Randi Toreskås Holta

Nosé-Hoover simulations of aqueous mixtures of small alcohols

Abstract

We have performed constant energy and constant temperature equilibrium molecular dynamics simulations of model liquids of water, methanol, and ethanol, and of the binary aqueous mixtures of the alcohols at mole fractions alcohol of 0.25, 0.50, and 0.75 at room temperature. The total number of molecules is 256 and the integration timestep is 0.5fs.

The molecules are modelled with rigid, nonpolarizable effective site potentials. The TIP4P model is applied for water, and the OPLS models for methanol and ethanol (*trans*).

The thermostatically controlled simulations are performed within the Nosé-Hoover formalism, with separate temperature control of the translational and rotational degrees of freedom. Reservoir coupling in the mixtures is determined as a weighted average of the liquid values.

We present enthalpies of vaporization and excess configurational energies for the mixtures. Further, we show all radial site-site correlation functions for the self- and cross-interactions. We also present results for self-diffusion coefficients for the components, and the velocity auto correlations functions.

The NVT simulations are investigated with respect to the possibility of reproducing the canonical ensemble. We find that the simulations are not entirely within the premises of the canonical ensemble. We suggest that this is due to a too weak coupling (too large heat bath mass). The results are nevertheless found to be in good agreement with experiments and published simulational studies.

The NVE simulations for the mixtures reproduce the desired temperature well.

The energy conservation is good and improves with decreasing water content.

Effects of using only trans-ethanol might be seen for the self-diffusion coefficients.

Structural results for both the model alcohols are found to be consistent with V-chains. Upon addition of water, methyl-methyl coordination numbers for both alcohols decrease less than the hydroxyl coordination numbers.

Preface

Dear reader,
what you hold in your hands is a result of more than seven years of work. I might not have the World Record in time consumption for completing a doctoral work, but it surely was a looong time.

I started out in 1989 with a background in astrophysics. It took some time to adjust from astronomical dimensions and states to the microscopic view of liquids, but I feel fortunate to have had the opportunity to study both the largest objects and some of the smallest.

The work has been performed at Telemark College's Department of Technology, under the supervision of Professor Bjørn Kvamme. The program was developed and tested on our CONVEX 200 supercomputer, and the production simulations were performed with our DEC Alpha 4000/300 workstations. The data are presented graphically with MATLAB, and the document is created with DECwrite.

By far the most important achievements during these years have been the births of my two children, Nils and Vigdis. Unfortunately, being both a mother and a dr. ing student appear to be mutually exclusive - at least I have felt so. It has been years of a constant feeling of falling short. I can therefore not recommend my choice to my fellow-females, even if I admit that I most likely would have done the same things all over again.

Thinking of my family, I look forward to spending a year off in Canada, entirely devoted to the welfare of these three people I care the most for.

Enjoy reading, I promise that this will be my one and only dr. ing. thesis!

Acknowledgements

First of all, I would like to thank my mother. She and my late father have always taken an interest in my school work and studies, and have encouraged me to fully utilize my abilities.

Then there is Olav, my husband of fourteen years. You have been the one and only person that never has had any doubt that I would complete this work. Thank you for all support, and for staying at my side for 'better and worse'!

My supervisor Professor Bjørn Kvamme, I have learned to know as a man of very fine qualities. I am particularly grateful because of your patience, and because you let me follow my own path through phase space (*sic*).

Thanks also to Professor Johan S. Høye At the Norwegian University of Science and Technology, for leading me into the mysteries of polyatomic liquids back in 1989.

Ole Kristian Førrisdal, my fellow dr. ing student, has worked hard during the past months to get an acknowledgement. And you do deserve a big hug for your thorough reading of my manuscript, thereby saving me a lot of embarrassment, and also for a wealth of useful information.

Senior engineer Eric Trøim of the local Department EDP, has helped me out more times than I am able to count. This man does not have the word 'no' in his vocabulary!

Professor Søren Toxvaerd and Professor Bill Hoover are greatly acknowledged for explaining me some tricks about thermostatting.

Contents

Chapter 1
Introduction

The topic of this work is computer simulation of three model liquids; water, methanol, and ethanol, both as pure species and as binary aqueous solutions at various concentrations. The simulation method is molecular dynamics with temperature control [1, 2]. The purpose of the simulations is twofold, the implementation and verification of this particular simulation method, and the calculation of thermodynamical, structural, and dynamical equilibrium quantities for the model mixtures.

To be able to calculate the various quantities, we need a model. The model consists of a mathematical description of the system together with simplifying assumptions regarding the physical nature of the real systems. We can therefore not claim to simulate the fluid itself, but only a model fluid who resembles the real fluid to some extent.

1.1 Background

1.1.1 The liquids

Methanol, ethanol, and water are important in a wide variety of industrial and life processes [3]. Their main use in industrial processes is as raw materials in chemical synthesis, for instance the production of acetic acid and formaldehyde from methanol and ethers from ethanol. They can also be used as motor fuels as an alternative to petroleum based fuels. Methanol has technical applicability as absorbing agent in gas scrubbers and as working fluid in refrigeration systems. Ethanol is also the next important solvent after water, and is used for instance in the cosmetic industry and for drug production.

Because of their importance, it is of general interest to understand and predict their behaviour. Even for such commonplace systems as aqueous solutions of alcohols, the mechanisms of solvation, and the internal structure, are not fully understood. Liquids are in general difficult to study and to model because they have long range disorder like a gas, and short range order like a solid. For the majority of molecular liquids, the molecular shape can not be regarded as spherical, so their individual orientations will be of importance. Orientations become even more important when the molecules possess dipole moments, ie. are polar. In addition to volumetric effects there will then also be electrostatic influence of both short and long range.

Water, methanol, and ethanol all belong to a particularly interesting group of polar molecules containing hydrogen and oxygen. For both alcohols and water, a common feature is the OH-group, giving raise to their hydrogen bonding ability. The hydrogen bonding is most pronounced in water, where it for example is responsible for the well-known decrease in density upon freezing, and the large enthalpy of vaporization [4]. The hydrogen bonding is also thought to be responsible for some anomalies of the aqueous mixtures of lower alcohols. When small amounts of alcohol are mixed with water, experimental results show a volume contraction, and an exothermic mixing with negative excess enthalpy [5].

From a simulational point of view, the main interest has been the hydrogen bonding, and how it influences the internal arrangement of molecules. The water liquid structure seem to be dominated by a hydrogen bonded network extending in all three dimensions [6, 7, 8], while the molecules of liquid methanol and ethanol are believed to associate in chains [9, 10, 11, 12]. Related questions are by what mechanism do the alcohols solvate in water, and what is the effect of the (nearly) non-polar methyl group in ethanol upon hydrogen bonding in the solution.

1.1.2 Getting information about matter

Within the area of statistical thermodynamics, a fundamental goal is to get information of physical systems, both in order to predict how their behaviour will be, and to explain the causes of their (observed) behaviour. Broadly speaking, three different fields of research collaborate in the establishment of such knowledge, see Figure 1.1:

- experimental measurements on real, macroscopic systems
- theoretical calculations where statistical mechanical theories are tested against or applied to model systems
- computer simulations where in principle exact results for model systems can be achieved

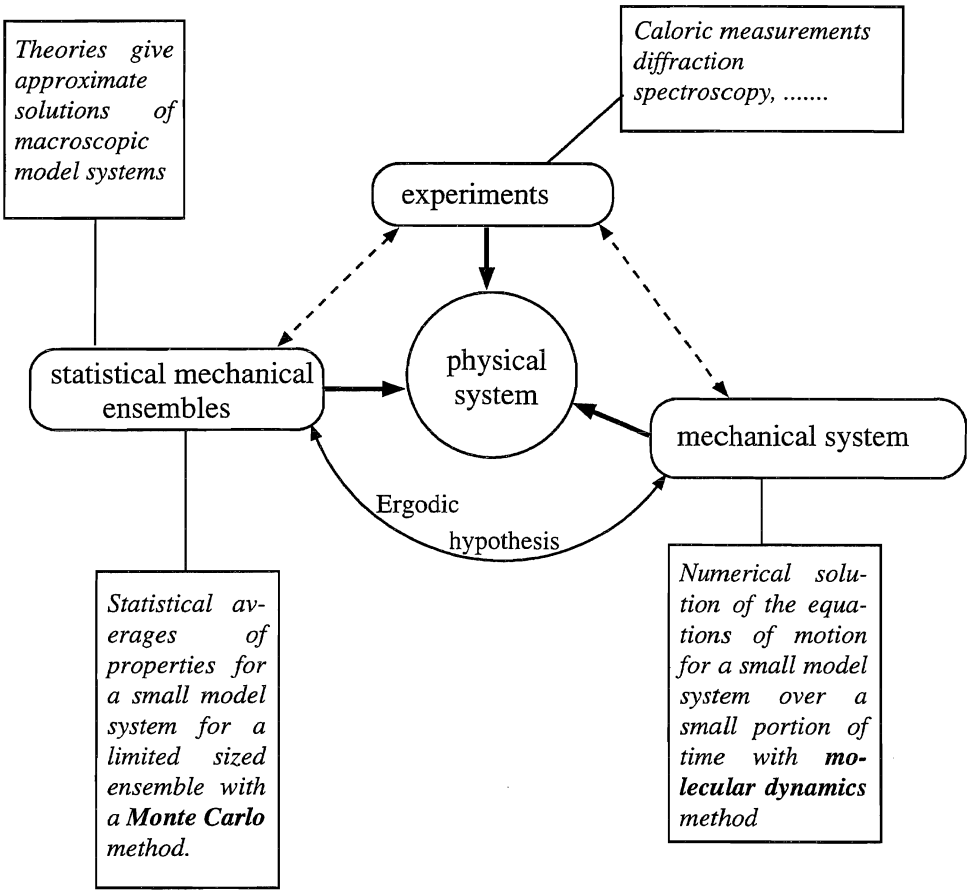


Figure 1.1 We all want to get information of matter, but the approaches are different.

There is an intimate interplay between the three approaches. A particular model of the physical system can for instance be evaluated by comparing the simulated results to experimental values, or a new theory of matter can be compared to simulation results for the same model. Also the interpretation of experimental results can improve with insight gained from theoretical calculations or simulations.

1.1.3 Microscopic vs. macroscopic view of matter

Fundamental to each of the three routes above is the microscopic view of matter as a collection of individual particles which each contribute to the macroscopic observables. Bulk properties are then merely statistical averages of the molecular properties [13, 14]. A well known example is the observable temperature, which is usually taken as the average of the sum of the kinetic energies of the molecules.

Whereas from a macroscopic point of view the state of a one-component system is determined when, say, its pressure P and temperature T are specified, the microscopic specification needs values for individual molecular positions and velocities. A particular macroscopic state at given P and T can, however, be realized through a variety of microscopic arrangements, thus making the notion of (thermo)dynamical equilibrium sensible. The bridge between these microscopic configurations and the macroscopic observables is provided by statistical mechanics together with a suitable model for the forces acting between the molecules.

1.1.4 Computer simulations

Molecular simulations can be conducted within two different strategies [15]:

- Monte Carlo methods, which rely on a random procedure to generate microscopic configurations
- molecular dynamics methods in which configurations are obtained by integrating the equations of motion for all particles in the system

There are several different methods to chose from within each of the two categories. Combinations of the basic ideas of the two approaches can also be favorable in some situations.

The basic difference between Monte Carlo methods and molecular dynamics is the stochastic vs. deterministic generation of configurations. A Monte Carlo simulation yields a property average over different microscopic states, or an ensemble, and a molecular dynamics simulation gives a dynamical average for the system over a finite time. The former solves the problem within the framework of statistical mechanics, while the second is a deterministic approach. The two simulation methods answer then the same questions from different points of view, and the numerical values are expected to be equal within the statistical precision.

The Monte Carlo methods are popular and relatively simple methods as compared to molecular dynamics. The advantage of molecular dynamics is the ability to compute time dependent quantities directly, since the motion of each particle is followed in time. For both methods, the system under study must be limited to a manageable number of molecules, and the molecular interactions must be specified.

1.1.5 Molecular models

Even if we aim at determining properties for the real liquids, the molecular interactions are not known in detail, and are also so complex that we need simplified models. The simplifying assumptions reduce the computing time, but we must stress that the results are only strictly valid for the hypothetical model liquid.

A common set of assumptions/approximations is [15]

- Classical mechanics applies.
- The molecules are rigid bodies with fixed geometries.
- Forces between molecules act between spherically symmetric interaction sites residing in the molecules. Not all atoms are represented by sites, and the sites need not coincide with the centre of mass of an atom. This is known as the site-site approximation.
- Only interactions between pair of sites are considered, which is the pairwise additive assumption.

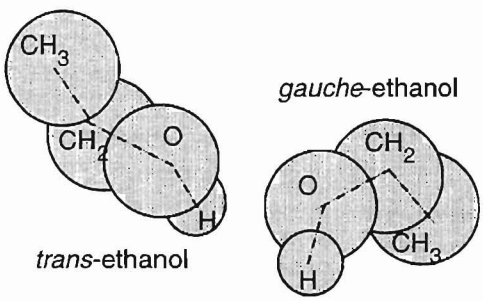


Figure 1.2
Representation of ethanol molecule. The molecule has four sites; CH_3, CH_2, O, and H, each described by parameters ε, σ and q (see text). Geometry of molecule is defined through bond lenghts and bond angles.

Within the site approximation, single atoms or groups of atoms in the molecule are regarded as sites of interaction and assigned separate potential parameters. The potential energy between two sites *a* and *b* in different molecules can be described by a Lennard-Jones potential for the short ranged part, and by a Coulombian part for the long ranged electrostatic interactions arising from the dipoles:

$$u_{ab}(r_{ab}) = 4\epsilon_{ab}\left(\frac{\sigma_{ab}^{12}}{r_{ab}^{12}} - \frac{\sigma_{ab}^{6}}{r_{ab}^{6}}\right) + k_C\frac{q_a q_b}{r_{ab}} \qquad (1.1)$$

where r_{ab} is the distance between the two sites a and b. ε_{ab} is the strength of the at-

traction between two sites, and σ_{ab} is the smallest distance the two sites can approach without repelling each other. q_a and q_b are fractional charges assigned to account for the dipole moment, and k_C is the Coulomb constant. The molecule as a whole is of course electrostatically neutral. This simple and schematic representation of the molecule can be remarkably powerful.

Usually the model parameters are adjusted to reproduce a few physical data. It is then not guaranteed that other properties follow correctly. Further, the parameters are only valid for a limited temperature range. The reliability of the results, whether a simulation or a theoretical calculation, then rely heavily on the accuracy of the interaction model.

1.2 Molecular dynamics - choice of surroundings

1.2.1 Isolated system versus closed system

Normally, molecular dynamics simulations are performed for isolated systems, that is the system can not exchange neither energy nor mass with its surroundings. According to the 1st law of thermodynamics, the total energy E, number of particles N, and volume V of the system must be constant. This simulation corresponds closely, but not exactly [16], to the microcanonical or NVE ensemble of statistical mechanics.

The total energy is a constant of the motion, that is, its value does not change with time, for the NVE system, but the distribution of energies among the different degrees of freedom is undetermined. A particular problem is the system temperature, which is not exactly known until a simulation is finished. This is impractical because experiments are normally conducted at fixed temperatures. (The same argument also goes for the system pressure. We will however not be concerned with pressures.)

To fix the temperature, one can use a closed system, see Figure 1.3. This is a system that has a constant number of particles, but can interact with a thermal reservoir in the surroundings by exchange of energy. If the thermal reservoir is assigned a temperature, this must also be the temperature of the system when thermal equilibrium is achieved.

Looking at the system and the reservoir together as an isolated system, the total energy of the two will still be a constant of the motion, but the total energy of the closed system is allowed to fluctuate. The contact between system and reservoir is maintained through the entire simulation, with energy flowing back and forth between the two of them. The closed system has its analogy in statistical mechanics with the canonical ensemble.

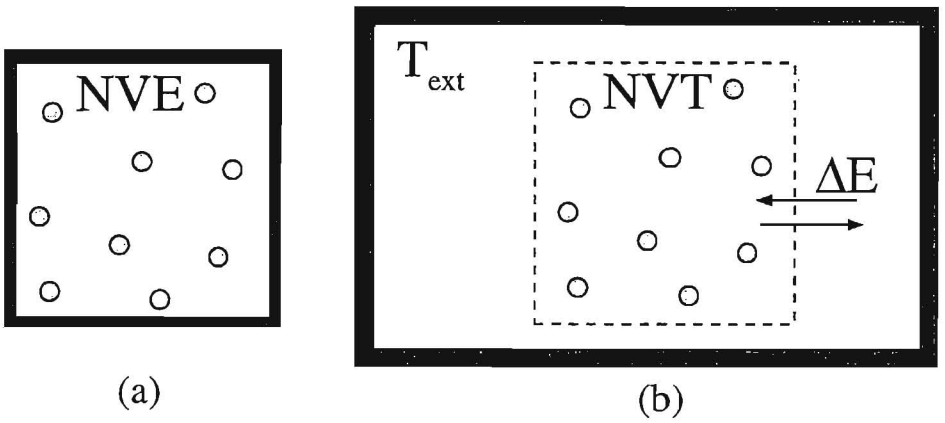


Figure 1.3 a) Isolated system of N particles with constant energy E, and volume V, and b) closed system (dashed lines) of N particles in contact with a thermal reservoir at constant temperature T, volume V. The heat bath is imaginary.

The interaction between the molecular system and the thermal reservoir is purely mathematical, but formulated according to the 1st law of thermodynamics. It appears as an additional force term in the equations of motion for the molecules. If average system temperature is too high, the force slows down the velocities of all molecules. If temperature is too low, all molecules are accelerated. The numerical cost to be paid is small: only one extra equation has to be integrated per timestep, and there are no expensive subcalculations.

The extended system dynamics is due to Nosé [1, 17], but the mathematical formulation was shortly after its publication developed further by Hoover [2]. We will use his form, which has become known as the Nosé-Hoover dynamics. This formulation permits the use of several thermal reservoirs [18] - which we will make use of.

The method is also extended to include control of pressure [17]. The Nosé dynamics or Nosé-Hoover dynamics is applied in non-equilibrium [19, 20] as well as in equilibrium simulations. Applications are also found in investigation of reactive dynamics [21] and in intracrystalline diffusion [22]. Also in quantum mechanical simulations [23] the method is used to keep electrons in their ground state. The proposal of the Nosé dynamics gave raise to a large activity within the field of thermostatted simulation, and a series of extensions have been proposed. A summary is found in Nosé 1991 [24].

1.2.2 A central question - canonical simulation?

Since the publication in 1984 the Nosé and Nosé-Hoover dynamics have been extensively studied from a theoretical point of view. The central question has been whether the extended system simulation corresponds to the canonical ensemble. Not only must property averages be reproduced, but their fluctuations must be within the canonical ensemble. It was proven by Nosé [1] that if the system and its surroundings is ergodic, the simulation will generate distributions (of properties) in the canonical ensemble. The problem is that ergodicity can in general not be proven easily - if at all! It is not sufficient that the system is chaotic - or 'complicated' as our relatively large system of Lennard-Jones and electrostatically interacting particles. Several extensions have been proposed, but until a very recent proposal by Hoover and Holian [25], none of them have been equally simple to implement as the Nosé-Hoover thermostat.

1.3 Description of our work

We will show results from both constant energy and constant temperature molecular dynamics simulations of water, methanol and ethanol as pure liquids and as binary mixtures of water and alcohol at mole fractions 0.25 0.50 and 0.75. The simulations are performed at 298K (water and methanol) and 293K (water and ethanol) at experimental densities. These temperatures are selected since experimental and simulated results occur most often at these temperatures.

We present thermodynamical data for the liquids/mixtures. The structure is represented with site-site radial correlation functions for all mixtures. We have also calculated the self-diffusion coefficient for both components of the mixtures.

We are restricted to the assumptions summarized in Subsection 1.1.5 above, and the molecular interactions are described with the site-site potentials TIP4P, Jorgensen *et al.* 1983 [26] for water and OPLS, Jorgensen, 1986 [27] for methanol and ethanol. These three- and four-site (ethanol) model potentials are attractive because they are simple to implement, and because the overall agreement with experimental results is good [26]. Simulated as a four-site molecule, ethanol can have two conformers, see Figure 1.2, of approximately equal abundance, [27]. Keeping within the rigid body approximation, we have deliberately ignored the presence of conformal equilibrium, and modelled the ethanol molecule as *trans* ethanol.

The total system size is taken to be 256 molecules in all the simulations, and the molecules are confined to a cubic box under periodic boundary conditions. The forces are cut at half box length, and the long-range part of the electrostatic forces is handled with the Ewald summation technique.

Code development and testing have been the most time consuming part of this work. Whenever possible, separate routines are borrowed with a varying degree of adjustment, mainly from Allen and Tildesley [15] and from Haile [28]. Example programs are usually written for the atomic case, and they have to be extended to several sites. The code is not included, but is available upon request.

With the thermostatted simulations, our system of molecules is in thermal contact with two reservoirs. One thermostat selects the translational degrees of freedom for all molecules and maintains the average value at the predefined value, the other controls the temperature of the rotational degrees of freedom. Separate thermostatting of different degrees of freedom is used to assure a proper thermalization of all degrees of freedom, particularly for mixtures. The use of multiple thermostats was suggested by Nosé in 1986 [18], to our knowledge multiple thermostatting has not been applied elsewhere.

Since we have implemented the code ourselves, and also use a multiple thermostat, it is necessary to verify the simulations carefully. Therefore we perform a thorough verification of the program by monitoring several quantities, and by comparing our results to published simulational studies, preferably for the same models. Verification of the use of separate translational and rotational thermostatting is done from comparison between NVE and NVT simulations. Mixtures of methanol and water have been studied by several authors (references appear in Chapter 5), though mainly with Monte Carlo methods and mainly dilute mixtures. There exist published data from simulation of TIP4P-water and OPLS-methanol, which are suited for verification of our simulations together with the existing experimental data. However, the combination of model forces and state conditions are several, and the quantities to calculate are *legio*. Not all of our calculated quantities appear elsewhere, as for instance the full combination of site-site correlation functions and their corresponding coordination numbers.

To our knowledge, there has not been published any molecular dynamics study of any kind of the mixtures of water/ethanol. Also the Monte Carlo and theoretical studies published does not cover the whole range of concentrations. For ethanol and its aqueous mixtures we only know of a handful of published simulation studies [29, 30, 31, 32]. These are all for very dilute mixtures.

We believe that the lack of published results is due to the computer consuming nature of these simulations, but a crucial ingredient is also the selection of a reliable potential model. Computing power is increasing almost daily, and the trend today in several fields of research and industry is that simulations substitute for experiments. Simulations are also a supplement to experiments as they can contribute to an understanding of observed phenomena by suggesting explanations.

1.4 Outline of thesis

The thesis is organized as follows: In Chapter 2 we summarize the basic principles of molecular dynamics (Section 2.1), and state what kind of information a molecular dynamics simulation can provide (Section 2.2). We present some highlights from the history of thermostatted simulations (Section 2.3), and describe the Nosé and Nosé-Hoover thermostats (Section 2.4). The mathematical formulation for multiple temperature control within the Nosé-Hoover framework is set up (Section 2.4) and the parameters are interpreted and discussed (Section 2.5). We discuss the possibility of having a canonical simulation, and the dependence upon heat bath characteristics. Finally (Section 2.6) the constants of the motion are discussed.

Chapter 3 starts with a brief outline of the program and the integration algorithm (Sections 3.1 and 3.2). Then we describe some of the technical details of the simulation (Sections 3.3 and 3.4), and which quantities we calculate (Section 3.5). Additional quantities calculated to verify the simulation are described in Section 3.6.

Chapter 4 describes the molecular models for water (Section 4.3) and for the alcohols (Section 4.4). We discuss briefly the applicability of pure liquid models in mixtures (Section 4.5) and conclude this chapter with a brief summary of polarizable and flexible models (Section 4.6).

In chapter 5 we present results from the simulations of water and methanol, and chapter 6 presents the results for water and ethanol. These chapters start with analysis and discussion of simple thermodynamic quantities (Sections 5.2 and 6.1), then we present results for the verification of the simulations (Sections 5.3 and 6.2), before we move on to the presentation and discussion of structure (Sections 5.4 and 6.3). Finally the results for self diffusion and the velocity auto correlation functions are presented (Sections 5.5 and 6.4).

Several figures and tables related to the discussions in chapter 5 and 6 appear in appendices (Appendix C -water/methanol and Appendix D -water/ethanol).

We conclude our work in Chapter 7, and Chapter 8 provides some ideas for future work.

1.4.1 A note about greek letters

Unfortunately, DECwrite represents greek letters with different fonts in plain text and in the equation editor. Also there has been no possibility of having bold greek letters symbolizing vectors. On the very few occasions where we need greek lettered vectors, they are symbolized with an arrow. We hope this will not lead to misunderstanding.

Chapter 2
Molecular dynamics

This chapter summarizes the general principles of molecular dynamics simulations, with particular emphasis on constant temperature simulations. Our starting point is however with a summary of the basic ideas of constant energy simulations. This is still the most frequently used method and also give a natural reference for discussions on the method of constant temperature simulation. Then, in Section 2.3 we briefly describe and compare four methods of temperature control, of which the Nosé-Hoover thermostat is one. These are not the only possible choices, but they have all been of major importance during the last one or two decades and can illustrate the progress of temperature controlled simulations. In Section 2.4 both the Nosé and the Nosé-Hoover equations of motion are presented. In Section 2.5 we will discuss the ability of the Nosé-Hoover thermostat to generate canonical distributions. This question is closely connected to the choice of heat bath characteristics. Finally, we discuss the constants of the motion.

The treatment will be restricted to classical simulations of liquid equilibrium mixtures of polyatomic mixtures. By classical is meant that the motion is governed by Newtonian or similar mechanics.

We distinguish between four kinds of systems. By **system** we mean the simulated portion of matter consisting of N molecules, either regarded as a mechanical system or as a statistical mechanical ensemble. By **extended system** we mean the system as defined above and its thermal reservoirs taken together. The **physical** or **model system** is a macroscopic amount of matter that interacts according to model forces. Finally, **real system** is the specific kind of bulk matter that we try to assign property values.

2.1 Mathematical model

2.1.1 Statistical mechanics and classical mechanics

When we do measurements on a physical system, the measured quantities are results of time-averaged action (influence) of all or a large portion of the molecules. The contribution from each molecule is a result of its individual position and momentum at a time. As these can vary widely from one moment to the next, all properties will in principle fluctuate around their mean values. For macroscopic systems the fluctuations are not measurable.

A physical system with f internal degrees of freedom can at every instant be regarded as a point in a $2f$-dimensional **phase space** where the particle positions $\mathbf{q}=(q_1,q_2,...,q_f)$ and momenta $\mathbf{p}=(p_1,p_2,...,p_f)$ constitute the coordinate axes [13, 14]. The physical system can move through a series of macroscopic states, where each macroscopic state can be realized by a variety of microscopic different arrangements of particle positions and momenta. Each point in phase space is a unique solution of Hamiltons equations. When the system evolves in time, the point representing the state moves in phase space, thus generating a **phase trajectory.** If the system is at equilibrium or is restricted in any other way, the allowed configurations and momenta are confined to parts of the phase space, and so is also the trajectory. For an infinite time, the trajectory formed will pass through all accessible points in phase space. Knowing the phase trajectory, nearly all information of the system is accessible through relations from classical thermodynamics or statistical mechanics.

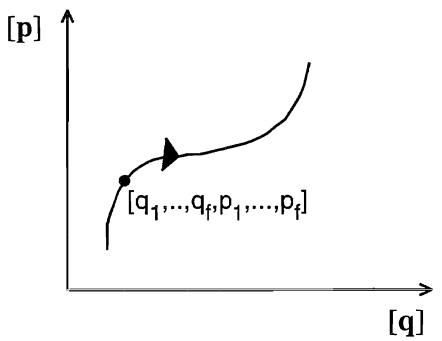


Figure 2.1

Part of a phase space trajectory of a system. **q** are positions of all particles, and **p** are their conjugated momenta.

In statistical mechanics the trajectory is represented as ensembles, a number of replicas of the system corresponding to the same macroscopic observables, but with different microscopic arrangements. If we aim at calculating trajectory averages, we must either have a mechanical model that can be integrated to give the individual motion of the constituent molecules, or we must have some theory that enables us to cal-

culate trajectory averages directly. The latter is the main subject of statistical mechanical theories. We relay on the first approach. The mathematical model is provided through equations of motion together with a model for the forces that cause the motion to change. If we can integrate this model, the system can be followed for as long as we wish.

The equality of results from either of the two approaches is postulated through the ergodic hypothesis, which states that given enough time the time average of a property will equal the ensemble average. Then simulations at, say constant energy, number of particles, and volume will provide results corresponding to the microcanonical ensemble.

In a macroscopically sized system, the number of particles is of the order of Avogadro's number. To keep track of the individual motions of such an amount of matter is not manageable with the computers of today. The system size must therefore be reduced drastically. But then one also reduces the dimensionality of the phase space and the number of available configurations that can generate the chosen macroscopic state. It is not within reach to follow the system for a macroscopic portion of time either, and we must be satisfied with a time interval of the order of nanoseconds. We have then arrived at the manageable task of simulating a tiny portion of a (model) system for a limited period of time.

The central questions to ask at this point are, how well does the simulated phase trajectory of the small model system compare to the phase trajectory of the full model system, and how well does the small portion of the calculated trajectory compare to the full trajectory of the small system? And, finally, how well does the model system resemble a real system?

The cost to be paid for looking at a small system is that the fluctuations of properties will be large: they decrease as $N^{1/2}$ with increasing molecule number N [28]. Also the trajectory averages can be systematically displaced. The trajectory being of finite length, cause a danger of not collecting a representative sample of the available microscopic states. Normally, a system size of more than ~100 molecules is regarded [33] as sufficient for determination of thermodynamic averages and structure. Casulleras and Guardia, 1992 [34] show however that an increase of system size varying from 125 to 512 molecules yield a systematic increase in the self-diffusion coefficient for liquid methanol. They also find the dielectric constant to be dependent of system size. Their study confirms that system size does not influence significantly upon structure.

2.1.2 Equations of motion

From the advent of molecular dynamics simulations in the fifties (Alder and Wainwright, 1957 [35]), calculations for isolated systems have been the standard approach. Then, to have no mass or energy transfer across system boundaries, molecule number N, volume V and total energy E must be constant. They are the macroscopic fixed quantities defining the state of the system in close relation to the statistical mechanical microcanonical ensemble, or the NVE-ensemble. A simulation for which these conditions apply is called an NVE-simulation. If we look at molecules composed of different atoms or group of atoms as constituting a rigid body, the equations of centre-of-mass motion for a classical system of N molecules follow from Newtons 2. law as

$$\frac{d\mathbf{r}_i}{dt} = \frac{\mathbf{p}_i}{m_i} \tag{2.1}$$

and

$$\frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i^{internal} \tag{2.2}$$

where $i=1,2,..,N$. $\mathbf{r}_i$ is centre-of mass position of ith molecule, $\mathbf{p}_i$ is linear momentum of ith particle, and m_i its mass. $\mathbf{F}_i^{internal}$ is the sum of forces acting upon the ith molecule from all molecules $j \neq i$. The equations are thus coupled through the forces, and the equations has to be integrated with time t for each molecule simultaneously. Since the system is isolated, no external forces act upon the system.

For molecules having anisotrope force-fields, one must also integrate the motion with respect to orientation in space, given by

$$\frac{d\mathbf{L}_i}{dt} = \frac{d(\mathbf{I}_i\vec{\omega}_i)}{dt} = \vec{\tau}_i \tag{2.3}$$

where $\mathbf{L}_i$ is angular momentum, $\vec{\omega}_i$ and $\vec{\tau}_i$ are the angular velocity of molecule i and the total torque on molecule i from the rest of the molecules. Since the moment of inertia of a molecule, $\mathbf{I}_i$ relative to a space fixed frame of reference will change as the body rotates, the rotation must be performed relative to axes fixed in the body. Through a standard coordinate transformation [36], Equation (2.3) be-

comes

$$\mathbf{I}_i \cdot \frac{d\vec{\omega}_i^b}{dt} + \vec{\omega}_i^b \times (\mathbf{I}_i \cdot \vec{\omega}_i^b) = \vec{\tau}_i \tag{2.4}$$

where ω^b is angular velocity relative to the body-fixed axes and $\mathbf{I}_i$ is the inertia tensor of molecule *i*. If the body-fixed axes are principal axes, the off-diagonal elements of the inertia tensor vanish, and the differential equations for the components of the angular velocity are Euler's equations. The centre of mass of the molecules is taken as the origin of the body axes.

$$\begin{aligned} I_{1i}\dot{\omega}_{1i}^b + (I_{3i} - I_{2i})\omega_{3i}^b\omega_{2i}^b &= \tau_{1i} \\ I_{2i}\dot{\omega}_{2i}^b + (I_{1i} - I_{3i})\omega_{1i}^b\omega_{3i}^b &= \tau_{2i} \\ I_{3i}\dot{\omega}_{3i}^b + (I_{2i} - I_{1i})\omega_{2i}^b\omega_{1i}^b &= \tau_{3i} \end{aligned} \tag{2.5}$$

where subscripts 1, 2, and 3 refer to the principal axes. The orientation is further described by integrating the Euler angles (ϕ, θ, ψ) [36, 37]

$$\begin{aligned} \dot{\phi}_i &= -\omega_{1i}^s \frac{\sin\phi_i \cos\theta_i}{\sin\theta_i} + \omega_{2i}^s \frac{\cos\phi_i \cos\theta_i}{\sin\theta_i} + \omega_{3i}^s \\ \dot{\theta}_i &= \omega_{1i}^s \cos\phi_i + \omega_{2i}^s \sin\phi_i \\ \dot{\psi}_i &= \omega_{1i}^s \frac{\sin\phi_i}{\sin\theta_i} - \omega_{2i}^s \frac{\cos\phi_i}{\sin\theta_i} \end{aligned} \tag{2.6}$$

where $\omega_{\zeta i}^s$ are are space fixed angular velocity of molecule *i* relative the ζ-axis.

Angular velocities in the space-fixed frame of reference are converted from the body-fixed variables through

$$\vec{\omega}_i^s = \mathbf{A}_i^T \cdot \vec{\omega}_i^b \tag{2.7}$$

where $\mathbf{A}^T$ is the transpose of the rotation matrix [15]

$$\mathbf{A} = \begin{bmatrix} \cos\phi\cos\psi - \sin\phi\cos\theta\sin\psi & \sin\phi\cos\psi + \cos\phi\cos\theta\sin\psi & \sin\theta\sin\psi \\ -\cos\phi\sin\psi - \sin\phi\cos\theta\cos\psi & -\sin\phi\sin\theta + \cos\psi\cos\theta\cos\psi & \sin\theta\cos\psi \\ \sin\phi\sin\theta & -\cos\theta\sin\theta & \cos\theta \end{bmatrix} \tag{2.8}$$

It is also possible to treat the subunits of polyatomic matter as individual objects, each obeying a separate equation of motion. The molecular geometry then enter as constraint forces that keep the molecule together. This constraint dynamics is particularly useful when the molecular model is flexible [15].

The average energy E of an isolated system of N rigid molecules subject to translations and rotations is given by the system Hamiltonian $\mathcal{H}$ as

$$E = \left\langle \mathcal{H} \right\rangle = \sum_i \frac{\mathbf{p}_i^2}{2m_i} + \sum_i \frac{1}{2}\omega_i \cdot \mathbf{I}_i \cdot \omega_i + \mathcal{U}(\mathbf{r}) \tag{2.9}$$

The first term is the kinetic energy due to the translational motion, the second term is the kinetic energy due to the rotational motion, while $\mathcal{U}$ (**r**) is the potential (configurational) energy in the system. Since the system is isolated, the total energy must be conserved in time according to 1st law of thermodynamics. If the model includes internal rotation or vibration, these energies must of course be included in Equation (2.9) to have a constant of the motion.

Having the mathematical model defined and the internal forces specified, the equations can be solved numerically by some finite-difference scheme, and the system trajectory be calculated for discrete times. A route to solution is described in Chapter 3, while model potentials are described in Chapter 4.

2.2 Information from MD simulations

For N molecules with three translational and three rotational degrees of freedom, $6N$ equations must be solved simultaneously. For each time t we get positions and orientations along with linear and angular velocities for each molecule. This information enables us to calculate quantities from the following categories, which does not constitute an exhaustive list:

- Simple thermodynamic properties, as temperature which is calculated directly from the velocities via the equipartition principle. Another example is pressure where the correction to ideal gas is calculated from the molecular virial, which is position and orientation dependent. The property values are obtained as time averages.
- Thermodynamic response functions, quantities defined through derivatives of simple thermodynamic properties, are obtainable from the fluctuations. Examples of such properties are isometric heat capacity which is calculated from RMS fluctuations in total energy, or isothermal compressibility which is calcu-

lated from fluctuation in volume. They can also be calculated by repeated simulations at systematically varied conditions.

- Entropic properties, Helmholtz or Gibbs free energies and entropy, are not defined as time-averages over a phase trajectory. They can be obtained as relative properties from several repeated simulations at different conditions (i.e. different total energy, temperature or chemical potential)
- Local density expressed as pair-correlation functions of positions and orientations.
- Time-correlation functions of thermodynamic quantities give information of transport properties. Examples are viscosity from pressure and self-diffusion from velocities.
- Time-space correlation functions where for example the time-development of local structure can be followed.

In Chapter 3 we give a closer description of the quantities we calculate. More extensive descriptions are found in [15, 28].

2.3 The development of thermostatted simulations

Assignment of initial velocities and positions fix the internal energy, which is kept constant throughout the simulation. With constant energy simulations, redistribution of the initial energy during the equilibration phase will normally produce a system temperature beyond our control, unless we choose very careful the initial configuration. The system temperature will in fact not be known until the simulation is finished. Choosing entropy S and volume V as independent variables for the total energy, the total derivative of $E(S,V)$ is

$$dE_{rev} = \left(\frac{\partial E}{\partial S}\right)_V dS + \left(\frac{\partial E}{\partial V}\right)_S dV \tag{2.10}$$

where the first gradient is identified with the temperature

$$T \equiv \left(\frac{\partial E}{\partial S}\right)_V \tag{2.11}$$

Even if E is constant, the gradient will vary when the entropy S in the system vary and the relative distribution of contributions vary. The gradient is unknown (not defined) at start.

This is of course a drawback of the method if one wishes to compare simulation results with experiments or theory. Experiments are normally conducted at fixed temperatures and pressures, and not at fixed total energy as is the case in a NVE-simulation. The need for a constant temperature is also obvious in the study of solutions where both solvent and solute should be studied at the same temperature (and pressure). This extends also to simulations of phase transitions, where both phases must be kept at the same temperature. Lately *ab initio* calculations [23] have also introduced the need for a method to keep the electrons in their ground states. When performing non-equilibrium molecular dynamics it can also be desirable to keep the temperature at a constant value or to maintain a temperature gradient.

Also several derived properties are most easily determined from isothermal simulations. Examples are entropic properties as determined through thermodynamic integration or the isothermal compressibility.

Thus a number of reasons for opening up the systems to thermal interactions with the surroundings exists. It is however appropriate to mention here that if only thermodynamical and structural quantities are of interest, the Monte Carlo method at constant temperature offers a simpler alternative to molecular dynamics simulation.

2.3.1 Velocity scaling

As a first step towards temperature control one can simply rescale the velocities [38]. All molecular velocities are multiplied with the same factor to yield a kinetic energy consistent with the required temperature. The scaling factor α is evaluated from the ratio of desired average velocities to actual average velocity.

$$\alpha = \sqrt{\frac{3Nk_BT}{\sum_i \mathbf{p}_i^2 / m_i}} \tag{2.12}$$

N is number of molecules, T is desired temperature, $\mathbf{p}_i$ is actual momentum of molecule i, and m_i is its mass. k_B is the Boltzmann constant. Energy is now added or removed until equilibrium states with the desired temperature are produced. The 'microscopic' effect will be an acceleration or retardation of the molecules. The objection against this approach is that the simulation is no longer that of an isolated system because of this addition/removal of energy. But the surroundings are not taken formally account of since this energy is transferred to/from nowhere. Consequently the Hamiltonian is not a constant of the motion but will have dis-

continuous points whenever the velocities are adjusted. The connection (see Equation 2.9) between the Hamiltonian and the internal energy of the system is no longer valid. When rescaling is used only during equilibration, that is, when no averages are calculated or no structural sampling is done, this is of no importance. It is however crucial to the (formal) validity of simulation results if velocity rescaling is used during the production period to maintain the temperature at a constant value, or when applied in simulations of systems away from equilibrium. Because of its simplicity the method has been extensively used, with the rescaling taking place at various intervals ranging from each timestep to as seldom as possible [24].

2.3.2 Stochastic method

In 1980 Andersen [39] proposed a deterministic method to maintain constant pressure in the NPH-ensemble (ie. constant molecule number N, pressure P and enthalpy H) by allowing the volume to fluctuate. This is done by letting the system interact with an external system. This interaction can be thought of as an exchange of energy in the form of work exerted by an imagined "piston", able to move in all directions. The coordinates and momenta of the system is controlled by the deviation of the internal pressure from the external value, which enters the equations of motion as a feedback to all the particles simultaneously. Andersen also suggested the possibility of adding one or more degrees of freedom to the system to introduce energy fluctuations, but was not able to find such a method for the NVT-ensemble. Instead, he proposed a stochastic temperature control where the particles were allowed to collide with a source or sink of energy in a random fashion, thereby altering the kinetic energy of the individual particle. After each stochastic collision, the particle is given a new velocity chosen at random from a Maxwell distribution at a predefined temperature. The total energy of the system is then not conserved. Andersens work is mainly important to the work presented here, since he was the first to introduce the idea of an extended system.

2.3.3 Gaussian dynamics

Within the area of non-equilibrium molecular dynamics (NEMD) there was also a growing demand for a way to control temperature based upon theoretical principles. Evans, 1983 [40] and Hoover *et al.*, 1982 [41] independently introduced a damping force into the equations of motion. The underlying idea is Gauss' principle of least constraint [42, 43], which states that a system subject to constraints in either coordinates or velocities will follow trajectories which deviates as little as possible from the Newtonian trajectories. In the case of constraints involving ve-

locities (nonholonomic constraints), the constraints act as driving forces that can perform work on the system. Gauss principle is one way of assuring that any force that would move the system off the constant kinetic energy hypersurface in phase space is compensated by a constraint force that project the system back onto the hypersurface again. For the constant temperature dynamics the constraint is that the system kinetic energy should equal a predefined energy

$$\sum_i \frac{\mathbf{p}_i^2}{2m_i} - \frac{f}{2} k_B T = 0 \tag{2.13}$$

where f is the degrees of freedom of the system. The time derivative of this constraint together with Gauss' principle give the equations of motion for isokinetic molecular dynamics [44]

$$\begin{aligned} \frac{d\mathbf{q}_i}{dt} &= \frac{\mathbf{p}_i}{m_i} \\ \frac{d\mathbf{p}_i}{dt} &= -\frac{\partial \mathcal{U}_i}{\partial \mathbf{q}_i} - \zeta \mathbf{p}_i \\ \zeta &= -\left(\sum_i \frac{\mathbf{p}_i}{m_i} \cdot \frac{\partial \mathcal{U}}{\partial \mathbf{q}_i}\right) / \left(\sum_i \frac{\mathbf{p}_i^2}{m_i}\right) = -\frac{d\mathcal{U}}{dt} / f k_B T \end{aligned} \tag{2.14}$$

This constraint force is proportional to the velocities, and to a friction coefficient, ζ, in the form of a Lagrange multiplier. The constraint force is therefore often interpreted as a friction force, but unlike the macroscopic apprehension of friction, the Gaussian friction can take on both positive and negative values. It can be shown [24, 44] that the isokinetic thermostat is canonical in the coordinates provided $f \rightarrow f - 1$, but not in the momenta [17]. Woodcock's method of velocity scaling has been established as an approximate solution to Equation (2.14), provided that the scaling takes place each timestep [24]. The total energy is, however, not conserved by this method either. The method finds its major application with nonequilibrium problems.

2.3.4 Extended system dynamics

In the early eighties, Nosé, 1984 [1] also was working with a method for equilibrium molecular dynamics simulation in the canonical ensemble. Probably inspired by Andersens [39] work, he looked for a method to open up the system to thermal interactions with the surroundings. He added an extra degree of freedom, a dimensional parameter s, to the system under study. Associated with s is a 'potential energy' defined to produce property averages equal to those of the canonical ensemble. The potential energy, $E_{p,res} \sim \ln s$, represents the energy of an external heat

reservoir, and s is its coordinate. s is coupled to the molecular momenta, and the system now interacts with its surroundings in a formal manner.

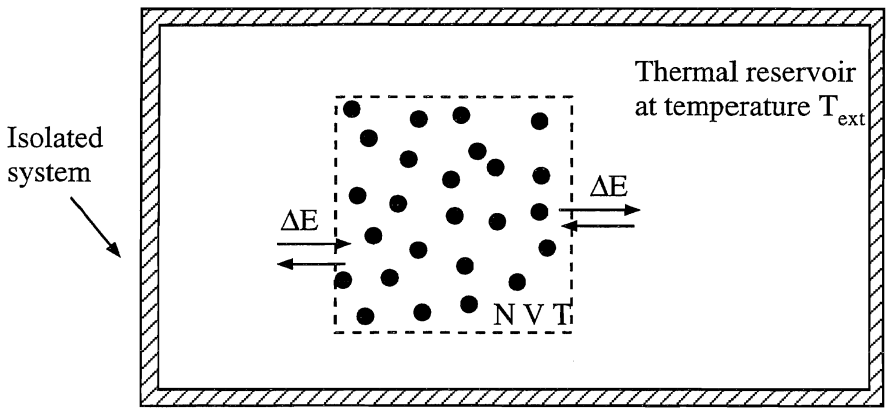


Figure 2.2 Closed system at fixed number of particles N, volume V, and temperature T in contact with a thermal reservoir at temperature T_{ext}. If system temperature T deviates from reservoir temperature by an amount ΔT, energy is transferred between reservoir and system. The reservoir is not physically placed in contact with system, but its effect is as if it was. Note also that the reservoir has no effect on volume or on the number of particles. The system and the thermal reservoir together, the extended system, does not communicate with the rest of the Universe.

The fundamental idea is that the N molecules of the system can exchange energy by heat with the thermal reservoirs through modifications of the momenta. The temperature in the system is then kept at a specified value. The interaction apply to all molecules simultaneously, and at the same rate. This rate of exchange of energy between system and reservoir is proportional to the difference between actual temperature and the preset temperature of the reservoir, and will therefore be a time dependent function with both positive and negative values. Because of thermal fluctuations, the total energy of the system is then allowed to fluctuate. Instead, the energy of the extended system has become a constant of the motion reflecting that the extended system is now an isolated system, see Figure 2.2.

In equilibrium simulations it is of great value to correlate the results with statistical mechanics. Since a system that is able to exchange energy with its surroundings at constant N, V, and T, is represented by the canonical ensemble, we hope

that the states generated during simulation are consistent with the canonical distribution. The extended system will then find its analogy in the microcanonical ensemble.

It has been shown [24] that provided the system is ergodic, the Nosé mechanics will generate trajectories which are canonical both in coordinate and momentum. Much effort has been put into the investigation of the ergodic nature of systems simulated with Nosé's equations of motion, and this important topic will be treated in Section 2.5.2.

The original Nosé equations can be cast in two different formulations. Starting from the Hamiltonian of the extended system Nosé first arrived at the equations of motion in virtual variables, Equation (2.16). But these equations are cumbersome when it comes to calculating dynamical quantities. s is a fluctuating parameter, and equal time intervals in virtual time will in general give unequal time intervals in real time. Dynamical variables will appear as functions of virtual time, and will thus not have any direct physical interpretation. Averages of thermodynamical properties and structure will however be unaffected. By interpreting s as a time scaling variable, Nosé [1] transformed the equations of motion to real variables, Equation (2.18). These transformations are however not canonical, and the equations can therefore not be obtained directly from a Hamiltonian. The equations in real variables are still dependent on both s and its conjugated momentum p_s. Hoover [2] reformulated the equations of motions by noting a redundancy of the parameters. He thus obtained a set of equations free of the time scaling variable s, which immediately became known as the Nosé-Hoover equations, Equation (2.24). A constant of the motion is still present. Hoover's formulation also has the advantage of permitting interaction with multiple thermostats, since the scaled time, which must be different for different thermostats, not appear directly. This is of importance to us, as we wish to do simulations on binary mixtures with two kinds of motion; translation and rotation. In the following section, we will take a closer look upon the mathematical formulation.

2.4 Mathematical formulation of the Nosé-Hoover thermostats

2.4.1 Nosé equations in virtual and real variables

We look at a system consisting of one or several components with a total of N rigid molecules, having three translational and three rotational degrees of freedom. The system is in contact with one thermal reservoir

Other degrees of freedom might easily be included, or the system might consist of atoms or be confined to only two dimensions.

Following Nosé [17], but including the rotation explicitly, a Hamiltonian in virtual variables (starred) for this extended system is postulated to be

$$\mathcal{H}(\mathbf{q}^*, \mathbf{p}^*, \mathbf{q}_\omega^*, \mathbf{p}_\omega^*, s^*, p_s^*) = \sum_{i=1}^{3N} \frac{(p_i^*)^2}{2m_i(s^*)^2} + \sum_{i=1}^{3N} \frac{p_{\omega i}^* I_i^{-1} p_{\omega i}^*}{2(s^*)^2} + \mathcal{U}(\mathbf{q}^*, \mathbf{q}_\omega^*) + \frac{(p_s^*)^2}{2Q} + gk_BT \ln s^* \quad (2.15)$$

where the coordinates $\mathbf{q}$ and $\mathbf{q}_\omega$ represent centre of mass position and molecular orientation in space respectively. $\mathbf{p}$ and $\mathbf{p}_\omega$ are their conjugated momenta. m_i is mass of molecule i, and I_i is a component of the inertia tensor of the molecule. $g=f+1$, where f is the degrees of freedom in the system ($6N$ for the system defined above). If other quantities besides total energy is conserved, g is reduced. The heat reservoir at temperature T, is represented by one degree of freedom, the dimensionless parameter s^*. The phase space is then extended by 2 dimensions to include the heat bath coordinate s^* and its conjugated momentum p_s^*. The appearance of s^* in the denominator of the molecular kinetic energy represents the interaction between the heat bath and the molecular motion. $\mathcal{U}$ is the potential energy of the system, depending upon the centre of mass positions and orientations. Associated with each s^* is a potential energy as given by the last term, and also a kinetic energy given by the fourth term in Equation (2.15). The constant Q in the kinetic energy is interpreted as the inertia or mass of the heat bath. In the limit of infinite mass, the interaction between system and surroundings is infinitely weak, and the heat bath is decoupled from the molecular momenta.

A reasonable value for s^* will be unity. The particular logarithmic form of the potential energy of the heat bath was shown by Nosé [17] to be necessary to generate a canonical distribution of states. Later Jellinek, 1988 [45] and Jellinek and Berry, 1988 [46] have proved that there exist an infinity of choices for the dynamics. Since this potential energy has no unique reference state, the numerical value of $\mathcal{H}$ has no physical meaning. The time average of the three first terms of (2.15) is however the internal energy of the system.

By the Hamiltonian formalism, the equations of motion in virtual variables become [17]:

$$\begin{aligned}
\frac{dq_i^*}{dt^*} &= \frac{\partial \mathcal{H}}{\partial p_i^*} = \frac{p_i^*}{m_i (s^*)^2} \\
\frac{dq^*{}_{\omega i}}{dt^*} &= \frac{\partial \mathcal{H}}{\partial p^*{}_{\omega i}} = \frac{p^*{}_{\omega i}}{I_i (s^*)^2} \\
\frac{dp_i^*}{dt^*} &= -\frac{\partial \mathcal{H}}{\partial q_i^*} = -\frac{\partial \mathcal{U}}{\partial q_i^*} \\
\frac{dp^*{}_{\omega i}}{dt^*} &= -\frac{\partial \mathcal{H}}{\partial q^*{}_{\omega i}} = -\frac{\partial \mathcal{U}}{\partial q^*{}_{\omega i}} \\
\frac{ds^*}{dt^*} &= \frac{\partial \mathcal{H}}{\partial p_s^*} = \frac{p_s^*}{Q} \\
\frac{dp_s^*}{dt^*} &= -\frac{\partial \mathcal{H}}{\partial s^*} = \left(\sum_{i=1}^{3N} \frac{(p_i^*)^2}{m_i (s^*)^3} + \sum_{i=1}^{3N} \frac{p_{\omega i}^* I_i^{-1} p_{\omega i}^*}{(s^*)^2} \right) - \frac{g k_B T}{s^*}
\end{aligned} \tag{2.16}$$

Because of the scaled time in the virtual variable formulation, these equations are impractical for use in simulation when dynamical variables are calculated, since these must be rescaled [47]. To obtain equations in real time with real momenta, the transformations [17]

$$\begin{aligned}
t &= \int dt^* / s^* \\
p_i &= p_i^* / s^* , \; p_{\omega i} = p^*{}_{\omega i} / s^* , \; p_s = p_s^* / s^* \\
q_i &= q_i^* , \; q_{\omega i} = q_{\omega i}^* , \; s = s^*
\end{aligned} \tag{2.17}$$

have to be performed. The real time equations of motions thus reads

$$
\begin{aligned}
\frac{dq_i}{dt} &= \frac{p_i}{m_i} \\
\frac{dq_{\omega i}}{dt} &= \frac{p_{\omega i}}{I_i} \\
\frac{dp_i}{dt} &= -\frac{\partial \mathcal{U}}{\partial q_i} - \frac{p_i}{s}\frac{ds}{dt} \\
\frac{dp_{\omega i}}{dt} &= -\frac{\partial \mathcal{U}}{\partial q_{\omega i}} - \frac{p_{\omega i}}{s}\frac{ds}{dt} \\
\frac{ds}{dt} &= \frac{p_s s^2}{Q} \\
\frac{dp_s}{dt} &= \frac{1}{s}\left(\sum_{i=1}^{3N}\frac{p_i^2}{m_i} + \sum_{i=1}^{3N}\frac{p_{\omega i}^2}{I_i} - gk_BT\right) - \frac{p_s}{s}\frac{ds}{dt}
\end{aligned}
\tag{2.18}
$$

The presence of the extra force term in the third (and fourth) of Equations (2.18) marks the difference between the Nosé mechanics and the Newtonian mechanics. As the equations of motion are not invariant during the transformation, these real variable equations are not canonical. Consequently there is no Hamiltonian corresponding to the real variable equations. Changing variables in the virtual variable Hamiltonian yield however an energy

$$
\mathcal{E}(\mathbf{q}, \mathbf{p}, \mathbf{q}_\omega, \mathbf{p}_\omega, s, p_s) = \sum_{i=1}^{3N}\frac{p_i^2}{2m_i} + \sum_{i=1}^{3N}\frac{p_{\omega i}^2}{2I_i} + \mathcal{U}(\mathbf{q}) + \frac{s^2 p_s^2}{2Q} + gk_BT\ln s \tag{2.19}
$$

which is a constant of the motion. Now g is reduced by one relative to the virtual variable formulation. The actual value of g is necessary for the dynamics in both virtual and real variables to generate canonical distribution of states [17]. Note that in applying the Hamiltonian formalism to Equation (2.19) the Equations (2.18) is not reproduced, but a set of equations for which the coupling between system and reservoir is vanishing.

Jellinek and Berry, 1989 [48] has clarified some delicate points of the Nosé-formulation. Despite the similarity of the scaling of phase space and time in Equation (2.17), they are completely independent. Thus s is not to be regarded as neither a time-scaling nor a mass-scaling parameter, but '..., the only role of s is as a source or sink of energy for the physical system; its coupling to the physical system regulates all the flow of energy.'

The dynamics generated by (2.16) and (2.18) are said [2] to traverse the same trajectories but at different rates. Then the time-averages calculated over the trajectories are not equal. To have equivalent time-averages, ie. to generated the same ensemble, one must have different Hamiltonians and hence different trajectories. This is reflected in the different values of g in the virtual and real variable formulations [48].

2.4.2 The Nosé-Hoover equations for one thermostat

Still the system dynamics described by Equations (2.18) is dependent on both the heat bath variable and its momentum. Hoover [2] noticed that one of these were redundant. By differentiating the fifth of Equations (2.18) with respect to t, one get

$$\frac{d^2s}{dt^2} = \frac{1}{Q}\left\{s^2\frac{dp_s}{dt} + 2p_s s\frac{ds}{dt}\right\} \tag{2.20}$$

where $p_s = (Q/s^2)(ds/dt)$. Rearranged and inserted into the last of equations (2.18), this yield

$$Q\left\{\frac{1}{s^2}\frac{d^2s}{dt^2} - \frac{1}{s^3}\left(\frac{ds}{dt}\right)^2\right\} = \frac{1}{s}\left(\sum_{i=1}^{3N}\frac{p_i^2}{m_i} + \sum_{i=1}^{3N}\frac{p_{\omega i}^2}{I_i} - gk_BT\right) \tag{2.21}$$

Then identifying a new variable

$$\eta = \frac{1}{s}\frac{ds}{dt} \tag{2.22}$$

and its time derivative

$$\frac{d\eta}{dt} = \frac{1}{s}\frac{d^2s}{dt^2} - \frac{1}{s^2}\left(\frac{ds}{dt}\right)^2 \tag{2.23}$$

enables the replacement of the two variables s and p_s by one variable η. The equations to be solved for the extended system are then the Nosé-Hoover equations [24], with the rotational motion explicitly included

$$
\begin{aligned}
\frac{dq_i}{dt} &= \frac{p_i}{m_i} \\
\frac{dq_{\omega i}}{dt} &= \frac{p_{\omega i}}{I_i} \\
\frac{dp_i}{dt} &= -\frac{\partial \mathcal{U}}{\partial q_i} - p_i\eta \\
\frac{dp_{\omega i}}{dt} &= -\frac{\partial \mathcal{U}}{\partial q_{\omega i}} - p_{\omega i}\eta \\
\frac{d\eta}{dt} &= \frac{1}{Q}\left(\sum_{i=1}^{3N}\frac{p_i}{m_i} + \sum_{i=1}^{3N}\frac{p_{\omega i}^2}{I_i} - gk_BT\right)
\end{aligned}
\tag{2.24}
$$

We note the similarity with the Gaussian method described in Equations (2.14). $p_i\eta$ act as an external force to the system, just as $p_i\zeta$ does, but η has a thermodynamically defined origin. Because of this similarity, η is sometimes interpreted as a friction coefficient. Just as ζ, η is time dependent and can take on both positive and negative values. But the relation between η and molecular velocities is through an integral, making this in the language of control theory an integral control. The particular form of η not arising from a constraint, imply that we must expect fluctuations in system kinetic energy - and in temperature - contrary to the constant kinetic energy in the Gaussian (isokinetic) method.

For a system in thermal equilibrium η represents the transfer of energy back and forth between system and surroundings, and is thus expected to be fluctuating around zero. The kinetic energy of the heat bath is the kinetic energy of one degree of freedom, and should by the equipartition principle be of the order of $\frac{1}{2}k_BT$.

The extended system energy corresponding to Equations (2.24) is [24]

$$
\mathcal{E}(\mathbf{q},\mathbf{p},\mathbf{q}_\omega,\mathbf{p}_\omega,\eta) = \sum_{i=1}^{3N}\frac{p_i^2}{2m_i} + \sum_{i=1}^{3N}\frac{p_{\omega i}^2}{2I_i} + \mathcal{U}(\mathbf{q}) + \frac{1}{2}Q\eta^2 + gk_BT\int_0^t \eta(t')dt' \tag{2.25}
$$

and is still a constant of the motion since all contributions to energy are included. If the energy is the only constant of the motion, $g=f=6N$ for translation and rotation of N rigid molecules, as with the real variable Nosé formulation.

2.4.3 The Nosé-Hoover equations for multiple control

Our particular interest is in binary mixtures of translating and rotating molecules, so what we want is a model that can handle the different kinds of motion individually. As pointed out by Nosé [1] and later investigated by others [47, 49, 50], that to produce canonical trajectories, the thermostat mass Q must have values appropriate for the actual motion and molecule to be controlled. The value assigned to Q should not disturb the natural progress of the different degrees of freedom with time. For the moment, let us be content with the statement [18] that it is not guaranteed that one and the same thermostat can control different motions satisfactorily since their characteristic timescales might be different. We then want to have the possibility of applying more than one thermostat to the system.

One could naively try to extend Equation (2.15) by introducing one extra (or more) parameters s_j and define interactions between the heat baths and the respective motion in the same manner as with one heat bath only. Then every s in the following equations are replaced by an s_j, and an additional summation over j is performed. It would look nice up to the point where the equations are transformed to real time variables. Since the thermostats oscillate at different frequencies, they scale time differently. Equation (2.17) then inevitably introduces multiple timescales, and we will have a set of equations where rotation and translation evolves according to different times. Either the Hamiltonian or the transformations or both are not applicable. So we are forced to look at equations (2.24) and postulate by analogy [18] that there exists a certain arrangement that our system evolve according to

$$\begin{aligned}
\frac{dq_{ij}}{dt} &= \frac{p_{ij}}{M_{ij}} \\
\frac{dp_{ij}}{dt} &= -\frac{\partial \mathcal{U}}{\partial q_{ij}} - p_{ij}\eta_j \\
\frac{d\eta_j}{dt} &= \frac{1}{Q_j}\left(\sum_{i=1}^{g_j}\frac{p_{ij}}{M_{ij}} - g_j k_B T_j\right)
\end{aligned} \tag{2.26}$$

where q_i and p_i now are the coordinates of each degree of freedom of a molecule i including all kinds of motion. M_{ij} is an inertia of the molecule i corresponding to one particular motion or kind of molecule affected by heat reservoir j, while Q_j is

the inertia of the jth heat reservoir. g_j is the number of degrees of freedom within the system affected by heat bath j. A constant of the motion still exists

$$\mathcal{E}(\mathbf{q},\mathbf{p},\mathbf{q}_\omega,\mathbf{p}_\omega,\eta_1,\ldots,\eta_J) = \sum_{j=1}^{J}\sum_{i=1}^{g_j}\frac{p_{ij}^2}{2m_{ij}} + \mathcal{U}(\mathbf{q},\mathbf{q}_\omega) + \sum_{j=1}^{J}\frac{1}{2}Q_j\eta_j^2 + \sum_{j=1}^{J}g_jk_BT_j\int_0^t \eta_j(t')dt' \quad (2.27)$$

The proof that this dynamics generates canonical distributions provided the system is ergodic, is included in Appendix A.

Not only can the different thermostats control specific categories of motion, such as translation, rotation or vibration, but also the motion in one specific coordinate direction can be thermostatted. It is also possible, at the cost of linear momentum conservation [24], to control different species separately. In the case of non-equilibrium simulations, the heat baths can be used to maintain a thermal gradient or to absorb dissipative heat created in fluid flow.

As a particular example, for a system in contact with two thermal reservoirs one thermostatting the translational motion, labelled *1*, the other the rotational motion, labelled *2*, the Nosé-Hoover equations of motion reads

$$\begin{aligned}
\frac{dq_{i1}}{dt} &= \frac{p_{i1}}{m_{i1}}\\
\frac{dq_{\omega,i2}}{dt} &= \frac{p_{\omega,i2}}{I_{i2}}\\
\frac{dp_{i1}}{dt} &= -\frac{\partial \mathcal{U}}{\partial q_{i1}} - p_{i1}\eta_1\\
\frac{dp_{\omega,i2}}{dt} &= -\frac{\partial \mathcal{U}}{\partial q_{\omega,i2}} - p_{\omega,i2}\eta_2\\
\frac{d\eta_1}{dt} &= \frac{1}{Q_1}\left(\sum_{i=1}^{g_1}\frac{p_{i1}}{m_{i1}} - g_1k_BT_1\right)\\
\frac{d\eta_2}{dt} &= \frac{1}{Q_2}\left(\sum_{i=1}^{g_2}\frac{p_{\omega i2}^2}{I_{i2}} - g_2k_BT_2\right)
\end{aligned} \quad (2.28)$$

q_{i1} and p_{i1} are the coordinates and conjugated momenta of the translational degrees of freedom of molecule i that are controlled by thermostat *1*, and $q_{\omega,i2}$ and $p_{\omega,i2}$ are orientations and momenta of molecule i, controlled by thermostat 2. g_j is the total number of degrees of freedom affected by thermostat j, which for this spe-

cific example of N translating and rotating molecules is $3N$ for each thermostat. Each reservoir can in principle have different temperatures T_j, but for a system in thermal equilibrium, they must all be equal. The motion of heat bath j is determined from the difference between the instantaneous values of the kinetic energy associated by the involved degrees of freedom and the preset value. As index 1 represents translation, m_i is molecular mass and I_i is molecular inertia. But because of the coupling in the forces in the equations of motion, the direct control of translation will indirectly influence the angular momenta $p_{\omega,i2}$ and the molecular orientations $q_{\omega,i2}$, and vice versa.

These mathematical thermostats are selective as distinct from commonplace heat baths. Firstly, they can identify one specific kind of motion to act upon, and secondly, they can act on distance thus having a closer resemblance to an external force field, and finally, they act upon all involved molecules simultaneously.

The total energy of this particular extended system is given by

$$\mathcal{E}(\mathbf{q},\mathbf{p},\mathbf{q}_\omega,\mathbf{p}_\omega,\eta_1,\eta_2) = \sum_{i=1}^{g_1}\frac{p_{i1}^2}{2m_i} + \sum_{i=1}^{g_2}\frac{p_{\omega,i2}^2}{I_i} + U(\mathbf{q},\mathbf{q}_\omega) + \sum_{j=1}^{2}\frac{1}{2}Q_j\eta_j^2 + \sum_{j=1}^{2}g_jk_BT_j\int_0^t\eta_j(t')dt' \tag{2.29}$$

Taking the time derivative of $\mathcal{E}$ and substituting the equations of motion, Equations (2.26), show that $\mathcal{E}$ is a conserved quantity.

2.5 Ergodic motion and the choice of Q

2.5.1 Ergodic motion

By **ergodicity** we mean that the phase space trajectory passes arbitrarily *close* to all points in available phase space over a sufficiently long time. This is a relaxation of the original definition due to Boltzmann where the trajectory should pass *through* all configurations [28], and is thus more correctly referred to as **quasi-ergodic** motion. We will nevertheless use the term ergodic in place of quasi-ergodic. Then the ergodic hypothesis states that if a system is ergodic the ensemble average of a quantity A equals the time average over the trajectory for the same quantity

$$\langle A(q,p)\rangle = \bar{A}(q,p) \tag{2.30}$$

Ergodic motion will be chaotic, but the converse is not true in general, as trajectories can be chaotic but **recurrent**. A recurrent trajectory will cross itself at some point and then the sequence of configurations will repeat. The trajectory will be confined to only a region of **available phase space** [51]. Available phase space for a microcanonical system at a given state is the hypersurface of points for which the Hamiltonian is constant.

An ergodic system is expected to be independent of initial equilibrium conditions, as ergodicity implies that wherever the system 'enters' a trajectory, all of phase space will be filled. The particular initial conditions then will only be one of the points on the trajectory, which also could have been passed through given other initial conditions. Ergodicity does however not imply that a system will evolve from nonequilibrium to equilibrium. When the initial conditions not corresponds to an equilibrium state, sufficient time must of course be provided to let the system relax to equilibrium, which is quite a different matter.

2.5.2 Ergodicity of Nosé-Hoover equations?

The necessary and sufficient conditions for the Nosé or Nosé-Hoover dynamics to generate canonical distributions, are that the number of degrees of freedom are set equal to[1] $g=3N+1$ (Nosé virtual), $g=3N$ (Nosé real and Nosé-Hoover) to satisfy the Liouville theorem [17], and that the trajectory is ergodic [48]. We emphasize the difference between the dynamics and its implementation in a molecular dynamics simulation with periodic boundaries. Under periodic boundaries condition, total linear momentum is conserved, and g must be reduced by 3. Cho *et al.* [52] have shown that an additional criterion for the generated trajectories to be ergodic, is that total linear momentum must be zero at start, see Section 2.6.2.

The Nosé and Nosé-Hoover thermostats have for a long time been known not to generate ergodic motion for all systems. The one-dimensional harmonic oscillator (ODHO) is a well-known example, and has been thoroughly studied by several authors (see below). The ODHO is well suited for studies like this because it has analytical solution and because it has physical significance as an approximation to several systems. Despite being integrable and periodic, the dynamics of an isolated ODHO is still ergodic [28]. Then it is also expected that the thermostatted ODHO should be ergodic.

The major outcome of these studies has been several extensions, of which some will be mentioned briefly below, to the Nosé or Nosé-Hoover formalisms which

[1] For rotation of N nonlinear molecules, 3N extra degrees of freedom must be added

seem to make the dynamic 'chaotic enough' to allow for a statistical mechanical description. In addition to the new computational schemes also deeper understanding of the Nosé thermostat has been furnished.

Kusnezov *et al.*, 1990 [53, 54] noted that the coordinates are too correlated with the momenta in the Nosé dynamics. They used the Nosé-Hoover equations as a starting point and introduced an additional friction coefficient (but no extra heat bath) to modify the velocities besides the modification of accelerations. With a nonlinear coupling of the friction coefficients in the equations of motion, they managed to get ergodic dynamics for several simple potentials including the ODHO-potential. They also investigated temperature dependence of the Nosé-Hoover scheme, and found clearly non-ergodic behaviour at low temperatures for an antisymmetric well potential, while ergodicity seemed to be present for the same system at higher temperatures.

Winkler, 1990 [55] applied the virtual variable formalism of Nosé, but made the equations more nonlinear by scaling the variables by $1/s^2$ in (2.17), and showed that this was enough to make the ODHO ergodic.

Martyna, Klein, and Tuckerman, 1992 [56] have applied a different approach to 'disturb' the extended system. They proposed the Nosé-Hoover chain method, where a chain of M thermostats is coupled to the system. The idea is that each thermostat is controlled by another thermostat so as to drive their fluctuations to be Gaussian distributed. The method seem to produce ergodic dynamics and has been used with success in protein simulations [57].

One might be tempted to say that non-ergodicity is only present for simple systems with very few degrees of freedom. A usual assumption [59] is indeed that the complexity of many-body problems, like the one we study, will produce ergodic trajectories. The picture is however not that simple. Posch *et al.*, 1986 [58] found no evidence for a two-dimensional system of two soft disks to be nonergodic with the Nosé thermostat, while Cho and Joannopoulos, 1992 [47] proved analytically that a system of hard spheres is nonergodic within the Nosé-formalism, regardless of system size. Nosé, 1993 [59] found that a system of N ODHO was ergodic if their frequencies were different, as they will be in a physical system. Recent studies of condensed LJ-systems [47, 49, 50, 60] have also shown a dependency upon Q for the generation of ergodic dynamics.

From the above brief summary of studies on ergodicity, it seems that not only the number of particles present but also the potential they are subject to can prevent ergodic behaviour. Also the actual mathematical coupling between system and reservoir can prevent ergodicity. Finally, the state of the system and parameters describing the system influence on ergodicity.

Then we conclude with the following assumptions: The Nosé-Hoover equations does not automatically guarantee ergodic dynamics. But the pair potential we use (see Sections 4.2 - 4.4) effectively redistribute energies among the molecules so that a representative sampling of phase space is achieved, and our system is large enough not to prevent ergodisity. Also the thermodynamical state has sufficiently high temperature (liquid) so that the potential is not approximated by a harmonic oscillator. The only free parameter left to vary is the heat bath mass, and there exist a range of values of Q within which the Nosé-Hoover equations can produce ergodic dynamics. Q has been subject to a handful of thorough studies [47, 49, 50] where the goal has been to study the influence of its numerical value. The conclusions have been somewhat diverging. The effects of and limitations upon Q will be the subject of the next section.

2.5.3 Thermostat mass Q

The parameter Q is usually interpreted as the thermostat mass, and controls the speed of the thermostat response. A massive thermostat is to be regarded as a weak regulator, and vice versa. In the limit of infinite mass Q, the system will decouple from the reservoir, and the motion will be that of a microcanonical ensemble with the system and the surroundings evolving separately [2]. If Q on the other hand is small, the temperature control will be intense, and the kinetic enery fluctuations will, in the limit of zero mass, be completely suppressed as with the Gaussian isokinetic method [24]. For increasing Q, the fluctuations in kinetic energy, or temperature, will also increase in amplitude, but the average is constant and equal to the reservoir temperature. This is a logical consequence of slow thermostatting; the system has time to deviate more from the required equilibrium temperature. In the thermodynamic limit, all fluctuations will be immeasurable small independent of Q.

Nosé [24] has obtained some useful expressions for Q in the small and large mass limits. For small values of Q,

$$Q \simeq 2gk_BT\left(\frac{\tau}{2\pi}\right)^2 \tag{2.31}$$

where g is number of degrees of freedom, k_B is Boltzmanns constant, T is the heat bath temperature and τ is the period of oscillation of the s-parameter. For large Q,

$$Q \simeq \frac{2g^2k_B^2T}{C_v}\left(\frac{\tau}{2\pi}\right)^2 \tag{2.32}$$

where C_v is the heat capacity of the system.

Nosé's original recommendation [1] for the choice of Q, was to use a value consistent with a timescale of the system in order to have ergodic motion. The argument was that if Q is too small, the decoupling of reservoir and system would give incomplete sampling of phase space, and if Q is too large the phase space will be ineffectively sampled. This has been confirmed by several authors [47, 49, 50], but doubt still exists to whether it is possible to have ergodic motion also when Q takes on extreme values. According to Cho and Joannopoulos,1992 [47] this is possible if we allow for very long simulations. They analysed a liquid Lennard-Jones system with the virtual variable formulation, Equation (2.16), and found that in case of Q small, the timestep had to be very small to integrate the equations of motion properly. Consequently the total simulation time had to be very long to get a proper sample of the trajectory. In case of large Q, when the thermostat variable changes slowly, a long simulation must be performed to include enough heat bath oscillations. In either case they interpreted their results as consistent with the canonical ensemble, but recommended nevertheless a thermostat period close to a characteristic frequency of the system to obtain a fast convergence to the canonical distributions.

At the same time Bylander and Kleinman [50] arrived at the opposite conclusion regarding small Q. They simulated liquid Na and found from calculations of specific heat that the fluctuations in energy were noncanonical. They interpreted their results as it was unlikely that small Q would ever give the canonical distribution.

Also Di Tolla and Ronchetti, 1993 [49] found that for large Q, the convergence to canonical distribution would eventually occur, but for small Q increasing the length of a simulation would not improve the results.

The solution to the divergence seem to be that even for small Q the system can be in equilibrium in the sense that averages are constant, but the extended system is not in equilibrium [50].

We have perhaps not seen the last contribution to this debate, but by now it is fairly well established that 1) even for large and complex systems the ergodicity for a Nosé-Hoover simulation is not guaranteed, so 2) to have fast convergence to canonical distributions, Q has to be within a specific range of values but 3) even if the distributions are highly noncanonical, the static averages can be reasonable.

All studies agree however in the recommendation of applying a value close to a characteristic frequency of the system (or of the degree of freedom one wishes to control). Also the very recent study of Holian *et al.* [60] arrive at this conclusion.

This is sensible with respect to the analysis of harmonic motion where the rate of energy transfer from an applied force to a forced oscillator is a maximum at resonance, and resonance occurs when the frequency of the applied force is equal to

the natural frequency for the undamped oscillator. This extends to a variety of physical situations where resonance is seen to occur whenever a system is subject to an external action that varies periodically with time.

If Q is too small, the thermostat frequency will show up as a slow modulation (envelope) of the temperature fluctuations, while the atomic frequencies will be present as fluctuations of the s-parameter. If, on the other hand, Q is too large, there will be a sinusoidal modulation due to the oscillations of s of the (faster) fluctuations of the temperature, while the atomic frequencies are not transmitted to the evolution of s, consistent with a microcanonical simulation. This is discussed, and also nicely illustrated in [50] and [60].

Upon closing this chapter, we want to mention the very recent suggestion by Hoover and Holian, 1996 [25], where they propose to control not only the kinetic temperature but also its fluctuations. This method is seen to exhibit ergodicity for the ODHO, and is also expected to be applicable to equilibrium many-body simulation.

2.5.4 Demonstrating ergodic motion

There is no general and conclusive *a priori* test of ergodicity [53, 58]. Calculation of Lyapunov exponents is one strategy [58]. The Lyapunov exponents are measures on how fast a perturbed trajectory will diverge from its parent trajectory, and thus a measure of chaos. It is however not clear what the exact numerical value of the exponent must be to assure ergodicity. Plots of trajectories in phase space have also been used, but as also chaotic trajectories can give space filling trajectories, this is not a proof of ergodicity [53]. Perhaps this method is better for demonstrating the converse. These methods is usually applied for systems with few degrees of freedom.

For practical use, we must resort to indicative tests. In the canonical ensemble, energies are distributed according to [13]

$$f_c(p) \propto e^{-mv^2/2k_BT}, \quad f_c(q) \propto e^{-U/k_BT}, \quad f_c(\eta) \propto e^{-Q\eta^2/2k_BT} \tag{2.33}$$

for velocity, configurational energy and reservoir kinetic energy, respectively. Calculated distributions for the simulated system can be compared to these theoretical results. Also one can compare trajectory averages to known canonical ensemble variables, as produced in a Monte Carlo simulation.

Also independence of numerical results upon initial values can be used as an indication of ergodicity [53].

It is further possible to analyse the degree of mixing by calculating cross- or auto-correlations of dynamical quantities. If such correlations vanish on a short timescale, ie. memory of previous configurations is lost quickly, the system has the mixing property, and ergodicity is guaranteed [28, 49].

Also the higher moments of kinetic energy and of reservoir kinetic energy can be calculated and compared to the analytical momenta [47]. The equations are given in Section 3.6.3, eqs. (3.39) and (3.40) page 73.

2.6 Constants of the motion

A particular choice of statistical mechanical ensemble imply that there can be some constants of the motion in addition to independent variables keeping the state fixed. The simulation method used for realizing this choice can however fail to conserve the correct quantities, or can restrict the sampling of the phase space by conserving too many quantities. Errors in the code can also destroy conservation. To be confident that our simulated results corresponds to the chosen ensemble, it is necessary to identify and examine these quantities. A conserved quantity will reduce the number of degrees of freedom of the system, so particularly for dilute mixtures where one species can be represented by a small number of particles, it is important that the correct number of degrees of freedom for the simulated system are implemented in the computer program. And, finally, an integral of the motion having a non-constant value must be taken as evidence of truncation errors or programming errors in the code.

From mechanics we know three independent kinds of integrals of the motion: the energy, the components of the linear momentum, and of the angular momentum. We will examine how these quantities behave in a NVT-simulation with one or more Nosé-Hoover thermostats. An important remark: the simulation should not conserve any additional variables since we then put restrictions on available phase space and consequently will have a different ensemble.

2.6.1 Conservation of extended system energy

In the microcanonical ensemble, the Hamiltonian is a conserved quantity. It is further identified with the time-dependent system energy. The canonical ensemble puts no restriction on system energy; it is allowed to fluctuate around an equilibrium value. But since the system together with its thermal surroundings constitute

an isolated system, the sum of their respective energies must be constant. This has the important consequence that in a molecular dynamics simulation with Nosé-Hoover thermostats, the total energy of the extended system given by Equation (2.29) must be strictly constant (within an acceptable limit of numerical accuracy).

2.6.2 Conservation of linear momentum

In the microcanonical ensemble neither linear nor angular momentum is conserved. This is not so obvious. If we introduce the microcanonical ensemble as a representation of the time development of an isolated system, it is clear that both linear and angular momentum, in addition to the energy, must be conserved. The isolated system has no knowledge of its surroundings, and can not be influenced by any external forces or torques. To be consistent with an isolated system, the microcanonical distribution function should then have been defined as [13]

$$\rho = constant \times \delta(E - E_0)\delta(\mathbf{P} - \mathbf{P}_0)\delta(\mathbf{L} - \mathbf{L}_0) \tag{2.34}$$

Where E_0, $\mathbf{P}_0$ and $\mathbf{L}_0$ are constant values of energy, linear momentum, and angular momentum. The linear and angular momenta are however excluded from the definition of the microcanonical ensemble realizing that any change in translational or rotational velocity of the whole system also will change its energy. Thus energy is the important quantity, and the microcanonical ensemble is then defined in the usual way

$$\rho = constant \times \delta(E - E_0) \tag{2.35}$$

This is justified by the fact that no isolated system can exist without being confined between walls [51]. With idealized reflecting walls, momentum conservation is destroyed, whereas energy is still conserved. Even with interacting walls, momentum will not be conserved while energy is conserved, if we take the walls to be part of the isolated system.

Since the canonical ensemble can be regarded as a subsystem of the microcanonical ensemble, the same assumption must also be valid in this ensemble. This becomes clear if we look at a microcanonical ensemble of N identical particles and one different. Let the one particle be a canonical ensemble and the N particles be the thermal reservoir. If the total momentum of this system of one particle has to remain constant, we can easily imagine situations where thermal equilibrium between the system and its reservoir never will be achieved.

The usual procedure of molecular dynamics simulation with periodic boundary conditions (see Subsection 3.3.3 and Figure 3.3, page 60) *does* however conserve linear momentum, regardless of ensemble. The system is confined between walls that reflects an incoming molecule from the opposite wall, so a translational invariance is imposed on the system. Therefore, all components of the total linear momentum will be conserved during simulations in the directions where these boundaries are applied. For an NVE simulation this means that the ensemble generated is not exactly the microcanonical, and is often called an molecular dynamics ensemble, or a NVEp-ensemble [61, 62, 63]. For an isothermal simulation with one or more Nosé-Hoover thermostats, the picture is a bit more complicated.

The time rate of change of total momentum in each coordinate direction[2] is equal to the total applied force in that direction

$$\frac{d}{dt}\sum_i p_i^\xi = F_i^\xi + F_e^\xi \tag{2.36}$$

where i enumerates molecules, F_i^ξ is internal forces due to interactions between molecules and F_e^ξ is an external force field. When there is no external force field, and if Newton's 3rd law is valid, the linear momentum becomes a constant of the motion [36]. But if the particles interact with a thermal reservoir, this can be regarded as an external force, and the conservation of momentum depends upon the specific form of the interaction between the system and the reservoir and the initial state of the system.

For the Nosé-Hoover multiple thermostat, integration of the total force give [24]

$$\frac{d}{dt}\sum_j^J\sum_i^{g_j} p_{ij}^\xi = F_{ij}^\xi - \sum_j^J \eta_j \sum_i^{g_j} p_{ij}^\xi = -\sum_j^J \eta_j \sum_i^{g_j} p_{ij}^\xi \tag{2.37}$$

since the sum of internal forces once again vanish with central forces. If $j=1$ the above equation can be integrated by separation of variables to yield

$$p^\xi = p_0^\xi e^{-\int \eta(t)dt} \tag{2.38}$$

where p_0 is total momentum at $t=0$. Then linear momentum is not conserved unless it is zero at the beginning [24, 62, 52]. But since the integral in the exponent is fluctuating around its initial value in an equilibrium simulation, p is expected to fluctuate around p_0 for a heat bath initial energy chosen equal to zero.

2 Since conservation of total momentum is conservation of both magnitude and direction, must the magnitude in each direction be conserved.

If $j>1$, the integration above can still be performed for different thermostats coupling to different coordinate directions, or for the same thermostat coupling to different kinds of molecules for the same kind of motion. In the latter case control of translational and rotational degrees of freedom can be performed separately, as translation and rotation confirm to independent conservation laws. But if one tries to thermostat translational motion for different species separately, one quickly gets into trouble with the integration of Equation (2.37), ie with two species [24]

$$\frac{d}{dt}\sum_{i}^{g_1} p_{i1}^{\xi} + \frac{d}{dt}\sum_{i}^{g_2} p_{i2}^{\xi} = -\eta_1 \sum_{i}^{g_1} p_{i1}^{\xi} - \eta_2 \sum_{i}^{g_2} p_{i2}^{\xi} \tag{2.39}$$

The multiple thermostat is now not seen to conserve total linear momentum, regardless of initial values chosen. The only solution of the integration of the forces that can give a conserved linear momentum is when each species separately has an initial state of zero momentum. This is equivalent to state that each species is a thermodynamic subsystem where thermodynamic and mechanical equilibrium hold separately, and that particles of different species does not interact. Then we obviously no longer have a mixture.

For a discussion of consequences of linear momentum conservation in the virtual variable formalism, see Cho *et al.*, 1993 [52]

It is perhaps worth a moment of afterthought, that the simulation can never generate the canonical ensemble since the conservation of linear momentum will restrict the phase space trajectory to a hypersurface of smaller dimensionality.

2.6.3 Non-conservation of total angular momentum.

Much of the discussion above also hold for the components of total angular momentum. By exactly the same arguments as used for linear momentum, it can be shown [24] that the total angular momentum with the Nosé-Hoover thermostats also is conserved only if its starting value is zero. Angular momentum can however never be conserved in a MD simulation with periodic boundary conditions, because when leaving and re-entering the box, a particle changes position and thereby also its contribution to the total angular momentum. Averaged over a long time, there will be approximately the same number of particles entering and leaving at the same edge, and the total angular momentum is expected to fluctuate around zero [28]. The molecular dynamics ensemble is therefore consistent with the corresponding statistical mechanics ensemble with respect to angular momentum.

Chapter 3
Simulation details

In this chapter we describe accurately how the molecular dynamics simulation is performed, which quantities we have calculated, and how this is done. We start with a brief outline of program structure, describe the central algorithms, and the start-up procedure for the mixture simulations. Then we move on to the calculation of forces. The remaining part of the chapter is dedicated to the computation of properties, and to how the reliability of the simulation is investigated. We have focused upon equilibrium liquid state verification, and methods to investigate the ergodic behaviour of the system.

A summary of the most important parameters is given in Table B.3.

3.1 Program outline

The main philosophy underlying the code is that it should be readable and easy to extend and maintain. Our goal has been to device a program that, firstly, could serve our own purposes as defined previously, and, secondly, could be used at our institution for solving related problems. This is reflected through the extended use of subroutines, where the criterion for creating a subroutine is the logical relationship between a set of statements. As the use is limited to our in-house computers, we have also taken full advantage of the ability to perform structured programming permitted by RISC FORTRAN.

Most of the data analysis are done within the program to avoid the large amount of output - but of course at the cost of some computer time. This also minimizes the

number of additional analysis-programs. An important exception is the output of translational velocity components needed to compute the velocity auto correlation function, and the accumulated centre of mass positions used for mean square displacement calculations. If inquired they are saved to disk at regular intervals, and analysed after program completion. Figure 3.1 show a flow chart for the program, while Figure 3.2 show an account of input requirements and output generation.

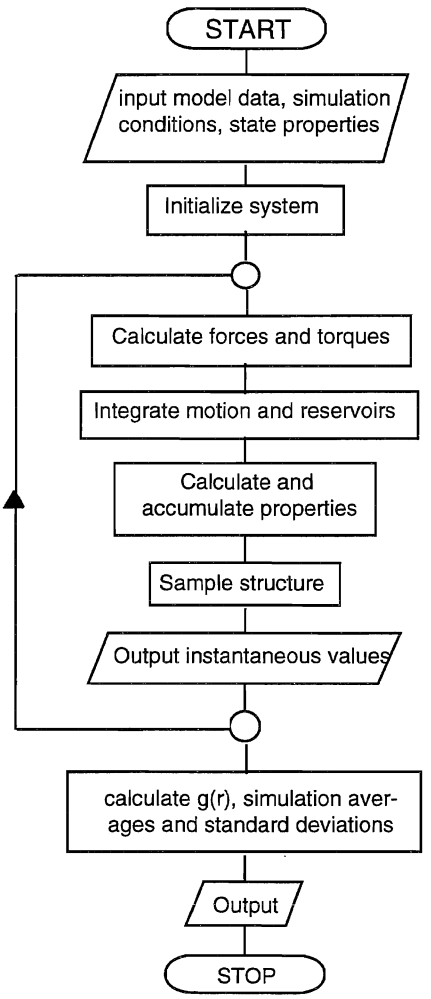


Figure 3.1 Simple flow chart of program

The program has been tested for a few different cases: pure species simulation, binary

mixture simulation, constant temperature or constant energy (with or without velocity scaling). In a NVT simulation the number of thermostats can vary from one to four by different choices of parameters.

INPUT

for each species

- moment of inertia, mass, number of sites
- potential parameters and partial charges, force-cutoff
- site-names
- local coordinates of molecular sites
- initial molecular orientations (Euler angles) and site-centre of mass lengths

for both species

- total number of molecules, and number molecules of one species
- number of unit cells in fcc-lattice
- temperature
- density
- simulation length
- timestep
- equilibration period
- output-frequency, sampling frequency, segment length
- input and output file names
- number and type of thermostats
- thermostat relaxation parameters

OUTPUT

at regular intervals

- temperatures, pressure, translational order parameter
- potential energy, kinetic energy, extended system energy, reservoir potential energy, reservoir kinetic energy, irreversible work
- energy transfer rate, heat bath parameters
- moments of kinetic energy and reservoir kinetic energy
- translational and angular total momentum
- translational velocities and uncorrected positions
- kinetic energies in coordinate directions

at the end

- radial site-site correlations
- linear velocity distributions
- averages and standard deviations of properties

SUBSEQUENT ANALYSIS

- velocity auto correlation function
- mean square displacement
- energy fluctuations
- site-coordination numbers

Figure 3.2 Overview of input requirements and output generation of molecular dynamics program.

3.2 Algorithms for integration of equations of motion

The heart of a molecular dynamics simulation is the integration of the equations of motion, and the standard route for the solution of these ordinary differential equations are the finite difference approach. Nearly all simulators use either a predictor-corrector type algorithm or a Störmer type algorithm [15].

In the predictor-corrector method, new positions, velocities and accelerations are predicted from old values of derivatives at step $t+h$ (where t is time and h is timestep) by a Taylor expansions about t of the position and its derivatives. Then the equations of motion are used to calculate the forces at time $t+h$, and the new accelerations are used to correct the predicted values. The correction step can be applied, at the expense of computing time, several times to refine the final values. The order of the method and hence the degree of energy conservation can be increased by increasing the number of derivatives. The accuracy of the computed trajectories relies on the accuracy of the predicted values. But as molecular interactions in a liquid can be greatly altered from one timestep to another, higher order predictor-corrector integrators does not necessarily lead to an improvement.

The Störmer method, as adopted by Verlet, 1967 [64] is a direct solution of the equations of motion, also through a Taylor expansion of position about t. The integrator bears a similarity to the predictor-corrector algorithm, but without the corrector [65]. The new positions are calculated from the accelerations at the previous timestep. Several modifications to the basic scheme devised by Verlet have been proposed (see references in [15, 28]). Beside having lower memory requirements, the Verlet methods lead generally to better energy conservation than predictor-corrector methods at long timesteps [15, 28].

The Verlet algorithm is stable and time-reversible, but both integrators do suffer from instability, caused by the stiffness of the differential equations, when the timestep exceeds some value. A typical upper bound of timestep is ~5fs, but varies with the system under study.

On some occasions Runge-Kutta methods have also been used as integrators. They suffer however from the inconvenience that forces has to be calculated several times per timestep; and force-calculation contributes ~95% to the total computing time in a standard molecular dynamics calculation. Recently there have been made some attempts to improve upon the speed of Runge-Kutta algorithms [66, 67]. Implicit Runge-Kutta methods seem to be stable for larger timesteps than the traditionally used methods. In combination with parallel computers they can offer an alternative to the standard strategies.

3.2.1 Algorithm of translational motion

To integrate the thermostatted equations of motion, Equation (2.28), we use an algorithm due to Toxvaerd [68]. We have extended it to handle several heat baths. This algorithm is based on the Verlet leap-frog algorithm [15], where the velocities are updated at mid-steps to generate the on-step positions. Inclusion of the friction term at time t and rewriting to avoid the velocity at time t, the finite difference version of the first of Equations (2.28) reads

$$v_i(t+\tfrac{h}{2}) = \frac{v_i(t-\tfrac{h}{2})[(1-\tfrac{h}{2}\eta_j(t)) + hf_i(t) / m_i]}{1+\tfrac{h}{2}\eta_j(t)} \tag{3.1}$$

where j identifies the heat bath in action, i runs over all molecules affected by heat bath j, and over all directions. v_i is molecular velocity, m_i is mass of molecule and η_j is friction parameter for heat bath j. h is the timestep and f_i is the total force on particle i exerted by all other molecules. A precise definition of f_i is found below. The positions are then found from the mid-step velocity

$$r_i(t+h) = r_i(t) + hv_i(t+\tfrac{h}{2}) \tag{3.2}$$

To calculate all energy contributions at the same time, we also need values for the velocities at time t,

$$v_i(t) = \tfrac{1}{2}[v_i(t+\tfrac{h}{2}) + v_i(t-\tfrac{h}{2})] \tag{3.3}$$

These velocities do not enter the integration. Finally, the thermostats are updated

$$\eta_j(t+h) = \eta_j(t) + \frac{h}{Q_j}(\sum_{i=1}^{3N}\frac{p_i^2(t+\tfrac{h}{2})}{m_i} - g_j k_B T_{ext}) \tag{3.4}$$

where Q_j is thermostat inertia, g_j is degrees of freedom connected to the specific ther-

mostat, and i counts the molecules affected by the thermostat. To be able to calculate total energy of extended system, the heat bath parameter is integrated through

$$\ln s_j(t+h) = \ln s_j(t) + \frac{h}{2}\left(\eta_j(t+h) + \eta_j(t)\right) \tag{3.5}$$

The Nosé-Hoover equations are by nature time-reversible. But time-symmetry is broken in Equations (3.1) and (3.4), and some irreversible work is introduced by the thermostat. The equations does not break time-symmetry because of their form, but because of the presence of the time-dependent energy transfer rate. If we at a time t suddenly reverse the signs of all velocities and timesteps, the new incremented (decremented) velocity will depend on a η_j that has not previously influenced the motion, and the velocity will thus be slightly different from the positive direction value. Hence the trajectories in the reversed direction will depart from the trajectories in the original direction.

3.2.2 Algorithm of rotational motion

The differential equations for rotational motion cannot be solved directly from the formulation in Equation (2.5), page 22. The sine-term i the denominator of Equation (2.6) introduces a singularity at angles 0 and π. Normally, the rotational motion is therefore integrated with the aid of quaternions [69]. The quaternions are four scalar parameters, regarded as constituting a vector $\mathbf{Q}$, defined for each molecule in terms of its Euler angles (ϕ,ψ,θ) [37].

$$\begin{aligned} q_0 &= \cos\tfrac{1}{2}\theta\cos\tfrac{1}{2}(\phi+\psi) \\ q_1 &= \sin\tfrac{1}{2}\theta\cos\tfrac{1}{2}(\phi-\psi) \\ q_2 &= \sin\tfrac{1}{2}\theta\sin\tfrac{1}{2}(\phi-\psi) \\ q_3 &= \cos\tfrac{1}{2}\theta\sin\tfrac{1}{2}(\phi+\psi) \end{aligned} \tag{3.6}$$

$$q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$$

Once used to define the quaternions, the Euler angles need not appear again during the integration, and the orientation of the body is completely described by the quaternions. Equation (3.6) combined with Equations (2.5), (2.6) and (2.7) then yield the equations of motion for each molecule in terms of quaternions:

$$\begin{aligned}
\frac{dq_0}{dt} &= \tfrac{1}{2}(-q_1\omega_x^b - q_2\omega_y^b - q_3\omega_z^b) \\
\frac{dq_1}{dt} &= \tfrac{1}{2}(\ q_0\omega_x^b - q_3\omega_y^b + q_2\omega_z^b) \\
\frac{dq_2}{dt} &= \tfrac{1}{2}(\ q_3\omega_x^b + q_0\omega_y^b - q_1\omega_z^b) \\
\frac{dq_3}{dt} &= \tfrac{1}{2}(-q_2\omega_x^b + q_1\omega_y^b + q_0\omega_z^b)
\end{aligned} \tag{3.7}$$

where ω_ζ^b is rotational frequency with respect to the ζ-axis in the body fixed coordinate system [15]. The rotation matrix in quaternions is [15]

$$\mathbf{A} = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 + q_0q_3) & 2(q_1q_3 - q_0q_2) \\ (q_1q_2 - q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 + q_0q_1) \\ 2(q_1q_3 + q_0q_2) & 2(q_2q_3 - q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix} \tag{3.8}$$

The quaternions can be integrated with a leap-frog-like scheme [15], where the positions and velocities are calculated at different times. But since the quaternion derivatives depend on both angular velocity, and upon the quaternions themselves calculated at the same timestep, this implies that **Q** has to be calculated at both mid-step and full-step. Fincham, 1992 [70] developed an implicit integration scheme that made the mid-step calculation of **Q** redundant, and also the energy drift as compared to the older explicit algorithm is claimed to be reduced. In this algorithm, the integration of the rotation is accomplished by first updating the components of the space fixed angular momentum, l_i

$$l_i(t) = [l_i(t - \tfrac{h}{2}) + \tfrac{h}{2}\tau_i(t)] \ / \ [1 + \tfrac{h}{2}\eta_j(t)] \tag{3.9}$$

where i as usual runs over all molecules affected by heat bath j, and over all directions, and τ_i is the torque on molecule i created by all the other molecules. Then the half-step angular momentum is calculated for use in next timestep

$$l_i(t + \tfrac{h}{2}) = [1 - \tfrac{h}{2}\eta_j(t)]l_i(t) + h\tau_i(t) \tag{3.10}$$

The body fixed angular velocities are calculated through the rotation matrix, Equation (3.8), and the time derivatives of the quaternions at time t follow from Equation (3.7). Based on an initial guess of full timestep quaternions, $q_{\gamma i}(t+h) = q_{\gamma i}(t) + h(dq_{\gamma i}/dt)$, the full timestep value is found by iteration of

$$q_{\gamma i}(t+h) = q_{\gamma i}(t) + \frac{h}{2}[\dot{q}_{\gamma i}(t) + \dot{q}_{\gamma i}(t+h)] \tag{3.11}$$

where γ=0,1,2,3 numbers the quaternion components. The iteration is repeated until the relative accuracy is better than 10^{-7}. For the first few hundred steps with an equimolar mixture of water and ethanol, the number of iterations needed varies between 2 and 5, with 3 occurring clearly most frequent. Implicit integration thus increase the computing time by only a minor amount.

The final values are normalized by the criterion appearing in Equation (3.6). Then the thermostat parameter for rotation is updated in an expression similar to Equation (3.4)

$$\eta_j(t+h) = \eta_j(t) + \frac{h}{Q_j}\left(\sum_i \vec{\omega}_i^b(t+\tfrac{h}{2}) \cdot \mathbf{I} \cdot \vec{\omega}_i^b(t+\tfrac{h}{2}) - g_j k_B T_{ext}\right) \tag{3.12}$$

where $\mathbf{I}$ is the inertia tensor for the molecule, Q_j is the heat bath mass (not to be confused with the quaternion parameter $\mathbf{Q}$). $\vec{\omega}_i^b$ is the body fixed angular velocity[3], and i running only over molecules, not directions. g_j is the number of degrees of freedom connected to heat bath j, and k_B is Boltzmann's constant. T_{ext} is the temperature of the heat baths. Since we are restricted to equilibrium simulations, the temperatures are equal for all reservoirs. By a proper choice of rotational axis, all the off-diagonal elements of the inertia tensor are zero in the body-fixed frame.

The heat bath coordinate s_j is updated as in the translational case according to

$$\ln s_j(t+h) = \ln s_j(t) + \frac{h}{2}\Big(\eta_j(t+h) + \eta_j(t)\Big) \tag{3.13}$$

For the particular case of two thermostats, one applied to the translational motion of all molecules regardless of species, and one applied to the rotational motion of all molecules, the index j will take on the value *1*, or the label *trans*, with the translational equations of motion, and the value *2*, or the label *rot*, with the rotational equations of motion. In this case index i runs over the total number of molecules for both translation and rotation.

[3] The arrow appears as a vector symbol because the equation editor does not permit boldfaec greek letters.

We use the implicit integration scheme for all species, not limited to water only. For a full description of the algorithm, we refer to Fincham [70].

3.3 Initial and boundary conditions

3.3.1 Starting configuration

All simulations are started according to the same procedure. The system is confined to a cubic fixed volume - a box without walls, where the molecules are arranged in a face centred cubic -lattice (fcc-lattice). The orientations of the molecular axes have a 45 degree slant in alternating directions [15].

To build a mixture lattice, two separate fcc-lattices of pure species A and B (alcohol and water) are constructed, each at the same, final mixture density. An amount of n_B molecular positions are randomly picked from species A lattice and then species B molecules is inserted at those positions. Finally all the molecules are renumbered to let species A occupy the first n_A elements and species B occupy the last n_B elements of all arrays of position, velocity, and so on. We then have a random (crystal) mixture of two species prior to the simulation. This is of course a very brutal mixing method. A typical concern is that bigger molecules inserted into the lattice of smaller molecules would participate in large repulsive overlap situations that could accelerate molecules beyond our solar system. We have not seen signs of such events, so more sophisticated methods, as bringing one box of each species in contact and letting them diffuse into one another, or adding one molecule at a time, then reshape the box and equilibrate, seem unnecessary complicated.

Translational and rotational velocities are assigned randomly to the molecules from a Gaussian distribution consistent with the preset temperature. The translational velocities are modified to give a total linear momentum of zero. As total angular momentum is not conserved in the simulation, there are no modifications of the angular velocities.

During the first 25000 steps, the system is allowed to relax from the unrealistic crystal lattice to a liquid state by the means of simple velocity scaling at each timestep. Then we switch from NVE simulation to NVT simulation by invoking the thermostat and, of course, stop the velocity scaling.

The initial values for energy transfer rate between system and reservoirs, η_j, are set equal to zero, and the potential energies of the heat baths are also zero, with $s=1$. The zero-values are preserved through the equilibration period, and initialize the reservoir equations at timestep 25001.

3.3.2 System size

We have used a simulation box with a total of 256 molecules for all compositions. Ensembles of more than ~100 molecules is normally recommended for prediction of thermodynamic quantities to be close to the thermodynamic limit [28].

Cassuleras and Guardia, 1992 [34] have performed simulations for liquid methanol with varying number of molecules. They found no noticeable differences in thermodynamic averages and structure for systems of 125 to 512. For velocity auto correlation function they found minor differences, but the self-diffusion as calculated from mean square displacement showed a systematic increase with increasing system size. This indicates that even a system of 512 particles is far from the thermodynamic limit. Our self-diffusion results for mixtures is thus expected to carry a substantially amount of systematic error due to the limited number of particles.

We should also keep in mind that for a 25% mole fraction, we only simulate 64 molecules of the actual species. This results in increased uncertainty with decreasing mole-fraction in the quantities calculated for separate components at constant time-step and simulation length. This should particularly be remembered when we present correlation functions and self-diffusion coefficients at decreasing mole fraction.

The limiting factor on the system size is computer resources. As time consumption in the calculations is proportional to N^2, this is of course a severe limitation. A reduction in computer time can be achieved by using neighbour lists [15].

The box lengths corresponding to 256 molecules of water is ~20Å, for methanol ~26Å, and for ethanol ~29Å. See Table B.3 for a list of all values.

3.3.3 Periodic boundary conditions

To define physical surroundings that can represent the bulk material, we use periodic boundaries [15, 28]. For a cubic simulation box, periodic boundaries are realized by copying the simulation box in all directions infinitely. Details of all these molecules need of course not be stored physically, as each and every one can be described by a translation of the corresponding simulation box molecule by the appropriate number of box lengths whenever needed.

Strictly speaking, the system is not isolated between timesteps as molecules can leave and enter the box. A molecule that leaves the box will appear at the opposite edge (see Figure 3.3, page 60) the next time-step, carrying all its characteristics except for the position relative to a fixed frame of reference. The total linear momentum is thus conserved. The total angular momentum is not, as the distance vector changes when the molecule leaves and re-enters the box. This contradicts the well known result

from classical mechanics that a system of particles has constant angular momentum if no external torque act on the system.

Note however that the velocity scaling applied during the equilibration phase does not conserve linear momentum in the mixtures because of different molecular masses. The degrees of freedom for translation are then *3N*. Also extra momentum must be removed before invoking the thermostats to ensure that the thermostatted simulation conserves linear momentum.

3.3.4 Time step and simulation length

The production runs have been performed with a timestep of 0.5fs. This is a very small timestep for a classical MD-simulation of rigid molecules, and accordingly the total simulation time also had to be small. With a total number of steps of 160000, a simulation of 256 4-site molecules with Ewald summation took ~10 days on our DEC Alpha 3000/400 stations. The production phase of each simulation only covered 67.5ps.

A small timestep does not necessarily mean an improvement of energy conservation because the round off errors become important at small timesteps [28]. We have not selected the timestep from analysis of algorithm performance, but rather the time step selected us once the thermostat masses were fixed. This will be clarified in the paragraph following this.

3.3.5 The thermostats and their characteristics

Due to non-conservation of linear momentum it is not recommended to control each component in a mixture separately [24]. Separate thermostatting of different degrees of freedom does however not destroy linear momentum conservation, and so we have applied one heat bath to control the translational motion of both components, and another heat bath to control the rotational motion.

We then need values for the heat bath masses Q_{trans} and Q_{rot}, or as we prefer, the relaxation parameters τ_{trans} and τ_{rot} for the two reservoirs.

From a trial and error procedure, we have estimated the relaxation parameters for thermostatting the pure liquids at room temperature. They are given in Table 3.1 below.

Table 3.1
Relaxation parameters for the heat reservoirs in the NVT simulations

	$\tau_{trans}[10^{-15}s]$	$\tau_{rot}[10^{-15}s]$
water	25	20
methanol	25	23
ethanol	50	50

The starting point for each liquid was estimates of the the rotational frequencies and of the mean free paths, then we varied the relaxation parameters above and below these values. The criterion for acceptance of a value was very simple, and was inspired by the results of Di Tolla and Ronchetti, 1993 [49]: Do we find, by inspection by eye, any suspicious oscillations in the time development of either of the *s*-parameters or of the temperatures, then the trial value is rejected.

If we accept that the optimum value for a relaxation parameter is equal to a characteristic frequency of the system studied, we must conclude that each mixture has its specific best value of the relaxation parameters. To perform a trial and error procedure for each and every mixture is very laborious, so instead we have applied a weighted average of the thermostat relaxation parameters for the pure liquids

$$\tau_{trans}^{mix} = x_1\tau_{trans}^1 + x_2\tau_{trans}^2 \quad \text{and} \quad \tau_{rot}^{mix} = x_1\tau_{rot}^1 + x_2\tau_{rot}^2 \tag{3.14}$$

So finally returning to the timestep as promised above: We have found the optimum relaxation times as presented in Table 3.1 to be approximately 20-50fs. As oscillations should be integrated with at least 100 points per oscillation to avoid stiff differential equations, this constraints the timestep to be at most 0.5fs.

We expect that averages of simple thermodynamic quantities are relatively insensitive to the heat bath mass [50], whereas properties derived from the fluctuations are highly sensitive to the heat bath characteristics. Studies on solid Lennard-Jones systems [49] show that the thermostat inertia must be within a factor of ~±5 of the characteristic frequencies of the system to avoid influence of thermostat oscillations on the fluctuations of properties. We do not use the fluctuations to calculate derivative properties. As far as we know, it is not clear how heat bath parameters affect entropic properties. Nosé [24] expects that the presence of a thermostat will have little effect upon single

particle dynamics. This is confirmed by the simulations of Pierleoni and Ryckaert, 1992 [71]. Collective dynamics is on the other hand likely to be affected [24].

3.4 Force and torque calculations

3.4.1 Short range forces

In the pairwise additive approximation, the force $F_{i_a j_b}$ between two sites a and b on different molecules i and j are found from

$$F_{i_a j_b} = -\frac{dU_{i_a j_b}}{dr_{i_a j_b}} \tag{3.15}$$

where $r_{i_a j_b}$ is distance between the two sites and U is the potential of site i_a felt by j_b, given in Equation (4.1). Due to symmetry, $F_{i_a j_b} = -F_{j_b i_a}$, which means that only half of the site-site forces need to be calculated. The total force on site i_a due to all sites on all other molecules is

$$F_{i_a} = \sum_j \sum_b F_{i_a j_b} \tag{3.16}$$

With a Lennard-Jones (12-6) potential plus a Coulomb term as given in Equation (4.1) the x-component, say, of the total force on site i_a to enter the program is

$$F_{i_a}^x = \sum_j \sum_b \left(\frac{A}{(r_{i_a j_b})^{13}} - \frac{B}{(r_{i_a j_b})^7} + \frac{C}{(r_{i_a j_b})^3} \right) x_{i_a j_b} \tag{3.17}$$

The A, B, and C are short for the parameters and constants given in Equation (4.1). $x_{i_a j_b}$ is x-component of distance between the two sites.

Likewise, the total torque τ_i^x in x-direction for molecule i caused by the forces acting upon the sites of molecule i, is

$$\tau_i^x = \sum_i \sum_a \left(\Delta y F_{i_a}^z - \Delta z F_{i_a}^y \right) \tag{3.18}$$

where Δy and Δz are distance in y and z-direction, respective, between the site i_a and the centre of mass of molecule i. The remaining components are found through a positive rotation of (x,y,z).

The virial W is calculated from the site forces as

$$W = -\frac{1}{3}\sum_{i}\sum_{j>i}\sum_{a}\sum_{b}\frac{d\mathcal{U}_{i_a j_b}}{dr_{i_a j_b}}\frac{\mathbf{r}_{i_a j_b}\cdot\mathbf{r}_{ij}}{r_{i_a j_b}} \tag{3.19}$$

where $\mathbf{r}_{ij}$ is distance between centre of mass positions of the molecules i and j.

The force calculation is very thoroughly described in [15], and we have followed the guidelines provided there.

3.4.2 Force cut-off

In principle both the Lennard-Jones and the Coulomb forces range to infinity, but only the Coulomb forces are significant beyond ~2.5σ. We therefore calculate the forces on one molecule as the sum of forces, both Lennard-Jones and Coulomb, from all molecules having their centre of mass within half a boxlength of the present molecule (Figure 3.3).

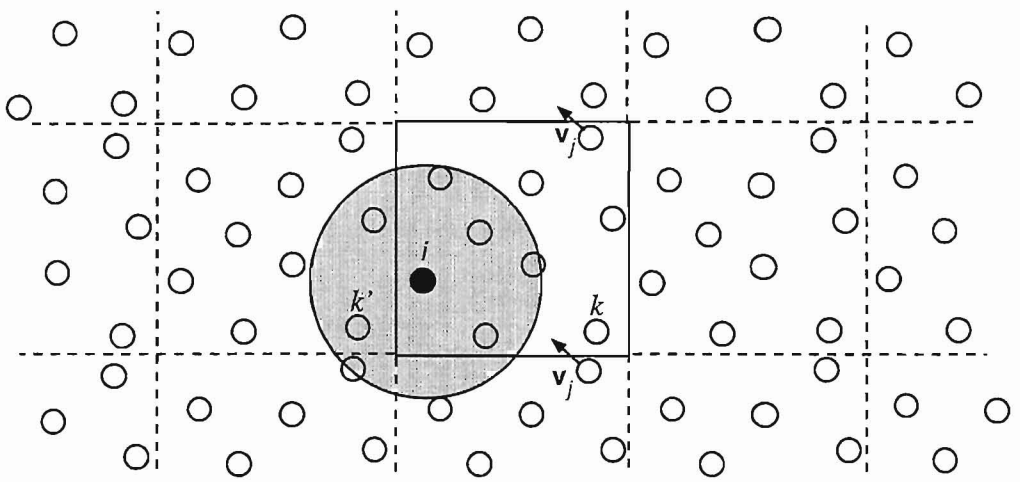


Figure 3.3 Figure above show periodic boundary conditions (PBC) and minimum image (MI) convention of a two-dimensional simulation box . Cut-off around central molecule (black circle) shown. Molecule *k* is real, but outside cut-off of molecule *i*. Molecule *i* interacts with closest copy of *k*, which is the imaginary, *minimum image* molecule, *k'*. All images are found by linear transformations with the boxlength *L* in all directions. Notice molecule leaving and re-entering the box.

If the sphere thus generated extends outside the simulation box, the periodic boundaries comes into effect through the minimum image convention [15]. A larger cut-off is not consistent with the minimum image convention, as a molecule could interact with several images of another molecule, as well as images of itself, simultaneously.

The cut-off introduces a discontinuity in the force field which lead to a discontinuity in the momentum of particle i when particle j enters or leaves the sphere. This show up as small fluctuations in the total energy of the system [28].

3.4.3 Long-range contributions to the forces

Application of a potential cut-off means that only a portion of the contributions to the physical properties is accounted for. We therefore add the Lennard-Jones contribution from the far-lying molecules directly to the energy and pressure by Equations (3.30) and (3.31) below. These are constant contributions calculated once and for all at the beginning of each simulation. For configurational energy this correction amounts to less than ~0.1J/g, so we are relative confident that the forces and hence the velocities are not seriously affected by the cut-off in the Lennard-Jones interactions.

Simple cut-off in the electrostatic forces is not advisable, at least for systems of less than ~1000 molecules [72]. These forces are of longer range than the Lennard-Jones forces, and a spherical cut-off can lead to increase in total energy [73, 74]. There are several methods [15,73] to include the long-range electrostatic forces into the force-sum. We have applied the Ewald summation technique, originally devised to calculate lattice sums for ions in a crystal [75, 76]. Despite its time-consumption, we chose to use the Ewald method, mainly because we deal with highly polar species. The standard method is extensively described elsewhere, see for example [15, 77].

The underlying philosophy is to replace the original electrostatic potential by a charge density $\sim e^{-r^2}$ with the same total charge q. The potential at each site is now of short range. Then a new potential for a lattice of point charges and a Gaussian charge distribution of opposite sign is subtracted. The result is a total potential with two main contributors, which is treated mathematically different. The damped interactions is calculated in r-space, while the cancelling distribution is calculated in k-space. The total potential between two sites i_a and j_b with fractional charges q_{i_a} and q_{j_b} is given by [79]

$$U_{i_a j_b} = cq_{i_a}q_{j_b}\frac{\mathrm{erfc}(r_{i_a j_b}\kappa)}{r_{i_a j_b}} + \frac{c}{\pi L^3}\sum_{k\neq 0} q_{i_a}q_{j_b}\frac{4\pi^2}{k^2}\mathrm{e}^{-k^2/4\kappa^2}\cos(\mathbf{k}\cdot\mathbf{r}_{i_a j_b}) - \frac{2c\kappa}{\sqrt{\pi}}q_{i_a}^2 - \frac{cq_{i_a}q_{i_b}}{r_{i_a i_b}}\mathrm{erf}(r_{i_a i_b}\kappa) \qquad (3.20)$$

where the first term is the **r**-space contribution, the second term is the **k**-space summation, and the final two terms are self contribution. c is the Coulomb constant $1/4\pi\varepsilon_0$, L is length of simulation box, κ is a convergence parameter. $\mathbf{k} = 2\pi\mathbf{n}/L$ is a reciprocal lattice vector, and $\mathbf{r}_{i_a j_b}$ is distance between sites. `erf` and `erfc` is the error function and the complementary error function [78].

The final equations [79, 80] for the contributions to the forces are given below for the ζ-direction. The real-space contribution is given by

$$F_{i_a j_b}^{R,\zeta} = \frac{c q_{i_a} q_{j_b}}{r_{i_a j_b}^2}\left(\frac{\mathtt{erfc}(r_{i_a j_b}\kappa)}{r_{i_a j_b}} + \frac{2\kappa}{\sqrt{\pi}} e^{-r_{i_a j_b}^2 \kappa^2}\right)\zeta_{i_a j_b} \tag{3.21}$$

The **k**-space contribution is given by

$$F_{i_a j_b}^{k,\zeta} = \frac{c}{\pi L^3}\sum_{k\neq 0} q_{i_a} q_{j_b} \frac{4\pi^2}{k^2} e^{-k^2/4\kappa^2} \mathtt{sin}(\mathbf{k}\cdot\mathbf{r}_{i_a j_b}) k_\zeta \tag{3.22}$$

and the self-term is given by

$$F_{i_a j_b}^{S,\zeta} = \frac{c q_{i_a} q_{i_b}}{r_{i_a i_b}^2}\left(\frac{\mathtt{erf}(r_{i_a i_b}\kappa)}{r_{i_a i_b}} - \frac{2\kappa}{\sqrt{\pi}} e^{-r_{i_a i_b}^2 \kappa^2}\right)\zeta_{i_a i_b} \tag{3.23}$$

We use a lattice parameter κ of 5.0/L, where L is the boxlength, and a total of 111 k-vectors. A variation of these parameters was not included in our study. The implementation of the k-space summation is performed with trigonometric functions.

During the last 5-10 years a number of faster performing implementations and alternatives to the Ewald sum have been proposed, mainly due to the time consumption involved in the calculation [81-86]. Discussions on possible effects of the use of the Ewald technique indicate that Ewald summation might influence on dynamical properties to give a lower diffusion constant than various cut-off methods [87]. On the other hand there is indication that Ewald summation give higher diffusion coefficients than the nearest image method [88].

3.5 Calculation of properties

3.5.1 Statistical analysis and sampling procedures

One of the main purposes of molecular dynamics simulation is to calculate mean values for properties. Rigorous conclusions from the simulations can however only be drawn within the statistical uncertainties of the properties in consideration. The error in a calculated property is caused by a combination of systematic error and statistical error.

Systematic errors displace the simulation average from the true trajectory average. As in laboratory experiments, systematic errors are caused by method or external influence. In a computer simulation truncation of finite difference equations, truncation in force-calculation, periodic boundary conditions, the finite number of molecules used, the sampling method, and hidden errors in the code can act as examples. Systematic errors does not improve with extended simulation. To reveal the presence, but not necessarily identify the sources, of any systematic errors, the following strategies are useful [28]:

For one particular problem

- run the same code on different computers with different compilers
- use the same algorithm, but different code
- use different algorithms for the same problem
- compare results from different simulation techniques

We have performed the first test by running our code at different optimization levels at Alpha 3000/400, and by running the same problem at CONVEX 220 at lowest optimization level. The results are shown in Table 3.2. The last three tests are performed by comparison with other published results from Monte Carlo or molecular dynamics simulations. As all details of a simulation seldom are reported, this approach can only give indications. Comparisons are discussed along with the presentation of our results in Chapter 5 and 6.

The representation of bulk material by a small number of particles introduces a systematic error that can be investigated if we increase the molecule number. Extrapolation of values from runs with increased molecule numbers will give the thermodynamic limit.

Statistical errors are the distribution of individual observations around the mean value and reflect the natural fluctuations of the observations. They are usually quantified as the variance or the standard deviation of the property. Thermodynamic equi-

librium properties are randomly drawn from a Gaussian distribution; the randomness are caused by the chaotic nature of the phase-space trajectory. Adjacent points on the phase space trajectory are however not uncorrelated. Each property has a definite correlation time within which their values depend on previous values. Such serial correlations can introduce systematic errors into the computed averages, unless special care is taken. A reliable calculation of averages can be done with one of the following methods [15, 28]:

- stratified systematic sampling
- stratified random sampling
- coarse-grained sampling

We have calculated the averages of thermodynamic properties with coarse-graining, while the static structure is sampled with stratified systematic sampling.

To start with, we define the full phase-space trajectory as the trajectory our system traces out within a macroscopic time. The computed phase-space trajectory consists of a finite number of points where we follow the system, so this sample is only a very small portion of the full trajectory. We want to approximate the averages of the computed trajectory to the full trajectory averages, and also to have a measure of the variance in the mean of the calculated trajectory with respect to the full trajectory mean.

In coarse-graining this is accomplished by dividing the computed trajectory into M segments each consisting of n observations. The values in each segment j are averaged according to

$$\bar{x}_j = \frac{1}{n}\sum_{k=1}^{n} x_k(k\Delta t) \tag{3.24}$$

where n is the number of observations in each segment, x_k is the observed value and Δt is the time increment. Then the expectation value $\langle x \rangle$ and its standard deviation σ is calculated from

$$\langle x \rangle = \frac{1}{M}\sum_{j=1}^{M} \bar{x}_j \tag{3.25}$$

and

$$\sigma = \frac{1}{\sqrt{M}}\sqrt{\frac{1}{M-1}\sum_{k=1}^{M}\left(\bar{x}_k - \langle x \rangle\right)^2} = \frac{1}{\sqrt{M(M-1)}}\sqrt{\sum_{k=1}^{M} x_k - \frac{1}{M}\left(\sum_{k=1}^{M} x_k\right)^2} \tag{3.26}$$

where the quadratic is expanded and recast for use within the program with the aid of Equations (3.24) and (3.25).

The calculation of standard deviations are done within the program, based on accumulated values. The calculated average will be within the trajectory average for 95% of the values if we assign an uncertainty of 2σ. Normally, we report the value corresponding to 1σ.

We stress once more that the above strategy relies on the fact that the segment duration must be larger than the correlation length for the property. Correlation length vary with property, so ideally one should calculate time-correlation functions for each property and be prepared to apply different segment lengths. We have calculated velocity auto correlation function, and use this correlation length to estimate both the systematic sampling of the static structure and the segment length of the coarse-grained sampling. From figures of the velocity auto correlation functions, see Figures 5.35, 5.36, page 137, and 6.31, 6.32, page 180, we see that correlations die after ~$0.6 \cdot 10^{-3}$ns. With a step length of 0.5fs, we should have segment lengths of about 1200 timesteps, with a total number of segments of 112 for 135000 steps of production. We have used a segment length of 300, which brings us to only $0.15 \cdot 10^{-3}$ns, where some correlation still exists. We have not evaluated the possible influence of this.

Even if the standard deviations are small, and the systematic errors are controlled, the distribution of observations around the mean can be wrong. A quantitative measure of this is the computation of χ^2 [28], which measures how well the actual distribution is correlated with the theoretical distribution. A qualitative measure is to accumulate values and compare to the theoretical distribution in a plot. We have done so for the internal energy and the linear velocities, discussed in Sections 5.3 and 6.2.

3.5.2 The simple thermodynamic quantities

Mixture pressures P, temperatures T, configurational energies E_p, kinetic energies E_k, and constants of the motion (i.e. internal energy U with the NVE simulations and extended system energy E_{tot} with the NVT simulations) were calculated each step and accumulated over segments of 300 steps in the production phase. Simulation averages and standard deviations were calculated at the end of each simulation.

In addition, time development of the above properties to be used in graphical representation were saved to disk as instantaneous values each 100 step. We also saved the translational and rotational temperatures of each component, and the total reservoir potential and kinetic energies, see Equation (2.29), at the same intervals.

Some of the properties were not averaged (coarse grained) during program execution, but were averaged after the simulation. These averages are based on the instantaneous output values (systematic sampling) with few (3-5) significant numbers. It will be evident from the respective tables to which properties this has been applied. Due to the different sampling method, the standard deviations reported are larger.

The temperatures are calculated from the kinetic energies of each component according to the law of equipartition of energies

$$T_{trans} = \sum_i m_i v_i^2 \,/\, g_{trans} k_B \tag{3.27}$$

and

$$T_{rot} = \sum_i I\omega_i^2 \,/\, g_{rot} k_B \tag{3.28}$$

where $g_{trans} = 3(N-1)$ and $g_{rot} = 3N$ are the translational and rotational degrees of freedom for each component. g_{trans} and g_{rot} differ because there is no constant of the motion associated with the rotation. The mixture averages are then found as weighted averages of the contributions from each component.

The pressure is calculated from the virial W, Equation (3.19), page 60, as

$$P = \frac{Nk_BT}{V} + \frac{W}{V} + P_{LR} \tag{3.29}$$

where N is number of molecules, V is the volume of the simulation box, T is the calculated temperature. The first term is thus the ideal gas contribution, and P_{LR} is the Lennard-Jones contribution from the molecules lying outside the cut-off sphere, given as [28]

$$\frac{P_{LR}}{\rho k_BT} = \frac{-2\pi\rho}{3k_BT}\int_{r_c}^{\infty} r\frac{du_{LJ}(r)}{dr}g(r)r^2dr \approx -\frac{16}{3}\frac{\pi\rho}{k_BTr_c^3}\sum_{a,b}\epsilon_{ab}\sigma_{ab}^6 \tag{3.30}$$

where the radial pair correlation function $g(r)$ is assumed to take the value 1 and u_{LJ} is the total site-site potential. ρ is the number density, T is calculated temperature, r_c is potential cut-off radius, and ε_{ab} and σ_{ab} are Lennard-Jones parameters for the sites a and b.

The long range contribution to configuration energy is included as [28]

$$U_{LR} = -2\pi\rho \int_{r_c}^{\infty} u_{LJ}(r) g(r) r^2 dr \approx -\frac{8}{3}\frac{\pi\rho}{r_c^3} \sum_{a,b} \epsilon_{ab}\sigma_{ab}^6 \tag{3.31}$$

3.5.3 Structural quantities

The normalized radial site-site correlation functions $g_{ab}(r_{ab})$ are calculated from

$$g_{ab}(r_{ab}) = \frac{\sum_k N_k(r_{ab}, \Delta r)}{\frac{4}{3}\pi\rho V(r, \Delta r) S n} \tag{3.32}$$

N_k is an array that accumulates neighbouring sites b lying within a shell of width Δr centred at a distance r_{ab} from site a. The total number of sites is divided by the number that would have been found in an ideal gas, by the number of steps S, and the number of sites n contributing to N_k. In the long distance limit, $g_{ab}(r_{ab})$ will approach unity.

The calculation is based on samples every 150th step, with a radial resolution of 0.05 Å. The structure is then sampled from 900 configurations. This number is fixed, so the statistical accuracy will vary with composition for various site-site correlations. The errors in peak heights are believed to be quite large, especially at short distances where the shells contain few molecules. A full analysis of standard deviations in g(r) is time-consuming, and is thus not performed. We have however applied Gaussian statistics to oxygen-oxygen self correlation function for water, g_{OO}, for one position near first maximum and another at long distance. The results are shown in Table 5.8. The standard deviations are found to be small.

Calculation of site-coordination number, n_{ab}, that is, the average number of nearest neighbouring sites of type b surrounding the central site a, is found by a Simpson integration of the site-site correlation functions out to the first minimum of each $g_{ab}(r)$.

$$n_{ab}(r_{min}) = 4\pi\rho_b \int_0^{r_{min}} g_{ab}(r) r^2 dr \tag{3.33}$$

ρ_b is number density of sites b in system.

3.5.4 Dynamical properties

The velocity auto correlation function $\Psi(t)$ is computed after the simulation is finished from all centre of mass velocities saved to disk at every 100th timestep. The output of all molecular velocities act only to improve the statistical precision since this is a single-particle correlation. The correlations are integrated up to 500 points, which means 50000 timesteps, or a delay time of 25ps. The storing requirements for the output is very large, but we wanted to have the possibility to do some variations in the final calculations.

The velocity auto correlation function $\Psi(t)$ is defined as [28]

$$\Psi(t) = \left\langle \mathbf{v}_i(t_0) \cdot \mathbf{v}_i(t_0 + t) \right\rangle \tag{3.34}$$

where $\mathbf{v}_i(t_0)$ is the velocity of molecule i at time origin t_0. The calculation of Ψ is done according to (but for each component separately)

$$\Psi(t) = \frac{1}{MN} \sum_{k}^{M} \sum_{i}^{N} \mathbf{v}_i(t_k) \cdot \mathbf{v}_i(t_k + t) \tag{3.35}$$

where M is number of available time origins, and N is the number of molecules. M is given by $M=L\text{-}t_{delay}/\Delta t$, where L is number of stored velocities. In our calculations we have 1350 stored velocities at intervals of 100 steps, and we have used a delay time of 50000 steps, which give us 850 origos.

The self-diffusion D of each component can either be calculated from the velocity auto correlation functions through the Green-Kubo relations [15, 28] or from the modified Einstein-relation [15, 28]

$$D = \frac{1}{6N} \lim_{t \to \infty} \frac{d}{dt} \left\langle \sum_{i=1}^{N} \left[r_i(t) - r_i(0) \right]^2 \right\rangle \tag{3.36}$$

based on accumulated positions for each species separately. We calculate the mean-square-displacement, represented by the brackets, for a delay of 5000 time steps (usually 500 positions). The positions are taken every 100th timestep. We have not calculated the statistical accuracy of the presented dynamical properties, but we discuss a probable estimate for the self-diffusion coefficients in Subsection 5.5.1.

3.6 Verification of the simulations

We can never guarantee that our code is free of errors, nor that the results meets with the theoretical restrictions. There are however several ways to gain increased confidence in the validity of the output. We have been concerned with elimination of code errors (of course!), verification of liquid equilibrium, and investigation of the ergodic behaviour of the system. In what follows, we will describe the methods and various tests we have applied to assess the reliability of the results. It has not been possible to verify each and every aspect of the simulation, but we have selected what we believe are the most important and the most interesting, balanced against the complexity in performance.

3.6.1 Code errors

We have compiled the code with three different compilers, the f77 compiler at DEC Alpha 3000/400, the fc compiler at Convex 220 and the Convex Application Compiler. The last compiler performs interprocedural analysis that tracks the flow of data and control between procedures. It is thus capable of, for instance, detecting errors that arise between different procedures, or inconsistencies in common blocks.

A system run under the same conditions at the Alpha and Convex computers would not generate identical trajectories as the organization of operations is different. Hence the accumulation of round-off errors is also different. We have compared the same code at two different computers and at two different levels of optimization at the same computer. Any differences in mean values will provide us with an insight to round-off errors. Table 3.2 give results for mixture temperature T, mixture pressure P, and mixture potential energy E_p for a equimolar mixture of water and ethanol.

Table 3.2

NVT-results for a water-ethanol mixture of 128 molecules of each, run for a total of 50000 steps at steplength 1fs and with properties averaged for the last 35000 steps. Standard deviation σ for each property is reported in parenthesis. All simulations are double precision. Ethanol potential slightly different from the OPLS-model applied in Chapter 6.

Computer	Compiler	<T> [K]	<P> [MPa]	<U> [J/g]
Convex 200	No optimization	292.7 (0.4)	81.5 (1.8)	-1287.0 (0.5)
DEC Alpha	No optimization	292.7 (0.4)	81.9 (1.9)	-1281.6 (0.5)
	Optimization	293.8 (0.4)	79.2 (1.9)	-1291.5 (0.6)

The deviations relative the non-optimized DEC Alpha simulation are largest for the pressure by nearly 3%, while the deviation in potential energy is less than 1%. We notice however that the deviation in potential energy is well outside $\pm 2\sigma$. Haile [28] present an analysis of a 1-dimensional Lennard-Jones system of 100 atoms simulated for 10000 and 20000 step, and our findings are of the same order as his. As we have simulated twice as long and also use a more complicated potential, we find our results to be satisfying, and conclude that the deviations are of numerical nature.

Different mean values show that the system follow different trajectories for different computers or compilers. From inspection of the time development of potential energy we find (not presented) that within 1500 timestep the differences has grown larger than the fourth digit for different computers. The same difference does not show up until 15000 timesteps for different compilers for the same computer.

From this brief investigation we also notice that the differences are larger with different optimization levels than with different computers. Note that the non-optimized and the optimized simulations with DEC Alpha stations yield different temperatures, which may in part be responsible for the differences in pressure and configurational energy.

All our production runs are performed with DEC Alpha stations at the highest optimization level.

We have rewritten central parts of the code to detect errors, and we have run our program for simple, atomic one component systems. The mixture simulation has been tested out for various compositions of water in water, and alcohol in alcohol. The results are the same as with one component.

As total energy of the extended system, Equation (2.29), and the Hamiltonian of the isolated system are strictly constant in an ideal simulation, monitoring these constants of the motion is a powerful check of code errors. To have a constant total energy, fluctuations in all contributions to total energy must cancel. Errors tend to show up quickly as a divergence. Some minor departures from ideality must be tolerated though, as the finite difference equations only are correct to a definite order in h. We are satisfied with an accuracy in the constants of motion to the fourth digit. The role of truncation can be evaluated by decreasing the timestep, then the accuracy should improve.

For each simulation we inspect the constant of the motion and estimate the drift as difference between lowest and highest values, normally appearing at the beginning and end of simulation, respectively.

3.6.2 Equilibrium liquid state and stability, general requirements

To have thermodynamical equilibrium, the system should be mechanically and thermal stable. That is, all thermodynamic properties of the system should fluctuate around stable mean values, and the 1st law of thermodynamics should be satisfied. It is also important that the system is in the intended state (liquid).

For all simulations we have therefore controlled the performance of the following quantities:

- The translational order parameter ρ given by [15]

$$\rho(k) = \frac{1}{N}\sum_{i=1}^{N} \cos\,(\mathbf{k}\cdot\mathbf{r}_i) \tag{3.37}$$

  where N is the number of molecules, $\mathbf{r}_i$ is the position of the ith molecule, and $\mathbf{k}$ is a reciprocal vector of the initial lattice. We start from a solid lattice where ρ has the value unity, but in the liquid state we expect r to fluctuate around zero with an amplitude of $1/N$.

- Energy conservation, the values of the constants of the motion defined in Equation (2.9) for the NVE simulations and in Equation (2.29) for the NVT simulations should be not only stable but also showing no fluctuations (see Subsection 3.6.1).
- Stability of time development of pressure and temperature.
- The velocity distribution initially taken from the Maxwell distribution should not depart from this. The x-component of linear velocity is sampled every 150th step, and the distribution is compared to the theoretical at the end of simulation.

$$p(v_{ix}) = \left(\frac{m_i}{2\pi k_B T}\right)^{\frac{1}{2}} e^{(-m_i v_{ix}^2/2k_B T)} \tag{3.38}$$

  where m_i is mass of molecule i, v_{ix} is velocity of molecule i in the x-direction and T is calculated temperature.

- The kinetic and rotational energies should be equally partitioned among all coordinate directions, and we save the square of rotational and translational velocities every 150th step.
- The total linear momentum is conserved in molecular dynamics simulations, the angular momentum is not. A sum over all components of linear and angular momenta is performed each 150th step, and saved to disk.

These and other tests are thoroughly discussed in [28].

3.6.3 Particular requirements for the extended system simulations

For the thermostatted simulations it is particularly interesting to see whether the simulations compare to the canonical ensemble. As discussed previously in Section 2.5, this is not guaranteed in advance. Also the extended system must satisfy the 1st law of thermodynamics, and there should be thermal equilibrium between the system and the reservoirs. We have thus looked at some quantities specific to the thermostatted simulation where the theoretical outcome is known.

- The reservoir parameters s_{trans} and s_{rot} are initialized to 1 at the beginning of the thermostatted simulation. Their average should not depart from this value, though they naturally will fluctuate. The parameters are saved to disk every 100th step, and coarse grain averages and standard deviations are calculated within the program.
- To have thermodynamical equilibrium between the system and the reservoirs, the rates of energy transfer η_{trans} and η_{rot} must fluctuate around zero. If not, there will be a net energy flux across the system boundaries. We therefore monitor the time-development of η_{trans} and η_{rot}. Not only should their mean values be close to zero separately, but the sum of their averages should also be closer to zero than the individual ones. Coarse grain averages and standard deviations are reported, and instantaneous values are saved to disk every 100th step
- From the equipartition principle, the kinetic energy of a thermostat should be proportional to $\frac{1}{2}k_BT$. The potential energies of the reservoirs should ideally fluctuate around an average of zero.
- The running averages of the higher moments of translational, rotational, reservoir translational and reservoir rotational energies. The first moment is simply the mean. If taken about the mean, the second is the variance, the third moment represents skewness, and the fourth is kurtosis which measures the degree of peakedness [89]. Also the fifth moment is included. In a canonical simulation they are expected to approach the values given below [47]:

$$
\begin{aligned}
\langle K_j \rangle &= \frac{g_j}{2} k_B T \\
\langle K_j^2 \rangle &= \frac{g_j}{2} (k_B T)^2 \\
\langle K_j^3 \rangle &= g_j (k_B T)^3 \\
\langle K_j^4 \rangle &= \frac{g_j}{2} [\frac{3g_j}{2} + 6](k_B T)^4 \\
\langle K_j^5 \rangle &= g_j [5g_j + 12](k_B T)^5
\end{aligned}
\tag{3.39}
$$

j specify the particular energy as listed above. g_j is the corresponding number of degrees of freedom, see Subsection 2.5.2. It takes the value unity for the reservoir energies. The simulational m'th moment is calculated from expansion of

$$
\langle K_j^m \rangle = \langle \, (K_j^{calc} - \langle K_j^{calc} \rangle \,)^m \rangle \tag{3.40}
$$

- Finally we analyse the distribution of simulation box energies (internal energies). In a canonical simulation the internal energy should be consistent with a Gaussian distribution.

$$
p(\mathcal{U}) = \left(\frac{m^2}{6\pi n (k_B T)^2} \right)^{\frac{1}{2}} e^{(-m^2(\langle \mathcal{U} \rangle - \mathcal{U})^2 / 6n(k_B T)^2)} \tag{3.41}
$$

Ergodicity in connection with the Nosé-Hoover thermostat was discussed in Section 2.5. To have reliable averages, the simulation of both the system and its sorroundings must sample a representative part of phase space. But even if the mean values are stable and give exellent values, the fluctuations around the mean can be highly non-canonical. The outcome of the two final tests above signals the canonical behaviour of the system.

Chapter 4
Molecular models

Today the most uncertain aspect of a computer simulation is related to the models of molecular interactions. Even with the most stable algorithm and an ergodic system, the computed results will not be closer to the real world than the potential model permits. Much effort has therefore been put into the task of constructing reliable models. So far no single empirical model of a real liquid is capable of reproducing all experimentally measured properties at all thermodynamical states of interest. With quantum mechanical calculations it is in principle possible to calculate energies accurate, but such *ab initio* calculations are computational expensive. We will not deal with that subject here.

An important consideration is the computational effort for a model, which means that a model must be simplified relative to reality. To make reasonable simplifications for the fluid, one must have an idea of the physical origin of the interactions. We will therefore briefly review the nature of molecular interactions and discuss some general considerations for the empirical models within the context of molecular simulations. Then we describe the particular models that we have chosen. A further complication is introduced when pure liquid models are combined in mixtures, so we conclude with a discussion of the applicability of pure species models in mixtures.

A description of the experimental results and proposed models from the literature for the structure of the bulk liquids and the aqueous mixtures of methanol and ethanol are deferred until Subsections 5.4.1, 6.3.1, and 6.3.2.

4.1 Molecular interactions

Molecular interactions are mainly of electrostatic and quantum mechanical origin. Magnetic interactions are usually negligible. The electrostatic interactions between the electron distributions of two molecules i and j sufficiently distant apart, are usually approximated with a series expansion in $\mathbf{r}_{ij}^{-1}$, where $\mathbf{r}_{ij}$ is their distance. This multipole expansion [90] increase the refinement of the force via single charge-, dipole-, quadrupole-, octopole-interactions, and so on. The extension and anisotropic distribution of the charge cloud is thereby taken into account. For a neutral molecule, say water, the first non-vanishing term is the dipole, but since water is non-linear and consists of unlike atoms, it also has a quadrupolar term which can contribute significantly.

Whether a molecule has a permanent dipole or not, it can always have an induced dipole (more generally an induced multipole) caused by the response of the electron distribution to the electrostatic field from the surroundings. The modification of the interactions caused by the changed electron distributions is termed induction interaction. The ratio between permanent dipole interactions and induced dipole interactions scales as r_{ij}^{-3} . Induced dipole-induced dipole interactions is the leading induced multipole interaction term, and is of the order of r_{ij}^{-6} . Induction forces between molecules are attractive for all orientations.

The dispersive interactions (London forces) are of quantum mechanical origin. They can however be viewed semi-classical as a result of fluctuating charge densities in the molecule caused by the electronic motion, which then creates instantaneous multipoles. In contrast with induction interactions, which are created by the surroundings, the dispersion interactions have their origin in the molecule itself. Dispersion interactions also scales with r_{ij}^{-6} to the first order (higher–order dispersion interactions is possible) and are attractive for all orientations.

Of interest to us are also overlap interactions, interactions between molecules coming so close that the Pauli exclusion principle becomes important. Overlap interaction is usually repulsive, but in some cases, as for instance when a proton donor like hydrogen and a proton acceptor like oxygen come close, it can be attractive and a bonding can occur. Because of the orientation of the quantum mechanical orbitals, these forces are directional dependent. There is no general theory [90].

In close connection to molecular interactions is the concept of polarizability. This is a measure of the extent to which multipoles are created or changed upon interactions from the surroundings. The polarizability of a molecule must therefore change with changing surroundings or state conditions. Polarizabilties are in general different for different directions in the molecule.

A detailed mathematical treatment of intermolecular forces is given by Gray and Gubbins [90].

4.2 Modelling the interactions

Polar molecules consists of atoms with different electronegativities. This causes a displacement of the electron distribution towards the more electronegative atom, imposing a permanent dipole moment upon the molecule. The molecule as a whole is however electrostatically neutral. Molecular liquids have anisotropic short range repulsion. Polar liquids also have a long ranged dipolar interaction.

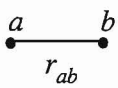
The various kinds of (anisotropic) forces acting between polar molecules, combined with the close packing of molecules in a liquid, makes it difficult to construct general theoretical models for polar fluids. For the same reason, it is also difficult to reproduce all the observed effects by a simple empirical potential.

The basic principles for classical founded models, are

- the assumption of pairwise additive interactions,
- and the idea of subdividing the molecules into interaction centres, or sites.

The sites are not necessarily coincident with the atomic centres of mass, between which central forces act. The pairwise additivity for interactions between molecules is thus extended to apply to parts of molecules via the site-site approximation.

a)



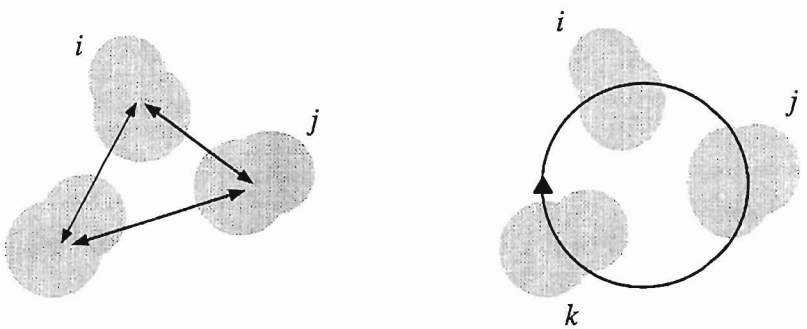
b)

Figure 4.1
Site-site approximation of two-atomic molecule.
a) Charge distribution
b) Site-representation, *a* and *b* located a distance r_{ab} apart, and characterized by 'some' parameters, see text.

The total potential U between the molecules can be expanded in a sum involving an increasing number of molecules $U = \Sigma U_{ij} + \Sigma U_{ijk} + ...$, where the first term adds up all pair interactions as if each pair was alone, the second term corrects for the mutual influence of all possible combinations of three isolated molecules, and so on [90]. The evaluation of the sum gets quickly very laborious, so the sum is usually truncated after the first term, leaving out the triplet and higher order interactions. Multipole interactions (electrostatic) are pairwise additive. The presence of a third molecule can

however have important influence on the charge distributions of the two molecules, and result in different interaction between the pair.

The assumption of pairwise additivity is to some extent compensated by the replacement of the pair potential with an effective potential, where the effects of the surroundings are taken into account in an average way. The molecule is thus to be regarded as non-polarizable.



a) Pair interactions b) Triplet interactions

Figure 4.2 Illustration of triplet forces. Presence of molecule *k* pulls electron cloud of molecule *i* and influence on orientation of molecule *j*. At the same time *i* and *k* are influenced by new orientation of *j*....

A common starting point for an effective site-site potential $U_{i_a j_b}$ is the choice of a Lennard-Jones (12-6) potential [91] for each site in combination with a Coulomb potential if the site has a partial charge.

$$U_{i_a j_b} = 4\pi\epsilon_{i_a j_b}\left[\left(\frac{\sigma_{i_a j_b}}{r_{i_a j_b}}\right)^{12} - \left(\frac{\sigma_{i_a j_b}}{r_{i_a j_b}}\right)^{6}\right] - \frac{1}{4\pi\epsilon_0}\frac{q_{i_a} q_{j_b}}{r_{i_a j_b}} \tag{4.1}$$

$\sigma_{i_a j_b}$ and $\varepsilon_{i_a j_b}$ are the Lennard-Jones parameters characterizing the repulsive core and the depth of the attractive potential well, respectively. $r_{i_a j_b}$ is the distance between to sites *a* and *b* on to different molecules *i* and *j*. q_{i_a} and q_{j_b} are the fractional charges of the sites *a* and *b*. $1/4\pi\varepsilon_0$ is the Coulomb constant in the metric system.

As the Lennard Jones form consists of a combination of a repulsive core (r^{-12}) and an attractive part (r^{-6}), effects of both dispersive and induction interactions are included to the first order, along with overlap forces.

For a potential of this kind, one has three unknown parameters for each site. As defined above, these are the electrostatic charge q, the contact distance σ between two sites, and the strength ε of the potential. In addition, a geometry for the molecule is also needed. A common strategy is to take the measured gas phase geometry represented by bond angles and bond lengths. The potential parameters are adjusted in an optimalization procedure, aiming to reproduce experimental values for some thermodynamic properties for the real, pure, liquid. Often the geometry or part of it is also adjusted from its gas phase values simultaneously. Then one has a model that can reproduce some of the features of the real fluid within a given accuracy, within a specified temperature range or density range, and for a subset of its properties. The model for the single molecule will necessarily be different from a real monomer since effects of higher order interactions are included only on average. As the calculation of site-site distances is computational expensive, the models with the least number of sites is in general preferred.

The use of an effective potential can have important consequences. As the effect of the surrounding molecules only is included on average, the short range structure is affected. Also the directional dependence on the polarizability is averaged. Studies on atomic liquids show that three-body interactions can have large effects on calculated pressure and surface tension. For polar molecules it is suggested that also higher order (three- and four-body) terms are significant [90, 92]. Also the potential usually is adjusted for a specific state of the pure species, and must be used with care at other densities and in mixtures where the surroundings are different. Another defect is that bonding charge density often is neglected. This can have consequences for the long range and multipole interactions, and can lead to wrong prediction of dimer structure [90].

In addition to the non-polarizable approximation also a rigid model where molecular bonds and angles are not subject to change is often assumed. Also the thermal vibrations of the molecules are usually not included. The bond vibrations are faster than the translational or rotational motion and must be integrated on a shorter timescale. This leads to a complication and increase in computer time. Also the vibrations are of quantum mechanical nature and might thus not be correctly integrated within a classical framework [93].

4.3 Models for water

4.3.1 Physical properties

Because of its puzzling properties and its importance in terrestrial life and activities, water has for the last 20 years or more, been the subject of intense research with the aid of theoretical, experimental and simulational techniques. Of vital importance to simulation and theoretical calculations is the model of the interactions, and numerous potentials has been suggested. So far no model has been found to capture all the features of liquid water.

Among some of its observed features, liquid water has a large specific heat, indicating that a great number of hydrogen bonds have to be broken when heated. Also water has high melting and boiling temperatures compared to other hydrides [94]. Other unusual properties are the expansion upon freezing, the existence of a maximum density at 4°C at atmospheric pressure, and the melting of ice with increasing pressure.

For a gas phase monomer, the intramolecular O-H bond length is 0.9575Å, and the HOH angle is 104.51° [95]. Of the molecular features of water is its high, compared to its size, dipole moment μ=1.85D and a polarizability of 1.46Å^3, both gas phase values [96]. Water also possesses quadrupole and octopole moments [90]. The non-vanishing moments indicate a directional preference for bonding. The polarizability imply that the pairwise additivity assumption might be substantially in error, since interactions between a pair of water molecules can not be seen isolated from their varying neighbours.

The polarity of water due to the covalent bond between oxygen and hydrogen, along with the nonshielded proton, provides for a intermolecular O — H · · · O bond; the hydrogen bond[4]. A neighbouring proton bonds to one of the lone-pairs of electrons on an oxygen atom. It is the ability of forming hydrogen bonds that is regarded as the main cause for the deviations of the behaviour of liquid water from that of a normal liquid. Not surprisingly, the hydrogen bonds have been/are particularly focused upon in simulations studies.

4.3.2 Molecular models for liquid water

An important class of empirical models originate from the site-site Lennard-Jones plus electrostatic potential model as described above. The most commonly used mod-

[4] Hydrogen bonds are mainly found to act between hydrogen and the elements oxygen, nitrogen and fluorine, but can also exist between other elements that differ in electronegativity.

els within this strategy, are the ST2 [97, 98], SPC [99] and the various TIPS [100, 101, 26] potentials. They are all rigid and non-polarizable. See Beveridge *et al.* 1983 [102] for a more detailed summary of potentials.

The ST2-potential is based on a potential of Ben-Naim and Stillinger, 1972 [103], which was one of the early attempts to model water interactions. Their idea was to use Lennard-Jones parameters for neon, which resembles water with respect to mass and electronic configuration. Water was then regarded as an atom with the hydrogen electrons filling the outer shell completely. Their model was refined by Rahman and Stillinger, 1973 [97]. The electrostatic interactions are modelled by four tetrahedrally arranged charge sites, with the fractional charges of ±0.23e. This model yields reasonable values for a broad range of properties, but give too much structure in the correlation functions.

In the early 80'ies, Jorgensen [100] introduced the Transferable Intermolecular Potential Functions (TIPS), model potentials for water and organic liquids. His idea was to find parametric representations for atoms or group of atoms which could be used to construct model potentials for other species, hence the 'transferability'. An important constraint was that the potential should have as few sites as possible, which resulted in the first three site version TIPS for liquid water. This potential yielded reasonable thermodynamic variables, but was not able to reproduce any structure behind the 1. maximum in the oxygen-oxygen correlation function.

The Simple Point Charge (SPC) model of Berendsen and coworkers [99] had a better adjustment of parameters than the original TIPS, and gave energy in closer accordance with experimental values, and also a small second peak in OO radial distribution function.

To get the 2. maximum in OO radial distribution function, Jorgensen had to include a fourth site representing the charge on oxygen located at the bisector of the HOH angle towards the hydrogen atoms. This site takes into consideration the increased electron density at the OH-bonds, and is an exception to the usual omission of bond densities (see page 78) with rigid and non-polarizable models. The four site model TIP2 [101] was later improved, and the final version TIP4P [26] is a frequently used site-site model. The effective dipolemoment is 2.18D [104]. It gives excellent energetic results at room temperature, and a predicted structure in agreement with experiments, though the height of 1st maximum is too large [105]. The least satisfactory values are with the the self-diffusion coefficient which comes out ~50% too high, and the dielectric constant ~30% too low [106]. This deviation from experimental values is also found with the SPC model, which has been refined (SPC/E) [107] to yield values in closer accord with experiments. The above models also have in common that the hydrogen

atoms have no Lennard-Jones interactions. For a comparison of various TIP models, the ST2 and SPC models, see [26].

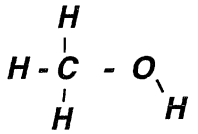
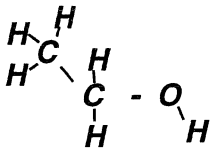
The TIP4P parameters for geometry and potential are summarized in Tables B.1 and B.2.

4.4 Models for methanol and ethanol

4.4.1 Physical properties

The geometrical description of methanol and ethanol monomers in the gas phase is summarized in Table 4.1 below.

Table 4.1
Gas phase geometry [95] for methanol and ethanol.

		Angles	Bond lengths
methanol		<COH = 108.53°	l_{CO} = 1.4246Å l_{OH} = 0.9451Å
ethanol		<COH = 105° <CCO = 107.8°	l_{OH} = 0.971Å l_{CO} = 1.431Å l_{CC} = 1.512Å

The dipole moments of methanol and ethanol are 2.87D and 1.66D respectively [96]. (However Gray and Gubbins [90] quote the values 1.70D and 1.44D, while Handbook [78] report 1.70D and 1.69D). Their polarizabilities are 3.3Å^3 and 5.11Å^3.

The polarity resides mainly with the hydroxyl group. The CH_3 group of methanol will probably have a lack of electrons due to the neighbouring oxygen, but the CH_3 group in ethanol is probably too far from the oxygen to have a fractional charge of any sig-

nificance. So ethanol is the simplest alcohol that have both a polar end and a nonpolar end.

The hydroxyl groups permits hydrogen bonding. Looking at the molecules one would expect one proton donor bond, and two proton acceptor bonds with the oxygen lone pairs. However, in the majority of situations they bond to only two neighbours, and the reason to this is most likely to be found with the physical hindrance from the methyl and ethyl groups.

The most interesting thermodynamical features of these alcohols, are related to their aqueous mixtures. A significant volume contraction upon addition of small amounts to water is observed. Also followed by a minimum in free energy, and a negative excess enthalpy. See Franks and Ives, 1966 [5] and Franks and Desnoyers, 1985 [108] for summaries.

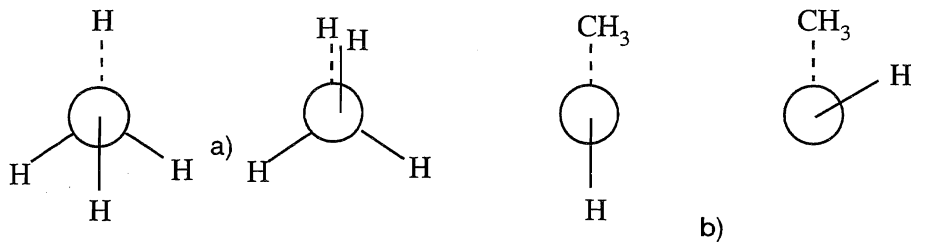


Figure 4.3 Methanol a) and ethanol b) viewed along the C-O bond. Carbon beneath paper plane. Staggered and eclipsed conformations of methyl group in methanol. Anti (or trans) and gauche conformers of ethanol.

Both methanol and ethanol have different stable conformers. In methanol the CH_3 group can rotate as shown in Figure 4.3 a), and the hydrogen atoms can be found in either a *trans* or a *cis* conformation. Ethanol can also have an internal rotation of the methyl group around the C-O bond, and the CH_3-group can be found either far from (*trans*) or close to (*gauche*) the hydroxyl group at an angle of 60°. (See Figure 4.3 b)) The relative amount of each conformer in the liquid phase is estimated [27] to be about 50-50, with the *trans* conformation slightly more preferred. This result comes however from MC-simulations of the pure liquid at room temperature.

4.4.2 Models for methanol-methanol and ethanol-ethanol interactions.

Of the rigid and non-polarizable models for methanol and ethanol, Jorgensens TIPS [100, 109, 110] and OPLS [27, 111] models have been dominant with liquid simula-

tions. But also the methanol models H1 and H2 of Haughney *et al.* [112] have been applied.

The underlying philosophy of the OPLS (Optimized Potentials for Liquid Simulations) are much the same as for the TIP4P. The CH_3 and CH_2 groups are each regarded as one site, and the carbon-hydrogens are not included separately [109]. This imply that the model can not recognize the different conformers of methanol, and of ethanol only the rotation of CH_3 around the C-O bond is distinguishable. Parameters for the CH_3 and CH_2 sites are as for alkyl group values for hydrocarbons [111]. The OPLS parameters for geometry and potential are summarized in Tables B.1 and B.2.

The OPLS for ethanol is parametrized from simulations where both *trans* and *gauche* conformers are included. We have chosen to simulate only the *trans* conformer. A result of this is that the molecule does not have the opportunity to 'flip over' to the other conformation and retain the hydroxyl-hydrogen position. Instead the whole molecule must be rotated, and free space must be found for the hydroxyl-hydrogen. It is possible that we will have some unfavourable configurations relative to a simulation with both conformers included.

4.5 Modelling interactions between unlike molecules

The potentials described above are adjusted for interactions between identical molecules. As emphasized above, the surroundings of the respective molecules are taken into the parameters of the effective potentials. But a single ethanol molecule alone in a container of water, will feel a different influence from the surroundings than in a container of ethanol. Nevertheless the ethanol-water interaction is modelled with a combinations of their parameters as if the polarizability was unaffected by the presence of the other component. This can for instance have effects upon the calculated mixture volumes. The volumes of model mixtures can not be expected to be equal to the stable volumes of the real mixtures when the effective potentials of the pure liquid models are transferred to the mixtures without modifications to reflect the different surroundings.

A second simplification is the use of Lorentz-Berthelot combining rules [122, 123] for Lennard-Jones interactions between unlike molecules.

$$\sigma_{i_a j_b} = \tfrac{1}{2}\left(\sigma_{i_a} + \sigma_{j_b}\right) \quad \text{and} \quad \epsilon_{i_a j_b} = \left(\epsilon_{i_a}\epsilon_{j_b}\right)^{\frac{1}{2}} \tag{4.2}$$

Tanaka and Gubbins, 1992 [113] have investigated the effects of changing the com-

bining rules upon excess enthalpy of mixing for mixtures of TIP4P-water and OPLS-methanol. They find that if the energy parameters for the Lennard-Jones interactions between unlike molecules are multiplied by a factor of 1.1, the results are in better agreement with experiment. The effects of modified combination rules on other properties are not examined so far.

4.6 Flexible and polarizable models

Along with the increase in computer power, there has also been a growth of more complicated, ie. computationally expensive, potentials introducing flexibility and polarizability. Flexibility is introduced by addition of an intramolecular potential to allow for intramolecular motion and distortion by neighbouring molecules. The results from the simulations with flexible models are not conclusive. Guàrdia and Padró, 1990 [114] find that molecular motion is enhanced when a flexible model of water is applied. Wallquist and Teleman, 1991 [115] find a reduction in self-diffusion for water with introduction of flexibility. Wallquist and Teleman find however the structure and potential energy to be relatively unaffected by the introduction of flexibility. Both three site (Pálinkás *et al.*, 1987 [116]) and six site (Hawlicka *et al.*, 1989 [117]) flexible model of methanol have been proposed. The results are in good agreement with experiments, and the increase from three sites to six sites does not seem to make large differences in the calculated quantities, except for the self-diffusion. Results from simulations of ionic aqueous solutions with a flexible water model does not clarify the role of flexibility [93].

Recently the inclusion of molecular polarizability has been a subject of great interest. Allowing for the distortion of electron clouds instead of simply using an enhanced dipole moment as is done in non-polarizable models, requires a smaller timestep. The potentials are in general quantum mechanically based. The greatest differences seem to be within calculation of dynamical properties, and structural and energetic properties are well reproduced. Wallquist *et al.* 1990 [118] find a lower self-diffusion with a polarizable model for water than experimental. Wallquist and Berne 1993 [119] compare two rigid models for water, one of the pairwise additive form, and one polarizable. They find overall agreement with experiment with both models, but the polarizable model exhibit a faster dynamics. Also Caldwell and Coldmann, 1995 [120] find self diffusion constants from polarizable models for water and methanol to be slightly larger than experiment.

van Belle *et al.* 1992 [121], compare SPC and a polarizable model derived from the SPC model (PSPC). They find no significant differences with thermodynamic and static structure, but a substantially lower self-diffusion (agreeing with experiment) and a faster reorientation than with the non-polarizable model.

The addition of polarizability will be even more important in mixtures, as the rigid non-polarizable potentials are parametrized for the pure species and the resulting effective potential will not represent the surroundings in a mixture correct.

Chapter 5
Mixtures of water and methanol - results and analysis

The main purpose of this part of the work is to verify the method and our algorithm by comparison to systems where both experimental and simulated data exists, preferably for the same model potentials. The focus of previous work has mainly been on excess quantities and hydrogen-bonding structure, and for bulk water and bulk methanol, there exist a wealth of experimental and simulation data. For mixtures there are less experimental data, but still a number of computational studies. These are however with few exceptions Monte Carlo simulations of dilute mixtures.

To judge whether any discrepancies arise from algorithm or method, we have performed both NVE and NVT simulations for all mixtures.

Below we present and discuss our results for some thermodynamical properties, all the site–site correlation functions (Section 5.4), and the self-diffusion coefficients from simulation of mixtures of water and methanol, and of the pure liquids. A separate section (Section 5.3) deals with the verification of equilibrium, and the discussion of whether the results from the NVT simulations are within the premises of the canonical ensemble. The chapter concludes with a summary of our findings.

Some of the results are presented in Appendix C.

5.1 Simulation conditions

Pure water, pure methanol, and binary mixtures of the two components with methanol mole fractions 0.25, 0.5 and 0.75 are simulated with NVE and NVT molecular dy-

namics at or near 25°C. We have selected this temperature because comparable results in literature are most frequent at this temperature. A description of the technical details of the simulations was given in Chapter 3. In Table B.3 we summarize the most important parameters and conditions.

All simulations start from a crystal lattice. This lattice melts quickly (see Figure 5.2, page 95) due to the high temperature. During equilibration, the temperature is forced to stay at 298.15K through velocity scaling each timestep. After 25000 steps the velocity scaling is turned off, and the simulation proceeds either with multiple Nosé-Hoover thermostatting, or unthermostatted as an isolated system. The production phases of both the NVT and the NVE simulations then have identical starting configurations. We emphasize that the only difference between our NVE and NVT algorithms is that in the NVE simulation the thermostats are never switched on. Any differences in properties for the two kinds of simulations must therefore be due to different propagation in phase space.

As molecular models we use the TIP4P model [26] for water and the OPLS model [27] for methanol. They are expressed as a Lennard–Jones part plus a Coulomb term, see Section 4.2,

$$U_{i_a j_b} = 4\pi\epsilon_{i_a j_b}\left[\left(\frac{\sigma_{i_a j_b}}{r_{i_a j_b}}\right)^{12} - \left(\frac{\sigma_{i_a j_b}}{r_{i_a j_b}}\right)^{6}\right] - \frac{1}{4\pi\epsilon_0}\frac{q_{i_a} q_{j_b}}{r_{i_a j_b}} \tag{4.1}$$

where i and j denotes different molecules, and a and b are sites on the molecules i and j respectively. In Table B.1 we list the geometry and physical data for water and methanol, and in Table B.2 we list the potential parameters. All Lennard-Jones interactions between unlike sites a and b on two molecules i and j are calculated with Lorentz-Berthelot combining rules [122], Equation (4.2) page 83. Jorgensen and coworkers have traditionally described the Lennard–Jones terms with A and C parameters containing both information of the potential well and the repulsive part of the potential. They also have used a geometric mean for interactions between different sites. Their parameters can however be expressed as σ and ε [27]. The OPLS parameters are reported in this form, and so we have chosen to express also the TIP4P model in terms of σ and ε. The application of Lorentz-Berthelot combination rules then yields slightly different potentials than the original potentials; particularly the repulsive distance between unlike sites is slightly larger with Lorentz-Berthelot combining rules than with a geometric mean. This will not affect pure water, since the water molecule has only one Lennard-Jones site.

The relaxation parameters will be different for each mixture, since we use a weighted average of the pure species values listed in Table 3.1, page 58.

5.2 Thermodynamical properties

Table 5.1 present calculated thermodynamical data from both the NVE and NVT simulations. The averages are coarse grain averages calculated during program execution. The average of 300 consecutive points enters as one contribution to the total simulation average, and the standard deviations are then calculated from the coarse grain averages. See reference [28] and Subsection 3.5.1.

Temperature: From Table 5.1 we see that the NVT-simulations preserve average temperature within 0.2K of preset value (although systematically too low). But also the mixture NVE-simulations are close (within ~2K) to predefined value. This must be due to the long equilibration period, where the temperature is forced to stay at preset value through velocity scaling, and to the short timestep.

The pure liquid NVE-simulations give the poorest agreement with temperature, particularly the pure water simulation, which give a temperature of 11K higher than the preset value. One contribution to this disagreement is seen within the first few hundred steps of the microcanonical production phase where average temperature increases by ~7K, see Figure C.1. This is followed by a slow drift of another ~6K through the rest of the production period. The jump/drift seem to be slightly more pronounced in the rotational temperature, Figure C.2. The same jump - also of ~7K - is seen for the pure methanol simulation, but the temperatures show no drift during the production period. This is consistent with the results for energy conservation, Figure C.5 and Table 5.4. The NVE simulations of pure water show a small positive drift in total energy, while the drift in total energy for pure methanol is practically zero.

It is also of interest to verify if the energy is equally distributed between the two components and their respective degrees of freedom. Tables C.1 and C.2 show translational and rotational temperatures of the two components. We see that the NVE simulations equilibrate the different degrees of freedom within 1.6K of mixture temperature, while the NVT-simulations show a difference of 1.1K at most. The majority of the NVT simulations, however, show no significant differences. We then conclude that it is appropriate to regard the systems to be in internal thermal equilibrium. The standard deviations give an indication of the magnitude of the temperature fluctuations. They are seen to decrease as $N^{1/2}$ with increasing number of particles for both kind of simulations. The fluctuations from NVT simulations are slightly higher than with the NVE simulations.

Table 5.1

Mixture temperatures T_{mix}, pressures P, and potential energies E_p from simulations. Numbers in italics are standard deviation 1σ. Experimental densities from [124], except for density for x_m=0.75 which is extrapolated to 25°C from values at 15°C and 20°C from Perry [125]. *Jorg* is Monte Carlo simulations of TIP4P-water [26] and OPLS-methanol [27], and *Ferr* is molecular dynamics simulations of Ferrario *et al.* [126] for mixtures of TIP4P-water and H1-methanol.
exp is calculated from from experimental enthalpy of vaporization, see [26, 27].

x_m		ρ [g/cm³][1)]	T_{mix} [K]	P [MPa]	$-E_p$ [J/g]	$-E_p$ [kJ/mol]
0.0	*NVE*	0.997	309.01 *(0.27)*	231.8 *(2.4)*	2271.58 *(0.39)*	40.93
	NVT	0.997	297.93 *(0.07)*	226.7 *(2.3)*	2305.08 *(0.81)*	41.54
	exp	0.997				41.51
	Jorg	0.999	298			42.13
0.25	*NVE*	0.937	300.38 *(0.26)*	85.2 *(2.0)*	1903.57 *(0.30)*	40.96
	NVT	0.937	297.95 *(0.08)*	93.1 *(1.9)*	1906.57 *(0.65)*	41.03
	exp	0.937				
	Ferr	0.938	298.2			41.1
0.50	*NVE*	0.882	296.72 *(0.25)*	70.2 *(1.7)*	1589.93 *(0.24)*	39.76
	NVT	0.882	297.98 *(0.08)*	57.4 *(1.5)*	1588.10 *(0.63)*	39.72
	exp	0.882	298			
	Ferr	0.882	298.3			39.8
0.75	*NVE*	0.833	299.80 *(0.25)*	55.3 *(1.3)*	1334.57 *(0.22)*	38.05
	NVT	0.833	298.01 *(0.08)*	52.8 *(1.3)*	1337.48 *(0.48)*	38.13
	exp	0.833	298			
	Ferr	0.831	298.3			37.8
1.00	*NVE*	0.786	305.59 *(0.26)*	50.3 *(1.1)*	1117.83 *(0.20)*	35.77
	NVT	0.786	298.05 *(0.09)*	55.1 *(1.1)*	1127.43 *(0.41)*	36.08
	exp	0.786				35.47
	Jorg	0.759	298			35.94
	Ferr	0.786	298.2			35.1

Pressure: The pressures are stable and positive for all simulations, indicating that the systems are in mechanical equilibrium. Figure 5.1 show a typical progress of pressure. There is agreement between the NVE and the NVT simulations. The specified thermodynamic states corresponds to atmospheric pressures, so our calculated pressures are too high, defining the mixtures as compressed liquids. Particularly the pure water simulations yield very high pressures.

The pressures for the mixtures and pure methanol are in the same range as the results of Ferrario *et al.* [126], who use TIP4P for water and H1 [127] for methanol. They report values of 60 - 90 MPa from NVT simulations of mixtures with methanol mole

fractions 0.25, 0.50, 0.75, and 1.0. Their pure water simulation is however in the NPT ensemble, and therefore yield a far lower pressure than our. Fincham [70] has performed simulations of SPC water with different algorithms and timesteps. His calculated pressures are in the range 36 - 50 MPa for thermostatted simulations.

The elevated pressures are easily explained for pure methanol, where we see that model density is more than 3% lower than experimental density (Table 5.1), which is the density we use. We are thus not alarmed by these values since small density variations in a condensed state will give large pressure variations [102]. A disagreement between model and experimental densities might also explain the pressures for the mixtures. Using TIP4P/OPLS in a NPT Monte Carlo simulation, Freitas [124] found mixture densities at 298K and 1atm pressure to be lower than experimental densities. Remember that potential parameters for the mixtures are found from Lorentz-Berthelot combining rules, and not from an optimization procedure, and also that effects of higher order interactions for *mixtures* are not included in the effective potentials, see Sections 4.2 and 4.5.

But the above explanation does not hold for the pure water results, where model and experimental densities are very close. The source might be found in the calculational procedure. The pressures are calculated from the molecular virial (see Subsection 3.4.1 and [15]). The virial derivation is not valid for systems with periodic boundaries because there is no external field to preserve the N-body shape [28]. To have a consistent expression, the pressure can be derived from the momentum flux. The numerical averages should nevertheless be valid, except that the statistical precision is lower. See Erpenbeck and Wood [129] for a thorough analysis.

Note also that the long range electrostatic contribution to the pressure is omitted due to a large increase in computing time. Only the long range Lennard-Jones contribution, Equation (3.30), is added. This amounts to −20MPa for pure water and is decreasing from −60MPa to −15MPa with increasing methanol content. We believe that inclusion of the electrostatic contribution beyond cutoff would stabilize the sample further by lowering the pressure. This applies of course to all compositions, but the effect of not including this contribution will probably be more pronounced for pure water than for pure methanol, as the former has a larger dipole moment. An estimate of this contribution can easily be made by calculating the k-space contribution for the last few hundred steps of a simulation. Our code showed however up to be extremely inefficient for this purpose.

The high pressures will not affect any of the calculated data, as the pressures are not used in further calculations. (Effects arising from disagreement between model and experimental densities will of course be present also in the other quantities.)

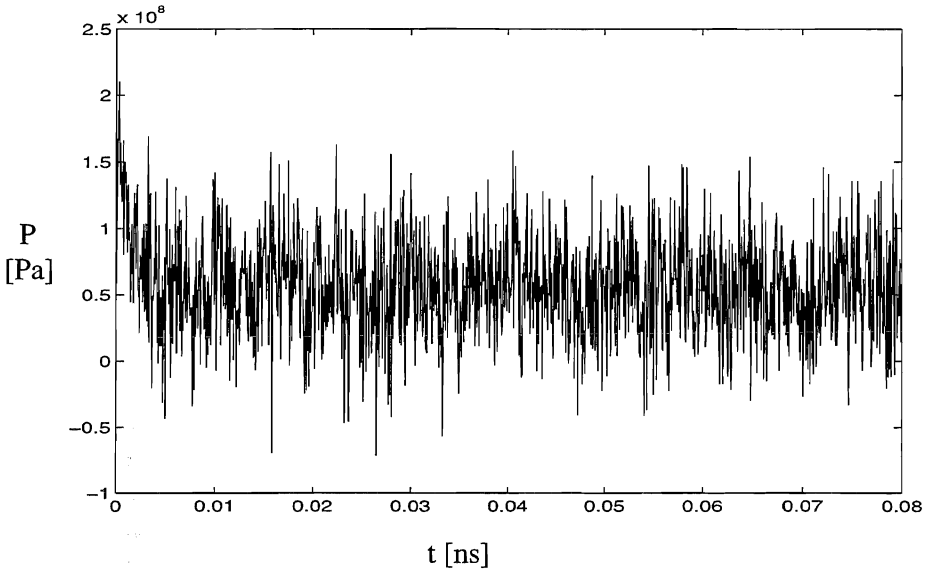


Figure 5.1 Instantaneous pressure, *P*, in a NVT simulation of pure methanol.

Enthalpy: The connection between enthalpy of vaporization and configurational energy is readily found from the definition of enthalpy

$$\begin{aligned} \Delta H^{vap} &= \Delta U^{vap} + \Delta(PV)^{vap} \\ &= \left(E_p(g) + E_{kin}(g) + PV(g)\right) - \left(E_p(l) + E_{kin}(l) + PV(l)\right) \end{aligned} \tag{5.1}$$

where E_p is configurational energy and E_{kin} consists of translational and rotational energies. *(g)* and *(l)* is saturated vapour and saturated liquid. *P* is pressure and *V* is volume. Then, making the following assumptions,

- the pressure is low and the liquid state can be approximated by saturated liquid at the same temperature, implying $E_p(l) \sim E_p(calc)$ and $PV(l)$ is small
- the saturated vapour can be regarded as an ideal gas, $PV(g) = RT$ and $E_p(g) = E_p(id) = 0$ (reference state)
- total kinetic energy is equal for saturated vapour and saturated liquid,

an estimate of ΔH^{vap} at 298K can be found from [27,131]

$$\Delta H^{vap}(298K) = -E_p + RT \tag{5.2}$$

where E_p is calculated configurational energy (Table 5.1) and R is the gas constant. For more accurate values, the enthalpy difference between saturated vapor and ideal gas must be subtracted. For pure methanol this amounts to 0.54 J/mol [27].

For NVT simulations of pure water and pure methanol Equation (5.2) gives respectively 44.02kJ/mol and 38.56kJ/mol, to be compared (Table 5.2) to experimental values of 44.04kJ/kg [96] and 37.43kJ/mol [96]. The calculated values agree very well with experiments considering the above simplifications, but we want to stress that they are only valid under the assumption that the calculated high pressures are results of the calculational procedure and not actual properties of the (model) liquid state.

Table 5.2

Calculated enthalpy of vaporization, ΔH^{vap} at temperature 298K compared to experimental results [96] and Monte Carlo simulations of TIP4P-water [26] and OPLS-methanol [27]. x_m is methanol mole fractions.

	x_m=0.0	x_m=0.25	x_m=0.50	x_m=0.75	x_m=1.0
$\Delta H^{vap}_{(calc)}$	44.02	43.51	42.2	40.61	38.56
$\Delta H^{vap}_{(exp)}$	44.04				37.43
Jorgensen	44.60				37.87

The excess enthalpies for the mixtures can further be found from [130]

$$\Delta H^E = \Delta H_{mix} - x_m \Delta H_m - (1 - x_m)\Delta H_w \tag{5.3}$$

where the enthalpies can be replaced with configurational energies E_p if the assumption of small PV-terms is valid. The excess *potential energies* of mixtures are shown in Table 5.3 along with experimental values [132, 133] for excess *enthalpies*. The experimental enthalpy values are interpolated to 298.15K, and to the proper composition. Our NVT calculations yield a minimum at 0.5, but figure C.3 suggests that a maximum deviation occurs in the range 0.25 - 0.50. An intermediate simulation

would have been very interesting, as experimental results [132] show a minimum in H^E near x_m=0.3. Also shown is Freitas' [124] results from Monte Carlo simulations of TIP4P/OPLS.

Table 5.3

Calculated excess potential energies E_p for the mixtures compared to experimental values for excess enthalpies H^E of Westmeier [132], of Lama and Lu [133] (*), and Monte Carlo values for TIP4P/OPLS from [124]. x_m is methanol mole fraction. All values at 298K.

	x_m=0.25	x_m=0.50	x_m=0.75
E_p(calc), [kJ/mol]	-0.85	-0.91	-0.68
H^E(exp), [kJ/mol]	-0.85	-0.79	-0.55
H^E(exp)*, [kJ/mol]	-0.85	-0.79	-0.58
H^E(MC), [kJ/mol]	-0.6	-0.4	-0.5

Configurational energy: From the simulations with x_m= 0.5 and 0.75, where temperatures are almost equal in the NVT and NVE simulations, we see from Table 5.1 that the resulting potential energies agree well. We also note that the standard deviations of the NVT results are twice the values from the NVE simulations. Segment averages are less stable with the NVT algorithm than with the NVE algorithm.

Configurational energies are not measured in experiments, but one can assign values from enthalpies of vaporization by inverting Equation (5.1). A commonly quoted value for pure water at 298K is 41.51kJ/mol [26], while for pure methanol we can use 35.47kJ/mol (calculated from data in [27]). Our pure water configurational energies, E_p, equals the 'experimental' value, while our pure methanol value is slightly higher. This is of course consistent with the results presented in Table 5.2 above.

It is also natural to compare to Jorgensen and coworkers [26, 27] Monte Carlo values. From Table 5.1 we find that our configurational energy for pure TIP4P-water is slightly lower than calculated by reference 27, and our configurational energy for OPLS-methanol is nearly equal. Note that no correction to computed quantities beyond cutoff is included in [26], while only Lennard-Jones correction is included in [27]. We apply an Ewald sum for the long range electrostatic forces, as discussed in

Subsection 3.4.3. For pure water we find a Lennard-Jones correction beyond cutoff of -21 J/g, or -0.38 kJ/mol. This makes however the discrepancy between our results and the reported results of [26] even larger.

We are only aware of the work of Freitas [124] on methanol–water mixtures using TIP4P and OPLS potential. He does not report potential energies, only excess enthalpies as referred in Table 5.3 above. Ferrario *et al.* [126] have used a combination of TIP4P and H1[127]. The H1 model for methanol is not unlike OPLS for methanol, both being rigid models with the same molecular geometry. Potential parameters are very close, with Lennard-Jones parameters of $\sigma(CH_3)$=3.861Å, $\sigma(O)$=3.083Å, $\varepsilon/k_B(CH_3)$=91.15K, $\varepsilon/k_B(O)$=87.94K, and partial charges 0.297 on methyl site and -0.728 on oxygen site. The close resemblance of the models explains why our configurational energies are very close to their values (se Table 5.1).

5.3 Theoretical analysis

We have also analysed a few variables that are of importance to the judgment of code performance and reliability of results.

We focus upon

- verification of liquid state equilibrium
- conservation of total energy and total linear momentum
- monitoring the heat bath parameters
- canonical sampling with the NVT simulation

The background is described in Subsections 3.6.2 and 3.6.3. Application of other tests are described in [28, 15].

Translational order is measured with the order parameter ρ. All mixtures for both NVE and NVT simulations become disordered within the first 1000 steps. An example is shown in Figure 5.2 below. The quick melting compared to the long equilibration time for the energy, which at least takes 5000—10000 steps (Figure C.5), show that the melting alone is a poor indicator of thermodynamic equilibrium for this system. The order parameter oscillates about zero with amplitude less than ±0.06, which agrees nicely with the expected amplitude of order $N^{-1/2}$ [64].

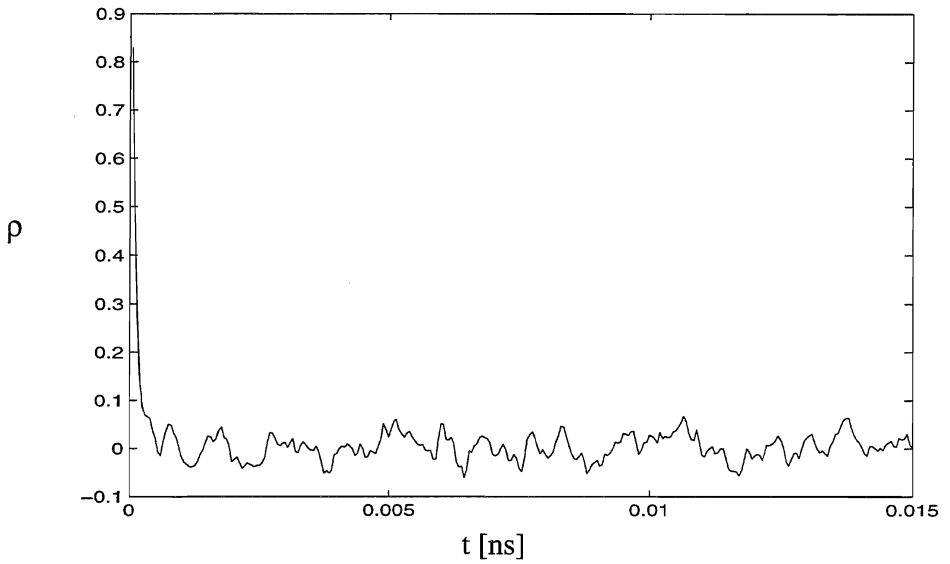


Figure 5.2 Translational order parameter ρ (dimensionless) for the first 30000 steps of a NVE simulation of methanol and water, x_m=0.75. Every 100th steps are plotted.

Drift in total energy: In an NVT simulation, the simulation box and the heat reservoirs together constitute an isolated system. The total energy of this extended system must then be constant, just like the system (box) energy for an NVE simulation. When added, fluctuations in all contributions to the energy must cancel exactly, regardless of simulation method (see Subsection 2.6.1). Due to the numerical integration, a perfect cancellation is not possible to achieve. The calculated energy is thus seen to have small fluctuations. Also a small drift during the course of the simulation is often observed.

All our simulations do show a positive drift. Values of drift in total energy are estimated from the difference between minimum value within the first ~five thousand steps of production period and maximum of the last ~five thousand steps. This should provide us with a rough estimate of maximum drift, see Table 5.4 below. The estimate is sufficient for our purposes, since we do not aim at making conclusive statements about algorithm performance. The sudden disappearance of the fluctuations in switching from velocity scaling to NVE or NVT simulations is demonstrated in Figures C.4 and C.5, respectively. On the scale of the plots, the fluctuations in the production phase are not visible. Figure C.6 show a close-up of the x_m=0.5 mixture total

energy, where both the drift and the remaining fluctuations are illustrated.

Table 5.4
Energy drift pr. step during production phase of simulation, ΔE and ΔE_{ext} for the NVE simulations and NVT simulations respectively. Total number of steps is 135000. The drift is only a rough estimate of difference between maximum and minimum values.

x_m	Δ E [J/(g step)] NVE	ΔE_{ext} [J/(g step)] NVT
0.0	$1.7 \cdot 10^{-4}$	$9.3 \cdot 10^{-5}$
0.25	$1.9 \cdot 10^{-5}$	$4.1 \cdot 10^{-5}$
0.50	$1.7 \cdot 10^{-5}$	$7.4 \cdot 10^{-6}$
0.75	$1.0 \cdot 10^{-5}$	$5.9 \cdot 10^{-6}$
1.00	$6.0 \cdot 10^{-6}$	$6.0 \cdot 10^{-6}$

From Figures C.4 and C.5, and Table 5.4 we see that the choice between NVE or NVT simulation does not influence significantly upon energy conservation. It seem however clear that conservation improves gradually with increasing methanol content, with pure water having the largest drift. Poorer energy conservation for water might be due to the energy jump at the cut-off [128]: a molecule leaving the cut-off sphere and entering as a minimum image, see Figure 3.3 page 60, will have a different orientation relative to the central molecule. Then the energy will be different, and this energy difference is not accounted for. Since water has the largest fractional charges, the ignorance of this energy will have the largest effect. For a dipolar system the energy jump is estimated to be of the order of $\mu^2/(2L^3)$, where μ is the dipole moment and L is the boxlength [128].

From Figures C.4 and C.5, we note that fluctuations are indeed present during the first 0.0125ns. These are caused by the initial velocity scaling procedure, which of course does not conserve the total energy, since energy is added or removed from/to nowhere. We also note that fluctuations in the equilibration period increase with increasing water content - ie. with decreasing system mass. Looking at these fluctuations, we would recommend the pure water system to be equilibrated for another five or ten thousand steps due to the slightly unstable progression. The mixtures with 25% and 50% methanol reach stability just within the equilibration period, while the two mixtures with 75% and 100% methanol stabilize very quickly.

Figure 5.3 show the contributions to extended system energy for pure water.

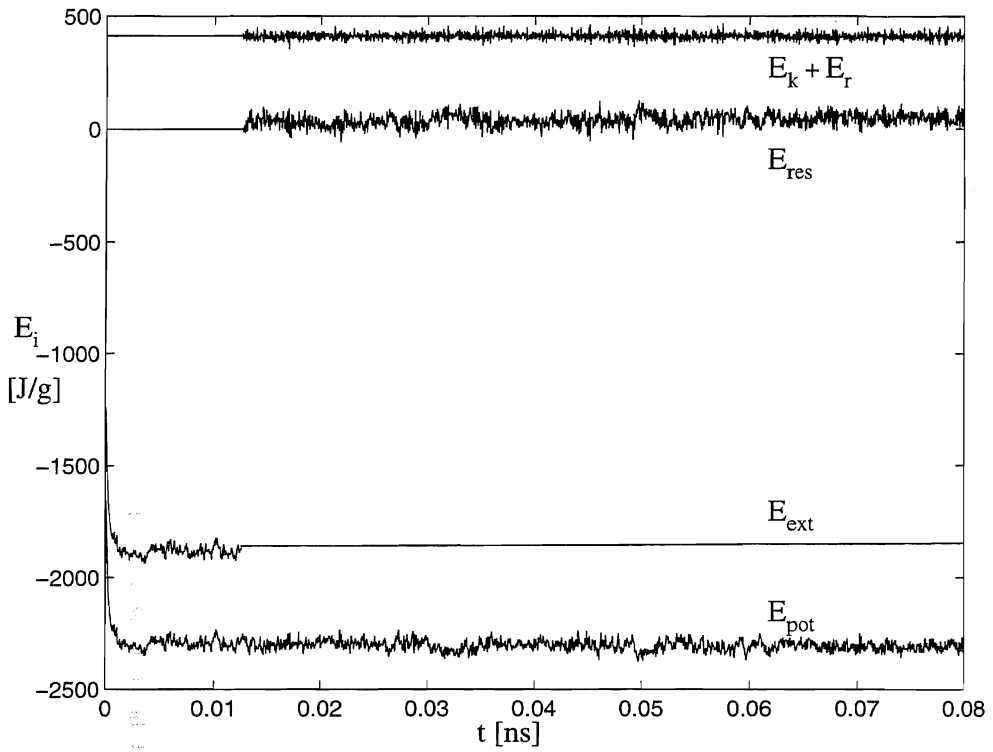


Figure 5.3 Contributions to energy, E_i, for pure water NVT simulation at 298K. Reservoir energy is displaced from zero by ~31kJ/g at start and increases to ~45kJ/g by the end of simulation, averages over first and last 5000 steps respectively. By comparison, potential energy increases only by ~5kJ/g between the same intervals. Kinetic energies of reservoirs not shown as they would appear as a horizontal line at zero energy, but is included in extended system energy E_{ext}. E_{res} is sum of potential energies of both reservoirs.

Velocity distribution: Figures C.7 – C.10 show that the initial velocity distribution is maintained during the production period. The calculated distribution gets smoother with increasing number of particles. Only selected NVT results are displayed, as results for the remaining part of the mixtures, and for the NVE simulations, do not reveal any surprises.

Conservation of linear momentum: Upper panel of Figure 5.4 below show a typical progress of linear momentum for an NVT simulation. Figure C.11 show the corresponding NVE result. Further, in Tables C.3 – C.4 the average linear momenta per

mass, p_ζ, for all coordinate directions ζ are tabulated. p_ζ is defined as

$$p_\zeta = \frac{m_w \sum_{i=1}^{k} v_{w,i} + m_m \sum_{j=k+1}^{n} v_{m,j}}{k m_w + (n-k) m_m} \tag{5.4}$$

where m_w is mass of water molecule and m_m is mass of methanol molecule. k is number of water molecules and n is total number of molecules. $v_{w,i}$ and $v_{m,i}$ are molecular linear velocities.

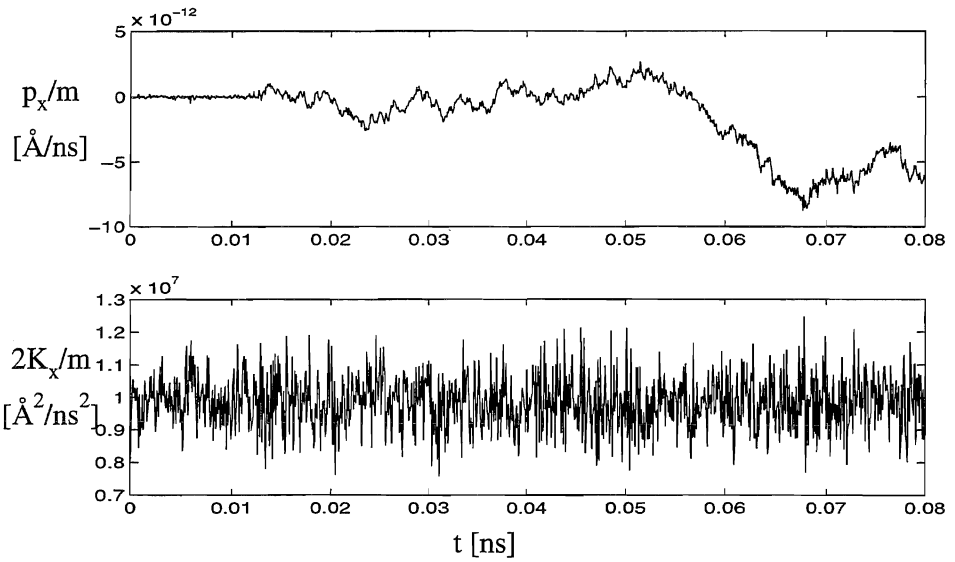


Figure 5.4 X-component of linear momentum pr. mass, p_x/m (top) and translational energy $2K_x/m$ from NVT simulation of an equimolar mixture of water and methanol. m is system mass.

For all mixtures, regardless of ensemble, p_ζ is of the same order of magnitude, and always less than ~10^{-12}m/s. This is to be compared to the component velocity of, say, water of v_{rms} ~400m/s found from the Maxwell distribution at the same temperature. The standard deviations are large, reflecting the drift in p_ζ. The fluctuations are very small. During the 135000 steps of production period, the drift in linear momentum is within a factor of 100 from the initial values. We are in the regime of rounding errors, and we must conclude that the linear momenta are conserved as expected. Note that during the equilibration period where velocity scaling is applied at each step, extra

momentum must also be removed. Linear momentum is not conserved automatically when velocity scaling is applied to mixtures because of different masses of the components (cf. Subsection 3.3.3, page 56).

Conservation of angular momentum: As expected, total angular momentum is not conserved. This is because of the use of periodic boundary conditions, as discussed in Subsection 2.6.3. The effect is clearly seen in Figure C.12, where angular momentum per mass in x-direction, L_x given by

$$L_\zeta = \frac{I_{w,\zeta}\sum_{i=1}^{k}\omega_{w,i} + I_{m,\zeta}\sum_{j=k+1}^{n}\omega_{m,j}}{km_w + (n-k)m_m} \tag{5.5}$$

is plotted for the entire simulation. $I_{w,\zeta}$ is ζ-component of moment of inertia of water and $\omega_{w,i}$ is angular velocity of water molecule *i*. Subscript *m* denotes methanol. L_ζ is stable around a mean value near zero, but oscillates with large amplitudes. The situation is similar for the NVT simulations (not shown). Averages and standard deviations for all components can be found in Tables C.5 and C.6. The averages can be compared to individual angular momenta of ~10^{-8}m^2/s (pure methanol). (Calculated from $k_BT/I_x=\Sigma\omega_x I\omega_x$, provided the components are mutually independent.)

Equipartition of energy: Tables C.3 and C.4, and Tables C.5 and C.6 show that translational and rotational energy are equally distributed among the three coordinate directions for both the NVE and the NVT simulations. Translational and rotational energy are also almost equal for the same mixture and simulation method. This is already seen from the temperatures, Tables C.1 and C.2. Inconsistencies between values from NVT and NVE simulations for the same mixtures are probably due to different mixture temperatures, as discussed previously. Examples of progress of kinetic energies are found in bottom panels of Figures 5.4 and C.11 – C.12.

Distribution of internal energies: Figure 5.5 below and Figures C.13 - C.16, show probability distributions for instantaneous internal energy from the NVT simulations. The distributions are based upon output data at each 100th step during production period (4 significant figures). For all compositions the energy distributions resemble the shapes of the theoretical, but are systematically too broad, and they do not reach the maximum probability density. There also seem to be a slight skewness in the distributions for the mixtures with 0.25 and 0.50 mole fraction methanol. The NVT simulations overestimate the fluctuations and hence the system does not follow a canonical path. Distributions based upon parts of the simulation reveal no significant differences with the full simulation distribution. We believe that the calculated distributions are results of a slightly too high heat bath parameter Q. This is also in accordance with the findings of Di Tolla and Ronchetti [49]. Too high Q will result in too weak -

or infrequent - temperature control, which means that temperature are not corrected until it has departed too far from the predefined value. At the opposite end of the scale - Q approaching zero - the fluctuations will be suppressed due to instantaneous control.

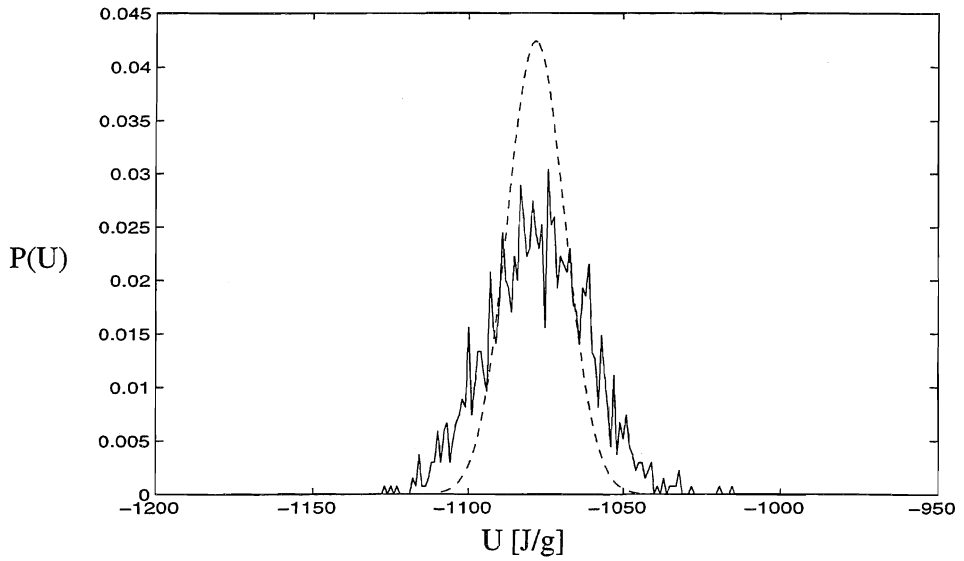


Figure 5.5 Distribution P(U) of internal energy (box energy) in NVT-simulation. Mole fraction methanol is 0.75. Dashed line is canonical distribution at 298.15K.

Moments of kinetic energies: Figs. C.17 - C.21 show 1st to 5th moments of system translational energies, Equation (3.40) page 73, and the moments of system rotational energy are displayed in figs. C.22 - C.26. All moments are divided by the theoretical values for fluctuations in the canonical ensemble, Equation (3.39). The moments are running averages based on instantaneous values of energy sampled every 150th step.

All figures show that the moments approach a stable value, but the approach is slow. Except for the 1st moment, none of the higher moments have reached their expected values by the end of the simulation. Values are typically of the order of 0.01, 10, 10, and 100, for 2nd, 3rd, 4th, and 5th moment respectively. Note however that the effect of using running averages is that any sudden changes are smoothed, and any large starting values are preserved in the averages. The slow approach towards unity might therefore in part be due to the averaging method.

Regarding the moments of reservoir kinetic energies, Figures C.27 - C.28 show that they also seem to stabilize during the last third of the production period. However, for the 2nd - 5th moments, their limiting values seem to be zero (of the order of 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6} respectively), and not unity as expected. The exception is the first moment, which agrees roughly with the theoretical value of 0.3866J/g for the same mixture. The same trend for moments of reservoir kinetic energies to fall below theoretical value is also found by Cho and Joannopoulos, 1992 [47].

As only one degree of freedom is connected with each reservoir, the approach to canonical distribution should necessarily take longer time than with the system kinetic energies [47].

Friction parameter η: The friction parameters η_{trans} and η_{rot} oscillate with large amplitudes about a value close to zero. Figure 5.6 (next page) provides an example for the equimolar mixture of water and methanol. Table 5.5 below show the average values for all simulations.

Table 5.5

Coarse grain averages of translational and rotational friction parameters from the NVT simulations. Standard deviations in parenthesis. x_m is methanol mole fraction.

	x_m=0	x_m=0.25	x_m=0.5	x_m=0.75	x_m=1.0
η_{trans}[ns^{-1}]	-9.8 (17.7)	-6.1 (18.3)	-1.1 (17.8)	-5.0 (18.5)	3.1 (18.7)
η_{rot}[ns^{-1}]	12.1 (14.5)	6.5 (17.8)	2.2 (19.7)	4.2 (20.1)	-2.9 (20.9)

Ideally, the averages shown in Table 5.5 should each amount to zero, since in a state of equilibrium there is no preferred direction of energy transfer. We note that except for pure methanol, the translational degrees of freedom need extra energy – on the average – while the rotational degrees of freedom need to bee calmed down. The friction parameters for water show the largest deviation from zero, while the equimolar mixture parameters are closest to expected value.

We also note that the sum of translational and rotational friction parameters are always closer to zero than the individual contributions.

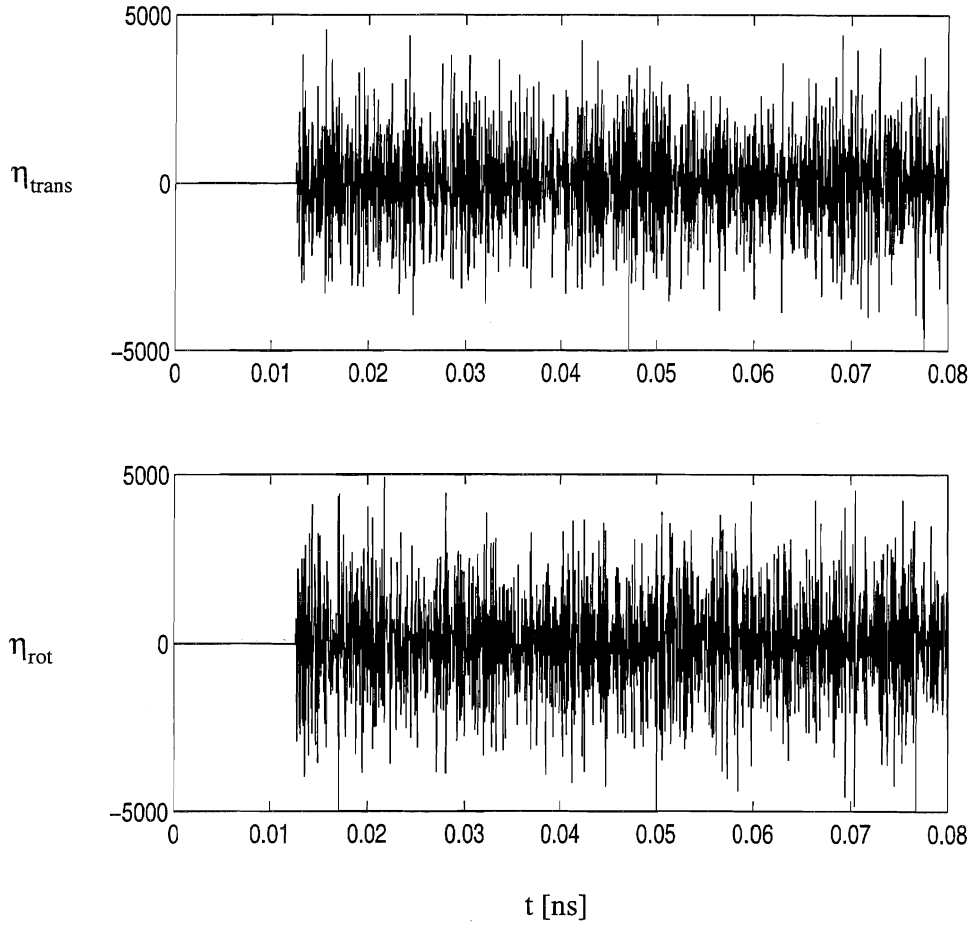


Figure 5.6 Friction parameters η_{trans} and η_{rot} for entire simulation of a equimolar mixture of water and methanol. instantaneous values each 100th step shown.

The sum of the kinetic energies of the reservoirs, see Equation (2.29), should amount to *2/(6N)* of the system kinetic energy. In Table 5.7 the calculated reservoir kinetic (translational plus rotational) energy are compared to theoretical values. We see that we underestimate the reservoir kinetic energy by ~15% for all mixtures.

Reservoir energies and *s*-variables: Figure 5.7 below show the development of the separate *s*-variables during the entire simulation. Figure C.29 show the sum of the *s*-variables. Notice the nearly symmetric ebbing and flowing of the translational and rotational degrees of freedom. This rolling behaviour imply that time scaling changes, and the translation and rotation develops according to two different timescales, which is conceptually somewhat difficult. We can not predict the continuation of the curves for larger times. Even if s_{rot} for the pure water rotational heat bath seem to diverge at the end of the simulation, it might as well have been decreasing again with a continued simulation.

The sum of *s*-parameters is shown to illustrate the overall stability of the two heat baths. We are however not convinced that only one reservoir would have resulted in the same stability.

Table 5.6

Coarse-grain averages of s-parameter for translational and rotational reservoirs. Same values for both water and methanol.

	x_m=0	x_m=0.25	x_m=0.5	x_m=0.75	x_m=1.0
s_{trans}	0.877(0.008)	0.827(0.006)	0.919(0.004)	0.940(0.006)	1.269(0.009)
s_{rot}	1.311(0.014)	1.267(0.009)	1.085(0.005)	1.101(0.007)	0.865(0.007)

Table 5.6 above show average translational and rotational *s*-parameters. We find that the sum of the averages are closer to the expected value of unity than each individual, just as we found for the friction parameter. This is also seen from Figures C.29 and 5.7.

As we have chosen an initial value of unity for both s-variables, the starting energy of each heat bath must be zero. For a system of thermal equilibrium, the reservoirs should act as mediums to transfer energy through, and their potential energies are expected to be constant on the average. We do however see that particularly for the pure water the potential energy departs from its starting value of zero, Table 5.7.

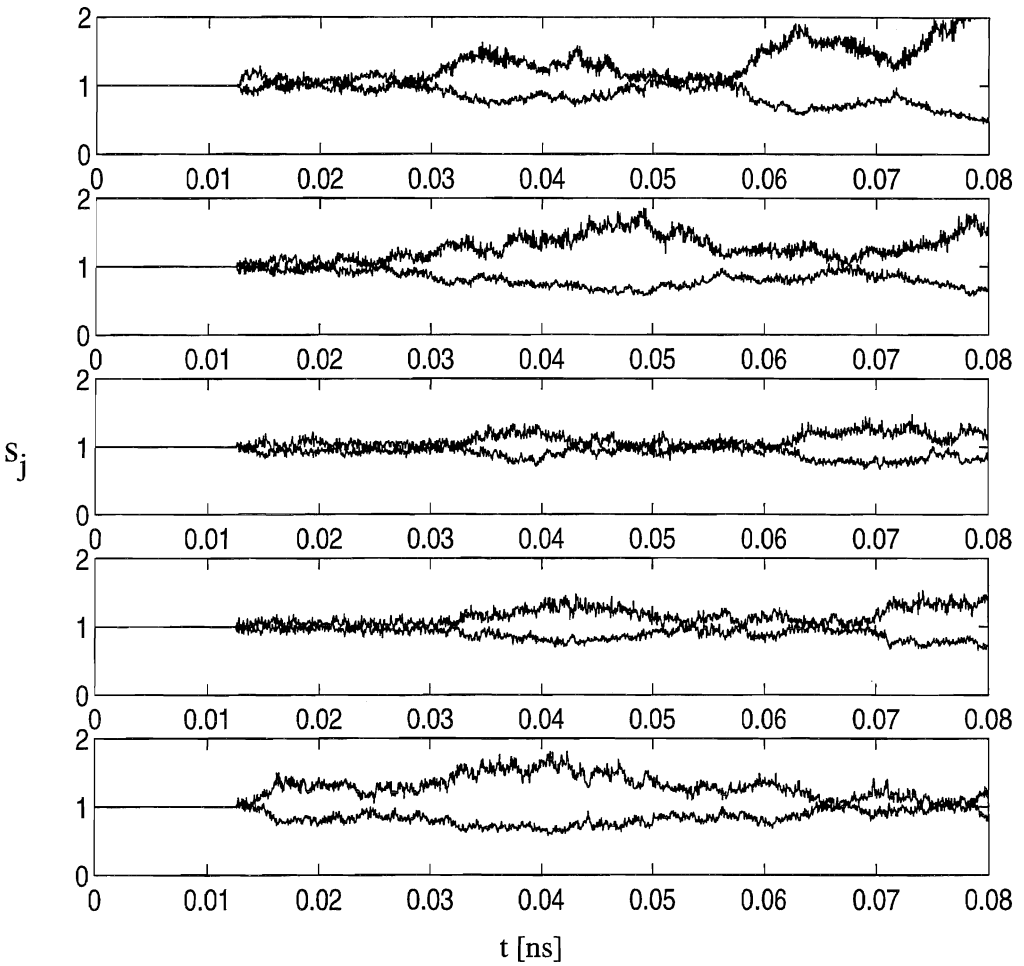


Figure 5.7 Heat bath variables s_{trans} and s_{rot} for all mixtures, ranging from pure water on top to pure methanol in the bottom panel. Lower curve is for translational except for bottom panel. Curves in the two lowest panels cross at 0.055-0.065 and 0.065 0.075 respective. Different amplitude of fluctuations is caused by different heat bath masses, see Table 3.1.

Table 5.7
Average potential and kinetic energy of the heat reservoirs, and average transferred energy pr. step. Standard deviations in parenthesis. Averages are based on output values with 4 significant digits taken each 100th step.

x_m	Reservoir pot. energy [J/g]		Transf. energy pr step [J/g]		Reservoir kin. energy [J/g]	
	average	expected	mean	expected	mean	expected
0	33.0 (29.2)	0	-0.0028 (0.45)	0	0.47 (0.56)	0.41
0.25	6.5 (20.1)	0	-0.0013 (0.35)	0	0.38 (0.45)	0.32
0.50	-3.4 (17.8)	0	-0.0009 (0.30)	0	0.33 (0.38)	0.28
0.75	3.1 (14.8)	0	0.0024 (0.26)	0	0.29 (0.33)	0.25
1.0	12.8 (13.7)	0	0.0007 (0.22)	0	0.25 (0.30)	0.21

5.4 Structural properties of methanol/water mixtures.

Below we present site-site radial correlation functions as defined in Subsection 3.5.3 for the self-correlations in the pure liquids and mixtures, and the cross-correlations in the mixtures. We have also calculated the site-coordination numbers by integrating all correlation functions to their respective first minima. The distance to first minimum is seen to vary only slightly with composition, but we have nevertheless used each individual minimum as upper integration limit. The coordination numbers appear in Tables 5.9—5.11.

Coordination number only provide an upper limit to the number of hydrogen bonds. To perform a hydrogen bond analysis, criteria upon bonding energies or/and bond angles and distances must be satisfied. We have not performed such an analysis, and are therefore prevented from drawing decisive conclusions about the hydrogen–bonding structure.

The calculated positions and heights of 1st and 2nd maxima, and of 1st minimum are given in Tables C.7 - C.9.

We have estimated [15] the standard deviation in the sampled values of oxygen-oxygen correlations between water molecules near 1st maximum and at one point in the ideal gas region. The results are divided by the same normalizing factor as g(r),

and give σ~0.1 units, for all mixtures, Table 5.8, except for the 0.50 mixture where σ = 0.04. If these values are independent (the sampling frequency is 150 steps), we will find 99% of the values within ±0.25 units of the plotted peaks. In the ideal gas region we find a uncertainty of less than ±0.1. A full calculation of variance for all distances and all correlations is impractical because of computing time.

The accuracy in the radial direction is limited by the bin widths of 0.05Å from the sampling procedure.

Table 5.8

Error in g_{oo} near 1.peak (2.80Å) and at long distance (7Å) from pure water simultaion. There is no difference between NVT and NVE simulation in absolute error.

x_m	Abs err 1σ	
	r=2.8Å	r=7.0Å
0.0	0.1	0.04
0.25	0.1	0.04
0.50	0.04	0.008
0.75	0.1	0.02

We show results from the NVT–simulations only, as these simulations yield nearly the same temperatures for all mixtures (see Table 5.1). Differences in correlation functions with the two simulation methods are however minor and relates only to peak heights. Radial positions of maxima/minima are unaltered. Figure 5.8 below show a comparison between $g_{HH}(r)$ from NVE and NVT simulations for the equimolar mixture of methanol and water. This correlation showed the largest difference with simulation method. The difference in peak height is 0.15, which is of the same order as the standard deviation calculated near 1st peak of oxygen-oxygen correlations for water (Table 5.8). Figure C.31 provides another example for cross correlations between the methyl-group and the hydrogen on water. Note that this time the NVE simulation gives the largest peak height.

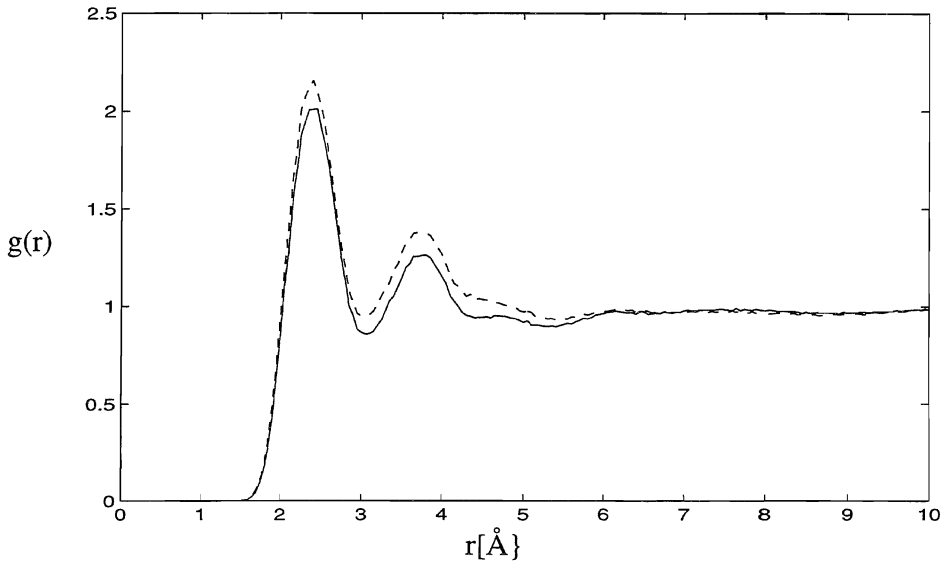


Figure 5.8 Radial self-correlation function $g_{HH}(r)$ for water in a equimolar mixture of water and methanol. NVE-simulation (———) and NVT-simulation (- - - -).

Correlations with electrostatic M-site on water is not shown. This site yields same pair–correlations as the oxygen Lennard-Jones site, except for the expected radial displacement of r_{OM} due to position of auxiliary site, see Figure C.31.

5.4.1 Brief review of present status on structure.

Before we look at the structural data, we will briefly review some of the main features of the local structure of liquid water, methanol, and their mixtures. We restrict the summary to pressures near 1 atm and temperatures near room temperature. For a thorough historical review on structural models, see Franks 1972 [4]. A more recent review on hydrogen-bonding liquids is given by Ladanyi and Skaf, 1993 [134].

The general view, supported by simulations during the last 20 years [97, 135, 136, 137, 138, 139, 140][5], is that bulk water consists of an infinite and random, three dimensional network [8] of hydrogen bonded molecules. The hydrogen–bonds are nearly linear [6]. Locally, the water molecules are arranged in a nearly tetrahedral

[5] This is not at all an exhaustive list!

structure, which is not stable but restructures on a picosecond timescale [142]. Each water molecule is observed to have an average of about 4.5 nearest neighbours [143], which is more than the maximum number of linear hydrogen bonds a water molecule can be engaged in. The fifth neighbour is explained by the existence of a bifurcated H-bond [144], where a proton is simultaneously shared between two oxygen atoms in the network. Not quite obvious from Figure 5.9 below, this allows a closer packing of water molecules. The bifurcated hydrogen-bond is not a hydrogen-bond within the traditional definition, but is more probably to be regarded as a transient state [141]. The presence of these weaker, bifurcated H-bonds is proposed to act as catalysts for the restructuring of the network [142].

Figure 5.9 a) Tetrahedral arrangement of water molecules. Only oxygen atoms shown except for hydrogen bonds of central molecule. Length of linear bond is 1.85Å.
b) Bifurcated hydrogen bond of length 2.3Å in tetrahedral arrangement. After Giguère [144].
Lines joining oxygen atoms does **not** represent bonds.

Recently, from analysis of density data, arguments for a mixture model have been proposed by Vedamuthu *et al.* 1994 [145]. Water is then regarded as a dynamical equilibrium mixture of (several) dense and bulky bonding forms. Curiously, mixture models have been abandoned for nearly 20 years. Neither network nor mixture models seem to be contradicted by experiments. See however Narten and Levy 1969 [146] for a discussion of validation of models from experimental results.

From hydrogen bond analysis, based upon both experiments [9, 10, 12] and simulations [109, 112, 116, 148, 149, 150], a methanol molecule is found to participate on the average in nearly two hydrogen bonds. The fraction of non-bonded molecules are low. From simulational studies of methanol with the OPLS [148] and TIPS [112] models is found that only ~2-4% of the molecules are non-bonded, and that less than

10% of the molecules are engaged in 3 bonds. The average number of hydrogen bonds pr. molecule from these simulations (1.7 - 1.9) is consistent with experiments (1.7-1.8 [12,9]). This leads to believe that pure methanol consist of branched, winding chains of assorted lengths [116, 147], se Figure 5.10 below. Average numbers of 10 [12] and 15.5 [148] chain molecules have been suggested. The two-bonding state is suggested to be relatively stable, and several times longer lived than one-bonding and three-bonding [148]. Structural order between chains are found to be small [12, 112]. In the two-bonding state a molecule has one acceptor bond and one donor bond, the second acceptor is less available due to steric hindrance of the methyl group. The preferred direction of the acceptor bond is between the lone pair directions of the oxygen atom [116].

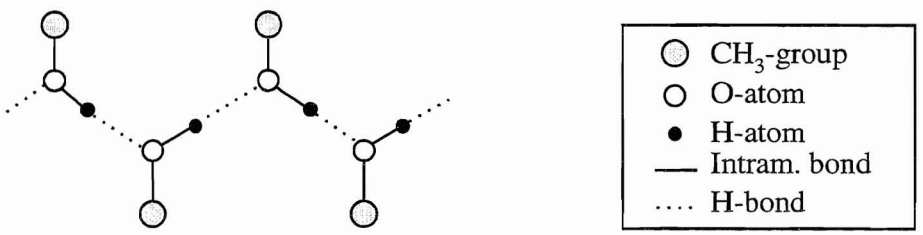


Figure 5.10 2-dimensional view of fragment of hydrogen bonded chain in bulk methanol. After Jorgensen,1981 [109].

Now turning to mixtures of water and methanol. Aqueous solutions of lower alcohol exhibits large negative excess properties [132, 133], with minimum occurring in the mole fraction range 0.1-0.3. A traditional explanation of excess properties in the mixtures has been as a result of entropy loss through enhancement of water structure near a solvated methanol molecule; the mechanism often termed hydrophobic solvation. The main focus has thus been on weak solutions of methanol in water, but also dilute solutions of water in methanol have been investigated. The tools have been simulations (mainly MC) [126, 151, 152, 153, 154, 155, 156, 113] and theory [158]. A few studies over the whole concentration range also exist [124, 131, 113]. Not until the recent neutron diffraction experiments of Soper and Finney 1993 [157], has there been any experimental results for mixture structure available.

Alcohols are interesting because the molecules have both an hydrophobic part (the organic group) and a hydrophilic part (the hydroxyl group). The central questions that have been asked, are

- how is the water-water and alcohol-alcohol interactions affected in mutual solutions
- in which way does water and alcohol interact when one is added to the other

The discussions on mixture structure have been based on results from simulation on model systems[6]. For weak methanol solutions, simulations predict formation of clathrate like water spheres around the hydrophobic end. The observed increase in intensity of pair correlation functions has unisonly been interpreted as promotion of water structure. Soper and Finney [157] do confirm the existence of 'a loose hydrogen bonded cage around the methanol molecule', but find however no evidence of enhanced structure in their neutron diffraction study of 0.1 mole fraction methanol solution. The water molecules are found to be oriented tangential to a sphere circumscribing the methyl group [152, 161].

Vaisman and Berkowitz 1992 [159] also question the traditional interpretation of increased peak heights, and points out that the pair correlation function is not an adequate measure of relative order at different concentrations. Locally, the solvent water molecules try to maintain their pure liquid structure, and so the local concentration decreases slower than overall concentration thus giving raise to an apparently increase in intensity.

Hydrogen–bond statistics from simulation of systems of 128 to 256 molecules [126, 156] show that in water–rich mixtures the number of two–bonded methanol decrease in favour of three–bonded methanol (cross-linked chains), while in methanol–rich mixtures the number of four–bonded water decrease in favour of three–bonded. Both water and methanol have a reduced acceptor character in the methanol–rich solution in comparison with the water rich solution, due to the larger size of the methanol molecules.

6 But see Franks 1966 [5] for a review on ideas and concepts from 'pre-simulation era'.

5.4.2 Water self-correlations

From Figures 5.14 —5.16, we see that the water - water correlations are characterized by

- an increase in height of g(r) with increasing methanol content
- no significant change of the position of 1st maximum upon changes in the composition
- slightly more structure at long range for mole fraction methanol of 0.75

Positions and heights of maxima and minima appear in Table C.7, while coordination numbers are given in Table 5.9.

Oxygen-oxygen correlations:

Figure 5.14 show the oxygen-oxygen correlations. The calculated 1st maximum at 2.8Å agree very well with the neutron diffraction data (2.875Å) of Soper and Phillips [143]. The second maximum at 4.55Å is clearly reproduced. With a methanol mole fraction of 0.75, the 1st minimum is deepened and the 2nd maximum is shifted to a distance of 5.15Å. The 2nd maximum give information of the O—O—O angle, and an increase in distance from 4.55Å to 5.15Å can be interpreted as an increase in this average angle from 107° to 137°. This can be taken as evidence of reduction in the thetrahedral structure of water in methanol-rich solutions, and is consistent with the loss of one acceptor–bond leaving the remaining bond situated in between the two lone–pair directions [126, 156]. We also notice that this 2nd peak is sharper for the mixture with highest methanol concentration, meaning that there is less variation in orientations.

Also a weak 3rd maximum at 6.5-7Å is present. This third maximum gets more pronounced with increasing methanol content, indicating more long range structure.

Integration of the first peak to 1st minimum at 3.5Å (see Table C.7) give a coordination number of 5.1 for pure water, in agreement with experimental value of ~4.5 within 3.3Å [143]. The coordination number decreases almost linearly with increasing number of methanol molecules (Table 5.9).

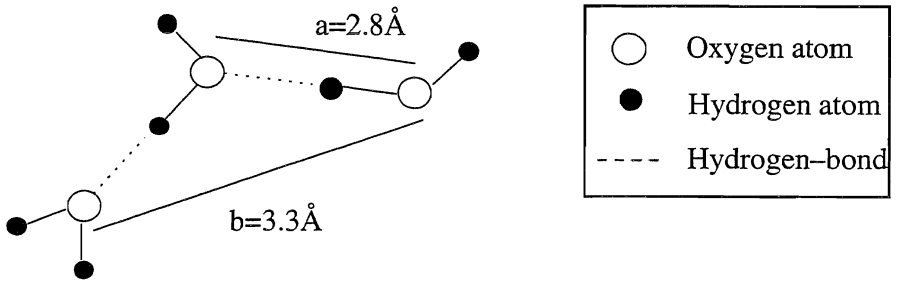


Figure 5.11 Approximate oxygen–oxygen distances between neighbouring molecules in a tetrahedral orientation. *a* and *b* both contribute to 1st peak of Figure 5.14. Figure after [7, 161].

Oxygen-hydrogen correlations:

Figure 5.15 show the two peaks of the oxygen-hydrogen correlation. The first peak at 1.85Å, in perfect agreement with experiment [143], is due to the hydrogen bond with oxygen as a proton acceptor. The second peak at 3.25Å is due to hydrogens of nearest neighbour oxygens, see Figure 5.12. Note that the 2nd peak has approximately the same height and shape for the two highest mole fractions of methanol, indicating that nearest neighbours are significantly affected in methanol–rich solutions. While pure water is flat beyond 2nd maximum, meaning that a molecule has no influence on the orientation of its 2nd nearest neighbours, the 0.75 mixture show a broad 2nd minimum at distance 4-5 Å followed by a shoulder and a 3rd maximum. Certain orientations must be less probable.

Integration of the first peak out to 1st minimum give a coordination number of 2.0 (Table 5.9), decreasing linearly with increasing methanol concentration.

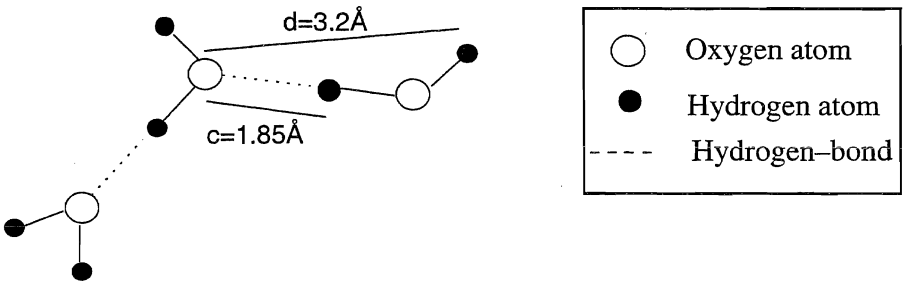


Figure 5.12 Approximate oxygen–hydrogen distances between neighbouring molecules in a tetrahedral orientation. *c* contribute to the 1st peak of Figure 5.15, and *d* to the second peak. Figure after [7, 161].

Hydrogen-hydrogen correlations:

Figure 5.16 show the hydrogen-hydrogen 1st peak at 2.45Å and the 2nd peak at 3.8Å, both peaks referring to hydrogen bonded neighbouring water molecules, see Figure 5.13. As with the oxygen-hydrogen correlations, the intensity of the first peak increases substantially more than the 2nd peak with increasing methanol content. Note the 'irregularity' for the 0.5 mole fraction mixture. Once again the 0.75 mixture is found to develop a deeper 2nd minimum than the rest of the mixtures.

Integration of the first peaks give a coordination number of 5.3 for pure water, in agreement with experiments [143]. Note from Table 5.9, the n_{OO} and n_{HH} follow each other closely; for each water molecule in first coordination shell, there is one oxygen site and one close hydrogen site.

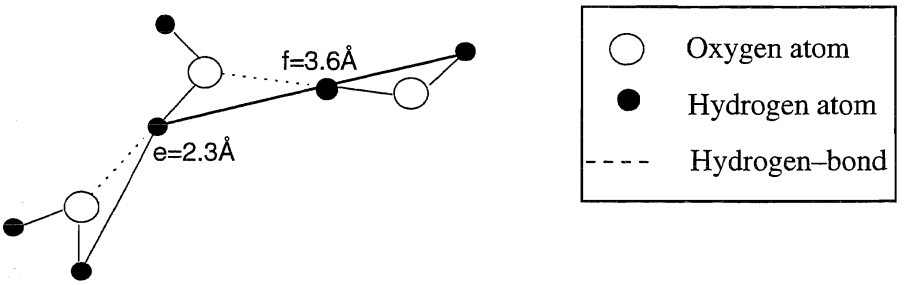


Figure 5.13 Approximate hydrogen–hydrogen distances between neighbouring molecules in a tetrahedral orientation. *e* contribute to the 1st peak of Figure 5.16, and *f* to the second peak. Figure after [7, 161].

Comparison with experiments:

Pure water structure have been investigated in several diffraction experiment, see for instance [6, 143, 146, 9, 7]. We have chosen to compare to the neutron diffraction results of Soper and Phillips, 1986, 1994 [143, 160]. There is overall agreement between our results and the data of [143], both regarding peak positions and peak heights. The only major difference is with peak heights of g_{OH}, where the difference in peak intensity of the two peaks (cf. Table C.7) is less than the experimental intensities [143].

According to the experimental results of [160], the peak heights of oxygen-oxygen correlation is however reduced from 3.1 to 2.3. This difference is believed [160] to be due to the uncertainty in the experiments. A comparison between those data and our results is shown in Figure 5.17. The experimental peak position is at 2.82Å, which agrees closely with the peak in our data at 2.8Å. The experimental results of the

oxygen-hydrogen correlation, Figure 5.18, show that the experimental first peak is markedly lower than our, while the second peak is equal to our within statistical uncertainty. The experimental height of the 1st peak is reduced from 1.385 [143] to 1.02 [160]. The experimental distance between the two peaks is slightly larger than our. The two peaks are positioned at 1.78Å and 3.30Å, while our peaks are at 1.85Å and 3.25Å. The best agreement is found with the hydrogen-hydrogen correlations, Figure 5.19, where the peaks are at exactly same positions, and also peak heights agree within statistical uncertainty.

For methanol–water solutions, we have not been able to find any experimental data on structure for mixtures at the selected compositions; we are in fact not aware of any such experimental study except for the above mentioned work of Soper and Finney [157, 161].

Comparison with simulations:

Numerous simulations of pure water have been performed with various potential functions. As a check upon reliability of simulation, we can compare to Jorgensens *et al.*'s [26] Monte Carlo results for TIP4P. It turns out that our calculated pair correlation functions show exactly the same features as their. Also peak heights agree, particularly the heights of g_{OH}, which disagreed slightly with experiment. Integrations of first peaks is reported to give 5.1 for oxygen-oxygen, and 3.9 for oxygen-hydrogen. The oxygen-oxygen peak height of ~3 agree with the results of [104].

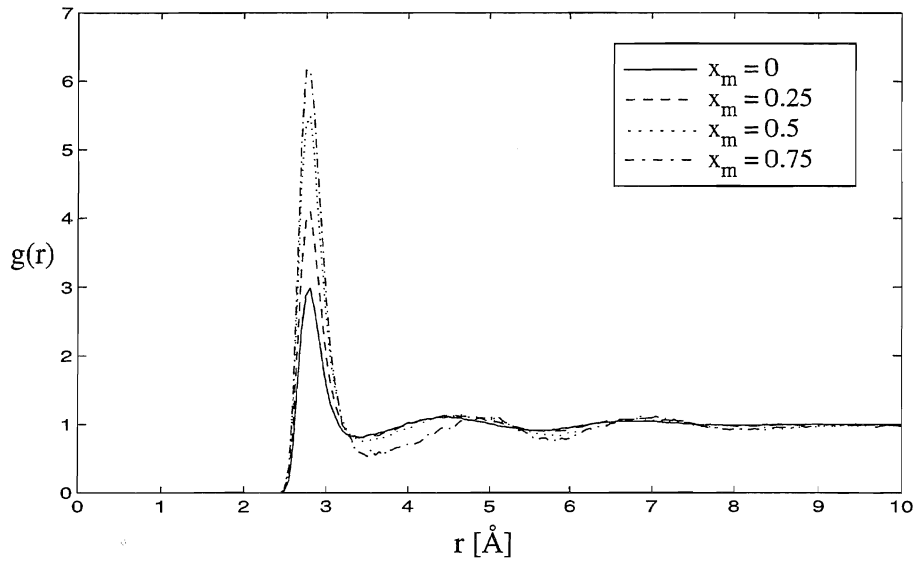


Figure 5.14 Radial correlation function g(r) for O_w - O_w in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

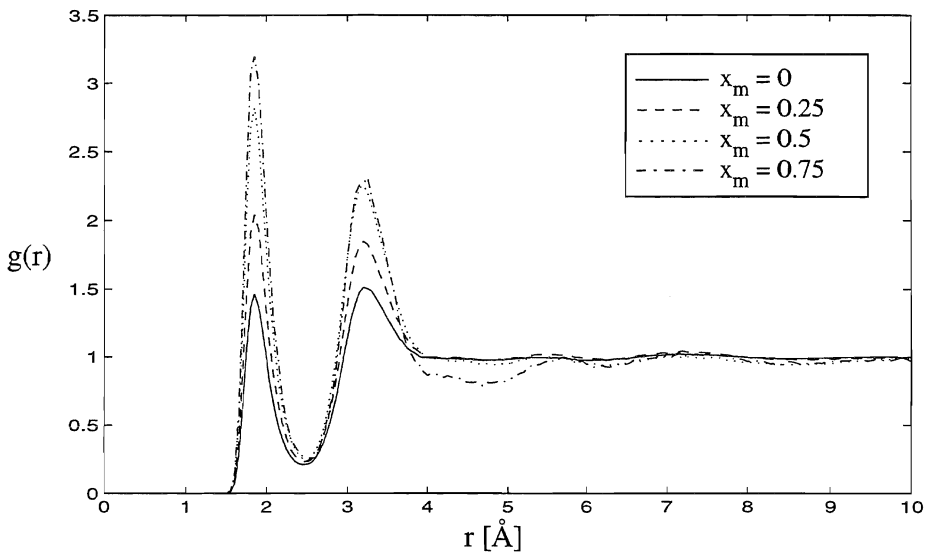


Figure 5.15 Radial correlation function g(r) for O_w - H_w in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

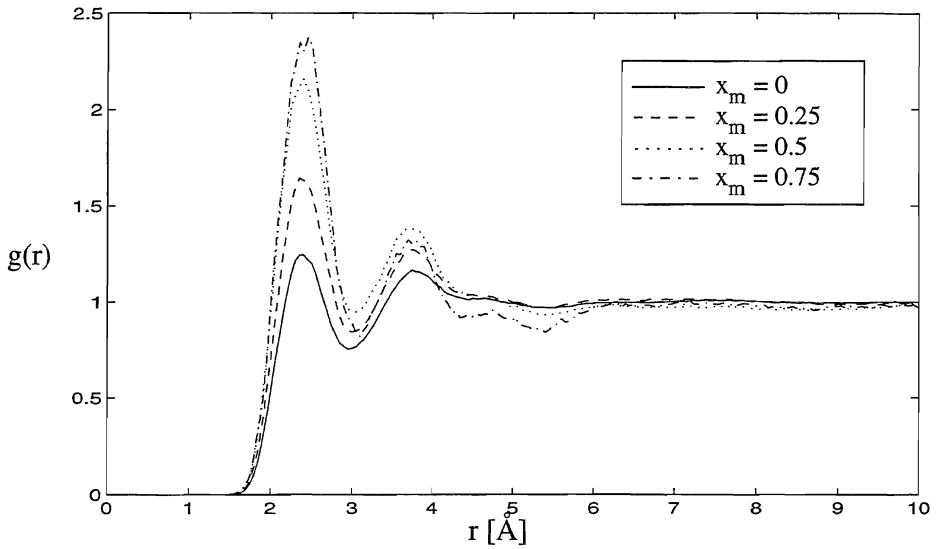


Figure 5.16 Radial correlation function g(r) for H_w - H_w in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

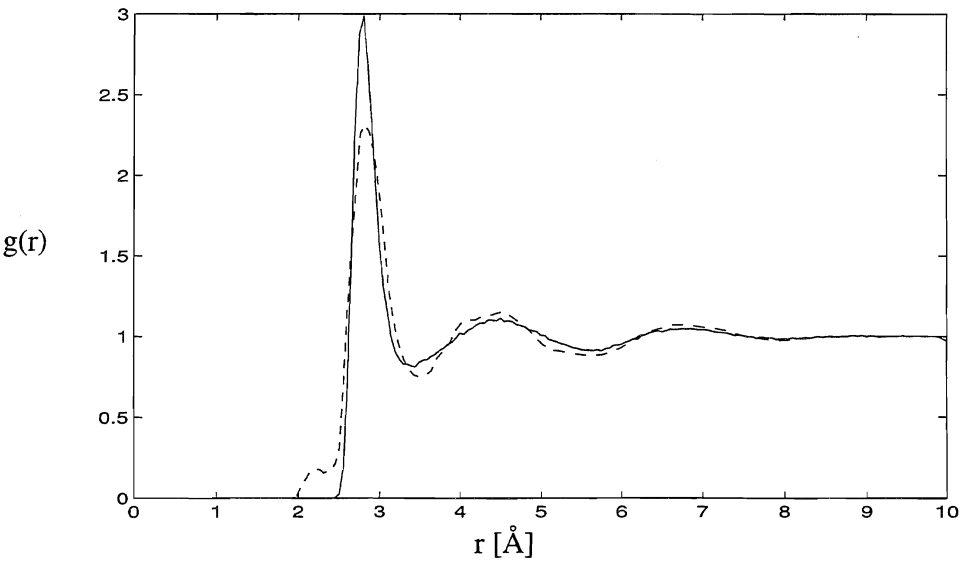


Figure 5.17 NVT-simuation of pure water (solid line) compared to data of [160] (dashed line) for oxygen-oxygen correlations. Statistical uncertainty near 1st peak is reported to be less than 0.2.

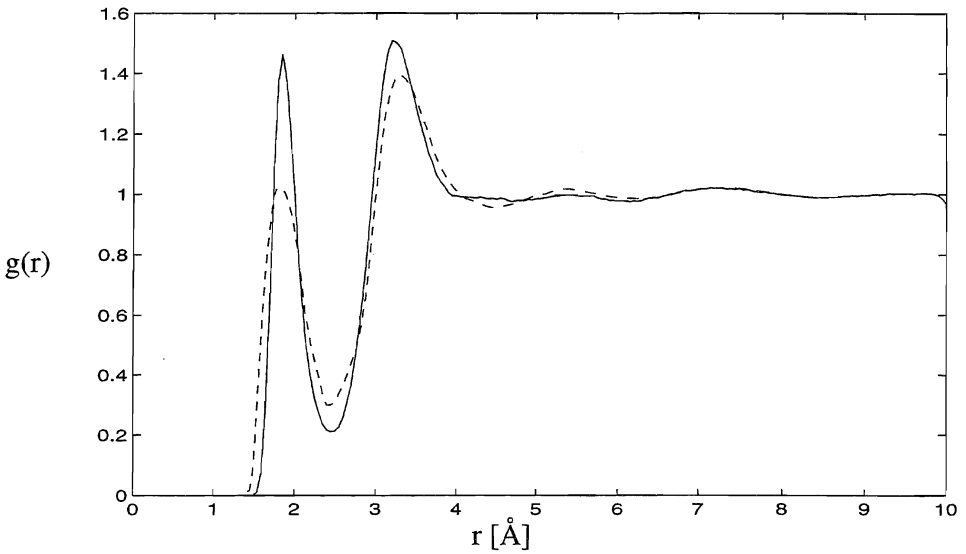


Figure 5.18 NVT-simuation of pure water (solid line) compared to data of [160] (dashed line) for oxygen-hydrogen correlations. Statistical uncertainty near 1st peak is reported to be less than 0.15.

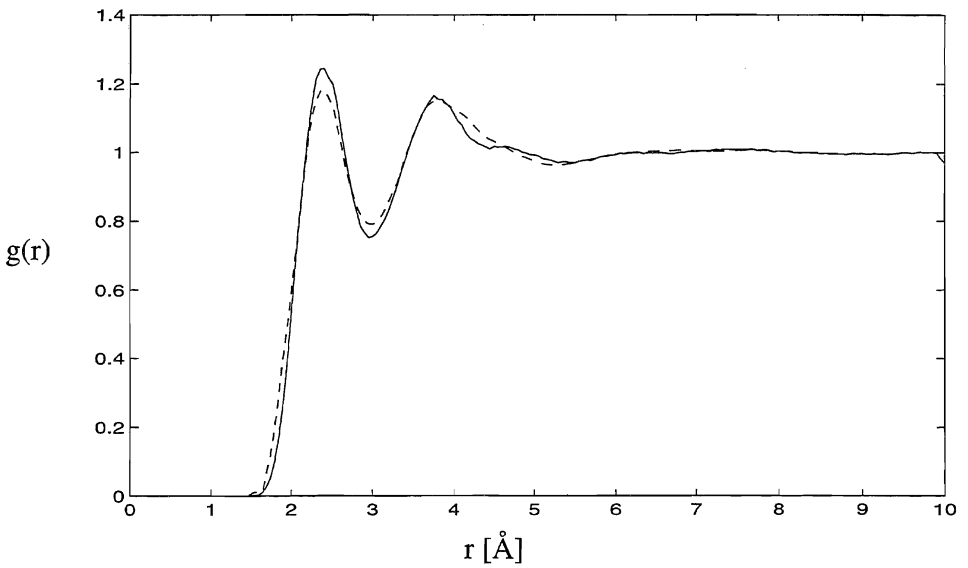


Figure 5.19 NVT-simulation of pure water (solid line) compared to data of [160] (dashed line) for hydrogen-hydrogen correlations. Statistical uncertainty near 1st peak is reported 0.1 approx.

5.4.3 Self-correlations for methanol

The self-correlations for methanol are shown in Figures 5.22 - 5.27. They are characterized by

- height of 1st maximum decrease with decreasing methanol content, except the methyl-methyl correlations, which are insensitive to composition
- positions of 1st peaks are unaltered with changing composition

As an aid to interpret the results we have calculated some distances in a straight V-chain polymer of methanol. The distances are given in Figures 5.20 and 5.21.

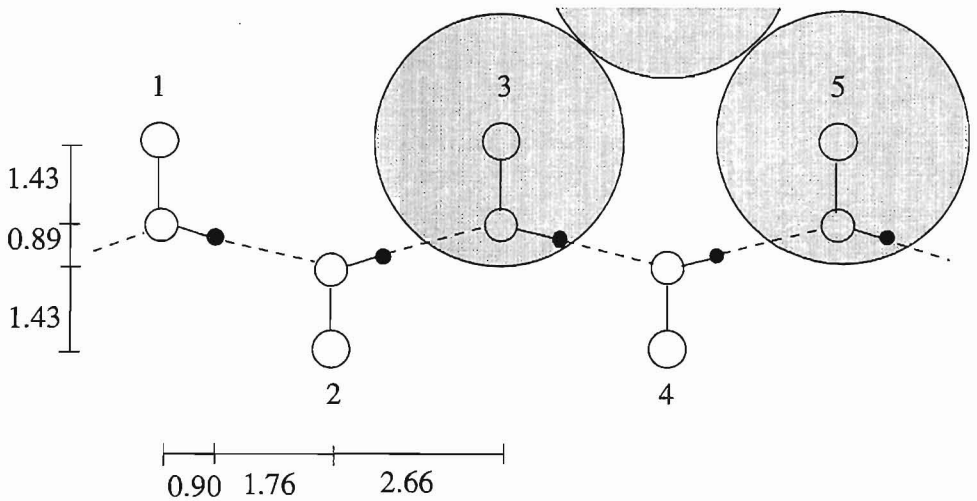


Figure 5.20 Intrachain distances in Ångstrøm in bulk methanol. Also shown is Lennard-Jones radius of methyl-group. Over-simplified view approximate to scale, to assist in visualising the pair–correlation functions.

Methyl-methyl correlations:

Figure 5.22 show the methyl-methyl self correlations. This site–site correlation function of pure methanol show a broad unsymmetric maximum at 4.1Å and a 2nd maximum at 8.0Å with an intervening minimum at 5.85Å. The 1st peak is sharpened slightly with decreasing methanol content shifting the minimum too higher separations. The position of 1st maximum is unaltered with changing composition. The second maximum is flattened and shifted towards lower separation. Both differences are small, and might be within statistical uncertainty.

With linear H-bonds and a planar chain, we see from Figure 5.21 that the closest intrachain nearest neighbour is at 4.6Å. (Could be both smaller and larger with winding chain or bent H-bond.) This configuration could be contributing to the right–hand side of 1st maximum. An interchain distance of 3.8Å (repulsive diameter of site) is also possible, and will then contribute to the left–hand side [27]. The minimum appearing at 5 -7Å, seem to preclude to some extent the intrachain next nearest neighbour at 5.3Å. The sharpening of first peak with decreased methanol content could be caused by a decrease in chain length.

Integration of first peak to 1st minimum yields a coordination number of nearly 12, decreasing - not linearly - with decreasing methanol content.

Methyl - oxygen correlations:

Pure methanol show (Figure 5.23) a sharp 1st maximum at 3.6Å due to intrachain nearest neighbours (see Figure 5.21) and a broad minimum at 6.45Å. Also a small minimum at 4.5Å grows to a shoulder for the two lowest methanol mole fractions, coincident with a decrease in intensity of 1st maximum. The shoulder is probably due to interchain contributions. Figure 5.21 show a possible arrangement with an interchain methyl–oxygen distance about 4.9Å. With the advantage of a 3-dimensional space, this distance could be lowered to about 4Å upon rotation of the neighbouring chain. Another possibility is that the nearest neighbours in a chain are substituted by water molecules. Then intrachain contributions to 1st peak is reduced, while the contributions from 2nd nearest neighbours are retained. Integration out to the minimum at 6.45Å yield a coordination number of 15.3 for pure methanol. If we use only the sharp peak, we find 4.9 nearest neighbours. The almost equal peak heights at the two lowest concentrations might be due to inaccurate sampling, but the difference in shoulder heights is significant.

Methyl - hydrogen correlations:

The methyl-group and the hydroxyl-hydrogen are both represented as positively charged sites, and have obviously no 'wish' of being close. The two distinct peaks at 2.9Å and 4.2Å, see Figure 5.24, must therefore be a consequence of a bonding of higher priority, namely the hydrogen-bond. From the simplified model in Figure 5.21 , we find an intrachain methyl-hydrogen distance of 2.7Å to the nearest chain neighbour, and a distance of 4.2Å to the next nearest chain neighbour. The first, low peak decreases clearly with decreasing methanol content, while the second minimum develops into a shoulder. The shoulder can be explained with interchain contributions. The reduction, and ultimately disappearing, of the first peak is consistent with replacement of methanol with water, both at high methanol concentration, and at low concentration where water is believed to build a cage around individual molecules. Integration of first peak to 3.3Å give a coordination number of 3.3, while extending

the integration to 4.8Å to include also the second peak yields 6.0 (both are pure methanol).

Oxygen - oxygen correlations:

The oxygen-oxygen correlations are shown in Figure 5.25. A sharp peak at 2.8Å due to the hydrogen–bond to the nearest neighbour, and a 2nd maximum at 4.95Å with a deep minimum at 3.45Å lying between are found. The 2nd maximum is at least in part due to next nearest chain neighbours. The minimum is shifted towards lower separation with decreasing methanol content and the 2nd maximum is broadened towards smaller separations. These distances (~4-5.5Å) are not consistent with the simple straight chain of Figure 5.21, but could originate from either interchain arrangement or intrachain branching/ring formation. Area under 1st peak is reduced from 2.0 in pure methanol to 0.5 in the most aqueous mixture. The reduction is consistent with nearest chain neighbours being replaced by water. A sort of chain must exist, since 2nd peak increase.

The self correlation for oxygen on methanol provide a control. The 1st peak is situated at a distance of 2.8Å, just as the oxygen-oxygen distance for self correlations of water. We notice that the peak heights are lower, but the heights for pure water and for pure methanol are roughly comparable. The 1st minima are also found at the same distances, though the depth is much larger with methanol than with water. The only difference between coordination numbers should then be number density, which is smaller by a factor of ~2.25 with methanol due to a larger simulation box, Table B.3. We therefore expect the coordination number for pure methanol to be a factor of 2.25 smaller than for pure water. From integration we find 2.0, which must be regarded as a sensible result from the discussion above.

Oxygen-hydrogen correlations:

These correlations, Figure 5.26, are characterized by a sharp peak at 1.9Å due to the hydrogen–bond, followed by a deep minimum at 2.65Å and a small 2nd maximum at 3.45Å due to intrachain contributions from hydrogen on the accepting neighbour (see Figure 5.21). The 1st minimum retains its depth with varying concentration, but the minimum near 4Å disappears with decreasing methanol concentration. The coordination number to 1st minimum is 1.0 for pure methanol, and decrease to 0.2 for the 0.25 mixture. Once again nearest neighbours are seen to disappear, while correlations beyond nearest neighbours increase.

Hydrogen-hydrogen correlations:

The self correlations for hydrogen on methanol show a 1st maximum at 2.5Å, which are the hydrogens on the nearest intrachain neighbours, followed by a deep minimum at 3.4Å. For pure methanol there is no second maximum, but there might be a weak maximum just above 5Å for the most water rich mixture. The minimum gets shal-

lower with decreasing methanol content. Integration to 1st minimum give a coordination number of 2.3, decreasing to 0.6 in the most water-rich mixture. A coordination number of 2.3 is larger than the expected 2-bondedness in the simple V-chain and also larger than the experimental value of 1.8, and might be a signal of 3-bondedness.

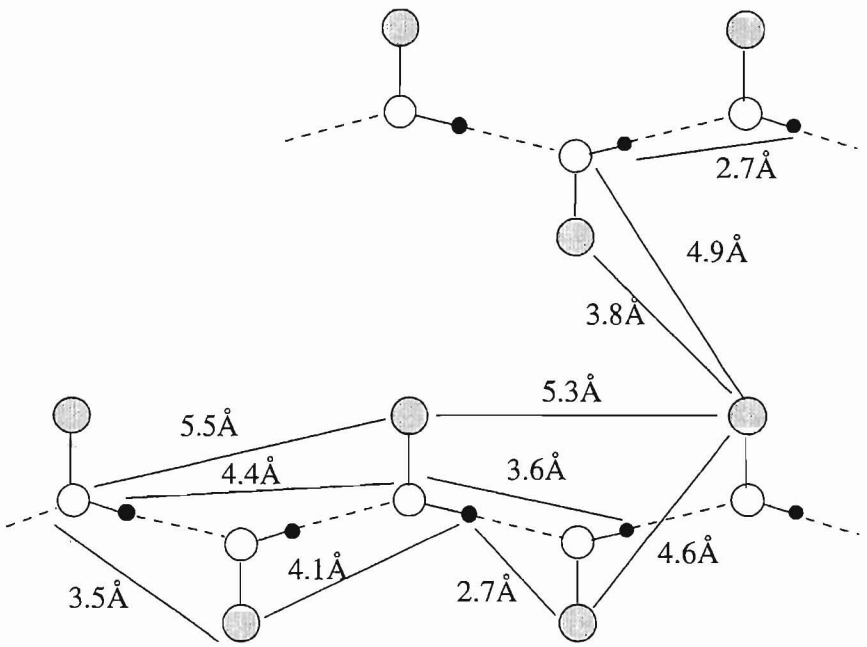


Figure 5.21 Some of the site-site distances with a planar non–branched methanol chain. Please imagine the effect of winding and branching upon the distances.

Comparison with experiments:

X-ray diffraction data of Narten and Habenschuss, 1984 [9], find that each hydroxyl group has on the average 1.8 hydroxyl neighbours at a distance of 2.8Å. They interpret their results as being caused by a chain structure, already suggested by Zachariasen [10] in 1934. Vahvaselkä *et al.*, 1995 [12] find the same distance and 1.7 hydroxyl neighbours. They find the closest possible C—C distance to be 3.5Å.

Comparison with simulations:

Due to the complexity of the molecules, all six different pair–correlation functions are not experimentally interpreted, and there is not any experimental results for mixtures at our selected compositions either. We must therefore resort to comparison with simulations.

For pure methanol we look at the MC–results of Jorgensen [27]. He finds the two maxima of g_{OO} at 2.75Å and 4.70Å, and the integration of the first peak to 3.4Å yield 2.0. The maxima of g_{OH} are at 1.82Å and 3.30Å. The first peak of g_{OH} integrate to 0.97 out to 2.6Å. The maximum of g_{HH} is at 2.35Å, and integration to 3.25 yield 2.1. All extrema appear at a slightly closer distance than our, but the agreement is good in view of the different simulation method and smaller system size of [27].

Wu *et al.* 1992 [131] use the older TIPS parameters [100] in their MC–simulation of dilute solution and equimolar mixtures of water and methanol. They also observe an increase in water-water and water-methanol pair-correlation functions and a corresponding decrease in methanol-methanol pair-correlation functions.

Haughney *et al.* [126] use TIP4P in mixture with model H1. This model is found to give results similar to the OPLS model [147]. They have focused on hydrogen-bonding analysis, but the structural features they report are in agreement with our.

Tanaka and Gubbins [113] report g_{CC} for x_m = 0.1, 0.3, 0.7. A slight increase in peak height and a shift of second maximum towards longer distances as methanol content increase, are observed.

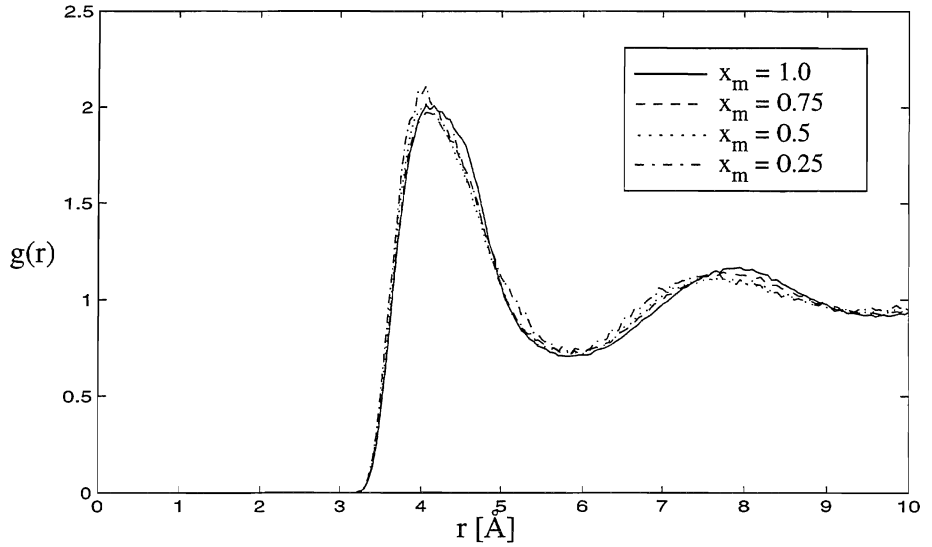


Figure 5.22 Radial correlation function g(r) for C_m - C_m in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

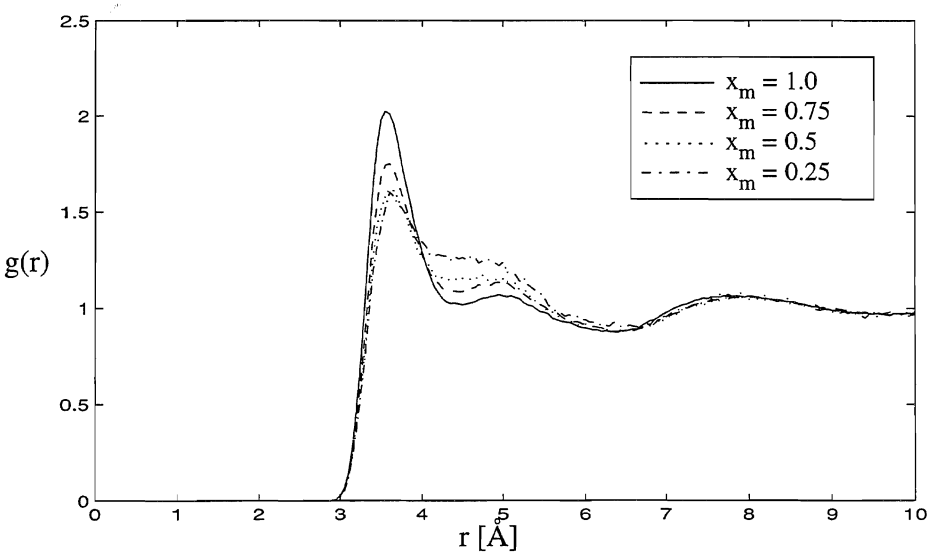


Figure 5.23 Radial correlation function g(r) for C_m - O_m in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

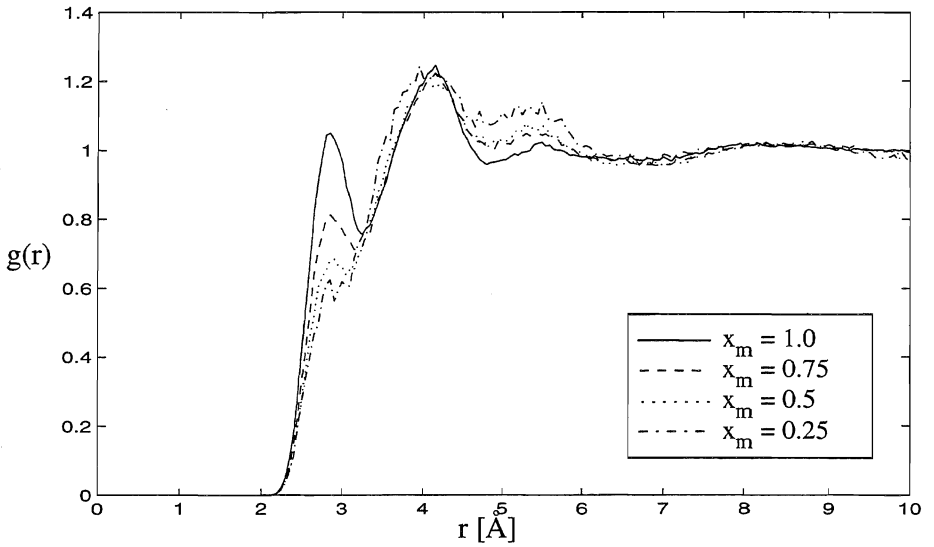


Figure 5.24 Radial correlation function g(r) for C_m - H_m in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

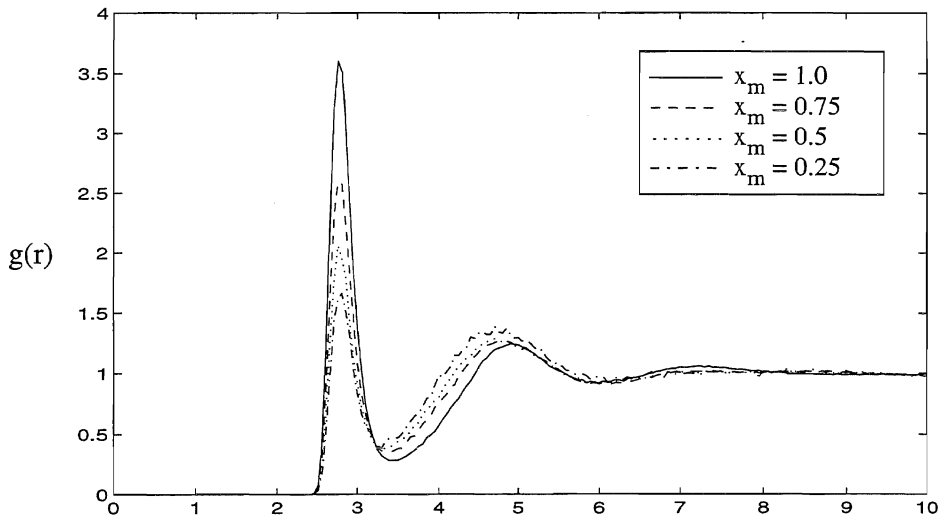


Figure 5.25 Radial correlation function g(r) for O_m - O_m in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

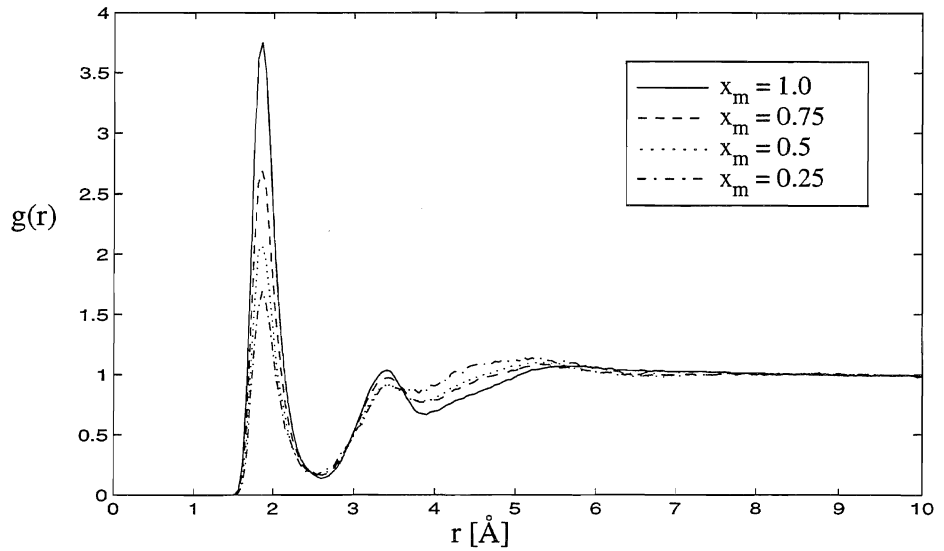


Figure 5.26 Radial correlation function g(r) for O_m - H_m in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

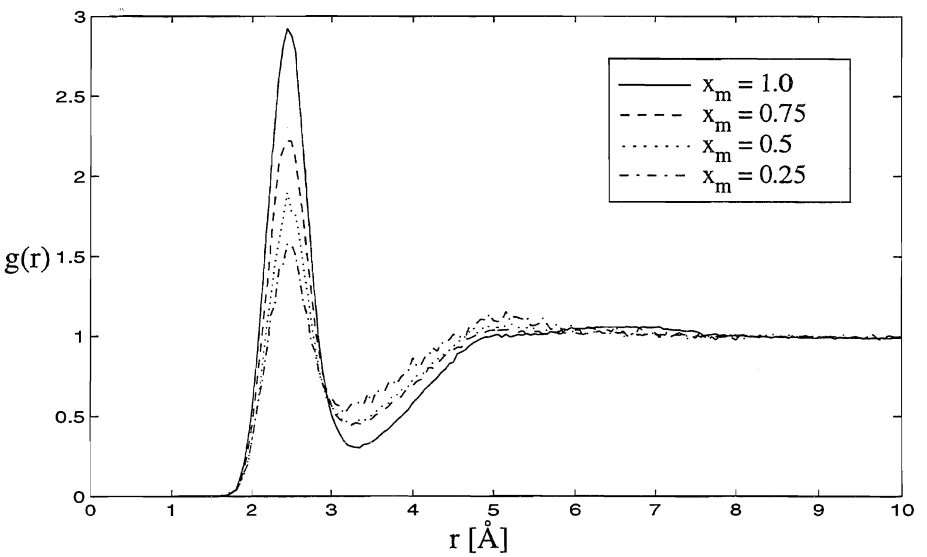


Figure 5.27 Radial correlation function g(r) for H_m - H_m in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

5.4.4 Cross-correlations for water and methanol.

The cross-correlations are presented in Figures 5.28 - 5.33, and they are characterized by

- increased peak height with increasing water content
- position of 1st maximum unchanged with composition
- only small variation with concentration beyond 1st maximum

Oxygen-methyl cross-correlations:
Figure 5.28 show a clear peak at 3.6Å followed by a broad minimum at 5.0–5.25Å. Position of 1st peak is consistent with water replacing methanol in chain. A very weak 2nd maximum, growing with increasing methanol molar ratio, at 5.35–5.5 Å can be seen. Coordination number increase from 3.3 to 6.9 with increasing methanol content.

Oxygen-oxygen(methanol) cross-correlations:
Figure 5.29 show a sharp 1st peak at 2.75Å and a broad 2nd peak at 5.25 Å. Both 1st and 2nd minima can be seen, and also a weak 3rd maximum. Coordination number increase from 0.7 to 2.6 with increasing methanol mole fraction. Compared to the oxygen-oxygen self correlations for both methanol and water, we see that the cross correlations are less sensitive to increasing methanol content beyond 1st maximum. In other words, the cross correlations for 2nd neighbours are less affected than the self-correlations.

Oxygen-hydrogen(methanol) cross-correlations:
Figure 5.30 show a 1st maximum at 1.9Å followed by a minimum at 2.6Å and a 2nd maximum at 3.4Å. The curve showing much of the same features as the corresponding self correlations for water at the same concentrations. First peak is due to water accepting a methanol proton. It integrates to one at the highest concentration of methanol, in agreement with the results of [156] for methanol mole fraction 0.9. The 2nd peak is due to H-atoms on the accepting neighbour. Coordination number increase from 0.3 to 1.0 with increasing methanol content.

Hydrogen-oxygen(methanol) cross-correlations:
Figure 5.32 show that the hydrogen-oxygen correlations are very similar to the O_w–H_m correlations described above, except for the first peak being a little higher in the former. Peaks appear at 1.8Å and 3.25Å, first peak integrating to 0.7 in the methanol richest solution. 1st peak must be due to water donating a proton to neighbouring methanol, while 2nd peak arises form the other hydrogen bonded methanol molecules.

Hydrogen-methyl cross-correlations:

This correlation, see Figure 5.31, show a clear similarity with the self-correlation for hydrogen-methyl at the same molar ratios up to ~4.5Å. An unsymmetric peak at 4.1Å is followed by a broad and shallow minimum near 5.6Å. A shoulder at 2.8Å develops to a small peak with increasing methanol concentration. The small 1st peak is due to water donating a proton to the hydroxyl group of methanol, while the large unsymmetric peak can be interpreted as water encageing the methyl–group. The steepness on the right side of the 2nd peak mean that little variation in position are found, while there are some randomness on the other side. Coordination numbers increase from 4.4 to 9.2 with increasing methanol concentration.

Hydrogen-hydrogen(methanol) cross-correlations:

A 1st peak at 2.45Å is followed by a minimum at 3.2Å. For the two methanol richest mixtures, a tiny maximum can be envisaged below 4Å. As for the two hydrogen-hydrogen self-correlations, the broadening of the 1st peak is symmetrical on both sides of the peak. Nearest neighbour interactions are however also here similar to the corresponding self–correlations, but the orientation of the other hydrogen can not be correlated, because of the missing 2nd peak. The coordination number increases from 0.6 to 2.0 with increasing concentration.

Comparison with experiments:

Soper and Finney, 1993, [157] have performed neutron diffraction experiments on 1:9 molar ratio mixtures in order to verify whether water generates a cage around the methyl group. They find a carbon-to-water(oxygen) distance of 3.7Å with approximately 10 neighbouring water molecules spherical arranged around the methanol molecule. The C-O_W distance is a little longer than our results for the 0.25 mixture. There seem to be no enhancement of water structure accompanying the addition of methanol.

Comparison with simulations:

Freitas [124] has calculated cross-correlations for TIP4P-water and OPLS-methanol at various mole fractions. His displayed pair-correlations for O_W-O_M, O_W-H_M, H_W-O_M, and O_W-Me at x_m = 0.25, 0.5, and 0.75, all agree very well with our result. So do also his calculated oxygen–oxygen coordination numbers.

Jorgensen and Madura [152] have simulated 'infinitely diluted' methanol in water. With the exception of peak heights, the shape and position of extrema of their displayed solute–solvent pair distribution functions are in agreement with our.

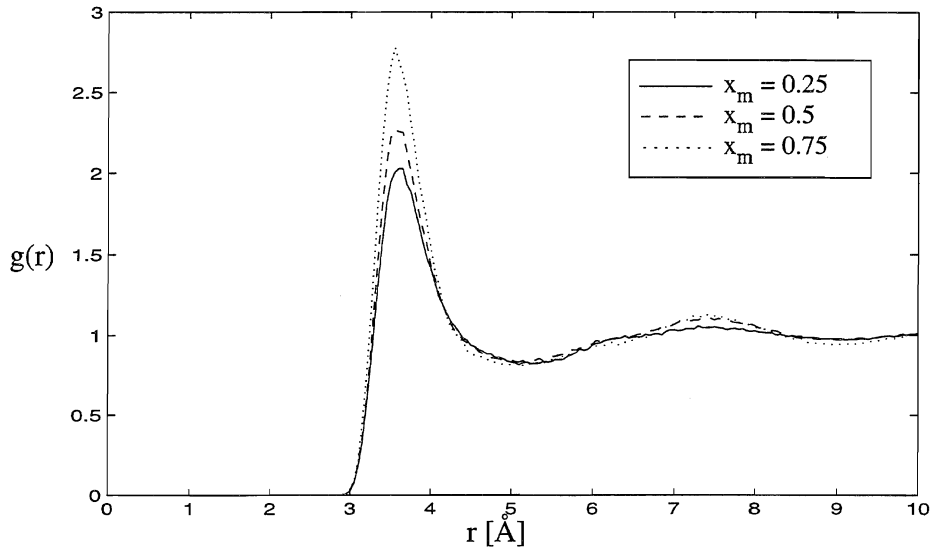


Figure 5.28 Radial cross-correlation function g(r) for O_w - C_m in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

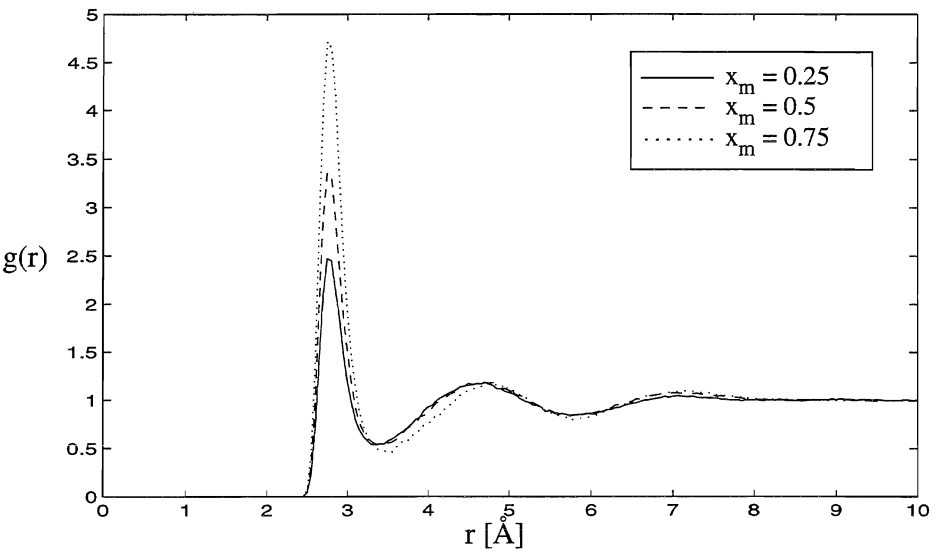


Figure 5.29 Radial cross-correlation function g(r) for O_w - O_m in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

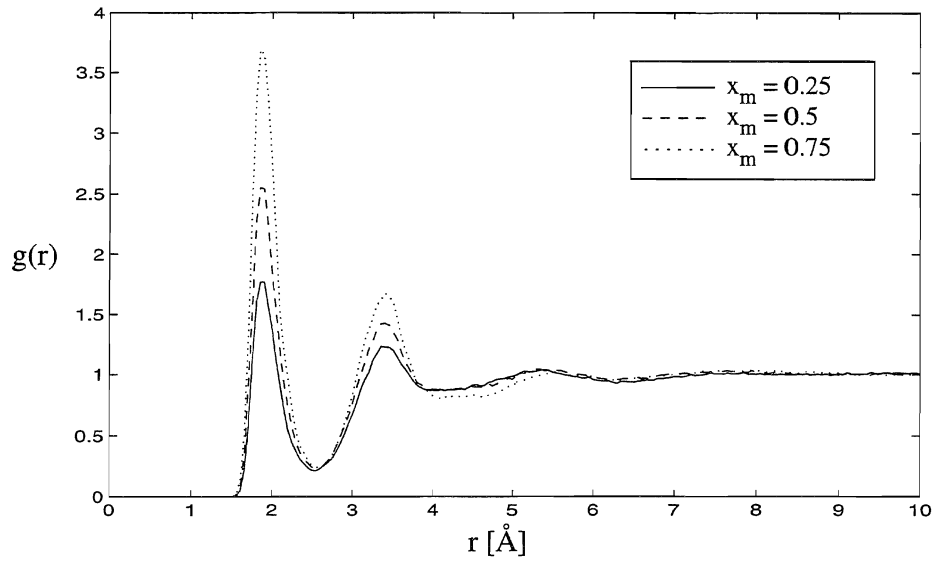


Figure 5.30 Radial cross-correlation function g(r) for O_w - H_m in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

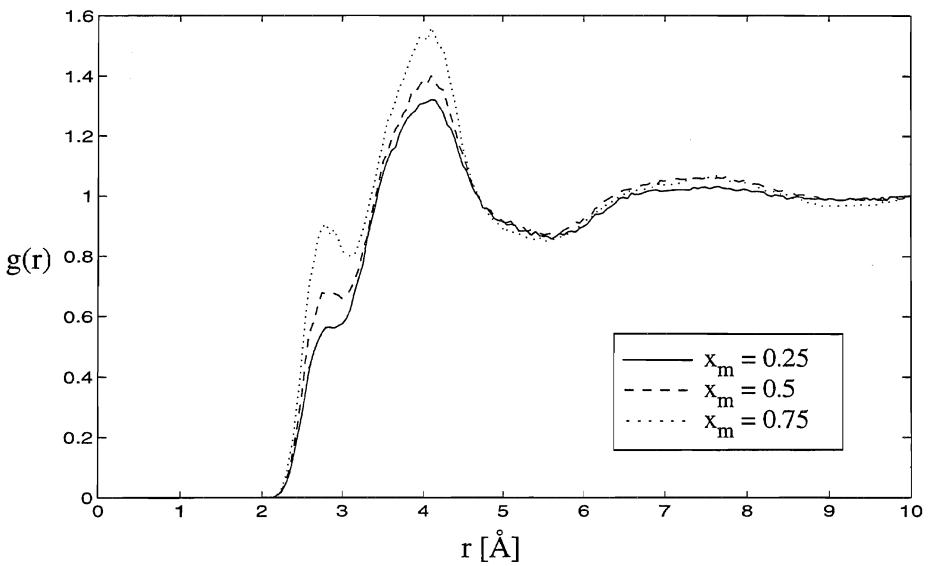


Figure 5.31 Radial cross-correlation function g(r) for H_w - C_m in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

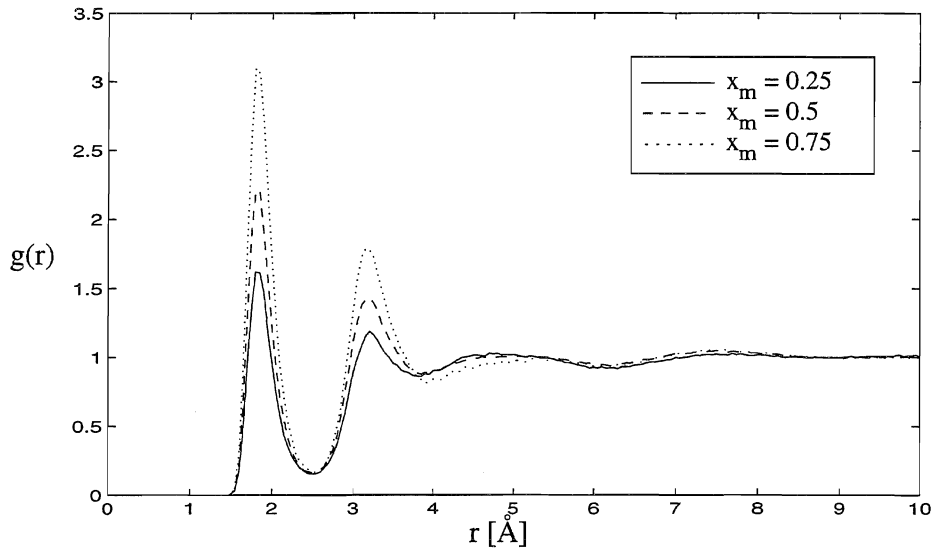


Figure 5.32 Radial cross-correlation function g(r) for H_w - O_m in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

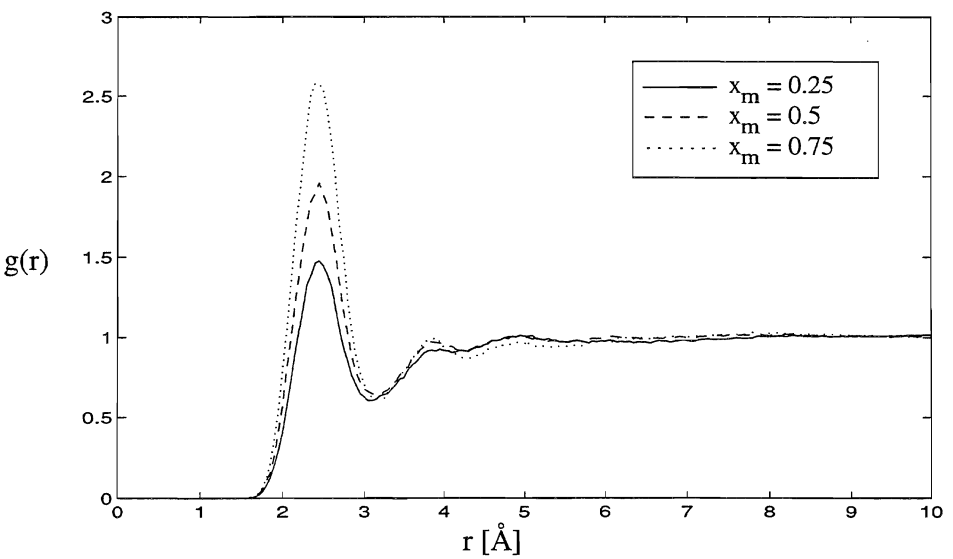


Figure 5.33 Radial cross-correlation function g(r) for H_w - H_m in water-methanol mixture sampled with NVT simulation. x_m is methanol mole fraction. Simulation conditions as given in table B.3.

5.4.5 Site-site coordination numbers.

The site-site coordination numbers presented below show the average number of neighbouring sites of a particular kind within the first coordination shell. To make conclusions about the number of H-bonds, one has to analyse the configurations with respect to criteria upon geometry (distance and angle) or require the energy to be below a specific value. Our site coordination numbers involving hydrogen and oxygen sites are therefore not measures of hydrogen bonds, since these arise from integration of first peak only.

Note also that site-coordination numbers between unlike sites on the same kind of molecules are a factor of two lower than the molecular coordination number. For coordination numbers between equal sites on the same kind of molecules, the site coordination will equal the molecular coordination. This applies regardless of composition.

From experiments, the coordination numbers of water is reported as ~4.5 (O_w-O_w) and ~1.8 (O_w-H_w) [143] at room temperature. For methanol, the coordination number (O_m-O_m) is reported to be 1.7 [12] and 1.8 [9]. Our corresponding site coordination numbers for the pure liquids, see Tables 5.9 and 5.10, are all a little larger than the experimental values.

From Monte Carlo simulations of mixtures of TIP4P-water and OPLS-methanol, Freitas [124] report the site-coordination numbers for oxygen-oxygen self- and cross-correlations. For water he finds 4.95, 3.49, 2.44, and 1.22, for methanol 0.46, 0.88, 1.37, and 2.07, and for water-methanol he finds 0.81, 1.59, and 2.51. All values are in very good agreement with our results, see Tables 5.9, 5.10, and 5.11.

From molecular dynamics simulations of flexible, three-site model mixtures, Pálinkás and Bako [155] find coordination numbers for O_w–O, O_m–O, and C–O at x_m=0.25 to be 3.81, 2.75, and 9.80. If we add our site-site coordination numbers, Tables 5.9, 5.10 , and 5.11, we get 4.3, 2.6, and 11.4, respectively.

The maximum number of hydrogen bonds per molecule is expected to occur near this concentration [124, 155].

Ferrario *et al.* [126] have calculated coordination numbers for O_m–O_w and C–O_w for dilute solutions x_m=0.125 and 0.875 of TIP4P-water and H1-methanol. They find that the methyl group is surrounded by ~16 water molecules and that the hydroxyl group has 2.6 water neighbours in the water-rich solution, while water oxygen is surrounded by 3 methanol oxygen and 8 methyl groups in methanol-rich solution. Their results, though at other compositions, fit nicely in at top and bottom of Table 5.11.

Table 5.9

Site coordination numbers for water-water correlations integrated to respective minima.

x_m	n_{OO}	n_{OH}	n_{HH}
0.0	5.1	2.0	5.3
0.25	3.6	1.5	3.8
0.50	2.5	1.1	2.9
0.75	1.1	0.5	1.3

Table 5.10

Site coordination numbers for methanol-methanol correlations integrated to respective minima.

[a] Numbers in parentheses from integration to 4.2Å(x_m=0.25 and 0.5) and 4.5Å.

[b] Numbers in parentheses from integration to 4.8Å.

x_m	n_{CC}	n_{OO}	n_{HH}	n_{CO} [a]	n_{CH} [b]	n_{OH}
0.25	5.8	0.5	0.6	1.5 (8.0)	0.2 (2.6)	0.2
0.50	8.9	0.9	1.2	2.4 (10.6)	0.4 (4.1)	0.4
0.75	10.7	1.4	1.7	4.1 (14.4)	0.8 (5.2)	0.7
1.0	11.9	2.0	2.3	4.9 (15.4)	1.1 (6.0)	1.0

Table 5.11

Site coordination numbers for water-methanol cross correlations, 1st index is water-sites. Methanol–water coordination numbers is found by multiplying the entries with N_w/N_m

x_m	n_{OC}	n_{OO}	n_{OH}	n_{HC}	n_{HO}	n_{HH}
0.25	3.3	0.7	0.3	4.4	0.2	0.6
0.50	5.9	1.5	0.6	7.0	0.4	1.3
0.75	6.9	2.6	1.0	9.2	0.7	2.0

5.5 Dynamical properties for methanol/water mixtures.

5.5.1 Self–diffusion for water and methanol.

Table 5.12 below show calculated and experimental values for self-diffusion of water and methanol calculated from centre of mass mean square displacement, Equation (3.36). Except for the NVE simulations of the pure liquids, all temperatures are within ±2K of 298K, cf. Table 5.1.

Table 5.12

Calculated and experimental self diffusion coefficients for water and methanol at 298K, except [a] at 309K and [b] at 306K. Cf. Table 5.1. Experimental results at 25°C from [162, 163]. Values of Partington [164] *et al.* in parentheses.

Molefraction methanol x_m	D[10^{-5}cm^2/s] NVE		D[10^{-5}cm^2/s] NVT		D[10^{-5}cm^2/s] exp	
	water	methanol	water	methanol	water	methanol
0.00	4.0[a]		2.9		2.26 (2.43)	
0.25	1.8	1.1	2.0	2.0	1.32	1.10
0.50	1.4	1.4	1.5	1.1	1.19	1.17
0.75	1.2	1.7	1.6	2.0	1.32	1.67
1.00		2.5[b]		2.4		2.44 (2.27)

Both the NVE and the NVT simulations show a decrease in the self-diffusion coefficients of water as the solution becomes more methanol-rich. An increase in self-diffusion of methanol with increasing x_m is also found. These trends are in agreement with experiments. Bifurcation (see Subsection 5.4.1) offers an explanation to the decreased mobility of water in solution [142]. The number of water molecules with 5 nearest neighbours is reduced, and the linear H-bonds require more energy to break than the bifurcated bonds. Consequently, dynamics of water are slowed down. Metha-

nol, on the other hand, experiences an increase in 3-bondedness in going from pure liquid to solution [126], and must also have a decreased mobility.

Nearly all values are higher than the experimental ones, but the disagreement between simulation and experiments are less the higher the methanol concentration becomes, for both components. From the NVT results we find that self-diffusion coefficient of pure methanol is in close agreement with experiment, while the self-diffusion coefficient of water is too high.

The NVE simulation temperature of the pure liquids are higher than requested by ~9K for water and ~7K for methanol (Table 5.1). Kida and Uedaira [166] have measured self-diffusion coefficients at 32°C, and find D_w~D_m=2.75·10^{-5}cm^2/s. Our calculated water value is still too high, particularly with the NVE simulation, while the methanol value is slightly underestimated.

The results of Hawlicka [162], shown in Figure 5.34, show that the self-diffusion for methanol has a flat minimum around 0.25 (extending between 0.2-0.5), and that the self-diffusion of water has a minimum around 0.5 (from 0.3-0.7). The same trend is also found at 32°C [166]. Our NVE simulation show a minimum in self-diffusion of water at x_m=0.5.

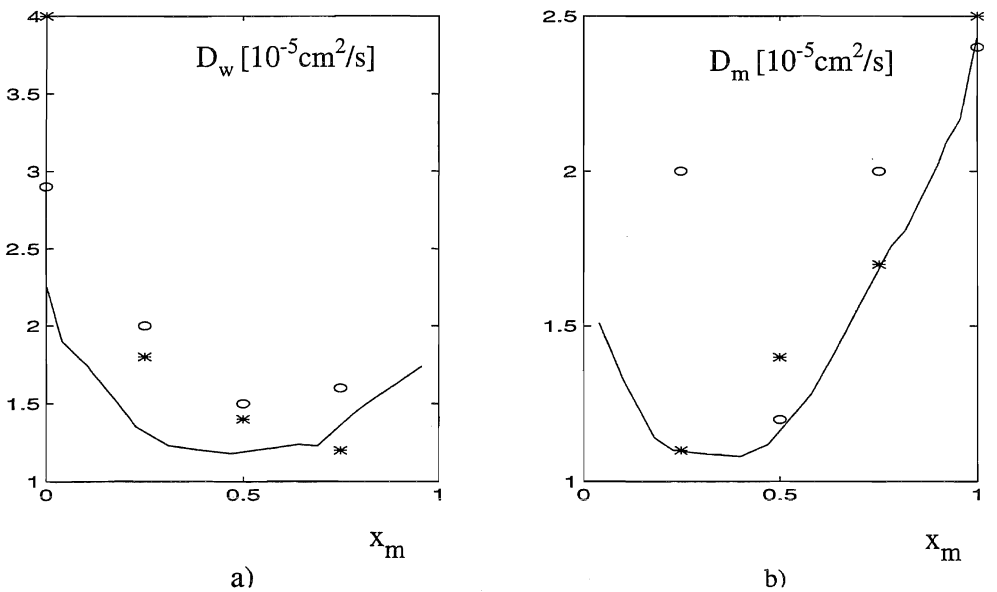


Figure 5.34 Self-diffusion for water, a) and methanol, b) calculated from NVE-simulation (*) and NVT-simulation (o) compared to experimental results of [162] (solid lines). Note that NVE-values for pure liquids are at higher temperature, see text.

We have no simulation results for the concentration range 0 – 0.25, and can therefore not confirm the minimum of methanol. The NVE simulations yield results for methanol in close agreement with experiments, see Figure 5.34. Formally, the NVT simulation yield a minimum for methanol at x_m=0.5. The NVT-value for D_m at mole fraction methanol 0.25 is an exception to the overall reasonable agreement with experiments. We find no indications in the results of the previous sections to why this value should be so high. Further simulations are needed to say anything conclusive about the source of this discrepancy.

We can also compare to a few simulation results. For OPLS at 298K and a system of 256 particles yield a self-diffusion of $2.33 \cdot 10^{-5}$cm^2/s [34], in perfect agreement with our result. Watanabe and Klein, 1989 [104], find the self diffusion of TIP4P to be $(3.3 \pm 0.5) \cdot 10^{-5}$cm^2/s at 298K, which compares to our results.

Haughney *et al.* [147] find that the OPLS model slightly underestimate self–diffusion in the range 268.2K to 338.2K. The calculated values of Ferrario *et al.* [126], for TIP4P-water – H1-methanol mixtures are typically a little lower than our results, and yield a pure water self–diffusion of $2.4 \cdot 10^{-5}$cm^2/s. This is in accordance with H1 underestimating self-diffusion in pure methanol more than OPLS. Considering their statistical uncertainties of ±0.15-0.3, the values are consistent except for the pure water value and the D_m for the lowest mole fraction methanol discussed above.

Matsumoto and Gubbins [148] find $2.7 \cdot 10^{-5}$cm^2/s at 300K with OPLS as an average over diffusivities for each bonding state.

We estimate the statistical significance in our calculated values to be about ±0.2 units. Also Haughney *et al.* [147] estimate the statistical uncertainty in their calculations to be 10%, which is almost the same as our. Our assumption is based upon results of occasional variation in delay time, and the observation that all mean square displacements not are straight lines at long times (see Figures C.32 and C.33). The tangent to the curves are therefore sensitive to the delay time at which it is taken. For all curves we have used the two last values of the curves to find the tangent. With a tangent to a larger portion of the curves, the values can differ by as much as 0.3 units. With an uncertainty of this magnitude, the NVE and NVT–results are within each others accuracy, except for the methanol values of the 0.25 mixture, and possibly the pure water values. As the simulation length and total system size are constant, the statistical precision of the data in Table 5.12 also varies with methanol mole fraction.

Casulleras and Guardia [34] have studied diffusivities for pure methanol (OPLS) at different system sizes, and found that even a system of 512 molecules – twice our size – is far from the thermodynamical limit. They find a small, but systematic, increase in the self-diffusion coefficient of methanol with increasing system size.

Wallquist and Teleman [115] find however that self-diffusion of SPC-water [99] (rigid, three site model) is not significantly enhanced by increasing the number of molecules from 216 to 1000 molecules. They find also no significant difference in self-diffusion coefficient whether they use the Ewald summation or a spherical truncation scheme.

We could perhaps turn the question of assigning uncertainties upside down: Formally, the two simulation methods should give equal self–diffusion coefficients. Therefore we can regard the two series of simulations as two different 'experiments'. If we use their spread around average values as a measure of accuracy, we can find the statistical uncertainty of each entry in Table 5.12.

For the reason to why the mean square displacements not are linear, we will point at the relative short length of simulations of only 80ps. The use of a Nosé-Hoover thermostat is expected [24] not to influence upon the dynamics of a system. But we found in Section 5.3 that our simulations are not strictly representatives of the canonical ensemble, we can not rule out the possibility that the simulation method affects the dynamics.

The trend of self–diffusivities of water to be overestimated, might be due to the potential model TIP4P. The results are in much better agreement with values at 32°C [166].

5.5.2 Velocity auto correlation functions for water and methanol

Velocity auto correlation functions for centre of mass motion relative to a laboratory frame of reference appear in Figures 5.35 - 5.36 below, and in Figures C.34 - C.37.

The negative region at intermediate times found for all mixtures is typical for liquids where density is so high that the average molecule change direction upon interaction with another molecule. The oscillations of both methanol and water, particularly clear in Figure C.38, is a fingerprint of the hydrogen–bond, which makes the molecule oscillate back and forth before an eventual escape from its neighbourhood [167]. These oscillations are seen to diminish with increasing temperature [98]. They are more pronounced with water than with methanol, perhaps not unexpected since water has more hydrogen bonds than methanol. Pure water also decorrelates more quickly initially, a decrease to 1/e within ~0.05ps, as compared to methanol which reaches 1/e within ~0.1ps. A complete decorrelation seem to occur within approximately the same time, ~1ps for both liquids.

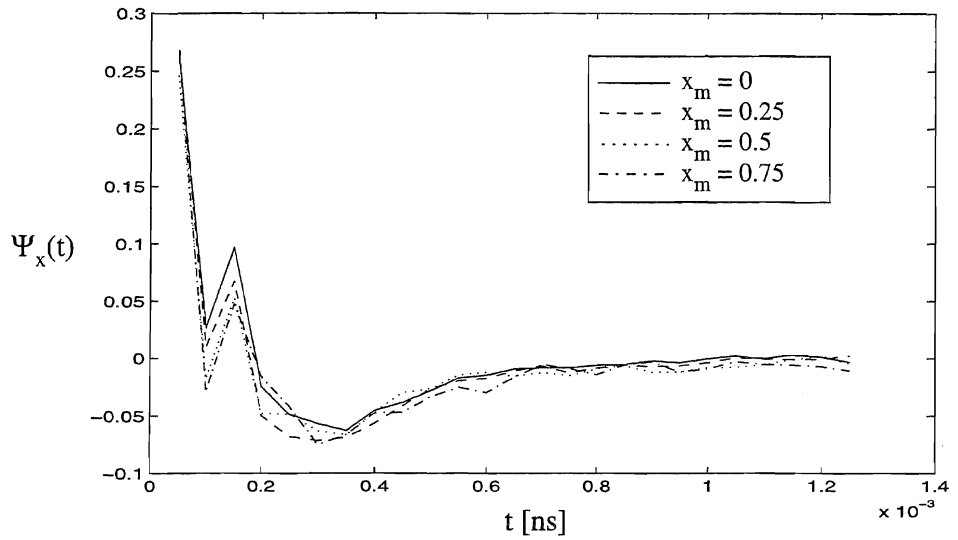


Figure 5.35 Normalized velocity auto correlation function $\Psi_x(t)$ for water in mixtures with methanol from NVT–simulations. x_m is methanol mole fraction.

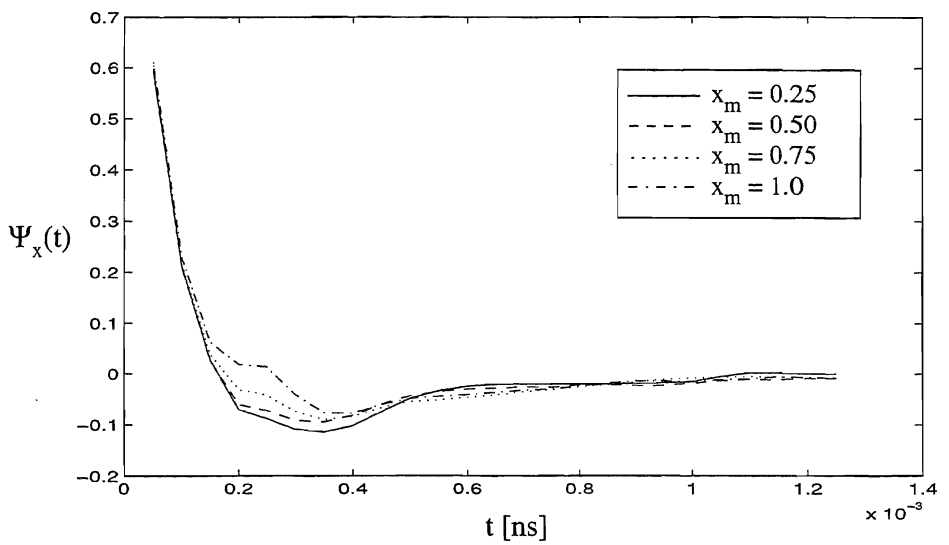


Figure 5.36 Normalized velocity auto correlation function $\Psi_x(t)$ for methanol in mixtures with water from NVT–simulations. x_m is methanol mole fraction.

Pure water is characterized by a correlation peak at $0.15 \cdot 10^{-3}$ns, which does not disappear with reduced mole fraction water, but instead is systematically lowered. In the negative region we find on the other hand no systematic variation with composition.

We have not been able to find any centre of mass velocity auto correlation functions for pure TIP4P-water in literature, despite its wide and frequent application in model simulations. Stillinger and Rahman [98] calculated velocity auto correlation function for ST2-water at different temperatures. Our calculated velocity auto correlation function is remarkably similar to their, except for being less oscillatory and having a local minimum with positive ordinate-value. We might however be misled by the lack of intermediate points.

Pure methanol has a small plateau at 0.2 - $0.25 \cdot 10^{-3}$ns, which disappears gradually for the mixtures. The negative region gets more negative with decreasing methanol content. Our pure methanol velocity auto correlation function is in good agreement with the the result of Guardia *et al.*, 1994 [168] for OPLS-methanol and the results of Haughney *et al.*, 1987 [147] and Alonso *et al.*, 1991 [150] for the H1-model.

The curves for the pure liquids are probably approaching zero faster than for the components in the mixtures.

Comparison between x-, y-, and z-components show only minor differences at long delay times, which are probably due to statistical noise.

Also a comparison between NVT and NVE results (Figures C.36 and C.37) reveal only small differences. The velocity auto correlations get less negative with the NVT–simulation. The difference with method in the backscattering region is only slightly larger than the difference with direction (not shown for this mixture) for water. For methanol the differences in direction is approximately equal to the difference with method.

Velocity auto correlation functions provide an alternative route to the self–diffusion coefficients [15, 169]. The use of this procedure in addition to the Einstein relation might have provided insight to whether differences in self-diffusion coefficients with the NVE and NVT simulations arise from method or are due to statistical uncertainties.

5.6 Summary

From the equilibrium molecular dynamics simulation of water and methanol, we summarize some of our results presented and discussed in the previous sections.

- The results for thermodynamic properties and structure are in good agreement with experiments and with simulations for the same models.
- There seem to be no difference between the NVE and the NVT results for thermodynamics and structure, except for differences caused by the NVE-temperatures being slightly higher than the NVT-temperatures.
- This disagreement of temperatures is largest for the pure liquids, and for the mixtures the NVE temperatures agree well with the predefined temperatures.
- The calculated pressures are very high, but it seem to be an effect of the simulational procedure/model potentials, since all other quantities correspond to values at 1 atmosphere. The pressures in the pure water simulations are a factor 3-4 higher than the pressures found for the mixtures. The pair correlation functions show no sign of a compressed liquid.
- Energy conservation is good, and improves with decreasing water content. Total linear momentum is also conserved.
- All tests that are performed and all variables that are investigated for both the NVE and the NVT methods, are consistent with systems in thermal and mechanical liquid equilibrium.
- The NVT method fails however to reproduce the canonical ensemble for the particular coupling parameters used. The possibility that the simulation would have become canonical if the simulation was continued, can however not be excluded.
- The Nosé-Hoover dynamics is proved to generate canonical distributions provided the trajectories are ergodic, and that the total linear momentum is conserved at zero. The failure of our simulations to generate canonical distributions, is then believed to be caused by slightly too large heat bath masses.
- The relaxation parameters used for mixtures were weighted averages of the separate water and methanol relaxation parameters. There seem to be no differences with the mixtures regarding the approach to the canonical ensemble, so a simple combination scheme is sufficient for mixtures.
- The radial positions of 1st maxima of the pair correlation function are not affected by varying concentrations, particularly we find no signs of increased cavities around the methyl site.

- The methyl-methyl coordination number decreases less than the other self coordination numbers with increasing water content, and less than the reduced number density. There is thus a relative increase in methyl-coordinated-methyl. The hydroxyl self coordination numbers decrease more than number density.
- The self correlations of water in the mixtures yield coordination numbers that decrease linearly with number density. Thus we find no drastic reduction or increment of the water structure.
- The pair correlation functions for methanol are consistent with a simple V-chain, but we have not considered rings or branched chains. These possibilities are then not excluded by our results.
- Upon addition of water to methanol, we find that the nearest neighbour interactions of methanol-methanol decrease, while second nearest interactions are less affected. This can be explained with water replacing methanol.
- The results for self-diffusion carry some uncertainty. The general trends known from experiments are reproduced. Also the tendency of the TIP4P-model to overestimate self-diffusion is confirmed. Except for the self-diffusion of methanol in the 0.25 mixture, the NVT and NVE results are equal within estimated uncertainties.

Though the simulations are not strictly microcanonical, the total of all results make us confident that we can proceed with simulations of the water-ethanol mixtures. The requirement of canonical distributions is particularly important for calculations of derivative properties. We do not calculate such properties.

Chapter 6
Mixtures of water and ethanol - results and analysis

Below we present the results from thermostatted molecular dynamics simulations of the model system water - ethanol (*trans*) at 293K. We have also performed an NVE simulation for one of the mixtures. The simulation conditions and molecular models are given in appendix B. The molecular models were discussed in Chapter 4. The calculational procedures and methodology applied are similar to the water - methanol simulations. Experimental data exist in varying abundance for thermodynamical data, structure and diffusion, but simulation results is scarce. Only a handful of Monte Carlo studies of liquid ethanol and of the infinitely diluted' aqueous solution are published. We are not aware of any molecular dynamics study of neither liquid ethanol nor its aqueous solutions at any concentrations.

The structure of this chapter is similar to the previous chapter on methanol-water mixtures. In Section 6.1 we present the thermodynamic properties, in Section 6.2 we discuss verification of the simulations, in Section 6.3 we review models for ethanol water mixture, and present our calculated correlation functions. In Section 6.4 we present results for self-diffusion coefficients and velocity auto correlation functions for water and ethanol. The chapter is concluded in Section 6.5, with a summary of our findings.

Much of the discussion of the results of water-methanol mixtures also apply to water-ethanol mixtures, and we will refer to the previous sections where appropriate.

6.1 Thermodynamic properties

Table 6.1 present calculated thermodynamical data from NVT simulations of water-ethanol mixtures. The results are briefly commented below, for discussions we refer to Section 5.2, page 88.

Temperature:

Table 6.1 show that the average temperature is very close to the desired temperature of 293.15K for all mixtures. Also the NVE simulation departs by only 4K from this temperature. We notice that all average NVT temperatures are slightly lower than the predefined, as were also observed for the methanol-water mixtures.

Figures D.1 and D.2 show translational and rotational temperatures for water and ethanol separately in a mixture of 0.75 mole fraction ethanol in water. Particularly the rotational degrees of freedom for water have an irregular behaviour. This is not particular to this mixture, but was seen for whichever component had the least number of molecules. It can be an indication of a heat bath with a frequency that is too small (ie. relaxation parameter = period is too large). Consequently Q is too large and the temperature control to weak, see Subsection 2.5.3, page 40.

For all mixtures the temperatures for the four different types of degrees of freedom show a slight difference, see Table D.1. The difference is largest with the equimolar mixture, while the rest of the simulations keep the rotational and translational temperatures within 1K of each other.

Pressure: All mixture pressures are stable, but the fluctuations are large. A tendency of negative pressures to occur more frequently than with the methanol-water mixtures is observed. The averages are large, see discussion in chapter 5.2, but lower than with methanol-water. There is also a marked difference between the pure water simulation and the rest of the simulations.

Table 6.1

Mixture temperatures T_{mix}, pressures P, and potential energies E_p from molecular dynamics simulations of water-ethanol mixtures. Numbers in italics is standard deviation 1σ. Experimental densities from Perry [125]. Jorg is Monte Carlo calculations of Jorgensen [27] with OPLS model included internal rotation at 298K.

x_e		ρ [g/cm³]	T_{mix} [K]	P [MPa]	$-E_p$ [J/g]	$-E_p$ [kJ/mol]
0.0	**NVT**	0.9982	292.92 *(0.067)*	240.6 *(2.3)*	-2323.92 *(0.89)*	-41.88
0.25	**NVT**	0.9226	292.97 *(0.081)*	40.1 *(1.7)*	-1688.56 *(0.60)*	-42.27
0.50	**NVT**	0.8631	293.03 *(0.118)*	14.4 *(1.3)*	-1294.56 *(0.50)*	-41.48
0.75	**NVE**	0.8219	296.93 *(0.29)*	11.8 *(1.1)*	-1043.73 *(0.19)*	-40.76
	NVT		293.08 *(0.16)*	17.5 *(1.1)*	-1051.46 *(0.50)*	-41.06
1.00	**NVT**	0.7893	293.13 *(0.21)*	26.5 *(1.0)*	-888.47 *(0.29)*	-40.92
	Jorg	0.785	298	0.1		-39.54±0.1

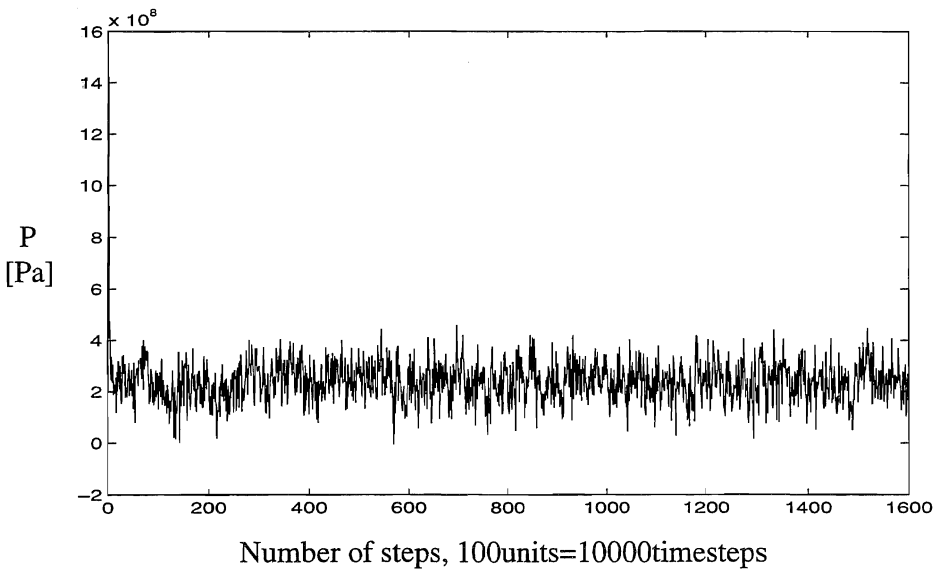


Figure 6.1 Instantaneous pressure P in a NVT simulation of pure water at 293K.

Enthalpy: Estimates of the enthalpies of vaporization are found from configurational energies with the aid of Equation (5.2), page 92. We have neglected corrections arising from the difference between enthalpy of real and ideal gas, and the differences in intramolecular rotational energies in gas and liquid. For pure OPLS-ethanol this amounts to a net energy of 0.2kJ/mol, to be subtracted, Jorgensen, 1986 [27].

Table 6.2
Calculated enthalpy of vaporization for ethanol-water mixtures at 293K compared to experimental results of 1) Smith and Van Ness [130] at 293K, 2) Marcus [96] at 298K, and simulation results of 3) Jorgensen [27] at 298K.

	x_e=0.0	x_e=0.25	x_e=0.50	x_e=0.75	x_e=1.0
ΔH^{vap}(calc) [kJ/mol]	44.26	44.64	43.86	43.44	43.30
ΔH^{vap}(exp) [kJ/mol]	44.22[1)]				42.30[2)]
MC-simulation	44.60[3)]				41.80[3)]

We see that the water value agrees very well with experiments at this temperature to. The liquid ethanol value is a little too high, even when we take the temperature difference into consideration. From the specific heat of liquid ethanol at 25°C, c_p=0.113kJ/molK [78], we find an enthalpy difference $\Delta H \sim \Delta E_p$=-0.57kJ/K for a decrease in liquid temperature of 5K. The 'experimental' value of ΔH^{vap} would then be found near 42.7kJ/mol, and the experimental configurational energy close to -40.5kJ/mol. Notice also that our molecular dynamics results for both methanol and ethanol are higher than Jorgensen's Monte Carlo results [27] for the same potential. See the discussion in Section 5.2.

We have also calculated the excess configurational energy for each mixture from Equation (5.3). They are compared to experimental values for excess enthalpy in Table 6.3. Experimental results show that ethanol-water mixtures at room temperature have a negative deviation from the ideal solution at a mole fraction ethanol near 0.15. At temperatures above ~50°C a positive deviation starts to develop. Our calculated values show a minimum of reasonable value at x_e=0.25, but at the highest mole fraction we also find a maximum. See however the discussion on page 145.

Table 6.3
Calculated excess potential energies E_p for the mixtures compared to experimental values for excess enthalpies H^E. x_e is ethanol mole fraction. All values at 293K.

	x_e=0.25	x_e=0.50	x_e=0.75
E_p(calc), [kJ/mol]	-0.63	-0.08	+0.10
H^E(exp), [kJ/mol]	-0.75	-0.47	-0.27

6.2 Theoretical analysis

The majority of the quantities analysed in this section show a strong similarity with the corresponding results for water-methanol mixtures. We therefore refer to Section 5.3. Additional figures are found in Appendix D.

Translational order: As with the methanol-water mixtures, the lattice is seen to melt well within the equilibration period in all simulations. We do however notice that mixtures with increasing ethanol content need more timesteps to get disordered than the corresponding methanol-water mixtures. The effect is not visible with pure water, so it is probably not the 5K temperature difference that is responsible. Both NVT and NVE show the same behaviour. An example is provided in Figure D.3.

Drift in total energy: Estimated drift in total energy is shown in Table 6.4, and we see that it is of the same order as the methanol-water NVT simulations (Table 5.4, page 96). We also observe the same trend of decreasing drift with increasing water content as for the methanol-water mixtures. Typical example of fluctuations in total energy is shown in Figure D.4. Figure 6.2 below show progress of total energy for whole simulation for all mixtures.

Plots of extended system energy and its contributions for selected systems are shown in Figures D.5, D.6, and D.7. The 0.75 mole fraction ethanol mixture displayed in Figure D.6 provides the only example where potential energy is not satisfactorily constant, but is clearly decreasing by ~25kJ/g during the production period. However, as system potential energy gradually decreases, reservoir potential energy increases by the same amount, and the total energy is kept constant. We see that conserved total energy and constant temperature are not sufficient requirements for a stable simulation. The decrease in configurational energy for this mixture imply that the results for this mixture must be questioned. Results will nevertheless be presented along with the rest of the mixtures. We doubt that this is an effect of the initial configuration, since

the same decrease in configurational energy is not found from the NVE simulation, Figure D.7, which starts from the same initial conditions. There is only a small decrease of 5kJ/g in the configurational energy from the NVE simulation. This is accompanied by an increase in total kinetic energy.

Table 6.4

Energy drift pr. step during production phase of NVT simulation - a total of 135000 steps. One example of NVE simulation also included. The drift is only a rough estimate of difference between maximum and minimum values.

x_m	ΔE_{ext} [J/(g step)]
0.0	$8.1 \cdot 10^{-5}$
0.25	$1.5 \cdot 10^{-5}$
0.50	$6.7 \cdot 10^{-6}$
0.75 ***NVE***	$5.2 \cdot 10^{-6}$ $3.7 \cdot 10^{-6}$
1.00	$3.0 \cdot 10^{-6}$

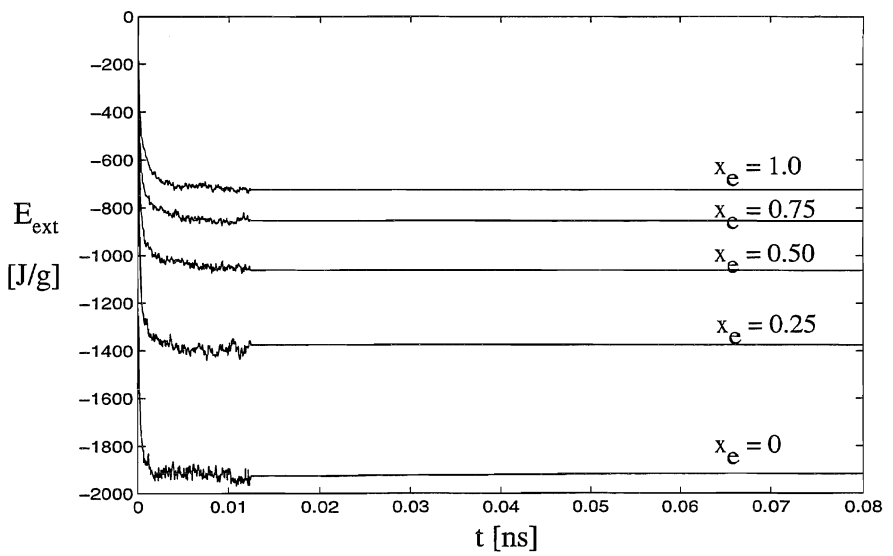


Figure 6.2 Conservation of extended system energy E_{ext} in NVT-simulation. x_e is ethanol mole fraction.

Velocity distribution: Looking at Figures. D.9 - D.15, we find that velocity distribution for the water molecules in the x_e=0.75 mixture (Figure D.11) show a peculiar behaviour near average velocity, particularly with the NVT simulation. There is of course a large statistical uncertainty with the distribution of the 64 molecular velocities, but comparing ethanol and water, we find that the distribution for the 64 ethanol molecules (Figure D.9) is closer to the expected one.

Conservation of linear momentum: Linear momenta are conserved in all directions, and their values are of the same order as for methanol-water mixtures, compare Tables D.2 and C.3 for methanol-water. See also Figure D.16.

Conservation of angular momentum: As expected, angular momenta is not conserved, but fluctuate with large amplitudes about zero, just as for methanol-water. Compare Tables D.3 and C.6 for methanol-water. See also Figure D.17.

Equipartition of energy: The kinetic energies are equally partitioned in each coordinate direction, see Tables D.2 and D.3.

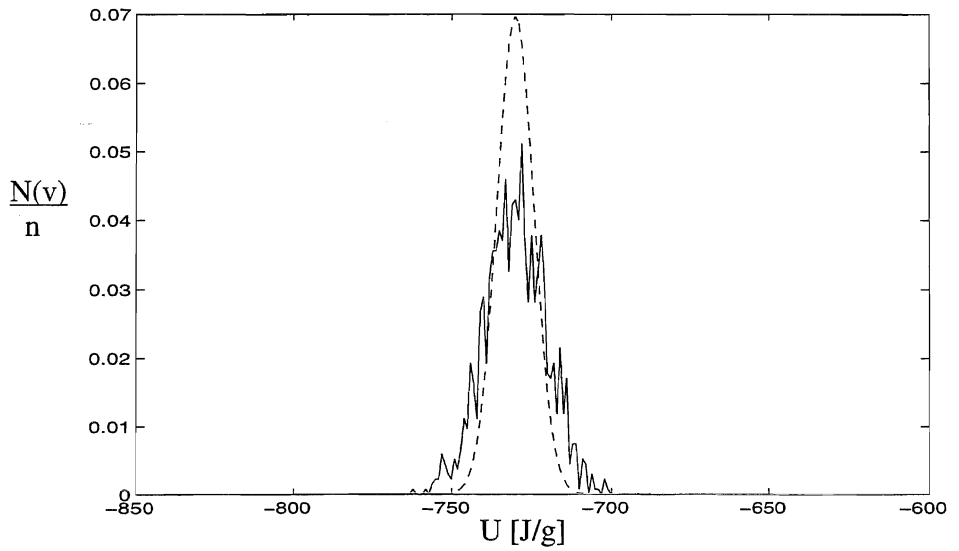


Figure 6.3 Distribution of internal energy, U, in NVT-simulation of pure ethanol. Dashed line is canonical distribution at 298.15K. See text of Figure D.18.

Distribution of internal energies: The fluctuations of internal energy is systematically overestimated for all mixtures, Figure 6.3 above, and Figures D.18 - D.21 This was also the observation for methanol-water mixtures.

Moments of kinetic energies: Inspecting Figures D.22 - D.31, we find the same kind of behaviour as with the methanol-water mixtures: The moments decrease/increase or fluctuate irregularly during the first half of the production period, then they seem to approach a limiting value. Two clear exceptions are found. The moments of translational energy of the x_e=0.5 and x_e=1.0 mixtures, Figures D.24 and D.26, continue to decrease at the same rate also in the last part of the production period.

Examples of the moments of reservoir kinetic energies are shown in Figures D.32 and D.33. They all seem to approach a limiting value, but this value is not unity as would be expected (all moments are divided by their canonical values).

Friction parameter η: Table 6.5 show averages of friction parameters for the translational and rotational heat baths. Theoretical average is zero. Averages of the sums of friction parameters are always found to be closer to zero than their separate averages. Note that friction parameters for ethanol-water mixtures are closer to zero than for the methanol-water mixtures, cf. Table 5.5, page 101.

Table 6.5
Coarse grain averages of translational and rotational friction parameters from the NVT simulations of water and ethanol at 293K. Standard deviation in parenthesis.

	x_e=0	x_e=0.25	x_e=0.5	x_e=0.75	x_e=1.0
$\eta_{trans}[ns^{-1}]$	-1.6 (17.4)	-1.7 (17.3)	-1.6 (18.5)	6.4 (20.3)	3.4 (19.2)
$\eta_{rot}[ns^{-1}]$	2.4 (13.8)	2.9 (18.3)	1.0 (21.7)	-3.7 (22.8)	-2.5 (22.0)

Figures D.34 and D.35 show friction parameters for translation and rotation for all mixtures. Decreasing amplitude of fluctuations with decreasing water content is due to increased relaxation parameters in mixtures τ_{trans} and τ_{rot}, see Table 3.1 and Equation (3.14), page 58. Table 6.7 give average transferred energy pr. step, and average reservoir kinetic energy of all mixtures. We see that total kinetic energy of reservoirs is underestimated also with water and ethanol by 15-20%. An example of reservoir kinetic energy is shown in Figure D.8.

Reservoir energies and s-variables: Reservoir potential energies, see Equation (2.29), for all mixtures are presented in Figure 6.4 and Table 6.7. Heat bath parameters, s_{trans} and s_{rot}, for all mixtures are shown in Figures D.36 and D.37, and their averages are given in Table 6.6 below.

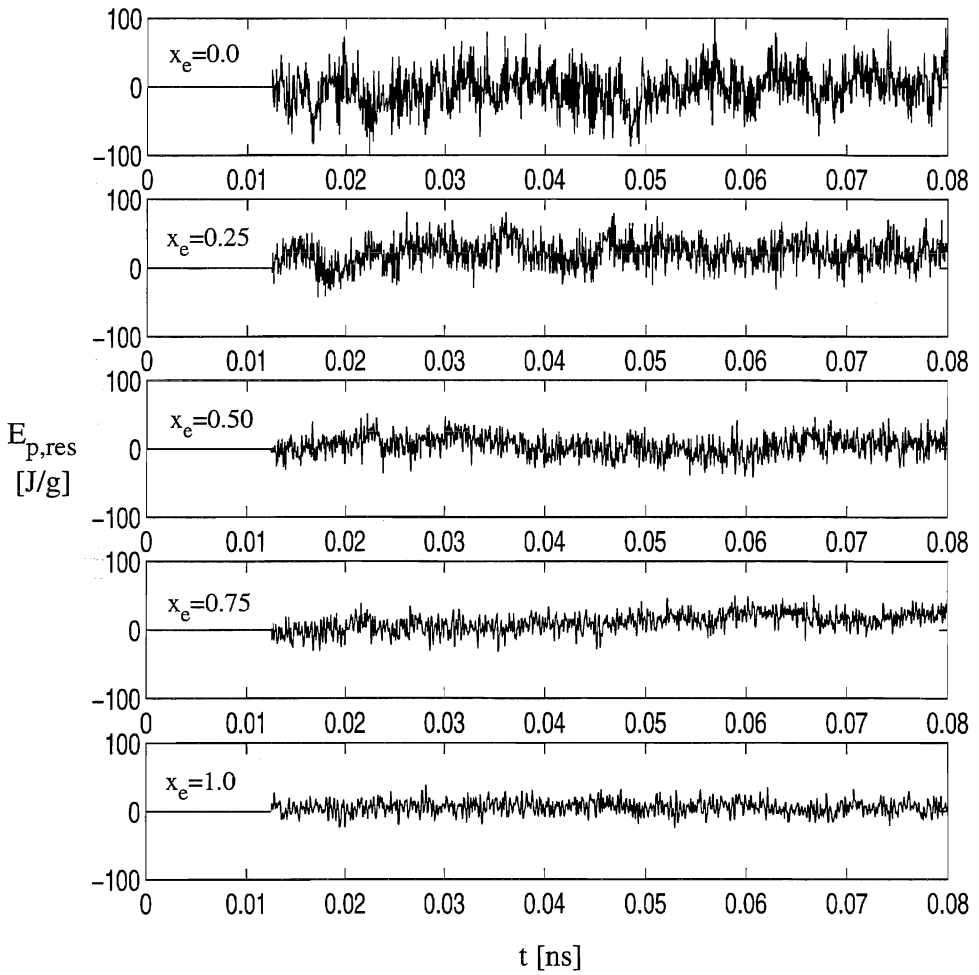


Figure 6.4 Reservoir potential energy $E_{p,res}$ for NVT simulations of mixtures of ethanol and water. x_e is ethanol mole fraction.

The departures of the reservoir potential energies from the theoretical values of zero are of the same order as with the methanol-water mixtures. The x_e=0.25 mixture shows the largest deviation from zero, this system has had energy in excess to dissipate to the sink. Peculiar though, is that this mixture has one of the two only nearly symmetrical pairs of heat bath parameters, see Figure D.37. On the other hand is reservoir potential energy of this mixture marked by slow irregular oscillations superimposed on the fluctuations. The same behaviour is also seen with the pure water system. Such oscillations are not found with the *s*-parameters or their sum, so it is probably an effect of adding together the two logarithmic energies with different amplitude of fluctuations. Note also that because s_{trans} and s_{rot} are not perfectly symmetrical about 1, their sums can not be perfectly stable either. Finally note that the reservoir potential energy of the x_e=0.75 mixture diverges slowly due to the gradual decrease of the system configurational energy, see Figure D.6.

Table 6.6

Coarse-grain averages of s-parameter for translational and rotational reservoirs for simulations of water and ethanol at 293K. Standard deviation in parenthesis.

	x_e=0	x_e=0.25	x_e=0.5	x_e=0.75	x_e=1.0
s_{trans}	0.749 (0.007)	0.891 (0.004)	0.943 (0.008)	1.625 (0.020)	1.308 (0.014)
s_{rot}	1.379 (0.012)	1.223 (0.006)	1.114 (0.008)	0.702 (0.009)	0.827 (0.008)

Finally observe that the x_e=0.5 and 1.0 mixtures which had the poorest approach to canonical moments, are the mixtures that satisfy the other criteria best.

Table 6.7
Average potential and kinetic energy of the heat reservoirs, and average transferred energy pr. step. Standard deviations in parenthesis. Averages are based on output values with 4 significant digits taken each 100th step.

x_e	Reservoir pot. energy [J/g]		Transf. energy pr step [J/g]		Reservoir kin. energy [J/g]	
	average	expected	mean	expected	mean	expected
0	-1.8 (25.7)	0	-0.002 (0.4)	0	0.43 (0.50)	0.53
0.25	18.4 (19.2)	0	0.001 (0.2)	0	0.31 (0.37)	0.38
0.50	3.9 (13.9)	0	-0.0001 (0.1)	0	0.25 (0.28)	0.30
0.75	8.9 (13.1)	0	-0.0001 (0.1)	0	0.21 (0.26)	0.24
1.0	4.5 (9.0)	0	-0.00004(0.08)	0	0.18 (0.20)	0.21

6.3 Structural properties of ethanol/water mixtures.

6.3.1 Structure of bulk ethanol

Harvey, 1939 [11] found from X-ray diffraction that each ethanol molecule was hydrogen bonded to about 2 neighbours with a hydroxyl distance of 2.9Å. The results of Narten and Habenshcuss, 1984 [9], confirm this. They refine the distance to 2.8Å, and the number of hydrogen bonds to 1.8 ±0.1.

Jorgensen, 1981, [110] performed simulation with the TIP potential for 128 ethanol molecules with internal rotation. He found winding chains and evidence of branching, like with pure methanol. From analysis of bond-energies, he found that each ethanol monomer is hydrogen-bonded to 1.75 neighbours, consistent with experiments. He also found the majority (~60%) of the monomers to be engaged in 2 bonds, and an estimated average chain length of 5-6 monomers. By analogy to solid ethanol, the chains in liquid ethanol, are believed to consist of U-shaped chains with alternating gauche and trans molecules as illustrated in Figure 6.5. This is unlike methanol, where peak positions are consistent with a V-shape, as shown in Figure 5.10. Vahvaselkä *et al.*, 1995 [12] suggest on the basis of X-ray scattering data that ethanol, like methanol, consists of irregular chains varying in length with an average of 10 hydrogen bonded monomers.

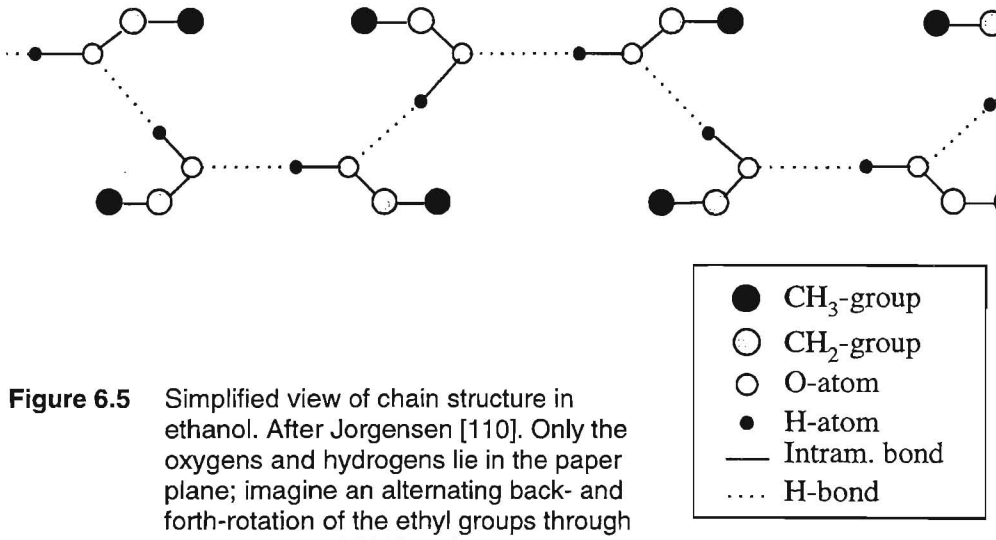


Figure 6.5 Simplified view of chain structure in ethanol. After Jorgensen [110]. Only the oxygens and hydrogens lie in the paper plane; imagine an alternating back- and forth-rotation of the ethyl groups through rotation around OH-bond.

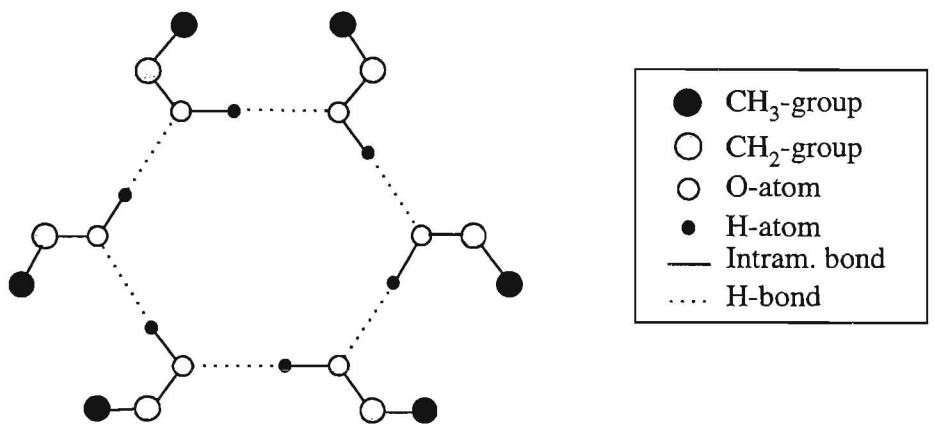


Figure 6.6 2-dimensional view of hexamer hydrogen bonded ring in ethanol. After Sarkar and Joarder,1993 [170]. The ethyl groups are not in the same plane.

Recently, Sarkar and Joarder, 1994 [170] from comparison of the experimental structure functions from [9] to a structure function for a hexamer model, suggested that there is 'on average a good possibility of hexagonal clustering of ethanol monomers in liquid state at room temperature', see Figure 6.6. This model give however to many two-bonded monomers compared to simulation results [27, 110]. Their similar analysis of U-shaped tetramer and closed tetramer structure gave less accordance with experiments.

6.3.2 Structure of ethanol-water solutions.

A prominent feature of the mixing of ethanol with water is the volume contraction upon addition of very small amounts of ethanol. The minimum volume occurs at mole fraction x_e=0.07. It has been suggested [5, 171] that the behaviour of ethanol-water mixtures over the whole range of composition is not due to one mechanism alone.

Very recently Nishi *et al.*, 1995 [172] have performed IR absorption measurements, mass spectrometry, and X-ray diffraction experiments on ethanol-water mixtures in the concentration range $0.001<x_e<0.03$.

They propose a structure for the dilute region where the solution is composed of several ethanol molecules stacked in a hydrophobic core and surrounded by a cage of strongly (linearly) hydrogen-bonded water molecules. Bulk water with weaker (linear and bent) hydrogen bonds is surrounding this hydrate structure. The motion of these large clusters might affect bulk water structure. This corresponds to the findings of Soper and Finney [157] for dilute water-methanol mixtures.

They also emphasize the resemblance of the structure factor of the dilute aqueous ethanol solution to that of bulk water at high pressures.

Matsumoto *et al.*, 1995 [173], have studied mixtures in the range $0.2<x_e<0.8$ with mass spectrograpy and X-ray diffraction. They find ethanol polymer formation and hydrate clusters to dominate. Water molecules do not prefer to bond to each other, but might act as stabilizers for the ethanol structure. They propose a 'sandwich-structure' where layers of ethanol polymers is separated by bridging water molecules. From their figures, it seem like the structure is composed of V-chain polymers of the *trans*-conformers. With all O-O distances fixed at 2.8Å, they have calculated a number of model distances which they find to be consistent with measurements, see Table 6.8, page 172.

Only very dilute mixtures are studied with Monte Carlo simulations. Alagona and Tani, 1982, 1988 [29, 31] have simulated a mixture of one *trans*-ethanol molecule and 123 and 215 water molecules with an *ab initio* potential for the cross interactions. They find water molecules to form a cage around the ethyl group of ethanol. On the

average two water molecules are hydrogen-bonded to the hydroxyl group. Lily *et al.* [32] have applied TIP4P and OPLS in a simulation of 1 ethanol molecule and 108 water molecules. They find 2.3 water molecules hydrogen-bonded to each ethanol monomer. Even if these calculations are far from the concentration we study, we expect to recognize some common trends with the pair-correlation functions.

It is not clear which structures dominate in liquid ethanol or its aqueous mixtures. It is though a general belief that internal rotation of the methyl group (or hydroxyl hydrogen) is likely, and that both *trans* and *gauche* conformers are present also in the liquid as they are in the gas and solid. Their ratio is not known, but is estimated from Monte Carlo simulations by Jorgensen [27] to nearly 1:1. The hexamer ring and U-shaped chain are suggested with reservations as bulk liquid structures. Central to both structures is the alternation of *trans* and *gauche* molecules.

As already mentioned on several occasions, our model liquid contains only the *trans* conformer. It is theoretically possible to construct both a hexamer ring and a U-shaped chain from only trans conformers, see Figure 6.7 and Figure D.38. We believe that particularly the chain would be more winding than seen from the figure, and probably subject to more branching than the trans-gauche chain. Note that in the trans-hexamer, only the CH_3-site is moved relative to the trans-gauche ring, while in the trans-chain, both CH_2 and CH_3 is moved. For *trans*-molecules it is perhaps more naturally to extend the V-chain for methanol by appending ethyl groups instead of methyl groups, as shown in Figure 6.29.

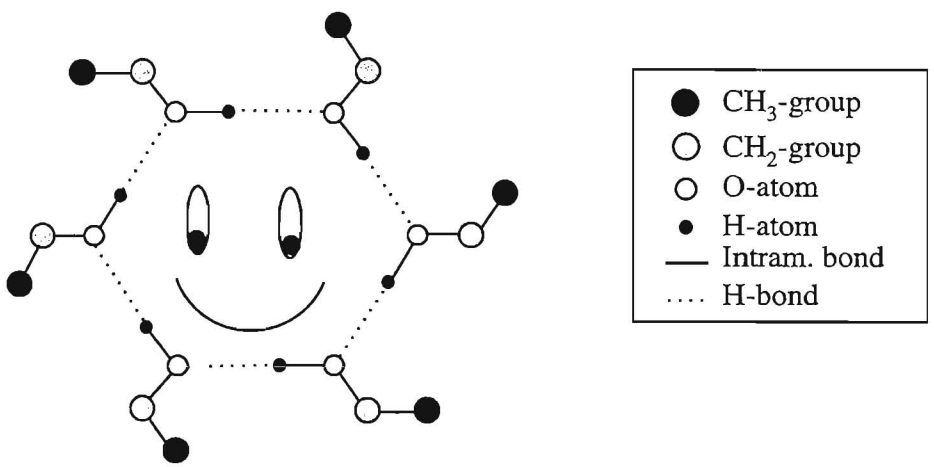


Figure 6.7 This cute little creature show how a hexamer ring can be composed only by trans conformers. This is not a suggestion for a new ethanol structure, but only visualize which pair correlation function will be different from experimental if a hexamer structure for *trans*-ethanol is found.

All attempts to visualize structure are oversimplified as much as they only show a two-dimensional possibility. It is left to the reader to vary angles and orientation to imagine more realistic structures. Calculated site-site distances are also based upon planar structures, but will hopefully give us ideas of ranges and limitations of correlations.

From the discussion above, we expect that the simulation of only trans-ethanol should have little effect upon the peak-positions for correlations containing combinations of oxygen and hydrogen. To some extent also the positions of peaks involving CH_2-groups are expected to be unaffected. The CH_3-interactions are on the other hand expected to disagree with experiments. Unfortunately, we have not found published experimental data concerning the ethyl group, but we might also see differences compared to the MC simulations of Jorgensen [27] regarding these correlations.

6.3.3 Self-correlations for water

The overall appearance of the self-correlations is quite similar to the water self-correlations in water-methanol mixtures. Nearest water neighbours to a central water molecule are not affected by the presence of ethanol in solution instead of methanol. Some differences can be seen beyond 1st minimum, particularly with increasing ethanol content. The water-water self correlations are thoroughly discussed in Subsection 5.4.2. We shall here only point out a few observations and some differences from the methanol-water mixtures. The site-site correlation functions are presented in Figures 6.8 - 6.10 and the coordination numbers appear in Table 6.9, page 176. The positions of maxima and minima are given in Tables D.4 - D.6. Self-correlations for water in the most ethanol-rich mixture from the NVE and NVT simulations are compared in Figures D.39 - D.41.

- Intensities of first maxima are higher than for the methanol-water mixtures, but only 0-5% less than expected from density ratios of the two mixtures.
- The positions of maxima and minima are in most cases not altered significantly (ie. more than 0.5Å), except for the 2nd maximum of g_{OO} for the x_e=0.5 and 0.75 mixtures.
- Water-water site coordination numbers are close (lower than or equal) to those found for methanol-water mixtures.
- The 5K difference in temperature does not seem to influence peak positions and heights for pure water significantly. We then assume that differences between water-methanol and water-ethanol pair correlation functions are not due to temperature.

- For the oxygen-oxygen and oxygen-hydrogen correlations the 1st and 2nd peaks are slightly broader with the two highest concentration than for the corresponding water-methanol mixtures.
- For the hydrogen-hydrogen correlations the depth of 1st minimum is less than with water-methanol mixtures.
- For the x_e=0.75 mixture, the 2nd minimum of all correlations are broader, deeper and at larger separations. This might due to the larger size of intervening ethanol molecules.

Regarding the final item of the list above, we found in Section 6.2 that this particular mixture was less stable than the rest of the mixtures. We have compared the results from the NVT simulation with the results from the NVE simulation for the same mixture. The temperatures of the simulations differ by only 4K. We find but small differences out to the first minimum, as can be seen from Figures D.39 - D.41. The correlation numbers will therefore be less affected. From ~3Å to 8Å, the NVT simulation depresses all self-correlations, but there is however still a clear difference between the NVE results for the 0.75 mixture and the results for the rest of the mixtures in this region.

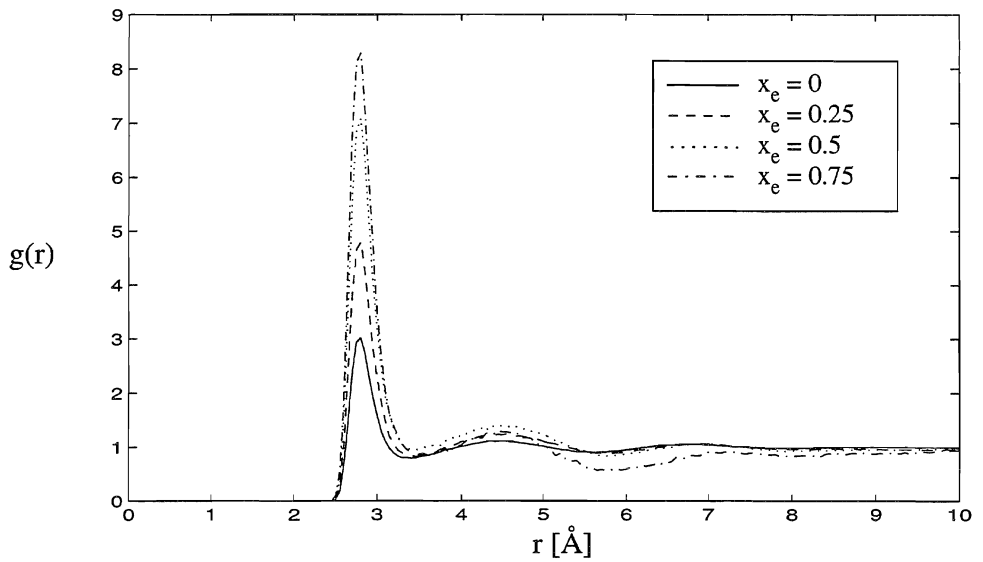


Figure 6.8 Radial self correlation function g(r) for O_w - O_w in water-ethanol mixture at 293K sampled with NVT simulation. x_e is ethanol mole fraction.

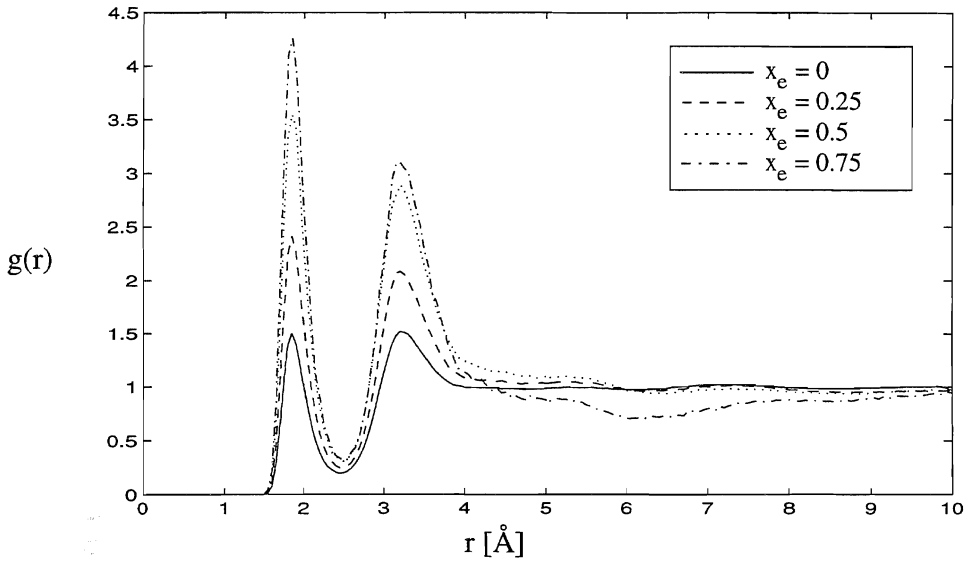


Figure 6.9 Radial self correlation function g(r) for O_w - H_w in water-ethanol mixture at 293K sampled with NVT simulation. x_e is ethanol mole fraction.

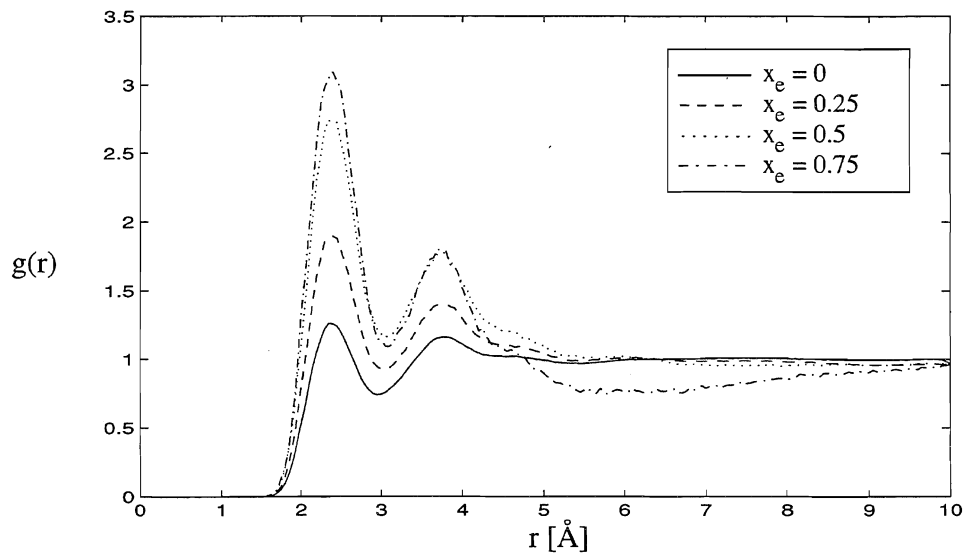


Figure 6.10 Radial self correlation function g(r) for H_w - H_w in water-ethanol mixture sampled with NVT simulation. x_e is ethanol mole fraction.

6.3.4 Self-correlations for ethanol

All site-site correlations for the self interactions of ethanol are presented below in Figures 6.11 - 6.20. The self correlations for ethanol are characterized by

- Except for the CH_3 and CH_2 correlations, all 1st peak intensities decrease with decreasing ethanol mole fraction.
- O-O, O-H, and H-H self-correlations for ethanol are very similar to the same correlations for methanol.
- Broad maxima for correlations between CH_3 and CH_2 groups indicate several contributions and varying orientation of ethyl group
- Bulk ethanol self-correlations for CH_3-CH_2, CH_3-O, CH_3-H, and CH_2-CH_2 all show clear and narrow peaks that diminish rapidly - and ultimately disappear - with increasing concentration.
- The oxygen-oxygen peak at 2.75Å agree with experiments [9, 12] Integration of 1st peak yield 2.1, slightly above experiment, but consistent with the calculations of Jorgensen, 1986 [27] who find 2.0, 0.97 and 2.1-2.2 for OO, OH, and HH respectively.
- The peak positions of oxygen-carbon and carbon-carbon agree nicely with the results of Jorgensen [110]. Largest disagreement is with the CH_3-CH_2 correlation, which is found at 0.2Å closer distance.
- The correlations for CH_3-CH_3, CH_3-CH_2, and CH_2-CH_2 all show broad and clear minima centred near 7Å. This is not in total agreement with the 'sandwich-model', if two layers should be present we find them to be separated by 8-9Å.
  Note however the discussion of effects of simulating only the trans-conformer, page 155.

CH_3 - CH_3 correlations:

The broadness of the unsymmetric first peak at 4.3Å, shown in Figure 6.11, indicate that several site-site contributions contribute. For pure ethanol there might be one dominant distance at 4.3Å and another near 5.1Å. The closest methyl-methyl distance in the trans V-chain is ~4.9Å (Figure 6.29, page 174) between non-hydrogen bonded neighbours. The closest distance can be intrachain contributions or caused by a winding chain. With decreasing content of ethanol, the largest distance becomes less pronounced and we observe a sharpening and tapering of the peak. This trend is also seen with methanol, but not nearly as clear. Along with the sharpening, the minimum and 2nd maximum are also pulled inward. The minimum near 7Å contradicts the 'sandwich-distance' found by [173]. If two layers should be present, their separation is about 9Å.

CH_3 - CH_2 correlations:

From Figure 6.12, we find a broad unsymmetric peak at 4.3Å and a narrow minimum at 6.5 - 7Å. A small peak at ~5.9Å for bulk ethanol gradually disappearing with decreasing ethanol concentration, indicate one distinct contribution. Notice the similarity with the CH_3-CH_3 correlations, not unexpected because of similarity in site-parameters.

CH_3 - O correlations:

Figure 6.13 show a small maximum at 3.9Å shrinking to a shoulder in aqueous mixtures. Also a sharp peak at 4.95Å is decreasing with decreasing ethanol concentration, while a small shoulder develops simultanously to the right of the second peak. The shape of the correlation function agree well with the TIP structure of bulk ethanol [110]. Also peak positions agree exactly. We do however find a substantially higher 2nd peak. The 1st peak integrates to 2.7 in pure ethanol.

CH_3 - H correlations:

From Figure 6.14, we see that bulk ethanol show three clear maxima at 3.15, 5.4, and 6.95. The first maximum disappears completely in the most aquatic solution, along with a decrease of the height of the 2nd maximum. This is parallell to the development for the CH_3-O correlations. The function show some kind of resemblance with the corresponding correlation for methanol, Figure 5.24, page 124. The TIP calculations of Jorgensen yield a nearly structureless pair correlation function for CH_3-H.

CH_2 - CH_2 correlations:

The peak at 4.65Å shown in Figure 6.15 seem to be a superposition of at least two distinct contributions of which the one at smallest separation dominate in bulk ethanol. Much of the structure beyond 1st peak is lost upon solution. The shape, height and peak position agree very well with the calculations of [27, 110].

CH_2 - O correlations:

A symmetric 1st peak at 3.7Å, which integrates to three for bulk ethanol, and a small broad maximum centred at 5.5Å are the dominant features of this correlation, see Figure 6.16. Positions of peaks agree once again very well with [110], but OPLS calculations [27] yield a coordination number of 'about two' for liquid ethanol, while we find exactly 3.0. 1st peak must be due to nearest neighbours. It decreases rapidly with decreasing ethanol mole fraction accompanied with an increase and broadening of 2nd peak.

CH_2 - H correlations:

Figure 6.17 show the two peaks arising from the hydrogen bond accepting neighbour and the hydrogen bond donating neighbour located at 3.0Å and 4.1Å respectively. The intensities of both peaks are drastically reduced with decreasing ethanol content. We observe a slight increase in intensity around 6.5Å as the 2nd peak disappears. The positions appear to agree with [110], but the height of 1st peak is significantly larger in our calculation, just as we found with CH_3-H above. However, Jorgensen, 1986 [27] and we agree that the 1st peak integrates to one for pure ethanol.

O - O correlations:

The curves shown in Figure 6.18 are very similar to the corresponding curves for methanol, with the 1st peak at nearly the same distance. Also the 1st minimum and the 2nd peak agree, though there is a slight variation with some of the mixtures, see Table D.5. The intensity of the 1st peak of bulk ethanol is significantly larger than for bulk methanol, and the intensity for x_e=0.25 is slightly lower than for the same mole fraction methanol. Also the 2nd peak is less pronounced with ethanol than with methanol; in other words - less structure beyond 1st minimum. The position of 1st peak agree well with both experiment [9] and simulation [27], and the area integrates to 2.1 for bulk ethanol, also in good agreement with experiment and simulation.

O - H correlations:

Figure 6.19 show a large peak at 1.85Å and a small peak at 3.35Å, both are close to methanol peak positions, and are also in good agreement with OPLS-ethanol calculations of Jorgensen [27]. 1st peak integrates to one as expected with only one donating neighbour. We notice the same differences with methanol in the intensities of 1st peaks as described for O-O correlations above. Also note that while the disappearance of second peak for ethanol is caused by a reduction in peak intensity, it is caused by a filling of the 2nd minimum for methanol mixtures.

H - H correlations:

Figure 6.20 show a sharp and symmetric 1st maximum just as for methanol, appearing at 2.4Å. The peak intensities of the two mixtures are clearly different, as described above. Notice the dissimilarity of bulk ethanol and bulk methanol beyond 1st minimum. Bulk ethanol is more structured than bulk methanol, but the structure is lost upon mixing.

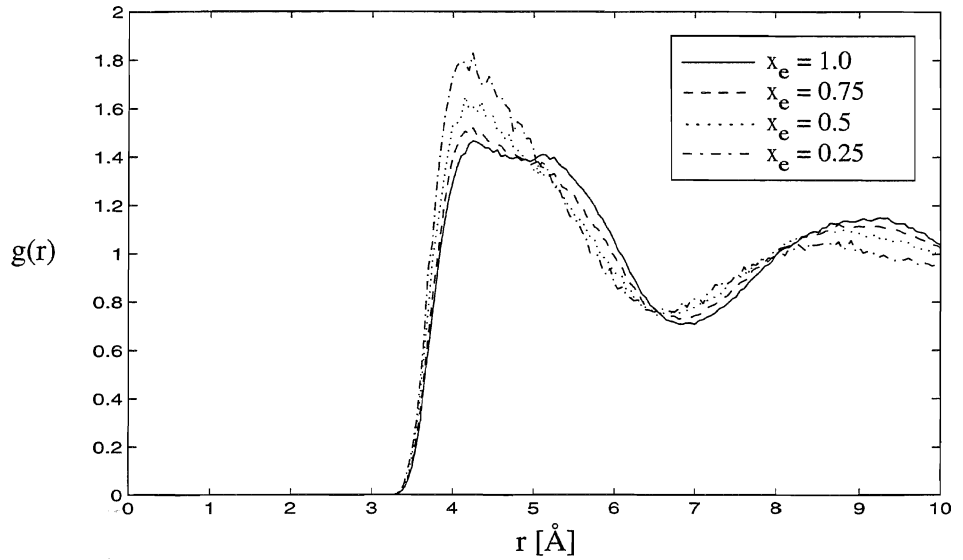


Figure 6.11 Radial self correlation function g(r) for CH_3 - CH_3 in water-ethanol mixtures sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

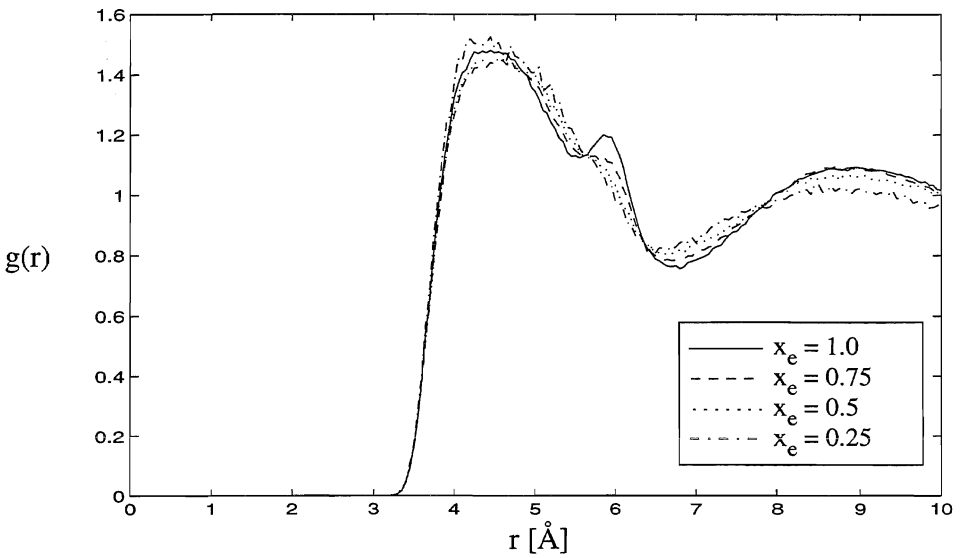


Figure 6.12 Radial self correlation function g(r) for CH_3 - CH_2 in water-ethanol mixtures sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

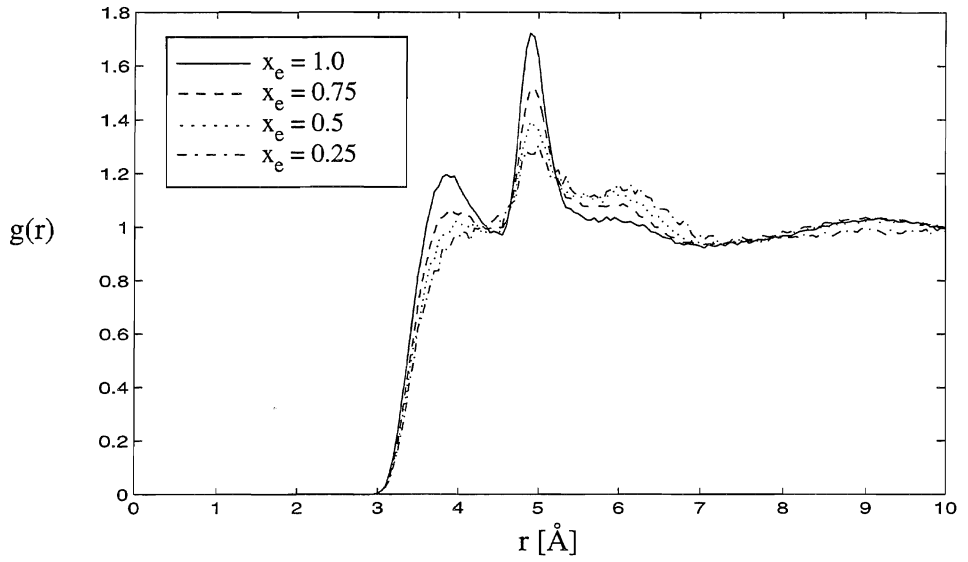


Figure 6.13 Radial self correlation function g(r) for CH_3 - O in water-ethanol mixtures sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

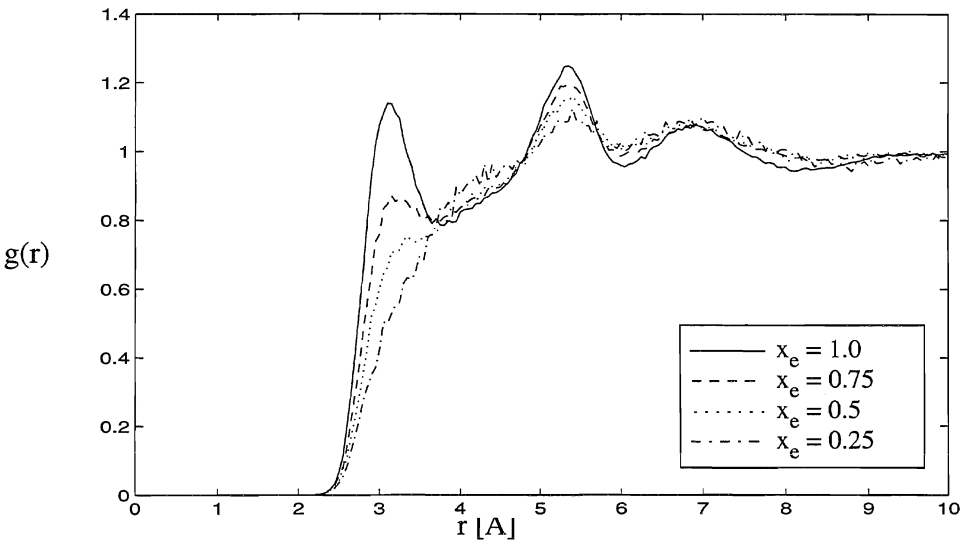


Figure 6.14 Radial self correlation function g(r) for CH_3 - H in water-ethanol mixtures sampled with NVT simulationat 293K. x_e is ethanol mole fraction.

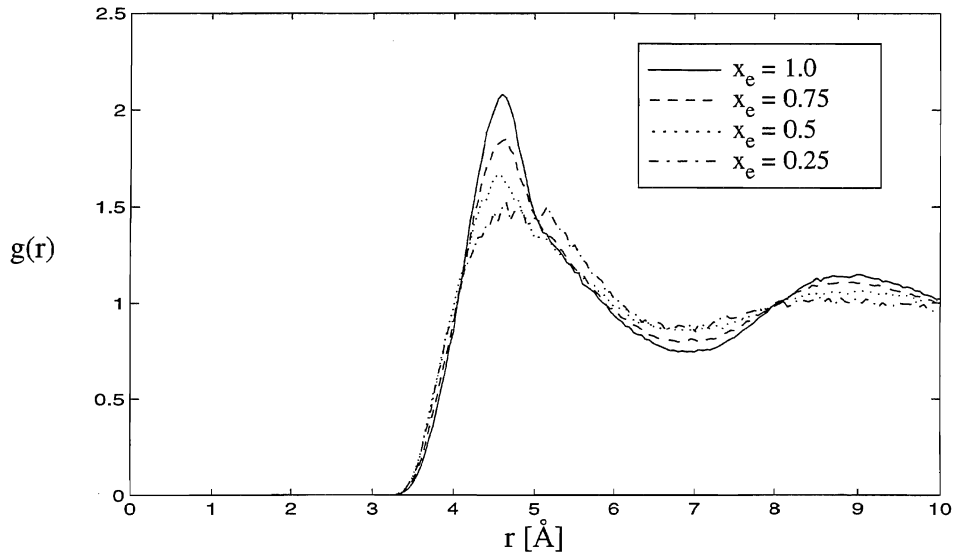


Figure 6.15 Radial self correlation function g(r) for CH_2 - CH_2 in water-ethanol mixtures sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

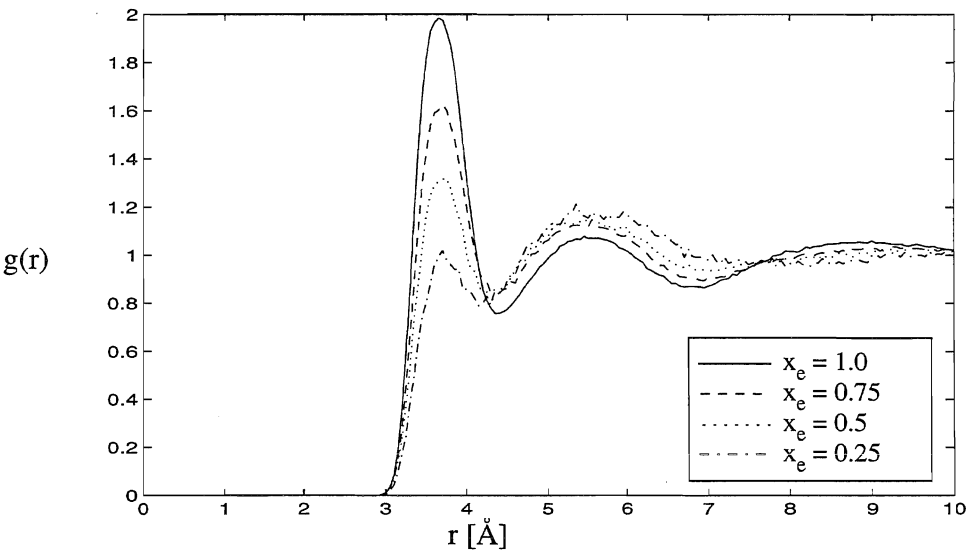


Figure 6.16 Radial correlation function g(r) for CH_2 - O_e in water-ethanol mixture sampled with NVT simulation. x_e is ethanol mole fraction.

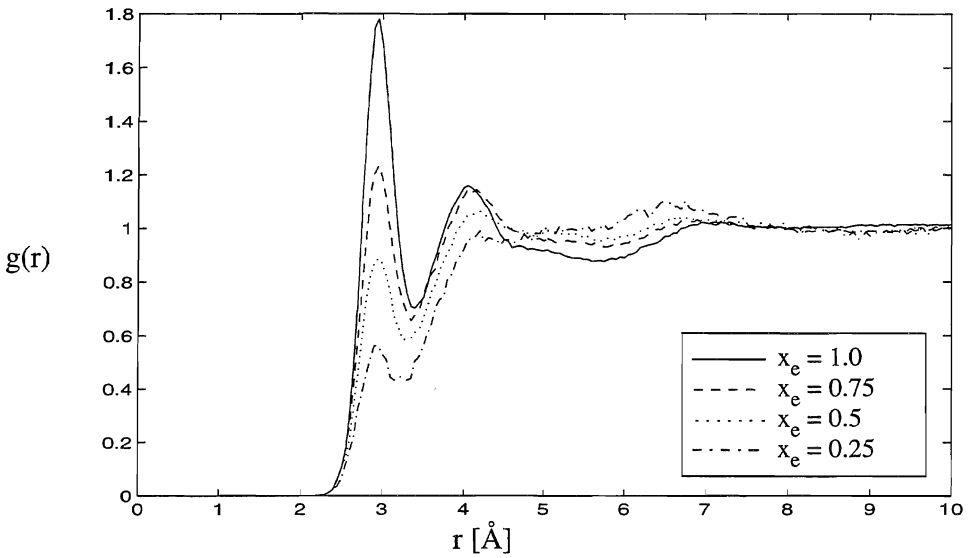


Figure 6.17 Radial self correlation function g(r) for CH_2- H_e in water-ethanol mixtures sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

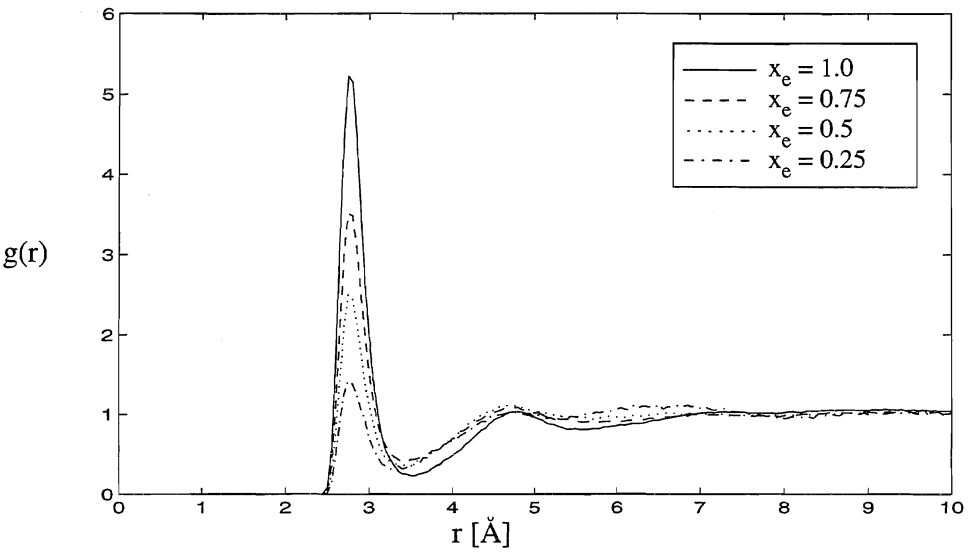


Figure 6.18 Radial self correlation function g(r) for O_e - O_e in water-ethanol mixture sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

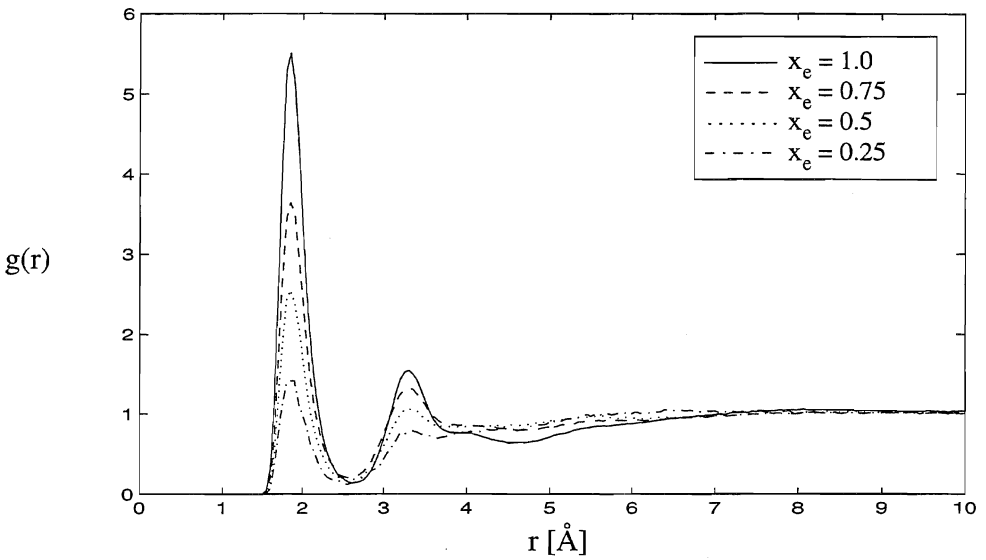


Figure 6.19 Radial self correlation function g(r) for O_e - H_e in water-ethanol mixtures sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

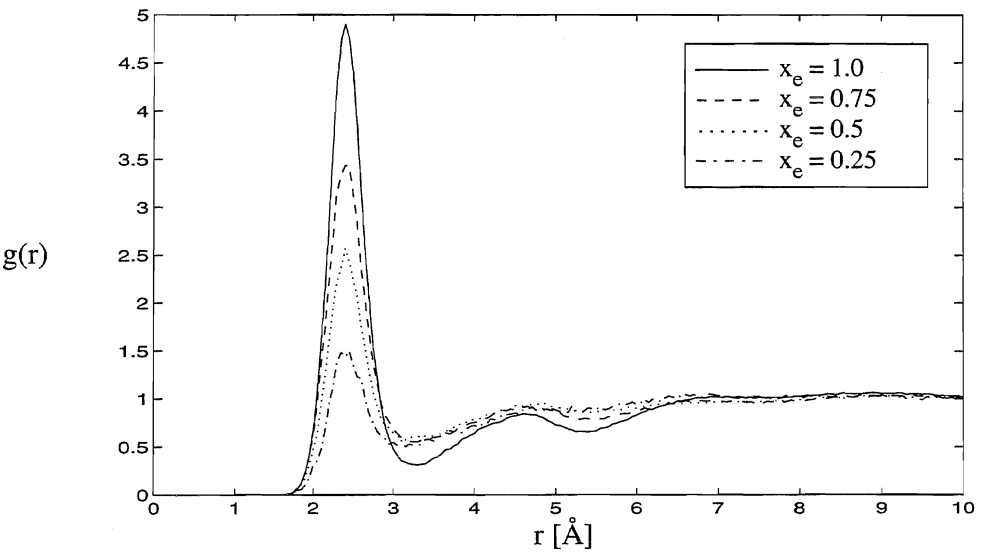


Figure 6.20 Radial self correlation function g(r) for H_e - H_e in water-ethanol mixtures sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

6.3.5 Cross-correlations for water and ethanol.

The cross-correlations are characterized by

- The peak heights increase with increasing ethanol content.
- The positions of maxima and minima are nearly invariant with composition.
- The peaks are higher than for methanol-water mixtures, but when corrected for the different number densities they vary within ±5% of the corresponding methanol-water mixtures. The exception is the x_e=0.25 mixture, where combinations of oxygen and hydrogen yield 1st peaks 10-15% higher than for the corresponding methanol-water mixture.
- There is a strong similarity between oxygen and hydrogen cross-correlations in water-ethanol and water-methanol mixtures.

O_w - CH_3 cross-correlations:

The two peaks at 3.75Å and 5.0Å seen in Figure 6.21 both decrease with decreasing ethanol content. We notice that the two peaks nearly coincide with the peaks of the CH_3-O self correlation, which also decrease with decreasing mole fraction ethanol. There is no similarity with the corresponding correlation function for methanol-water; the peak appears at a larger separation, and where there is a broad minimum in methanol mixtures, the ethanol cross-correlation functions show a 2nd peak.

Comparing to Alagona and Tani [31], who use a different model and a very dilute solution, we find that the position of the 1st peak agree fairly well. They find also a shoulder/small maximum at only a slightly shorter (~0.3Å) separation than our 2nd peak.

O_w - CH_2 cross-correlations:

The peak position at 3.7Å, Figure 6.22, correspond almost to the position of the peak in the dilute solution [31], but while the dilute solution show a clear shoulder to the right of the peak, we only find a small remnant. This correlation function has very little structure beyond the 1st maximum. Once again we observe that both CH_2-O self correlations, which we found at the same distance, and cross correlations for CH_2 decrease with decreasing ethanol content.

O_w - O_e cross-correlations:

As can be seen from Figure 6.23, the 1st peak appear at 2.75Å, exactly the same position as for cross-correlations of methanol and water. The heights are larger with ethanol, but the two alcohols show a strong qualitative similarity.

O_w - H_e cross-correlations:
These curves shown in Figure 6.24, are easily mistaken with the corresponding curves for methanol. The 1st maximum appear at the same position, 1.85Å, the minimum at 2.6Å and the 2nd maximum at 3.35, only slightly to the left of the 2nd minimum for methanol. The major difference is with the peak heights, which once again appear higher for both peaks. Also there is a complete lack of structure beyond 2nd maximum where methanol correlations show a diminutive 3rd maximum.

H_w - CH_3 cross-correlations:
This pair correlation, Figure 6.25, is nearly structureless at all composistions, except from a minimum at 6.15Å at all concentrations, and the two small maxima for the most ethanol-rich mixture. This is very unlike the methanol-mixtures. As seen from hydrogen on water, the methyl groups of ethanol are nearly randomly distributed.

H_w - CH_2 cross-correlations: The maxima at 2.95Å and at 4.2Å shown in Figure 6.26, might be compared to the self correlations of ethanol hydrogen and CH_2 at 3.0Å and 4.1-4.25Å. Note also the similarity with the corresponding cross correlation function of methanol mixtures.

H_w - O_e cross-correlations:
The correlation functions displayed in Figure 6.27, show a strong resemblance with the O_w-H_e correlation, with the O_e-H_e self correlation, and with the H_w-O_m cross correlation for water and metanol. Typically we find higher peaks than for methanol-water, but lower than with the O_e-H_e correlation.

H_w - H_e cross-correlations:
The 1st maximum at 2.4Å shown in Figure 6.28, corresponds to the maximum position for the methanol-water mixtures, but the intensity of the peak is higher as usual. There are also some minor differences with the 1st minimum beeing shallower and the 2nd peak clearer.

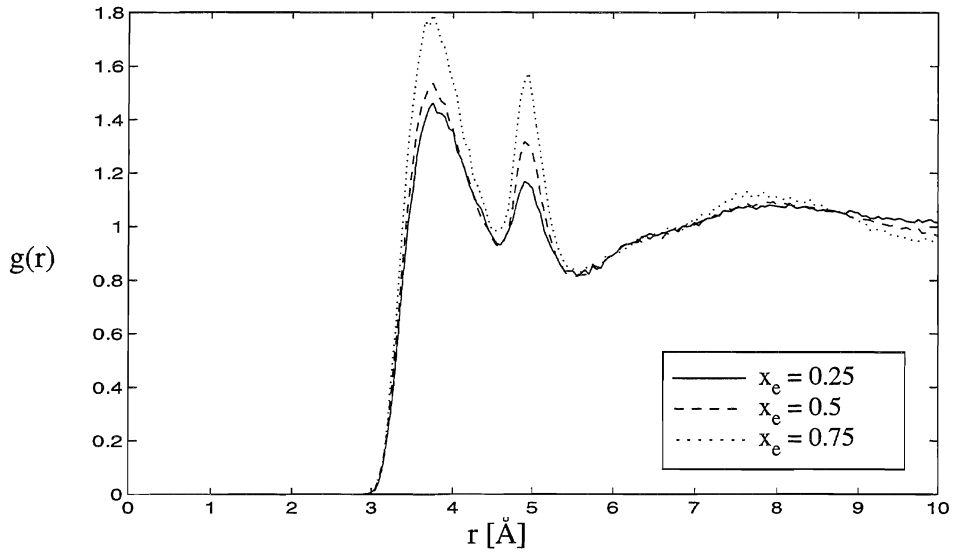


Figure 6.21 Radial cross-correlation function g(r) for O_w - CH_3 in water-ethanol mixtures sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

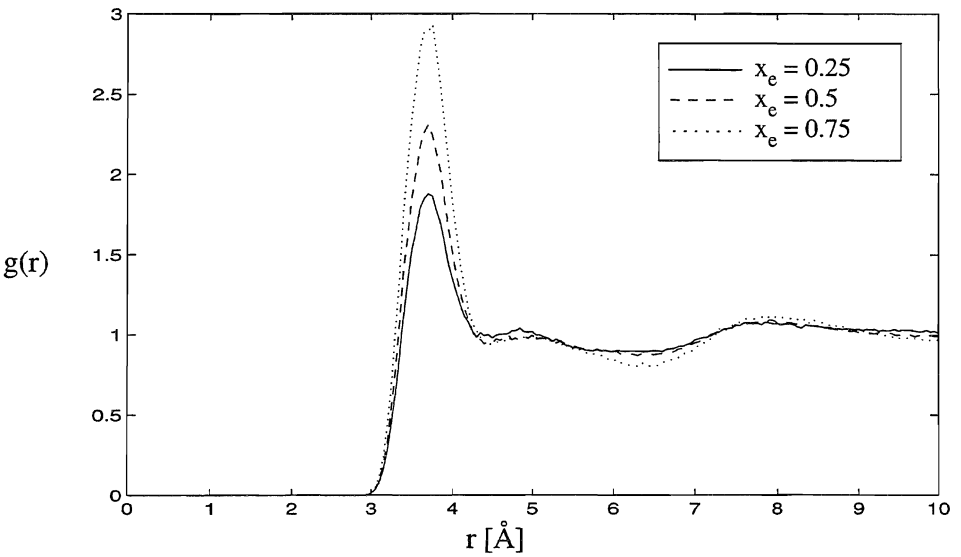


Figure 6.22 Radial cross-correlation function g(r) for O_w - CH_2 in water-ethanol mixtures sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

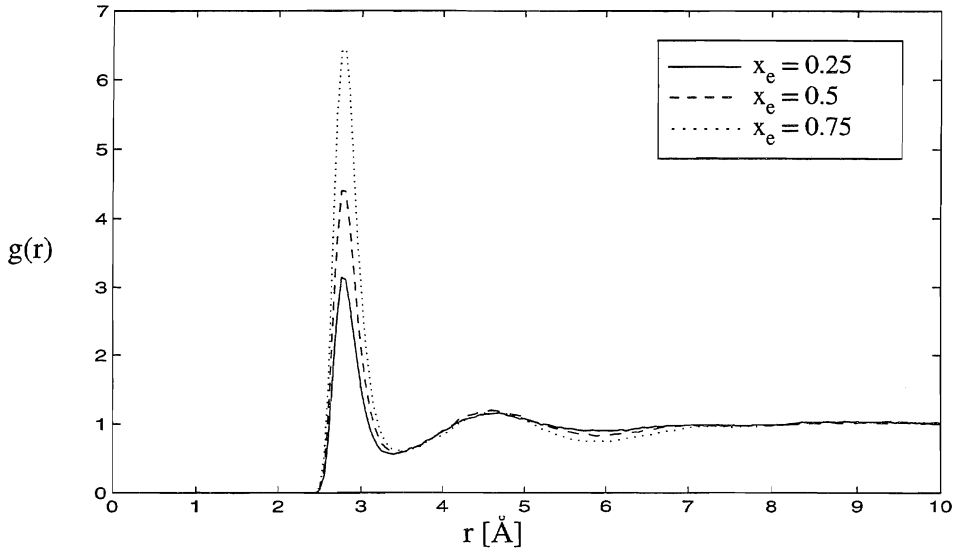


Figure 6.23 Radial cross-correlation function g(r) for O_w - O_e in water-ethanol mixtures sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

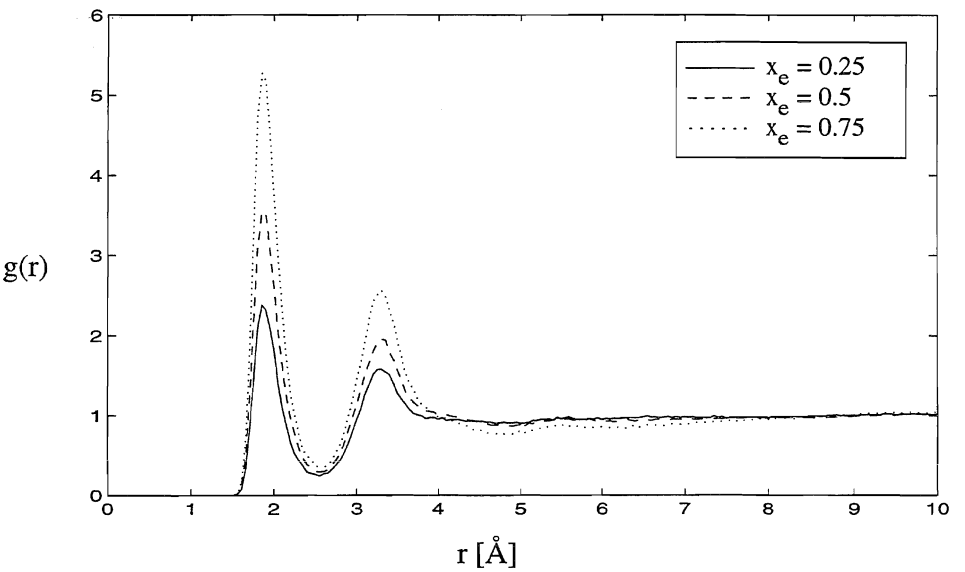


Figure 6.24 Radial cross-correlation function g(r) for O_w - H_e in water-ethanol mixtures sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

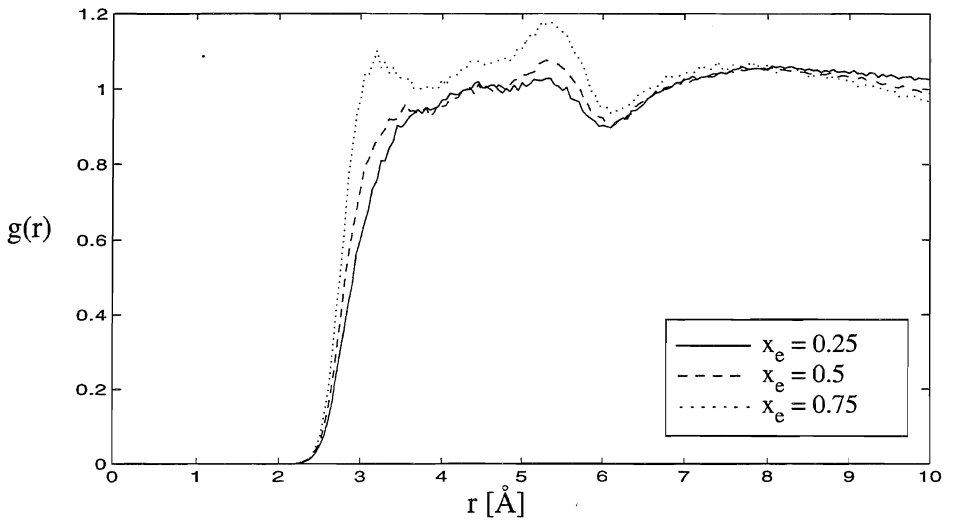


Figure 6.25 Radial cross-correlation function g(r) for H_w - CH_3 in water-ethanol mixtures sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

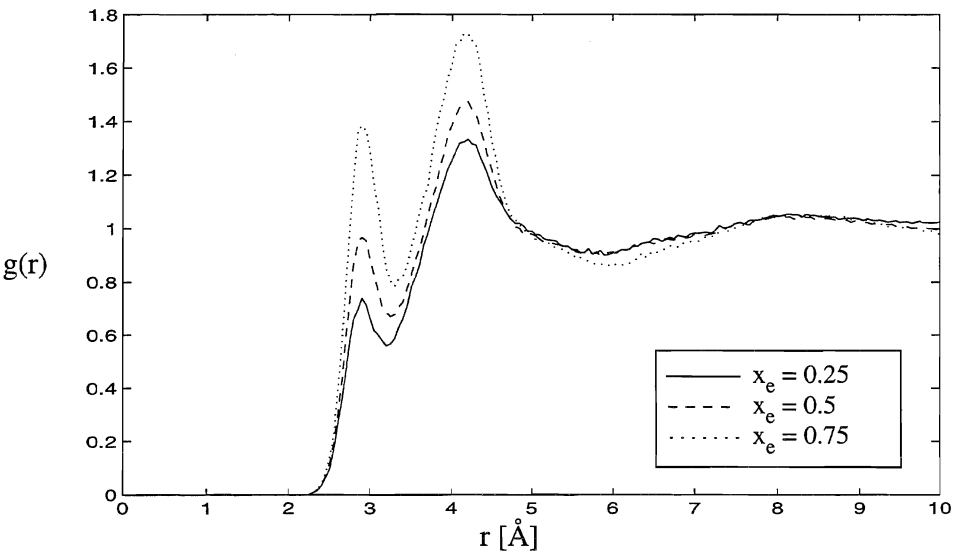


Figure 6.26 Radial cross-correlation function g(r) for H_w - CH_2 in water-ethanol mixtures sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

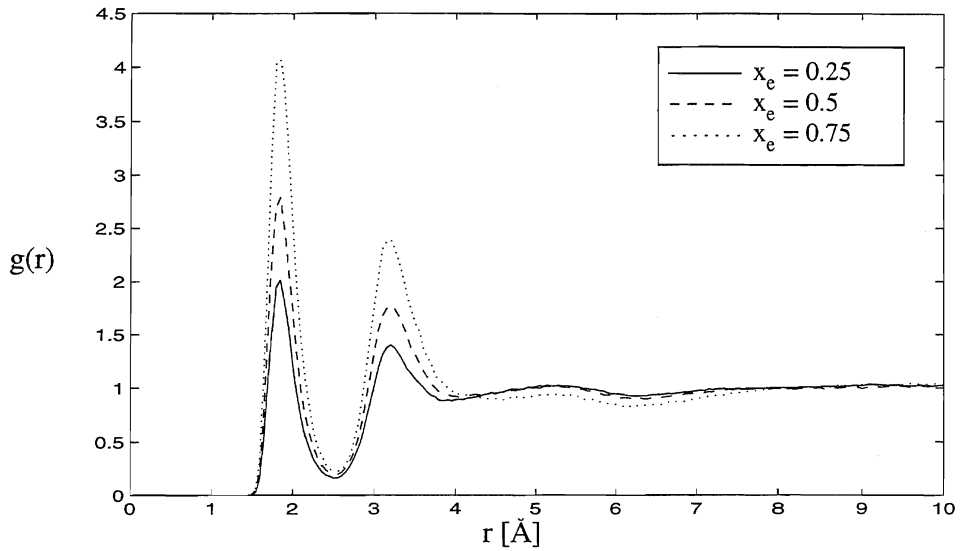


Figure 6.27 Radial cross-correlation function g(r) for H_w -O_e in water-ethanol mixtures sampled with NVT simulation at 293K. x_e is ethanol mole fraction.

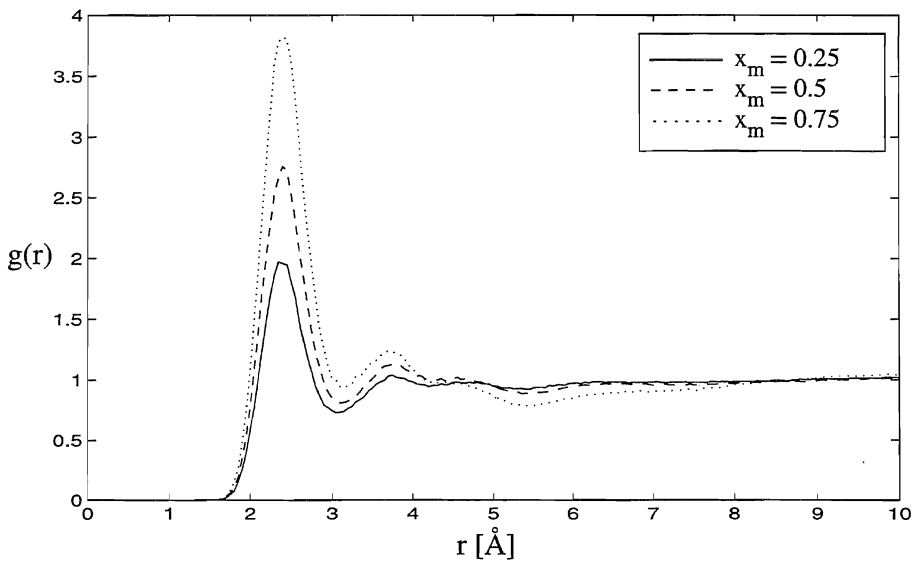


Figure 6.28 Radial cross-correlation function g(r) for H_w -H_e in water-ethanol mixture sampled with NVT simulation. x_e is ethanol mole fraction.

The assignment of one particular structure (or preferably several) to this amount of information is difficult as there are so many distances to sort out. We start by comparing to the proposed 'sandwich-model' [173], we see that we do not find the interlayer distance of 7.0Å. In fact we only find ethanol-hydrogen and water-oxygen in abundance at this distance from a central ethanol molecule.

Except for the interlayer distance, our results are in reasonable agreement with the proposed model. It would perhaps have been very peculiar if we had perfect agreement given all the possible variations in bending and tilting of the ethyl group. We see perhaps signs of a V-shaped trimer as in the model, but the parallel layers are - if they exist - at a larger separation.

Table 6.8

Some intermolecular ethanol-ethanol distances with a 'sandwich model' for ethanol-water mixtures from [173] compared to our findings

pairs	r_{max} (from [173])	r_{max} (our calculations)
O_e-O_e	4.95 (interlayer)	4.8 (2nd max)
CH_3-CH_3	7.0 (interlayer)	9.0 (2nd max)
CH_3-CH_3	4.5	4.3 (broad max)
CH_2-CH_2	4.0 and 4.5	4.6 (max from 4 to 5.5)
CH_3-CH_2	4.2 and 5.3	max from 4 to 5-6
CH_2-O_e	3.6 and 4.8	3.7 and 5.3 (broad)

Let us try to examine the trimer closer from crude estimates of site-site distances as given in Figure 6.29.

From 2nd maxima of the oxygen-hydrogen and the oxygen-oxygen self-correlations for ethanol, we find a trimer angle of 122° and 124° respectively. The close agreement tells us that both hydrogen bonds are linear, but the peak width nevertheless suggest a variation between ~110° and ~135°. We will use a trimer angle of 123°. The O_e-$(OH)_e$ distance is constant with decreasing ethanol mole fraction while the O-O-O angle gets slightly less. This imply that the H-bonds will become a little bent in aqueous solutions.

Sitting at a CH_3-site, we find a neighbouring methyl within a distance of 4.0 - 4.5Å, closer the more diluted the mixtures are. At nearly the same distance we also find a CH_2-site. The 1st peaks are both broad, due to several contributions, and cover the same radial region. Their distances are too close to be the hydrogen bonded neighbours, they must both be situated either on the same side of a chain, and/or belonging to different polymers (or a neighbouring monomer).

The possibly shortest distance between a CH_3 and its nearest hydrogen bonded oxygen neighbour is 4.9Å from Figure 6.29. This is the 2nd peak of the CH_3-O correlation. A 2nd CH_3-H peak is found at 5.4Å, and could very well be due to the hydrogen-companion of this oxygen-site.

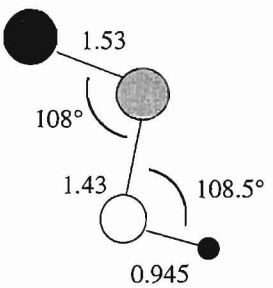
The shortest CH_3-O distance is 3.7Å and is found from Figure 6.29 between two hydrogen bonded trans-conformers in the same plane. If the peaks in CH_3-O at 3.9Å and in CH_3-H at 3.15Å are from nearest donating neighbour, the ethyl groups must be rotated out of the O-O-O plane.

As seen from the CH_2-site, another CH_2 is found about 4.7Å apart. Also the CH_3-site is found at nearly the same separation. Since also the CH_3-correlations show abundance of CH_3 and CH_2 at the same distances, the two sites must belong to the same ethyl group, and be roughly parallel to the central ethyl group. The closest hydrogen and oxygen disappear in the mixtures, but looking at the cross correlations we find peaks at these positions, which could mean that ethanol is replaced by water. A shoulder build at the right side of 2nd CH_2-O peak and also the 2nd minimum of CH_3 is filled.

Finally, placing ourselves at the oxygen site, we see a close and a far-lying hydrogen at 1.85Å and 3.35Å respectively. The close hydrogen is the proton donor and the far belong to proton accepting nearest neighbour. CH_2 is observed at 3.7Å (decreasing with increasing water) and 5.5Å, and CH_3 is at distances 3.9 (disappearing with increasing water) and 4.95Å. The closest pair of distances to CH_2 and CH_3 can be the nearest H-bonded neighbour in a chain. The outer pair can be a next nearest neighbour. The CH_2-O correlation at 3.7Å are replaced partly by a CH_2-O_w correlation, and the CH_3-O correlation at 3.9Å by a CH_2-O_w correlation in aqueous mixtures. This is a further indication of ethanols in chain being substituted by water molecules.

The distances contain detailed information about the geometry of a chain or ring, but it is not easy to extract this in a manual process. It would of course have been smarter to calculate angular correlation functions.

We particularly wonder why there seem to be a lack of sites around 7.Å from the ethyl-carbons in the bulk ethanol. The region is filled with hydrogen. For mixtures, the region seems to be filled with water oxygens.



trans-ethanol

Trimer distances with
O-O-O angle 123°
Based on a planar structure

a - 6.1
b - 4.8
c - 4.85
d - 4.9
e - 4.9
f - 3.8
g - 3.7
h - 5.3
i - 4.2

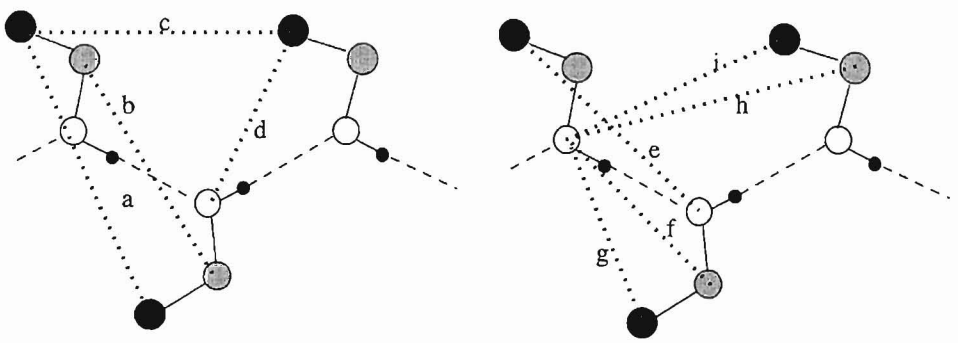


Figure 6.29 Trans ethanol in trimer with some of the distances calculated. Imagine a back and forth rotation of ethyl groups around OH-bond. Ethyl groups can not lie in the same plane as the oxygens, because of overlap between CH_3 and O.

6.3.6 Site-site coordination numbers

Tables 6.9 - 6.11 below show site-site coordination number calculated from integration to 1st minima of the respective pair correlation functions, see Tables D.4 - D.6.

An oxygen-oxygen coordination number of 2.1 for bulk ethanol give preference for long chains or rings, or branched shorter chains. Experimental value is 1.7 [9], which allows for more chain ends/monomers.

The total number of oxygens coordinated to a central ethanol-oxygen increase from 2.1 in bulk ethanol to 2.7 in the x_e=0.25 mixture, where nearly two of the oxygens come from the surrounding water. The same increase was seen with methanol. With increasing water content ethanol, like methanol, must accept an extra proton.

For water the total number of oxygen coordinated oxygens (regardless where they belong) decrease from 4.9 in bulk water to 3.7 in the most ethanol rich mixture. This is almost the same figures as for water in methanol mixtures, see Tables 5.9 - 5.11, page 132.

The sum of all oxygen-oxygen coordinations is found to have a maximum at ethanol mole fraction 0.25.

In general, the site-site coordination numbers for ethanol are very similar to the corresponding correlations for methanol mixtures. The differences relate to the CH_3 and CH_2 sites, where n_{CH_3O} for methanol is larger than both n_{CH_3O} and n_{CH_2O} for ethanol. Also the cross correlations of methanol give n_{OCH_3} larger than both n_{OCH_3} and n_{OCH_2} for ethanol, but the cross correlations for H_w and CH_3 in ethanol yield the same coordination numbers as for methanol. Finally the cross correlations yield n_{CH_3O} for methanol equal to n_{CH_2O} for ethanol. The above observations hold for all mole fractions of ethanol.

We find good agreement with the reported coordination numbers of Jorgensen[27,110], except possibly for n_{CH_2O} which is 'about two' with Jorgensen and 3.0 in our calculation.

Table 6.9

Site coordination numbers for water-water correlations in ethanol-water mixtures at 293K integrated to their respective minima.

x_e	n_{OO}	n_{OH}	n_{HH}
0.0	4.9	2.0	5.3
0.25	3.4	1.5	4.2
0.50	2.4	1.1	2.8
0.75	1.1	0.5	1.2

Table 6.10

Site coordination numbers for ethanol-ethanol correlations in ethanol-water mixtures at 293K integrated to respective minima.

x_e	n_{C3C3}	n_{C3C2}	n_{C3O}	n_{C3H}	n_{C2C2}	n_{C2O}	n_{C2H}	n_{OO}	n_{OH}	n_{HH}
0.25	6.9	6.3	0.9	0.4	7.6	0.8	0.2	0.3	0.2	0.5
0.50	9.3	9.6	1.4	0.7	11.0	1.6	0.4	0.9	0.4	1.1
0.75	11.9	11.5	2.1	1.0	12.5	2.4	0.8	1.4	0.7	1.7
1.0	14.0	13.1	2.7	1.5	13.5	3.0	1.1	2.1	1.0	2.2

Table 6.11

Site coordination numbers for water-ethanol cross correlations in binary mixtures at 293K, 1st index is water-sites. Ethanol–water coordination numbers are found by multiplying the entries with N_w/N_m

x_e	$\bar{n}_{OC3}$	$\bar{n}_{OC2}$	$\bar{n}_{OO}$	$\bar{n}_{OH}$	$\bar{n}_{HC3}$	$\bar{n}_{HC2}$	$\bar{n}_{HO}$	$\bar{n}_{HH}$
0.25	1.8	1.8	0.8	0.3	4.7	0.2	0.2	0.6
0.50	2.7	2.8	1.6	0.6	7.0	0.5	0.4	1.3
0.75	3.2	4.0	2.6	1.0	9.0	0.8	0.7	2.0

6.4 Dynamical properties for ethanol/water mixtures.

6.4.1 Self–diffusion for water and ethanol.

Table 6.12 and Figure 6.30 show calculated and experimental values [164, 174] for self diffusion coefficients for water and ethanol. The experimental values are at 298K, while the calculated values are at 293K. Comparing to the experimental results for methanol-water mixtures at 25°C and 32°C, we expect that the qualitative features of the curves for the ethanol-water mixtures will not be altered with a temperature reduction of only 5K, but that all values will be lowered.

Partington *et al.* [164] have reported values for D_w and D_e in the liquids for the temperature range 15°C to 45°C. Their values at 25°C for water are higher ($2.44 \cdot 10^{-5} cm^2/s$) and for ethanol slightly lower ($1.011 \cdot 10^{-5} cm^2/s$) than the values of Hertz and Leiter [174]. Interpolation to 20°C yield the values $D_w = 2.16 \cdot 10^{-5} cm^2/s$ and $D_e = 0.89 \cdot 10^{-5} cm^2/s$. We therefore expect that the experimental curves of Figure 6.30 would have been lowered by roughly 10% if taken at 20°C.

Table 6.12

Calculated and experimental self diffusion coefficients for water and ethanol. The NVE simulation is at 296K. Experimental values from Hertz and Leiter, 1982 [174] at 298K, calculated values at 293K, but see text.

Molefraction ethanol x_e	D[$10^{-5}cm^2/s$] NVT (at 293K)		D[$10^{-5}cm^2/s$] exp at (298K)	
	water	ethanol	water	ethanol
0.00	3.1		2.25	
0.25	1.1	0.9	0.93	0.60
0.50	1.0	1.1	0.87	0.73
0.75 NVE	0.8 0.8	0.8 0.7	0.93	0.87
1.00		0.7		1.09

Compared to the interpolated liquid values of Partington *et al.* [164], and also to the results of Hertz and Leiter [174], our liquid values are too high by ~40% for water and too low by ~20% for ethanol.

A very prominent feature of D_w in the ethanol-mixtures as well as in the methanol-mixtures is the steep descent in dilute alcohol solution. Then the D_w in ethanol-mixtures proceed nearly constantly with a weak minimum at x_e=0.5, and increase slowly in the ethanol rich mixtures. In methanol solutions on the other hand, D_w increase to nearly its liquid water value in the very dilute water (methanol-rich) solution.

The self-diffusion coefficient of ethanol is on the other hand more similar to that of methanol in aqueous solutions, with a clear minimum near $x_{e\,or\,m}$ =0.2.

Turning to our calculated values, we see that for water all values, except possibly for the x_e=0.75 mixture, are slightly higher than experimental values. The general trend of a steep decrease followed by a nearly constant development is reproduced. The weak experimental minimum is not reproduced, but considering the uncertainties in the values, they show a good agreement with experiments.

For ethanol our figures are reasonable, but we see an increase in D_e with reduced ethanol mole fraction from x_e=1.0 to x_e=0.5 instead of a slow decrease.

The accuracy of the calculations is assumed to be of the same order as for the methanol-water calculations. This is probably not good enough to reveal the finer details of the variation with concentration, see discussion of accuracy in Subsection 5.5.1.

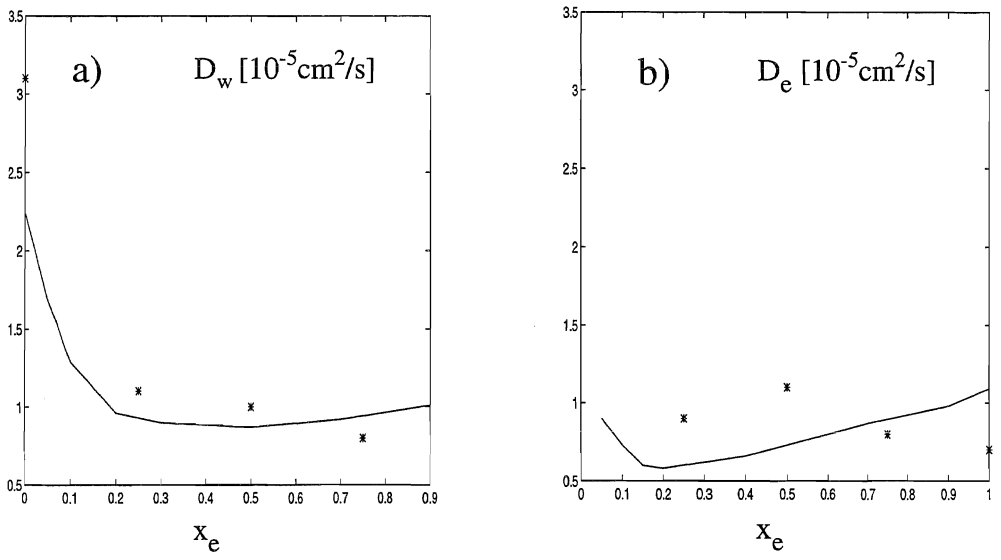


Figure 6.30 Self-diffusion coefficients D_w for water (a) and D_e for ethanol (b) calculated from NVT–simulation (*) at 293K compared to experimental results [174] (solid lines) at 298K.

The TIP4P-water yield too high values, also when mixed with ethanol, while the OPLS-ethanol yield too low values. It is tempting to ascribe the lowering of D_w in ethanol rich mixtures and the increase of D_e in water rich mixtures to the influence of the solvent-solute interactions. Again, the effect of only simulating a *trans*-ethanol molecule must be taken into consideration. This might very likely be a source of error in the self-diffusion of ethanol. Intuitively one would believe that a molecule capable of changing internal orientation would diffuse more readily than a molecule with a completely rigid structure.

From Figure D.42 and D.43, we see that all the mean square displacements for water yield straight and smooth curves, while the same curves for ethanol are not smooth. The curves for mole fractions 0.5 and 0.25 are clearly convex also at long times, which explain the higher diffusivities calculated for these mixtures. The validity of the Einstein-relation for those curves is questionable.

In fig D.44, the mean square deviations of the NVT and NVE simulations at mole fraction x_e=0.75 is compared. Even if the NVT simulation of this mixture is questionable with respect to stability, see Section 6.1, the NVT and NVE simulations yield values in close agreement.

6.4.2 Velocity auto correlation functions for water and ethanol

Figure 6.31 show the x-components of normalized velocity auto correlation function for water in mixtures with ethanol. Velocity auto correlation function for pure water is similar to that at 298K, but the minimum near 0.1ns is deeper in mixtures with ethanol than in mixtures with methanol. Also the behaviour in the back-scattering region is different with ethanol: there is a larger variation in depth with varying concentration than with methanol. The molecular velocities in the x_e=0.25 mixture seem to change sign more often, and the velocities in the x_e=0.75 mixture more seldom than in the corresponding methanol mixtures. But remember that the 0.75 mixture has an anomalous velocity distribution for water, Section 6.2.

For ethanol the minimum near 0.3ns grows deeper with decreasing ethanol content. Pure ethanol has a shallow and broad minimum. Notice a very small shoulder in the same position as for methanol, but weaker.

The variation with direction is small, particularly for times less than 0.4 - 0.5, as is seen from Figure 6.33.

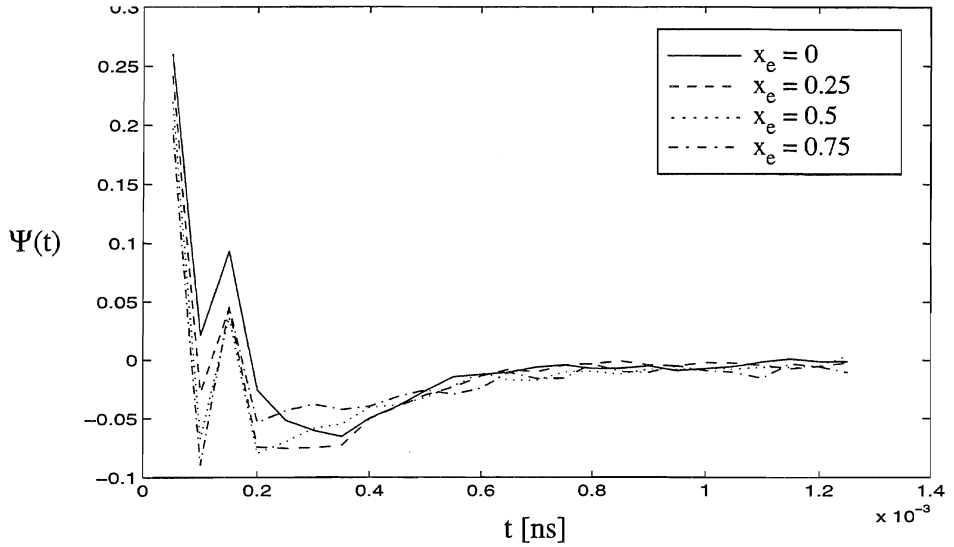


Figure 6.31 Normalized velocity auto correlation function (x-component) for water from NVT–simulations of water-ethanol mixtures at 293K.

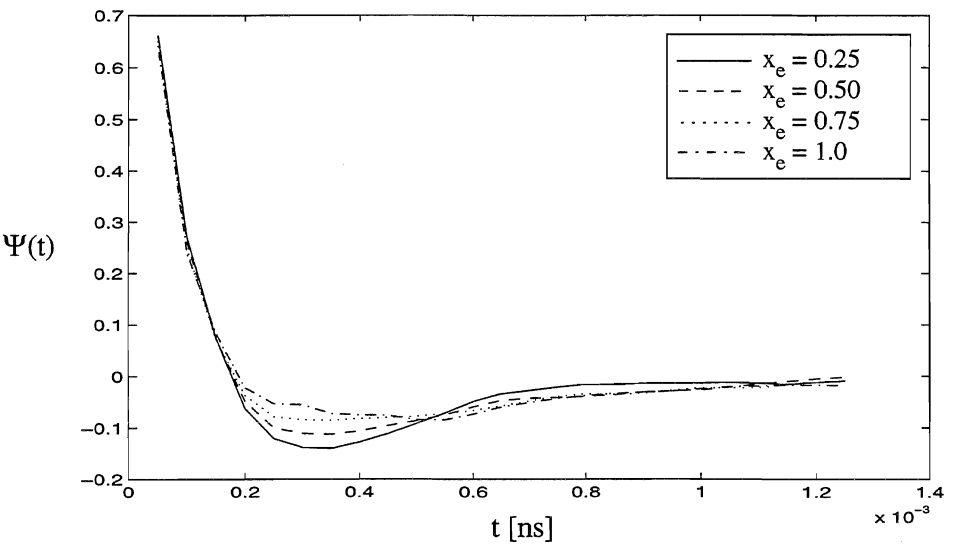


Figure 6.32 Normalized velocity auto correlation function (VACF) (x-component) for methanol from NVT–simulations of water-ethanol mixtures at 293K.

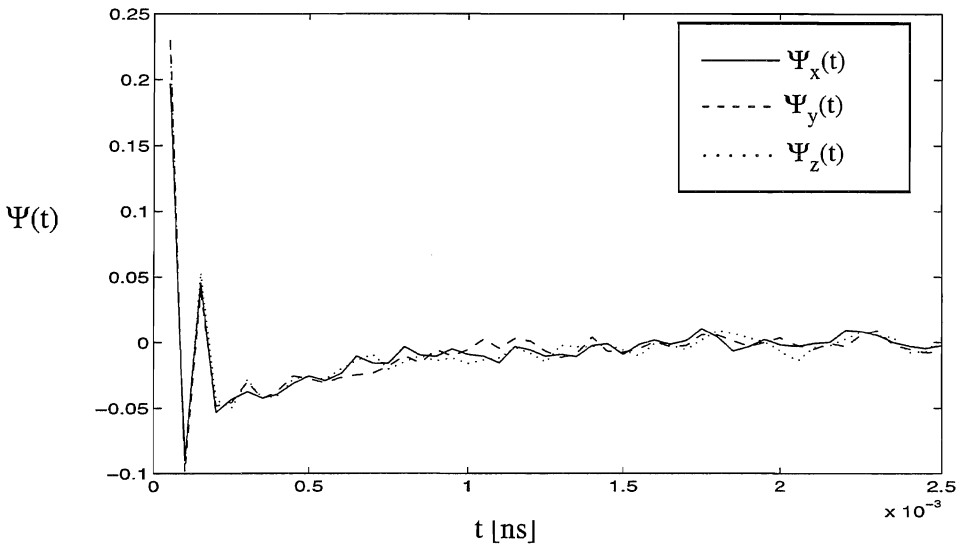


Figure 6.33 x, y, and z-components of normalized velocity auto correlation function for water in a 0.75 ethanol mixture from NVT–simulation.

6.5 Summary

We summarize below some of our results from the NVT equilibrium simulations of water and ethanol and their mixtures. The results are presented and discussed in the previous sections. An important premise of our simulations, is that ethanol is modelled as the *trans*-conformer.

- None of the simulations are strict representatives of the canonical ensemble, but they seem to improve with time, as we also found for the methanol-water simulations. In both cases this seem to be due to a slightly to weak temperature control (too large Q). With a smaller Q, a reduction in steplength must however be considered.
- Our thermodynamic calculations are in general in good agreement with experiments and simulations. The exception is the mixture with 0.75 mole fraction ethanol, which failed to reach a stable conformational energy in the NVT-simulation. Thus the excess potential energy for this mixture has a positive value for this mixture.

- An additional simulation at constant energy was performed for this mixture. This simulation showed a more stable configurational energy.
- The calculated pressures are high, as for the methanol-water simulations.
- Also the structure data for the pure liquids agree with the experimental data.
- The radial positions of 1st maxima are not changed with varying composition.
- Site coordination numbers for methyl-methyl correlations decrease only by a factor of two when the ethanol mole fraction is reduced from 1.0 to 0.25. This was also observed for methanol. The hydroxyl interaction decrease more than number density.
- For the mixtures we find evidence that water is replacing ethanol upon addition of water to ethanol, as we found from the methanol-water simulations.
- We found our calculated structure of the mixtures to be in reasonable agreement with a proposed 'sandwich-model' for mixtures [173], but the interlayer distance was not in agreement.
- For the structure, we could not expect to find the hydration cages observed in experiments, simply because we did not simulate the dilute mixtures.
- The self-diffusion coefficients of water in mixtures of ethanol and water are found to be reasonably reproduced, but the diffusion is, as for the methanol-water mixtures, overestimated.
- The results for self-diffusion of ethanol does not reproduce experimental trends, even if the values not depart unreasonable far from experimental values. One possible reason for this behaviour might be the use of the *trans*-conformation only.

Chapter 7
Conclusion

We have implemented a multiple Nosé-Hoover thermostat for separate temperature control of translational and rotational degrees of freedom in equilibrium molecular dynamics simulations. We have performed equilibrium molecular dynamics simulations at constant energy (NVE) and constant temperature (NVT) for model liquids of water, methanol, and ethanol, and for binary aqueous mixtures of the alcohols as described in the introductory chapter. The results for thermodynamics, structure and dynamics have been presented.

The stability of pressure and energy, along with the numerical values of potential energy show that the systems are equilibrated. From the conserved quantities, and from the distribution of energy between the various kind of degrees of freedom, the mixtures are in internal equilibrium. Finally, the radial correlation functions are consistent with a liquid state.

Conservation of total energy is found to improve with increasing alcohol mole fraction for both NVE and NVT simulations.

The agreement between our NVT and NVE simulations is very good. Our results are in overall good agreement with experiments, and with published Monte Carlo and molecular dynamics simulations of the same model liquids.

Temperatures for mixtures calculated in the NVE simulations are also very close to the desired temperatures. Thus, if accurate temperatures are not needed, NVE simulations are sufficient.

We deliberately do not name the thermostatted simulations 'canonical', since we find that none of the simulations are strictly canonical. They show a 'slightly noncanonical' behaviour, which is due to a too weak coupling between the system and the reser-

voirs. They seemed however to be approaching the canonical distribution, but whether this would ever have happened for the particular relaxation parameters we applied is not clear.

The estimation of appropriate relaxation parameters turned out to be a tedious trial-and-error procedure. From the simulated results, our values are however not very much out of range.

Comfortingly, it also seems that a simple weighted average of relaxation parameters for the mixtures is sufficient. The analysis with respect to the canonical distribution show no large differences between the liquids and their mixtures.

Except for the overestimation of fluctuations, the thermodynamical and structural data are not affected by the choice of surroundings. With the exception of self-diffusion of methanol in the most water-rich mixture, the self-diffusion coefficients are not affected by the thermostats either. The application of two separate heat baths is therefore quite successful.

The heats of vaporizations are in good agreement with the experimental results. Excess configurational energies are compared to experimental excess enthalpies. They are in good agreement for the methanol-water mixtures. For the ethanol-water mixtures the agreement get worse with increasing ethanol mole fraction, the curve showing both a maximum and a minimum.

We find the self-diffusion constants of water and methanol to be in reasonable accordance with experiments. The self-diffusion coefficients of ethanol disagree most with experiments. This might be due to the rigid potential applied for ethanol. Water in mixtures with ethanol, seem on the other hand not to be affected by its mixture companion.

We found that the structures of methanol-water and ethanol-water mixtures were similar. This not surprising, as the potential models are similar. They were found to be consistent with a V-chain, but the presence of other structures is however not ruled out.

We also found that upon mixing, water tended to replace the nearest neighbors of methanol/ethanol in the chains. The ethanol results must be interpreted with some caution, though, since only one conformer is present.

Methyl-methyl coordination numbers decrease less than hydroxyl coordination numbers upon increasing the water content for both methanol and ethanol mixtures.

Chapter 8
Suggestions for future work

Contrary to be at the finishing line, we feel that the subject is not exhausted at all, and that several topics remain unexplored. Given a fresh start, we would have continued with some of the following:

For the method of thermostatting:

- An analysis of the effects of using one thermostat contra using several thermostats. We have not included this in a formal manner, but except for the adjustment of heat bath masses, it is straightforward and easily performed within the existing program.
- Finding a smoother method to select values for thermostat parameters. The trial-and-error method usually applied is far to laborious if one wants to between models and liquids. The general recommendation is to select frequencies close to characteristic frequencies. The connection between frequencies and heat bath masses are on the other hand only through the approximate relations of Nosé.
- We have assumed without discussion that the potential is sufficiently chaotic to generate ergodic trajectories. The periodic boundaries condition together with the Ewald sum might introduce restrictions upon the available phase space. Previously, only model systems of harmonic oscillators and non-polar Lennard-Jones systems have been investigated with respect to ergodic behaviour. Any effect can for instance be investigated by studying larger systems, different handling of long range forces, or applying a potential with fractional charges increasing from zero.
- A refinement of the values we have found is not necessary for calculation of thermodynamics and structure, but if one wants to calculate quantities derived from

fluctuations, the simulations must be canonical. However, one might also consider the application of other methods that is claimed to be canonical regardless of thermostat coupling, for instance the recently proposed method of Hoover and Holian, 1996 [25].

For the molecular dynamics calculational procedure:

- Longer simulations are needed in the calculation of the diffusional properties. Then the implementation of a faster version of the Ewald summation method, or an alternative method is needed. An alternative method would still have to meet the requirements of a simulation that is close to canonical. The simulation times applied in this study are however sufficient for the calculation of the other properties.
- A rigid calculation of the statistical errors in the self-diffusion constants is important. It is dangerous to interpret the causes to disagreement between calculated and experimental results when the accuracy is not known. To a first approximation this can be done with a calculation of the variance in tangent to the mean square displacement curves.

For the systems:

- Smaller mole fractions of the alcohols must be simulated in order to reproduce the minimum in excess energy, or to reproduce (if possible) the experimentally observed hydration cages. This might require larger number of molecules, and for the hydration structure, different sampling techniques.

Other calculations we would have enjoyed:

- A hydrogen-bond analysis for ethanol and ethanol-water mixtures, both with respect to total number of hydrogen bonds per molecule, and fractions of n-mers in the mixtures. This is not performed for for these systems.
- A calculation of angular correlation functions. These functions makes it easier to analyze the structures.
- The calculation of reorientational correlation functions. They can be compared to spectroscopic experiments, and then provide a mean for validation of the model.
- Calculations of free energy and entropy. Entropic quantities are to our knowledge not studied in simulations for water-ethanol mixtures, and could provide useful information to the mechanisms of solvation.

Regarding the potential models:

- On the basis of the above investigations for diffusion, excess energies and structure, it should be possible to get a stronger confidence to whether a flexible model should be used for ethanol. As for the moment, we have a strong feeling that both conformers are needed in order to get realistic results for dynamics and structure. If so, a Monte Carlo calculation is easier, but then the dynamical results are not available. A faster calculation of long range forces is then needed to allow for the increased computing time associated with flexible models due to the multiple timescale.
- For methanol and ethanol there are few model potentials to choose from. Indications of a possible underestimation of diffusion are seen for the OPLS models. Also the TIP4P model has a weakness regarding the self-diffusion, and is known to yield too much structure in the oxygen-oxygen self correlations. We also found extremely high pressures in the liquid water simulations at room temperature. These deficiencies suggest that one still should look for improved model potentials for water, preferably rigid, and for the alcohols.

References

[1] S. Nosé; Mol. Phys., 52 (1984) 255

[2] W.G. Hoover; Phys. Rev. A, 31 (1985) 1695

[3] *'Ullmann's Encyclopedia of Industrial Chemistry'*, exec. ed. W. Gerhartz, VCH, Weinheim, 1987

[4] F. Franks; *'Water. A Comprehensive Treatise'*, Vol 1, Plenum Press, New York, 1972

[5] F. Franks, D.J.G. Ives; Quarterly Reviews, 20 (1966) 1

[6] A.H. Narten, W.E. Thiessen, L. Blum; Science, 217 (1982) 1033

[7] J.C. Dore; *'Water Science Reviews'*, vol 1, ed F. Franks, Cambridge University Press, Cambridge, 1985

[8] S.A. Rice, M.G. Sceats; J. Phys. Chem., 85 (1981) 1108

[9] A.H. Narten, A. Habenschuss; J. Chem. Phys., 80 (1984) 3387

[10] W.H. Zachariasen; J. Chem. Phys., 3 (1935) 158

[11] G.G. Harvey; J. Chem. Phys., 7 (1939) 878

[12] K.S. Vahvaselkä, R. Serimaa, M. Torrkeli; J. Appl. Cryst., 28 (1995) 189

[13] L.D. Landau, E.M. Lifshitz; *'Statistical Physics. Part 1'*, Pergamon, Oxford, 1980, 3rd ed.

[14] B.K. Agarwal, M. Eisner; *'Statistical Mechanics'* John Wiley & Sons, New Dehli, 1988

[15] M.P. Allen, D.J. Tildesley; *'Computer Simulations of Liquids'*, Clarendon Press, Oxford, 1990

[16] W.W. Wood, *'Fundamental Problems in Statistical Mechanics'*, ed. E.G.D. Cohen, North-Holland Publishing Company, Amsterdam, 1975

[17] S. Nosé; J. Chem. Phys., 81 (1984) 511

[18] S. Nosé; Mol. Phys., 57 (1986) 187

[19] D.J. Evans, B.L. Holian; J. Chem. Phys., 83 (1985) 4069

[20] W.G. Hoover; Nuclear Physics, A545 (1992) 523c

[21] S. Baratham, I. L'Heureux, R. Kapral; J. Chem. Phys., 91 (1989) 5602

[22] D. Keffer, A.V. McCormick, H.T. Davis; Mol. Phys., 87 (1996) 367

[23] G. Galli, A. Pasquarello; *'Computer Simulation in Chemical Physics'*, eds. M.P. Allen, D.J. Tildesley, NATO ASI Series C, Vol 397, Kluwer Academic Publishers, Dordrecht, 1992

[24] S. Nosé; Progress of Theoretical Physics. Supplement, 103 (1991) 1

[25] W. G. Hoover, B.L. Holian; Phys. Lett. A, 211 (1996) 253

[26] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein; J. Chem. Phys., 79 (1983) 926

[27] W.L. Jorgensen; J. Phys. Chem., 90 (1986) 1276

[28] J.M. Haile; *'Molecular Dynamics Simulation. Elementary Methods'*, John Wiley & Sons, New York, 1992

[29] G. Alagona, A. Tani; Chem. Phys. Lett., 87 (1982) 337

[30] A. Tani; Mol. Phys., 53 (1984) 1289

[31] G. Alagona, A. Tani; J. Mol. Struct. (Theochem), 166 (1988) 375

[32] A. C. Lilly,Jr, F.J. Beers, J.C. Schug; J. Mol. Struct. (Theochem), 209 (1990) 69

[33] J.P. Hansen, I.R. McDonald, *'Theory of Simple Liquids'*, 2nd ed., Academic Press, London, 1990

[34] J. Casulleras, E. Guardia; Mol. Sim., 8 (1992) 273

[35] B.J. Alder, T.E. Wainwright; J. Chem. Phys., 27 (1957) 1208

[36] K.R. Symon; *'Mechanics'*, Addison-Wesley, Reading, 1973, 3rd ed.

[37] H. Goldstein; *'Classical Mechanics'*, Addison-Wesley, Tokyo, 1964

[38] L.V. Woodcock; Chem. Phys. Lett., 10 (1971) 257

[39] H.C. Andersen; J. Chem. Phys., 72 (1980) 2384

[40] D.J. Evans; J. Chem. Phys., 78 (1983) 3297

[41] W.G. Hoover, A.J.C. Ladd, B. Moran; Phys. Rev. Lett., 48 (1982) 1818

[42] K.F. Gauss; Journal für die reine und angewandt Mathematik, IV (1829) 232

[43] D.J. Evans, W.G. Hoover, B.H. Failor, B. Moran, A.J.C. Ladd; Phys. Rev. A, 28 (1983) 1016

[44] W.G. Hoover; *'Computational Statistical Mechanics'*, Elsevier, Amsterdam, 1991

[45] J. Jellinek; J. Phys. Chem., 92 (1988) 3163

[46] J. Jellinek, R.S. Berry; Phys. Rev. A, 38 (1988) 3069

[47] K. Cho, J.D. Joannopoulos; Phys. Rev. A, 45 (1992) 7089

[48] J. Jellinek, R.S. Berry; Phys. Rev. A, 40 (1989) 2816

[49] F.D. Di Tolla, M. Ronchetti; Phys. Rev. E, 48 (1993) 1726

[50] D.M. Bylander, L. Kleinman; Phys. Rev. B, 46 (1992) 13756

[51] R. Fowler, E.A. Guggenheim; *'Statistical Thermodynamics: a version of statistical mechanics for students of physics and chemistry'*, Cambridge University Press, Cambridge, 1939

[52] K. Cho, J.D. Joannopoulos, L. Kleinman; Phys. Rev. E, 47 (1993) 3145

[53] D. Kusnezov, A. Bulgac, W. Bauer; Annals of Physics, 204 (1990) 155

[54] A. Bulgac, D. Kusnezov; Phys. Rev. A, 42 (1990) 5045

[55] R.G. Winkler; Phys. Rev. A, 45 (1992) 2250

[56] G.J. Martyna, M.L. Klein, M. Tuckerman; J. Chem. Phys., 97 (1992) 2635

[57] D.J. Tobias, G.J. Martyna, M.L. Klein; J. Phys. Chem., 97 (1993) 12959

[58] H.A. Posch, W.G. Hoover, F.J. Vesely; Phys. Rev. A, 33 (1986) 4253

[59] S. Nosé; Phys. Rev. E, 47 (1993) 164

[60] B.L. Holian, A.F. Voter, R. Ravelo; Phys. Rev. E, 52 (1995) 2338

[61] F. Lado; J. Chem. Phys., 75 (1981) 5461

[62] T. Çagin, J.R. Ray; Phys. Rev. A, 37 (1988) 4510

[63] R. Lustig; J. Chem. Phys., 100 (1994) 3048

[64] L. Verlet; Phys. Rev., 159 (1967) 98

[65] H.J.C. Berendsen, W.F. van Gunsteren; *'Molecular Liquids - Dynamics and Interactions'*, eds. A.J. Barnes *et al.*, NATO ASI series C135, Reidel, New York, 1984

[66] D. Janezic, B. Orel; Int. J. of Quantum Chem., 51 (1994) 407

[67] D. Janezic, R. Trobec; J. Chem. Inf. Comput. Sci., 34 (1994) 641

[68] S. Toxvaerd; Mol. Phys., 72 (1991) 159

[69] D.J. Evans; Mol. Phys., 34 (1977) 317

[70] D. Fincham; Mol. Sim., 8 (1992) 165

[71] C. Pierleoni, J.-P. Ryckaert; Mol. Phys., 75 (1992) 731

[72] Z. A. Rycerz; Mol. Sim., 9 (1992) 327

[73] K. Tasaki, S. McDonald, J.W. Brady; J. Comp. Chem., 14 (1993) 278

[74] D.B. Kitchen, F. Hirata, J.D. Westbrook, R. Levy, D. Kofke, M. Yarmush; J. Comp. Chem., 11 (1990) 1169

[75] P. Ewald; Ann. Phys. Leipzig, 64 (1921) 253

[76] H. Kornfeld; Z. Phys., 22 (1924) 27

[77] C. Kittel; *'Introduction to Solid State Physics'*, 6th ed., John Wiley & Sons, New York, 1986

[78] *'CRC Handbook of Chemistry and Physics. 62 edition'*, ed R.C. Weast, CRC Press, Inc., Boca Raton, 1981

[79] D.M. Heyes, CCP5 Quarterly, 8 (1982) 29

[80] D.M. Heyes, J. Chem. Soc. Faraday Trans., 90 (1994) 3039

[81] S. Bhowmick, D. Roy, R. Bhattacharya; Chem. Phys. Lett., 148 (1988) 317

[82] S. Kuwajima, A. Warshel; J. Chem. Phys., 89 (1988) 3751

[83] B. Cichocki, B.U. Felderhof, K. Hinsen; Phys. Rev. A, 39 (1989) 5350

[84] J.M. Caillol; J. Chem. Phys., 96 (1992) 1455

[85] R.K. Kalia, S. de Leeuw, A. Nakano, P. Vashishta; Comp. Phys. Comm., 74 (1993) 316

[86] D. Fincham; Mol. Sim., 13 (1994) 1

[87] O. Teleman, A. Wallqvist; Int. J. Quantum Chem: Quantum Chem. Symp. 24, (1990) 245

[88] S.-H. Suh, K. Ha, J.-S. Kim, C.-Y. Park, N.H. Heo; Korean J. Chem. Eng., 9 (1992) 135

[89] M.R. Spiegel; *'Schaum's Outline of Theory and Problems of Statistics'*, McGraw-Hill, New York, 1972

[90] C.G. Gray, K.E. Gubbins; *'Theory of Molecular Fluids. Volume 1: Fundamentals'*, Clarendon Press, Oxford, 1984

[91] J.E. Lennard-Jones; Proc. Roy. Soc. (London), 106A (1924) 441 and 463

[92] J.P. Hansen; in *'Hydrogen-Bonded Liquids'*, eds. J.C. Dore, J. Teixeira, NATO ASI Series C, Vol 329, Kluwer Academic Publishers, Dordrecht, 1991

[93] D.E. Smith, A.D.J. Haymet; Fluid Phase Equilibria, 88 (1993) 79

[94] A.G. Sharpe; *'Inorganic Chemistry'*, Longman, London, 1981

[95] *'CRC Handbook of Chemistry and Physics. 74 edition'*, ed D.R. Lide, CRC Press, Inc., Boca Raton, 1993

[96] Y. Marcus; *'Ion solvation'*, Wiley and Sons Ltd., Chichester, 1985

[97] A. Rahman, F.H. Stillinger; J. Am. Chem. Soc., 95 (1973) 7943

[98] F.H. Stillinger, A. Rahman; J. Chem. Phys., 60 (1974) 1545

[99] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans; *'Jerusalem Symposia on Quantum Chemistry and Biochemistry'*, ed. B. Pullman, Reidel, Dordrecht, 1981

[100] W.L. Jorgensen; J. Am. Chem. Soc., 103 (1981) 335

[101] W.L. Jorgensen; J. Chem. Phys., 77 (1982) 4156

[102] D.L. Beveridge, M. Mezei, P.K. Mehrotra, F.T. Marchese, G. Ravi-Shanker, T. Vasu, S. Swaminathan; *'Molecular-based Study of Fluids'*, eds, J.M Haile, G.A. Mansori, American Chemical Society, Washington, D.C., 1983

[103] A. Ben-Naim, F.H. Stillinger; *'Structure and Transport Processes in Water and Aqueous Solutions'*, ed. R.A. Horne, Wiley Interscience, New York, 1972

[104] K. Watanabe, M.L. Klein; Chem. Phys., 131 (1989) 157

[105] M. Ferrario, A. Tani; Chem. Phys. Lett., 121 (1985) 182

[106] M. Sprik; *'Computer Simulation in Chemical Physics'*, eds. M.P. Allen, D.J. Tildesley, NATO ASI Series C, Vol 397, Kluwer Academic Publishers, Dordrecht, 1992

[107] H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma; J. Phys. Chem., 91 (1987) 6269

[108] F. Franks, J.E. Desnoyers; *'Water Science Reviews 1'*, ed F. Franks, Cambridge University Press, Cambridge, 1985

[109] W.L. Jorgensen; J. Am. Chem. Soc., 103 (1981) 341

[110] W.L. Jorgensen; J. Am. Chem. Soc., 103 (1981) 345

[111] W.L. Jorgensen; J.D. Madura, C.J. Swenson, J. Am. Chem. Soc., 106 (1984) 6638

[112] M. Haughney, M. Ferrario, I.R. McDonald; Mol. Phys., 58 (1986) 849

[113] H. Tanaka, K.E. Gubbins; J. Chem. Phys., 97 (1992) 2626

[114] E. Guàrdia, J.A. Padró; J. Phys. Chem., 94 (1990) 6049

[115] A. Wallquist, O. Teleman; Mol. Phys., 74 (1991) 515

[116] G. Pálinkás, E. Hawlicka, K. Heinzinger; J. Phys. Chem., 91 (1987) 4334

[117] E. Hawlicka, G. Pálinkás, K. Heinzinger; Chem. Phys. Lett., 154 (1989) 255

[118] A. Wallquist, P. Ahlström, G. Karlström; J. Phys. Chem., 94 (1990) 1649

[119] A. Wallquist, B.J. Berne; J. Phys. Chem., 97 (1993) 13841

[120] J.W. Caldwell, P.A. Kollman; J. Phys. Chem., 99 (1995) 6208

[121] D. van Belle, M. Froeyen, G. Lippens, S.J. Wodak; Mol. Phys., 77 (1992) 239

[122] H. A. Lorentz; Ann. Phys., 12 (1881) 127

[123] D. Berthelot; Comptes Rendus Acad. Sci., 126 (1898) 1703

[124] L.C.G. Freitas; J. Mol. Struct. (Theochem), 282 (1993) 151

[125] R.H. Perry, D.W. Green; *'Perry's Chemical Engineers' Handbook'*, McGraw-Hill, 1984

[126] M. Ferrario, M. Haughney, I.R. McDonald, M.L. Klein; J. Chem. Phys., 93 (1990) 5156

[127] M. Haughney, M. Ferrario, I.R. McDonald; J. Phys. Chem., 91 (1987) 4934

[128] A.R. Tindell, D.J. Tildesley, J. Walton; CCP5 Quarterly, 4 (1982) 26

[129] J.J. Erpenbeck, W.W. Wood; *'Statistical Mechanics'*, ed. B.B. Berne, Plenum Press, New York, 1977

[130] J.M. Smith, H.C. Van Ness; *'Introduction to Chemical Engineering Thermodynamics'*, McGraw-Hill, Singapore, 1987, 4th ed

[131] X.-W. Wu, Y.-G.Li, J.-F.Lu, T.Teng; Fluid Phase Equilibria, 77(1992) 139

[132] S. Westmeier; J. Chem. Techn, 28 Bd 6 (1976) 350

[133] R.F. Lama, B.C.-Y. Lu; J. Chem. Eng. Data, 10 (1965) 216

[134] B.M. Ladanyi, M.S. Skaf; Annu. Rev. Phys. Chem., 44 (1993) 335

[135] A. Rahman, F.H. Stillinger; J. Chem. Phys., 55 (1971) 3336

[136] F.H. Stillinger, A. Rahman; J. Chem. Phys., 57 (1972) 1281

[137] M.G. Sceats, S.A. Rice; J. Chem Phys., 72 (1980) 3236

[138] R.J. Speedy, J.D. Madura, W.L. Jorgensen; J. Phys. Chem., 91 (1987) 909

[139] A. Geiger, F.H. Stillinger, A. Rahman; J. Chem. Phys., 70 (1979) 4185

[140] R.L. Blumberg, H.E. Stanley, A. Geiger, P. Mausbach; J. Chem. Phys., 80 (1984) 5230

[141] A. Geiger, P. Mausbach; '*Hydrogen-Bonded Liquids*', eds. J.C. Dore and J. Teixeira, Kluwer Academic Publishers, Dordrecht, 1991

[142] F. Sciortino, A. Geiger, H.E. Stanley; J. Chem. Phys., 96 (1992) 3857

[143] A.K. Soper, M.G. Phillips; Chem. Phys., 107 (1986) 47

[144] P.A. Giguère; J. Chem Phys., 87 (1987) 4835

[145] M. Vedamuthu, S. Singh, G.W. Robinson; J. Phys. Chem., 98 (1994) 2222

[146] A.H. Narten, H.A. Levy; Science, 165(1969) 447

[147] M. Haughney, M. Ferrario, I.R. McDonald; J. Phys. Chem., 91 (1987) 4934

[148] M. Matsumoto, K.E. Gubbins; J. Chem. Phys., 93 (1990) 1981

[149] P.F.W. Stouten, B.P.van Eijck, J. Kroon; J. Mol. Struct., 243 (1991) 61

[150] J. Alonso, F.J. Bermejo, M. García-Hernández, J.L. Martínez, W.S. Howells; J. Mol. Struct., 250 (1991) 147

[151] G. Bolis, G. Corongiu, E. Clementi; Chem. Phys. Lett., 86 (1982) 299

[152] W.L. Jorgensen, J.D. Madura; J. Am. Chem. Soc., 105 (1983) 1407

[153] S. Okazaki, K. Nakanishi, H. Touhara; J. Chem. Phys., 78 (1983) 454

[154] S. Okazaki, H. Touhara, K. Nakanishi; J. Chem. Phys., 81 (1984) 890

[155] G. Pálinkás, I. Bakó; Z. Naturforsch., 46a (1991) 95

[156] G. Pálinkás, E. Hawlicka, K. Heinzinger; Chem. Phys., 158 (1991) 65

[157] A.K. Soper, J.L. Finney; Phys. Rev. Letters, 71 (1993) 4346

[158] C.A. Koh, H. Tanaka, J.M. Walsh, K.E. Gubbins, J.A. Zollweg; Fluid Phase Equilibria, 3 (1993) 51

[159] I.I. Vaisman, M.L. Berkowitz; J. Am. Chem. Soc., 114 (1992) 7889

[160] A.K. Soper, M.G. Phillips; J. Chem. Phys., 101 (1994) 6888

[161] J. L. Finney, A.K. Soper; Chem. Soc. Rev., 23 (1994) 1

[162] E. Hawlicka; Ber. Bunsenges. Phys. Chem., 87 (1983) 425

[163] W. Reimschüssel, E. Hawlicka; Radiochimica Acta, 31 (1982) 157

[164] J. R. Partington, R.F. Hudson, K.W. Bagnall; Journal de Chimie Physique et de Physico-chimie biologique, 55 (1958) 77; Nature, 169 (1952) 583

[165] P.A. Johnson, A.L. Babb; J. Phys. Chem., 60 (1956) 15

[166] J. Kida, H. Uedaira; J. Magn. Res., 27 (1977) 253

[167] F.H. Stillinger in *'Advances in Chemical Physics'*, Wiley, New York, 1975

[168] E. Guàrdia, G. Sesé, J.A. Padró; J. Mol. Liq., 62 (1994) 1

[169] D.C. Rapaport; *'The Art of Molecular Dynamics Simulation'*, Cambridge University Press, Cambridge, 1995

[170] S. Sarkar, R.N. Joarder; J. Chem. Phys., 100 (1994) 5118

[171] A. Coccia, P.L. Inodovia, F. Podo, V. Viti; Chem. Phys., 7 (1975) 30

[172] N. Nishi, S. Takahashi, M. Matsumoto, A. Tanaka, K. Muraya, T. Takamuku, T. Yamaguchi; J. Chem. Phys., 99 (1995) 462

[173] M. Matsumoto, N. Nishi, T. Furusawa, M.Saita, T. Takamuku, M. Yamagami, T. Yamaguchi; Bull. Chem. Soc. Jpn., 68 (1995) 1775

[174] H.G. Hertz, H. Leiter; Zeitschrift für Physikalische Chemie Neue Folge, 133 (1982) 45

Appendix A
Proof of canonical distribution

The proof that Equation (2.26), page 35 generates canonical states is due to Nosé [18, 24] and is included for completeness. For simplicity, q and p represents both translational and rotational degrees of freedom, and M_i symbolizes the either the mass or the moment of inertia of the system.

Let $f=f(\boldsymbol{q},\boldsymbol{p},\eta_1,...,\eta_J)$ be the distribution of states in extended phase space $(\mathbf{q},\mathbf{p},\eta_1,...,\eta_J)$. Each phase point move with velocity $\boldsymbol{v}=(\dot{\boldsymbol{q}},\dot{\boldsymbol{p}},\dot{\eta}_1,...,\dot{\eta}_J)$. The total time derivative of f is

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \mathbf{v} \cdot (\nabla f) \tag{A.1}$$

where $\nabla = [\frac{\partial}{\partial q_{11}},\cdots,\frac{\partial}{\partial q_{fk}},\frac{\partial}{\partial p_{11}},\cdots,\frac{\partial}{\partial p_{fk}},\frac{\partial}{\partial \eta_1},\cdots,\frac{\partial}{\partial \eta_k}]$ is the gradient. If the flow of phase points is regarded as a fluid flow, the continuity equation for a volume in phase space applies

$$\frac{\partial f}{\partial t} = -\nabla \cdot (\mathbf{v}f) = -f(\nabla \cdot \mathbf{v}) - \mathbf{v} \cdot (\nabla f) \tag{A.2}$$

Combining Equations (A.1) and (A.2) and expanding in components yield

$$\frac{df}{dt} = -\sum_{ij}\left(\frac{\partial}{\partial q_{ij}} \cdot \dot{q}_{ij} + \frac{\partial}{\partial p_{ij}} \cdot \dot{p}_{ij}\right)f - \sum_{j}\left(\frac{\partial}{\partial \eta_j} \cdot \dot{\eta}_j\right)f \tag{A.3}$$

Now, substituting the equations of motion (2.26)

$$\frac{df}{dt} = -\sum_{ij}\left(\frac{\partial}{\partial q_{ij}}\cdot\frac{p_{ij}}{M_{ij}} + \frac{\partial}{\partial p_{ij}}\cdot\left\{-\frac{\partial \mathcal{U}}{\partial q_{ij}} - p_{ij}\eta_j\right\}\right)f$$
$$-\sum_{j}\left(\frac{\partial}{\partial \eta_j}\cdot\frac{1}{Q_j}\left\{\sum_i \frac{p_{ij}}{M_{ij}} - g_j k_B T\right\}\right)f \qquad \text{(A.4)}$$

where the only surviving term is

$$\frac{df}{dt} = \sum_{ij}\left(\frac{\partial}{\partial p_{ij}}\cdot p_{ij}\eta_j\right)f = f\sum_j g_j\eta_j \qquad \text{(A.5)}$$

For an isolated system no 'frictional force' appears, and (A.5) reduce to Liouville's theorem, *df/dt =0.* The microcanonical distribution function then follow as a consequence. Next defining a function corresponding to the total energy of the system

$$\tilde{\mathcal{E}}(\mathbf{q},\mathbf{p},\eta) = \sum_{ij}\frac{p_{ij}^2}{2M_{ij}} + \mathcal{U}(\mathbf{q}) + \frac{1}{2}\sum_j Q_j\eta_j^2 \qquad \text{(A.6)}$$

with time derivative

$$\frac{d\tilde{\mathcal{E}}}{dt} = \sum_{ij}\left(\frac{\partial\tilde{\mathcal{E}}}{\partial q_{ij}}\cdot\dot{q}_{ij} + \frac{\partial\tilde{\mathcal{E}}}{\partial p_{ij}}\cdot\dot{p}_{ij}\right) + \sum_j\left(\frac{\partial\tilde{\mathcal{E}}}{\partial\eta_j}\cdot\dot{\eta}_j\right)$$
$$= \sum_{ij}\left(\eta_j\frac{\partial\mathcal{U}^*}{\partial q_{ij}}\frac{p_{ij}}{M_{ij}} - \frac{p_{ij}}{M_{ij}}\left\{\frac{\partial\mathcal{U}^*}{\partial q_{ij}} + p_{ij}\eta_j\right\}\right) + \sum_j\eta_j\left(\sum_i\frac{p_{ij}^2}{M_{ij}} - g_j k_B T\right)$$
$$= -\sum_j \eta_j g_j k_B T \qquad \text{(A.7)}$$

Equating (A.5) and (A.7) yield

$$\frac{1}{f}\frac{df}{dt} = -\frac{1}{k_B T}\frac{d\tilde{\mathcal{E}}}{dt} \qquad \text{(A.8)}$$

which integrates to

$$f = const \times e^{-\tilde{\mathcal{E}}/k_B T} \qquad \text{(A.9)}$$

which has the form of a canonical distribution function. To make the formal identification with the canonical ensemble, we must however assume that the trajectories in the extended system as generated by Equation (2.26), is ergodic. This is necessary if we are to say that ensemble averages equals time averages, the essence of statistical mechanics.

Appendix B
Model data and simulation details

Table B.1
Geometry and physical data of TIP4P-water, OPLS-methanol, and OPLS-ethanol. X denotes CH_3/CH_2 for methanol/ethanol and H for water.

	water	**methanol**	**ethanol**
< X-O-H	104.52	108.5	108.5
< C-C-O	-	-	108
L_{OM} [Å]	0.15	–	-
L_{OH} [Å]	0.9572	0.945	0.945
L_{CO} [Å]	–	1.43	1.43
L_{CC} [Å]			1.53
Moment of inertia **I_{xx} [gÅ²]** **I_{yy} [gÅ²]** **I_{zz} [gÅ²]**	 $1.02054 \cdot 10^{-23}$ $2.93870 \cdot 10^{-23}$ $1.91816 \cdot 10^{-23}$	 $2.823468 \cdot 10^{-23}$ $2.945919 \cdot 10^{-23}$ $0.122451 \cdot 10^{-23}$	 $1.212316 \cdot 10^{-23}$ $9.099199 \cdot 10^{-23}$ $7.886886 \cdot 10^{-23}$
Mass of molecule, [g]	$2.99167 \cdot 10^{-23}$	$5.32082 \cdot 10^{-23}$	$7.64995 \cdot 10^{-23}$
Mass of molecule[amu]	18.02	32.0	46.1

Table B.2
Potential parameters for model of water (TIP4P) and methanol (OPLS).

model	site	σ [Å]	ε/k_B [K]	q/e
TIP4P	O	3.154	78.05	0
	H	0	0	0.52
	M	0	0	-1.04
OPLS-methanol	CH_3	3.775	104.22	0.265
	O	3.070	85.59	-0.700
	H	0	0	0.435
OPLS-ethanol	CH_3	3.905	88.11	0
	CH_2	3.905	59.41	0.265
	O	3.07	85.59	-0.700
	H	0	0	0.435

Table B.3
Simulation conditions for water-ethanol mixtures.

	NVT	NVE
Timestep [fs]	0.5	
Total steps	160000 (= 80ps)	
Equilibration steps	25000 velocity scaling	
System geometry	Cubic simulation box with PBC	
Total number of molecules	256	
Mole fraction alcohol	0.0, 0.25, 0.50, 0.75, 1.0	
No. ethanol molecules	0, 64, 128, 192, 256	
Lengths of box, L [Å]	**W/M:** 19.7305, 21.3739, 22.9346, 24.4200, 25.8781	
	W/E: 19.7233, 22.5931, 25.0833, 27.2344, 29.1666	
Long range forces	Ewald sum, $\kappa = 5/L$	
Initial start	fcc-lattice Maxwell distributed linear/angular velocities	
Numbers of heat res.	2 (trans and rot)	
τ_{trans} [fs] water methanol ethanol	25 25 50	
τ_{rot} [fs] water methanol ethanol	20 23 50	
Data output	each 100th step	
Structural sampling	each 150th step	
Coarse grain intervals	450, each based on 300 steps	
Potential model	TIP4P-water/OPLS-alcohol	

Appendix C
Further results water-methanol

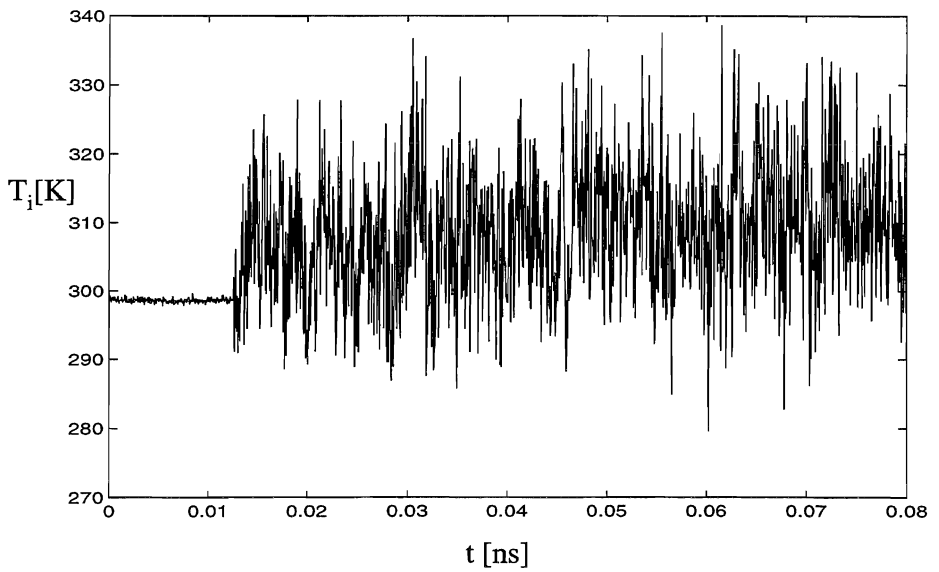


Figure C.1 Instantaneous total temperature T_i of pure water from NVE simulation. Each 100th step is plotted.

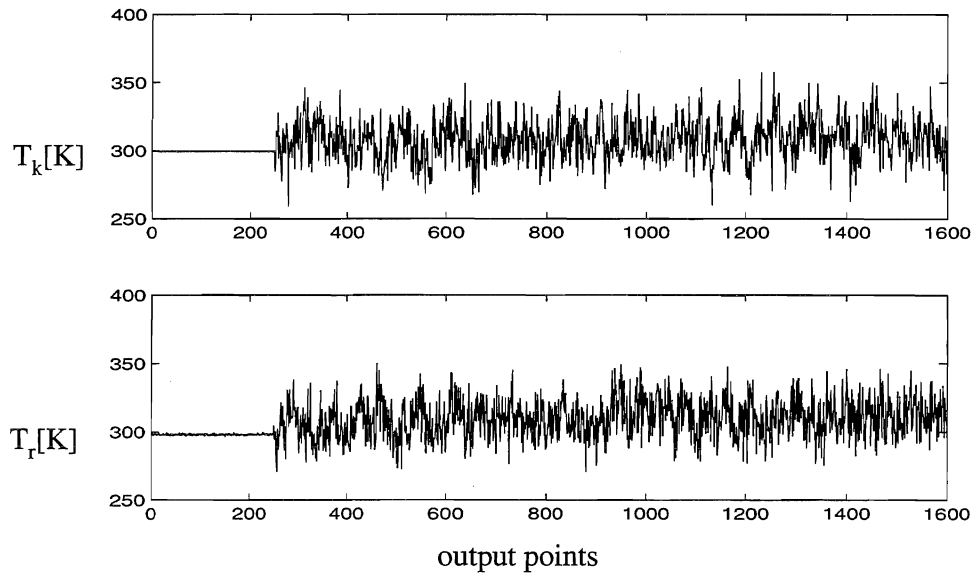


Figure C.2 Instantaneous translational (T_k) and rotational (T_r) temperatures from NVE simulation of pure water. 100 units on abcissa = 10000 timesteps.

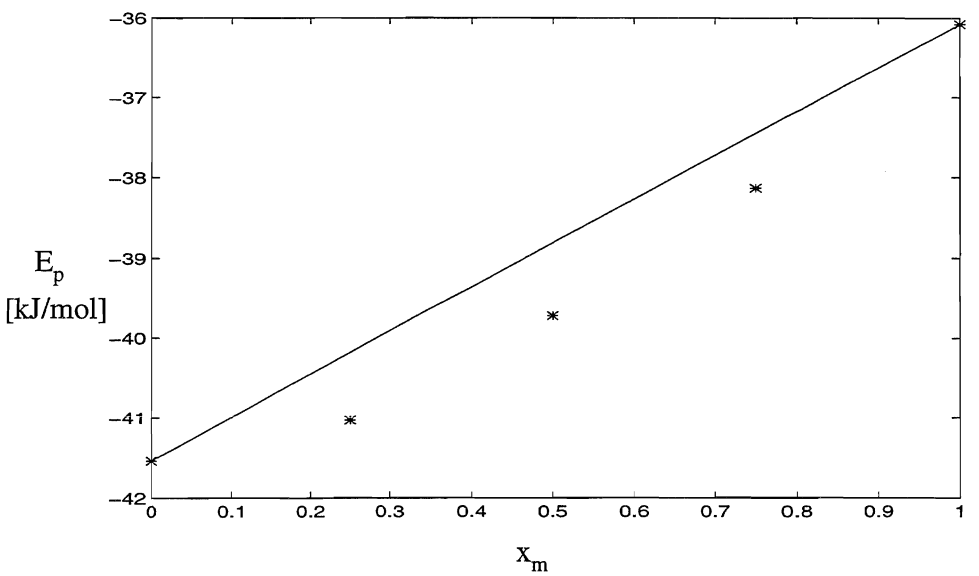


Figure C.3 Exess potential energy E_p for water-methanol mixtures at 298K as function of methanol mole fraction x_m. Straight line is for ideal solution, * are calculated values.

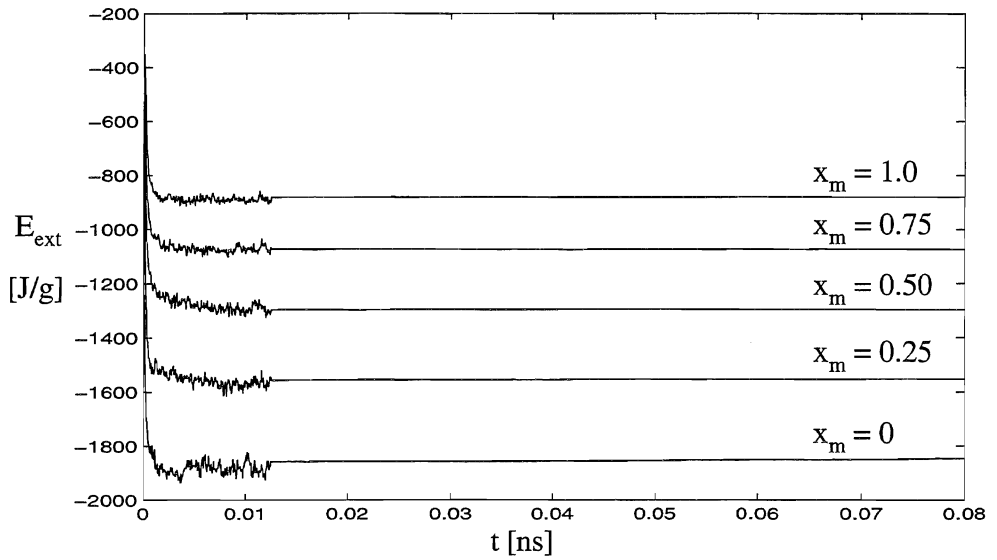


Figure C.4 Conservation of extended system energy E_{ext} in NVT-simulation. x_m is methanol mole fraction.

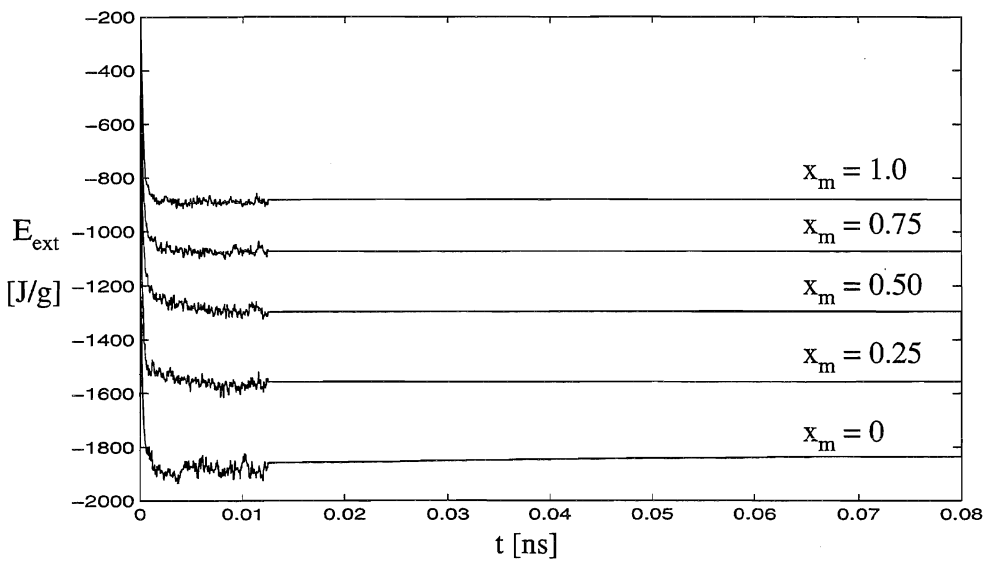


Figure C.5 Conservation of extended system energy in NVE-simulation. x_m is methanol mole fraction.

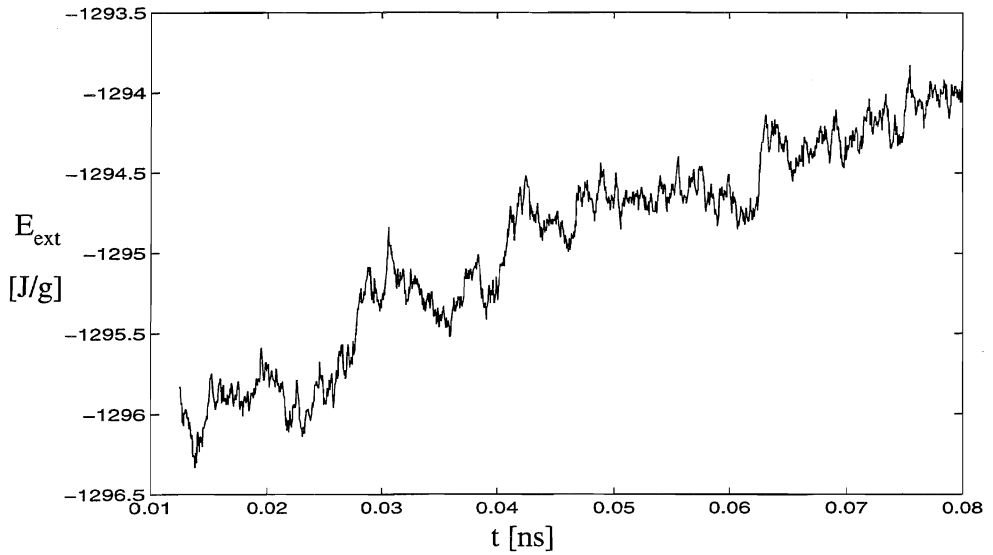


Figure C.6 Typical progress of extended system energy. Shown above is the last 135000 steps of the NVE-simulation of the equimolar mixture of water and methanol displayed in Figure C.5. Note the finer vertical scale.

Table C.1

Rotational and translational temperature for each species from NVE simulation. Averages and standard deviations in round brackets calculated from tabulated values with one decimal taken at intervals of 100 timesteps. There is thus a slight disagreement between T_{mix} above and mixture temperature as given in Table 5.1. Standard deviation is for distribution of single points.

x_m	water		methanol		$\langle T_{mix} \rangle$
	$\langle T_k \rangle$	$\langle T_r \rangle$	$\langle T_k \rangle$	$\langle T_r \rangle$	
0.0	307.8 (14.8)	309.5 (14.1)			308.6 (9.3)
0.25	301.6 (16.8)	298.4 (16.7)	300.2 (29.0)	300.3 (31.1)	300.0 (9.2)
0.50	297.9 (20.1)	295.5 (19.9)	297.1 (20.4)	296.2 (19.8)	296.6 (8.5)
0.75	299.5 (29.9)	299.8 (27.6)	299.4 (16.6)	300.3 (15.8)	299.7 (8.7)
1.0			305.4 (13.5)	305.4 (14.2)	305.3 (8.9)

Table C.2

Rotational and translational temperature for each species from NVT simulation. See text of Table C.1 above.

x_e	water		methanol		$\langle T_{mix} \rangle$
	$\langle T_k \rangle$	$\langle T_r \rangle$	$\langle T_k \rangle$	$\langle T_r \rangle$	
0.0	298.1 (15.2)	297.0 (15.3)			297.5 (10.9)
0.25	298.1 (17.7)	297.6 (17.7)	298.6 (29.4)	297.7 (29.6)	297.8 (10.6)
0.50	297.6 (21.7)	296.8 (21.7)	298.1 (21.5)	298.8 (21.3)	297.7 (10.6)
0.75	297.5 (30.4)	297.1 (31.1)	298.6 (17.8)	298.6 (16.9)	298.2 (10.5)
1.0			298.2 (15.4)	297.9 (15.3)	298.0 (11.0)

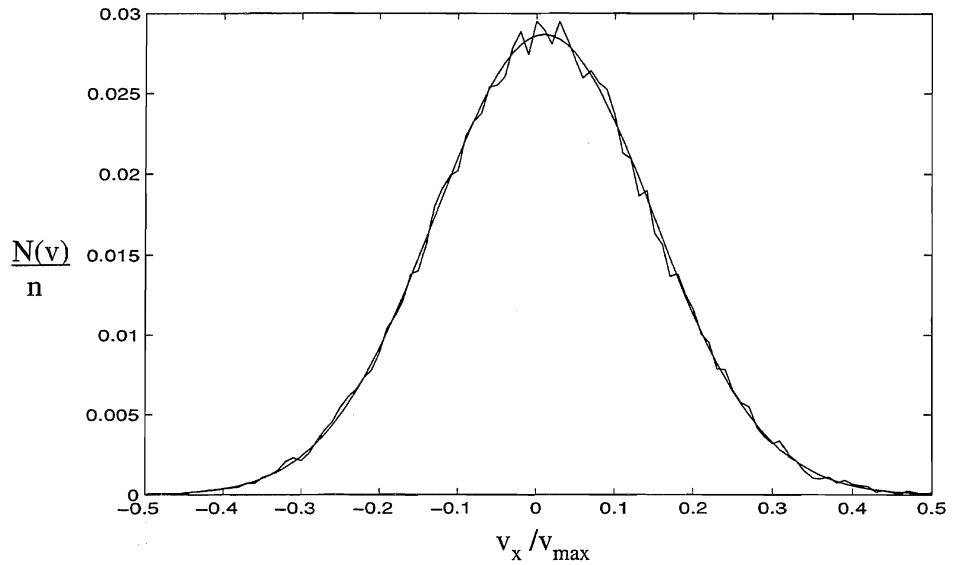


Figure C.7 Velocity distribution N(v)/n of 64 methanol molecules in a NVT-simulation of a 25% mixture with water at 298K. n is number of methanol molecules. Smooth line is Maxwell distribution. v_{max}=2000m/s. Structural sampling each 150th step. Uncertainty in x-direction is 0.01.

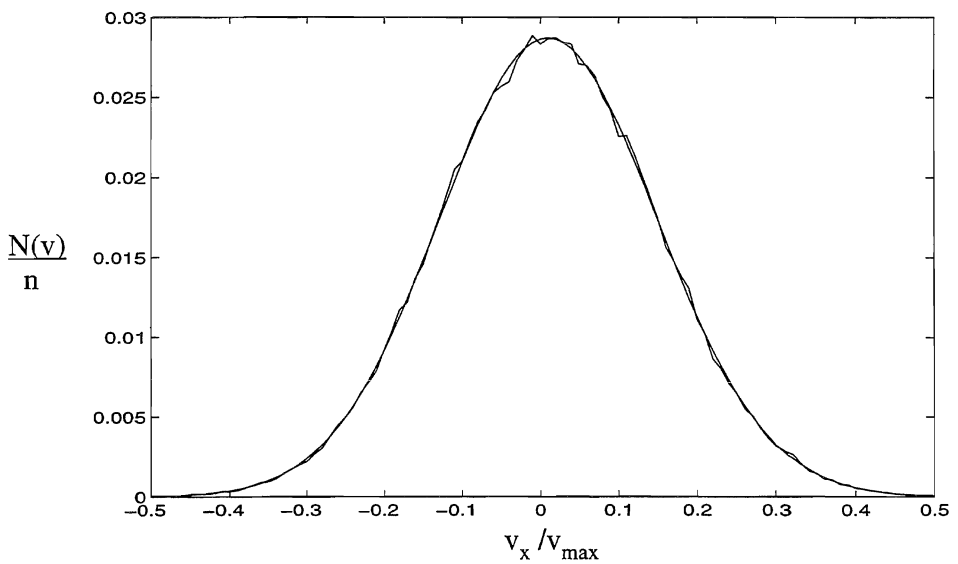


Figure C.8 Velocity distribution N(v) of 256 methanol molecules in a NVT-simulation of pure methanol. See text to Figure C.7.

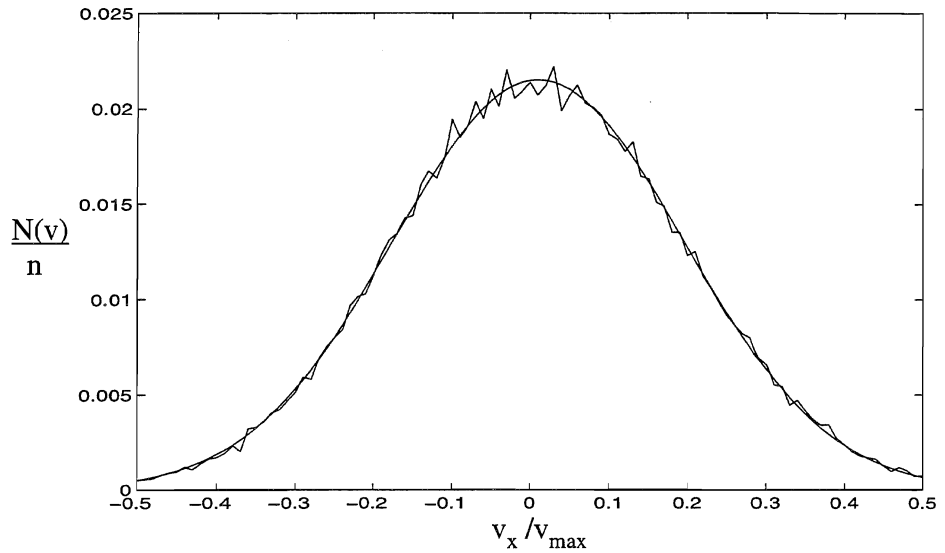


Figure C.9 Velocity distribution N(v) of 64 water molecules in an NVT-simulation of a 75% mixture of metanol in water. n is number of water molecules. See text to Fig C.7

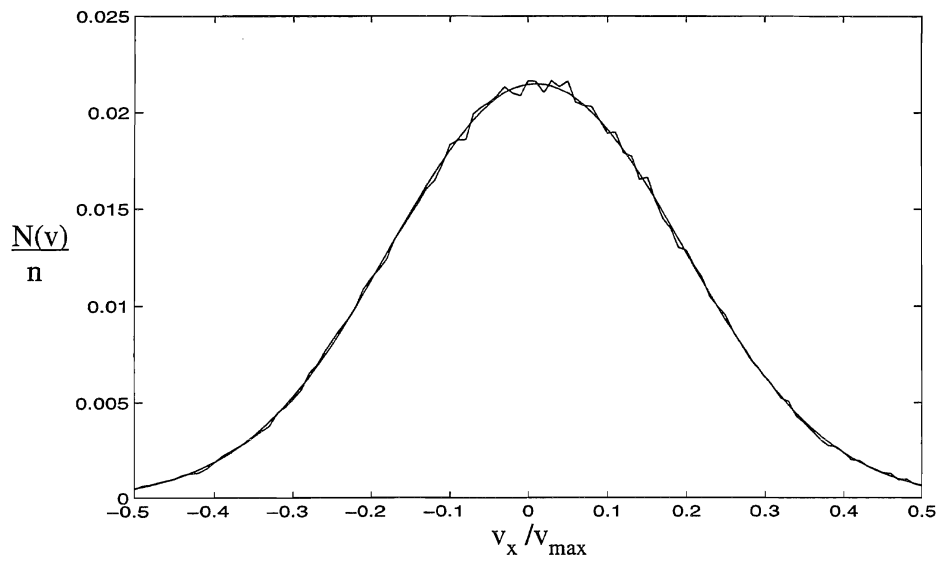


Figure C.10 Velocity distribution N(v) of 256 water molecules in an NVT-simulation of pure water. n is number of water molecules. See text to Figure C.7.

Table C.3
Linear momentum per mass, p_ζ, and translational energy per mass, K_ζ, for each direction with NVT-simulations.

x_m	x-direction				y-direction				z-direction			
	$p_x[10^{-13}m/s]$		$2K_x$ [J/g]		$p_y[10^{-13}m/s]$		$2K_y$ [J/g]		$p_z[10^{-13}m/s]$		$2K_z$ [J/g]	
0	-2.501	*2.302*	136.9	*11.8*	3.021	*2.548*	136.9	*11.9*	7.938	*6.622*	137.2	*11.5*
0.25	-3.245	*3.398*	115.2	*9.9*	0.178	*2.696*	114.4	*9.8*	3.548	*3.652*	115.0	*10.0*
0.50	-1.460	*2.609*	98.5	*8.2*	-3.024	*2.172*	98.5	*8.3*	5.662	*4.071*	98.8	*8.2*
0.75	-0.968	*2.328*	86.5	*7.4*	-1.118	*1.860*	87.1	*7.6*	-1.302	*1.328*	86.4	*7.5*
1.0	-1.057	*1.699*	77.1	*6.6*	-2.578	*3.458*	77.0	*6.3*	0.661	*0.997*	77.1	*6.7*

Table C.4
Linear momentum per mass, p_ζ, and translational energy per mass, K_ζ, for each direction from NVE-simulations.

x_m	x-direction				y-direction				z-direction			
	$p_x[10^{-13}m/s]$		$2K_x$ [J/g]		$p_y[10^{-13}m/s]$		$2K_y$ [J/g]		$p_z[10^{-13}m/s]$		$2K_z$ [J/g]	
0	1.096	2.120	140.8	11.7	1.942	2.969	140.6	11.8	3.030	2.771	141.2	12.2
0.25	2.855	2.646	115.8	9.4	-2.636	2.427	116.0	10.0	4.411	2.474	115.4	9.6
0.50	-6.140	4.204	98.3	8.1	1.236	2.142	98.1	7.9	1.112	1.431	99.1	7.8
0.75	0.137	1.246	86.8	7.2	2.775	2.290	86.8	7.4	0.492	0.699	86.9	7.4
1.0	1.266	1.325	78.9	6.7	-3.624	2.073	78.8	6.3	-0.981	1.113	78.5	6.6

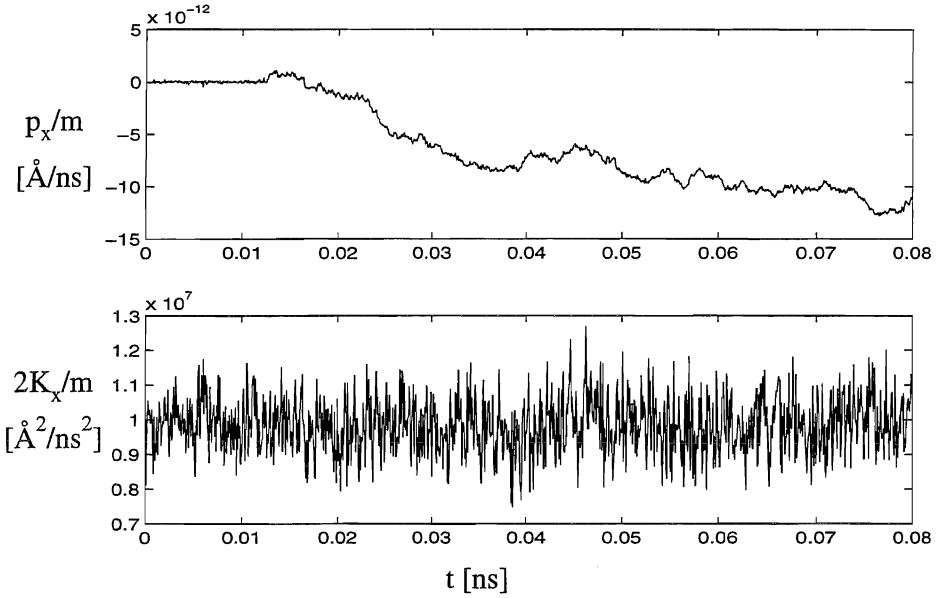


Figure C.11 X-component of linear momentum (top) and translational energy (bottom) from NVE simulation of an equimolar mixture of water and methanol. m is system mass.

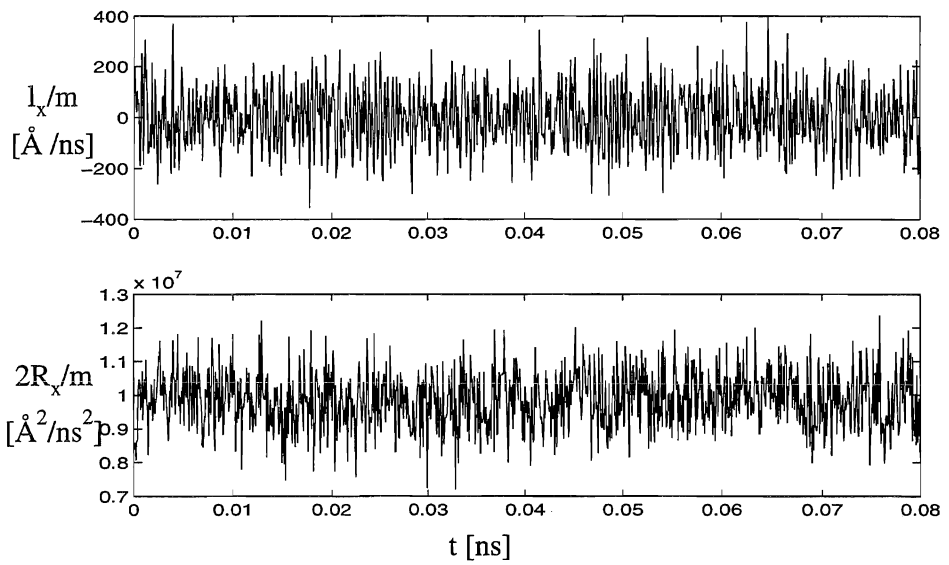


Figure C.12 X-component of angular momentum(top) and rotational energy (bottom) for NVE simulation of an equimolar mixture of water and methanol. m is system mass.

Table C.5
Mean and standard deviation of angular momentum per mass, L_ζ and rotational energy per mass, R_ζ in NVE-simulations. x_m is mole fraction methanol.

x_m	x-direction				y-direction				z-direction			
	$L_x[10^{-11}m^2/s]$		$2R_x[J/g]$		$L_y[10^{-11}m^2/s]$		$2R_y[J/g]$		$L_z[10^{-11}m^2/s]$		$2R_z[J/g]$	
0	0.713	*44.7*	141.9	*11.9*	1.384	*71.9*	143.2	*11.9*	2.254	*60.0*	141.4	*11.6*
0.25	0.072	*104.6*	115.7	*9.8*	-0.219	*113.3*	115.2	*9.5*	-0.258	*45.6*	115.6	*9.6*
0.50	1.136	*116.6*	98.6	*8.3*	-0.463	*119.9*	97.9	*8.3*	-1.378	*38.5*	98.7	*8.0*
0.75	0.591	*125.7*	87.5	*7.1*	1.042	*125.6*	87.5	*7.4*	2.322	*32.1*	86.7	*6.9*
1.0	-0.406	*130.8*	78.9	*6.9*	-0.106	*131.1*	78.7	*6.6*	0.206	*26.8*	78.9	*6.6*

Table C.6
Mean and standard deviation of angular momentum per mass, L_ζ and rotational energy per mass, R_ζ in NVT-simulations. x_m is mole fraction methanol.

x_m	x-direction				y-direction				z-direction			
	$L_x[10^{-11}m^2/s]$		$2R_x[J/g]$		$L_y[10^{-11}m^2/s]$		$2R_y[J/g]$		$L_z[10^{-11}m^2/s]$		$2R_z[J/g]$	
0	0.515	*41.6*	137.1	*11.6*	0.323	*70.3*	137.4	*11.9*	1.084	*57.2*	137.5	*11.7*
0.25	-3.469	*100.2*	114.8	*10.3*	-3.798	*105.7*	115.6	*9.8*	1.972	*46.2*	114.9	*9.8*
0.50	-2.065	*113.8*	99.0	*8.5*	-1.373	*121.5*	98.9	*8.6*	-0.472	*37.8*	98.6	*8.4*
0.75	-1.773	*118.4*	86.6	*7.4*	1.950	*130.6*	86.3	*7.4*	1.721	*33.2*	86.9	*7.5*
1.0	-2.855	*127.5*	77.4	*6.5*	2.784	*136.5*	77.7	*6.7*	0.505	*24.8*	77.2	*6.9*

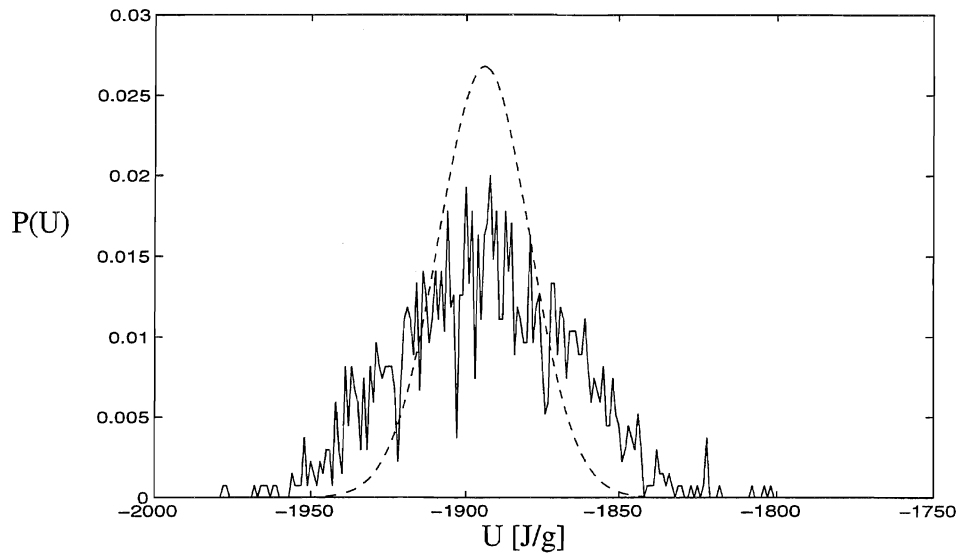


Figure C.13 Distribution of internal energy P(U) for pure water from NVT simulation. Dashed line is canonical distribution at 298.15K.

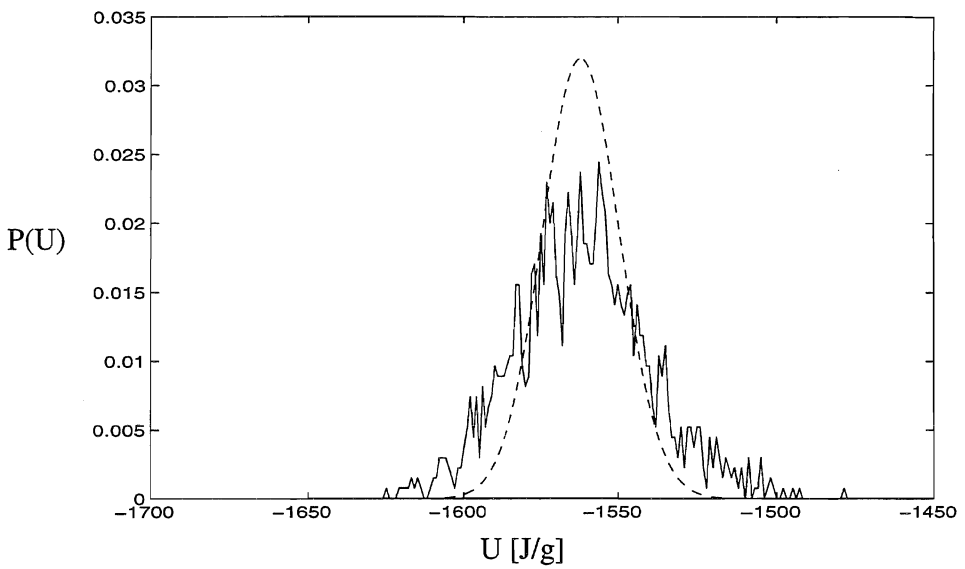


Figure C.14 Distribution of internal energy P(U) for a mixture with x_m=0.25 in NVT-simulation. Dashed line is canonical distribution at 298.15K.

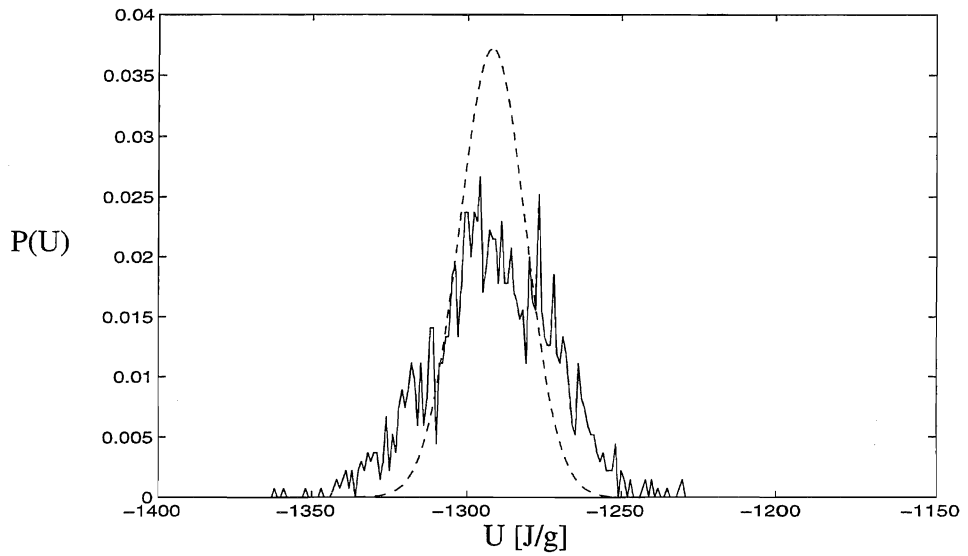


Figure C.15 Distribution of internal energy P(U) for a mixture with x_m=0.5 from NVT-simulation. Dashed line is canonical distribution at 298.15K.

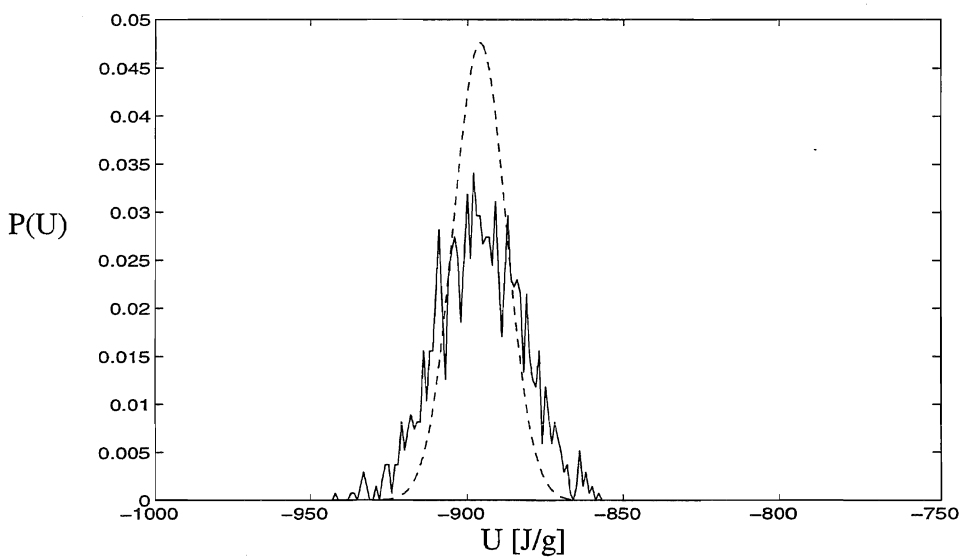


Figure C.16 Distribution of internal energy P(U) in NVT-simulation for pure methanol. Dashed line is canonical distribution at 298.15K.

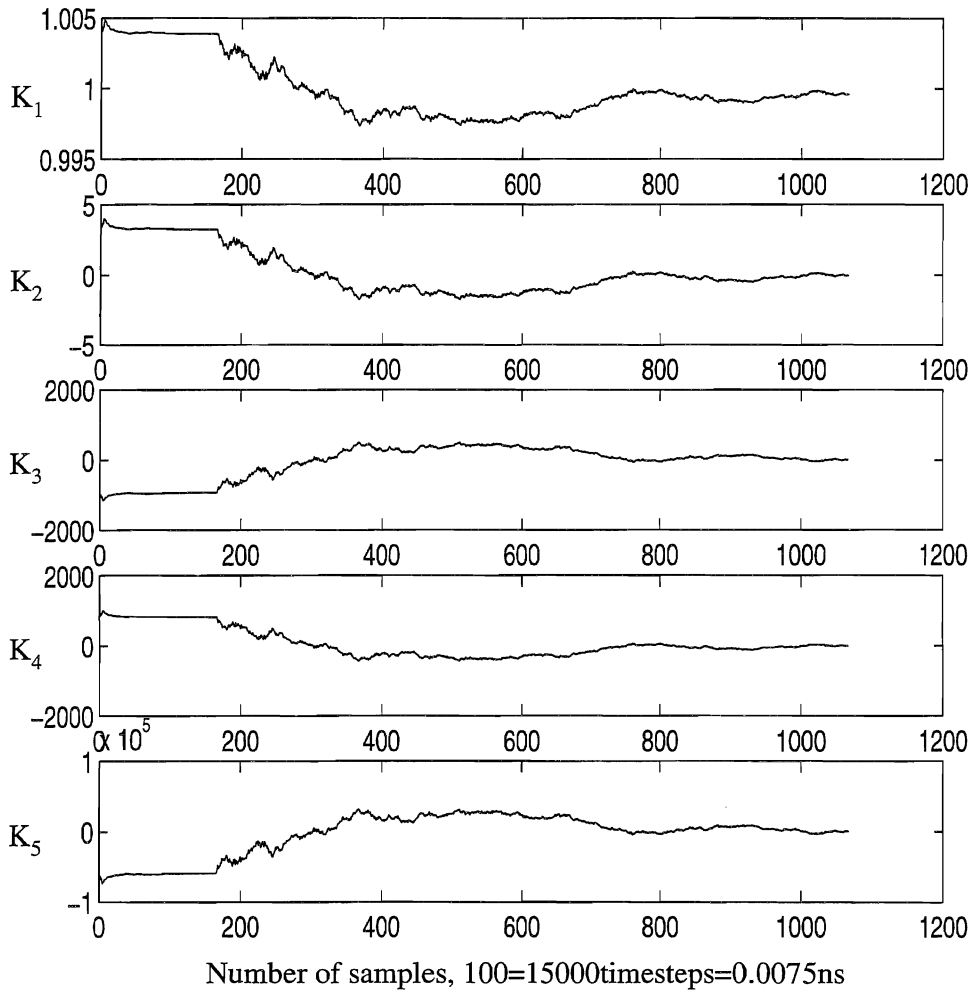


Figure C.17 Running averages of 1st (top) to 5th (bottom) moments of translational energy for a NVT simulation of pure water. Simulation conditions as in table B.3, page 3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see Equation (3.39).

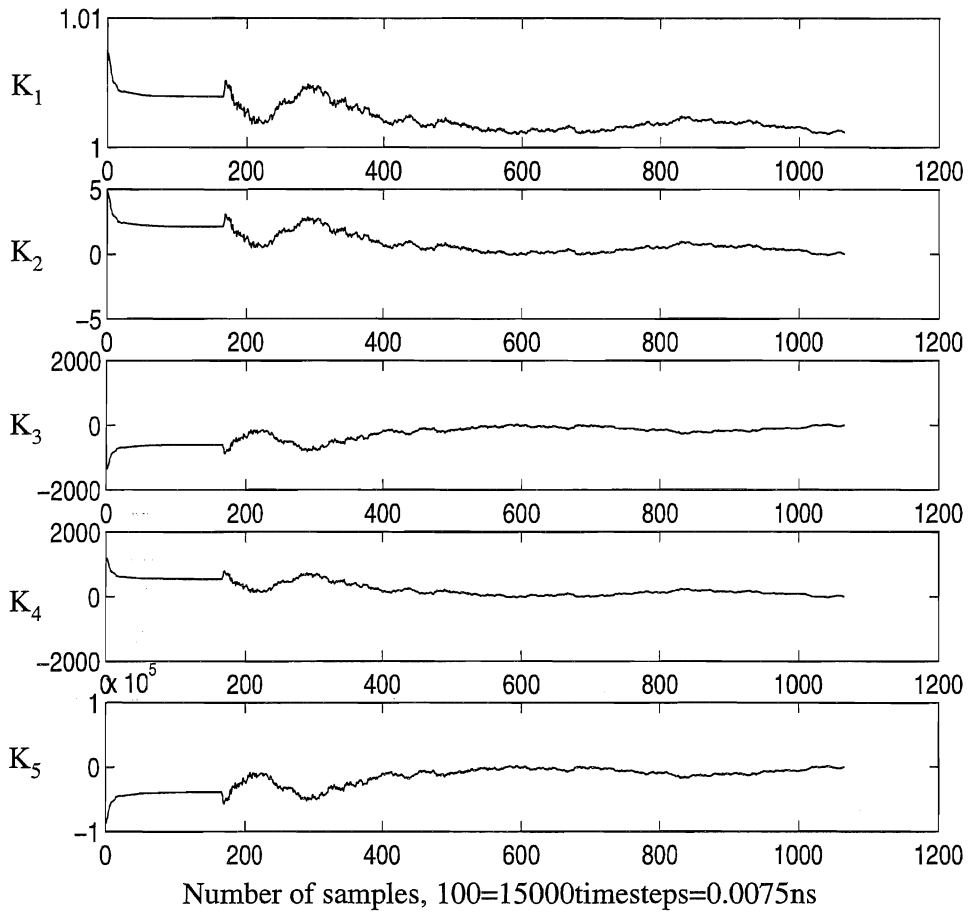
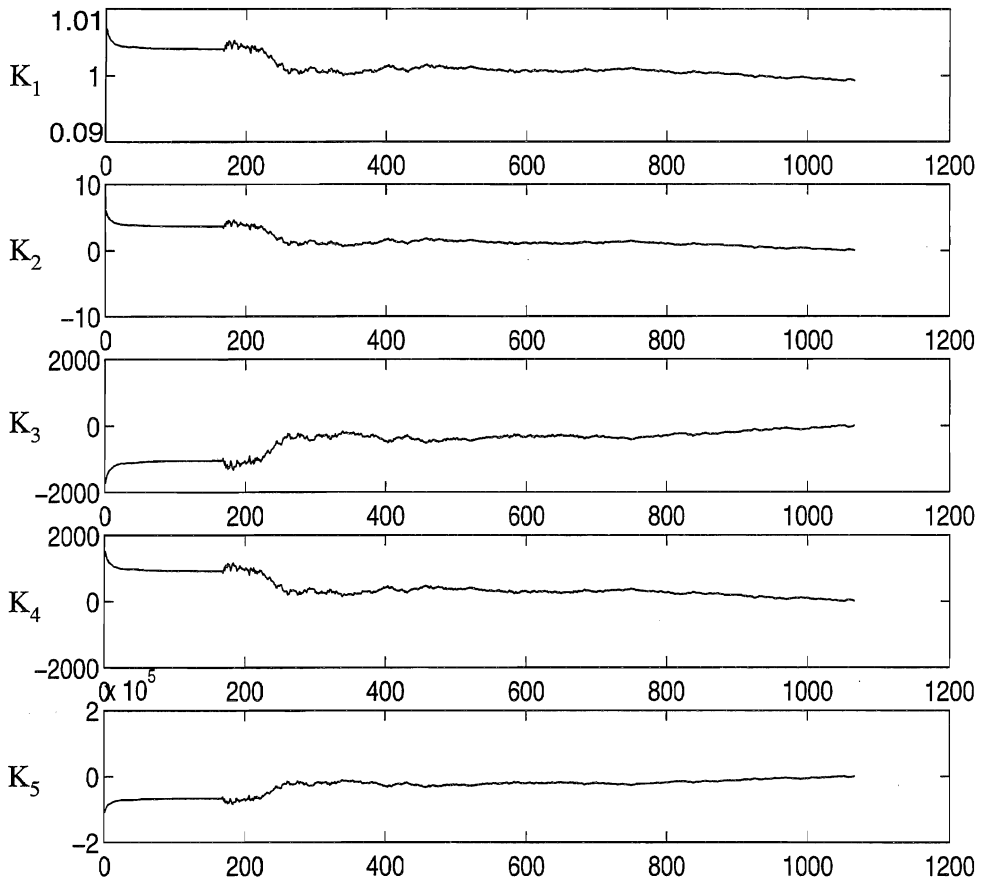


Figure C.18 Running averages of 1st (top) to 5th (bottom) moment of translational energy for a NVT simulation of 0.25 mole fraction methanol in water. Simulation conditions as in table B.3, page 3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see Equation (3.39).



Number of samples, 100=15000timesteps=0.0075ns

Figure C.19 Running averages of 1st (top) to 5th (bottom) moment of translational energy for a NVT simulation of 0.50 mole fraction methanol in water. Simulation conditions as in table B.3, page 3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see Equation (3.39).

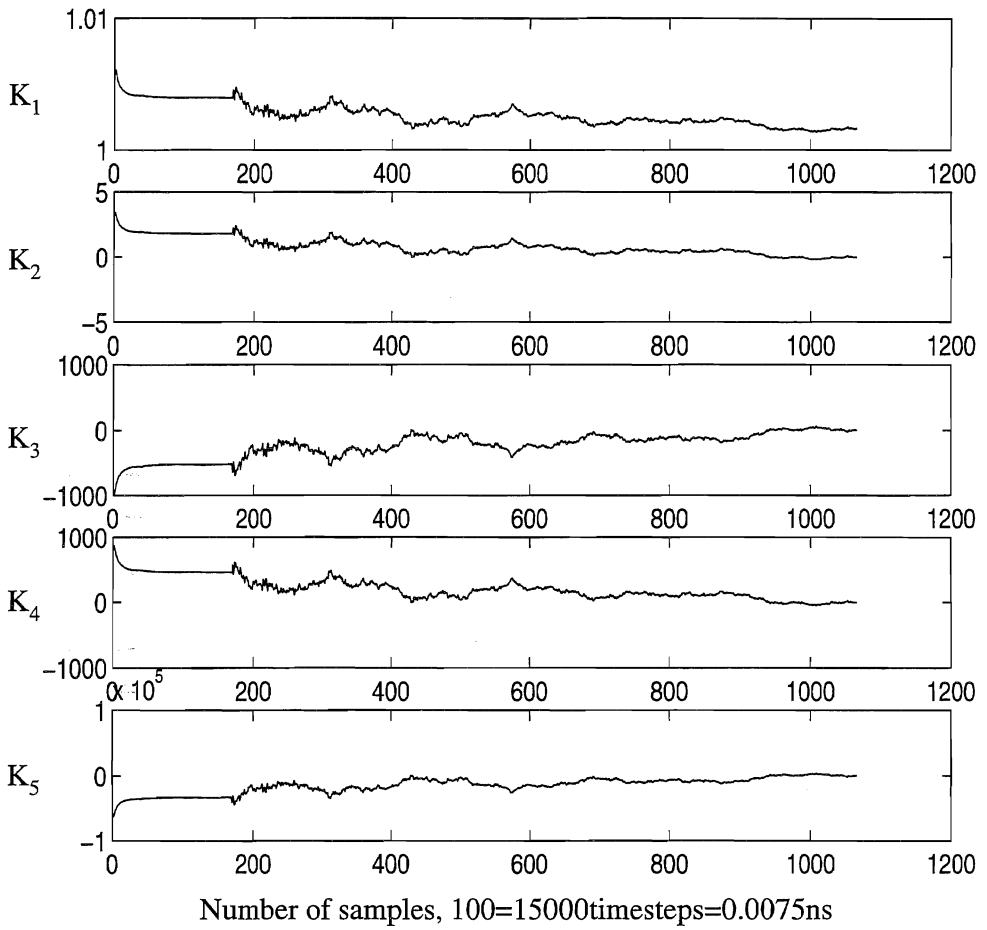


Figure C.20 Running averages of 1st (top) to 5th (bottom) moment of translational energy for a NVT simulation of 0.75 mole fraction methanol in water. Simulation conditions as in table B.3, page 3.Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see Equation (3.39).

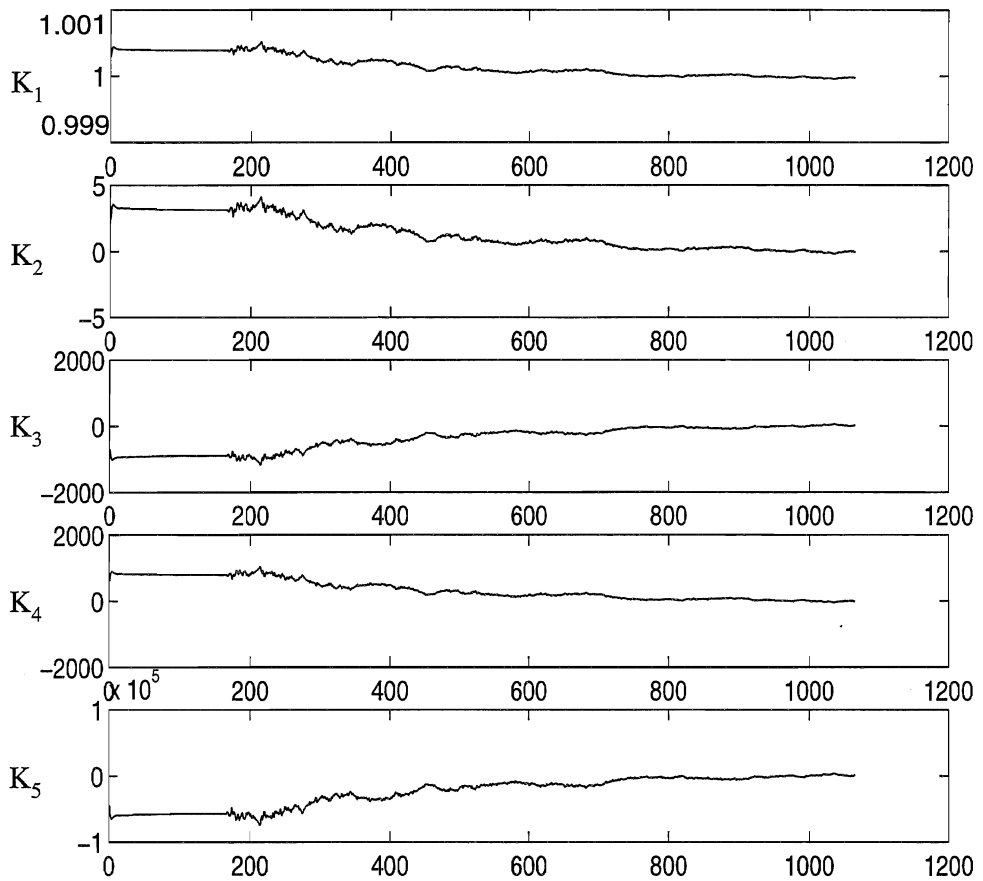


Figure C.21 Running averages of 1st (top) to 5th (bottom) moment of translational energy for a NVT simulation of pure methanol. Simulation conditions as in table B.3, page 3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see Equation (3.39).

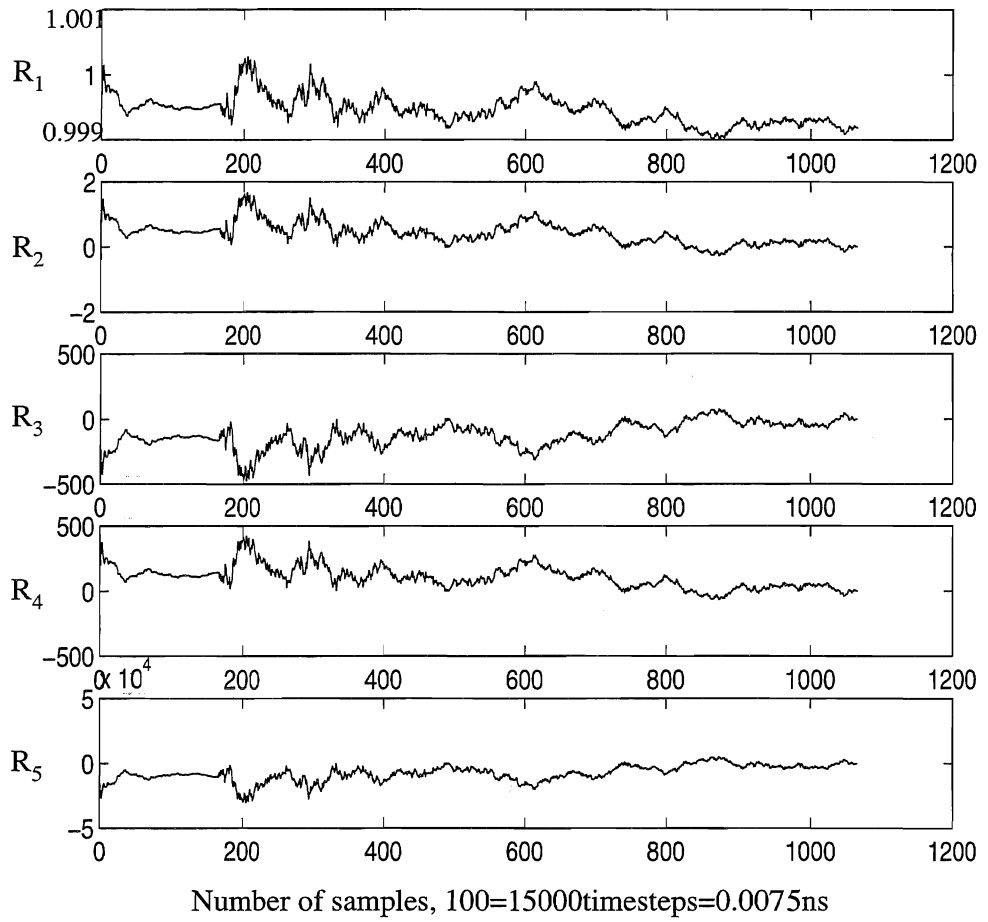


Figure C.22 Running averages of 1st (top) to 5th (bottom) moments of rotational energy for a NVT simulation of pure water. Simulation conditions as in table B.3, page 3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see Equation (3.39).

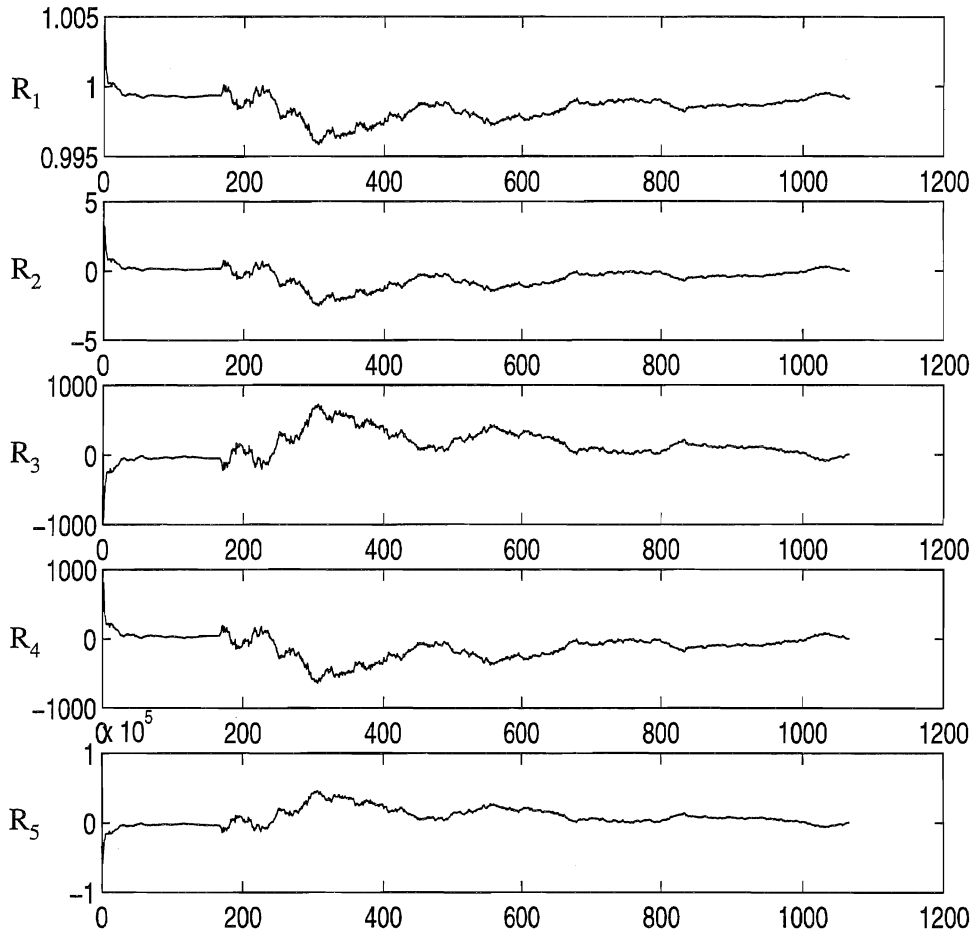


Figure C.23 Running averages of 1st (top) to 5th (bottom) moment of rotational energy for a NVT simulation of 0.25 mole fraction methanol in water. Simulation conditions as in table B.3, page 3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see Equation (3.39).

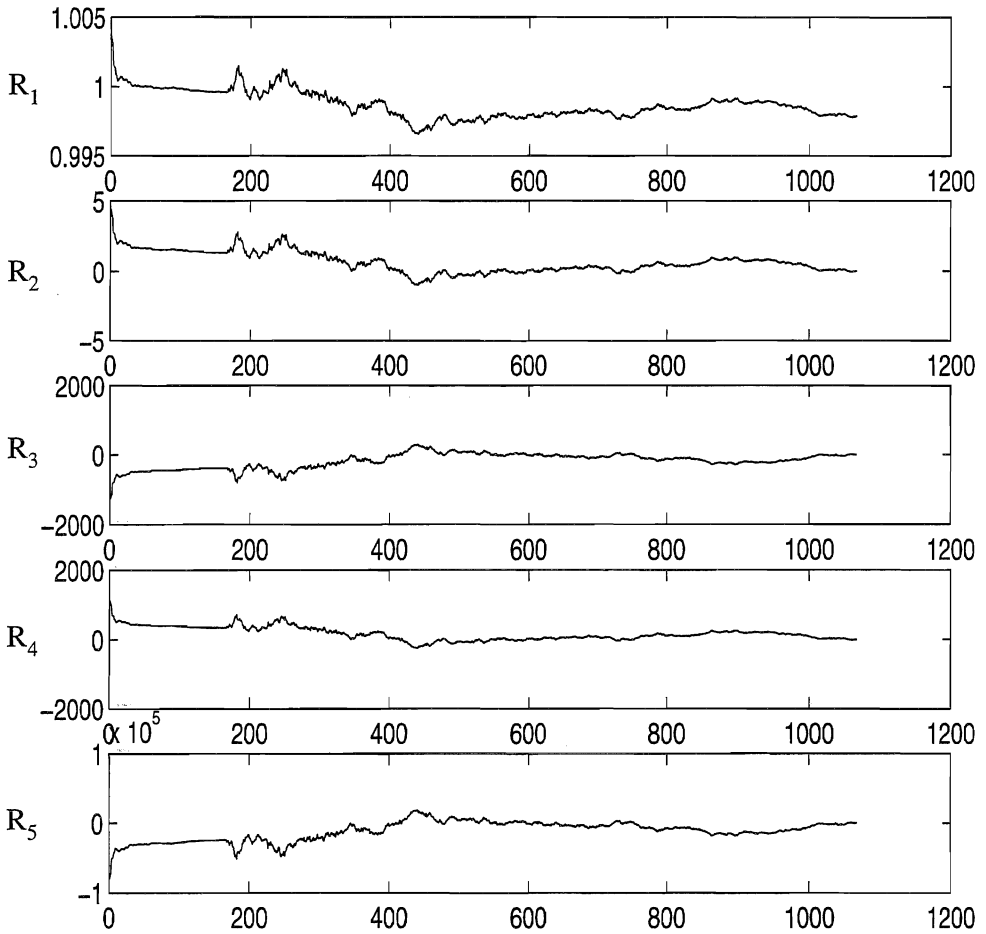


Figure C.24 Running averages of 1st (top) to 5th (bottom) moment of rotational energy for a NVT simulation of 0.50 mole fraction methanol in water. Simulation conditions as in table B.3, page 3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see Equation (3.39).

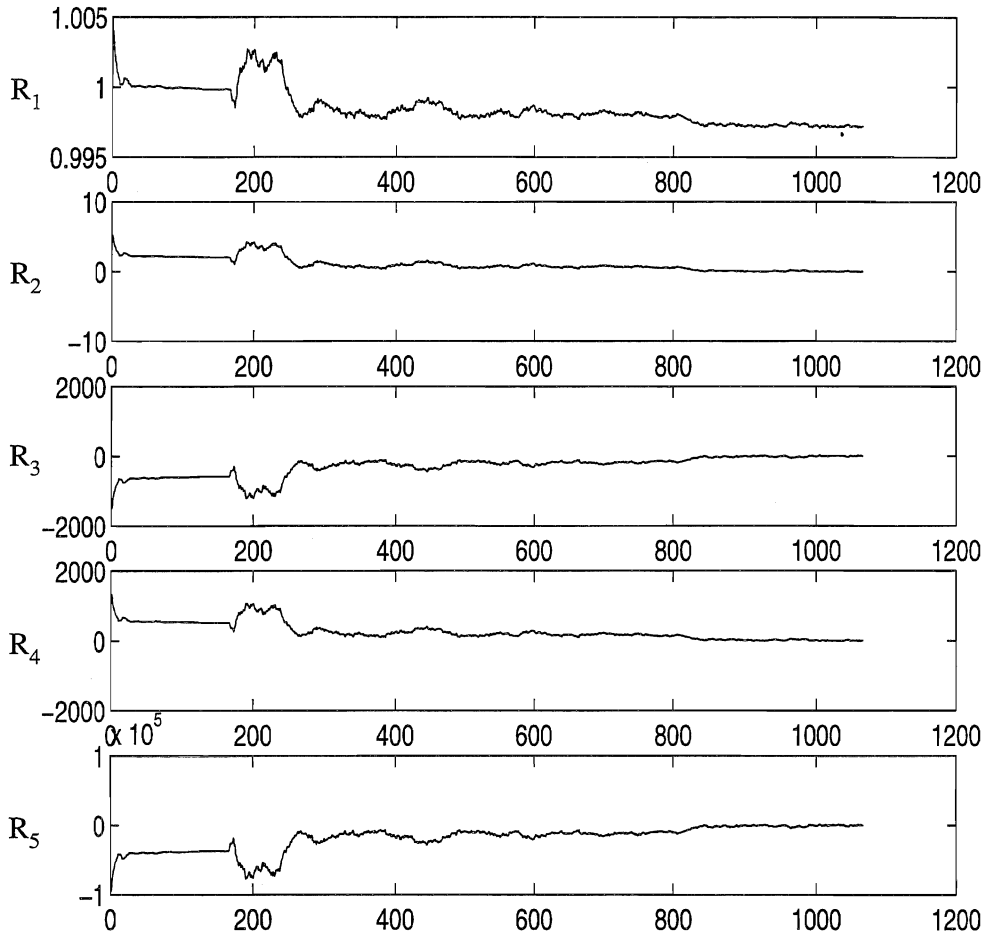


Figure C.25 Running averages of 1st (top) to 5th (bottom) moment of rotational energy for a NVT simulation of 0.75 mole fraction methanol in water. Simulation conditions as in table B.3, page 3.Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see Equation (3.39).

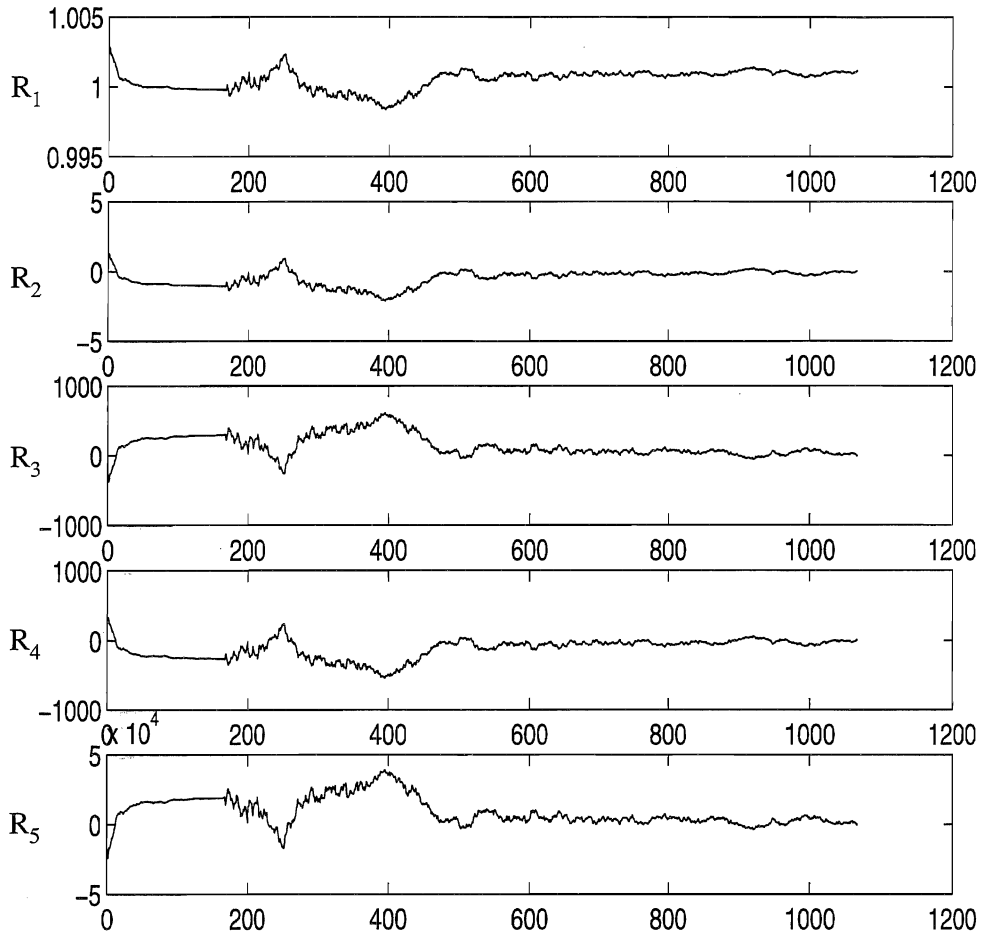


Figure C.26 Running averages of 1st (top) to 5th (bottom) moment of rotational energy for a NVT simulation of pure methanol. Simulation conditions as in table B.3, page 3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see Equation (3.39).

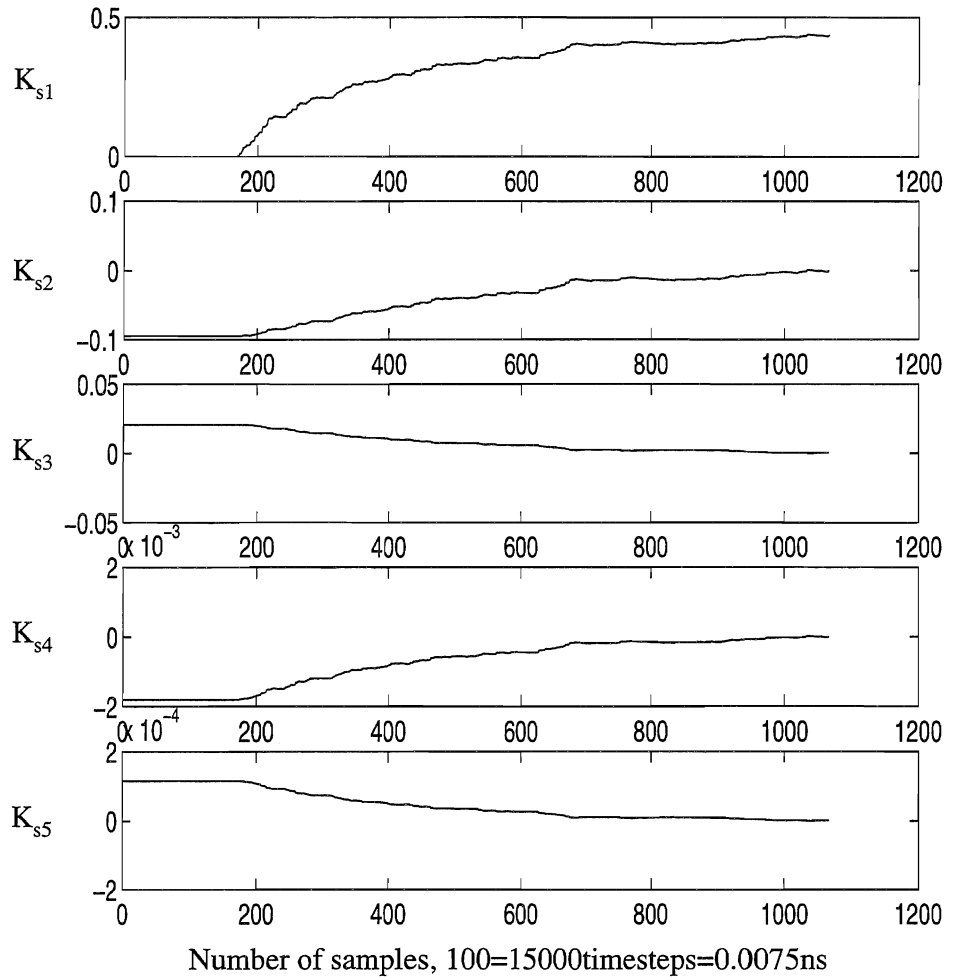


Figure C.27 Running averages of 1st (top) to 5th (bottom) moment of reservoir translational energy for a NVT simulation of 0.50 mole fraction methanol in water. Simulation conditions as in table B.3, page 3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and the values are divided by theoretical values, see Equation (3.39).

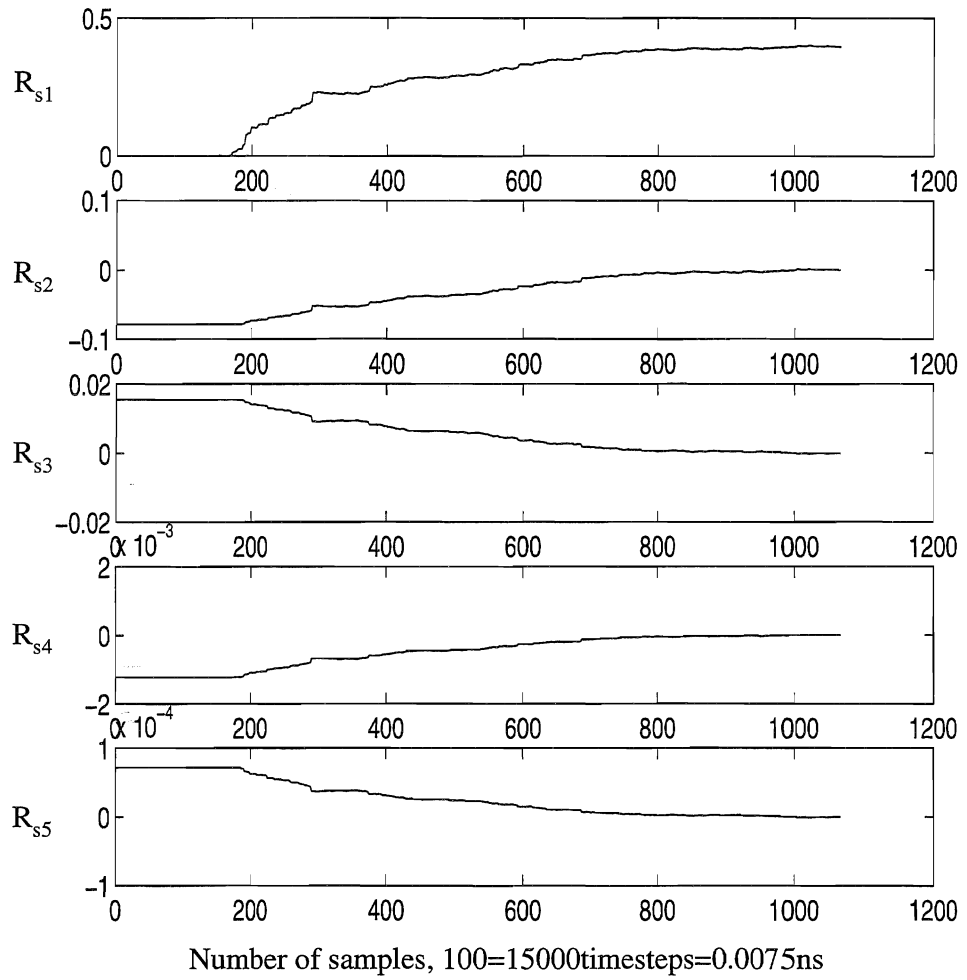


Figure C.28 Running averages of 1st (top) to 5th (bottom) moment of reservoir rotational energy for a NVT simulation of 0.50 mole fraction methanol in water. Simulation conditions as in table B.3, page 3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see Equation (3.39)

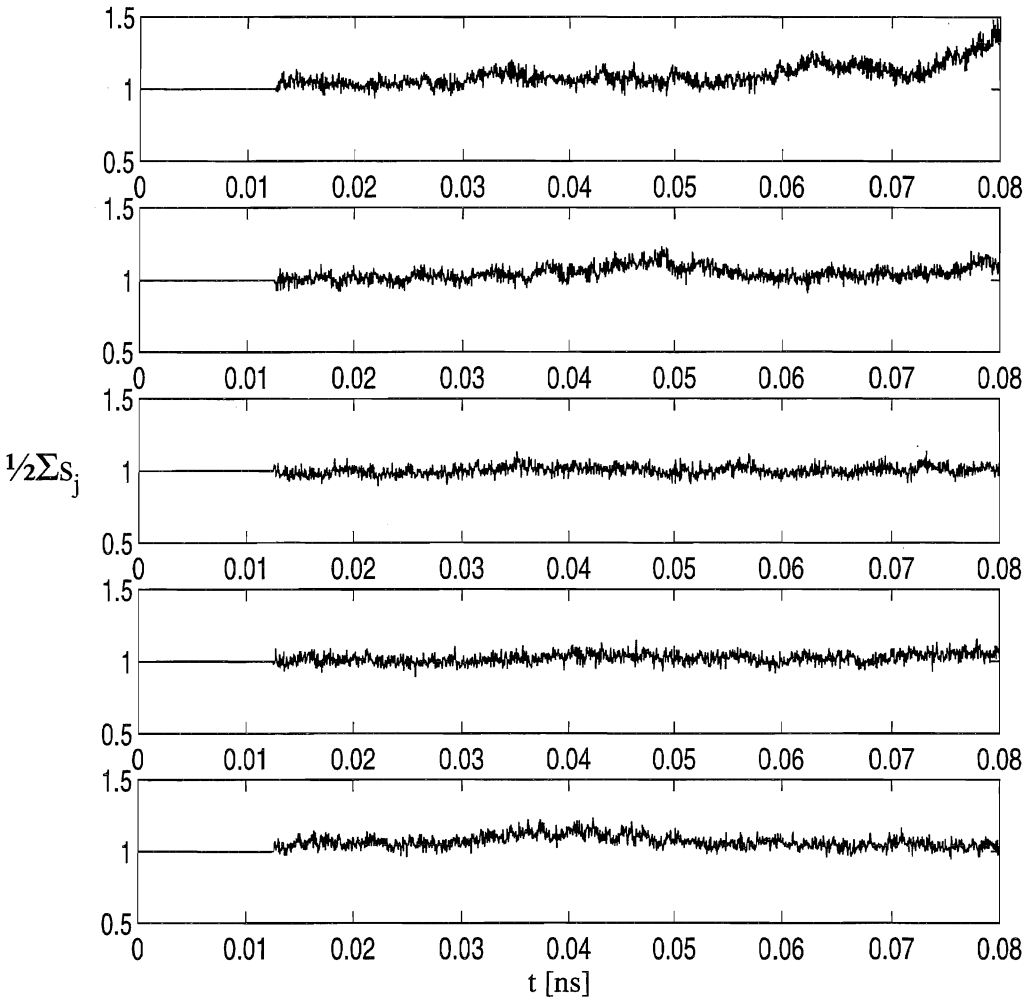


Figure C.29 Average $(s_{trans} + s_{rot})/2$ of heat bath variables for all mixtures, ranging from pure water (top) to pure methanol (bottom). Figure cover whole simulation.

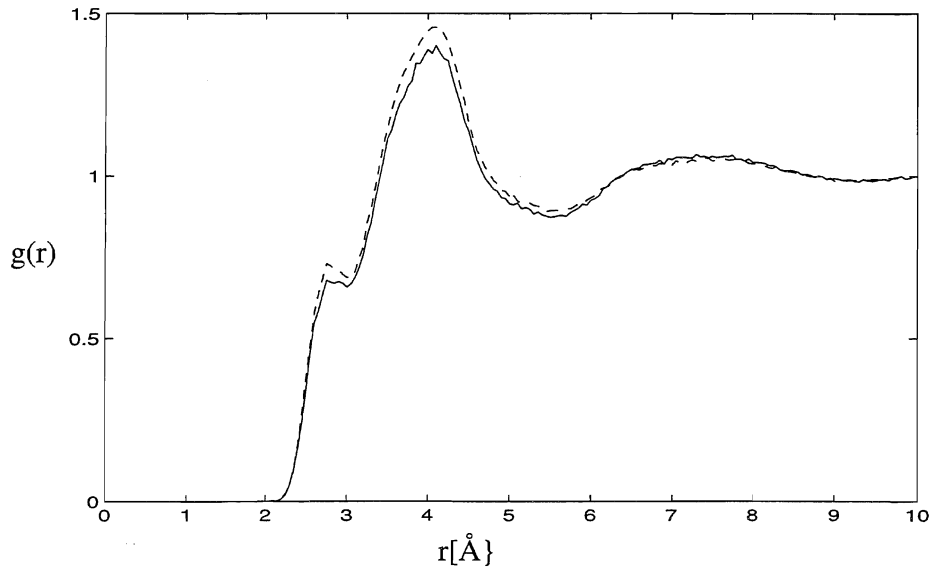


Figure C.30 NVT (———) and NVE (- - - -) results for g_{HC}, the crosscorrelations between water and methanol, for an equimolar mixture.

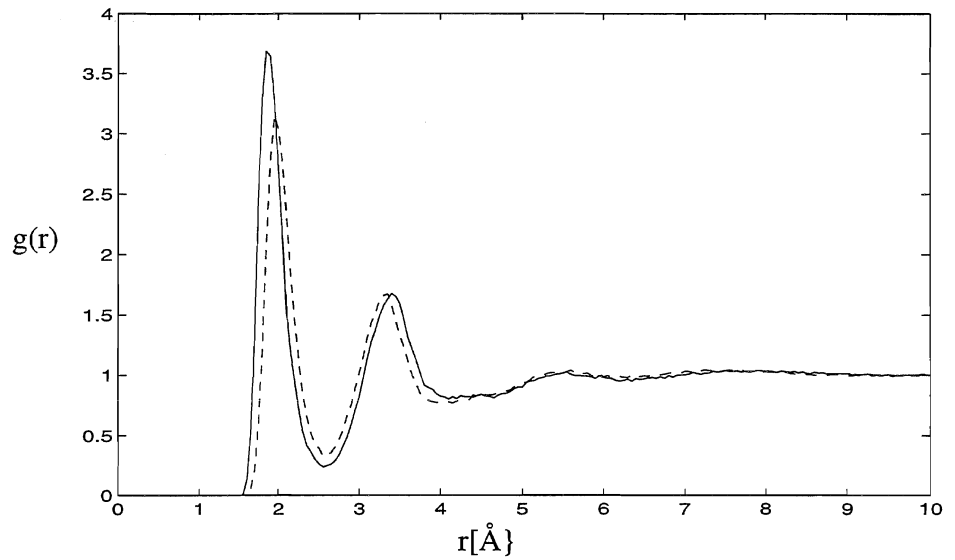


Figure C.31 Water-methanol crosscorrelations, g_{OH} (____) and g_{MH} (- - - -) in a 0.75 mole fraction methanol mixture.

Table C.7

Radial distance and height of 1st and 2nd maximum and of 1st minimum for site-site self correlations of water with varying methanol content. Results for 298K from NVT-simulations with simulation conditions as given in table B.3.

x_m	g	1.maximum		1.minimum		2.maximum	
		r[Å]	Height	r[Å]	Depth	r[Å]	Height
0	g_{OO}	2.8	2.99	3.5	0.81	4.55	1.11
	g_{OH}	1.85	1.46	2.5	0.21	3.25	1.51
	g_{HH}	2.4	1.24	3.0	0.75	3.8	1.17
0.25	g_{OO}	2.8	4.09	3.45	0.80	4.55	1.13
	g_{OH}	1.85	2.04	2.55	0.23	3.25	1.85
	g_{HH}	2.35	1.64	3.05	0.84	3.8	1.27
0.50	g_{OO}	2.8	5.48	3.5	0.75	4.7	1.14
	g_{OH}	1.85	2.82	2.55	0.25	3.25	2.25
	g_{HH}	2.40	2.16	3.1	0.95	3.75	1.38
0.75	g_{OO}	2.75	6.16	3.55	0.53	5.15	1.12
	g_{OH}	1.85	3.20	2.55	0.23	3.3	2.31
	g_{HH}	2.45	2.37	3.15	0.82	3.75	1.32

Table C.8

Radial distance and height of 1st and 2nd maximum and of 1st minimum for site-site self correlations of methanol with varying methanol content. Results for 298 K from NVT-simulations with simulation conditions as given in table B.3.
A weak maximum appearing in g_{CO} near 5 Å not regarded as a separate maximum, but see text on page 119.

x_m	g	1.maximum		1.minimum		2.maximum	
		r[Å]	Height	r[Å]	Depth	r[Å]	Height
0.25	g_{CC}	4.1	2.11	6.05	0.71	7.75	1.13
	g_{CO}	3.65	1.60	6.75	0.89	7.95	1.08
	g_{CH}	2.9	0.62	2.95	0.56	4.0	1.24
	g_{OO}	2.85	1.66	3.35	0.38	4.75	1.39
	g_{OH}	1.9	1.69	2.6	0.17	3.5	0.91
	g_{HH}	2.55	1.58	3.2	0.52	—	—
0.50	g_{CC}	4.1	2.02	6.0	0.72	7.65	1.11
	g_{CO}	3.65	1.62	6.35	0.87	8.0	1.06
	g_{CH}	2.95	0.68	3.1	0.64	4.05	1.20
	g_{OO}	2.8	2.06	3.35	0.37	4.75	1.28
	g_{OH}	1.9	2.07	2.6	0.16	3.45	0.91
	g_{HH}	2.5	1.90	3.3	0.45	—	—
0.75	g_{CC}	4.1	1.97	5.9	0.71	7.8	1.14
	g_{CO}	3.65	1.75	6.55	0.87	7.75	1.06
	g_{CH}	2.90	0.81	3.25	0.71	4.2	1.22
	g_{OO}	2.8	2.59	3.45	0.35	4.85	1.27
	g_{OH}	1.9	2.68	2.55	0.16	3.45	0.97
	g_{HH}	2.5	2.22	3.3	0.44	—	—
1.00	g_{CC}	4.1	2.01	5.85	0.71	8.0	1.17
	g_{CO}	3.60	2.02	6.45	0.88	7.75	1.07
	g_{CH}	2.90	1.05	3.3	0.75	4.2	1.25
	g_{OO}	2.8	3.60	3.45	0.28	4.95	1.24
	g_{OH}	1.9	3.75	2.65	0.14	3.45	1.03
	g_{HH}	2.5	2.92	3.4	0.30	—	—

Table C.9
Radial distance and height of 1st and 2nd maximum and of 1st minimum for site-site cross-correlations of water with varying methanol content. Results for 298 K from NVT-simulations with simulation conditions as given in table B.3.

x_m	g	1.maximum		1.minimum		2.maximum	
		r[Å]	Height	r[Å]	Depth	r[Å]	Height
0.25	g_{OC}	3.60	2.03	5.15	0.81	7.35	1.06
	g_{OO}	2.75	2.47	3.40	0.53	5.25	1.18
	g_{OH}	1.90	1.77	2.60	0.21	3.40	1.23
	g_{HC}	4.10	1.32	5.65	0.86	7.70	1.03
	g_{HO}	1.80	1.62	2.55	0.15	3.25	1.19
	g_{HH}	2.45	1.48	3.15	0.60	-	-
0.50	g_{OC}	3.55	2.27	5.25	0.83	7.50	1.11
	g_{OO}	2.75	3.36	3.45	0.54	5.20	1.19
	g_{OH}	1.85	2.55	2.60	0.24	3.45	1.43
	g_{HC}	4.10	1.40	5.55	0.87	7.35	1.07
	g_{HO}	1.80	2.21	2.55	0.16	3.25	1.43
	g_{HH}	2.45	1.96	3.25	0.64	(3.80)	(0.97)
0.75	g_{OC}	3.55	2.78	5.00	0.81	7.45	1.13
	g_{OO}	2.75	4.72	3.60	0.46	5.30	1.18
	g_{OH}	1.85	3.69	2.60	0.24	3.45	1.68
	g_{HC}	4.10	1.56	5.55	0.85	7.70	1.07
	g_{HO}	1.80	3.10	2.55	0.17	3.20	1.79
	g_{HH}	2.45	2.58	3.20	0.61	(3.95)	(0.99)

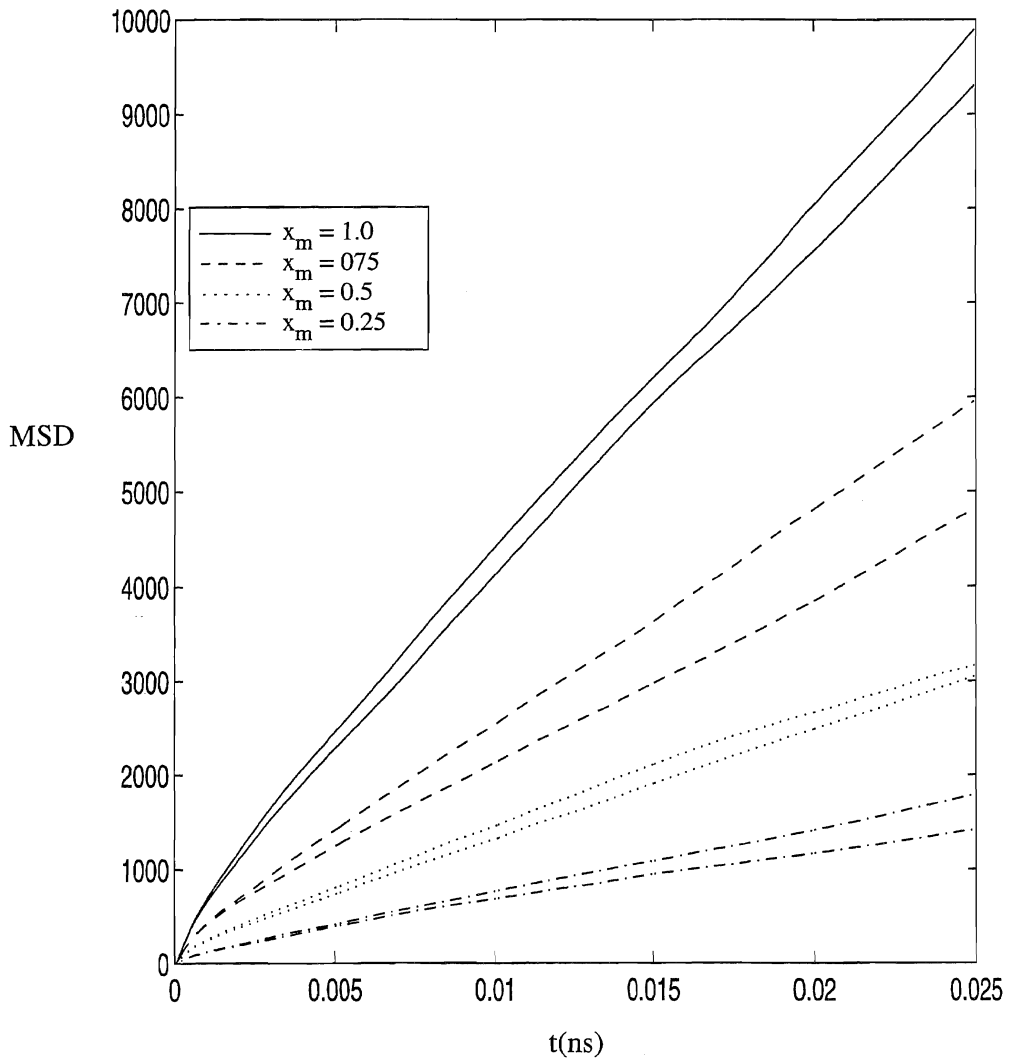


Figure C.32 Mean square displacement for methanol in different environment. Upper (——) is NVE, upper (- - -) is NVT, upper (····) is NVT, upper (- · - · -) is NVT

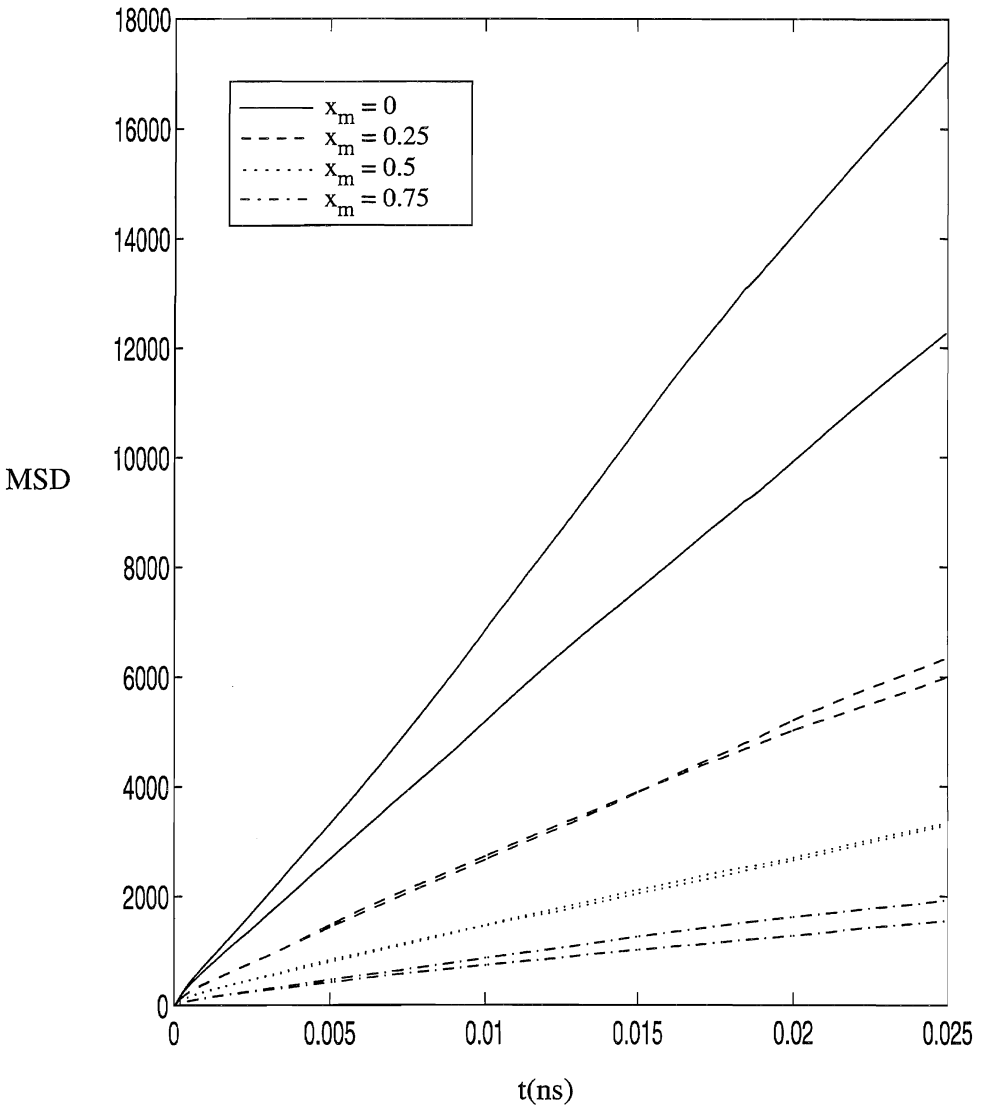


Figure C.33 Mean square displacement for water in different environment. Upper (—) is NVE, upper (- - -) is NVT, upper (- · - · -) is NVT

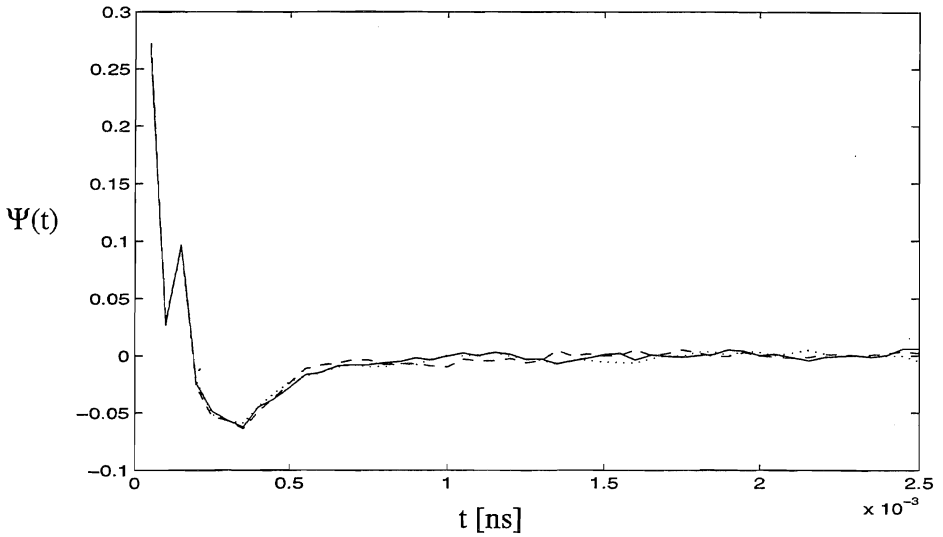


Figure C.34 x, y, and z-components of normalized velocity auto correlation function Ψ(t) for pure water from NVT–simulation.

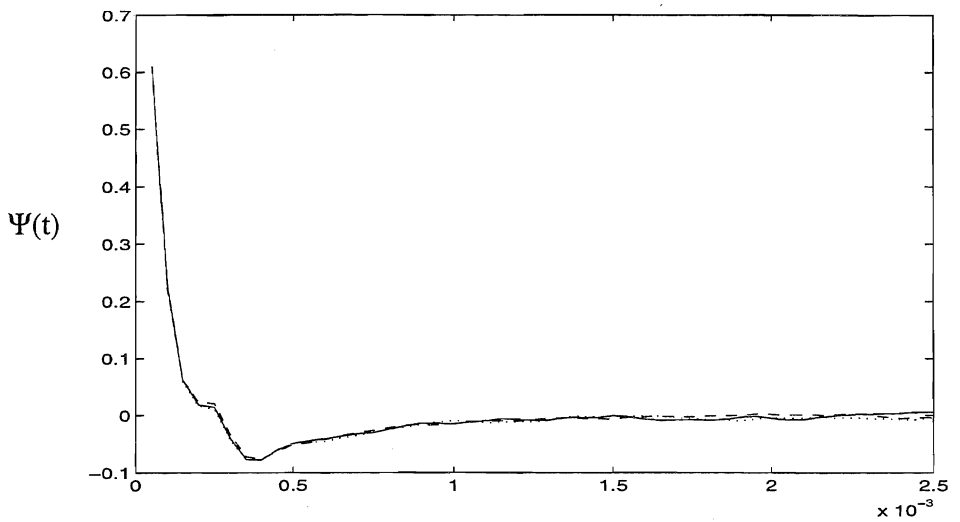


Figure C.35 x, y and z-components of normalized velocity auto correlation function Ψ(t) for pure methanol from NVT–simulation.

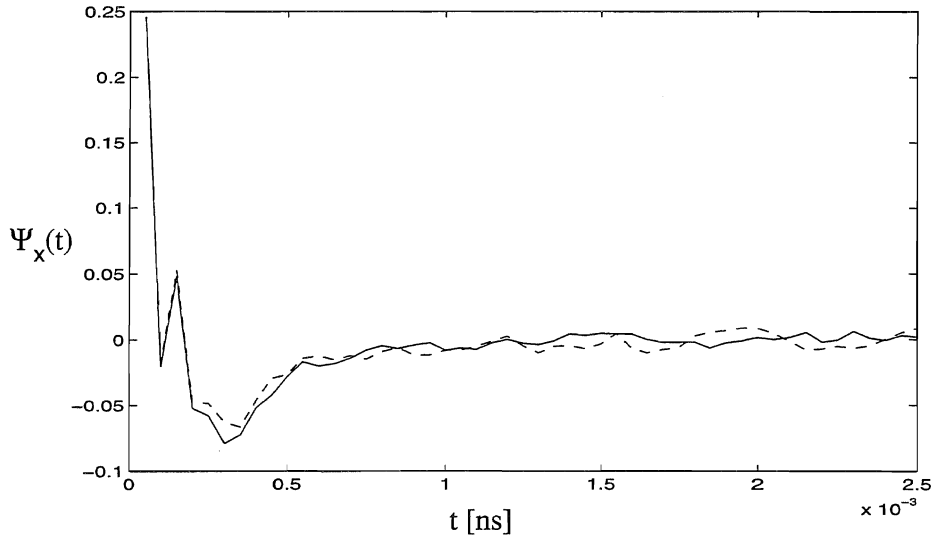


Figure C.36 Normalized velocity auto correlation function $\Psi_x(t)$ for water in an equimolar mixture with methanol. NVE: ———, and NVT: - - - -

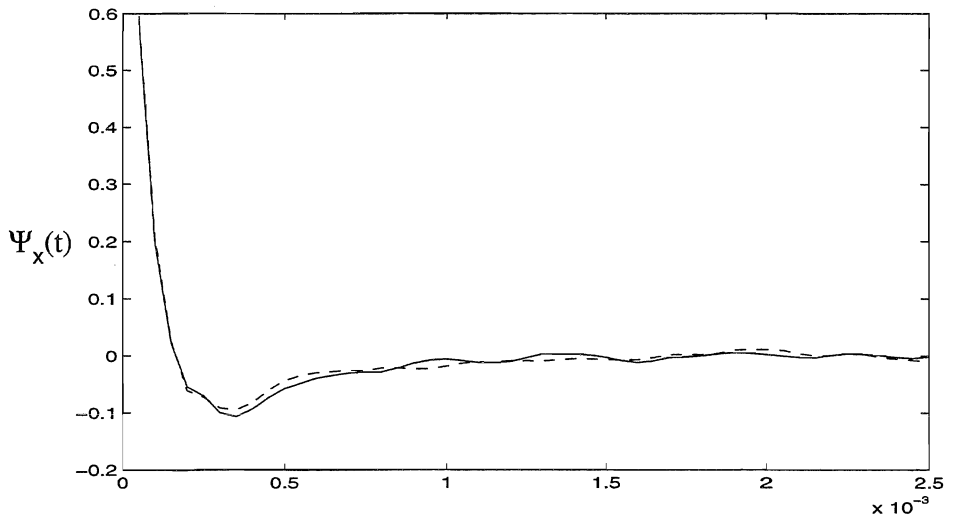


Figure C.37 Normalized velocity auto correlation function $\Psi_x(t)$ for methanol in an equimolar mixture with water. NVE: ———, and NVT: - - - -

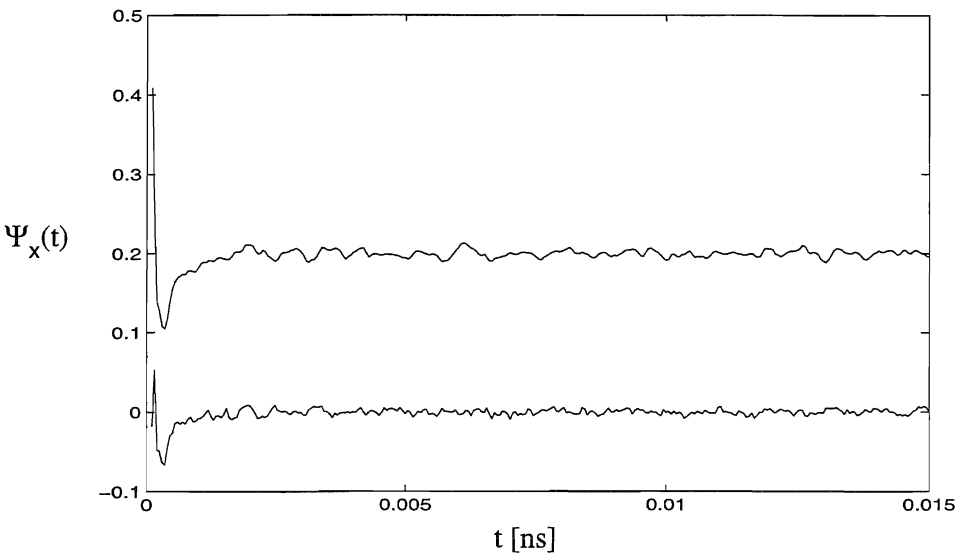


Figure C.38 Normalized centre of mass velocity auto correlation function $\Psi_x(t)$ for water (lower) and methanol (displaced by 0.2 units) in an equimolar mixture of water and methanol.

Appendix D
Further results water-ethanol

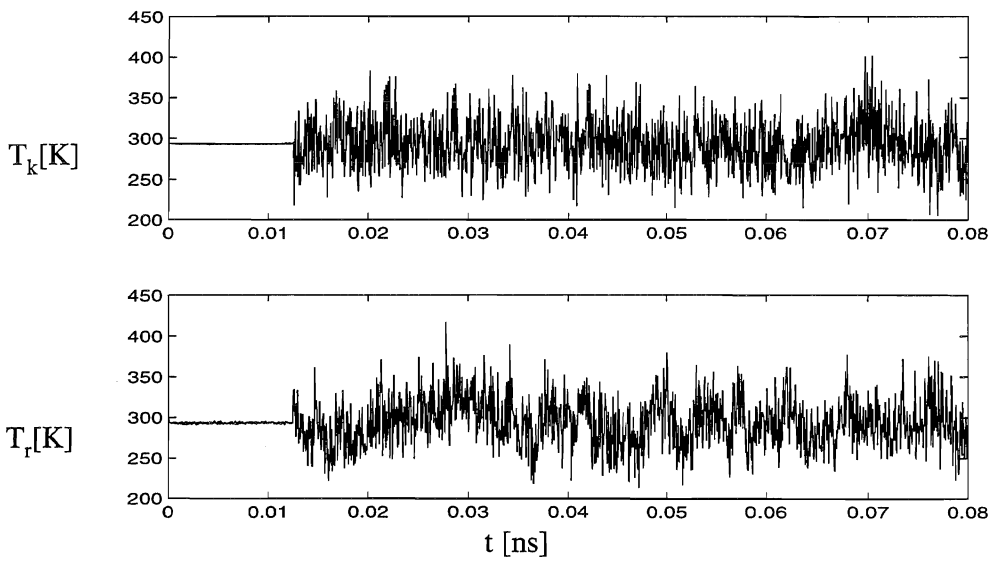


Figure D.1 Instantaneous translational (T_k) and rotational (T_r) temperatures for water from NVT simulation of 0.75 mole fraction ethanol in water.

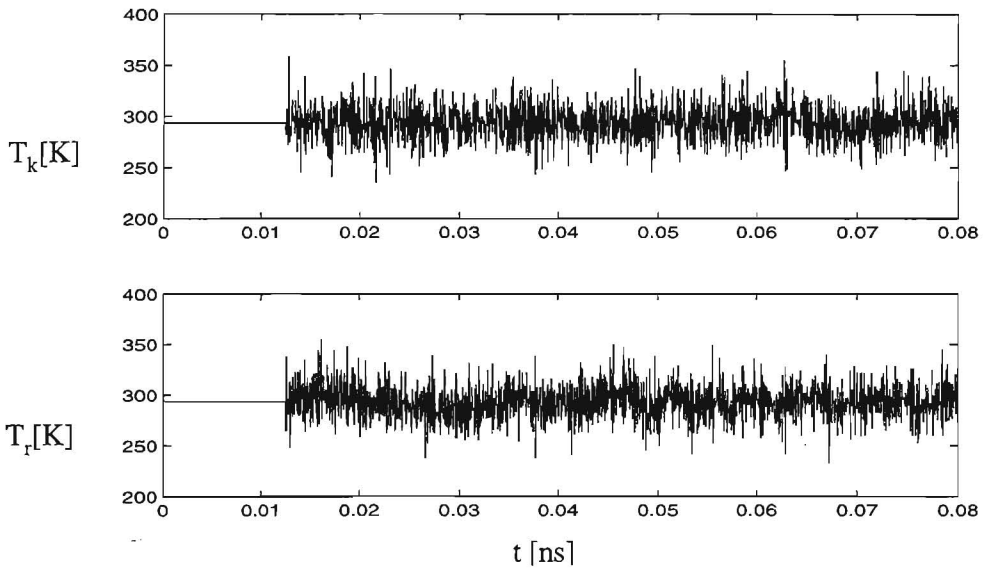


Figure D.2 Instantaneous translational (T_k) and rotational (T_r) temperatures for ethanol from NVT simulation of 0.75 mole fraction ethanol in water.

Table D.1

Rotational and translational temperature for each species from NVT simulation. Averages and standard deviations in round brackets calculated from tabulated values with one decimal taken at intervals of 100 timesteps. There is thus a slight disagreement between T_{mix} above and mixture temperature given in table. Standard deviation for distribution of single points

x_e	water		ethanol		$<T_{mix}>$
	$<T_k>$	$<T_r>$	$<T_k>$	$<T_r>$	
0.0	293.3 (13.9)	292.6 (13.8)			292.9 (9.9)
0.25	293.1 (15.7)	292.2 (15.8)	293.5 (27.7)	294.4 (27.2)	292.9 (9.8)
0.50	293.8 (19.7)	290.3 (19.4)	292.8 (19.3)	295.9 (19.3)	293.2 (9.9)
0.75 ***NVE***	292.6 (27.9) *295.2 (28.4)*	293.9 (27.6) *293.6 (27.5)*	293.3 (16.3) *296.3 (15.5)*	292.6 (16.1) *297.4 (15.0)*	293.0 (9.7) *296.1 (8.3)*
1.0			293.3 (13.9)	293.1 (13.9)	293.1 (9.9)

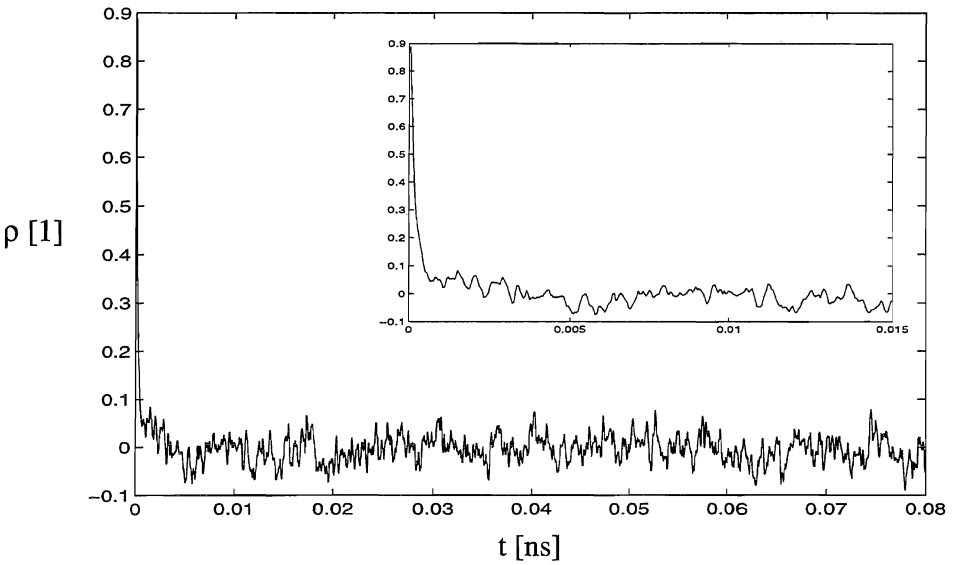


Figure D.3 Translational order parameter ρ for a NVT simulation of ethanol and water, x_e=0.75. Every 100th steps are plotted. Inset shows the first 30000 steps.

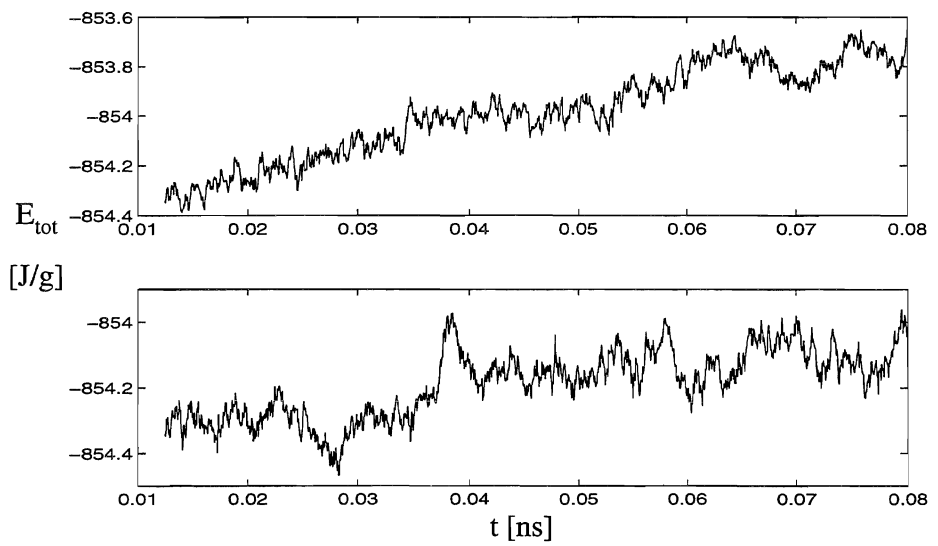


Figure D.4 Progress of total energy in NVT (top) and NVE simulation of 0.75 mixture of ethanol in water. The last 135000 steps is displayed.

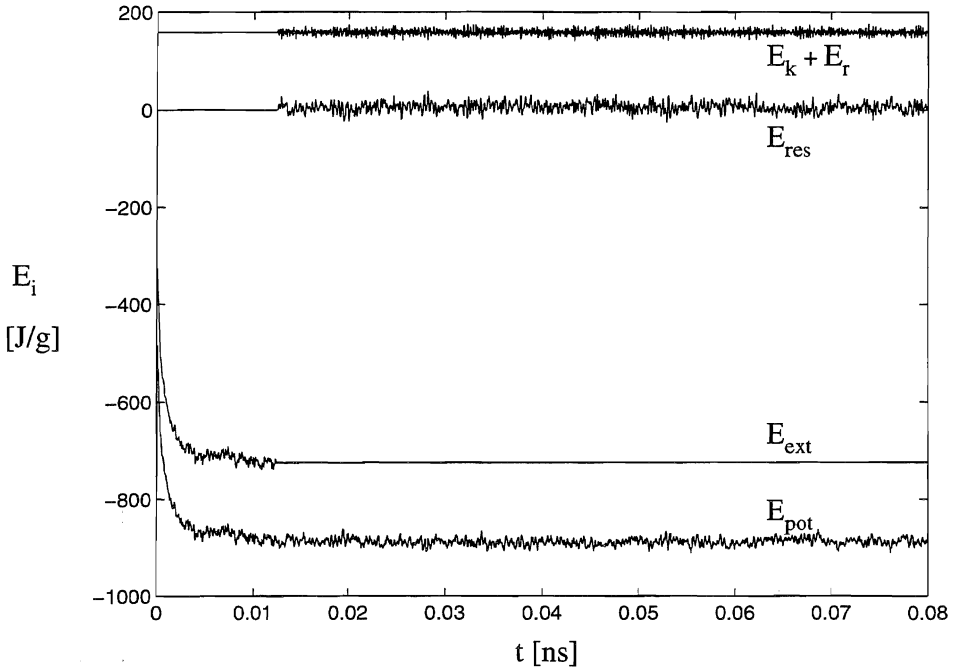


Figure D.5 Contributions to energy, E_i, for pure ethanol NVT simulation. $E_k + E_r$ is total kinetic energyfor the molecules, E_{res} is sum of potential energy of the two reservoirs, E_{pot} is total configurational energy, and E_{ext} is total energy of extended system. Reservoir kinetic energy not shown, but included in E_{ext}.

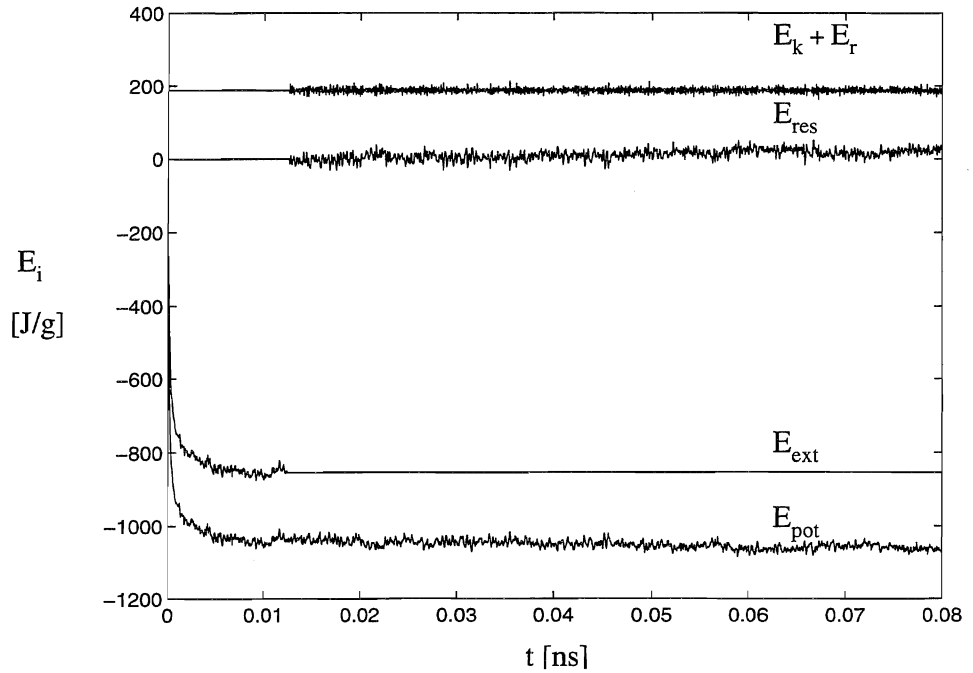


Figure D.6 Contributions to energy for a NVT simulation of 0.75 mole fraction ethanol mixture with water.

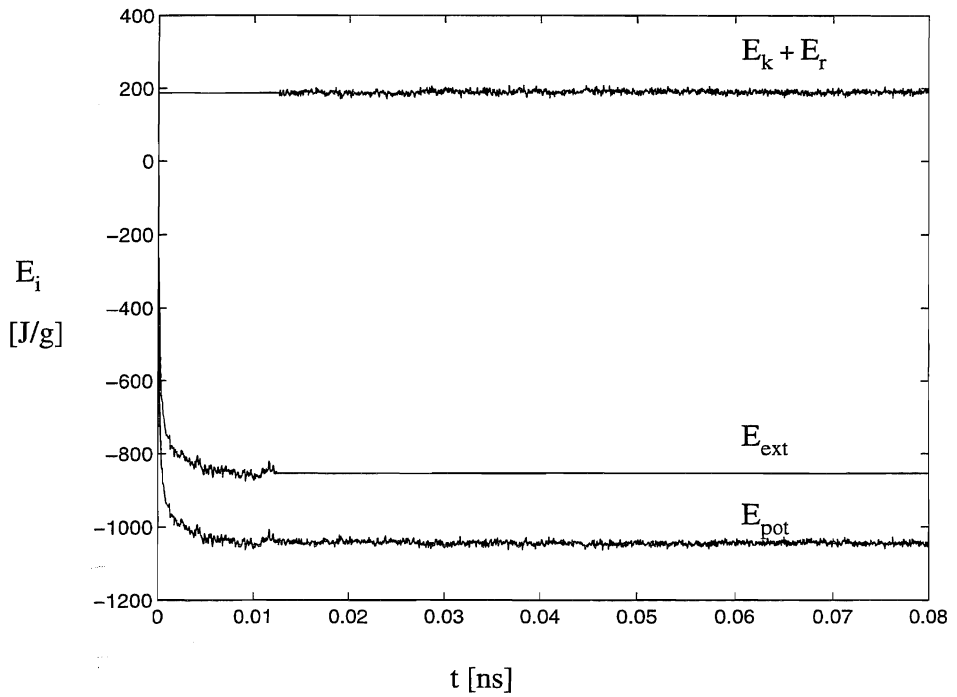


Figure D.7 Contributions to energy for a NVE simulation of 0.75 mole fraction ethanol mixture with water. See text of Figure D.5 above.

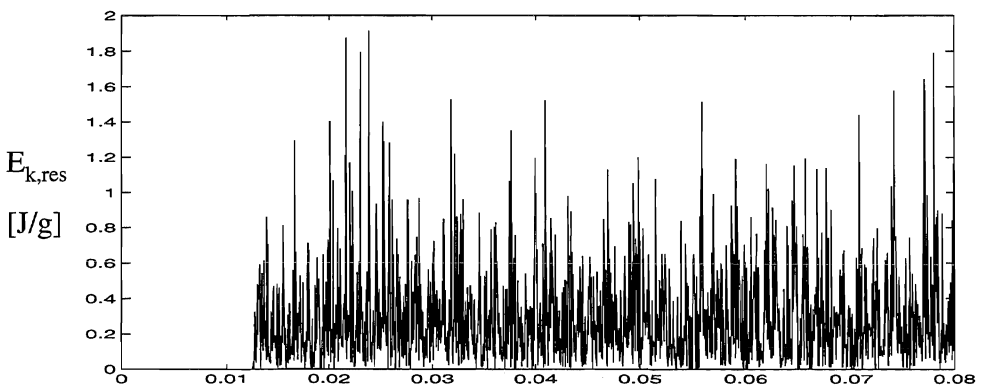


Figure D.8 Sum of resevoir kinetic energy $E_{k,res}$ for quimolar mixture of water and ethanol.

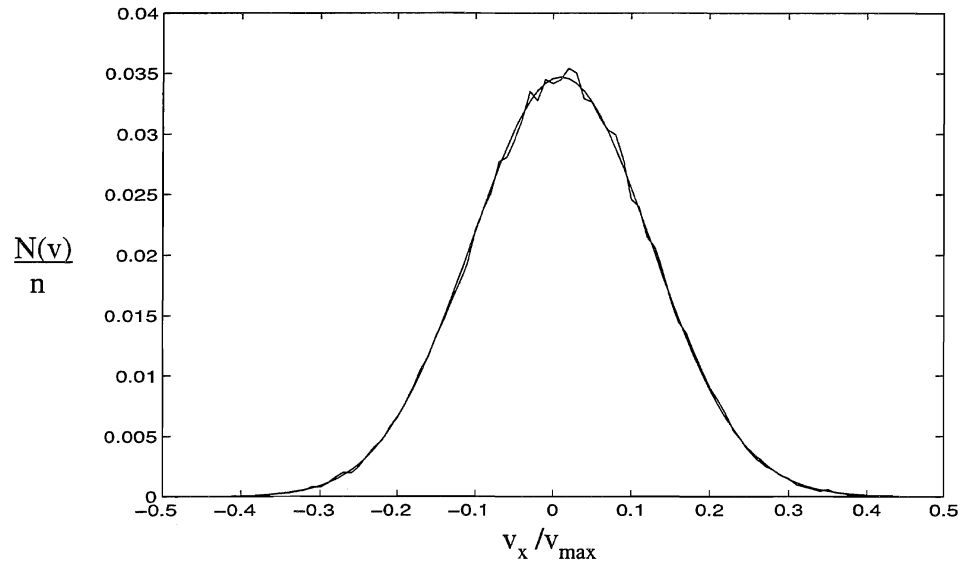


Figure D.9 Velocity distribution N(v)/n of 64 ethanol molecules in a NVT simulation of a 25% mixture with water. n is number of methanol molecules. Smooth line is Maxwell distribution at T=293K. v_{max}=2000m/s. Structural sampling each 150th step. Uncertainty in x-direction is 0.01.

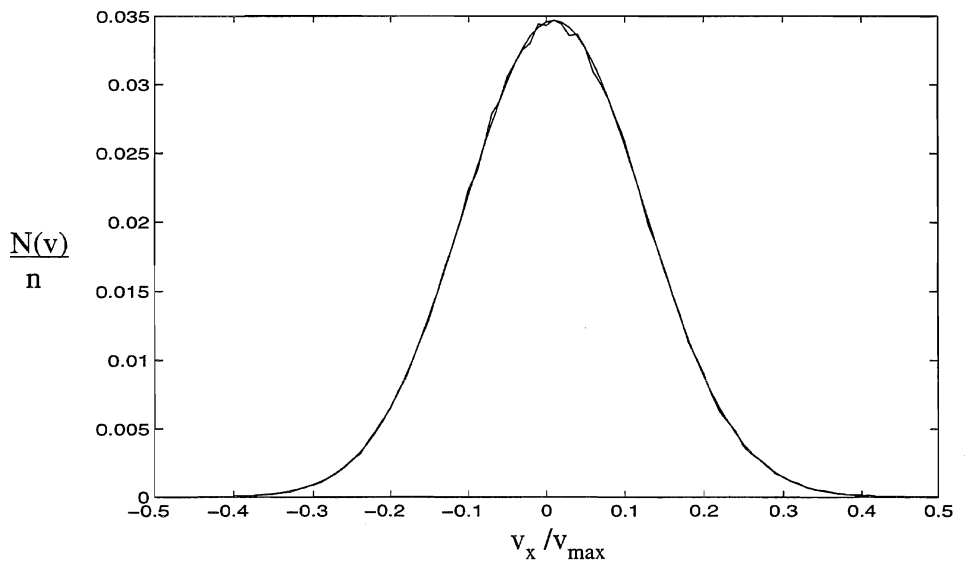


Figure D.10 Velocity distribution N(v) of 256 ethanol molecules in a NVT simulation of pure ethanol. n is number of methanol molecules. See text to Figure D.9.

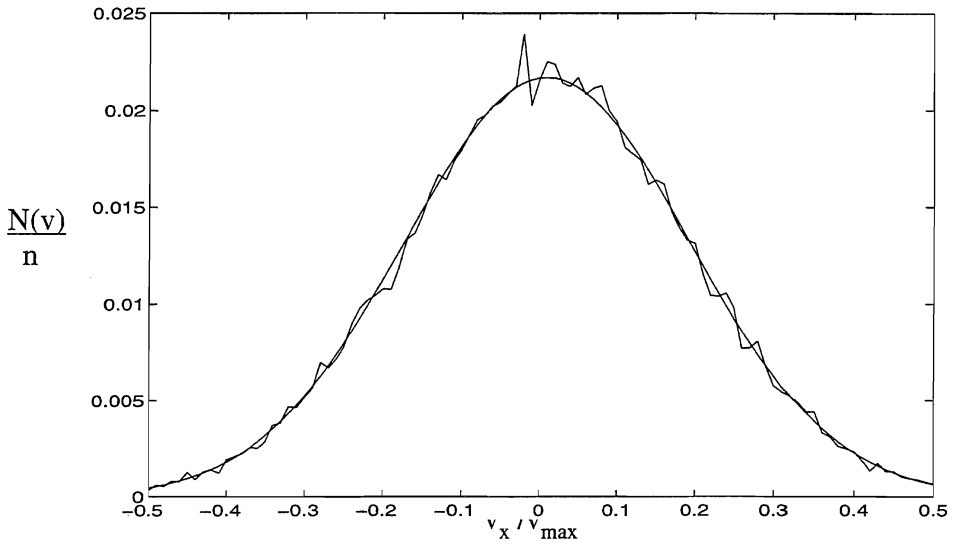


Figure D.11 Velocity distribution N(v) of 64 water molecules in a NVT simulation of a 75% mixture of ethanol in water. n is number of water molecules. See text to Fig D.9

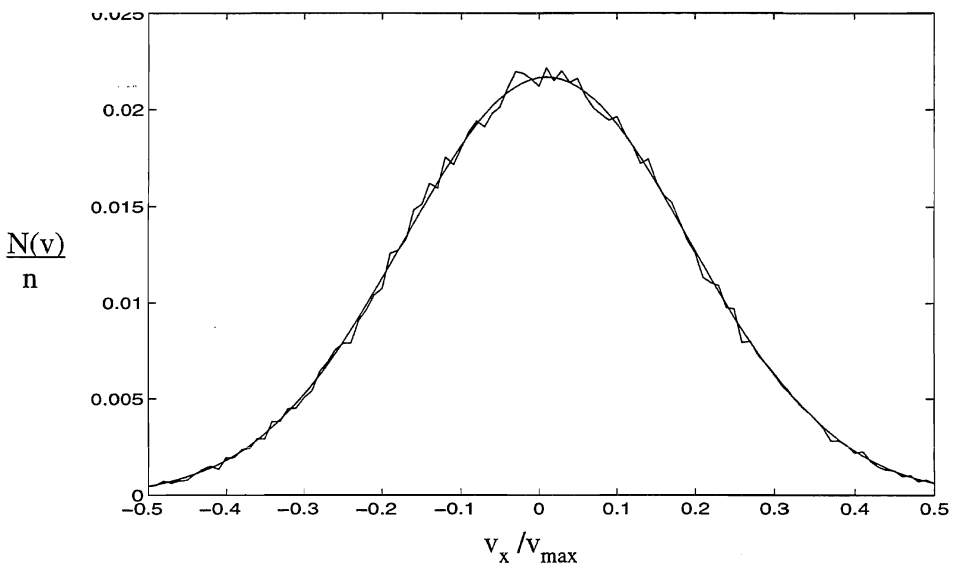


Figure D.12 Velocity distribution N(v) of 128 water molecules in a NVT simulation of a equimolar mixture of ethanol in water. n is number of water molecules. See text to Fig D.9.

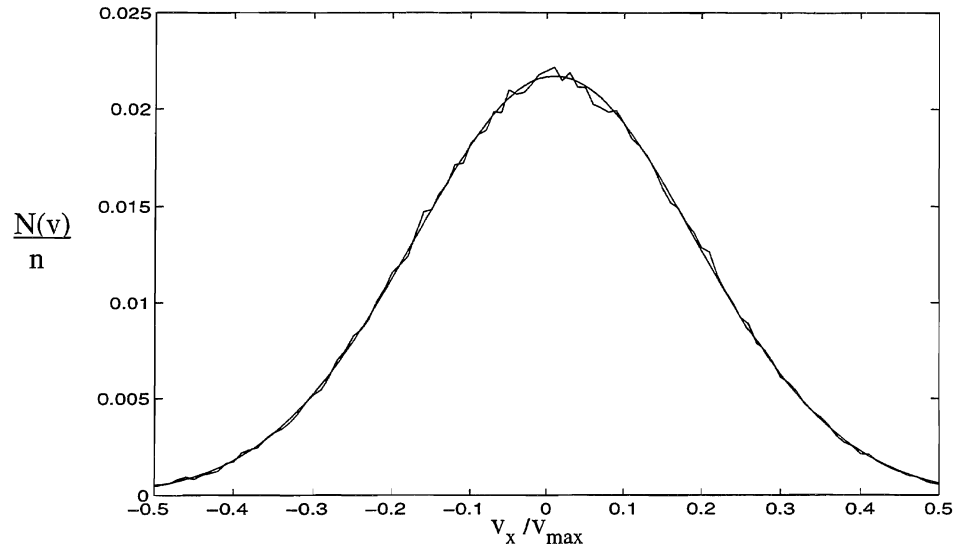


Figure D.13 Velocity distribution N(v) of 192 water molecules in a NVT simulation of a 25% mixture of ethanol in water. n is number of water molecules. See text to Fig D.9.

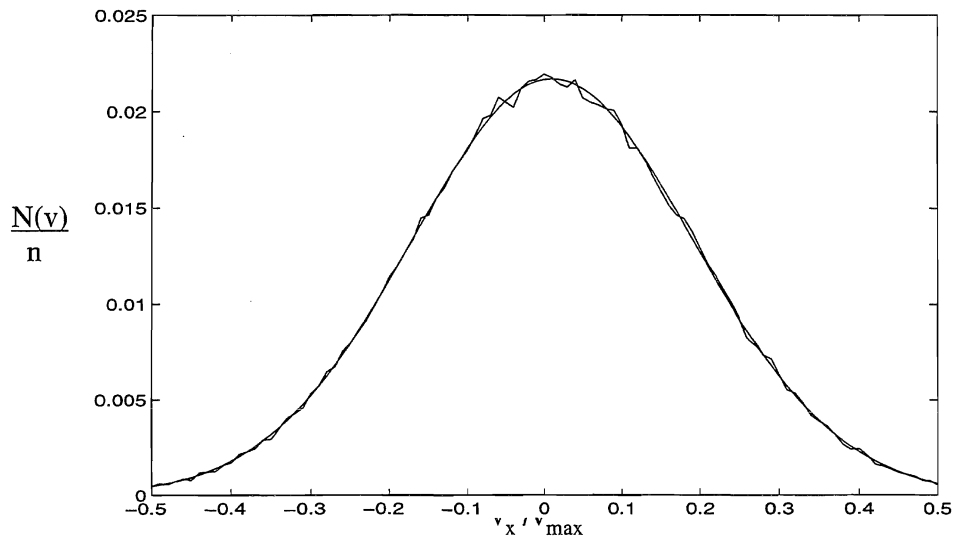


Figure D.14 Velocity distribution N(v) of 256 water molecules in a NVT simulation of pure water at 293K. n is number of water molecules. See text to Figure D.9

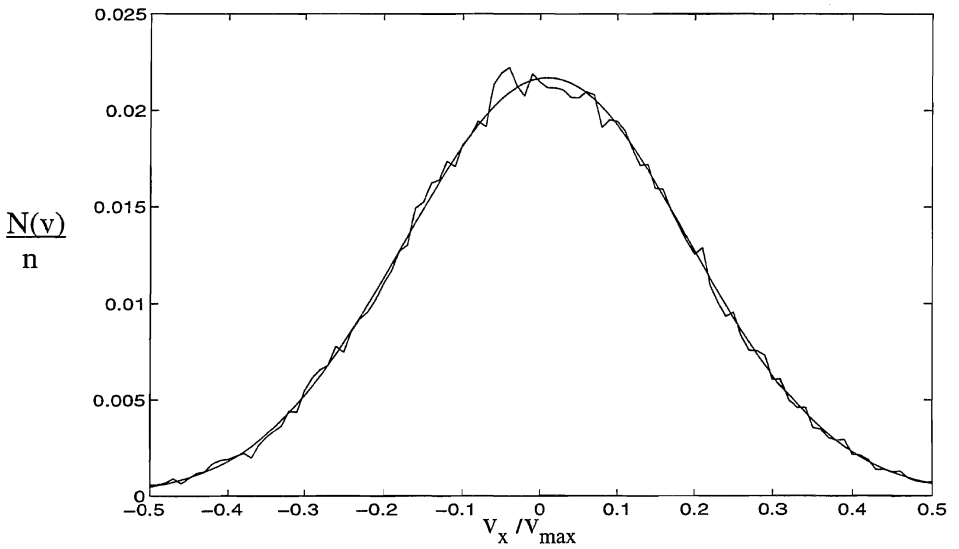


Figure D.15 Velocity distribution N(v) of 64 water molecules in a NVE simulation of a 75% mixture of ethanol in water. n is number of water molecules. See text to Figure D.9

Table D.2
Linear momentum per mass, p_ζ, and translational energy per mass, K_ζ, for each direction for NVT-simulations of water and ethanol.

x_e	x-direction		y-direction		z-direction	
	$p_x[10^{-13}m/s]$	$2K_x$ [J/g]	$p_y[10^{-13}m/s]$	$2K_y$ [J/g]	$p_z[10^{-13}m/s]$	$2K_z$ [J/g]
0	5.032 (4.633)	135.3 (11.6)	3.853 (2.833)	135.1 (11.5)	-8.355 (6.810)	134.6 (11.9)
0.25	-7.500 (3.908)	97.1 (8.2)	-1.540 (2.205)	96.8 (8.2)	-4.589 (3.203)	97.2 (8.1)
0.50	1.628 (2.416)	75.7 (6.4)	0.871 (2.616)	75.9 (6.4)	1.449 (2.355)	75.6 (6.5)
0.75	-2.149 (1.943)	62.1 (5.4)	0.609 (1.258)	62.3 (5.4)	-3.312 (3.905)	62.0 (5.2)
NVE	1.076 (1.329)	62.9 (5.5)	0.082 (0.516)	63.4 (5.4)	0.603 (1.332)	62.6 (5.5)
1.0	2.066 (1.595)	53.0 (4.6)	0.774 (1.695)	52.4 (4.6)	-2.689 (1.871)	52.8 (4.7)

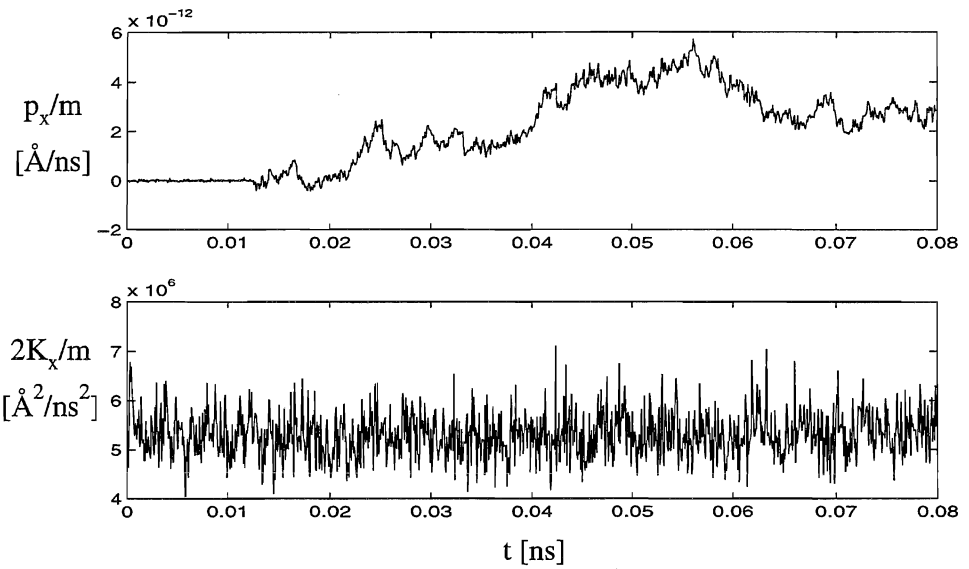


Figure D.16 X-component of linear momentum (top) and translational energy from NVT simulation of ethanol. *m* is system mass.

Table D.3
Mean and standard deviation of angular momentum per mass, I_ζ and rotational energy per mass, R_ζ in NVT-simulations. x_e is mole fraction ethanol.

x_e	x-direction		y-direction		z-direction	
	$I_x[10^{-11}m^2/s]$	$2R_x[J/g]$	$I_y[10^{-11}m^2/s]$	$2R_y[J/g]$	$I_z[10^{-11}m^2/s]$	$2R_z[J/g]$
0	0.906 (40.6)	134.1 (11.9)	-1.89 (71.7)	135.3 (11.2)	0.40 (57.5)	135.0 (11.8
0.25	0.337 (60.4)	97.3 (8.5)	0.43 (154.5)	97.1 (8.1)	2.21 (144.0)	97.2 (8.1)
0.50	-0.704 (60.4)	75.8 (6.5)	-5.76 (167.5)	76.2 (6.5)	3.31 (148.5)	75.8 (6.3)
0.75	-0.833 (63.5)	62.5 (5.6)	-3.64 (173.2)	62.5 (5.3)	-1.31 (150.5)	62.1 (5.5
NVE	*-0.976 (63.3)*	62.9 (5.3)	-3.34 (167.1)	63.3 (5.5)	-2.57 (146.3)	63.0 (5.5
1.0	-1.370 (69.7)	52.9 (4.9)	-5.24 (163.1)	52.8 (4.5)	1.15 (153.1)	52.8 (4.7

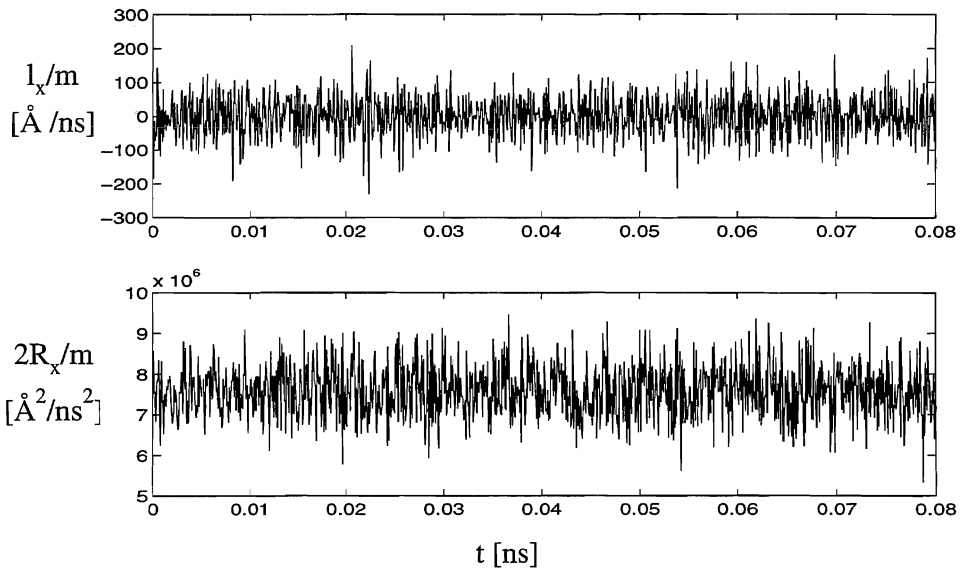


Figure D.17 X-component of angular momentum (top) and rotational energy for NVT simulation of an equimolar mixture of water and ethanol. *m* is system mass.

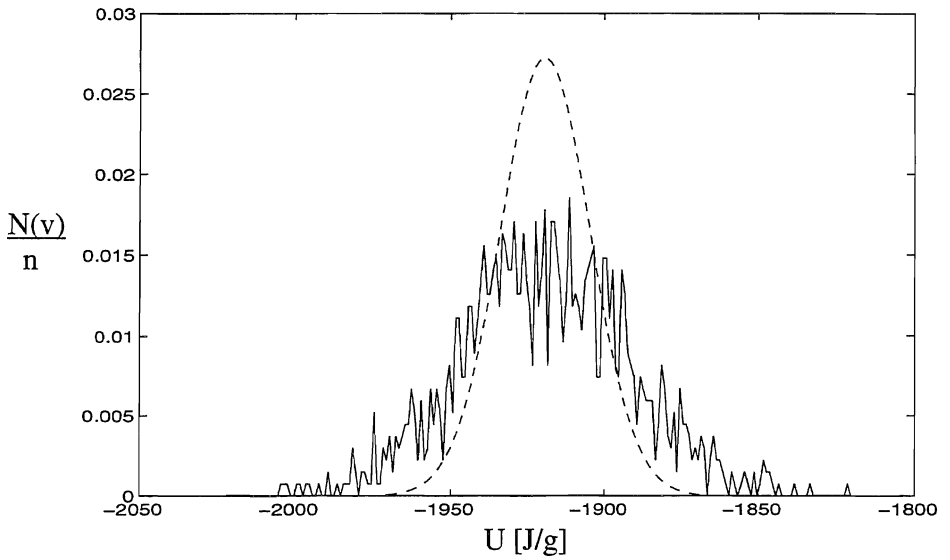


Figure D.18 Distribution of internal energy, U, of pure water. Dashed line is canonical distribution at 293.15K. Calculated distribution from output data with 5 significant digits each 100th step.

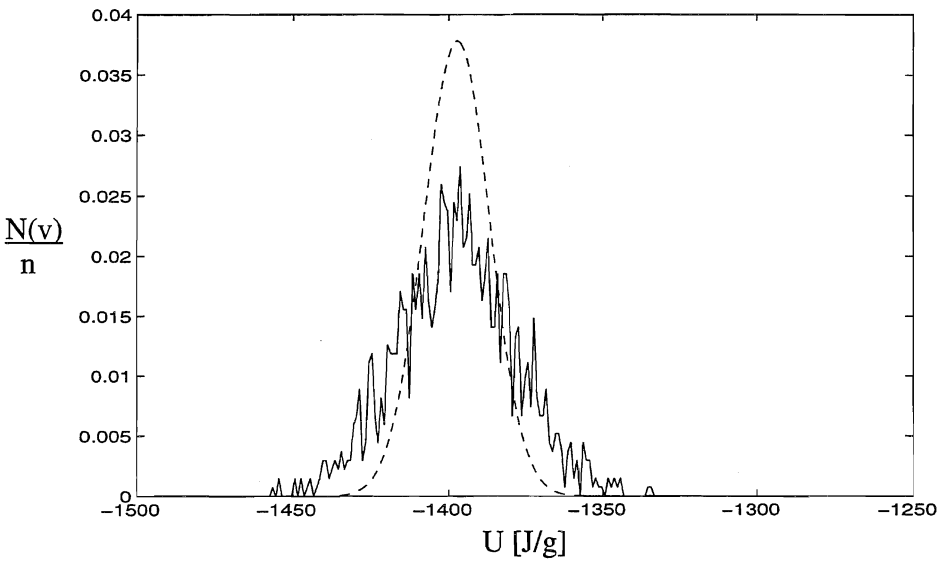


Figure D.19 Distribution of internal energy, U, in NVT-simulation. Molefraction ethanol is 0.25. Dashed line is canonical distribution at 293.15K. See text of fig D.18.

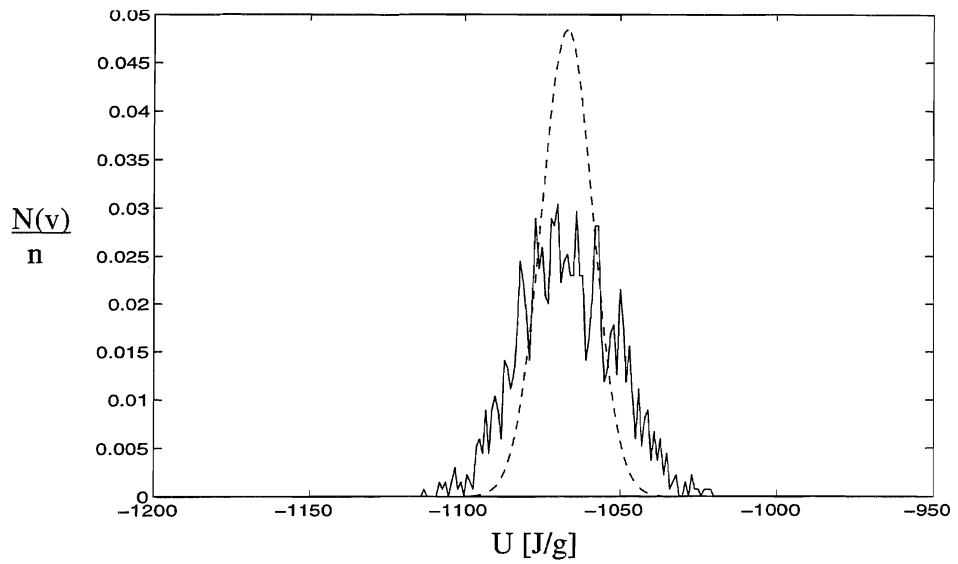


Figure D.20 Distribution of internal energy, U, in NVT-simulation. Molefraction ethanol is 0.5. Dashed line is canonical distribution at 293.15K. See text of Figure D.18.

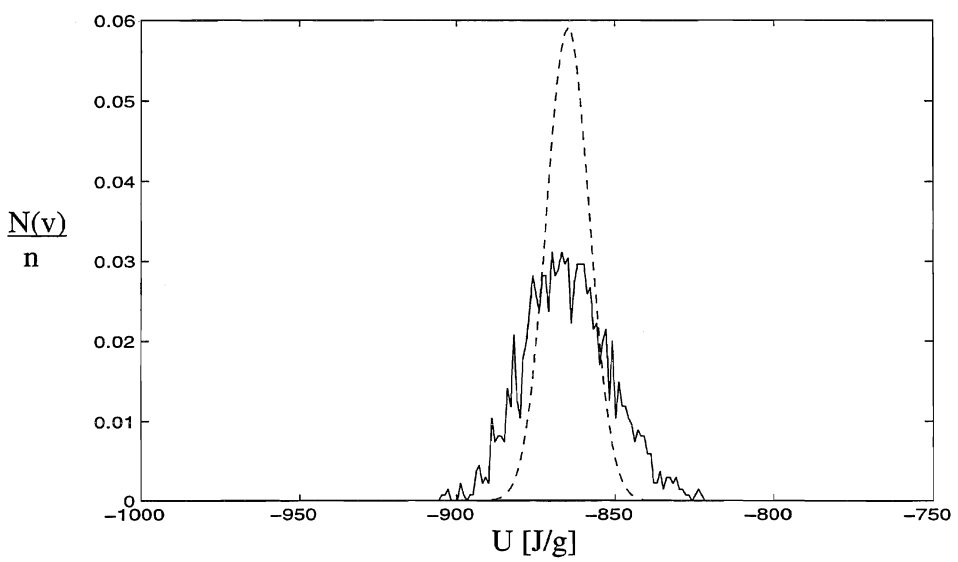


Figure D.21 Distribution of internal energy, U, in NVT-simulation. Molefraction ethanol is 0.75. Dashed line is canonical distribution at 293.15K.See text of Figure D.18

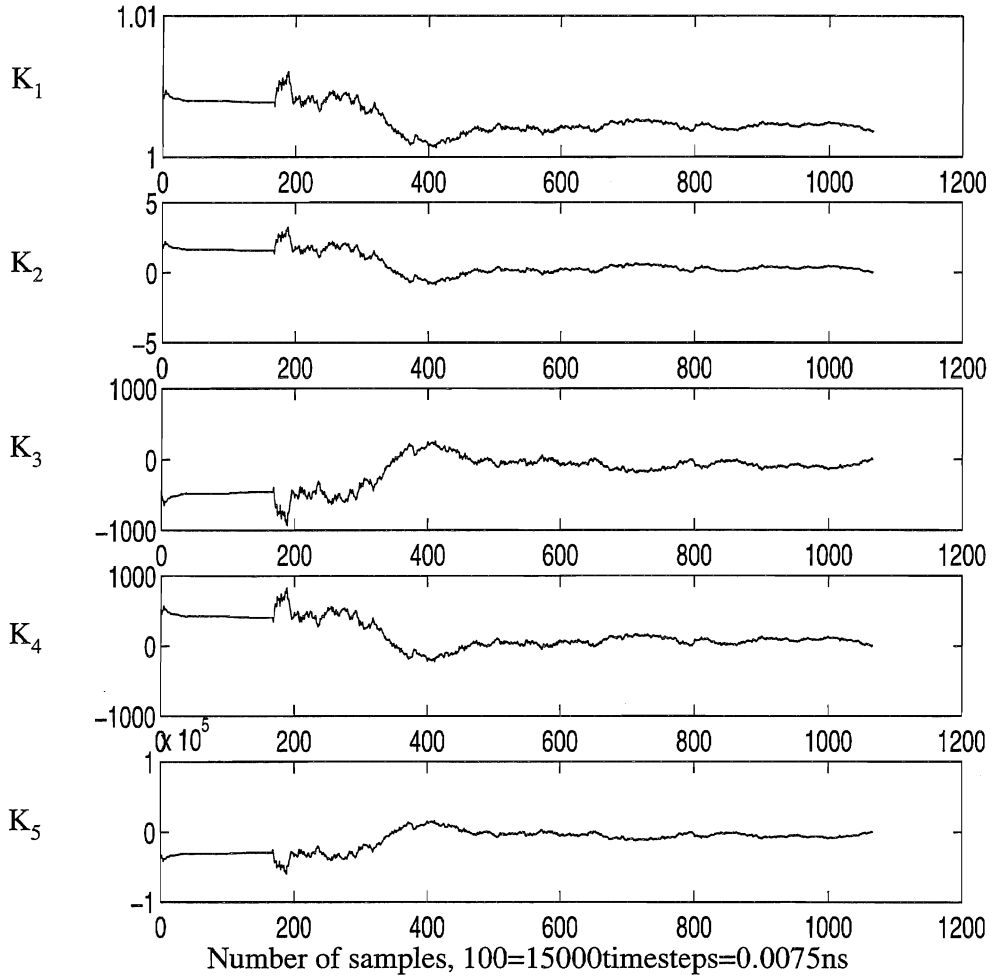


Figure D.22 Running averages of 1st (top) to 5th (bottom) moments of translational energy for a NVT simulation of pure water at 293K. Simulation conditions given in table B.3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see chapter 5.

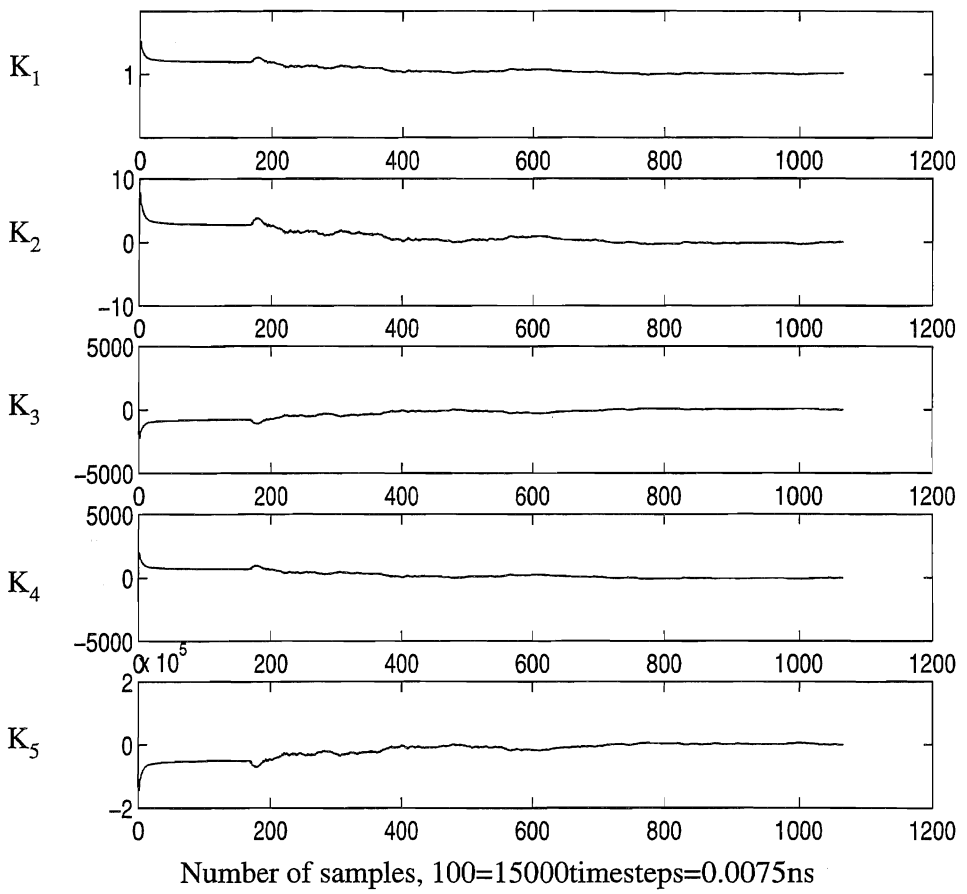


Figure D.23 Running averages of 1st (top) to 5th (bottom) moment of translational energy for a NVT simulation of 0.25 mole fraction ethanol in water. Simulation conditions as in table B.3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see chapter 5.

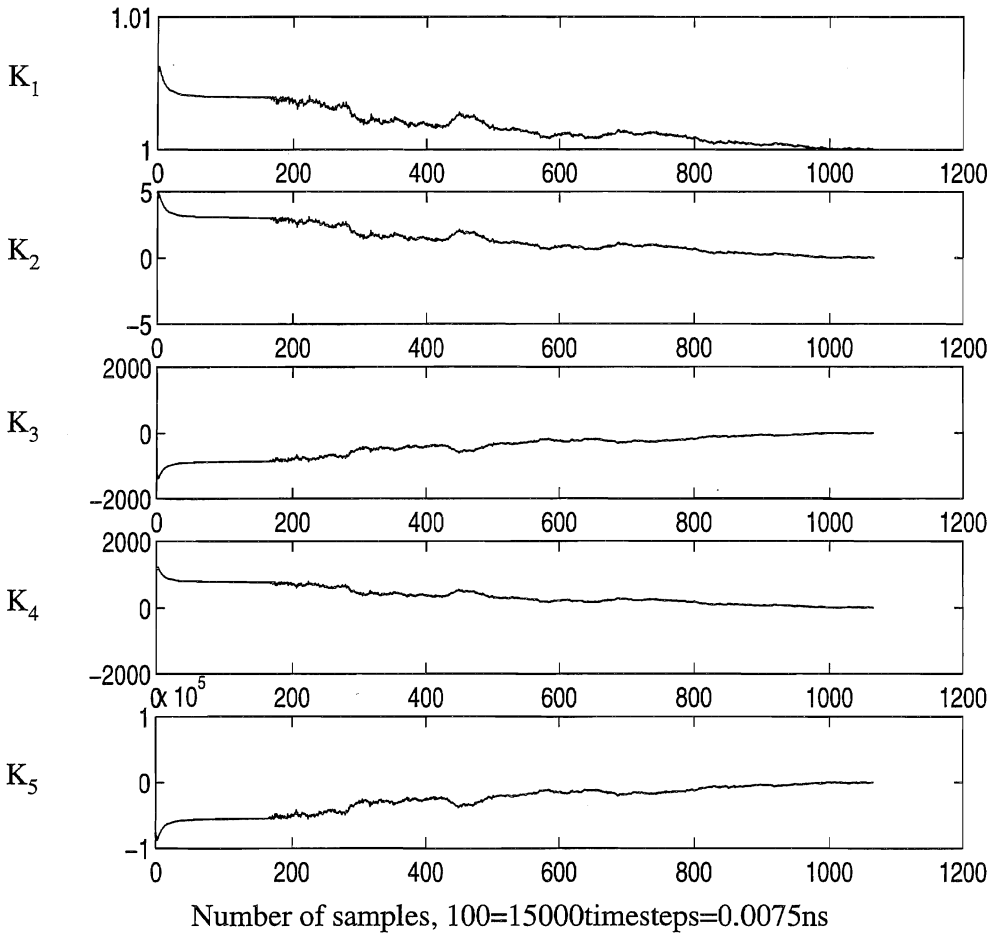


Figure D.24 Running averages of 1st (top) to 5th (bottom) moment of translational energy for a NVT simulation of 0.50 mole fraction ethanol in water. Simulation conditions as in table B.3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see chapter 5.

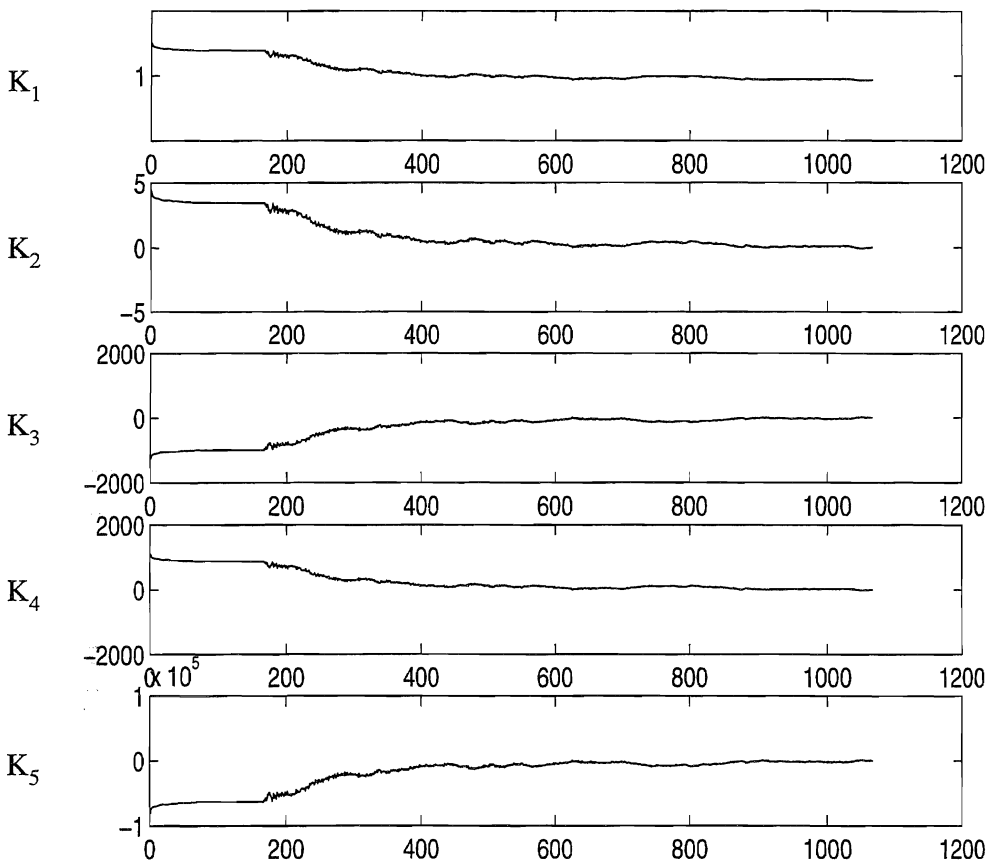


Figure D.25 Running averages of 1st (top) to 5th (bottom) moment of translational energy for a NVT simulation of 0.75 mole fraction ethanol in water. Simulation conditions as in table B.3, page 3.Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see chapter 5.

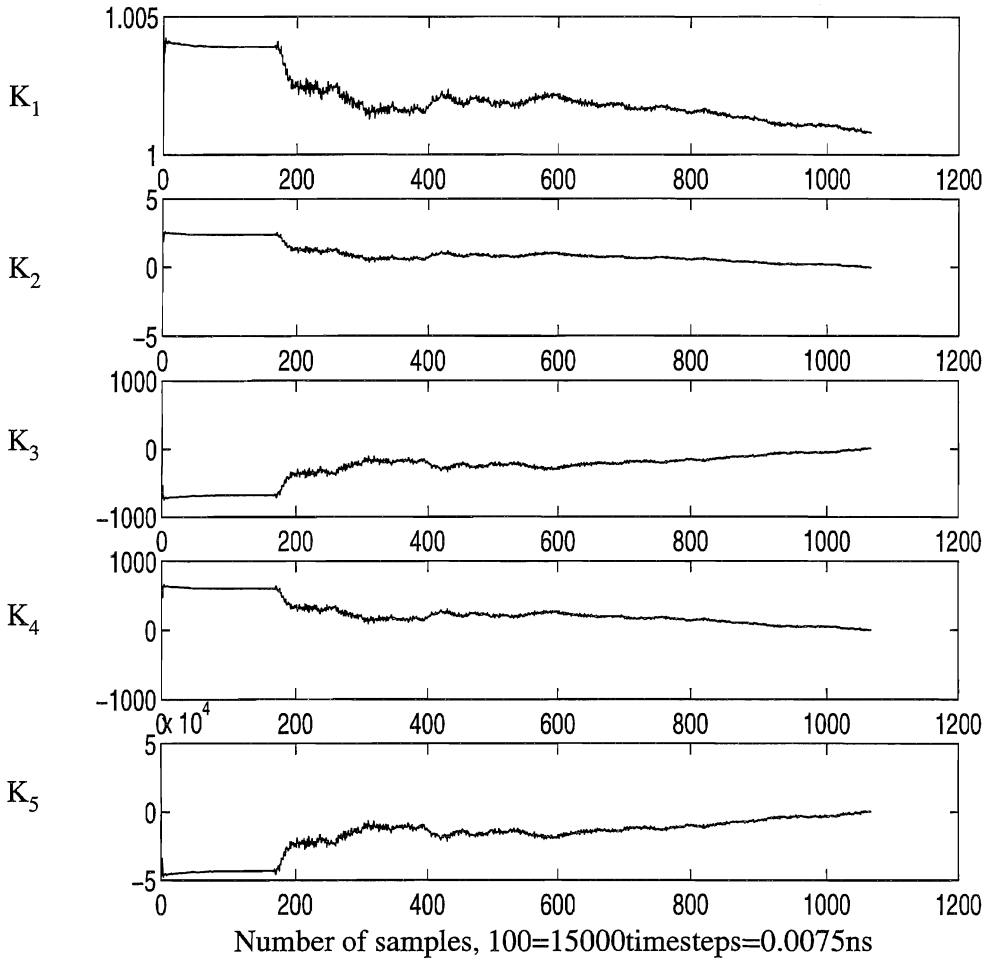


Figure D.26 Running averages of 1st (top) to 5th (bottom) moment of translational energy for a NVT simulation of pure ethanol at 293K. Simulation conditions as in table B.3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see chapter 5.

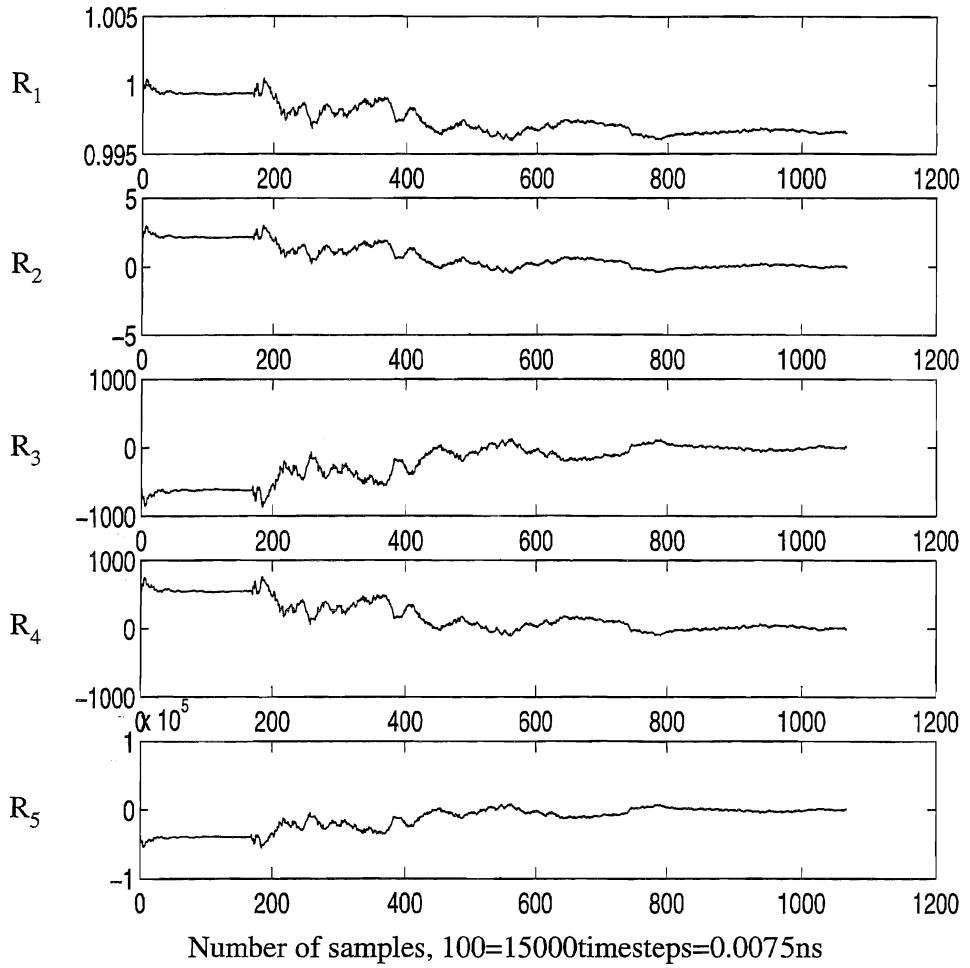


Figure D.27 Running averages of 1st (top) to 5th (bottom) moments of rotational energy for a NVT simulation of pure water. Simulation conditions as in table B.3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see chapter 5.

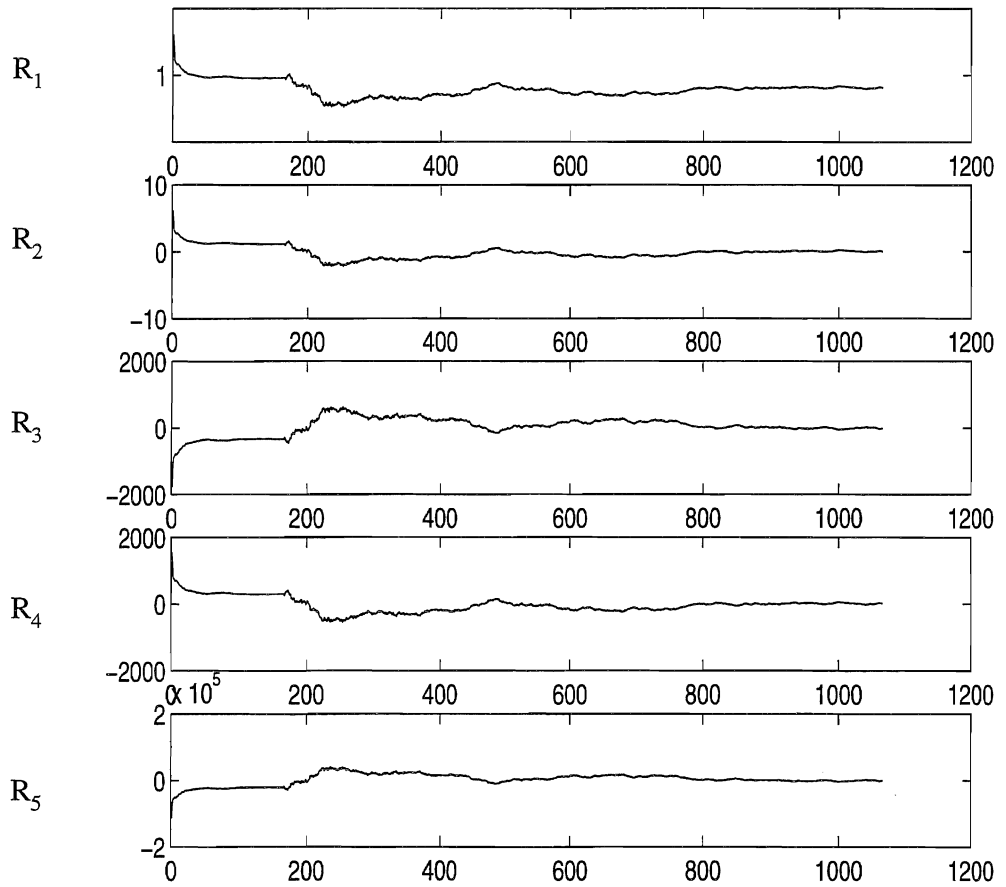


Figure D.28 Running averages of 1st (top) to 5th (bottom) moment of rotational energy for a NVT simulation of 0.25 mole fraction ethanol in water. Simulation conditions as in table B.3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see chapter 5.

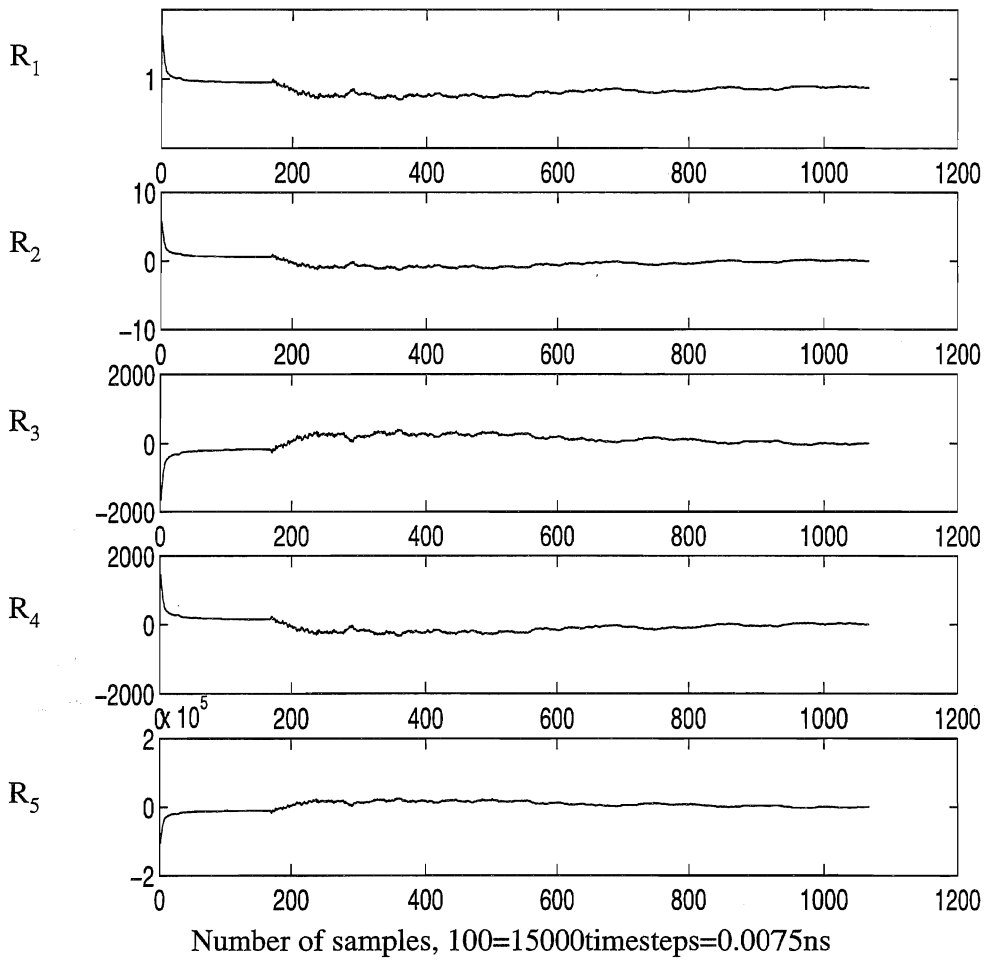


Figure D.29 Running averages of 1st (top) to 5th (bottom) moment of rotational energy for a NVT simulation of 0.50 mole fraction ethanol in water. Simulation conditions as in table B.3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see chapter 5.

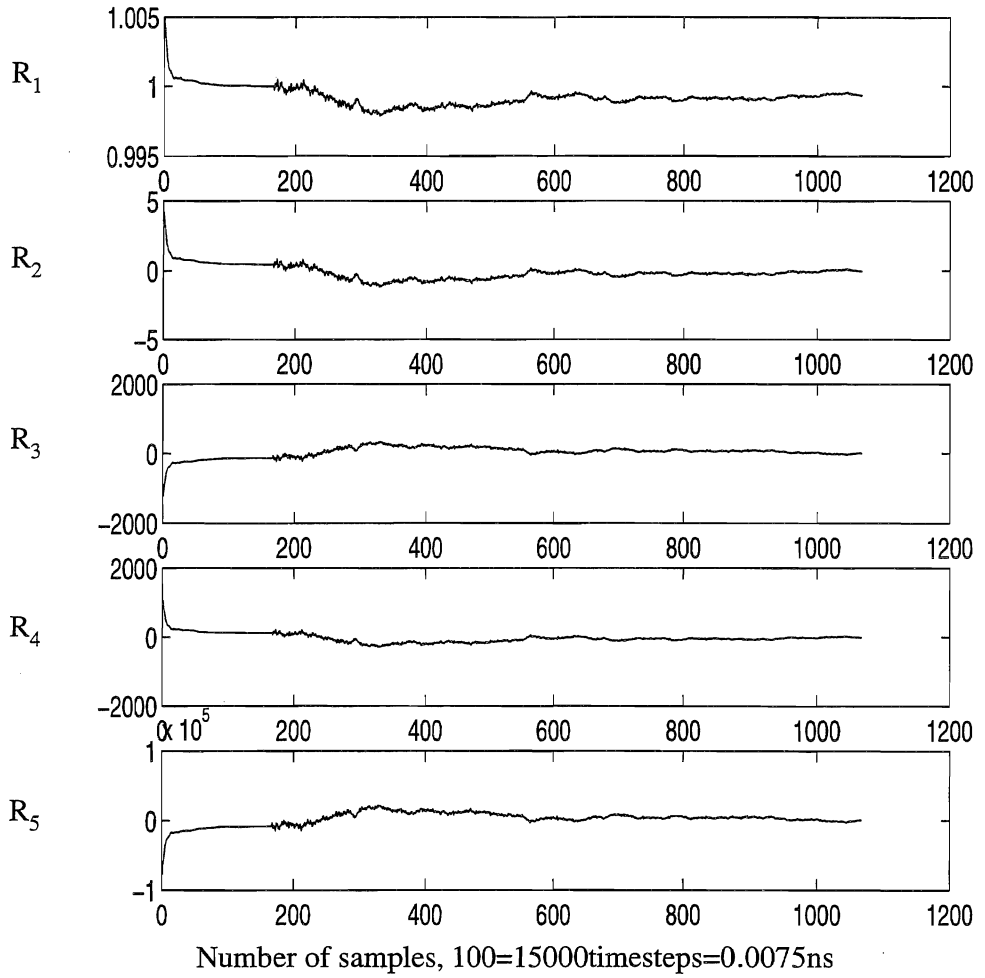


Figure D.30 Running averages of 1st (top) to 5th (bottom) moment of rotational energy for a NVT simulation of 0.75 mole fraction methanol in water. Simulation conditions as in table B.3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see chapter 5.

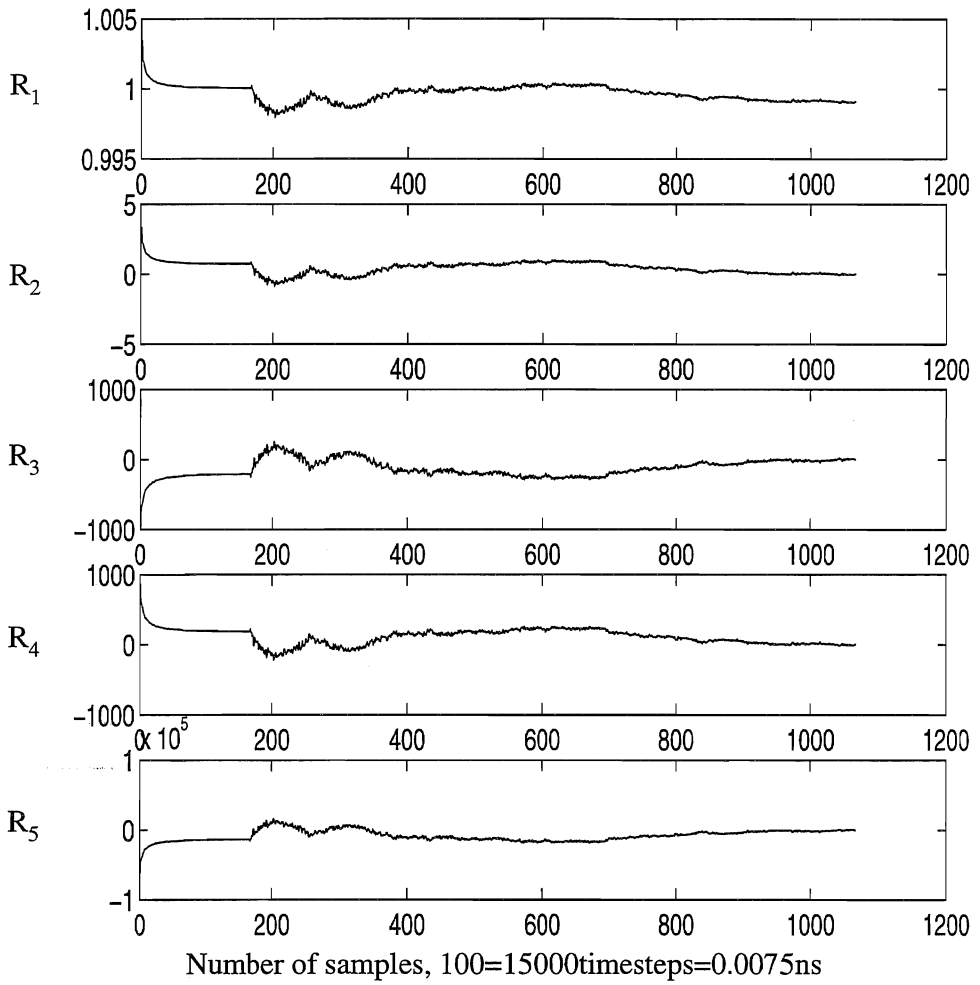


Figure D.31 Running averages of 1st (top) to 5th (bottom) moment of rotational energy for a NVT simulation of pure ethanol. Simulation conditions as in table B.3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see chapter 5.

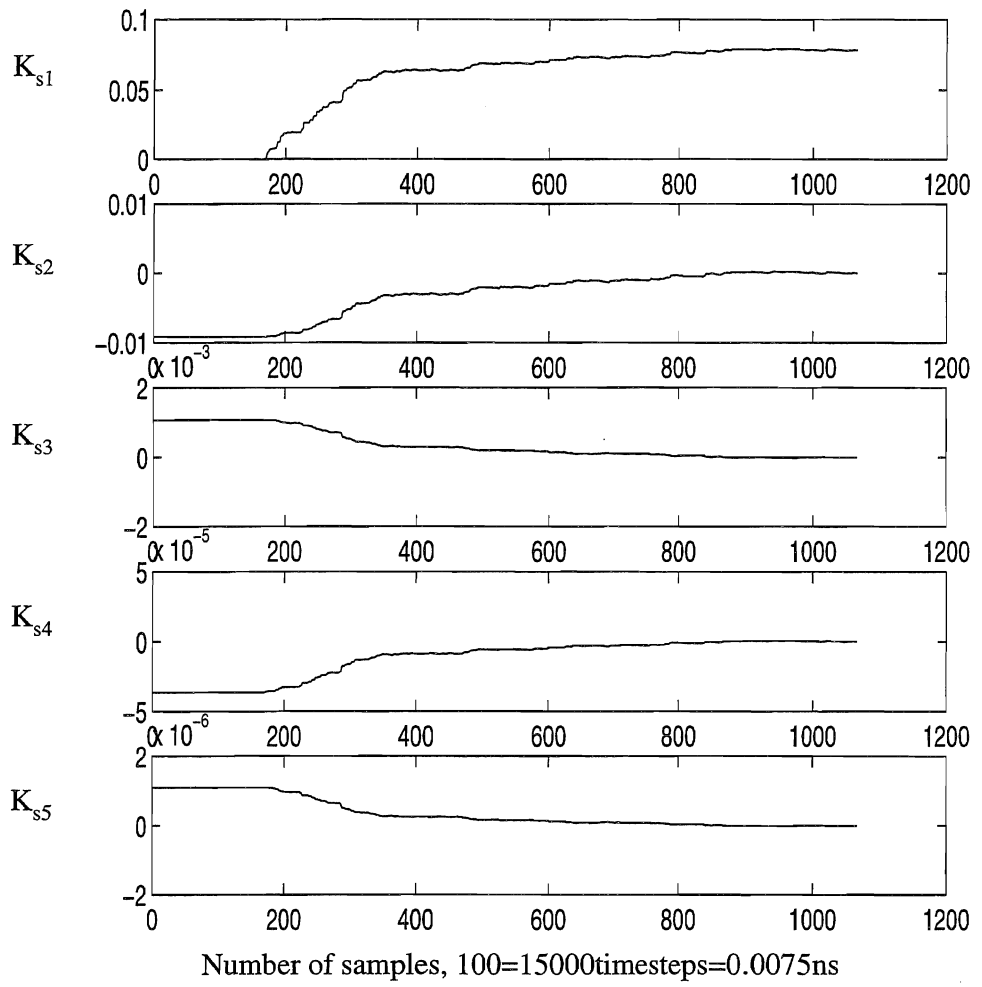


Figure D.32 Running averages of 1st (top) to 5th (bottom) moment of reservoir translational energy for a NVT simulation of 0.75 mole fraction ethanol in water. Simulation conditions as in table B.3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see chapter 5.

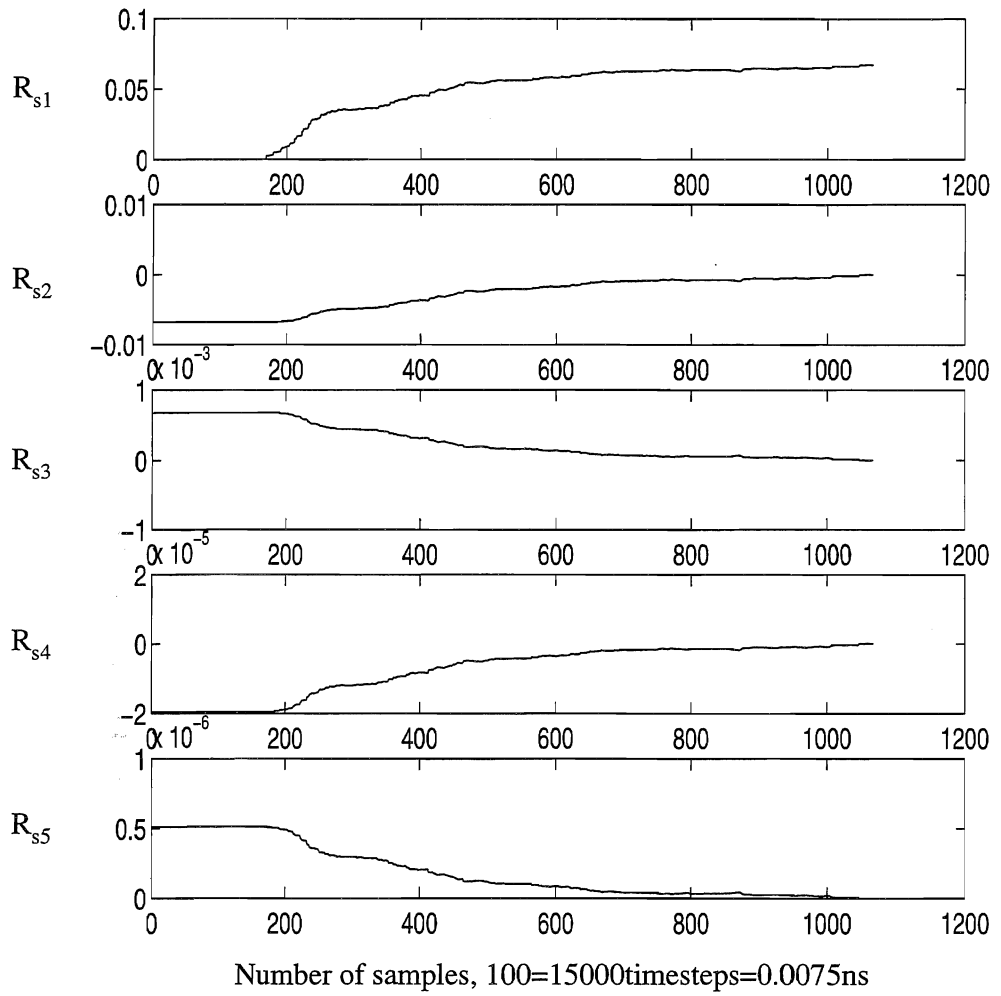


Figure D.33 Running averages of 1st (top) to 5th (bottom) moment of reservoir rotational energy for a NVT simulation of 0.75 mole fraction ethanol in water. Simulation conditions as in table B.3. Averages based on instantaneous values of energy each 150th step. Figure cover whole simulation, and values are divided by theoretical values, see chapter 5.

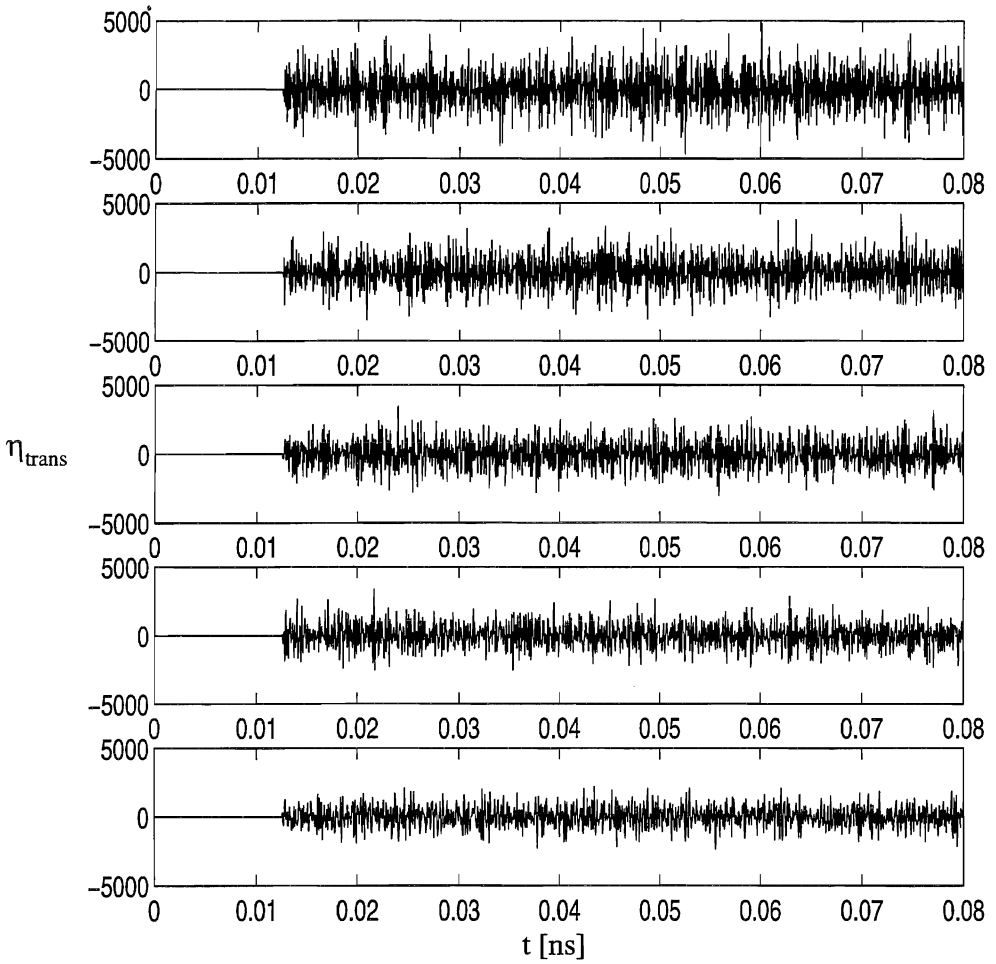


Figure D.34 Friction parameter η_{trans} simulation for mixtures of water and ethanol at 293K. Panels ranging from pure water (top) to pure ethanol (bottom).

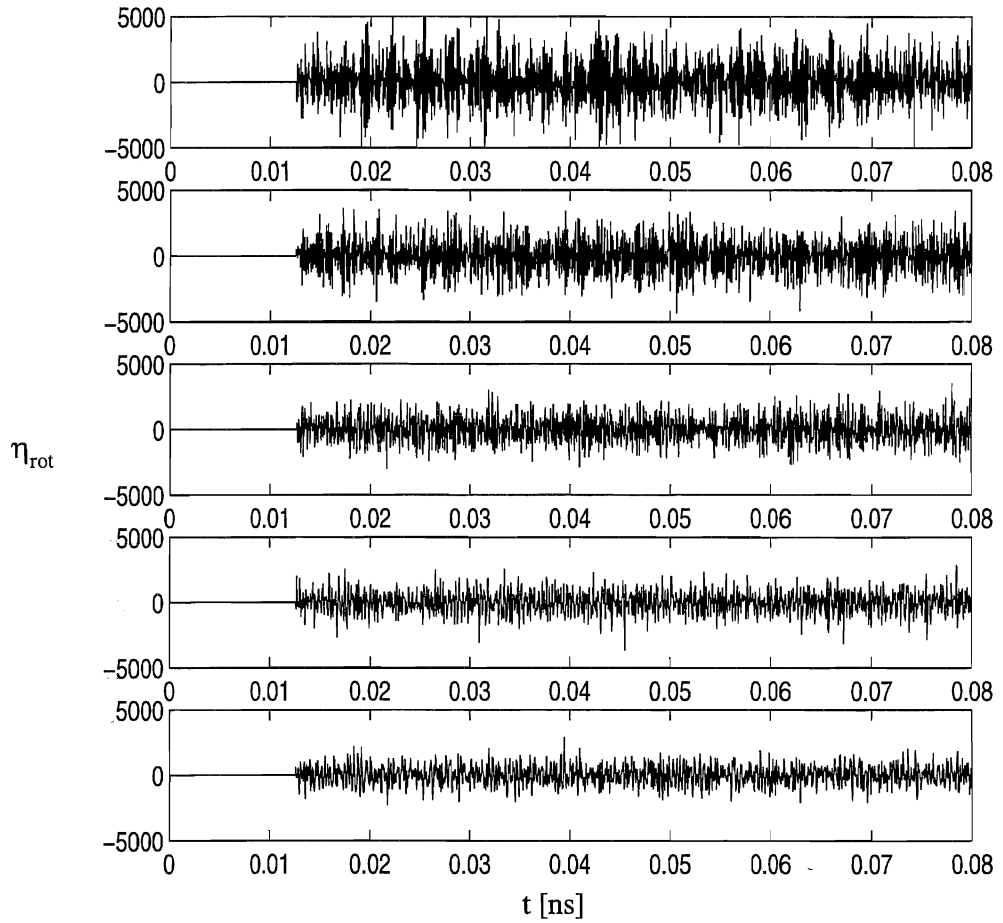


Figure D.35 Friction parameter η_{rot} for entire simulation of mixtures of water and ethanol. instantaneous values each 100th step shown. Panels ranging from pure water (top) to pure ethanol (bottom).

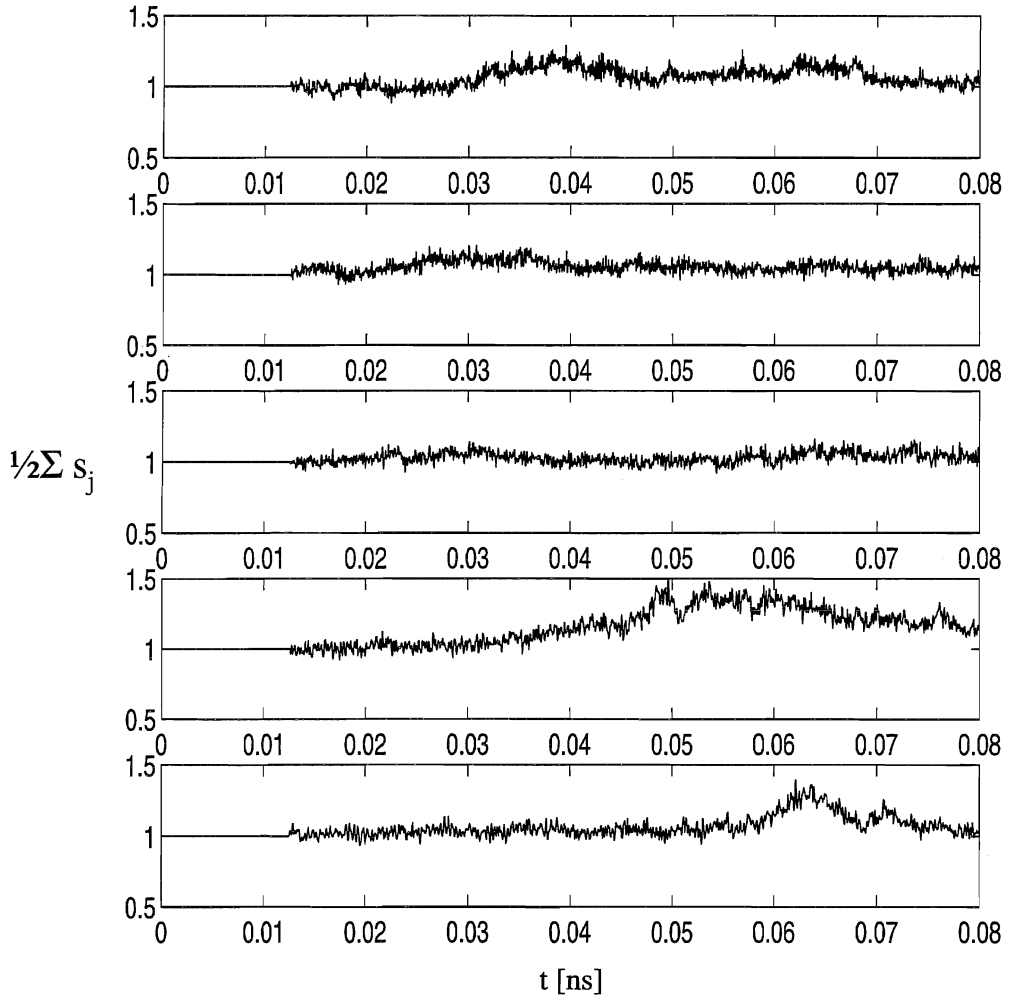


Figure D.36 Average $(s_{trans} + s_{rot})/2$ of heat bath variables for all mixtures, ranging from pure water (top) to pure ethanol (bottom).

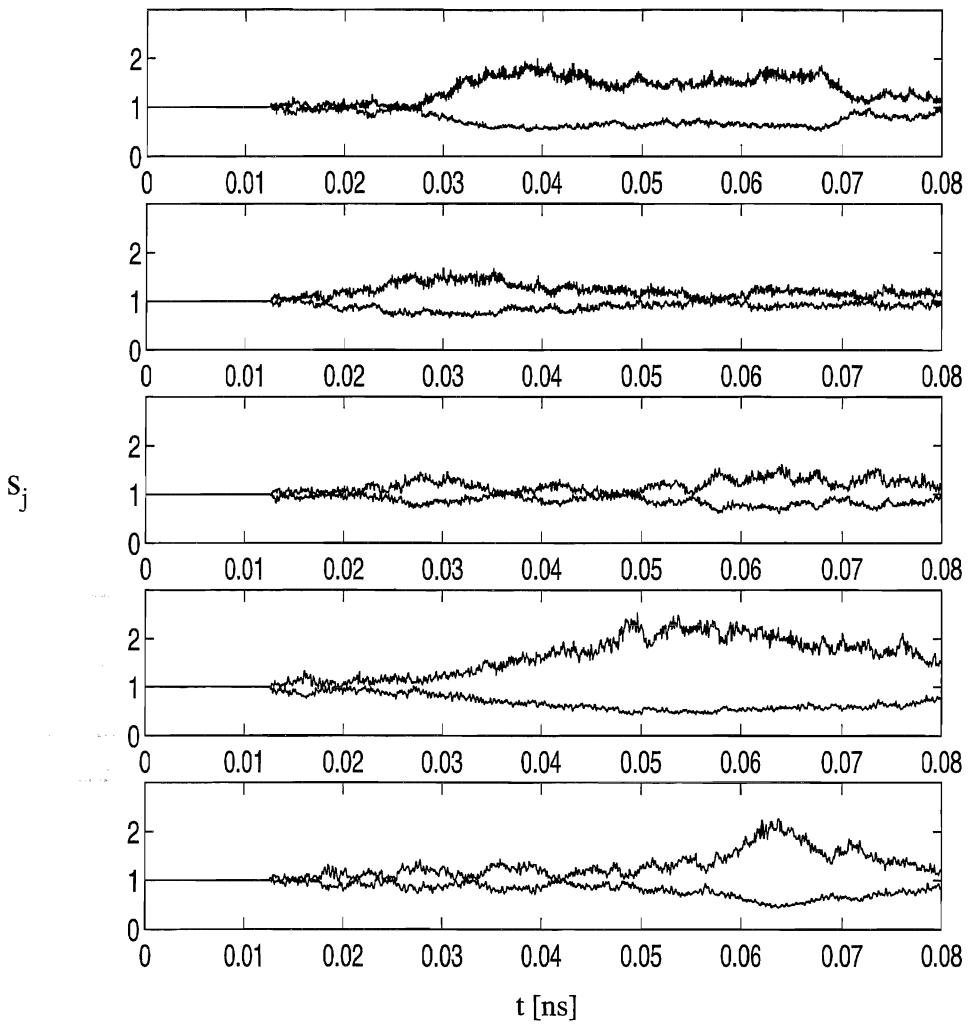


Figure D.37 Heat bath variables s_{trans} and s_{rot} for all mixtures, ranging from pure water on top to pure ethanol in the bottom panel.
Panel 1 and 2: Lower curve is s_{trans}
Panel 3: s_{trans} on top to t~0.035ns bottom for the rest. Thr curves touch at 0.045ns
Panel 4: s_{trans} on top, except at the very beginning. s_{rot} has now the samllest amplitude.
Panel 5: s_{trans} on top, except around 0.02ns. The curves touch at several places.

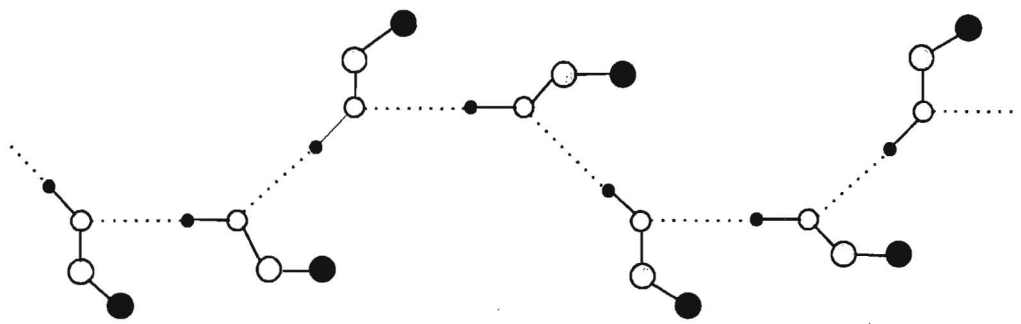


Figure D.38 This chain show how the u-shape could look like when the chain consists only of trans conformers. We have only varied site posistions in the paper plane, but a utilization of the full thre dimensional space is more attractive. Symbols as in Figure 6.7.

Table D.4
Radial distance and height of 1st and 2nd maximum and of 1st minimum for site-site self correlations of water with varying ethanol content. Results for 293K from NVT-simulations.

x_e	g(r)	1.maximum		1.minimum		2.maximum	
		r[Å]	Height	r[Å]	Depth	r[Å]	Height
0	g_{OO}	2.85	3.02	3.45	0.80	4.4	1.12
	g_{OH}	1.9	1.50	2.5	0.20	3.25	1.52
	g_{HH}	2.4	1.26	3.0	0.74	3.8	1.16
0.25	g_{OO}	2.85	4.79	3.45	0.83	4.6	1.24
	g_{OH}	1.9	2.41	2.55	0.24	3.25	2.09
	g_{HH}	2.45	1.89	3.1	0.92	3.85	1.40
0.50	g_{OO}	2.85	7.08	3.5	0.95	4.5	1.40
	g_{OH}	1.9	3.54	2.55	0.31	3.25	2.89
	g_{HH}	2.4	2.74	3.1	1.16	3.8	1.78
0.75	g_{OO}	2.85	8.29	3.5	0.81	4.5	1.30
	g_{OH}	1.9	4.25	2.55	0.31	3.2	3.10
	g_{HH}	2.45	3.09	3.1	1.09	3.75	1.80

Table D.5 (Next page)
Radial distance and height of 1st and 2nd maximum and of 1st minimum for site-site self correlations of ethanol in aqueous mixures with varying ethanol content. Results for 298 K from NVT-simulations

x_e	g	1.maximum		1.minimum		2.maximum		2nd min	3rd max
		r[Å]	Height	r[Å]	Depth	r[Å]	Height	r/height	r/height
	g_{C3C3}	4.3	1.83	6.8	0.76	8.95	1.06		
	g_{C3C2}	4.5	1.53	6.55	0.81	8.5	1.04		
	g_{C3O}	-	-	-	-	5.05	1.30	7.3/0.94	
	g_{C3H}	-	-	-	-	5.45	1.12	6.1/1.00	7.05/1.0
0.25	g_{C2C2}	4.7	1.52	7.05	0.85	8.85	1.03		
	g_{C2O}	3.75	1.02	4.2	0.79	5.4	1.21		
	g_{C2H}	3.0	0.56	3.2	0.43	4.25	0.99		
	g_{OO}	2.75	1.42	3.3	0.31	4.8	1.02		
	g_{OH}	1.85	1.41	2.6	0.13	-	-		
	g_{HH}	2.45	1.50	3.2	0.50	-	-		
	g_{C3C3}	4.2	1.65	6.6	0.74	8.65	1.10		
	g_{C3C2}	4.65	1.46	6.7	0.80	8.75	1.07		
	g_{C3O}	4.05	1.02	4.35	0.97	4.95	1.39		
	g_{C3H}	-	-	-	-	5.45	1.16	6.05/1.1	6.9/1.08
0.50	g_{C2C2}	4.6	1.67	7.05	0.84	9.1	1.07		
	g_{C2O}	3.8	1.32	4.3	0.85	5.5	1.15	7.1/0.93	
	g_{C2H}	3.0	0.89	3.3	0.58	4.25	1.06		
	g_{OO}	2.75	2.50	3.55	0.34	4.7	1.11		
	g_{OH}	1.85	2.55	2.6	0.17	3.35	1.06	5.5/0.85	
	g_{HH}	2.4	2.57	3.2	0.57	4.9	0.96		
	g_{C3C3}	4.3	1.52	6.85	0.73	8.8	1.12		
	g_{C3C2}	4.75	1.45	6.75	0.78	8.75	1.10		
	g_{C3O}	3.95	1.07	4.5	0.99	5.0	1.52		
	g_{C3H}	3.2	0.87	3.7	0.80	5.4	1.20	6.0/0.99	6.9/1.08
0.75	g_{C2C2}	4.7	1.85	6.95	0.79	8.85	1.11		
	g_{C2O}	3.75	1.63	4.4	0.83	5.6	1.15	7.0/0.89	
	g_{C2H}	3.0	1.23	3.4	0.66	4.1	1.15		
	g_{OO}	2.75	3.51	3.45	0.39	4.85	1.09		
	g_{OH}	1.85	3.64	2.65	0.20	3.35	1.32		
	g_{HH}	2.4	3.43	3.35	0.55	4.7	0.92	5.35/0.9	
	g_{C3C3}	4.3	1.47	7.05	0.41	9.3	1.15		
	g_{C3C2}	4.5	1.48	6.85	0.76	8.95	1.09		
	g_{C3O}	3.9	1.20	4.6	0.97	4.95	1.72		
	g_{C3H}	3.15	1.14	3.85	0.78	5.4	1.25	6.1/0.96	6.95/1.8
1.00	g_{C2C2}	4.65	2.08	6.95	0.74	9.05	1.15		
	g_{C2O}	3.7	1.99	4.4	0.76	5.5	1.08	7.0/0.86	
	g_{C2H}	3.0	1.78	3.45	0.70	4.1	1.16		
	g_{OO}	2.75	5.22	3.6	0.23	4.85	1.03		
	g_{OH}	1.85	5.51	2.7	0.14	3.35	1.54		
	g_{HH}	2.4	4.90	3.35	0.31	4.65	0.84	5.45/0.6	

Table D.6
Radial distance and height of 1st and 2nd maximum and of 1st minimum for site-site cross-correlations of water with varying ethanol content. Results for 293 K from NVT-simulations. Water first indices.

x_e	g	1.maximum		1.minimum		2.maximum		2.maximum
		r[Å]	Height	r[Å]	Depth	r[Å]	Height	r[Å]/height
0.25	g_{OC3}	3.75	1.46	4.65	0.93	4.95	1.17	5.6/0.82
	g_{OC2}	3.7	1.88	(4.55)	(0.98)	(4.9)*	(1.04)*	6.05/0.89
	g_{OO}	2.75	3.14	*	*	4.75	1.16	6.15/0.90
	g_{OH}	1.85	2.38	3.45	0.56	3.35	1.58	-
	g_{HC3}	-	-	2.6	0.24	5.35	1.03	6.15/0.90
	g_{HC2}	2.95	0.74	-	-	4.25	1.33	5.95/0.90
	g_{HO}	1.85	2.01	3.25	0.56	3.25	1.40	5.5/0.92
	g_{HH}	2.35	1.97	2.55	0.16	3.8	1.04	-
				3.1	0.73			
0.50	g_{OC3}	3.75	1.54	4.65	0.92	4.95	1.32	5.6/0.81
	g_{OC2}	3.7	2.31	(4.45)	(0.94)	(4.75)*	(1.00)*	6.35/0.87
	g_{OO}	2.75	4.40	*	*	4.65	1.20	5.95/0.83
	g_{OH}	1.85	3.54	3.5	0.57	3.35	1.95	-
	g_{HC3}	-	-	2.6	0.29	5.35	1.08	6.15/0.90
	g_{HC2}	2.95	0.97	-	-	4.25	1.47	5.9/0.90
	g_{HO}	1.85	2.79	3.3	0.67	3.25	1.77	5.5/0.88
	g_{HH}	2.4	2.76	2.6	0.20	3.85	1.13	-
				3.15	0.81			
0.75	g_{OC3}	3.75	1.79	4.55	0.98	5.0	1.57	5.6/0.82
	g_{OC2}	3.75	2.93	(4.5)*	(0.94)	5.55	1.00	6.3/0.80
	g_{OO}	2.8	6.43	3.55	*	4.65	1.20	5.85/0.74
	g_{OH}	1.85	5.27	2.65	0.59	3.35	2.56	-
	g_{HC3}	3.25	1.10	3.9	0.35	5.3	1.18	6.15/0.94
	g_{HC2}	2.95	1.38	3.35	1.00	4.2	1.73	6.2/0.86
	g_{HO}	1.85	4.07	2.6	0.78	3.25	2.39	5.6/0.78
	g_{HH}	2.35	3.82	3.15	0.23	3.8	1.24	-
					0.94			

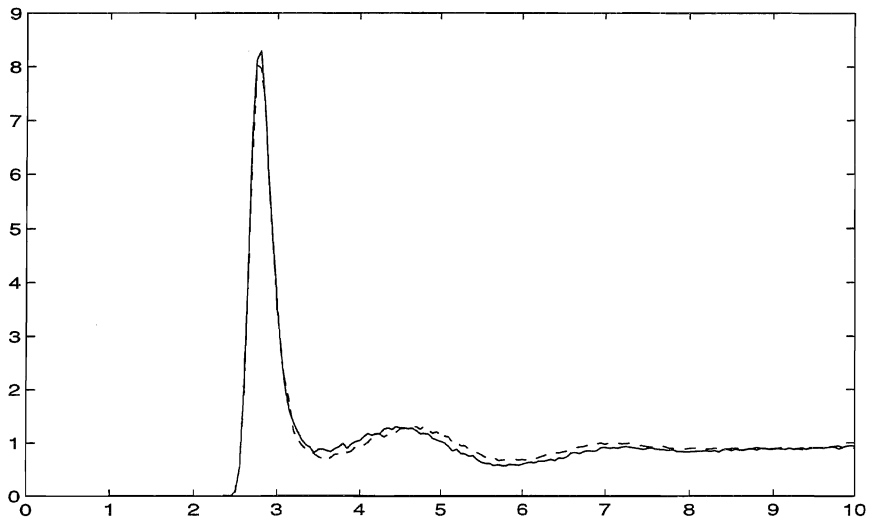


Figure D.39 Comparison self correlations g_{OO} for water from NVE (dashed line) and NVT (solid line) simulations of x_e=0.75 mixture of ethanol and water.

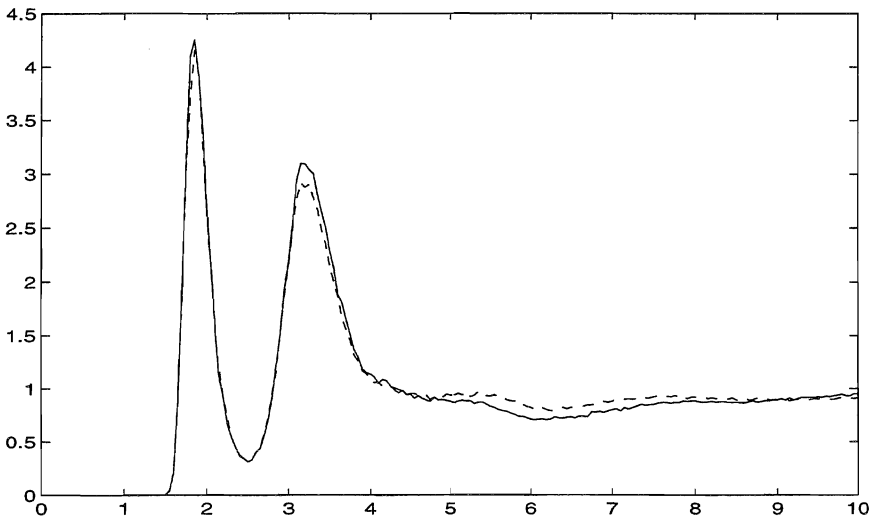


Figure D.40 Comparison self correlations g_{OH} for water from NVE (dashed line) and NVT (solid line) simulations of x_e=0.75 mixture of ethanol and water.

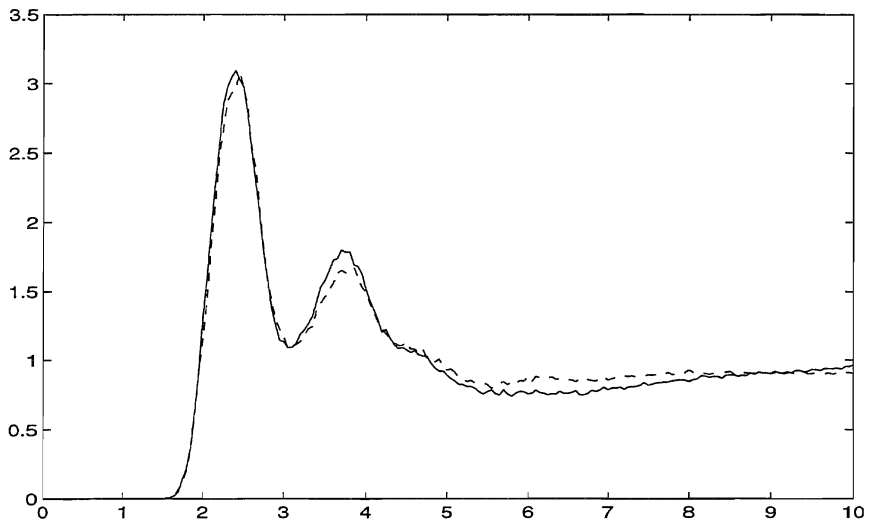


Figure D.41 Comparison self correlations g_{HH} for water from NVE (dashed line) and NVT (solid line) simulations of x_e=0.75 mixture of ethanol and water.

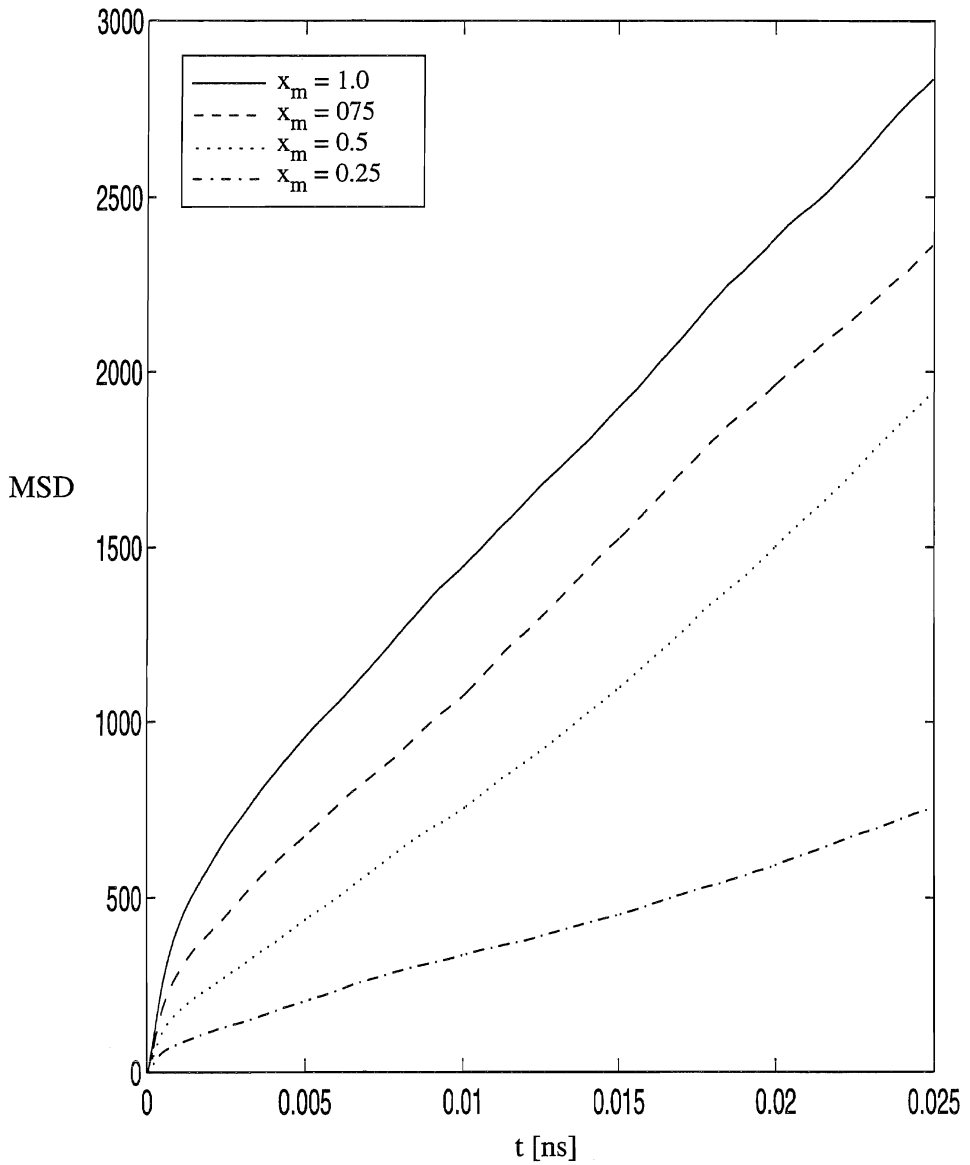


Figure D.42 Mean square displacement (MSD) for ethanol in mixtures with water at 293K.

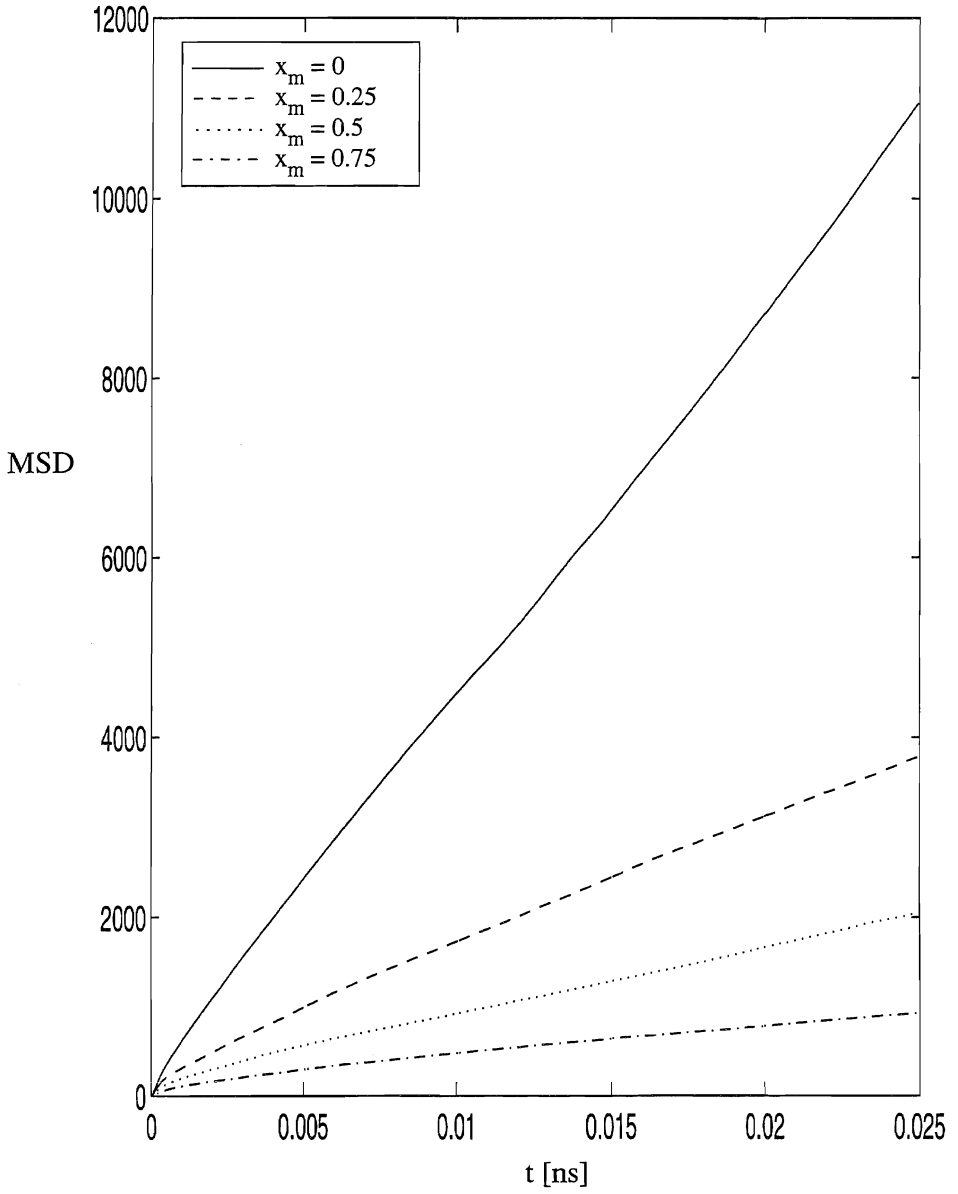


Figure D.43 Mean square displacement MSD for water in mixtures with ethanol at 293K

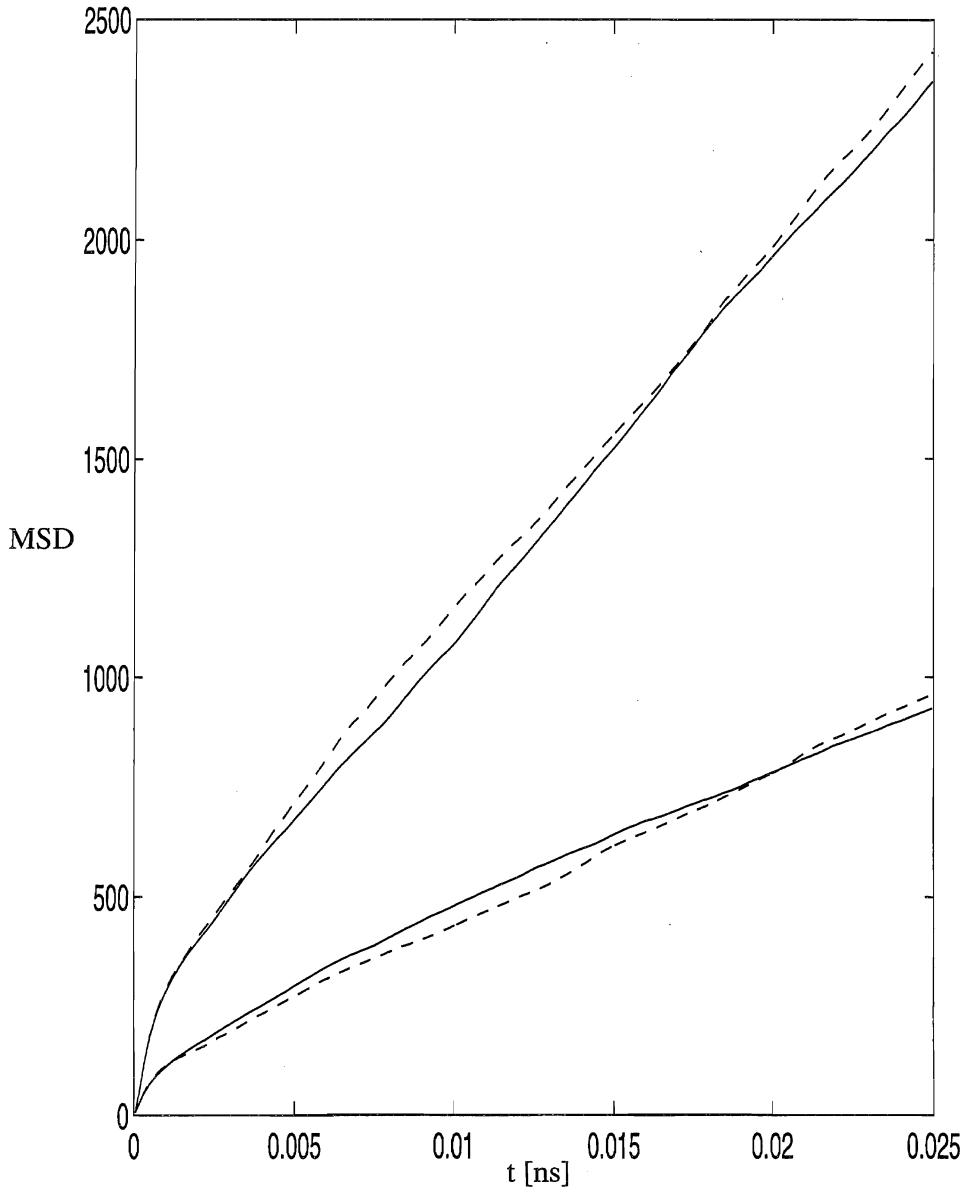


Figure D.44 Difference of NVT (———) and NVE (- - - - -) results for mean square displacements (MSD) in an ethanol-water mixture with mole fraction ethanol 0.75. Upper pair ethanol, lower pair water.