of configuration space that is accessible to system 1 and the part that is accessible to 0. In order to achieve this in a single simulation, one should modify the Markov chain that governs the sampling of configuration space. This is achieved by replacing the Boltzmann factor of the system by a non-negative weight function  $\pi(\mathbf{q}^N)$ . As a result, the probability of visiting a point  $\mathbf{q}^N$  in configuration space is now proportional to  $\pi(\mathbf{q}^N)$ . The expression for  $\langle \exp(-\beta\Delta U) \rangle_0$  becomes:

$$\langle \exp(-\beta\Delta U) \rangle_0 = \frac{\int d\mathbf{q}^N \pi(\mathbf{q}^N) (\exp(-\beta U_1(\mathbf{q}^N))/\pi(\mathbf{q}^N))}{\int d\mathbf{q}^N \pi(\mathbf{q}^N) (\exp(-\beta U_0(\mathbf{q}^N))/\pi(\mathbf{q}^N))} , \quad (55)$$

or, introducing the notation  $\langle \dots \rangle_\pi$  to denote an average over a probability distribution proportional to  $\pi(\mathbf{q}^N)$ :

$$\langle \exp(-\beta\Delta U) \rangle_0 = \frac{\langle \exp(-\beta U_1)/\pi \rangle_\pi}{\langle \exp(-\beta U_0)/\pi \rangle_\pi} \quad (56)$$

In order for both the numerator and the denominator in this equation to be non-zero,  $\pi(\mathbf{q}^N)$  should have an appreciable overlap with both the regions of configuration space that are sampled by system 0 and system 1. This ‘bridging’ property of  $\pi$  is responsible for the name ‘umbrella’ sampling.

Although umbrella sampling is, in principle, a powerful technique, one drawback is that the function  $\pi$  is not known *a priori*. Rather, it must be constructed using the available information about the Boltzmann weights of systems 0 and 1. Constructing a good sampling distribution used to require skill and patience. However, during the past few years there have been several developments that should take some of the black magic out of umbrella sampling. In fact, there have been a fairly large number of independent papers that arrive at similar results. Below, I briefly review some of the more recent developments.

### 3.2.1. Recent developments

First of all, Valleau [89, 90, 91] has proposed a systematic scheme to compute the weight function  $\pi$  that is needed if umbrella sampling is used to compute the properties of a system over a wide range of densities and temperatures. This problem was already considered in 1967 by McDonald and Singer [92, 93, 94] who developed a pre-cursor of the umbrella-sampling scheme.

In the case of temperature scaling alone, the umbrella sampling scheme is fairly straightforward. Suppose that we wish to gather statistics on the configurational properties of a system in a temperature range between upper and lower limits  $T_U$  and  $T_L$ . We can then use as our sampling function a linear combination of Boltzmann weights at a discrete set of points in the temperature interval of interest:

$$\pi(\mathbf{q}^N) = \sum_{i=i_L}^{i_U} w_i \exp(-\beta_i U(\mathbf{q}^N)) \quad (57)$$

where  $\beta_i = 1/k_B T_i$  and  $w_i$  is a non-negative weight that remains to be specified. In order to get the sampling as uniform as possible, a logical choice for  $w_i$  is

$$w_i \sim \exp[+\beta_i F_{ex}(T_i, V)]$$

because this choice insures that every term of Eqn. 57 yields the same contribution to the configuration-space integral of  $\pi$ . Of course, we do not know  $F_{ex}(T_i, V)$  a priori and hence we have to make a reasonable guess (on basis of the equation of state). But this is not particularly difficult.

Things become more subtle when we try to apply the same approach to include sampling over a range of densities. At first sight it would seem that we can almost use the same expression as in Eqn. 57 (but for the fact that we have to use scaled coordinates,  $\mathbf{s} = \mathbf{q}/V^{1/d}$ ). The weight function  $\pi$  would then become

$$\pi(\mathbf{s}^N) = \sum_{i=i_L, j=j_L}^{i_U, j_U} w_{i,j} \exp(-\beta_i U([V_j^{1/d} \mathbf{s}]^N)) \quad (58)$$

where  $V_j$  is a volume between the limits  $V_U$  and  $V_L$ . Again, we could choose

$$w_{i,j} \sim \exp[+\beta_i F_{ex}(T_i, V_j)] .$$

But this will *not* work. The reason is that the sampling function is 'too tolerant' in the sense that it allows particles to approach one another up to a distance that corresponds to a 'hard-core' diameter in the system with the lowest density. Let us call this distance  $s_{min} = \sigma/(V_U^{1/d})$ . If we sample

configuration space with the function given in Eqn. 58 then, for all but the smallest systems, there will be at least one pair of particles with a separation near  $s_{min}$ . For systems with a volume  $V < V_U$ , this separation will correspond to a very high energy configuration because the real pair distance would be:  $r_{min} = \sigma(V/V_U)^{1/d}$ . As a consequence, our distribution  $\pi$  is not at all a bridging distribution. Valleau [89] has suggested to take care of this problem by making the weight function  $\pi(\mathbf{q}^N)$  depend explicitly on the shortest pair distance in the system. Although this fix works for the example discussed by Valleau, it is case-specific. By this I mean that if our umbrella sampling had been designed to span another class of Hamiltonians (not necessarily temperature- or density-scaled), then we might run into a similar problem and we would again need to *think*. The essence of a general-purpose computational tool is that one can use it with a minimum of thought.

Below, I describe a class of algorithms that achieve essentially the same as umbrella sampling, but that are more robust (*i.e.* require less thought).

### 3.2.2. Parameter hopping

Let us consider the general case of a sequence of Hamiltonians  $H_1, H_2, \dots, H_n$ . The umbrella sampling weight function would be of the form

$$\pi(\mathbf{q}^N) = \sum_{i=1}^n w_i \exp(-H_i(\mathbf{q}^N)) \quad (59)$$

For convenience, we have absorbed the temperature in the Hamiltonian. Recall that the bridging distribution  $\pi$  in umbrella sampling is ideally designed such that every term in the sum 59 contributes equally to the integral of  $\pi$ . There is another way in which we can achieve the same result. Rather than sampling over  $\mathbf{q}^N$  with a fixed distribution function  $\pi$ , we perform Boltzmann sampling of the system with a given Hamiltonian  $H_i$ . But unlike in normal Monte Carlo, we also allow moves where the system changes its Hamiltonian. The probability of acceptance of such a move is

$$P_{acc}(H_i \rightarrow H_j) = \min \left( 1, \frac{w_j \exp(-H_j(\mathbf{q}^N))}{w_i \exp(-H_i(\mathbf{q}^N))} \right) \quad (60)$$

We adjust the transition rates for Hamiltonian hops, *i.e.* the  $w_i$ , such that every Hamiltonian is sampled about equally. Clearly, this will be achieved if  $w_i \sim \exp(+F_i)$ . Hence, as far as sampling configuration space is concerned, this sampling scheme achieves the umbrella-sampling ‘tiling’, but in a more robust way. Moreover, whereas the configurations sampled by umbrella sampling are *unphysical*, in the sense that the distribution  $\pi$  does not correspond to any real Hamiltonian, the present scheme will, at every stage, sample

according to a real Boltzmann distribution. The method has been called the method of ‘extended ensembles’ by Lyubartsev et al. [95], ‘simulated tempering’ by Marinari and Parisi [96], ‘pseudo-Bayes Metropolis-Coupled Markov Chain Monte Carlo’ by Geyer [97] (who, fortunately, decided to save paper and adopted the term ‘simulated tempering’ in a subsequent manuscript [98]). Closely related schemes had been proposed by Nezbeda and Kolafa [99] in the context of the Widom particle insertion method, while Shing and Azadipour and later Vega, Shing and Rull proposed a (mildly approximate) version of this method to grand-canonical MC [100] and MD [101] simulations.

### 3.2.3. Sampling ‘glassy’ Hamiltonians

Much of the interest in ‘parameter’ hopping is not free-energy related but derives from the fact that it promises to be a really useful tool in studying the equilibrium behavior of systems with many local free energy minima, such as (spin) glasses. A system that would be trapped in a single local minimum at low temperatures, is now able to make an ‘excursion’ to a higher temperature and come back in a completely different configuration. In fact, this aspect of the method was propagated even earlier in a slightly different (and, again, slightly approximate) manner by Frantz *et al*, who called the method ‘J-walking’ (where  $J$  stands ‘jump’) [102].

A very interesting development in the area of simulation of ‘sticky’ problems is the ‘swapping’ algorithm by Geyer [97]. In this algorithm there is a real system associated with every Hamiltonian. Occasionally, however, the systems may *swap* Hamiltonians. In this way, there will always be one system per Hamiltonian. The acceptance of swaps can be quite high. There is no need to know, or fix, the constants  $w_i$  in the acceptance rate (they are simply not there). And yet, we can compute all desired free energy differences using a multi-histogram (or related) method. An important advantage of the swapping method is that it makes it possible to suppress the slow decay of fluctuations in the glassy systems: whenever a swap is accepted, the system continues in a completely different and independent configuration. By the time the original configuration returns from its excursion, it has had the opportunity to equilibrate at a higher temperature (or, in any event, under more ‘ergodic’ conditions). This makes it possible to measure the free energy of glassy systems.

## 3.3. RARE EVENTS

The techniques described above are eminently suited to sample regions of configuration space that cannot be accessed by normal MD or MC simulations. In particular, we can use umbrella sampling in combination with

Molecular Dynamics simulations to study processes for which the rate-limiting step is a barrier crossing. The basic idea behind such calculations is that the rate at which a barrier crossing proceeds is determined by the product of a static term, namely the probability of finding the system at the top of the barrier and a dynamic term that describes the rate at which systems at the top of the barrier move to the other valley [103, 104, 105]. Let us consider a system that can exist in two states  $A$  and  $B$  that are separated by a free-energy barrier. A ‘reaction’ coordinate  $q$  connects states  $A$  and  $B$ . If we perturb the initial concentration of systems in state  $A$ , then the approach to equilibrium will depend on the rate at which the free-energy barrier is crossed. The rate constant for this barrier crossing  $k_{BA}$  is

$$k_{BA} = \lim_{t \rightarrow \infty} \frac{\langle \dot{q} \delta(q^* - q(0)) \theta(q(t) - q^*) \rangle}{\langle \theta(q^* - q) \rangle} \quad (61)$$

where  $q^*$  denotes the value of the reaction coordinate at the top of the barrier.  $\langle \theta(q^* - q) \rangle$  denotes the probability to find the system somewhere in valley  $A$  (the ‘reactant’ side, to use the language of chemical reactions). We can rephrase Eqn. 61 in a form that makes the separation in statics and dynamics more obvious. We do this by multiplying and dividing and multiplying the right hand side of Eqn. 61 by  $\langle \delta(q^* - q) \rangle$ . We then obtain

$$k_{BA} = \lim_{t \rightarrow \infty} \frac{\langle \dot{q}(0) \delta(q^* - q(0)) \theta(q(t) - q^*) \rangle}{\langle \delta(q^* - q(0)) \rangle} \times \frac{\langle \delta(q^* - q) \rangle}{\langle \theta(q^* - q) \rangle} \quad (62)$$

The second term on the right-hand side of Eq. 62, *i.e.*  $\langle \delta(q^* - q) \rangle / \langle \theta(q^* - q) \rangle$  denotes the probability density to find the system at the top of the barrier, divided by the probability that the system is on the reactant side of the barrier. This ratio, is completely determined by the height of the free-energy barrier and can therefore be measured by umbrella sampling. For instance, van Duijneveldt and Frenkel [106] have used umbrella sampling to estimate the free-energy barrier for crystal nucleation.

The first term on the left hand side of Eq. 62 is a conditional average, namely the average of the product  $\dot{q}(0) \theta(q(t) - q^*)$ , given that  $q(0) = q^*$ . We can write this term as

$$\lim_{t \rightarrow \infty} \langle \dot{q}(0) \theta(q(t) - q^*) \rangle_{q(0)=q^*} \quad (63)$$

This dynamic quantity can be computed by Molecular Dynamics. In practice, constraint MD will be used to compute this average, in which the system is prepared under the initial constraint  $q(0) = q^*$  and then released. It should be noted that the use of a constraint in the MD sampling does have a (slight) effect on the average in Eqn. 63 [104].

## 4. Smart Sampling schemes

### 4.1. INTRODUCTION

One of the main problems in the simulation of complex liquids is that the 'natural' dynamics of these systems may be quite slow. As a consequence, normal Monte Carlo or Molecular Dynamics methods may not be adequate to achieve a good sampling of configuration space. For this reason, many schemes have been considered that can lead to an enhanced equilibration. Most of these techniques rely on the freedom one has in Monte Carlo simulations to carry out unphysical moves. To give an example: in the simulation of large molecules of biological interest, it is often difficult to explore the different possible conformations of such molecules. To alleviate this problem, it has been suggested to allow the molecule to explore pathways from one conformation to another that do not *cross* potential energy barriers, but by-pass them. This is achieved by allowing the molecule to make excursions into a higher 'embedding' dimension [107].

### 4.2. HYBRID MC

In Molecular Dynamics simulations, all particle coordinates are updated simultaneously. In conventional MC simulations, only a few coordinates are changed in a trial move. As a consequence, collective molecular motions are not well represented by Monte Carlo and this may adversely affect the rate of equilibration. The advantage of Monte Carlo is that, unlike MD, we can carry out unphysical moves. Moreover, in MC the system is not constrained to move on a hyper-surface where some Hamiltonian is conserved. The time-step in Molecular Dynamics is limited by the fact that energy must be conserved. Clearly, no such constraint applies to Monte Carlo. For this reason, many authors have attempted to combine the 'natural dynamics' of MD with the large jumps in configuration space that are possible in MC. The book by Allen and Tildesley [82] describes a number of such techniques ('force-bias MC', 'Langevin dynamics', 'smart Monte Carlo') that basically work by including some or all information about the intermolecular forces in the construction of a collective MC trial move.

A technique that has achieved much attention during the past few years is the Hybrid Monte Carlo scheme [108]. The basic idea behind this scheme is that one can use Molecular Dynamics to generate Monte Carlo trial moves. At first sight, there is no advantage in doing so. However, the criteria for what constitutes a good Monte Carlo trial move are more tolerant than the specifications of a good Molecular Dynamics time step. In particular, one can take a time step that is too long for MD. Energy will not be conserved in such a trial move. However, as long as one uses an algorithm

that is time-reversible and ‘area preserving’ (*i.e.* that conserves volume in phase space), such a collective move can be used as a Monte Carlo trial move. Fortunately, there now exists a systematic way to construct time-reversible, area-preserving MD algorithms, using the multiple-timestep MD scheme of Tuckerman *et al.* [109]. The usual Metropolis algorithm can then be used to decide on the acceptance or rejection of the move (see e.g. [110, 111]). For every trial move, the particle velocities are chosen at random from a Maxwell distribution. In fact, it turns out that it is often advantageous to construct a trial move that consists of a sequence of MD steps. The reason is that, due to the randomization of the velocities, the ‘diffusion constant’ of the system becomes quite low if the velocities are randomized well before the natural decay of the velocity autocorrelation function. Yet one cannot make the time step for a single hybrid MC move too long, because then the acceptance would become very small. As a consequence, the performance of hybrid MC is not dramatically better than that of the corresponding molecular dynamics. Moreover, the acceptance probability of hybrid MC moves of constant length decreases with the system size, because the root-mean-square error in the energy increases with  $N^{1/2}$ . MD does not suffer from a similar problem. That is to say, the ‘noise’ in the total energy increases with  $N$ , but stability of the MD algorithm does not deteriorate. Hence, for very large systems, MD will always win. For more normal system sizes, hybrid MC may be advantageous. However, it is certainly a scheme that requires fine-tuning [111].

Forrest and Suter have devised a hybrid MC scheme that samples polymer conformations, using a fictitious dynamics of the generalized coordinates [112]. This scheme leads to an improved sampling of the polymer conformations compared to normal MD. However, much of the gain seems to be due to the choice of a suitable fictitious dynamics for the generalized coordinates. The gain due to the choice of hybrid MC rather than MD would seem to be less significant.

### 4.3. CONFIGURATIONAL BIAS MC

The sampling of equilibrium conformations of polymers is usually time-consuming. The main reason is that the natural dynamics of polymers is dominated by topological constraints (chains cannot cross) and hence any algorithm that is based on the real motion of macromolecules will suffer from the same problem. For this reason, many algorithms have been proposed to speed up the Monte Carlo sampling of polymer conformations (see *e.g.* ref. [113]). Here I shall focus on the Configurational Bias Monte Carlo (CBMC) scheme [73, 114, 76, 115]. This is a relatively recent simulation technique that is of particular use in cases where it is not possible

to change the conformation of a macromolecule by successive small steps. An example where such large-scale conformation changes are essential is in the simulation of phase-equilibria using the Gibbs ensemble or in grand-canonical simulations of chain molecules. However, although I introduce the scheme in the context of polymers, the CBMC method is not limited to chain molecules. This will be discussed at the end of this section.

Before explaining the CBMC scheme, it is useful to recall the general recipe to construct a Monte Carlo algorithm. One always starts (or should start) from the condition of detailed balance:

$$P_o \times P_{gen}(o \rightarrow n) \times P_{acc}(o \rightarrow n) = P_n \times P_{gen}(n \rightarrow o) \times P_{acc}(n \rightarrow o), \quad (64)$$

where  $P_o$  ( $P_n$ ) is the Boltzmann weight of the old (new) conformation,  $P_{gen}$  denotes the *a priori* probability to generate the trial move from  $o$  to  $n$ , and  $P_{acc}$  is the probability that this trial move will be accepted. From Eqn. 64 it follows that

$$\frac{P_{acc}(o \rightarrow n)}{P_{acc}(n \rightarrow o)} = \exp(-\beta\Delta U) \frac{P_{gen}(n \rightarrow o)}{P_{gen}(o \rightarrow n)}, \quad (65)$$

where  $\exp(-\beta\Delta U)$  is the ratio of the Boltzmann weights of the new and old conformations. If we use the Metropolis rule to decide on the acceptance of MC trial moves, then Eqn 77 implies

$$P_{acc}(o \rightarrow n) = \min \left( 1, \exp(-\beta\Delta U) \frac{P_{gen}(n \rightarrow o)}{P_{gen}(o \rightarrow n)} \right). \quad (66)$$

Ideally, by biasing the probability to generate a trial conformation in the right way, we could make the term on the right-hand side of Eqn. 66 always equal to unity. In that case, every trial move will be accepted. In section 4.4 we shall see that it is sometimes possible to achieve this ideal situation. However, in general, biased generation of trial moves is simply a technique to enhance the acceptance of such moves without violating detailed balance. Configurational bias Monte Carlo was designed to enhance the acceptance probability of trial moves that involve large conformational changes.

The starting point for configurational bias Monte Carlo is the Rosenbluth particle-insertion scheme described in section 2.7. At first sight, this may not appear to lead to a new algorithm but to a very old one. After all, the original Rosenbluth scheme itself was designed as a method to sample polymer conformations. However, as explained in section 2.7.3, the Rosenbluth scheme suffers from the drawback that it generates an unrepresentative sample of all polymer conformations: i.e. the probability to generate a particular conformation  $\Gamma$  using the Rosenbluth scheme, is *not* proportional to the Boltzmann weight of that conformation. The Rosenbluth weight  $W$ ,

was introduced to correct for this biasing in the computation of thermal averages. In section 2.7.3 it was explained that this correction procedure only works for relatively short chains. The solution of this problem is to bias the Rosenbluth sampling in such a way that the correct (Boltzmann) distribution of chain conformations is generated in a Monte Carlo sequence. In the configurational bias scheme to be discussed below, the Rosenbluth weight is used to bias the *acceptance* of trial conformations that are generated with the Rosenbluth procedure. As a consequence, all conformations are generated *with their correct Boltzmann weight*. This removes the main drawback of the original Rosenbluth scheme. Moreover, I shall show that the configurational bias scheme can also be extended to continuously deformable molecules.

#### 4.3.1. Sampling discrete conformations

Consider a trial move from a chain-conformation  $\Gamma_1$  with Rosenbluth weight  $w_1$  to some other conformation  $\Gamma_2$  with Rosenbluth weight  $w_2$ . Let us denote the probability (*i.e.* the normalized Boltzmann weight) of finding the chain in conformation  $\Gamma_1$  by  $P_1$  and the corresponding probability for  $\Gamma_2$  by  $P_2$ . In equilibrium, the number of chains in conformation 1(2),  $N_1$  ( $N_2$ ) is obviously proportional  $P_1$  ( $P_2$ ). We wish to perform a Monte-Carlo move such that detailed balance is satisfied, *i.e.* the rate  $K_{12}$  with which conformations of type 1 are transformed into type 2 equals the reverse rate:

$$K_{12} = K_{21}. \quad (67)$$

Let us try to compute  $K_{12}$  and  $K_{21}$ .  $K_{12}$  is equal to  $N_1$  times  $P_{gen}(\Gamma_2)$ , the probability of generating chain 2 by Rosenbluth sampling, times the acceptance probability  $P_{acc}$ . We assume that  $P_{acc}$  is an, as yet unspecified function of the ratio  $w_2/w_1$ ,  $g(w_2/w_1)$ .  $P_{gen}(\Gamma_2)$  is equal to

$$P_{gen}(\Gamma_2) = \prod_{i=1}^{\ell} \frac{\exp(-\beta u^{(i)}(\Gamma_2(i)))}{Z_i}, \quad (68)$$

where, as before, we use the notation  $Z_i \equiv \sum_{j=1}^b \exp(-\beta u^{(i)}(\Gamma_2(j)))$ . The sum in  $Z_i$  runs over all  $b$  possible directions of the  $i$ -th segment of the polymer. With this notation, Eq. 67 becomes

$$N_1 P_{gen}(\Gamma_2) g(w_2/w_1) = N_2 P_{gen}(\Gamma_1) g(w_1/w_2). \quad (69)$$

Equation 69 can be simplified by using the result that the Rosenbluth weight  $w_2$  times the generating probability  $P_{gen}(\Gamma_2)$  is proportional to  $N_2$  (and similarly for 1). Hence multiplying equation 69 on the left by  $w_2/w_1$  and on the right by  $w_1/w_1$  yields

$$N_1 N_2 \frac{g(w_2/w_1)}{w_2} = N_2 N_1 \frac{g(w_1/w_2)}{w_1} \quad (70)$$

or

$$\frac{g(w_2/w_1)}{g(w_1/w_2)} = \frac{w_2}{w_1}. \quad (71)$$

There are many choices of  $g$  that satisfy this condition. One obvious choice is the Metropolis form [54]

$$g(x) = \text{Min}(x, 1) \quad (72)$$

In words, the Configurational Bias Monte Carlo (CBMC) scheme works as follows:

1. Generate a trial conformation by using the Rosenbluth scheme (i.e. eq. 68) to regrow the entire molecule, or part thereof.
2. Compute the Rosenbluth weights  $w_{trial}$  and  $w_{old}$  of the trial conformation *and of the old conformation*.
3. Accept the trial move with a probability  $\text{Min}(w_{trial}/w_{old}, 1)$ .

#### 4.3.2. Continuously deformable chain

Next consider Configurational Bias Monte Carlo sampling of flexible chains with or without internal bending energy. Again, we first consider the probability to generate a trial configuration  $\Gamma_j$  using the (extended) Rosenbluth scheme. As before, we only consider the expression for one of the  $\ell$  segments, to keep the equations simple.

$$P_{gen}(\{j\}) = d\Gamma_j P_{id}(\Gamma_j) \left[ \prod_{j' \neq j}^k d\Gamma_{j'} P_{id}(\Gamma_{j'}) \right] \frac{\exp(-\beta u_{ext}(j))}{\sum_{j'=1}^k \exp(-\beta u_{ext}(j'))} \quad (73)$$

In order to compute the acceptance probability of this move, we have to consider what happens in the reverse move. Then we start from conformation  $j$  and generate a set of  $k$  trial directions that includes  $i$ . When computing the acceptance probability of the forward move, we have to impose detailed balance. However, detailed balance in this case means not just that in equilibrium the number of moves from  $i$  to  $j$  is equal to the number of reverse move, but even that the rates are equal *for any given set of trial directions for the forward and reverse moves*. This condition we will call 'super-detailed balance'. Super-detailed balance implies that we can only decide on the acceptance of the forward move if we also generate a set of  $k - 1$  trial directions around the old conformation  $i$ . We denote the probability to generate this set of  $k - 1$  trial orientations by  $P_{rest}(\{i\})$ , where the sub-script 'rest' indicates that this is the set of orientations around, but excluding,  $i$ . This allows us to compute the ratio  $w_j^{(t)}/w_i^{(o)}$  of the Rosenbluth weights for forward and reverse moves:

$$w_j^{(t)} = \frac{\exp(-\beta u_{ext}^{(t)}(j) + \sum_{j' \neq j}^{k-1} \exp(-\beta u_{ext}^{(t)}(j'))}{k}$$

and

$$w_i^{(o)} = \frac{\exp(-\beta u_{ext}^{(o)}(i)) + \sum_{i' \neq i}^{k-1} \exp(-\beta u_{ext}^{(o)}(i'))}{k}$$

The superscript ( $t$ ) and ( $o$ ) distinguish the trial conformation from the old conformation. Again, we shall devise a Metropolis criterion to determine the acceptance probability of a trial move from  $i$  to  $j$ . As before, this acceptance probability is determined by the ratio  $x \equiv w_j^{(t)}/w_i^{(o)}$  (actually, for a molecule of  $\ell$  segments, we should compute a product of such factors). Let us assume that  $w_j^{(t)} < w_i^{(o)}$ . In that case,  $P_{acc}(i \rightarrow j) = x$  while  $P_{acc}(j \rightarrow i) = 1$ . Next, let us check whether detailed balance is satisfied. To do so, we write down the explicit expressions for  $K_{ij}$  and  $K_{ji}$ .

$$K_{ij} = N_i P_{gen}(\{j\}) P_{rest}(\{i\}) w_j^{(t)} / w_i^{(o)}$$

and

$$K_{ji} = N_j P_{gen}(\{i\}) P_{rest}(\{j\}) 1$$

In addition, we use the fact that

$$P_{gen}(\{j\}) P_{rest}(\{i\}) w_j^{(t)} \sim N_j P_{rest}(\{j\}) P_{rest}(\{i\})$$

and

$$P_{gen}(\{i\}) P_{rest}(\{j\}) w_j^{(t)} \sim N_i P_{rest}(\{i\}) P_{rest}(\{j\})$$

It then follows immediately that

$$K_{ij} w_i^{(o)} = N_i P_{gen}(\{j\}) P_{rest}(\{i\}) w_j^{(t)} \quad (74)$$

$$= \text{constant} \times N_i N_j P_{rest}(\{i\}) P_{rest}(\{j\})$$

$$= N_j P_{gen}(\{i\}) P_{rest}(\{j\}) w_i^{(o)}$$

$$= K_{ji} w_i^{(o)} \quad (75)$$

Hence,  $K_{ij}$  is indeed equal to  $K_{ji}$ . This completes the proof that the above Configurational Bias MC scheme satisfies detailed balance. Note that, in this derivation, the number of trial directions,  $k$ , was arbitrary.

The CBMC scheme for completely flexible molecules has been applied by Frenkel et al. [76] to study conformational changes of chain molecules in solution. The method was found to be particularly efficient in generating large conformational changes of such a molecule. More interestingly, the CBMC scheme makes it possible to perform direct simulations of phase equilibria in polymer mixtures, using the Gibbs-ensemble method [116, 117, 118]. This aspect of the technique will be discussed in the lectures of Panagiotopoulos, elsewhere in this volume.

### 4.3.3. How many trial directions?

What is the optimum choice of the number of trial directions in a CBMC simulation of a continuously deformable chain? In this section I give a crude, but very simple, rule to arrive at an estimate. A more detailed description can be found in ref [119, 120]. To arrive at an estimate, I define the efficiency  $E(\ell)$  of the algorithm to grow a polymer of length  $\ell$  as the average acceptance of a trial move,  $P_{acc}(\ell)$ , divided by the average cost  $c(\ell)$  of a trial move. The cost is assumed to be proportional to the total number of trial directions that is generated before a move can be rejected. For a polymer of length  $\ell$  and with  $k$  trial directions per segment, the computational cost should be between  $k$  if a trial move is rejected at the first segment and  $\ell \times k$  if a trial move is accepted. In fact, as we must compute both the new and the old Rosenbluth weight, we should distinguish between a situation where we compute the old and new weights *segment-by-segment* or *first* for the trial move and, only if the latter results in an acceptable configuration, of the old configuration. This will make a (slight) difference. For the moment, I assume that we generate old and new trial directions simultaneously.

I use the following notation:  $k$  is the number of trial directions per segment,  $\ell$  is the number of segments in the polymer and  $P_R$  is the probability that a given trial segment will be rejected because it results in a hard-core overlap.

It is obvious that  $k$  should, in principle depend on the segment index  $n$ . This is so because the amount of ‘work’ that one should spend to grow the remainder of a chain from segment  $n + 1$  to  $l$  will depend on the amount of work that was already spent to arrive at segment  $n$ . The more work has been ‘invested’ the larger the amount of work that we should be prepared to spend to complete the chain.

Consider the ratio  $E(\ell + 1)/E(\ell)$ :

$$E(\ell + 1)/E(\ell) = \frac{P_{acc}(\ell + 1)/c(\ell + 1)}{P_{acc}(\ell)/c(\ell)}$$

The probability  $P_{acc}(\ell + 1)$  is given by

$$P_{acc}(\ell + 1) = P_{acc}(\ell) \times (1 - P_R^{k(\ell+1)})$$

where  $k$  depends explicitly on  $\ell$ . In addition we have the following recursion expression for  $c(\ell + 1)$ , the cost to grow a chain of length  $\ell + 1$ :

$$c(\ell + 1) = c(\ell) + P_{acc}(\ell) \times (2k(\ell + 1) - 1)$$

We thus arrive at the following very simple recursion relation for  $E$ :

$$E(\ell + 1)/E(\ell) = \frac{1 - P_R^{k(\ell+1)}}{1 + E(\ell) \times (2k(\ell + 1) - 1)}$$

In order to solve for  $E(\ell + 1)$ , we only have to specify  $P_R$  and  $E(0)$ , the efficiency to start a regrowth. Both can be computed numerically in a short simulation on a fluid of monomers or dimers. When we cut a chain, virtually no work is involved in generating the new starting point and we can consider the limit  $E(0) \rightarrow \infty$ . If we move the chain to a new starting point,  $E(0)$  will be less. In our units, the cost will be 1, and the acceptance 1 (in case of an excluded volume map) or less (e.g.  $1 - P_R$ ). We then compute  $k(1)$  such that  $E(1)$  is maximized. Next  $k(2)$  to maximize  $E(2)/E(1)$ , and so on. This scheme is extremely efficient and will indeed predict values for  $k(\ell)$  that are initially quickly, but later slowly, increasing functions of  $\ell$ .

A final refinement that can be made is to improve on the assumption that the probability of acceptance of a trial segment is simply equal to  $1 - P_R^k$ . Again, this can be done numerically, using a simple statistical model. It does result in slightly different values for  $k(\ell)$ . But the overall trend is much the same as before.

#### 4.3.4. From polymers back to atoms

Up to this point, I have introduced the configurational-bias MC scheme as a method to sample conformations of straight chain molecules. However, the method is applicable to branched polymers and ring polymers. In the latter case, the condition that the  $N$ -th segment is connected to the first, is accounted for by biasing the generation of trial directions [121, 122]. The *a priori* probability to generate a given trial direction  $\mathbf{F}_i$  as the  $m$ -th segment, is proportional to the number of ideal (*i.e.* non-(self)-avoiding) conformations that connect this segment to the first segment in  $N - m$  steps.

Thus far, the configurational-bias scheme has been presented exclusively as a method to generate *polymer* conformations. In fact, the method is more general than that. It can be used as a scheme to perform collective rearrangements of *any* set of labeled coordinates. In fact, the scheme can be used to carry out Monte Carlo moves to swap  $n$  small particles within a volume  $\Delta V$  with one large particle that occupies the same (excluded) volume. This application of the CBMC scheme has been exploited by Biben [123] to study mixtures of large and small hard spheres. Gibbs-ensemble simulations of mixtures of spherical colloids and rodlike polymers were performed by Bolhuis [51] using CBMC-style particle swaps and a closely related approach was employed by Dijkstra to study phase separation [121, 122] in mixtures of large and small hard-core particles on a lattice. An application of CBMC to improve the sampling of ionic solutions has been proposed by Shelley and Patey [124].

#### 4.4. CLUSTER MOVES

As explained at the beginning of section 4.3, the crucial step in configurational bias Monte Carlo is that the ‘bias’ in the generation of trial conformations results in an enhanced acceptance of these trials moves. Ideally, we should like to bias the generation of trial moves in such a way that *every* move is always accepted. Surprisingly, it turns out that this is sometimes possible. Swendsen and Wang [125] (for a review, see ref. [8]) have shown that, at least for certain classes of lattice problems, it is possible to perform ‘cluster moves’ that have an acceptance probability of 100%.

The central idea behind the Swendsen-Wang (SW) scheme and subsequent extensions and modifications is to generate trial configurations with a probability that is proportional to the Boltzmann weight of that configuration. As a result, the subsequent trial moves can be accepted with 100% probability. I use a somewhat simplified derivation of the Swendsen-Wang scheme for cluster moves based, again, on the condition for detailed balance. Consider an ‘old’ configuration (labeled by a superscript  $o$ ) and a ‘new’ configuration, denoted by a superscript  $n$ . Detailed balance is satisfied if the following equality holds:

$$\begin{aligned} P_B^o P_{Gen}^o(\{cluster\}) P^{\{cl\}}(o \rightarrow n) P_{acc}(o \rightarrow n) &= \\ P_B^n P_{Gen}^n(\{cluster\}) P^{\{cl\}}(n \rightarrow o) P_{acc}(n \rightarrow o) & \end{aligned} \quad (76)$$

where  $P_B^{o(n)}$  is the Boltzmann weight of old (new) configuration,  $P_{Gen}^o(\{cluster\})$  denotes the probability of generating a specific cluster, starting from the old configuration of the system.  $P^{\{cl\}}(o \rightarrow n)$  is the probability to transform the generated cluster from the old to the new situation, and finally  $P_{acc}(o \rightarrow n)$  is that acceptance probability of a given trial move. We can simplify eq. 76 in two ways: first of all, we impose that the *a priori* probability  $P^{\{cl\}}(o \rightarrow n)$  is the same for the forward and reverse moves. Moreover, we wish to impose  $P_{acc} = 1$  for both forward and reverse moves. This may not always be feasible. However, for the simple case that I discuss below, this is indeed possible. The detailed balance equation then becomes:

$$P_B^o P_{Gen}^o(\{cluster\}) = P_B^n P_{Gen}^n(\{cluster\}) \quad (77)$$

or

$$\frac{P_{Gen}^o(\{cluster\})}{P_{Gen}^n(\{cluster\})} = \frac{P_B^n}{P_B^o} = \exp(-\beta \Delta E), \quad (78)$$

where  $\Delta E$  is the difference in energy between the new and the old configurations. The trick is then to find a recipe for cluster generation that will satisfy eq.78. In order to illustrate how this works, consider the Ising model. The extension to many other models is straightforward.

*Ising model* For the construction of the Ising SW algorithm, the dimensionality of the model is irrelevant. Consider a given configuration of the spin system, with  $N_p$  spin-pairs parallel and  $N_a$  spin-pairs anti-parallel. The total energy of that configuration is:

$$E = (N_a - N_p)J$$

where  $J$  denotes the strength of the nearest-neighbor interaction. The Boltzmann weight of that configuration is

$$P_B^o = \exp(-\beta J(N_a - N_p))/Z$$

where  $Z$  is the partition function of the system. In general,  $Z$  is unknown, but that is unimportant. We only use the fact that  $Z$  is a constant. Next, we construct clusters by creating bonds between spin pairs according to the following recipe:

- If nearest neighbors are *anti-parallel*, they are not connected.
- If nearest neighbors are parallel, they are connected with probability  $p$  and disconnected with probability  $(1 - p)$ .

Here, I have assumed that  $J$  is positive. If  $J$  is negative (anti-ferromagnetic interaction), parallel spins are not connected, while anti-parallel spins are connected with a probability  $p$ .

In the case that we consider, there are  $N_p$  parallel spin pairs. The probability that  $n_c$  of these are connected and  $n_b = N_p - n_c$  are 'broken' is:

$$P_{Gen}^o(\{cluster\}) = p^{n_c}(1 - p)^{n_b} .$$

Note that this is the probability of connecting (*c.q.* break) a *specified* subset of all links between parallel spins. Once the connected bonds have been selected, we can define the clusters in the system. A cluster is a set of spins that are at least singly connected by bonds. We now choose our subset of clusters to flip. After the cluster flipping, the number of parallel and anti-parallel spin pairs will have changed, e.g.

$$N_p^{new} = N_p^{old} + \Delta$$

and (hence)

$$N_a^{new} = N_a^{old} - \Delta .$$

Therefore, the total energy of the system will have changed by an amount  $-2J\Delta$ :

$$E_{new} = E_{old} - 2J\Delta .$$

Let us now consider the probability to make the *reverse* move. In order to do this, we should generate the *same* cluster structure, but now starting

from a situation where there are  $N_p + \Delta$  parallel spin pairs and  $N_a - \Delta$  anti-parallel pairs. As before, the bonds between anti-parallel pairs are assumed to be broken (this is compatible with the same cluster structure). We also know that the new number of connected bonds,  $n'_c$  must be equal to  $n_c$ , *because the same number of connected bonds is required to generate the same cluster structure*. The difference appears when we consider how many of the bonds between parallel spins in the new configuration should be broken ( $n'_b$ ). Using

$$N_p^{new} = n'_c + n'_b = n_c + n'_b = N_p^{old} + \Delta = n_c + n_b + \Delta$$

we see that

$$n'_b = n_b + \Delta$$

If we insert this in eq. 78, we obtain

$$\begin{aligned} \frac{P_{Gen}^o(\{cluster\})}{P_{Gen}^n(\{cluster\})} &= \\ \frac{p^{n_c}(1-p)^{n_b}}{p^{n_c}(1-p)^{n_b+\Delta}} &= \\ (1-p)^{-\Delta} &= \frac{P_B^n}{P_B^o} = \exp(2\beta J \Delta). \end{aligned} \quad (79)$$

In order to satisfy this equation, we must have:

$$1 - p = \exp(-2\beta J)$$

or

$$p = 1 - \exp(-2\beta J),$$

which is the Swendsen-Wang rule.

#### 4.5. GENERAL CLUSTER MOVES

In general, it is not possible to design clusters such that trial moves are always accepted. However, it is often convenient to perform clustering to enhance the acceptance of trial moves. For instance, in molecular systems with very strong short-range attractions, trial moves that pull two neighboring particles apart are very likely to be rejected. It is therefore preferable to include trial moves that attempt to displace the tightly bound particles as a single cluster. In order to do this, we have to specify a rule to generate clusters. Let us assume that we have such a rule that tells us that particles  $i$  and  $j$  belong to a single cluster with probability  $p(i, j)$  and are

disconnected with a probability  $1 - p(i, j)$ .  $p(i, j)$  depends on the state (relative distance, orientation, spin etc.) of particles  $i$  and  $j$ . Moreover, we impose that  $p(i, j)$  is unchanged in a cluster move if both  $i$  and  $j$  belong to the cluster, and also if neither particle belongs to the cluster. For instance,  $p(i, j)$  could depend on the current distance of  $i$  and  $j$  only. If we denote the potential energy of the old (new) configuration by  $U_0$  ( $U_1$ ), the detailed balance condition requires that

$$\begin{aligned} \exp(-\beta U_o) \prod_{kl} (1 - p^f(k, l)) P_{acc}(o \rightarrow n) &= \\ \exp(-\beta U_n) \prod_{kl} (1 - p^r(k, l)) P_{acc}(n \rightarrow o) & \end{aligned} \quad (80)$$

where  $k$  denotes a particle *on* the cluster, and  $l$  a particle outside it. The superscripts  $f$  and  $r$  denote forward and reverse moves. In writing Eqn. 80, we have used the fact that the probability to form bonds completely within, or completely outside, the cluster is the same for forward and reverse moves. From Eqn. 80, we derive an expression for the ratio of the acceptance probabilities:

$$R \equiv \frac{P_{acc}(0 \rightarrow 1)}{P_{acc}(1 \rightarrow 0)} = \exp(-\beta(U_1 - U_0)) \prod_{kl} \frac{(1 - p^r(k, l))}{(1 - p^f(k, l))}. \quad (81)$$

Clearly, many choices for  $p_{kl}$  are possible. A particularly simple form was chosen by Wu, Chandler and Smit [126] who assumed that  $p(i, j) = 1$  for  $r_{ij}$  less than a critical distance  $r_c$  and  $p(i, j) = 0$  beyond that distance. Note that the acceptance rule in Eqn. 80 guarantees that two particles that did not belong to the same cluster in the old configuration will not end up at a distance less than  $r_c$ .

## 5. Tricks with lattice models

### 5.1. ATHERMAL MIXTURES

Phase separation in binary mixtures is the example that is used in many textbooks to illustrate the competition between energy and entropy in a phase transformation. For a mixture at constant total volume  $V$ , the Helmholtz free energy  $F$ , should be minimal. As a first approximation, the entropy of mixing of a mixture of two species  $A$  and  $B$ , is replaced by the entropy of mixing of an ideal mixture

$$S_{id}(X) = -Nk_B[X \ln X + (1 - X) \ln(1 - X)] \quad (82)$$

where  $X$  denotes the mole-fraction of one component (say  $A$ ):  $X = N_A/(N_A + N_B)$ . The entropy of mixing given by eqn. 82 is a convex function of  $X$ .

As a consequence,  $S_{id}(X)$  will always decrease if phase separation takes place. This implies that phase separation can only take place if the resulting decrease in energy  $E$  outweighs the increase in  $-TS_{id}$ . In a hard-core mixture, there is no energy change upon mixing. Hence, if eqn. 82 were exact, we should never observe phase separation in a hard-core mixture.

The naive picture sketched above is not correct. From computer simulation [121, 122, 127, 51, 42] we know of many examples of de-mixing transitions in systems with purely hard-core interactions. Below, I discuss one particular class of de-mixing transitions in hard-core mixtures, namely depletion flocculation in colloids [128, 49]. The reason why I discuss this example is that it allows me to introduce another technique to study phase transitions in complex fluids by computer simulation. This technique is based on an exact enumeration approach to grand-canonical simulations. In particular, we consider a mixture of two components, one of which is held at a fixed chemical potential. In some cases, we can ‘integrate out’ the degrees of freedom of the species at fixed chemical potential.

## 5.2. DEPLETION FLOCCULATION

Experimentally, it is well known that the addition of a small amount of free, non-adsorbing polymer to a colloidal suspension induces an effective attraction between the colloidal particles and may even lead to coagulation. This effect has been studied extensively [129] and is well understood, at least qualitatively. The polymer-induced attraction between colloids is an *entropic* effect: when the colloidal particles are close together, the total number of accessible polymer conformations is larger than when the colloidal particles are far apart. However, although the physical mechanism responsible for polymer-induced coagulation is understood *qualitatively*, a quantitative description of this phenomenon is more difficult. This is so because the polymer-induced attraction between the colloidal particles is non-pairwise additive. Moreover, it depends both on the osmotic pressure of the polymer and on the concentration of the colloid. A direct analytical evaluation of the (grand-canonical) partition function is impossible, even when one considers only the very simplest model, *viz.* that of a mixture of hard-core colloidal particles with ideal chain molecules with conformations that are restricted to a lattice. It is therefore desirable to carry out ‘exact’ numerical simulations to investigate the phase behavior. Yet, the computational problems are still formidable. What is required is a numerical scheme that samples the positions of the colloidal particles while averaging over all possible conformations of a large (and fluctuating) number of chain molecules.

Fortunately, it is possible to construct a rigorous and efficient Monte

Carlo scheme to study this model. Our approach relies on the fact that we can recursively compute the partition function of an ideal (non-self avoiding) chain on a lattice in an arbitrary external potential [80, 128]. This is most easily seen by considering a chain of length  $\ell - 1$  on a lattice. Suppose that every lattice site has  $b$  neighbors, that the polymer consists of  $p$  segments and that there are  $N$  lattice sites. Clearly, for an ideal polymer in the absence of any obstacles, the total number of allowed conformations is  $\Omega_{id} \equiv N \times b^p$ . If obstacles are present, some random walks are blocked, and the total number of allowed conformations,  $\Omega_T$ , is less than  $\Omega_{id}$ . The ‘brute-force’ method to determine the ratio  $\Omega_T/\Omega_{id}$  would be to attempt a large number of insertions of chains with arbitrary conformation at random points on the lattice. The ratio of the number of ‘accepted’ trial moves to the total number of attempted insertions can be used to compute the excess chemical potential of an ideal polymer chain in this system (see section 2.5.1)

$$\mu_{excess} = -k_B T \log \langle P_{acceptance} \rangle \quad (83)$$

This is the usual ‘Widom’ expression for the excess chemical potential [53]. However, unless the density of obstacles is quite low, the relative statistical error in the insertion probability will be quite large

Next, consider the recursive approach. To this end, let us first compute all  $N$  Boltzmann factors associated with the insertion of a point particle at any lattice site. Clearly, the sum of these Boltzmann factors is simply the partition function of a point-particle on a lattice, in an external potential. Let us denote the Boltzmann factor associated with site  $i$  as  $f_i^{(0)} \equiv \exp(-u(\mathbf{r}_i)/k_B T)$ . The partition function for a one-segment polymer (2 points) is computed as follows. Starting from site  $i$ , we have  $b$  ways to grow one segment. But all  $b$  directions will, in general, have different Boltzmann weights  $f_j^{(0)}$ , where  $j$  denotes one of the nearest neighbors of  $i$ . The total Boltzmann weight associated with the addition of a 1-segment polymer at site  $i$  is then:

$$f_i^{(1)} \equiv f_i^{(0)} \times \left( \sum_j^b f_j^{(0)} \right) \quad (84)$$

where the sum runs over the nearest neighbors of  $i$ . The partition function of a 1-segment polymer on a lattice is then simply:

$$\Omega_1 = \sum_{i=1}^N f_i^{(1)}, \quad (85)$$

(where, for convenience, I have assumed that the polymer ‘head’ and ‘tail’ are distinguishable). Repeating the same argument  $p$  times, it is clear that

the Boltzmann factor associated with *all possible conformations* of a  $p$ -segment polymer starting at site  $i$  is given by:

$$f_i^{(p)} = f_i^{(0)} \times \left( \sum_j^b f_j^{(p-1)} \right). \quad (86)$$

And the total partition function is given by:

$$\Omega_p = \sum_{i=1}^N f_i^{(p)}. \quad (87)$$

The important fact to note is that the computation of  $\Omega_p$ , a quantity that depends on  $N \times b^p$  conformations, requires only  $p$  iterations of  $b \times N$  local ‘propagations’ of real numbers. For a fixed external potential, the partition function thus computed is *exact*. In a practical situation, as with the addition of polymer to a colloidal dispersion, the external potential is not fixed, but depends on the (continuous) coordinates of all  $M$  colloidal particles,  $\{\mathbf{r}^M\}$ . Hence, the total partition function of the polymer also depends on these coordinates:  $\Omega_p = \Omega_p(\mathbf{r}^M)$ .

We now use this scheme to perform simulations of depletion flocculation.

Up to this point we have not specified the nature of the ‘external’ potential. We now assume that this potential is due to the presence of  $N$  hard, spherical colloidal particles each of which occupies many lattice sites. The polymer partition function clearly depends on the coordinates  $\mathbf{r}^N$  of the colloidal particles:  $\Omega_\ell(\mathbf{r}^N)$ . The configurational part of the partition function of the system of  $N$  colloids plus one polymer of length  $\ell$  in volume  $V$  is then given by:

$$Z(V, N, 1) = \int_V d\mathbf{r}^N e^{-U_{hs}(\mathbf{r}^N)} (\Omega_\ell(\mathbf{r}^N)), \quad (88)$$

where  $U_{hs}(\mathbf{r}^N)$  denotes the hard-sphere interaction. Next, we make use of the fact that we are considering *ideal* polymers. In that case we can immediately write down the corresponding partition function for  $N$  colloids and  $M$  ideal polymers:

$$Z(V, N, M) = \int_V d\mathbf{r}^N e^{-U_{hs}(\mathbf{r}^N)} (\Omega_\ell(\mathbf{r}^N))^M / M!, \quad (89)$$

where the factor  $1/M!$  accounts for the fact that the polymers are indistinguishable. Using eq. 89 it is straightforward to transform to an ensemble where the polymer chemical potential (i.e. the osmotic pressure) is kept

fixed. The corresponding grand-canonical partition function is given by:

$$\begin{aligned}
 \Xi(V, N, \mu) &= \sum_{M=0}^{\infty} e^{M\mu_{ch}} Z(V, N, M)/M! \\
 &= \int_V d\mathbf{r}^N e^{-U_{hs}(\mathbf{r}^N)} \sum_{M=0}^{\infty} e^{M\mu_{ch}} (\Omega_{\ell}(\mathbf{r}^N))^M / M! \\
 &= \int_V d\mathbf{r}^N e^{-U_{hs}(\mathbf{r}^N)} e^{z\Omega_{\ell}(\mathbf{r}^N)}. \tag{90}
 \end{aligned}$$

In the last line of eq. 90, we have introduced the polymer fugacity  $z \equiv \exp(\mu_{ch})$ , where  $\mu_{ch}$  denotes the chemical potential of the chain molecules.

The important point to note is that eq. 90 allows us to evaluate the properties of the colloidal particles in osmotic equilibrium with a polymer reservoir. In particular, it shows that we can perform Monte Carlo sampling of the colloidal particles. The polymers only affect  $U_{eff}(\mathbf{r}^N)$ , the effective interaction between the colloidal particles:

$$U_{eff}(\mathbf{r}^N) \equiv U_{hs}(\mathbf{r}^N) - z\Omega_{\ell}(\mathbf{r}^N). \tag{91}$$

$z\Omega_{\ell}(\mathbf{r}^N)$  measures the entropic interaction between the colloids due to *all* possible polymer conformations. This entropic interaction is, in principle, not pairwise additive. In fact, it is shown in ref. [49], that for all but the shortest chain molecules, this non-additivity of the polymer-induced interaction between the colloids, has a pronounced effect on the structure and stability of the mixture. For more details, we refer the reader to ref. [49].

It turns out that the same approach can be used in other contexts. For instance, van Duijneveldt and Lekkerkerker [127] applied the method to the study of a hard-core mixture on a triangular lattice (the corresponding problem on a square or cubic lattice can be mapped onto the Ising model [130] and is therefore ‘trivial’). However, the technique is not limited to complex fluids, or even classical systems. In particular, we can apply the same approach to the demixing of a system of ideal point fermions (“non-interacting electrons”) and classical fluid particles. The analogy between this system and the previous ones rests on the fact that, in Feynman’s path-integral formulation of quantum mechanics, there is an isomorphism between the partition-function of a  $d$ -dimensional quantum-mechanical system, and that of a  $d + 1$ -dimensional, “polymer” system. Ideal quantum particles in  $d$  dimensions can be described by ideal, directed polymers in  $d + 1$  dimensions. This suggests that we may expect that mixtures of ideal quantum particles and atoms (effectively, hard spheres) behave much like a mixture of ideal polymers and colloids. In fact, the situation in the quantum case is complicated by the fact that one must take exchange into

account. This is relatively straightforward in the case of numerical simulations of bosons [131], but it creates serious problems for fermions (the ‘sign’-problem). However, in ref. [132], it is shown how, at least for a mixture of atoms and ideal fermions, these problems may be overcome. Alavi and Frenkel [132] show that it is, in fact, possible to evaluate numerically (but exactly) the grand-canonical partition function of ideal fermions in the presence of arbitrary “obstacles” (in particular, atoms). This makes it possible to perform Monte Carlo sampling of the atomic positions in the presence of the fermion ‘bath’. Here, I will not discuss the technical aspects of these simulations. Rather, I will summarize the main conclusion, which is that there is indeed strong evidence for a demixing transition as the chemical potential of the fermions is increased beyond a certain critical value. Again, there are no attractive forces in this system. Hence, this fermion-induced phase separation appears to be yet another example of “attraction through repulsion”. For more details, the reader is referred to ref. [132].

### 5.3. DYNAMICS OF LATTICE GASES

There exists a rather unexpected relation between the scheme to compute the entropy of an ideal polymer on a lattice and a method to study the decay of velocity fluctuations in lattice gases. Lattice gases can be considered to be the very simplest representation of an atomic fluid: the ‘atoms’ are constrained to move along the bonds joining the lattice sites. No two particles can move along the same bond in the same direction. The state of the lattice is completely specified by indicating which links are occupied and which are empty. This implies that lattice-gas particles are indistinguishable.

The time evolution of the system is governed by the following rules:

1. Propagation: all particles move in one time step (for convenience we choose  $\Delta t = 1$ ) from their initial lattice position (say  $\mathbf{X}$ ) to a new position ( $\mathbf{X}' = \mathbf{X} + \mathbf{c}_\alpha$ ; where  $\mathbf{c}_\alpha$  is the velocity of species  $\alpha$ ). The velocities  $\mathbf{c}_\alpha$  are such that at the end of the propagation steps all particles are once more positioned at lattice sites.
2. Collision: the particles at all sites on the lattice undergo a collision that conserves the total number of particles and the total momentum at each site. These local collision rules may, or may not, be deterministic.

Provided that the lattice has a sufficiently high symmetry (e.g. triangular in 2 dimensions) and the collision rules are sufficiently isotropic (for a discussion, see [133]), it can be shown that the equation that governs the time evolution of the ‘flow field’ of such a lattice gas becomes equivalent to the Navier-Stokes equation for an incompressible fluid in the limit that the flow velocity is much less than the particle velocity, and all spatial variations in

the system occur on a scale that is large compared to the mean free path of the lattice gas particles. In this respect lattice gases model atomic fluids.

Because lattice gases provide a computationally cheap model of a fluid, several attempts have been made to use lattice-gases as a model of the solvent in problems where both Brownian motion and hydrodynamic effects are important (see *e.g.* [134, 135]). Below, I shall describe one such application of lattice gases: namely the computation of the long-time ‘hydrodynamic’ tail in the velocity correlation of a particle in a fluid.

In the late sixties, Alder and Wainwright [136] showed that, at long times, the velocity autocorrelation function (VACF) of a particle in a moderately dense fluid of hard spheres or hard disks decays algebraically. This observation had far-reaching implications because it demonstrated the breakdown of the ‘Molecular Chaos’ assumption, originally introduced by Boltzmann as the ‘Stoßzahlansatz’: Molecular Chaos immediately implies exponential decay of the VACF. The algebraic long-time tails are hydrodynamic in origin, as they are caused by the coupling between particle diffusion and shear modes in the fluid[137, 138, 139, 140, 141].

To a first approximation the leading term in the long-time tail of the velocity ACF is given by:

$$\langle v_x(0)v_x(t) \rangle \approx \frac{d-1}{d} \frac{1}{\rho(4\pi(D_0 + \nu_0)t)^{d/2}}, \quad (92)$$

where  $\rho$  is the number density,  $D_0$  the ‘bare’ self-diffusion constant,  $\nu_0$  the kinematic viscosity and  $d$  the dimensionality. Hydrodynamic tails are very weak and hence a quantitative comparison between theory and simulation was difficult [142, 143, 144].

However, using lattice-gas models it is possible to compute hydrodynamic tails with high accuracy [145]. The technique is based on the observation that lattice gas particles are indistinguishable. Hence, even if we label a particle before a collision, we have no way to identify the tagged particle after the collision. This fact may be exploited by averaging the velocity auto-correlation function over all possible post-collisional labelings of the particles. This is analogous to the summing over all possible polymer conformations that was described in the previous section. Using this approach it is possible to increase the computational efficiency of the method by 10 orders of magnitude and arrive at extremely accurate measurements of the hydrodynamic tails in lattice gases. For more details, see [135].

## Appendices

### A. Overlapping-distribution method for polymers

Let us first consider how the basic idea behind the overlapping-distribution method can be applied to the Rosenbluth-insertion scheme. The simplest approach would be to consider the histogram of the potential energy change upon addition or removal of a chain molecule. This approach has been pursued by de Pablo *et al.* [79]. However, it turns out that for chain molecules this approach differs from the original Shing-Gubbins approach in that it has little, if any, diagnostic value. For instance, if we consider the chemical potential of hard-core chain molecules, the distributions of  $\Delta U$  will always overlap (namely at  $\Delta U=0$ ), even in the regime where the method cannot be trusted. Here, I shall describe an overlapping distribution method that is based on histograms of Rosenbluth weights [84]. This method will prove to be a useful diagnostic tool.

Consider again systems with *internal* potential energy  $u_{int}$  and ‘external’ potential energy  $u_{ext}$ . In what follows, I shall compare two systems. The first, denoted by “0”, contains  $N$  chain molecules ( $N \geq 0$ ). The second system, denoted by “1”, contains  $N + 1$  chain molecules. In addition, both systems may contain a fixed number of other (“solvent”) molecules. Let us first consider system 1. Around every segment  $j$  of a particular chain molecule (say  $i$ ), we can generate  $k - 1$  trial directions according to an “internal” probability distribution given by Eqn 35. Note that the set does *not* include the actual orientation of segment  $j$ . We denote this set of trial orientations by

$$\{\gamma_{rest}(j)\} \equiv \prod_{j'=1}^{k-1} \{\gamma\}_{j'}$$

where the subscript ‘rest’ indicates that this set excludes the actual segment  $j$ . The probability to generate this set of trial directions is given by  $P_{rest}(j)$  (given by Eqn. 35). Having thus constructed an “umbrella” of trial directions around every segment  $1 \leq j \leq \ell$ , we can compute the Rosenbluth weight  $w_i$  of molecule  $i$ . Clearly,  $w_i$  depends on all coordinates of the remaining  $N$  molecules (for convenience, I assume that we are dealing with a neat liquid), on the position  $\mathbf{r}_i$  and conformation  $\Gamma_i$  of molecule  $i$ , and on the  $\ell$  sets of  $k - 1$  trial directions,

$$\{\Gamma_{rest}\} \equiv \prod_{j=1}^{\ell} \{\gamma\}_{rest}(j)$$

We now define a quantity  $x$  through

$$x \equiv \ln w_i(\mathbf{Q}^{N+1}, \{\Gamma_{rest}\})$$

where I use  $\mathbf{Q}$  to denote the translational coordinates  $\mathbf{r}$  and conformational coordinates  $\Gamma$  of a molecule. Next, consider the expression for the probability density of  $x$ ,  $p_1(x)$ :

$$p_1(x) = \frac{\int d\mathbf{Q}^{N+1} d\{\Gamma_{\text{rest}}\} \exp(-\beta U(\mathbf{Q}^{N+1})) \prod_{j=1}^{\ell} P_{\text{rest}}(j) \delta(x - \ln w_i)}{Z_{N+1}} \quad (\text{A-1})$$

where

$$\begin{aligned} Z_{N+1} &= \int d\mathbf{Q}^{N+1} d\{\Gamma_{\text{rest}}\} \exp(-\beta U(\mathbf{Q}^{N+1})) \prod_{j=1}^{\ell} P_{\text{rest}}(j) \\ &= \int \dots \int d\mathbf{Q}^{N+1} \exp(-\beta U(\mathbf{Q}^{N+1})) \end{aligned} \quad (\text{A-2})$$

The second line of this equation follows from the fact that all  $P_{\text{int}}(j)$  are normalized. We shall now try to relate  $p_1(x)$  to an average in system 0 (*i.e.* the system containing only  $N$  chain molecules). To this end, we write  $U(\mathbf{Q}^{N+1})$  as  $U(\mathbf{Q}^N) + u_{\text{ex}}(\mathbf{Q}^N, \mathbf{Q}_i) + u_{\text{int}}(\mathbf{Q}_i)$ . Secondly, we make use of the fact that

$$\exp(-\beta u_{\text{int}}(i)) = Z_{id} \times \prod_{j=1}^{\ell} P_{\text{int}}(j)$$

where

$$Z_{id} \equiv \int d\Gamma_i \prod_{j=1}^{\ell} \exp(-\beta u_{\text{int}}(j))$$

Our expression for  $p_1(x)$  now becomes

$$\begin{aligned} p_1(x) &= \frac{Z_{id}}{Z_{N+1}} \times \int d\mathbf{Q}^N d\mathbf{r}_i d\{\Gamma_{\text{trial}}\} \exp(-\beta U(\mathbf{Q}^N)) \\ &\quad \times \prod_{j=1}^{\ell} P_{\text{trial}}(j) \exp(-\beta u_{\text{ex}}(j)) \delta(x - \ln w_i). \end{aligned} \quad (\text{A-3})$$

We use the symbol  $\{\Gamma_{\text{trial}}\}$  to denote the set of *all* trial segments, *i.e.* the 'umbrella' of trial directions around all segments of the chain molecule, *plus the segments themselves*. Next, we multiply and divide every term  $\exp(-\beta u_{\text{ex}}(j))$  by  $Z_j$ , defined as

$$Z_j \equiv \sum_{j'=1}^k \exp(-\beta u_{\text{ext}}(j')).$$

This allows us to write eqn. A-3 as

$$p_1(x) = \frac{V Z_{id} \sum_{\text{trials}} \int d\mathbf{Q}^N ds_i d\{\Gamma_{\text{trial}}\} \exp(-\beta U(\mathbf{Q}^N)) P_{sel}(\mathbf{Q}_i) w_i \delta(x - \ln w_i)}{Z_{N+1}} \quad (\text{A-4})$$

where we have transformed from real coordinates  $\mathbf{r}_i$  to scaled coordinates  $\mathbf{s}_i$  by factoring out  $V$ , the volume of the system.  $P_{sel}(\mathbf{Q}_i)$  denotes the probability of selecting the actual conformation of the molecule from the given set of trial segments according to the rule given in Eqn. 36. Finally, we multiply and divide by  $Z_N$  and employ the fact that the  $\delta$ -function ensures that  $w_i = \exp(x)$ :

$$p_1(x) = e^x \left( \frac{V Z_{id} Z_N}{Z_{N+1}} \right) \times \frac{\sum_{\text{trials}} \int d\mathbf{Q}^N ds_i d\{\Gamma_{\text{trial}}\} \exp(-\beta U(\mathbf{Q}^N)) P_{sel}(\mathbf{Q}_i) \delta(x - \ln w_i)}{Z_N} \quad (\text{A-5})$$

Finally, we obtain

$$p_1(x) = e^x \frac{V Z_{id} Z_N}{Z_{N+1}} p_0(x) \quad (\text{A-6})$$

or

$$\ln p_1(x) = x + \beta \mu_{ex} + \ln p_0(x)$$

Hence, by constructing a histogram of  $\ln w$  both in system 0 (with  $N$  chains) and in system 1 (with  $N + 1$  chains), we can derive the excess chemical potential of the chain molecules by studying  $\ln p_1(x) - \ln p_0(x)$ . As in the original Bennett/ Shing-Gubbins scheme [65, 66, 67], the method only works if there is a range of  $x$  values where we have good statistics on both  $p_1(x)$  and  $p_0(x)$ . The advantage of this ‘overlapping-distribution’ scheme over the simple ‘Rosenbluth’ particle-insertion method is that, with the present method, sampling problems for long chains will manifest themselves as a breakdown of the overlap of  $p_0$  and  $p_1$ .

## B. Recursive sampling

Using the notation of section 2.7.4, the average number of off-spring of chain  $i$  in generation  $m$  is  $\langle n_m(i) \rangle$ . We wish to compute the average number of copies of a chain molecule that grow to form a *specified* conformation  $\{\mathbf{q}, \Gamma_1, \Gamma_2, \dots, \Gamma_\ell\}$ , where  $\mathbf{q}$  is the position of the first monomer of the chain, and  $\Gamma_m$ , the unit vector specifying the orientation of the  $m$ -th

segment. We denote the *a priori* probability to select this sequence of trial coordinates by:

$$P_0(\mathbf{q}) \prod_{m=1}^{\ell} P_m(\Gamma_m). \quad (\text{B-1})$$

For polymers with discrete conformations, the probabilities  $P_m$  are finite, whereas for continuously deformable molecules the  $P_m$  are proportional to a probability-density  $p_m(\Gamma_m)$  times an infinitesimal volume element  $d\Gamma_m$ . For what follows, the distinction is unimportant. The number of copies that has the correct orientation in the  $m$ -th generation is best computed using the generating function  $\mathcal{F}_m(z)$ :

$$\mathcal{F}_m(z) = \sum_{k=0}^{\infty} z^k \mathcal{P}_m(k) \quad (\text{B-2})$$

where  $\mathcal{P}_m(k)$  is the probability that  $k$  copies of the specified conformation occur in generation  $m$ . Let us compute  $\mathcal{F}_m$  recursively. First consider the monomer:

$$\mathcal{F}_0(z) = P_0(\mathbf{q})z + (1 - P_0(\mathbf{q})). \quad (\text{B-3})$$

In the next generation, we will generate on average  $n_0 \equiv \pi_0 \exp(-\beta u_0(\mathbf{r}^N))$  dimers. As in section 2.7.4,  $n_0$  is written as the sum of an integer part  $i_0$  and a fractional part  $f_0$ . The generating function  $\mathcal{F}_0(z)$  is modified by replacing the argument  $z$  by  $\mathcal{F}_1(z)$ . This takes into account that every monomer may generate a number of dimers:

$$\begin{aligned} \mathcal{F}_0(z) &= \mathcal{F}_0(\mathcal{F}_1(z)) \\ &= \mathcal{F}_0\left((1 - f_0)[P_1(\Gamma_1)(z) + (1 - P_1(\Gamma_1))]^{i_0} \right. \\ &\quad \left. + f_0[P_1(\Gamma_1)(z) + (1 - P_1(\Gamma_1))]^{i_0+1}\right) \end{aligned} \quad (\text{B-4})$$

We can subsequently insert  $\mathcal{F}_2(z)$  in the expression for  $\mathcal{F}_1(z)$ , and so on. The final expression that we need is

$$\begin{aligned} \mathcal{F}_0(z) &= \mathcal{F}_0\left(\mathcal{F}_1\left(\dots\left((1 - f_{\ell-1})[P_{\ell}(\Gamma_{\ell})(z) + (1 - P_{\ell}(\Gamma_{\ell}))]^{i_{\ell-1}} \right. \right. \right. \\ &\quad \left. \left. + f_{\ell-1}[P_{\ell}(\Gamma_{\ell})(z) + (1 - P_{\ell}(\Gamma_{\ell}))]^{i_{\ell-1}+1}\right)\dots\right) \end{aligned} \quad (\text{B-5})$$

We can now compute the average number of chains with the correct conformation:

$$\langle N(\mathbf{q}, \Gamma_1, \Gamma_2, \dots, \Gamma_{\ell}) \rangle = \left( \frac{z d\mathcal{F}_0(z)}{dz} \right)_{z=1}.$$

We use the chain rule to carry out the differentiation to find

$$\langle N(\mathbf{q}, \Gamma_1, \Gamma_2, \dots, \Gamma_\ell) \rangle = \left( \prod_{i=0}^{\ell} \pi_i \right) P_0(\mathbf{q}) \prod_{m=1}^{\ell} P_m(\Gamma_m) \exp(-\beta U_\ell(\mathbf{r}^N)) \quad (\text{B-6})$$

where we have used the fact that  $\mathcal{F}_m(z=1)=1$ , for all  $m$ . Eqn. B-6 demonstrates that the number of copies of a given polymer conformation is directly proportional to its Boltzmann weight. This result can be used to prove the relation between the excess chemical potential that was used in section 2.7.4

## Acknowledgments

The work of the FOM Institute is part of the research program of FOM and is supported financially by the Netherlands Organization for Research (NWO). I gratefully acknowledge help from: Marjolein Dijkstra, Peter Bolhuis, Maarten Hagen, Maria-Jose Ruiz, Peter Bladon, Matt Glaser, Thierry Biben, Christopher Lowe, Martin van der Hoef, Germonda Mooij, Evert Jan Meijer, Pieter-Rein ten Wolde, Thanasis Panagiotopoulos, Sanat Kumar and Berend Smit. I thank dr. Grassberger for sending me re- and pre-prints on the recursive sampling scheme and John Valleau for drawing my attention to ref. [98].

## References

1. Gibson, J.B., Goland, A.N., Milgram, M. and Vineyard, G.H. (1960) *Phys. Rev.*, **120**,1229–1253.
2. Kremer, K., Grest, G.S. and Carmesin, I. *Phys. Rev. Lett.*, **61**,566.
3. Kremer, K. (1993) In M. P. Allen and D. J. Tildesley, editors, *Computer Simulation in Chemical Physics*, pages 397–459. Kluwer, Dordrecht.
4. Smit, B., Hilbers, P.A.J., Esselink, K., Rupert, L.A.M., van Os, N.M. and Schlijper, A.G. (1991) *Nature*, **348**,624.
5. Smit, B. (1993) In M. P. Allen and D. J. Tildesley, editors, *Computer Simulation in Chemical Physics*, pages 461–472. Kluwer, Dordrecht.
6. Binder, K. (1979) *Monte Carlo Methods in Statistical Physics*, Springer, Berlin.
7. Binder, K. (1984) *Applications of the Monte Carlo Method in Statistical Physics*, Springer, Berlin.
8. Binder, K. (1992) *The Monte Carlo Method in Condensed Matter Physics*, Springer, Berlin.
9. Mouritsen, O.G. (1984) *Computer Studies of Phase Transitions and Critical Phenomena*, Springer, Berlin.
10. Alder, B.J. and Wainwright, T.E. (1962) *Phys. Rev.*, **127**,359.
11. Toxvaerd, S. and Praetgaard, E. (1977) *J.Chem. Phys.*, **67**,5291.
12. Ladd, A.J.C. and Woodcock, L.V. (1978) *J. Phys. C.*, **11**,3565.
13. Ladd, A.J.C. and Woodcock, L.V. (1978) *Mol. Phys.*, **36**,611.
14. Cape, J.N. and Woodcock, L.V. (1978) *Chem. Phys. Lett.*, **59**,271.
15. Hiwatari, Y., Stoll, E. and Schneider, T. (1978) *J.Chem. Phys.*, **68**,3401.
16. Ueda, A., Takada, J. and Hiwatari, Y. (1981) *J. Phys. Soc. Jpn.*, **50**,307.
17. Panagiotopoulos, A.Z. (1987) *Mol. Phys.*, **61**,813.

18. Panagiotopoulos, A.Z. (1987) *Mol. Phys.*, **62**,701.
19. Panagiotopoulos, A.Z. , Quirke, N., Stapleton, M. and Tildesley, D.J. (1988) *Mol. Phys.*, **63**,527.
20. Smit, B., de Smedt, Ph. and Frenkel, D. (1989) *Mol. Phys.*, **68**,931.
21. Kofke D.A., (1993) *Mol. Phys.*, **78**,1331.
22. Kofke D.A., (1993) *J. Chem. Phys.*, **98**,4149.
23. Agrawal, R., Mehta, M. and Kofke, D.A. preprint.
24. Parrinello, M. and Rahman, A. (1980) *Phys. Rev. Lett.*, **45**,1196.
25. Parrinello, M. and Rahman, A. (1981) *J. Appl. Phys.*, **52**,7182.
26. Hoover, W.G. and Ree, F.H. (1967) *J. Chem. Phys.*, **47**,4873.
27. Hansen, J.P. and McDonald, I.R. (1986) *Theory of Simple Liquids*, Academic Press, London, (2nd edition).
28. van Gunsteren, W.F., Weiner, P.K. and Wilkinson, A.J. (1993) *Computer simulation of biomolecular systems: theoretical and experimental applications*, Escom (Leiden).
29. Watanabe, M. and Reinhardt, W.P. (1990) *Phys. Rev. Lett.*, **65**,3301-3304.
30. Actually, the relation between adiabatic transformations in *mechanics* and in *thermodynamics* is subtle. Any time-dependent external perturbation of a classical many-body system will not change the total volume in phase space that is occupied by the system. However, if the external perturbation proceeds too rapidly, then the final energy of the system will depend on the initial conditions. As a consequence, a system that was initially characterized by the thermodynamic variables  $N, V$  and  $E$ , ends up in a state that does not correspond to any known ensemble. In contrast, if the change of the Hamiltonian is truly adiabatic, then the energy of the system is a unique function of time ( $E(t)$ ). The system can then be characterized at every time  $t$  by a set of parameters  $N, V, E(t)$ , while the entropy of the system remains constant (because of the conservation of phase-space volume).
31. Straatsma, T.P., Berendsen, H.J.C. and Postma, J.P.M. (1986) *J. Chem. Phys.*, **85**,6720.
32. Frenkel D. and Ladd, A.J.C. (1984) *J. Chem. Phys.*, **81**,3188.
33. Meijer, E.J., Frenkel, D. LeSar, R.A. and Ladd, A.J.C. (1990) *J. Chem. Phys.*, **92**,7570.
34. Frenkel, D. (1986) In G. Ciccotti and W. G. Hoover, editors, *Molecular Dynamics Simulations of Statistical Mechanical Systems*, pages 151-188, North-Holland, Amsterdam.
35. Meijer, E.J. and Frenkel, D. (1991) *J. of Chemical Physics*, **94**,2269.
36. Frenkel, D. and Mulder, B.M. (1985) *Mol. Phys.*, **55**,1171.
37. Stroobants, A., Lekkerkerker, H.N.W. and Frenkel, D. (1987) *Phys. Rev.*, **A36**,2929.
38. Frenkel, D., Lekkerkerker, H.N.W. and Stroobants, A. (1988) *Nature*, **332**,822.
39. Veerman, J.A.C. and Frenkel, D. (1991) *Phys. Rev.* , **A43**,4334.
40. Veerman, J.A.C. and Frenkel, D. (1990) *Phys. Rev.* , **A41**,3237.
41. Veerman, J.A.C. and Frenkel, D. (1992) *Phys. Rev.* , **A45**,5633.
42. Kranendonk, W.G.T. and Frenkel, D. (1991) *Mol. Phys.*, **72**,679.
43. Eldridge, M.D., Madden, P.A. and Frenkel, D. (1993) *Nature*, **365**,35-37.
44. Kranendonk, W.G.T. and Frenkel, D. (1991) *Mol. Phys.*, **72**,699.
45. Eldridge, M.D., Madden, P.A. and Frenkel, D. (1993) *Mol. Phys.*, **79**,105-120.
46. Eldridge, M.D., Madden, P.A. and Frenkel, D. (1993) *Mol. Phys.*, **80**,987-996.
47. Bolhuis, P.G., Hagen, M.H.J. and Frenkel, D. *Phys. Rev. E* (in press).
48. Hagen, M.H.J., Meijer, E.J., Mooij, G.C.A.M., Frenkel, D. and Lekkerkerker, H.N.W. (1993) *Nature*, **365**,425-426.
49. Meijer, E.J. and Frenkel, D. (1994) *J. Chem. Phys.*, **100**,6873-6887.
50. Hagen, M.H.J. and Frenkel, D. *J. Chem. Phys.* (in press).
51. Bolhuis, P.G. and Frenkel, D. *J. Chem. Phys.* (in press).
52. Dijkstra, M. and Frenkel, D. (to be published). See also: Dijkstra, M. (1994) thesis,

Utrecht.

53. Widom, B. (1963) *J. Chem. Phys.*, **39**,2808.
54. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953) *J. Chem. Phys.*, **21**,1087.
55. Shing, K.S. (1985) *Chem. Phys. Lett.*, **119**,149.
56. Shing, K.S. (1986) *J. Chem. Phys.*, **85**,1105.
57. Sindzingre, P., Ciccotti, G., Massobrio, C. and Frenkel, D. (1987) *Chem. Phys. Lett.*, **136**,35.
58. Lustig, R. (1994) *J. Chem. Phys.*, **100**,3048–3059.
59. Lustig, R. (1994) *J. Chem. Phys.*, **100**,3060–3067.
60. Adams, D.J. (1974) *Mol. Phys.*, **28**,1252.
61. Heinbuch, U. and Fischer, J. (1987) *Mol. Simul.*, **1**,109.
62. Siepmann, J.I., McDonald, I.R. and Frenkel, D. (1992) *J. Phys. Condensed Matter*, **4**,679.
63. Shing, K.S. and Chung, S.T. (1987) *J. Phys. Chem.*, **91**,1674.
64. Sindzingre, P., Massobrio, C., Ciccotti, G. and Frenkel, D. (1989) *Chemical Physics*, **129**,213.
65. Shing, K.S. and Gubbins, K.E. (1982) *Mol. Phys.*, **46**,1109.
66. Shing, K.S. and Gubbins, K.E. (1983) *Mol. Phys.*, **49**,1121.
67. Bennett, C.H. (1976) *J. Comput. Phys.*, **22**,245.
68. Ferrenberg, A.M. and Swendsen, R.H. (1989) *Phys. Rev. Lett.*, **63**,1195.
69. Rosenbluth, M.N. and Rosenbluth, A.W. (1955) *J. Chem. Phys.*, **23**,356.
70. Müller, M. and Paul, W. (1994) *J. Chem. Phys.*, **100**,719–724.
71. Kumar, S.K., Szeifer, I. and Panagiotopoulos, A.Z. (1991) *Phys. Rev. Lett.*, **66**,2935.
72. Mon, K.K. and Griffiths, R.B. (1985) *Phys. Rev. A*, **31**,956.
73. Harris, J. and Rice, S.A. (1988) *J. Chem. Phys.*, **88**,1298.
74. Siepmann, J.I. (1990) *Mol. Phys.*, **70**,1145.
75. Frenkel, D. and Smit, B. (1992) *Mol. Phys.*, **75**,983.
76. Mooij, G.C.A.M., Frenkel, D. and Smit, B. (1992) *J. Phys. Condensed Matter*, **4**,3053.
77. Smit, B. and Siepmann, J.I. (1994) *Science*, **264**,1118–1120.
78. Smit, B. and Siepmann, J.I. (1994) *J. Phys. Chem.* (in press).
79. de Pablo, J.J., Laso, M. and Suter U.W. (1992) *J. Chem. Phys.*, **96**,6157–6162.
80. Frenkel, D. (1990) *J. Phys. Condensed Matter*, **2**(SA),265.
81. Mooij, G.C.A.M. and Frenkel, D. (1991) *Mol. Phys.*, **74**,41.
82. Allen, M.P. and Tildesley, D.J. (1987) *Computer Simulation of Liquids*, Clarendon, Oxford.
83. Batoulis, J. and Kremer, K. (1988) *J. Phys. A: Math. Gen.*, **21**,127–146.
84. Mooij, G.C.A.M. and Frenkel, D. (1994) *J. Phys. Condensed Matter*, **6**,3879–3888.
85. Grassberger, P. (1993) *J. Phys. A: Math. Gen.*, **26**,1023–1036.
86. Grassberger, P. (1993) *J. Phys. A: Math. Gen.*, **26**,2769–2776.
87. Hegger, R. and Grassberger, P. preprint.
88. Torrie, G.M. and Valleau, J.P. (1977) *J. Comp. Phys.*, **23**,187.
89. Valleau, J.P. (1991) *J. Comp. Phys.*, **96**,193–216.
90. Valleau, J.P. (1991) In M. Meyer and V. Pontikis, editors, *Computer Simulation in Materials Science*, pages 67–84. Kluwer, Dordrecht.
91. Valleau, J.P. (1977) *J. Chem. Phys.*, **99**,4718–4728.
92. McDonald, I.R. and Singer, K. (1967) *Discuss. Faraday Soc.*, **43**,40.
93. McDonald, I.R. and Singer, K. (1967) *J. Chem. Phys.*, **47**,4766.
94. McDonald, I.R. and Singer, K. (1969) *J. Chem. Phys.*, **50**,2308.
95. Lyubartsev, A.P., Martsinovski, A.A., Shevkunov, S.V. and Vorontsov-Vel'yaminov, P.N. (1992) *J. Chem. Phys.*, **96**,1776–1783.
96. Marinari, E. and Parisi, G. (1992) *Europhysics Letters*, **19**,451–458.
97. Geyer, C.J. (1991) Markov chain monte carlo maximum likelihood. In *Computing*

- Science and Statistics: Proceedings of the 23rd Symposium on the interface*, pages 156–163.
98. Geyer, C.J. and Thompson, E.A. (1994) Technical Report, School of Statistics, University of Minnesota No.589.
  99. Nezbeda, I. and Kolafa, J.A. (1991) *Molec. Sim.*, **5**,391–403.
  100. Shing, K.S. and Azadipour, S.R. (1992) *Chem. Phys. Lett.*, **190**,386.
  101. Vega, L.F., Shing, K.S. and Rull, L.F. (1994) *Mol. Phys.*, **82**,439–453.
  102. Frantz, D.D., Freeman, D.L. and Doll, J.D. (1990) *J. Chem. Phys.*, **93**,2769–2784.
  103. Bennett, C.H. (1975) In A. S. Nowick and J. J. Burton, editors, *Diffusion in solids: Recent developments*, pages 73–113. Academic Press, New York.
  104. Ciccotti, G. (1991) In M. Meyer and V. Pontikis, editors, *Computer Simulation in Materials Science*, pages 119–137. Kluwer, Dordrecht.
  105. Chandler, D. (1987) *Introduction to Modern Statistical Mechanics*, Oxford University Press, New York.
  106. van Duijneveldt, J.S. and Frenkel, D. *J.Chem.Phys.*, **96**,4655.
  107. Crippen, G.M. (1982) *J. Comput. Chem.*, **3**,471–476.
  108. Duane, S., Kennedy, A.D., Pendleton, B.J. and Roweth, D. (1987) *Phys. Lett.*, **195 B**,216–222.
  109. Tuckerman, M., Berne, B.J. and Martyna, G.J. (1992) *J. Chem. Phys.*, **97**1990–2001.
  110. Mehlig, B., Heermann, D.W. and Forrest, B.M. (1992) *Mol. Phys.*, **76**,1347–1357.
  111. Mehlig, B., Heermann, D.W. and Forrest, B.M. (1992) *Phys. Rev. B*, **45**,679–685.
  112. Forrest, B.M. and Suter, U.W. (1994) *Mol. Phys.*, **82**393–410.
  113. Kremer, K. and Binder, K. (1988) Monte carlo simulations of lattice models for macromolecules. *Computer Physics Reports*, 7:259–310.
  114. Siepmann, J.I. and Frenkel, D. (1992) *Mol.Phys.*, **75**,59.
  115. de Pablo, J.J., Laso, M. and Suter U.W. *J. Chem. Phys.*, **96**,2394–2403.
  116. Mooij, G.C.A.M., Frenkel, D. and Smit, B. (1992) *J. Phys. Condensed Matter*, **4**,L255.
  117. Siepmann, J.I., Karaborni, S. and Smit, B. (1993) Vapor-liquid equilibria of model alkanes. *J. Am. Chem. Soc.*, **115**,6454–6455.
  118. Siepmann, J.I., Karaborni, S. and Smit, B. (1993) *Nature*, **365**,330–332.
  119. Mooij, G.C.A.M. (1993) Thesis, Utrecht University.
  120. Mooij, G.C.A.M. and Frenkel, D. to be published.
  121. Dijkstra, M. and Frenkel, D. (1994) *Phys. Rev. Lett.*, **72**,298–300.
  122. Dijkstra, M., Frenkel, D. and Hansen, J.-P. (1994) *J. Chem. Phys.*, **101**,3179.
  123. Biben, T. (1993) Thesis, Université Claude Bernard, Lyon, France.
  124. Shelley, J.C. and Patey, G. (1994) *J. Chem. Phys.*, **100**,8265–8270.
  125. Swendsen, R.H. and Wang, J.S. (1987) *Phys. Rev. Lett.*, **58**,86.
  126. Wu, D., Chandler, D. and Smit, B. (1992) *J. Phys. Chem.*, **96**,4077–4083.
  127. van Duijneveldt, J.S. and Lekkerkerker, H.N.W. (1993) *Phys. Rev. Lett.*, **71**,4264–4266.
  128. Meijer, E.J. and Frenkel, D. (1991) *Phys. Rev. Lett.*, **67**,1110.
  129. Napper, D.H. (1983) *Polymeric Stabilization of Colloidal Dispersions*, Academic Press.
  130. Frenkel, D. and Louis, A.A. (1992) *Phys. Rev. Lett.*, **68**,3363.
  131. Ceperley, D.M. and Pollock, E.L. (1986) *Phys. Rev. Lett.*, **56**,351.
  132. Alavi, A. and Frenkel, D. (1992) *J.Chem. Phys.*, **97**,9249–9257.
  133. Frisch, U., d’Humières, D., Hasslacher, B., Lallemand, P., Pomeau, Y. and Rivet, J.-P. (1987) *Complex Systems*, **1**,649.
  134. Ladd, A.J.C. (1991) In M. Meyer and V. Pontikis, editors, *Computer Simulation in Materials Science*, page 481. Kluwer, Dordrecht.
  135. van der Hoef, M.A. and Frenkel, D. (1992) In M. Mareschal and B.L. Holian, editors, *Microscopic Simulations of Complex Hydrodynamic Phenomena*, pages 205–220. Plenum, New York.

136. Alder, B.J. and Wainwright, T.E. (1970) *Phys. Rev. A*, **118**.
137. Dorfman, J.R. and Cohen, E.G.D. (1970) *Phys. Rev. Lett.*, **25**,1257.
138. Dorfman, J.R. and Cohen, E.G.D. (1972) *Phys. Rev.*, **A6**,776.
139. Dorfman, J.R. and Cohen, E.G.D. (1975) *Phys. Rev.*, **A12**,292.
140. Ernst, M.H., Hauge, E.H. and van Leeuwen, J.M.J. (1971) *Phys. Rev. A*, **4**2055.
141. Pomeau, Y. and Résibois, P. (1975) *Physics Reports*, **19**,63.
142. Levesque, D. and Ashurst, W.T. (1974) *Phys. Rev. Lett.*, **33**,277.
143. Erpenbeck, J.J. and Wood, W.W. (1982) *Phys. Rev.*, **A26**,1648.
144. Erpenbeck, J.J. and Wood, W.W. (1985) *Phys. Rev.*, **A32**,412.
145. Ernst, M.H. (1990) In J.P.Hansen and D.Levesque, editors, *Liquids, Freezing and Glass Transition*, pages 43–143. North-Holland, Amsterdam.